Transactions of the ASME®

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 58 • NUMBER 3 • SEPTEMBER 1991

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) for $160.00 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 Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.

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

PRICES: To members, $29.00, annually: to nonmembers, $160.00. Add $20.00 for postage to countries outside the United States and Canada.

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc. Canadian Goods & Services Tax Registration #126148048.

(Contents continued on Inside Back Cover)

CONTENTS (CONTINUED)

BRIEF NOTES

BOOK REVIEWS

ANNOUNCEMENTS AND SPECIAL NOTES

A Reformation of the Equations of Anisotropic Poroelasticity

M. Thompson
Schlumberger Cambridge Research,
Cambridge CB3 0HG, U.K.

J. R. Willis
School of Mathematical Sciences,
University of Bath,
Bath BA2 7AY, U.K.

The constitutive equations of linear poroelasticity presented by Biot (1955) and Biot and Willis (1957) extended the description of rock behavior into the realm of saturated porous rocks. For isotropic material behavior, Rice and Cleary (1976) gave a formulation which involved material constants whose physical interpretation was particularly simple and direct; this is an aid both to their measurement and to the interpretation of predictions from the theory. This paper treats anisotropic poroelasticity *in terms of material tensors with interpretations similar to those of the constants employed by Rice and Cleary. An* effective stress principle *is derived for such anisotropic material. The material tensors are defined, rigorously, from the stress field and pore fluid content changes produced by boundary displacements compatible with a uniform mean strain and uniform pore pressure increments. Such displacements and pore pressure increments lead to homogeneous deformation on all scales significantly larger than the length scale of microstructural inhomogeneities. This macroscopic behavior is related to the microscopic behavior of the solid skeleton. The tensors which describe the microscopic behavior of the solid skeleton would be difficult, even impossible, to measure, but their introduction allows relationships between measurable quantities to be identified. The end product of the analysis is a set of constitutive equations in which the parameters are all measurable directly from well-accepted testing procedures. Relationships exist between measurable quantities that can be used to verify that the constitutive equations described here are valid for the rock under consideration. The case of* transverse isotropy *is discussed explicitly for illustration.*

1 Introduction

The background to the work presented here lies in rock mechanics so it is convenient to describe the problem in those terms, although the application area for the work is much broader than that.

When carrying out mechanical testing of rocks, even dry nonporous rocks, there can be problems of repeatability and inconsistency in results over different size scales. Some of the difficulties arise through large-scale inhomogeneity and from the distribution of flaws which can result in small complete specimens being stronger than large ones. The testing techniques themselves can also present difficulties. After great care has been taken to deal with such problems, if there remain wide variations between tests then the tendency is to test many samples and average the results. However, there can be unforeseen bias in the tests which is not helped by averaging. The material constitutive description may be inadequate such that there are some constitutive parameters which lie outside experimental control. It was just this sort of inadequacy that occurred in the testing of saturated porous media until the full importance was realized of the water content in specimen preparation and storage and of pore pressure during testing. The linear poroelastic description of Biot (1955) and Biot and Willis (1957) was crucial in explaining the relevance of pore pressure. Biot and Willis outlined experiments sufficient for identification of the constants in their theory but experimental determination is simpler relative to the formulation of Rice and Cleary (1976), since the constants introduced therein have very direct physical interpretations.

When testing shales, even while restricting attention to mechanical behavior independent of chemistry, there are still problems remaining. Shales are often bedded and behave in an anisotropic, nonlinear manner and include inhomogeneities on the microscale. With such a rock the behavior lies outside the descriptive range of isotropic linear poroelasticity. This paper gives the constitutive equations describing the *incremental* behavior of an anisotropic poroelastic material. The incremental nature of the constitutive equations allows them to be applied incrementally to nonlinear behavior. The parameters that appear in these equations can be obtained by standard triaxial tests familiar to rock mechanics experimentalists. A valuable aspect of the work is the derivation of an effective stress principle, even in this case of anisotropy.

Figure 1 illustrates the relationship of the skeleton material to the saturated porous elastic medium. The term *skeleton*

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Sept. 26, 1989; final revision, Mar. 10, 1990.

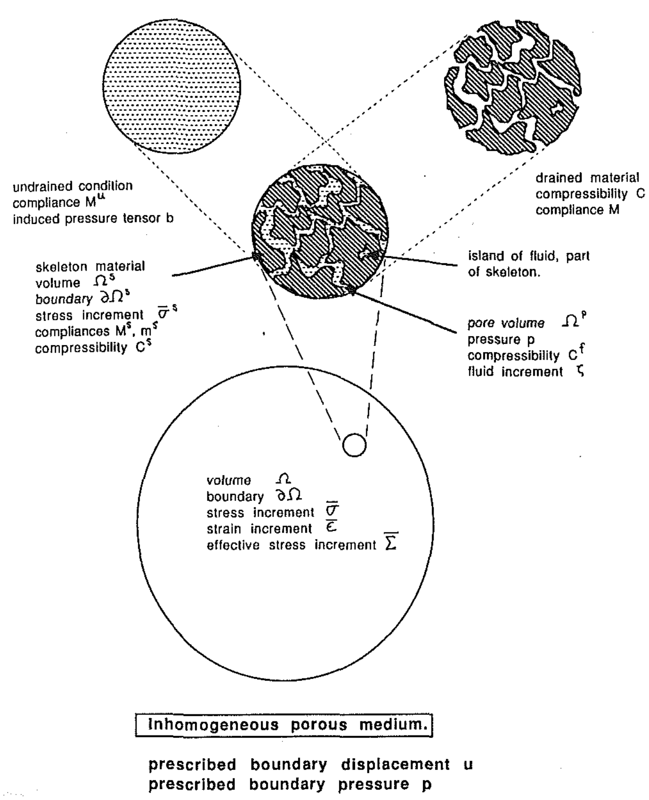


Fig. 1 Representation of the porous material at different scales, with an indication of the associated mechanical properties

material does not mean an intact piece of rock with no water rather like a dry sponge, but instead refers to a representative piece of the material that makes up the skeleton. The skeleton material will be regarded as nonporous, even when there is immobile water such as bound water in shales. The mechanical properties of the skeleton are not expected to be within our capability of measuring. In our derivation the skeleton material need not be homogeneous and bedding could well occur on a large scale across the structure of the skeleton.

A departure from the usual test procedure is to prescribe the boundary displacements rather than controlling the boundary stresses. Testing the rock involves the determination of:

- the drained compliances **M**. *Drained* here refers to a saturated specimen at atmospheric pressure, not a dry material.
- the undrained compliances $\mathbf{M}^u$ and the tensor **b** which relates pore pressure change to stress increment. In such a test no fluid is allowed to cross the boundary of the sample. It is difficult to prevent fluid crossing the boundaries while also measuring the pore pressure but it can be done with care. (See, for example, Mesri, Adachi, and Ullrick (1976)).
- the strain field generated by isotropic stress equal to pore pressure magnitude.

It is presumed that during tests the pressure distribution throughout the sample is uniform so that in some low permeability rocks, the testing time will be long. The testing scale imposes a form of averaging which then decides the spatial scale over which the constitutive equations are applicable. Suppose the rock has inhomogeneities on a scale L_i and the measurement scale is L_m. If $L_m \gg L_i$, then it is acceptable to apply the results to larger scales of rock. Furthermore, the results will be independent of the type of boundary condition that is applied. If $L_m = 0(L_i)$, then each sample will be unique, the results cannot be scaled up and will apply only for boundary conditions of the proposed displacement type. Note that, unless assumptions are made about the homogeneity of the material, the scale of application cannot be smaller than the testing scale for no information has been obtained on the smaller scale. It is important that time scales of application are slow enough to allow inertia effects to be ignored for these are not included in this constitutive behavior. This rules out the application of these constitutive equations to acoustic wave travel.

The tensors $\mathbf{M}^u$ and **b**, which can be measured, are expressible in terms of the unmeasurable properties of the skeleton material; the tensor **b** generalizes the parameters, A and B, introduced by Skempton (1954) and we will refer to it as the "Skempton tensor." Elimination of the tensors for the solid skeleton between the equations results in relations between drained and undrained compliances and the Skempton tensor: The structure of these relations shows that not all components of the drained and undrained compliance tensors can be independent. The tensor which appears in the "effective stress" principle is also identified and expressed in terms of drained and undrained compliances and the Skempton tensor.

The outcome of the analysis is a set of constitutive relations (for strain increment versus stress and fluid pressure increments, and for increment in fluid content) in which the parameters that appear are given explicitly in terms of drained and undrained compliances and the components of the Skempton tensor. They are illustrated by an explicit treatment of transverse isotropy, with full isotropy following as a limiting case.

2 Basic Relations

Throughout this work, σ will denote an *increment* of stress and ϵ will denote the corresponding increment of strain, superimposed upon some known initial state of deformation of a fluid-saturated porous medium. The medium, as initially deformed, occupies a region of space denoted by Ω, whose bounding surface is $\partial\Omega$ (Fig. 1). The solid skeleton occupies Ω_s, while the pores occupy Ω_p, so that Ω is the union of Ω_s and Ω_p. The porosity φ is the ratio of the volumes of Ω_p and Ω.

For the purpose of specifying test conditions by which the constitutive parameters are defined, only quasi-static increments are considered, so that σ is self-equilibrated over Ω_s and the fluid pressure increment p is uniform over Ω_p. It is thus assumed that all points of Ω_p are connected to $\partial\Omega$ by paths lying wholly within Ω_p. "Islands" of fluid isolated from $\partial\Omega$ may be present but these are considered as part of Ω_s. We could specify traction or displacement boundary conditions on $\partial\Omega$. However, to be definite, an idealized mode of loading is envisaged in which the boundary $\partial\Omega$ is regarded as a permeable surface which can be subjected to incremental displacements of the form $u_i = \overline{a}_{ij}\, x_j$. Here, $\overline{a}_{ij}$ are constants and x_j are components of the position vector **x** of a generic point of $\partial\Omega$. Since $\partial\Omega$ is permeable, the fluid pressure increment p is controllable independently of the constant tensor $\overline{\mathbf{a}}$. The average displacement gradient over Ω is $\overline{\mathbf{a}}$ exactly. The average strain $\overline{\epsilon}$ is the symmetric part of $\overline{\mathbf{a}}$. Since rigid rotations have no effect on constitutive behavior, $\overline{\mathbf{a}}$ will be taken as symmetric, and hence identical to $\overline{\epsilon}$, in the sequel. Application of such displacement boundary conditions is well established in the context of the micromechanics of composite materials (Hill, 1963).

The average stress increment in the porous medium has (ij) component

$$\overline{\sigma}_{ij} = \frac{1}{|\Omega|} \int_{\partial\Omega} \sigma_{ik}\, n_k\, x_j\, ds \tag{1}$$

exactly, where $|\Omega|$ denotes the volume of Ω and the unit normal **n** is directed out of Ω. Constitutive behavior is taken as incrementally linear, so increments in mean stress, strain, and fluid pressure are related linearly and are thus additive. Thus, it is legitimate to *define* tensors **M**, **m** so that

$$\overline{\epsilon} = \mathbf{M}:\overline{\sigma} + (\mathbf{M}:\delta - \mathbf{m})p \tag{2}$$

during any loading of the class described above. Here, **M** denotes the fourth-order tensor of compliances under drained conditions ($p=0$), while the second-order tensor $\mathbf{m}$ defines the response to an "all-round" pressure change p, so that $\bar{\sigma}=-p\boldsymbol{\delta}$. The tensor $\boldsymbol{\delta}$ has components δ_{ij}. Notation is employed, in (2) and the sequel, in which a dot implies contraction over a single repeated suffix. Thus, $\mathbf{M}:\bar{\sigma}$ denotes a second-order tensor with components $M_{ijkl}\,\bar{\sigma}_{kl}$.

An "effective stress" principle follows immediately from (2). Introducing $\mathbf{L}$ as the tensor inverse to $\mathbf{M}$, it follows that (2) is equivalent to

$$\bar{\epsilon}=\mathbf{M}:\bar{\Sigma}, \tag{3}$$

where the effective stress increment $\bar{\Sigma}$ is given by

$$\bar{\Sigma}=\bar{\sigma}+\alpha p, \tag{4}$$

with

$$\alpha=\boldsymbol{\delta}-\mathbf{L}:\mathbf{m}. \tag{5}$$

This form for α is sufficient to define it experimentally from the measurables $\mathbf{M}$ and $\mathbf{m}$. One of the objectives of this work will be to relate α to other physical properties.

If now $\bar{\sigma}^s$ denotes the average value of σ over Ω_s, it follows that

$$\bar{\sigma}=(1-\varphi)\bar{\sigma}^s-\varphi p\boldsymbol{\delta}, \tag{6}$$

since the stress in the pores is $-p\boldsymbol{\delta}$, exactly. Correspondingly, if the mean value of ϵ over Ω_s is $\bar{\epsilon}^s$, tensors $\mathbf{M}^s$, $\mathbf{m}^s$ may be *defined* so that

$$\bar{\epsilon}^s=\mathbf{M}^s:\bar{\sigma}^s+(\mathbf{M}^s:\boldsymbol{\delta}-\mathbf{m}^s)p. \tag{7}$$

$\mathbf{M}^s$ and $\mathbf{m}^s$ are difficult to measure directly but it should be noted that (7) defines them, even when the skeleton is inhomogeneous. If the skeleton is, in fact, homogeneous, then $\mathbf{M}^s$ is its tensor of compliances and $\mathbf{m}^s=\mathbf{M}^s:\boldsymbol{\delta}$. This follows because in a homogeneous medium, any stress field with average value $\bar{\sigma}^s$ generates a strain field with average value $\mathbf{M}^s:\bar{\sigma}^s$ and so, in particular, when $\bar{\sigma}^s=-p\boldsymbol{\delta}$, $\bar{\epsilon}=-\mathbf{M}^s:\boldsymbol{\delta}p$ directly, or $-\mathbf{m}^s p$ from (7). The introduction of $\mathbf{M}^s$ and $\mathbf{m}^s$ *as definitions* is necessary when the skeleton is inhomogeneous.

It is also true, generally, that

$$\mathbf{m}=\boldsymbol{\delta}:\mathbf{M}^s, \tag{8}$$

subject only to the condition that the skeleton has local compliances with the "usual" symmetry: In suffix notation, $M'_{ijkl}=M'_{klij}$, where here $\mathbf{M}'$ denotes the local compliance of the skeleton material. The proof is given in the Appendix. It follows that the expression (5) for the tensor α related to the effective stress reduces to that given by Carroll (1979) so long as (8) holds. Carroll's derivation applied only to a homogeneous skeleton.

Before moving on, it is noted that elimination of $\bar{\sigma}^s$ between (6) and (7) yields

$$\bar{\epsilon}^s=\mathbf{M}^s:(\bar{\sigma}+p\boldsymbol{\delta})/(1-\varphi)-\mathbf{m}^s p. \tag{9}$$

An expression is now developed for the increment ζ in fluid content (defined as increment of fluid mass divided by initial fluid density), per unit volume of the porous medium, induced by the average stress and pressure increments $\bar{\sigma}$ and p. Let v represent the fractional volume change of the pores. Then, analogously to (6),

$$\boldsymbol{\delta}:\bar{\epsilon}=(1-\varphi)\boldsymbol{\delta}:\bar{\epsilon}^s+\varphi v. \tag{10}$$

But also,

$$v=-C^f p+\zeta/\varphi, \tag{11}$$

where C^f denotes fluid compressibility. Therefore, from (10) and (11),

$$\zeta=\boldsymbol{\delta}:\bar{\epsilon}-(1-\varphi)\boldsymbol{\delta}:\bar{\epsilon}^s+\varphi C^f p, \tag{12}$$

and so, from (2) and (9),

$$\zeta=\boldsymbol{\delta}:(\mathbf{M}-\mathbf{M}^s):\bar{\sigma}+[C-C^s-\boldsymbol{\delta}:(\mathbf{m}-\mathbf{m}^s)+\varphi(C^f-\boldsymbol{\delta}:\mathbf{m}^s)]p, \tag{13}$$

where the overall drained compressibility C is defined as

$$C=\boldsymbol{\delta}:\mathbf{M}:\boldsymbol{\delta}, \tag{14}$$

with a similar definition for C^s.

When the skeleton is composed of homogeneous material, so that $\mathbf{m}=\mathbf{m}^s=\mathbf{M}^s:\boldsymbol{\delta}$, (13) reduces to

$$\zeta=\boldsymbol{\delta}:(\mathbf{M}-\mathbf{M}^s):\bar{\sigma}+[C-C^s+\varphi(C^f-C^s)]p. \tag{15}$$

In the particular case of undrained deformation, $\zeta=0$ and (13) then yields

$$p=-\mathbf{b}:\bar{\sigma}, \tag{16}$$

where

$$\mathbf{b}=\boldsymbol{\delta}:(\mathbf{M}-\mathbf{M}^s)/[C-C^s-\boldsymbol{\delta}:(\mathbf{m}-\mathbf{m}^s)+\varphi(C^f-\boldsymbol{\delta}:\mathbf{m}^s)]. \tag{17}$$

The tensor $\mathbf{b}$ generalizes the Skempton parameters A, B, which apply to transversely isotropic porous media (Skempton, 1954).

It follows now that, during undrained deformation,

$$\bar{\epsilon}=\mathbf{M}^u:\bar{\sigma}, \tag{18}$$

where

$$\mathbf{M}^u=\mathbf{M}-(\mathbf{M}:\boldsymbol{\delta}-\mathbf{m})\times\mathbf{b} \tag{19}$$

or, in suffix notation,

$$M^u_{ijkl}=M_{ijkl}-(M_{ijpp}-m_{ij})b_{kl}. \tag{20}$$

Equation (19) is sufficient to determine $\mathbf{b}$ from the measureables $\mathbf{M}$, $\mathbf{M}^u$, and $\mathbf{m}$ although it is defined also through (16) directly. It may be noted that $\mathbf{M}^u$ has the usual symmetry when the skeleton does, so that (8) holds.

The basic constitutive relations are (2) and (13). The latter equation was developed via explicit consideration of the micromechanics and, in consequence, involves tensors $\mathbf{M}^s$, $\mathbf{m}^s$ which are not susceptible to direct measurement. The approach via micromechanics has previously been followed by Nur and Byerlee (1971), Carroll (1979), Carroll and Katsube (1983) and others[1] but only for the case of a skeleton composed of homogeneous material. The phenomenological approach initiated by Biot pays less explicit attention to the behavior of the skeleton (though the parameters introduced are motivated through appreciation of micromechanical processes). The approach of Biot introduces, in addition, thermodynamic potentials which impose *a priori* symmetry restrictions which need not always be realized. A result of the form of (19), with (8) incorporated, was derived by Rudnicki (1985), through consideration of Maxwell relations for a thermodynamic potential. Rudnicki's equation was given for moduli rather than compliances, but similar arguments are applicable for the present formulation: The relevant thermodynamic potential is the Helmholtz function ψ for which

$$d\psi=\bar{\sigma}:d\bar{\epsilon}+pd\zeta \tag{21}$$

(Rudnicki, personal communication). The argument given in the Appendix has the advantage of precision: Equation (8) follows rigorously, even for a sample whose macroscopic dimensions are of the order of the microstructural dimensions, so long as the skeleton material has the requisite local symmetry *and* the boundary conditions are of the type assumed, but not necessarily for other boundary conditions.

Throughout the remainder of this paper, (8) will be assumed to hold.

3 Elimination of Unmeasurable Quantities

It has been remarked already that the tensors $\mathbf{M}^s$, $\mathbf{m}^s$ are not susceptible to direct measurement. They can, however, be

[1]A minor flaw in the derivation of Carroll (1979), repeated in later work, is that a boundary condition $\sigma\cdot\mathbf{n}=\bar{\sigma}\cdot\mathbf{n}$ is applied over the whole of $\partial\Omega$, even though the fluid can sustain only hydrostatic stress. His derivation thus applies strictly to a medium, none of whose pores intersect the outer boundary. Use of the displacement boundary condition removes this difficulty.

eliminated between the preceding equations in a variety of ways. First, it follows from (19) (or (20)) that

$$(\mathbf{M}:\boldsymbol{\delta}-\mathbf{m})_{ij}=(M_{ijkl}-M^{u}_{ijkl})/b_{kl}, \tag{22}$$

with no sum on k, l, for any values of k, l for which $b_{kl}\neq 0$. Thus, there are several equations relating components of $\mathbf{M}^u$ to components of $\mathbf{M}$. One explicit equation, equivalent to (22), is obtained by noting from (19) that

$$(\mathbf{M}-\mathbf{M}^u):\boldsymbol{\delta}=(\mathbf{M}:\boldsymbol{\delta}-\mathbf{m})(\mathbf{b}:\boldsymbol{\delta}),$$

so that

$$\mathbf{M}:\boldsymbol{\delta}-\mathbf{m}=(\mathbf{M}-\mathbf{M}^u):\boldsymbol{\delta}/(\mathbf{b}:\boldsymbol{\delta}). \tag{23}$$

Also, contracting (23) with $\boldsymbol{\delta}$,

$$C-C^s=(C-C^u)/(\mathbf{b}:\boldsymbol{\delta}), \tag{24}$$

while, from (17),

$$\mathbf{b}:\boldsymbol{\delta}=(C-C^s)/[C-C^s-\boldsymbol{\delta}:(\mathbf{m}-\mathbf{m}^s)+\varphi(C^f-\boldsymbol{\delta}:\mathbf{m}^s)]. \tag{25}$$

Hence,

$$[C-C^s-\boldsymbol{\delta}:(\mathbf{m}-\mathbf{m}^s)+\varphi(C^f-\boldsymbol{\delta}:\mathbf{m}^s)]=(C-C^u)/(\mathbf{b}:\boldsymbol{\delta})^2. \tag{26}$$

Thus, the general relation (13) can be written

$$\zeta=\frac{\boldsymbol{\delta}:(\mathbf{M}-\mathbf{M}^u):\boldsymbol{\sigma}}{(\mathbf{b}:\boldsymbol{\delta})}+\frac{(C-C^u)p}{(\mathbf{b}:\boldsymbol{\delta})^2} \tag{27}$$

having used (8), and (5) can be written

$$\boldsymbol{\alpha}=(\boldsymbol{\delta}-\mathbf{L}:\mathbf{M}^u:\boldsymbol{\delta})/(\mathbf{b}:\boldsymbol{\delta}). \tag{28}$$

The expression (27) can also be given in terms of $\boldsymbol{\alpha}$, by noting that, from (28),

$$\mathbf{M}:\boldsymbol{\alpha}=(\mathbf{M}-\mathbf{M}^u):\boldsymbol{\delta}/(\mathbf{b}:\boldsymbol{\delta}) \tag{29}$$

and hence that

$$\boldsymbol{\delta}:\mathbf{M}:\boldsymbol{\alpha}=(C-C^u)/(\mathbf{b}:\boldsymbol{\delta}). \tag{30}$$

Thus, (27) gives

$$\zeta=\boldsymbol{\alpha}:\mathbf{M}:\bar{\boldsymbol{\sigma}}+[(\boldsymbol{\delta}:\mathbf{M}:\boldsymbol{\alpha})/(\mathbf{b}:\boldsymbol{\delta})]p. \tag{31}$$

It is possible also to eliminate $\bar{\boldsymbol{\sigma}}$ in favor of $\bar{\boldsymbol{\epsilon}}$: The neatest result is obtained by using the effective stress principle embodied in (3)–(5), to give

$$\zeta=\boldsymbol{\alpha}:\bar{\boldsymbol{\epsilon}}-[\boldsymbol{\alpha}:\mathbf{M}:\boldsymbol{\alpha}-(\boldsymbol{\delta}:\mathbf{M}:\boldsymbol{\alpha})/(\mathbf{b}:\boldsymbol{\delta})]p. \tag{32}$$

4 Transverse Isotropy

If the material is transversely isotropic, with its axis of symmetry parallel to $0x_3$, the stress-strain relations (3) can be given in the form

$$\epsilon_{11}=\frac{1}{E}\Sigma_{11}-\frac{\nu}{E}\Sigma_{22}-\frac{\nu'}{E'}\Sigma_{33},$$

$$\epsilon_{22}=\frac{-\nu}{E}\Sigma_{11}+\frac{1}{E}\Sigma_{22}-\frac{\nu'}{E'}\Sigma_{33},$$

$$\epsilon_{33}=\frac{-\nu''}{E}(\Sigma_{11}+\Sigma_{22})+\frac{1}{E'}\Sigma_{33},$$

$$\epsilon_{12}=\frac{1}{2G}\Sigma_{12},\ \epsilon_{23}=\frac{1}{2G'}\Sigma_{23},\ \epsilon_{13}=\frac{1}{2G'}\Sigma_{13}, \tag{33}$$

recalling that Σ refers to the effective stress (4). In (33), G is the shear modulus in the plane of isotropy and is equal to $E/[2(1+\nu)]$. The usual symmetry (as employed in the Appendix) implies

$$\frac{\nu''}{E}=\frac{\nu'}{E'}. \tag{34}$$

This will be assumed throughout this section. The constants in (33) are otherwise independent. The tensors $\mathbf{M}^u$ and $\mathbf{M}^s$ can be expressed in similar explicit forms, although there is not necessarily any relation corresponding to (34) for $\mathbf{M}^s$.

Second-order tensors such as $\mathbf{m}$ have the form

$$\mathbf{m}=\mathrm{diag}\ (m,\ m,\ m'). \tag{35}$$

It follows from (8) that

$$m=(1-\nu_s-\nu_s'')/E_s,\ m'=(1-2\nu_s)/E_s'. \tag{36}$$

Equation (14) gives

$$C=2(1-\nu-\nu'')/E+(1-2\nu')/E'. \tag{37}$$

The tensor $\mathbf{b}$ has the form

$$\mathbf{b}=\mathrm{diag}\ (b,\ b,\ b'), \tag{38}$$

where, in terms of the parameters A and B of Skempton (1954),

$$b=(1-A)B/2,\ b'=AB, \tag{39}$$

so that $B=2b+b'=\mathbf{b}:\boldsymbol{\delta}$.

The relations (19) can now be given in the explicit form

$$\frac{1}{E}-\frac{1}{E_u}=\left(\frac{1-\nu-\nu''}{E}-\frac{1-\nu_s-\nu_s''}{E_s}\right)b,$$

$$\frac{-\nu}{E}+\frac{\nu_u}{E_u}=\left(\frac{1-\nu-\nu''}{E}-\frac{1-\nu_s-\nu_s''}{E_s}\right)b,$$

$$\frac{-\nu'}{E'}+\frac{\nu_u'}{E_u'}=\left(\frac{1-\nu-\nu''}{E}-\frac{1-\nu_s-\nu_s''}{E_s}\right)b',$$

$$\frac{-\nu''}{E}+\frac{\nu_u''}{E_u}=\left(\frac{1-2\nu'}{E'}-\frac{1-2\nu_s'}{E_s'}\right)b,$$

$$\frac{1}{E'}-\frac{1}{E_u'}=\left(\frac{1-2\nu'}{E'}-\frac{1-2\nu_s'}{E_s'}\right)b', \tag{40}$$

together with

$$\frac{1}{G}-\frac{1}{G_u}=0,\ \frac{1}{G'}-\frac{1}{G_u'}=0. \tag{41}$$

The first of equations (41) is implied by the first two of equations (40). It allows E_u to be expressed in terms of E and the relevant Poisson's ratios. Also, the relation (34) (which applies also to $\mathbf{M}^u$), allows E_u' to be expressed in terms of E (or E') and Poisson's ratios. Thus,

$$E_u'=\frac{\nu_u'E_u}{\nu_u''}=\left(\frac{\nu_u'}{\nu_u''}\right)\left(\frac{1+\nu_u}{1+\nu}\right)E=\left(\frac{\nu''}{\nu_u''}\right)\left(\frac{\nu_u'}{\nu'}\right)\left(\frac{1+\nu_u}{1+\nu}\right)E'. \tag{42}$$

Calculation of $\boldsymbol{\alpha}$ from (28) is tedious because of the need to find the inverse $\mathbf{L}$ of $\mathbf{M}$. The result is that

$$\boldsymbol{\alpha}=\mathrm{diag}\ (\alpha,\ \alpha,\ \alpha'), \tag{43}$$

where

$$\alpha=\left\{1-\frac{E}{1-\nu-2\nu'\nu''}\left[\frac{(1-\nu_u-\nu_u'')}{E_u}+\nu'\,\frac{(1-2\nu_u')}{E_u'}\right]\right\}/(2b+b'),$$

$$\alpha'=\left\{1-\frac{E'}{1-\nu-2\nu'\nu''}\left[2\nu''\,\frac{(1-\nu_u-\nu_u'')}{E_u}+(1-\nu)\,\frac{(1-2\nu_u')}{E_u'}\right]\right\}/(2b+b'). \tag{44}$$

In view of the long expressions for $\boldsymbol{\alpha}$, there is some advantage in using the forms (31), (32) for ζ. First,

$$\mathbf{M}:\boldsymbol{\alpha}=\mathrm{diag}\left\{\frac{1}{E}[(1-\nu)\alpha-\nu''\alpha'],\ \frac{1}{E}[(1-\nu)\alpha-\nu''\alpha'],\ \frac{1}{E'}(\alpha'-2\nu'\alpha)\right\} \tag{45}$$

and hence,

$$\zeta = [(1-\nu)\alpha - \nu''\alpha'](\bar{\sigma}_{11}+\bar{\sigma}_{22})/E + (\alpha' - 2\nu'\alpha)\bar{\sigma}_{33}/E'$$
$$+ [2(1-\nu-\nu'')\alpha/E + (1-2\nu')\alpha'/E']p/(2b+b'), \quad (46)$$

having also used (34). Also, using (45), equation (32) can be given as

$$\zeta = \alpha(\bar{\epsilon}_{11}+\bar{\epsilon}_{22}) + \alpha'\bar{\epsilon}_{33}$$
$$- \{2\alpha[(1-\nu)\alpha - \nu''\alpha']/E + \alpha'(\alpha' - 2\nu'\alpha)/E'$$
$$- (2[(1-\nu)\alpha - \nu''\alpha']/E + (\alpha' - 2\nu'\alpha)/E')/(2b+b')\}p. \quad (47)$$

It is emphasized that a variety of different forms for (44), (46), and (47) could be derived, by making use of (34) and (40) as, for example, in (42). The interdependencies implied by (34) and (40) might perhaps also provide checks on experimental measurements.

5 Isotropy

When the overall structure of the porous medium is isotropic, primed variables are equal to corresponding unprimed ones and equations (40) yield

$$\frac{1-2\nu}{E} - \frac{1-2\nu_u}{E_u} = \left(\frac{1-2\nu}{E} - \frac{1-2\nu_s}{E_s}\right)B \quad (48)$$

and

$$\frac{1+\nu}{E} - \frac{1+\nu_u}{E_u} = 0, \quad (49)$$

where $B = 3b$, since $b' = b$. Equation (49) is also equivalent to (41). Equations (44) both imply

$$\alpha = \left\{1 - \frac{E(1-2\nu_u)}{E_u(1-2\nu)}\right\}/B \quad (50)$$

or, using (49)

$$\alpha = \frac{3(\nu_u - \nu)}{(1+\nu_u)(1-2\nu)B}. \quad (51)$$

This last form also follows easily from formulae given by Rice and Cleary (1976). Equation (46) simplifies to

$$\zeta = \frac{(1-2\nu)}{E}\alpha\{\bar{\sigma}_{11}+\bar{\sigma}_{22}+\bar{\sigma}_{33}+3p/B\}, \quad (52)$$

in agreement with Rice and Cleary (1976) and (47) gives

$$\zeta = \alpha\left\{\bar{\epsilon}_{11}+\bar{\epsilon}_{22}+\bar{\epsilon}_{33} - \frac{3}{B}(B\alpha - 1)\frac{(1-2\nu)}{E}p\right\}. \quad (53)$$

Rearrangement of (53) using $2G = E/(1+\nu)$ and (51) gives

$$p = \frac{2GB^2(1+\nu_u)^2(1-2\nu)\zeta}{9(\nu_u-\nu)(1-2\nu_u)} - \frac{2GB(1+\nu_u)}{3(1-2\nu_u)}(\bar{\epsilon}_{11}+\bar{\epsilon}_{22}+\bar{\epsilon}_{33}), \quad (54)$$

in agreement with Detournay (1986).

6 Concluding Remarks

The main results of this work are contained in equations (19) and (31) (or (32)) together with (28) which gives the tensor $\boldsymbol{\alpha}$ that appears in the effective stress principle. The equations should be useful in practice because they are expressed in terms of the tensors of compliances $\mathbf{M}$, $\mathbf{M}^u$ under drained and undrained conditions respectively, and the "Skempton tensor" $\mathbf{b}$, all of which are obtainable directly from experiments. The equations (19) yield relations between components of $\mathbf{M}$ and $\mathbf{M}^u$, whose verification might provide checks on experimental measurements. The micromechanical considerations on which the derivations are based introduce tensors $\mathbf{M}^s$, $\mathbf{m}^s$ which relate to the properties of the solid skeleton. This part of the analysis is related to that of Nur and Byerlee (1971) or Carroll (1979) in the case that the skeleton is homogeneous. The present analysis, however, is applicable to an inhomogeneous skeleton and hence also to incremental deformation of a nonlinear porous medium, in which the increment is superimposed upon a deformation which is inevitably inhomogeneous on the microscale.

Acknowledgment

Thanks are due to Dr. J. R. A. Pearson for helpful discussions and to a referee who made numerous suggestions for improving the presentation.

References

Biot, M. A., 1955, "Theory of Elasticity and Consolidation for a Porous Anisotropic Solid," *J. Appl. Phys.*, Vol. 26, pp. 182–185.

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.

Carroll, M. M., 1979, "An Effective Stress Law for Anisotropic Elastic Deformation," *J. Geophys. Res.*, Vol. 84, pp. 7510–7512.

Carroll, M. M., and Katsube, N., 1983, "The Role of Terzaghi Effective Stress in Linearly Elastic Deformation," *ASME Journal of Energy Resources Technology*, Vol. 105, pp. 509–511.

Detournay, E., 1986, "Poroelastic Solution of a Borehole in a Non-hydrostatic Stress Field," unpublished report.

Hill, R., 1963, "Elastic Properties of Reinforced Solids: Some Theoretical Principles," *J. Mech. Phys. Solids*, Vol. 11, pp. 357–372.

Nur, A., and Byerlee, J. D., 1971, "An Exact Effective Stress Law for Elastic Deformation of Rock with Fluids," *J. Geophys. Res.*, Vol. 76, pp. 6414–6419.

Mesri, G., Adachi, K., and Ullrick, C. R., 1976, "Pore Pressure Response in Rock to Undrained Change in all Round Stress," *Geotechnique*, Vol. 26, pp. 317–330.

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

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

Skempton, A. N., 1954, "The Pore Pressure Coefficients A and B," *Geotechnique*, Vol. 4, pp. 143–147.

APPENDIX

The relation (8) is proved by noting that, when the skeleton has compliances with the 'usual' symmetry, Betti's theorem implies

$$\int_{\partial\Omega_s} (\boldsymbol{\sigma}_1\cdot\mathbf{n})\cdot\mathbf{u}_2\,ds = \int_{\partial\Omega_s} (\boldsymbol{\sigma}_2\cdot\mathbf{n})\cdot\mathbf{u}_1\,ds, \quad (A1)$$

where $\partial\Omega_s$ denotes the boundary of Ω_s, $\mathbf{u}_1$ is the displacement associated with $\boldsymbol{\sigma}_1$ and $\mathbf{u}_2$ is the displacement associated with $\boldsymbol{\sigma}_2$. Now let $\boldsymbol{\sigma}_1$ be associated with $\bar{\boldsymbol{\epsilon}}_1$ and p_1, and $\boldsymbol{\sigma}_2$ be associated with $\bar{\boldsymbol{\epsilon}}_2$ and p_2, in the manner defined in Section 2. Thus, (A1) can be expanded to

$$\int_{\partial\Omega} (\boldsymbol{\sigma}_1\cdot\mathbf{n} + p_1\mathbf{n})\cdot(\bar{\boldsymbol{\epsilon}}_2\cdot\mathbf{x})\,ds - p_1\int_{\partial\Omega_s}\mathbf{n}\cdot\mathbf{u}_2\,ds$$
$$= \int_{\partial\Omega} (\boldsymbol{\sigma}_2\cdot\mathbf{n} + p_2\mathbf{n})\cdot(\bar{\boldsymbol{\epsilon}}_1\cdot\mathbf{x})ds - p_2\int_{\partial\Omega_s}\mathbf{n}\cdot\mathbf{u}_1\,ds, \quad (A2)$$

since $\boldsymbol{\sigma}_i\cdot\mathbf{n} + p_i\mathbf{n} = 0$ wherever $\partial\Omega$ bounds fluid. The relation (1) allows (A2) to be written

$$(\bar{\boldsymbol{\sigma}}_1 + p_1\boldsymbol{\delta}):\bar{\boldsymbol{\epsilon}}_2 - (1-\varphi)p_1\boldsymbol{\delta}:\bar{\boldsymbol{\epsilon}}_2^s$$
$$= (\bar{\boldsymbol{\sigma}}_2 + p_2\boldsymbol{\delta}):\bar{\boldsymbol{\epsilon}}_1 - (1-\varphi)p_2\boldsymbol{\delta}:\bar{\boldsymbol{\epsilon}}_1^s, \quad (A3)$$

the forms for the integrals over $\partial\Omega_s$ following from Gauss' theorem. Substituting for $\bar{\boldsymbol{\epsilon}}_i$ and $\bar{\boldsymbol{\epsilon}}_i^s$ using (2) and (9) reduces (A3) to

$$(\bar{\boldsymbol{\sigma}}_1:\mathbf{M}:\bar{\boldsymbol{\sigma}}_2 - \bar{\boldsymbol{\sigma}}_2:\mathbf{M}:\bar{\boldsymbol{\sigma}}_1)$$
$$+ (\boldsymbol{\delta}:\mathbf{M}:(p_1\bar{\boldsymbol{\sigma}}_2 - p_2\bar{\boldsymbol{\sigma}}_1) - (p_1\bar{\boldsymbol{\sigma}}_2 - p_2\bar{\boldsymbol{\sigma}}_1):\mathbf{M}:\boldsymbol{\delta})$$
$$+ (\mathbf{m} - \boldsymbol{\delta}:\mathbf{M}^s):(p_1\bar{\boldsymbol{\sigma}}_2 - p_2\bar{\boldsymbol{\sigma}}_1) = 0, \quad (A4)$$

and this implies both the symmetry of $\mathbf{M}$ and the relation (8), since p_i, $\bar{\boldsymbol{\sigma}}_i$ are arbitrary. Symmetry of $\mathbf{M}^s$ is not, however, guaranteed.

An Expression of Elastic-Plastic Constitutive Law Incorporating Vertex Formation and Kinematic Hardening

Moriaki Goya
Associate Professor,
Faculty of Engineering,
University of the Ryukyus,
Nishihara-Cho, Okinawa 903-01,
Japan
Assoc. Mem. ASME

Koichi Ito
Associate Professor,
Faculty of the Department of Precision Engineering,
Tohoku University,
Aramaki, Sendai, Miyagi 980,
Japan

A phenomenological corner theory was proposed for elastic-plastic materials by the authors in the previous paper (Goya and Ito, 1980). The theory was developed by introducing two transition parameters, $\mu(\alpha)$ and $\beta(\alpha)$, which, respectively, denote the normalized magnitude and direction angle of plastic strain increments, and both monotonously vary with the direction angle of stress increments. The purpose of this report is to incorporate the Bauschinger effect into the above theory. This is achieved by the introduction of Ziegler's kinematic hardening rule. To demonstrate the validity and applicability of a newly developed theory, we analyze the bilinear strain-path problem using the developed equation, in which, after some linear loading, the path is abruptly changed to various directions. In the calculation, specific functions, such as $\mu(\alpha) = Cos(.5\pi\alpha/\alpha_{max})$ and $\beta(\alpha) = (\alpha_{max} - .5\pi)\alpha/\alpha_{max}$, are chosen for the transition parameters. As has been demonstrated by numerous experimental research on this problem, the results in this report also show a distinctive decrease of the effective stress just after the change of path direction. Discussions are also made on the uniqueness of the inversion of the constitutive equation, and sufficient conditions for such uniqueness are revealed in terms of $\mu(\alpha)$, $\beta(\alpha)$ and some work-hardening coefficients.

1 Introduction

The classical J_2-flow theory is a widely used constitutive equation for plastic deformation analyses, while the J_2-deformation theory, which is interpreted as a constitutive equation with a singular point on the yield surface, may not be suitable for deformation analyses because of its failures in taking account of unloading and neutral loading. However, these theories result in quite contrary conclusions about the applicability to bifurcation analyses of shell-type structures: The classical flow theory gives unobservable predictions for buckling or necking problems of shell-type structures, while J_2-deformation theory gives rather realistic values. Motivated by this paradox, several new theories of plasticity have been discussed (Christoffersen and Hutchinson, 1979; Ito, 1982; Gotoh, 1985).

Christoffersen and Hutchinson (1979) have proposed a new phenomenological corner theory of plasticity, which takes advantage of the properties of a singular point on the yield surface. The theory is called J_2-corner theory, which was developed to agree with J_2-deformation theory for nearly proportional loadings. The J_2-corner theory has been used in a number of finite element method analyses (Tomita and Shindo, 1986; Tevergaard et al., 1981), since this theory is the first corner theory which has a general but applicable form and also incorporates a smooth transition to elastic unloading for increasingly nonproportional stress increments.

Gotoh (1985) has discussed, from the viewpoint of tensor algebra, a general form of plastic constitutive equations which ensures one-to-one correspondence between plastic strain increments and stress increments. The simplest form of Gotoh's theory was reduced to a corner theory, which is different from the one presented by Christoffersen and Hutchinson.

The above two theories, however, include somewhat and ambiguous transition parameters from the viewpoint of physical interpretation, such as $f(\Theta)$ in Christoffersen and Hutchinson's theory and $P(\Theta)$ in Gotoh's theory.

Using the Kröner-Budiansky-Wu model for polycrystalline metals, Ito has numerically studied a constitutive relation for stress paths abruptly changing their direction from a proportional loading direction. In the discussion of the calculated results, Ito introduced two new independent transition parameters such as $\mu(\alpha)$ and $\beta(\alpha)$. These have been defined as the parameters with some physical interpretation: $\mu(\alpha)$ determines the dependence of the normalized magnitude of plastic strain increments on an angular measure α, which denotes the direction angle of deviatoric stress increments and is defined to

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 10, 1988; final revision, Feb. 21, 1990.

be measured from the proportional loading direction, and $\beta(\alpha)$ determines the dependence of plastic strain increment directions on the angular measure α.

Recently the authors (1990) proposed a new simple corner theory for plasticity by assuming that the plastic strain increment is considered to be decomposed into two tensor components: the one projected onto the direction of the deviatoric stress increment and the other onto the direction normal to the plastic potential surface. It was shown that the two parameters, $\mu(\alpha)$ and $\beta(\alpha)$, could naturally be embodied in the proposed theory. The research elucidated that those two parameters were essential in describing the property of the corner theory. The sufficient conditions on $\mu(\alpha)$ and $\beta(\alpha)$ for ensuring the existence of the unique inversion of the constitutive equation are also revealed.

In this report, we develop the above theory to incorporate the Bauschinger effect, which must be considered for analyses of plastic deformations subject to reverse loading. The research on corner theory, which takes into account the Baushinger effect, seems to be scant. Only Tomita et al. (1986) have tried to develop J_2-corner theory, proposed by Christoffersen and Hutchinson, by introducing a virtual smooth surface, which was supposed to translate according to Ziegler's kinematic hardening rule during plastic loading and also to be activated as a subsequent yield surface after elastic unloading once occurs from a preceding plastic state. The developed law, however, seems to include somewhat uncertain points, since the discussions of the relationship between the initial yield surface and so introduced subsequent yield surfaces were not contained in the report.

All the stresses and strains discussed in this report are defined in the rectangular Cartesian coordinates system. Furthermore, the choice of the appropriate rate (or increment) measure, canceling the effect of material rotation, will be not alluded in the paper, since the authors believe that this rate measure problem is not perfectly solved yet, though several papers have been published concerning this problem (Dafalias, 1983).

2 General Description of Goya and Ito's Theory (1990)

A new constitutive equation was developed based on the fundamental assumption that the plastic potential exists and stress increments $\dot{\sigma}$ effect plastic strain increments $\dot{\epsilon}^p$. By the use of scalar weighting functions $K_F(\alpha,\beta)$ and $K_D(\alpha,\beta)$, the equation is expressed as an averaging rule incorporating both effects by plastic potential and deviatoric stress increments: the plastic increment $\dot{\epsilon}^p$ is assumed to consist of two different incremental strains as shown in the following expression:

$$\dot{\epsilon}^p = \dot{\epsilon}^p{}_D + \dot{\epsilon}^p{}_F \tag{1}$$

where

$$\dot{\epsilon}^p{}_D = K_D |\dot{\epsilon}^p| \dot{\sigma}'/|\dot{\sigma}'| \tag{2}$$

$$\dot{\epsilon}^p{}_F = K_F |\dot{\epsilon}^p| \sigma'/|\sigma'| \tag{3}$$

$$K_F = \sin(\alpha-\beta)/\sin\alpha \tag{4}$$

$$K_D = \sin\beta/\sin\alpha \tag{5}$$

$$\alpha \underline{\underline{d}} \cos^{-1}\{\dot{\sigma}':\sigma'/(|\dot{\sigma}'||\sigma'|)\} \tag{6}$$

$$\beta \underline{\underline{d}} \cos^{-1}\{\dot{\epsilon}^p:\sigma'/(|\dot{\epsilon}^p||\sigma'|)\}. \tag{7}$$

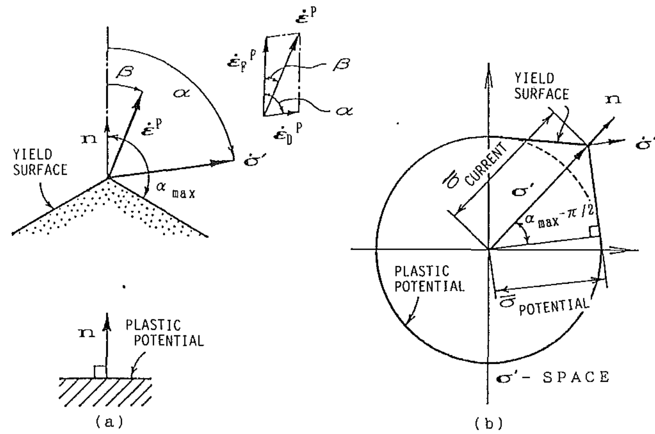


Fig. 1 Geometrical relations among a plastic potential, yield surface, angular measures, stress increment, and strain increments: (*a*) magnified view of the vicinity of a plastic loading point, (*b*) general view

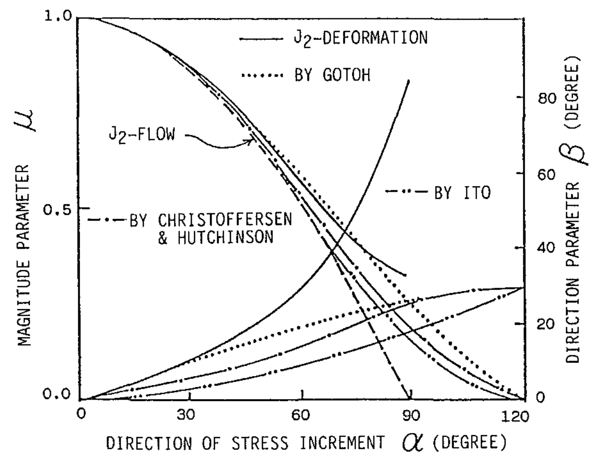


Fig. 2 Variation of $\mu(\alpha)$ and $\beta(\alpha)$ calculated from several already established theories

Figure 1(*a*) illustrates the fundamental relationships, in the vicinity of the loading point, among the strain increments, angular measures, yield surface, and plastic potential with a smooth surface. Note that the yield surface does not necessarily coincide with the plastic potential surface when the corner is formed at the loading point. Figure 1(*b*), however, shows that the plastic potential can be identical to the yield surface, as has been assumed in J_2-flow theory, when the current stress point is located exactly on or inside the plastic potential surface. As the main hypotheses introduced in the theory, the following relations are assumed between the stress increment and corresponding strain increment:

$$\bar{\dot{\sigma}} = H'_{tot}\bar{\dot{\epsilon}}^p/\mu(\alpha) \tag{8a}$$

$$\beta = \beta(\alpha) \tag{8b}$$

where $\bar{\dot{\sigma}} \underline{\underline{d}} \sqrt{(3/2)\dot{\sigma}':\dot{\sigma}'}$, $\bar{\dot{\epsilon}} \underline{\underline{d}} \sqrt{(2/3)\dot{\epsilon}':\dot{\epsilon}'}$ and $H'_{tot}(>0)$ is the slope of a stress-strain curve for proportional loading. Figure

Nomenclature

- σ = stress tensor
- ϵ = strain tensor
- δ = Kronecker delta
- α = translation of plastic potential
- ϵ_v = change of volume
- $\dot{(\ )}$ = increments of ( )
- $(\ )'$ = deviatoric component of ( )
- $|\ |$ = magnitude of tensors
- $:$ = inner product of tensors
- α = angular measure of stress increment direction
- β = angular measure of strain increment direction
- μ = nondimensional magnitude parameter of strain increments
- E = Young's modulus
- G = shear modulus
- K = bulk modulus
- ν = Poisson's ratio

Superscripts

- $(\ )^p$ = plastic component of ( )
- $(\ )^e$ = elastic component of ( )

2 shows the variations of $\mu(\alpha)$ and $\beta(\alpha)$ calculated from several established theories.

Then, assuming the Mises-type plastic potential and noting the relation

$$\bar{\dot{\sigma}}\cos\alpha = \dot{\bar{\sigma}}, \tag{9}$$

we can obtain, after some manipulation, the following equations about the two components of the plastic strain increment in Eq. (1):

$$\dot{\epsilon}^p{}_F = (3/2)K_F \frac{\sigma' \mu(\alpha)\dot{\bar{\sigma}}}{\bar{\sigma}\cos\alpha H'_{\mathrm{tot}}} \tag{10}$$

$$\dot{\epsilon}^p{}_D = (3/2)K_D \frac{\mu(\alpha)\dot{\sigma}'}{H'_{\mathrm{tot}}}. \tag{11}$$

Assuming that Hooke's law governs the relation between elastic strain increments and stress increments, and that a total strain increment consists of both elastic and plastic parts, we finally obtain the following constitutive equation for plastic loading:

$$\dot{\sigma} = 2G^* \left\{ \dot{\epsilon} + \left(\frac{K}{2G^*} - \frac{1}{3} \right) \dot{\epsilon}_v \delta - \frac{(\sigma' : \dot{\epsilon})\sigma'}{S^*} \right\} \tag{12}$$

where G ; elastic shear modulus

K ; elastic bulk modulus

$$G^* = GH'_{\mathrm{tot}} / (H'_{\mathrm{tot}} + 3GK_D\mu) \tag{13}$$

; instantaneous effective shear modulus

$$S^* = 2\bar{\sigma}^2 \{ H'_{\mathrm{tot}}\cos\alpha/(3G) + \mu\cos\beta \} / (3K_F\mu) \tag{14}$$

It has been proved that Eq. (12) ensures one-to-one correspondence between the stress increment and strain increment, if the transition parameters are assumed to satisfy the following conditions within the range $0 \leqq \alpha \leqq \alpha_{\max}$.

(i) $\mu(0) = 1$ and $\mu(\alpha_{\max}) = 0$, and
$\mu(\alpha)$ decreases monotonically with respect to α.

(ii) $\beta(0) = 0$ and $\beta(\alpha_{\max}) = \alpha_{\max} - \pi/2$, and
$\beta(\alpha)$ increases monotonically with respect to α.

As has been proposed by Christoffersen and Hutchinson, $\alpha_{\max}$ is defined through the geometrical relation between the radius of the plastic potential and the location of a current loading point, see Fig. 1(b). Then, the following equation holds:

$$\sin\alpha_{\max} = \frac{\bar{\sigma}(\text{potential})}{\bar{\sigma}(\text{current})}, \tag{15}$$

where it is assumed that the singular part of the yield surface forms an enveloping surface of hyperplanes, which are tangential to the plastic potential and also intersecting with each other at the current loading point, see Fig. 1(b). Therefore, when a stress increment or a strain increment causes elastic unloading from a current plastic stress state, the conical part of the yield surface reduces its vertex angle: The conical part of the yield surface is deformable during unloading, though the rest part of the yield surface is left unchanged.

The criteria have been revealed for judging whether an applied strain increment causes plastic loading or elastic unloading, and were written as follows:

(iii) If $\dfrac{\sigma' : \dot{\epsilon}'}{\bar{\sigma}\dot{\bar{\epsilon}}} \geqq \cos(\alpha_{\max})$ then plastic loading occurs,
and Eq. (12) should be used.

(iv) If $\dfrac{\sigma' : \dot{\epsilon}'}{\bar{\sigma}\dot{\bar{\epsilon}}} < \cos(\alpha_{\max})$ then elastic unloading occurs,
and Hooke's law should be used.

3 Derivation of Corner Theory Incorporating Kinematic-Hardening Rule

By the application of a kinematic-hardening rule to the transition of plastic potential, a new constitutive equation is directly derived from the corner theory briefly described in the preceding section.

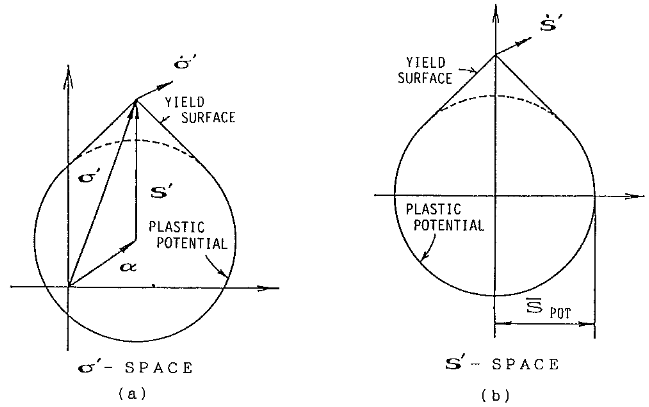


Fig. 3 Schematic of a translating plastic potential and a yield surface with vertex formation

In Fig. 3(a), a typical situation at a general stress state is illustrated for the plastic potential, yield surface, and stresses in the deviatoric stress space, namely σ'-space. Some fundamental ideas are drawn out through the consideration of geometrical relations among these quantities. To facilitate the incorporation of the kinematic-hardening rule with the corner theory, let us introduce a new stress defined by the following equation:

$$\mathbf{S} = \sigma - \alpha, \tag{16}$$

where α is a tensor denoting the translation of the plastic potential.

Figure 3(b) shows the schematic relations among some basic quantities in $\mathbf{S}'$-space, whose origin moves with the center of the translating plastic potential. For the translation of the origin, following Ziegler's kinematic rule is assumed in this report:

$$\dot{\alpha} = H'_k \dot{\bar{\epsilon}} \mathbf{S}/\bar{S}, \tag{17}$$

where $H'_k(\geqq 0)$ denotes the kinematic-hardening coefficient.

Let us assume that the corner theory discussed in Section 2 similarly holds in the $\mathbf{S}'$-space: The σ'-space is replaced by $\mathbf{S}'$-space in this section and $\mathbf{S}'$ plays the role of σ' in Section 2. Thus, we simply obtain the following equations for Eqs. (5)–(8).

$$\dot{\bar{S}} = H' \dot{\bar{\epsilon}}^p / \mu(\alpha) \tag{18}$$

$$\dot{\bar{S}} = \bar{\dot{S}}\cos(\alpha) \tag{19}$$

$$\dot{\epsilon}^p{}_F = (3/2)K_F \frac{\mathbf{S}' \mu(\alpha)\dot{\bar{S}}}{\bar{S}H'\cos\alpha} \tag{20}$$

$$\dot{\epsilon}^p{}_D = (3/2)K_D \frac{\mu(\alpha)\dot{\mathbf{S}}'}{H'}. \tag{21}$$

A total strain increment is expressed as the sum of elastic and plastic parts:

$$\dot{\epsilon}' = \dot{\epsilon}^{e\prime} + \dot{\epsilon}^p \tag{22}$$

with

$$\dot{\sigma}' = 2G\dot{\epsilon}^{e\prime}. \tag{23}$$

Substituting $\mathbf{S}' = \sigma' - \alpha'$ and Eq. (23) into Eq. (22), we obtain the following equations

$$\dot{\epsilon}' = \frac{\dot{\mathbf{S}}'}{2G^*} + \frac{(3K_F G + H'_k)\mu\dot{\bar{S}}}{2GH'_k\bar{S}} \mathbf{S}'. \tag{24}$$

Taking the inner product of the deviatoric stress $\mathbf{S}'$ to both sides of Eq. (24), then solving the result with respect to $\dot{\bar{S}}$, we obtain the following equation:

$$\dot{\overline{S}} = \frac{\mathbf{S}':\dot{\epsilon}'}{\overline{S}\left\{\dfrac{1}{2G^*}+\dfrac{\mu(K_FG+H'_k/3)}{GH'\cos\alpha}\right\}}. \tag{25}$$

The substitution of Eq. (25) into Eq. (24) yields:

$$\frac{\dot{\mathbf{S}}'}{2G^*} = \dot{\epsilon}' - \frac{3\mu(3K_FG+H'_k)\mathbf{S}'(\mathbf{S}':\dot{\epsilon}')}{2\overline{S}^2(H'\cos\alpha+\mu H'_k+3\mu G\cos\beta)}. \tag{26}$$

Now, substituting Eqs. (17) and (26) into the relation $\dot{\sigma}' = \dot{\mathbf{S}}' + \dot{\alpha}'$ and using Eq. (18), we finally obtain the following constitutive relation:

$$\dot{\sigma} = 2G^*\left\{\dot{\epsilon} + \frac{\nu^*}{1-2\nu}\dot{\epsilon}_v\delta - \frac{(\mathbf{S}':\dot{\epsilon})\mathbf{S}'}{\Sigma^*}\right\}, \tag{27}$$

where

$$\Sigma^* \underline{\underline{d}} \frac{2\overline{S}^2H'(H'\cos\alpha+\mu H'_k+3\mu G\cos\beta)}{9\mu G(K_FH'-\mu K_DH'_k)} \tag{28}$$

$$\nu^* \underline{\underline{d}} \nu + EK_D\mu/(2H'). \tag{29}$$

For the verification of the applicability of the Eq. (27), it is indispensable to ensure that, assuming a current stress is at a plastic loading state, we can judge whether a given total strain increment $\dot{\epsilon}$ causes plastic loading or elastic unloading. Furthermore, it must be proved that Eq. (27) gives a unique stress increment $\dot{\sigma}$ for the given strain increment $\dot{\epsilon}$: one-to-one correspondence must exist between strain and stress increments. These two requirements are deeply connected with each other and the details of the discussions are described in the Appendix. The conclusion for the first requirement can be described as follows:

If $\mathbf{S}':\dot{\epsilon}'/(\overline{S}\,\dot{\overline{\epsilon}}) \geqq \cos\alpha_{\max}$, then plastic loading is caused, however, if $\mathbf{S}':\dot{\epsilon}'/(\overline{S}\,\dot{\overline{\epsilon}}) < \cos\alpha_{\max}$, then elastic unloading is caused.

The uniqueness of Eq. (27) can be proved if the parameters, $\mu(\alpha)$ and $\beta(\alpha)$, satisfy the conditions (i) and (ii) described in Section 2, and $H' > H'_k$ holds. The last inequality condition is too sufficient to ensure the uniqueness comparing with the precise condition given in the Appendix, though it is useful due to its simplicity.

Some fundamental features of the developed theory will be revealed through the application to an uniaxial tension problem. Especially, it is necessary to investigate the mutual relationships among the work-hardening coefficients introduced in this report, since the coefficients have been used so far without discussing in the details. First, we can easily obtain the following relations for an uniaxial problem:

$$K_F+K_D=1,\ \overline{S}=S_1,\ S'_1=-2S'_2=-2S'_3=2S_1/3$$
$$G^*=GH'/(H'+3K_DG),$$
$$\Sigma^*=2S_1{}^2(H'+H'_k+3G)H'/\{9G(K_FH'-K_DH'_k)\}.$$

Substituting these relations into Eq. (27), then arranging the result, we obtain the following equation:

$$\dot{\sigma}_1 = (H'+H'_k)\dot{\epsilon}_1{}^p.$$

Thus, we know that the sum of H' and H'_k equals to the slope of $\sigma_1 - \epsilon_1{}^p$ curve obtained by the uniaxial tension test. Denoting the slope by H'_{tot}, we obtain the following relation:

$$H'_{\text{tot}} = H' + H'_k. \tag{30}$$

Furthermore, we can derive another relation concerning the coefficient H' by assuming several specific material properties. Let us suppose that a material has no vertex at the loading point on the yield surface, then the coefficient H' becomes perfectly identical to the value, denoted by H'_{pot}, for the plastic potential, which is expanded isotropically in $\mathbf{S}'$-space by added plastic work. If we suppose that the vertex is formed at the loading point on the yield surface, then the relation among the coefficients H', H'_{pot} and H'_v can be written as:

$$H' = H'_{\text{pot}} + H'_v, \tag{31}$$

where $H'_v(\geqq 0)$ is a coefficient in relation to the vertex formation.

From the analogy to J_2-flow theory, it is reasonable to assume following relation for the growth of the plastic potential due to plastic work-hardening:

$$\dot{\overline{S}}_{\text{pot}} = H'_{\text{pot}}\dot{\overline{\epsilon}}^p. \tag{32}$$

The kinematic corner theory, developed in this report, regards the yield surface as the one different from the plastic potential in the vicinity of a current plastic loading point (see Figs. 1 and 3). The yield surface can be singular at the plastic loading point, though the plastic potential is assumed to be regular wherever, and Ziegler's kinematic rule is assumed to the translation of the plastic potential. Substituting $\mu(0)=1$, the proportional loading condition, into Eq. (18) and comparing the results with Eq. (32), we easily notice that the emergence or the growth of the vertex is possible only for nearly proportional loading with positive H'_v.

4 Examination of the Developed Constitutive Equation

Several experimental and/or theoretical studies have been made on the problems of changing strain paths. Ohashi et al. (1981) have studied experimentally the constitutive properties of a brass subject to an abrupt change of strain-path direction after uniaxial tension. They revealed that the decrease of the effective stress $\overline{\sigma}$ was observed just after the change of direction, though a given strain increment, added after the change, still caused a plastic deformation. The same phenomena were similarly observed in other studies as well (Tokuda et al., 1986). The phenomena, however, are not explainable from the viewpoint of J_2-flow theory: In J_2-flow theory, the decrease of the effective stress $\overline{\sigma}$ immediately implies unloading and never causes plastic deformation.

To investigate the essential features of the developed constitutive equation without intermixing the effect of material rotation, we suppose that the stress principal axis be fixed relative to the material. Now, disregarding the question of choice of appropriate rate measure, we can analyze the above problem for the several paths illustrated in Fig. 4. It should be noticed that, in this problem, only pure-shear deformation is allowable due to the restriction on the rotation of the stress principal axis. In Fig. 4, the angular measure η determines changed directions of strain paths. The material is proportionally deformed in the first loading stage 0-A along the strain path of $\epsilon_{32}=0$ up to $\epsilon_{31}=0.1$. Then, at point A, the direction of the strain path is suddenly changed to a direction with $\epsilon_{32}\neq 0$. In the second loading stage A–B, the path is subjected to retain this direction. The effective stress and effective plastic strain are calculated by the following equations:

$$\overline{\sigma} = \{3(\sigma_{31}{}^2+\sigma_{32}{}^2)\}^{1/2} \tag{33}$$

$$\overline{\epsilon}^p = \int \dot{\overline{\epsilon}}^p. \tag{34}$$

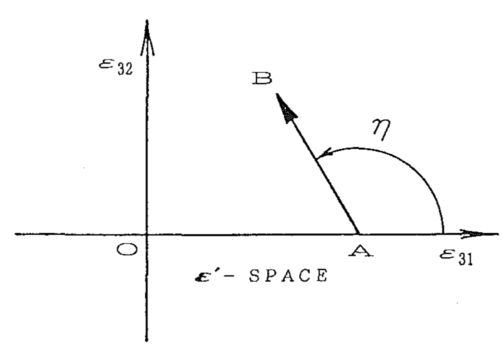


Fig. 4 A strain path subject to abrupt change of this direction

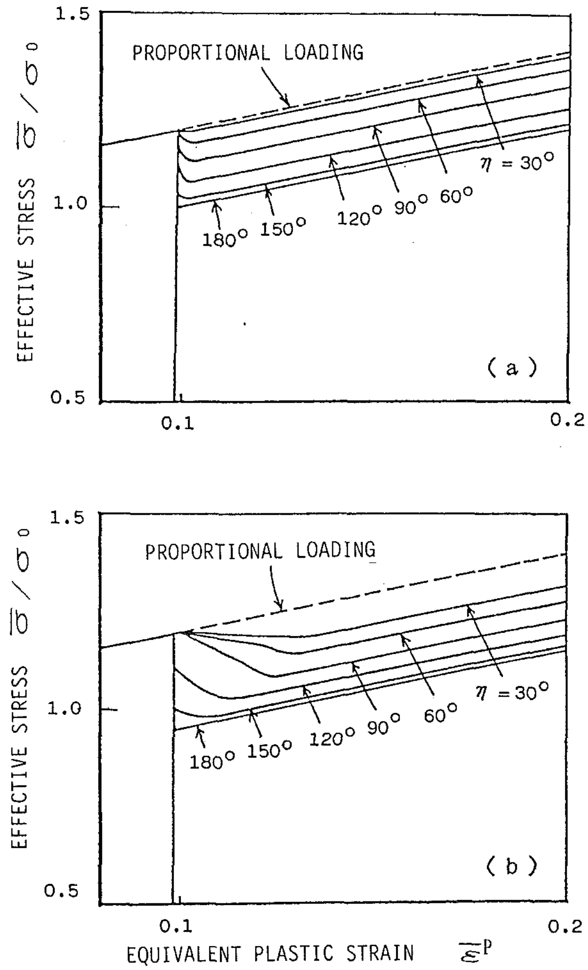


Fig. 5 Variation of effective stress for strain-path changing problems $\sigma_0 = 100$Mpa, $H'_{tot} = 200$MPa, $H' = 100$Mpa; (a) $H'_k = 100$MPa, $H'_v = 0$ MPa; only with kinematic hardening, (b) $H'_k = 100$MPa, $H'_v = 50$MPa; with both kinematic hardening and vertex formation

As the transition parameters, we choose the following functions due to their simplicity but also their good agreement with the data obtained from polycrystalline model analyses (Ito, 1982), though it is possible to choose any pair of functions that satisfy the conditions (i) and (ii):

$$\text{(i)} \quad \mu o(\alpha) = \cos\left(\frac{\pi\alpha}{2\alpha_{\max}}\right) \tag{35}$$

$$\text{(ii)} \quad \beta o(\alpha) = \frac{(\alpha_{\max} - \pi/2)\alpha}{\alpha_{\max}}. \tag{36}$$

The calculations are made by setting elastic constants as: Young's modulus $E = 75$GPa and Poisson's ration $\nu = 0.3$. To draw out essential features of the developed constitutive equation, let us suppose that all the work-hardening coefficients are constant: The material is assumed to harden according to the linear relation $\bar{\sigma} = \sigma o + H'_{\text{tot}}\bar{\epsilon}^p$ for proportional loading where σo is the initial yield stress. The following two sets of the work-hardening coefficients are chosen as typical examples:

(*a*) $H'_k \neq 0, H'_V = 0$; without vertex formation effect
(*b*) $H'_k \neq 0, H'_V \neq 0$; with both kinematic hardening and vertex formation effects.

Figure 5(*a*) illustrates the relations between the effective stress and effective plastic strain of the case (a) for various strain paths defined in Fig. 4. The plastic-hardening coefficients are chosen as: $H' = 100$MPa, $H'_k = 100$MPa and $H'_V = 0$ MPa. In the figure, the broken line shows the curve for proportional loading and the solid lines are those for the paths subject to abrupt change. The sudden decrease of $\bar{\sigma}$ can be clearly observed in any solid curves just after the change of the path. This means that the drop is explainable by only introducing the kinematic-hardening effect.

Figure 5(*b*) shows the curves for the case (b), where the hardening coefficients are chosen as $H' = 100$MPa, $H'_k = 100$MPa and $H'_V = 50$MPa. The results in Fig. 5(*b*) show that, when η is chosen as 30 deg or 60 deg, the slope just after the change of the direction becomes more gentle than the rest of the data obtained choosing other values for η. It can be said, therefore, that the vertex formation causes gradual decrease of the effective stress for the abrupt change of the strain path. The numerical results for the case where η is chosen as 120 deg, 150 deg, or 180 deg show the similar results as obtained in Fig. 4(*a*): This similarity is inferable, since we have assumed that the vertex can be diminished after the loading point once passes inside the plastic potential.

5 Concluding Remarks

The corner theory, previously proposed by Goya and Ito, was naturally developed to incorporate the Ziegler's kinematic-hardening rule.

For the verification of the applicability of the developed theory, the theory was applied to the problem of changing strain paths. A qualitative agreement with experimental data was demonstrated through numerical simulation. The sharp decrease of the effective stress $\bar{\sigma}$ was simulated by only introducing the kinematic-hardening effect. The vertex formation on the yield surface, therefore, can not be simply proved by the emergence of this decrease. It was demonstrated, however, that the vertex formation makes the slope of this decreasing part gentle for the relatively small change of the strain-path direction. This may be a useful information of experimentally confirm the formation of the vertex at the current loading point.

It should be noticed that the developed theory allows the vertex to disappear from the yield surface whenever the current loading point is on or within the plastic potential surface and then the yield surface becomes precisely identical to the plastic potential until the next emergence of the vertex. It was proved that the derived equation has a unique inversion, and the sufficient conditions for the uniqueness were given with simple expression in terms of $\mu(\alpha)$, $\beta(\alpha)$, H', and H'_k.

The developed corner theory can be interpreted as a naturally generalized expression of J_2-flow theory, J_2-deformation theory, and several other J_2-corner theories, since the developed theory can be coincident with any other established J_2-theory simply by choosing specific functions or values for the introduced parameters such as $\mu(\alpha)$, $\beta(\alpha)$, K_f, K_D, H'_k, and H'.

References

Christoffersen, J., and Hutchinson, W. J., 1979, "A Class of Phenomenological Corner Theories of Plasticity," *J. Mech. Phys. Solids*, Vol. 27, pp. 465–487.

Dafalias, Y. F., 1983, "Corotational Rates for Kinematic Hardening at Large Plastic Deformations," ASME Journal of Applied Mechanics, Vol. 50, pp. 561–565.

Gotoh, M., 1985, "A Class of Plastic Constitutive Equations with Vertex Effect—II. Discussions on the Simplest Form," *Int. J. Solids Structures*, Vol. 21, pp. 1117–1129.

Goya, M., and Ito, K., 1990, "An Expression of Constitutive Law of Elastic-Plastic Material—Initially Isotropic Materials with Mises Type Plastic Potential," *JSME International Journal*, Ser. I, Vol. 33, No. 1, pp. 57–63.

Ito, K., 1982, "Reduction of Elastic-Plastic Buckling Stress Due to Mobility of a Plastic Strain Rate Vector in the Flow Rule. Improvement of Plastic Buckling Theory in Problems of Plate and Shell," *J. Japan Society Tech. Plasticity*, Vol. 23, No. 256, pp. 410–417 (in Japanese).

Ohashi, Y., Tokuda, M., Suzuki, T., and Kurita, Y., 1981, "Stress-Strain Relation of Brass for the Plastic Deformation along Bi-Linear Strain Trajectories with Various Corner Angle," *Bulletin of the JSME*, Vol. 24, No. 197, pp. 1909–1915.

Tokuda, M., Ohno, N., and Krantochvil, J., 1986, "Unified Constitutive Equations for the Inelastic Behavior of Polycrystalline Metals, Based on a Semi-

Micro Approach," *Transactions of Japan Society of Mechanical Engineers* (Ser. A), Vol. 52, No. 475, pp. 758–764 (in Japanese).

Tomita, Y., Shindo, A., Kim, S. Y., and Michiura, K., 1986, "Deformation Behaviour of Elastic-Plastic Tubes Under External Pressure and Axial Load," *Int. J. Mech. Sci.*, Vol. 28, pp. 263–274.

Tvergaard, V., Needleman, A., and Lo, K. K., 1981, "Flow Localization in the Plane Strain Tensile Test," *J. Mech. Phys. Solids*, Vol. 29, No. 2, pp. 115–142.

APPENDIX

From Eq. (24), it follows that

$$\frac{\mathbf{S}':\dot{\epsilon}'}{\bar{S}\dot{\bar{S}}}=\left(\frac{\cos\alpha}{3G^*}+\mu\frac{K_fG+H'_k/3}{GH'}\right). \tag{A1}$$

By use of the definition $\mathbf{n}\underline{\underline{d}}\mathbf{S}'/|\mathbf{S}'|$, Eq. (A1) is written as

$$\frac{\mathbf{n}:\dot{\epsilon}'}{\dot{\bar{S}}}=\frac{\cos\alpha}{\sqrt{6}G^*}+\mu\frac{\sqrt{3}(K_fG+H'_k/3)}{\sqrt{2}GH'}. \tag{A2}$$

The substitution of the following relation

$$\cos\beta=K_D\cos\alpha+K_f \tag{A3}$$

and the definition of G^* into Eq. (A2) yields

$$\frac{\mathbf{n}:\dot{\epsilon}'}{\dot{\bar{\epsilon}}^p}=\frac{H'\cos\alpha}{\sqrt{6}\mu G}+\frac{\sqrt{3}\cos\beta}{\sqrt{2}}+\frac{H'_k}{\sqrt{6}G}. \tag{A4}$$

On the other hand, from Eq. (22), we get

$$(3/2)\dot{\bar{\epsilon}}^2=\dot{\epsilon}^{e\prime}:\dot{\epsilon}^{e\prime}+2\dot{\epsilon}^{e\prime}:\dot{\epsilon}'+(3/2)\dot{\bar{\epsilon}}^{p2}. \tag{A5}$$

Because of the relation $\dot{\sigma}'=\dot{\mathbf{S}}'+\dot{\alpha}'$, the elastic strain increment $\dot{\epsilon}^{e\prime}$ can be interpreted as the sum of two components, $\dot{\epsilon}^e_\alpha{}'$ and $\dot{\epsilon}^e_s{}'$, which are, respectively, defined as $\dot{\epsilon}^e_\alpha{}'\underline{\underline{d}}\dot{\alpha}'/(2G)$ and $\dot{\epsilon}^e_s{}'\underline{\underline{d}}\dot{\mathbf{S}}'/(2G)$.

Then, we get

$$\begin{aligned}\dot{\epsilon}^{e\prime}:\dot{\epsilon}^p&=(\dot{\epsilon}^e_\alpha{}'+\dot{\epsilon}^e_s{}'):\dot{\epsilon}^p\\&=\left(\frac{H'_k\dot{\bar{\epsilon}}^p}{\sqrt{6}G}\mathbf{n}+\frac{H'\dot{\bar{\epsilon}}^p}{\sqrt{6}G\mu}l\right):\dot{\epsilon}^p,\ \text{where } l\underline{\underline{d}}\dot{\mathbf{S}}'/|\dot{\mathbf{S}}'|\\&=(2G)^{-1}\{H'_k\cos\beta+(H'/\mu)\cos(\alpha-\beta)\}\dot{\bar{\epsilon}}^{p2}\end{aligned} \tag{A6}$$

and

$$\dot{\epsilon}^{e\prime}:\dot{\epsilon}^{e\prime}=\frac{1}{4G^2}\left(\frac{2H'^2}{3\mu^2}+\frac{4H'H'_k\cos\alpha}{3\mu}+\frac{2H'^2_k}{3}\right)\dot{\bar{\epsilon}}^{p2}. \tag{A7}$$

Substituting Eqs. (A6) and (A7) into Eq. (A5) and manipulating the results, we obtain the following equation:

$$\dot{\bar{\epsilon}}^p=\left[\frac{1}{9G^2}\left(\frac{H'^2}{\mu^2}+2\frac{H'H'_k\cos\alpha}{\mu}+H'^2_k\right)+\frac{2}{3G}\left\{\frac{H'\cos(\alpha-\beta)}{\mu}+H'_k\cos\beta\right\}+1\right]^{-1/2}\dot{\bar{\epsilon}}. \tag{A8}$$

The substitution of Eq. (A8) into Eq. (A4) yields

$$\frac{\mathbf{n}:\dot{\epsilon}'}{\sqrt{3/2}\,\dot{\bar{\epsilon}}}=F(\alpha)\ \text{or}\ \cos\Theta n\dot{\epsilon}'=F(\alpha), \tag{A9,10}$$

where

$$F(\alpha)\underline{\underline{d}}\left(\frac{H'\cos\alpha}{3\mu G}+\cos\beta+\frac{H'_k}{3G}\right)*\left[\frac{1}{9G^2}\left(\frac{H'^2}{\mu^2}+2\frac{H'H'_k\cos\alpha}{\mu}+H'^2_k\right)+\frac{2}{3G}\left\{\frac{H'\cos(\alpha-\beta)}{\mu}+H'_k\cos\beta\right\}+1\right]^{-1/2} \tag{A11}$$

and

$$\Theta n\dot{\epsilon}'\underline{\underline{d}}\cos^{-1}\left\{\frac{\mathbf{n}:\dot{\epsilon}'}{\sqrt{3/2}\,\dot{\bar{\epsilon}}}\right\}. \tag{A12}$$

From the conditions that $\mu(0)=1$, $\beta(0)=0$, $\mu(\alpha_{\max})=0$ and $\beta(\alpha_{\max})=\alpha_{\max}-\pi/2$, it is straightforward to show that

$$F=1\ \text{at}\ \alpha=0,\ \text{and}\ F=\cos\alpha_{max}\ \text{at}\ \alpha=\alpha_{max}. \tag{A13}$$

Therefore, from Eq. (A9), it is concluded that if $\alpha=0$, then $\Theta n\dot{\epsilon}'=0$ and if $\alpha=\alpha_{\max}$ then $\Theta n\dot{\epsilon}'=\alpha_{\max}$.

Furthermore, we need to show that $F(\alpha)$ is monotone decreasing with α within the range $0\leqq\alpha\leqq\alpha_{\max}$. Differentiating Eq. (A11) with respect to α we get

$$\frac{dF}{d\alpha}=\frac{F_{\text{num}}}{F_{\text{den}}}, \tag{A14}$$

where

$$F_{\text{den}}\underline{\underline{d}}\frac{1}{9G^2}\left(\frac{H'^2}{\mu^2}+2\frac{H'H'_k\cos\alpha}{\mu}+H'^2_k\right)+\frac{2}{3G}\left\{\frac{H'\cos(\alpha-\beta)}{\mu}+H'_k\cos\beta\right\}+1 \tag{A15}$$

$$\begin{aligned}F_{\text{num}}\underline{\underline{d}}&(H'\sin\alpha+3\mu G\sin\beta)\\&*[-H'\{H'+\mu H'_k\cos\alpha+3G\mu\cos(\alpha-\beta)\}\\&+(d\mu/d\alpha)\{3GH'\sin(\alpha-\beta)+H'_kH'\sin\alpha\}\\&-(d\beta/d\alpha)3G\mu\{3G\mu+\mu H'_k\cos\beta+H'\cos(\alpha-\beta)\}].\end{aligned} \tag{A16}$$

The numerator, F_{num}, is simply shown to be positive for any α. Now, to check the sign of $dF/d\alpha$, the consideration is required with respect to the sign of each term in the square brackets of Eq. (A16), since $(H'\sin\alpha+3\mu G\sin\beta)\geqq 0$ for any α.

From the first term,

$$-H'\{H'+\mu H'_k\cos\alpha+3G\mu\cos(\alpha-\beta)\}<-H'(H'-H'_k), \tag{A7}$$

since $H'>0$, $H'_k\geqq 0$, $0\leqq\mu\leqq 1$ and $0\leqq\alpha-\beta\leqq\pi/2$ for any α.

From the second term,

$$(d\mu/d\alpha)\{3GH'\sin(\alpha-\beta)+H'H'_k\sin\alpha\}\leqq 0 \tag{A18}$$

since $d\mu/d\alpha\leqq 0$ and $0\leqq\alpha-\beta\leqq\pi/2$ for any α.

From the third term,
since $d\beta/d\alpha\geqq 0$ and $0\leqq\beta\leqq\alpha_{\max}-\pi/2$ for any α.

Now, it is clear that $F(\alpha)$ is monotone decreasing with α within the range $0<\alpha<\alpha_{\max}$, if $H'>H'_k$:$F(\alpha)$ is a function with one-to-one correspondence with respect to α. Therefore, considering the relation given by Eqs. (A9), (A11) and just proved condition about $F(\alpha)$, the plastic loading range determined from α is identical to the one from $\Theta n\dot{\epsilon}'$. Then, we can conclude as follows:

If $\dfrac{\mathbf{n}:\dot{\epsilon}'}{\sqrt{3/2}\,\dot{\bar{\epsilon}}}\geqq\cos\alpha_{\max}$, then $\dot{\epsilon}'$ falls within plastic loading range.

If $\dfrac{\mathbf{n}:\dot{\epsilon}'}{\sqrt{3/2}\,\dot{\bar{\epsilon}}}<\cos\alpha_{\max}$, then $\dot{\epsilon}'$ falls within elastic unloading range.

Description of Nonproportional Cyclic Plasticity of Stainless Steel by a Two-Surface Model

Yukio Takahashi
Senior Research Engineer.

Takashi Ogata
Research Engineer.

Central Research Institute of Electric Power Industry,
Komae-shi, Tokyo 201, Japan

A simple elastic-plastic constitutive model based on the two-surface theory is developed to describe deformation behavior of austenitic stainless steels under multiaxial cyclic loading. Dependency of saturated stress range both on strain range and the proportionality of loading is considered. To establish a precise procedure for determination of material constants for nonproportional loading, the intervariable relation in the axial-torsional circular strain-path condition is studied in detail. A full procedure is then developed for determination of all material parameters. Finally, the effectiveness of the present model is demonstrated by application to axial-torsional cyclic tests for type 304 stainless steel at 550°C.

1 Introduction

Austenitic stainless steels (type 304 and type 316) are extensively used as structural materials in Liquid Metal Reactor (LMR) plants because of their various advantageous characteristics, especially their high-temperature strength. Therefore, the characterization of their mechanical behavior in the operating temperature range (200–550°C) of LMR plants becomes very important for the structural design of these plants. Behavior under cyclic loadings is of special importance because the principal concern in the structural design of LMR plants is the failure caused by repeated cycles of thermal transient loadings.

The results of extensive research carried out in the past indicate that these materials exhibit various complicated behaviors in these temperature ranges. They show significant cyclic hardening in uniaxial isothermal tests at these temperatures, in contrast to a small amount of hardening at room temperature. Moreover, the stress range at a stabilized cycle in constant strain-range cyclic tests increases with the imposed strain range. Behavior under nonisothermal conditions is also so complicated that straightforward application of the constitutive models developed under isothermal conditions often gives inaccurate results (Ohno et al., 1988). Finally, in the multiaxial stress state, it has been observed in many tests that the materials show different hardening behavior in nonproportional loading than that in proportional loading. The effect of this should be properly accounted for because the loadings given to each part of the real components are not proportional, although the deviation from proportionality could largely depend on the specific condition of the structures.

Various kinds of nonproportional cyclic loading tests were conducted recently using axial-torsional loading systems with thin-wall cylinder specimens. Krempl et al. (1984) conducted cyclic loading tests for type 304 stainless steel at room temperature controlling axial strain and torsional strain in both in-phase and out-of-phase modes. Similar experiments were conducted by Cailletaud et al. (1984) on 316L stainless steel at room temperature. McDowell (1985) performed cyclic loading tests on type 304 stainless steel at room temperature with various strain paths in axial-torsional strain space, which include circular, elliptical paths and more irregular ones consisting of several linear segments. Similar tests were carried out by Benallal and Marquis (1987) for type 316 stainless steel at room temperature. Tanaka et al. (1985) also studied the cyclic hardening behavior of type 316 stainless steel at room temperature using several kinds of paths specified in plastic (not total) strain space.

All these experiments, for either type 304 or type 316 stainless steel, show a similar trend and clearly indicate that cyclic hardening behavior largely depends on the proportionality of the loadings. It was also generally observed in these tests that the circular strain path defined in the equivalent strain plane yields the greatest amount of cyclic hardening among the various types of strain paths with the same equivalent strain ranges. Based on this finding, Murakami et al. (1987) conducted a series of uniaxial cyclic tests and circular strain-path tests for type 316 stainless steel at several temperatures between room temperature and 700°C. Their study shows that the difference between stress ranges at a stabilized cycle in these two kinds of tests is significant throughout the temperature range expected for structural components of LMR plants.

As for the uniaxial cyclic behavior of these stainless steels, many research works have been done on modeling with elastic-plastic or elastic-viscoplastic constitutive equations. Some recent models (Ohno and Kachi, 1986; Nouailhas, 1987) have the capability of representing deformation behavior under uniaxial cyclic loading, including cyclic hardening processes and stress-strain hysteresis shape in each cycle, with fairly good

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, July 29, 1988; final revision, May 9, 1990.

accuracy. However, these models yield poor predictions for multiaxial cyclic loading if there is some degree of nonproportionality due to ignorance of the above-mentioned characteristic of the materials. This fact has led several investigators to the effort of improving constitutive models for better representation of the material behavior under nonproportional cyclic loadings.

McDowell (1985a) made a detailed examination of the experimental stress-strain relations at the stabilized state in several axial-torsional tests with different strain paths and showed that the Mroz-type kinematic hardening rule (Mroz, 1967) gives the best description of the rate as well as the direction of movement of the yield surface compared with other rules. Then he constructed, based on this observation, one set of elastic-plastic constitutive equations based on the two-surface theory for describing the cyclic hardening behavior of the material under nonproportional cyclic loadings (McDowell, 1985b).

Tanaka et al. (1986a,b) developed an elastic-plastic constitutive model for nonproportional cyclic loading by modifying the model of Ohno and Kachi (1986). Their model also uses the two-surface theory with two kinds of cyclic nonhardening regions assumed in the plastic strain space.

Benallal and Marquis (1987) proposed a modification of the elastic-viscoplastic constitutive model of Chaboche (1977) based on their experimental results. In their model cyclic hardening is expressed by the expansion of the yield surface toward an asymptotic value which changes according to the degree of nonproportionality of the loading.

These new models give much better predictions of the hardening behavior under nonproportional loading than the classical models. However, further work is necessary for more detailed evaluation of the capability of these models. At the same time the possibility of simpler formulations should be continuously sought in order to make practical application easier.

In this paper, a simple elastic-plastic constitutive model will be presented with an example of successful application to type 304 stainless steel at a high temperature. The elastic-plastic constitutive model presented by Ohno and Kachi (1986) was used as a starting point for development, and several modifications were made to obtain the desired accuracy. The present model employs the two-surface theory to describe the nonlinear stress-strain relation during one cycle. Cyclic hardening is represented mainly by expansion of the bounding surface. Dependency of the stress range at the saturated state on the strain range is accounted for by using a progressively growing surface assumed in the plastic strain space.

Following the establishment of the fundamental structure of the constitutive equations, an intervariable relation existing at the steady-state cycle in the circular strain-path condition was examined in detail. This examination led to an exact procedure for determining the material parameters relevant to nonproportional loading. A detailed procedure was then developed for determination of all material parameters from the results of cyclic uniaxial and circular strain-path loading tests. Finally, the present model was applied to type 304 stainless steel at 550°C.

2 Constitutive Model

2.1 Background. The present model is based on the model proposed by Ohno and Kachi (1986) and utilizes the concepts of the two-surface model and the cyclic nonhardening region. The idea of the two surface model was initially proposed by Krieg (1975) as well as Dafalias and Popov (1975). The use of this model enables one to describe the nonlinearity of the stress-strain relation during cyclic loadings without ad hoc rules. It should be noted that the nonlinear kinematic hardening model presented by Chaboche (1977) has a very strong similarity to the two-surface model as pointed out in (Chaboche and Rousselier, 1983) and (Benallal and Marquis, 1987). It can be easily shown that both models become completely equivalent in a special case.

Cyclic hardening can be represented by the expansion of the bounding surface and the yield surface in the two-surface theory. The expansion of the former surface leads to a general increase of the stress range and that of the latter surface brings about the expansion of the purely elastic domain. Ohno et al. (1988) showed that the stress-strain relation in uniaxial cyclic loadings can be described with good accuracy for type 304 stainless steel at various temperatures by only the expansion of the bounding surface.

The cyclic nonhardening region defined in the plastic strain space is the most important concept developed by Ohno (1982). The use of this concept enables one to describe cyclic hardening behavior under irregular loading with more reality than when using the simpler maximum plastic strain range criterion previously suggested in (Chaboche et al., 1979).

2.2 Description of the Present Model. The present model assumes both the yield (loading) surface and the bounding surface to be hyperspheres in deviatoric stress space. They are represented as

$$f=(3/2)(s_{ij}-\alpha_{ij})(s_{ij}-\alpha_{ij})-\kappa^2, \tag{1}$$

$$F=(3/2)(s_{ij}-\alpha^*_{ij})(s_{ij}-\alpha^*_{ij})-\kappa^{*2}, \tag{2}$$

where s_{ij} is the stress deviator, α_{ij} is the center of the yield surface, and α^*_{ij} the center of the bounding surface. κ and κ^* represent the sizes of the yield surface and the bounding surface, respectively. The image point on the bounding surface is given according to the following Mroz-type mapping rule:

$$s^*_{ij}=\alpha^*_{ij}+(s_{ij}-\alpha_{ij})\kappa^*/\kappa. \tag{3}$$

It is assumed that the evolution of α_{ij}, α^*_{ij}, κ and κ^* takes place according to the following equations in which superimposed dots represent the derivatives with respect to time:

$$\dot{\alpha}_{ij}=A(\delta)\dot{p}(s^*_{ij}-s_{ij})/\delta, \quad \delta=\sqrt{(3/2)(s^*_{ij}-s_{ij})(s^*_{ij}-s_{ij})}, \tag{4}$$

$$\dot{\alpha}^*_{ij}=(2/3)K\dot{\epsilon}^p_{ij}, \tag{5}$$

$$\kappa=\kappa_0+\hat{\kappa}, \ \dot{\hat{\kappa}}=C(Q(\rho,q)-\hat{\kappa})\dot{p}, \tag{6}$$

$$\kappa^*=\kappa^*_0+\hat{\kappa}^*, \ \dot{\hat{\kappa}}^*=C^*(Q^*(\rho,q)-\hat{\kappa}^*)\dot{p}, \tag{7}$$

$$\dot{p}=\sqrt{(2/3)\dot{\epsilon}^p_{ij}\dot{\epsilon}^p_{ij}}. \tag{8}$$

Equation (4) means that the center of the yield surface moves in the direction of $(s^*_{ij}-s_{ij})$ by the rate $A(\delta)\dot{p}$ depending on that norm, δ. The functional form of $A(\delta)$ can be determined on the basis of the observation of stress-strain behavior as shown in the later section. Equations (6) and (7) mean that the sizes of the yield surface and bounding surface change with plastic strain, approaching the values of $\kappa_0+Q(\rho,q)$ and $\kappa^*_0+Q^*(\rho,q)$, respectively. $\kappa_0+Q(\rho,q)$ and $\kappa^*_0+Q^*(\rho,q)$ represent the sizes of these surfaces at the stabilized state in cyclic loading with plastic strain range, $\Delta\epsilon^p=2\rho$ and the proportional coefficient, q.

The hardening index surface, whose definition is the same as that of the cyclic nonhardening region (Ohno, 1982) but used in a somewhat different way, is defined as a hypersphere in plastic strain space as follows:

$$g=(2/3)(\epsilon^p_{ij}-\beta_{ij})(\epsilon^p_{ij}-\beta_{ij})-\rho^2. \tag{9}$$

Here, β_{ij} and ρ represent the center and the size of the surface and they develop according to the following equations:

$$\dot{\beta}_{ij}=\sqrt{3/2}(1-c)\Gamma\dot{p}\nu_{ij}, \quad \nu_{ij}=\sqrt{2/3}(\epsilon^p_{ij}-\beta_{ij})/\rho, \tag{10}$$

$$\dot{\rho}=c\Gamma\dot{p}, \tag{11}$$

$$\Gamma=\begin{cases}\sqrt{2/3}\nu_{ij}\dot{\epsilon}^p_{ij}/\dot{p} & g=0 \text{ and } \nu_{ij}\dot{\epsilon}^p_{ij}\geqq 0\\ 0 & g<0 \text{ or } \nu_{ij}\dot{\epsilon}^p_{ij}<0,\end{cases} \tag{12}$$

where c is a constant which determines the rate of the expansion of the surface. A large c value accelerates the cyclic hardening process, although the saturated state is not affected by the choice of this value, provided that the value is not too large.

The variable q is used for representing the effect of the proportionality of the loading. q becomes 1 or 0 for the proportional loading or circular strain-path loading condition, respectively, and assumes an intermediate value for other general loadings. Various expressions are possible to satisfy this requirement, and the best one should be chosen based on extensive study on the cyclic hardening behavior under various loading conditions. In the literature, various proposals have been made by several investigators for the expression of similar parameters in their models to describe the difference in hardening behavior under different loading conditions (McDowell, 1985b; Tanaka et al., 1986a,b; Benallal and Marquis, 1987). In this study, however, no evaluation was made concerning the effectiveness of these expressions since we only dealt with two limit states, i.e., uniaxial and circular strain-path loadings, where the choice of the expression makes no difference.

Finally, the model assumes the conventional normality rule in order to determine the direction of plastic strain increments:

$$\dot{\epsilon}_{ij}^{p} = \lambda (s_{ij} - \alpha_{ij}), \tag{13}$$

where λ is a scaler variable to be determined from the consistency condition $f = 0$ and the elastic stress-strain relationship.

2.3 Comparison With Other Models. As mentioned in the Introduction, several constitutive models have been proposed for the same purpose as the present model. A brief discussion will be given here concerning the similarities and differences between the present model and others.

(*a*) *Model of McDowell* (*1985b*). The two-surface theory is used in the McDowell model as in the present model. Although the two models differ in various detailed aspects, especially in the forms of the kinematic hardening functions, the most noticeable difference can be found in the means used to account for the strain-range dependency of cyclic hardening. McDowell used the maximum plastic strain-range criterion modified by introduction of a fading memory term to cope with the hardening behavior under variable strain-range cycling. On the other hand, the hardening index surface, which is fundamentally the same as the cyclic nonhardening region proposed by Ohno (1982), is used in the present model for the same purpose. However, it is not clear whether a reasonable prediction can be made by the McDowell's modification for a variety of strain histories, even in the uniaxial stress state, for which the effectiveness of the cyclic nonhardening region has been presented (Ohno, 1982; Ohno and Kachi, 1986).

(*b*) *Model of Tanaka et al.* (*1986a,b*). This model utilizes the concepts of the two-surface theory and the cyclic nonhardening region. In their model, it is assumed that the sizes of the yield and the bounding surfaces directly depend on the size of the cyclic nonhardening regions. Hence, the use of two cyclic nonhardening regions (called proportional and perfect nonhardening regions) at the same time is required for distinguishing cyclic hardening behavior under proportional and nonproportional loadings, which makes the model structure rather complicated. Indirect dependence of isotropic hardening on the size of the hardening index surface assumed in the present model (see equations (6) and (7)) makes the use of two regions unnecessary and results in a simpler structure of the constitutive model.

(*c*) *Model of Benallal and Marquis* (*1987*). Although formally expressed in a different way, time-independent limit of their model is equivalent to a particular form of the two surface model, in which the function $A(\delta)$ is a linear function of δ and $K = 0$ in the notation in the present study. These assumptions make the model simpler, but at the same time they limit the capability of the description of the stress-strain relation to some extent. It was not intended in their model to represent the strain-range dependency of cyclic hardening, although this can be easily done by assuming some variables to be functions of the size of the hardening index surface as well as of the proportionality coefficient, as is done in the present model.

3 Analysis of the Intervariable Relation at Stabilized Circular Strain-Path Cycle

Here, special attention is given to the intervariable relation existing at the stabilized cycle in circular strain-path loading, where the axial strain ϵ_{zz} and the engineering shear strain γ_{rz} change according to the following equations:

$$\epsilon_{zz} = (\Delta\epsilon_{zz}/2)\cos(\omega t), \tag{14}$$

$$\gamma_{rz} = (\Delta\gamma_{rz}/2)\cos(\omega t - \pi/2), \tag{15}$$

where ω is angular frequency of the loading cycle, t is the time, and the ratio of the amplitudes $\Delta\gamma_{rz}/\Delta\epsilon_{zz} = \sqrt{3}$.

For convenience, the following vectors are defined:

$$\boldsymbol{\epsilon} = (\epsilon_1 = \epsilon_{zz},\ \epsilon_2 = \gamma_{rz}/\sqrt{3}) = \boldsymbol{\epsilon}^e + \boldsymbol{\epsilon}^p, \tag{16}$$

$$\boldsymbol{\epsilon}^e = (\epsilon_1^e = \epsilon_{zz}^e,\ \epsilon_2^e = \gamma_{rz}^e/\sqrt{3}), \tag{17}$$

$$\boldsymbol{\epsilon}^p = (\epsilon_1^p = \epsilon_{zz}^p,\ \epsilon_2^p = \gamma_{rz}^p/\sqrt{3}), \tag{18}$$

$$\boldsymbol{\sigma} = (\sigma_1 = \sigma_{zz},\ \sigma_2 = \sqrt{3}\tau_{rz}), \tag{19}$$

$$\boldsymbol{\alpha} = (\alpha_1 = (3/2)\alpha_{zz},\ \alpha_2 = \sqrt{3}\alpha_{rz}), \tag{20}$$

$$\boldsymbol{\alpha}^* = (\alpha_1^* = (3/2)\alpha_{zz}^*,\ \alpha_2^* = \sqrt{3}\alpha_{rz}^*). \tag{21}$$

In the circular strain-path condition, the trace of $\boldsymbol{\epsilon}$ forms a circle of the diameter, $\epsilon_a = \Delta\epsilon_{zz} = \Delta\gamma_{rz}/\sqrt{3}$.

If one assumes the Poisson's ratio of 0.5 for convenience (in the actual calculation example shown in Section 5, 0.3 was assumed), Hooke's law can be simply stated as

$$\boldsymbol{\sigma} = E\boldsymbol{\epsilon}^e, \tag{22}$$

where E is the Young's modulus. It should be noted that under this condition, all the vectors defined above have a constant norm throughout the cycle, i.e., form a complete circular trace, in the stabilized condition ($\dot{\kappa} = \dot{\kappa}^* = 0$).

Equation (21) leads to the following equation:

$$E\boldsymbol{\epsilon} = \boldsymbol{\sigma} + E\boldsymbol{\epsilon}^p. \tag{23}$$

The normality rule requires

$$\dot{\boldsymbol{\epsilon}}^p = \lambda(\boldsymbol{\sigma} - \boldsymbol{\alpha}). \tag{24}$$

Considering $\dot{\boldsymbol{\epsilon}}^p \boldsymbol{\epsilon}^p = 0$ due to the circular movement of $\boldsymbol{\epsilon}^p$,

$$\boldsymbol{\epsilon}^p \bullet (\boldsymbol{\sigma} - \boldsymbol{\alpha}) = 0. \tag{25}$$

Here, $\mathbf{a} \bullet \mathbf{b}$ means an inner product of two vectors, $\mathbf{a}$ and $\mathbf{b}$. The yield condition can be stated as

$$|\boldsymbol{\sigma} - \boldsymbol{\alpha}| = \kappa. \tag{26}$$

Here, $|\mathbf{a}|$ means a norm of the vector, $\mathbf{a}$.

The image point on the bounding surface is given by

$$\boldsymbol{\sigma}^* = \boldsymbol{\alpha}^* + \kappa^*(\boldsymbol{\sigma} - \boldsymbol{\alpha})/\kappa. \tag{27}$$

Here, $\boldsymbol{\alpha}^*$ is given simply as

$$\boldsymbol{\alpha}^* = K\boldsymbol{\epsilon}^p. \tag{28}$$

Finally, movement of the center of the yield surface is described by

$$\dot{\boldsymbol{\alpha}} = A(\delta)\,\frac{\boldsymbol{\sigma}^* - \boldsymbol{\sigma}}{\delta}\,|\dot{\epsilon}_p|, \quad \delta = |\boldsymbol{\sigma}^* - \boldsymbol{\sigma}|. \tag{29}$$

The condition $\dot{\boldsymbol{\alpha}} \bullet \boldsymbol{\alpha} = 0$ requires that

$$\boldsymbol{\alpha} \bullet (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}) = 0. \tag{30}$$

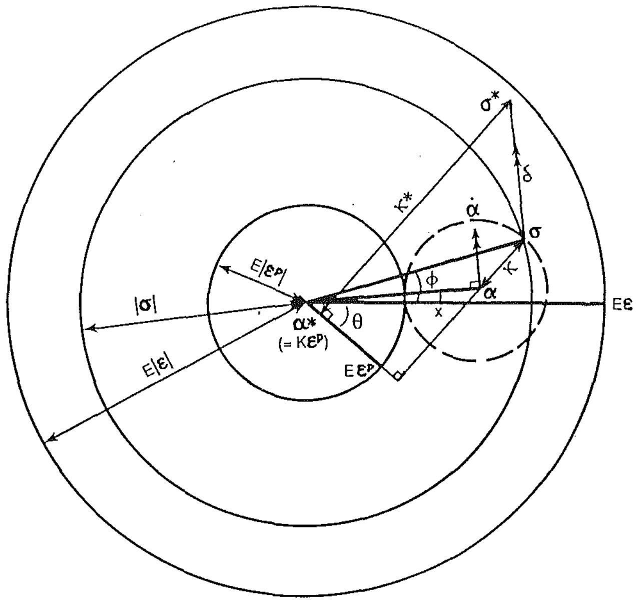


Fig. 1 Intervariable relation in steady-state circular strain-path condition

The relation of these vectors can be obtained graphically using three angles, ϕ, θ, and x as shown in Fig. 1. Here these angles give the phase differences of the stress, the center of the yield surface and the plastic strain vectors against total strain vectors, respectively. Afterwards, scaler symbols without subscripts will be used to represent the norms of various vector variables.

From equation (23), ϕ and θ can be easily obtained as

$$\phi = \cos^{-1}\{(\epsilon^2 + \epsilon^{e2} - \epsilon^{p2})/2\epsilon\epsilon^e\}. \tag{31}$$

$$\theta = \cos^{-1}\{(\epsilon^2 + \epsilon^{p2} - \epsilon^{e2})/2\epsilon\epsilon^p\}. \tag{32}$$

Consideration of equations (25) and (26) leads to the following equations:

$$\kappa = \sigma\sin(\theta + x)/\cos(\theta + x), \tag{33}$$

$$\alpha = \kappa\cos(\theta + \phi)/\sin(\phi - x). \tag{34}$$

Moreover, the following equations can be derived from equations (26), (28), and (30).

$$\kappa^* = \sigma\cos(\phi - x)/\sin(\theta + x) - K\epsilon^p/\tan(\theta + x), \tag{35}$$

$$\delta = |\sigma^* - \sigma| = \{\sigma\cos(\theta + \phi) - K\epsilon^p\}/\sin(\theta + x). \tag{36}$$

Finally, equation (29) can be used to derive the following scaler relation:

$$\alpha = A(\delta)\epsilon^p. \tag{37}$$

Equations (33)–(37) can be used for obtaining the values of κ and κ^* as well as x if the function for the hardening modulus $A(\delta)$ is given. Since these equations are related to each other, some iteration procedure needs to be used for the general functional forms of $A(\delta)$. One method which is used in this study is shown in Table 1.

These equations are quite general so that they can be used for determination of the material constants of any other constitutive models if they are based on the two-surface theory or an equivalent theory. In particular, it is expected that consideration of the above relationship will produce a better method of constants determination than the method described by Tanaka et al. (1986b) for their constitutive model.

4 Procedure for Determination of Material Constants and Functions

The present model has six material constants and three functions to be determined. The six constants are κ_0, κ_0^*, K, C, C^*, and c whereas the three functions are $A(\delta)$, $Q(\rho,q)$, and $Q^*(\rho,q)$. A general procedure for their determination from the test results of the constant strain range tests under uniaxial and circular strain-path conditions will be given below. For the function $Q(\rho,q)$ and $Q^*(\rho,q)$, only a procedure for determination of the functional forms at $q = 0$ and $q = 1$ will be shown in this paper. Of course, the simplest forms of these functions are given as follows:

Table 1 Iteration procedure for determination of *x*

1)	Assume $x = 0$
2)	$\kappa = \sigma\sin(\phi - x)/\cos(\theta + x)$
3)	$\delta = \lvert \underset{\sim}{\sigma} - \underset{\sim}{\sigma}^* \rvert = \{\sigma\cos(\phi+\theta) - K\epsilon^p\}/\sin(\theta + x)$
4)	$\alpha = A(\delta)\epsilon^p$
5)	$x_1 = \phi - \sin^{-1}\{\kappa\cos(\phi+\theta)/\alpha\}$
6)	If $\lvert x - x_1 \rvert \leq 10^{-3}$ stop
	If $\lvert x - x_1 \rvert > 10^{-3}$ $x = (x + x_1)/2$ go to 2)

$$Q(\rho,q) = (1-q)Q(\rho,0) + qQ(\rho,1), \tag{38}$$

$$Q^*(\rho,q) = (1-q)Q^*(\rho,0) + qQ^*(\rho,1). \tag{39}$$

(i) Determination of κ_0: The initial size of the yield surface κ_0 can be determined most easily as a point of departure from the linear stress-strain relationship in initial loading although some ambiguity exists depending on the precision of the measurements.

(ii) Determination of K, $Q(\rho,1)$, and $A(\delta)$: At the stabilized cycle in the constant strain range uniaxial cyclic tests, ρ becomes equal to $\Delta\epsilon^p/2$, where $\Delta\epsilon^p$ represents the plastic-strain range observed at the stabilized cycle. Moreover, κ and κ^* can be given as

$$\kappa = \kappa_0 + Q(\Delta\epsilon^p/2, 1), \tag{40}$$

$$\kappa^* = \kappa_0^* + Q^*(\Delta\epsilon^p/2, 1). \tag{41}$$

κ, at each stabilized condition, can be estimated as the half of the stress range of the linear portion in the hysteresis loops. Then the function $Q(\rho,1)$ can be determined by fitting each value of $Q(\Delta\epsilon^p/2,1) = \kappa - \kappa_0$ as a function of $\rho = \Delta\epsilon^p/2$.

The following equations are useful for determining the value of K and the function $A(\delta)$ from the relation between the axial stress, σ, and the axial plastic strain, ϵ^p, in the stabilized cycle.

$$H' = d\sigma/d\epsilon^p = A(\delta), \; \delta = |\mathrm{sgn}(\dot{\sigma})\kappa^* + K\epsilon^p - \sigma|. \tag{42}$$

Based on the plotting of the relation between H and $(\sigma - K\epsilon^p)$ for several cases of the strain range, we can optimize the value of κ^* for each value of $\Delta\epsilon^p$ as well as the value of K and the function $A(\delta)$. More detailed explanation of the procedure used in this study will be given in the next section.

(iii) Determination of κ_0^*, C, C^*, c, and $Q^*(\rho, 1)$: Next, the values of κ_0^*, C, C, and c can be determined based on the experimental data on the cyclic hardening process in the constant strain range uniaxial cyclic tests. The following knowledges are helpful for determining these constants.

(*a*) κ_0^* determines the stress level in the initial phase of the cyclic hardening process.

(*b*) c determines the number of cycles to saturation of the stress range. As in the Ohno model (Ohno et al., 1988), the following expression is useful for obtaining an approximate value of c from the number of cycles required for the attainment of saturated behavior, N_s.

$$N_s c \approx 1 \tag{43}$$

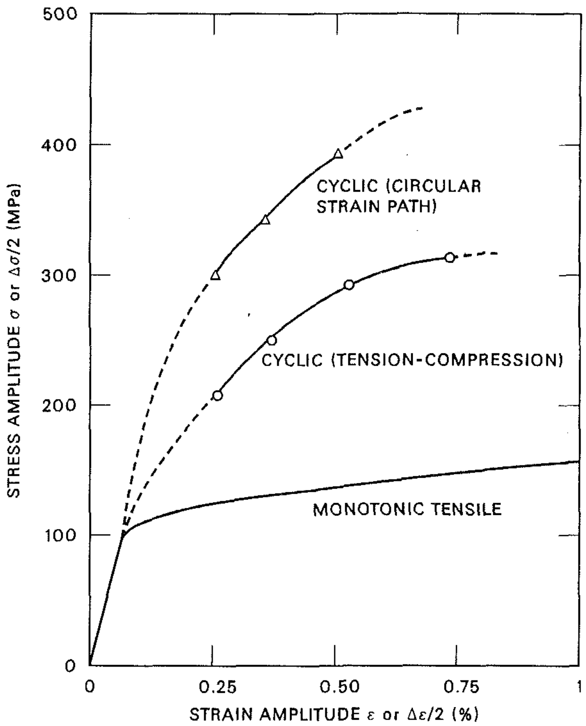


Fig. 2 Stress-strain relation in various loadings

(*c*) C^* and C determine the growth rates of the total stress range and the elastic response range, respectively, in the cyclic hardening process.

The function $Q^*(\rho,1)$ can be determined by fitting each value of $Q^*(\Delta\epsilon^p/2,1) = \kappa^* - \kappa_0^*$ as a function of $\rho = \Delta\epsilon^p/2$.

(iv) Determination of $Q(\rho,0)$ and $Q^*(\rho,0)$: At the stabilized cycle in the constant strain-range circular strain-path tests, ρ becomes equal to $\Delta\epsilon^p/2$, where $\Delta\epsilon^p$ means the axial plastic strain range at the stabilized cycle. The following equations hold under this condition:

$$\kappa = \kappa_0 + Q(\Delta\epsilon^p/2,0), \qquad (44)$$

$$\kappa^* = \kappa_0^* + Q^*(\Delta\epsilon^p/2,0). \qquad (45)$$

Analysis of the intervariable relation at the stabilized cycle shown in the last section gives a precise procedure for determining κ and κ^*. For example, substitution of the total strain range, the stress range and the plastic strain range at the stabilized cycle makes an exact determination of κ and κ^* possible by solving equations (31)-(37) with the function $A(\delta)$ previously determined from the results of uniaxial tests in the procedure (ii).

Finally, the functions $Q(\rho,0)$ and $Q^*(\rho,0)$ can be determined by fitting each value of $Q(\Delta\epsilon^p/2,0) = \kappa - \kappa_0$ and $Q^*(\Delta\epsilon^p/2,0) = \kappa^* - \kappa_0^*$, respectively, as a function of $\rho = \Delta\epsilon^p/2$.

5 Application to Type 304 Stainless Steel at 550°C

5.1 Outline of Experimental Program. The test material is annealed type 304 stainless steel. Uniaxial tension-compression cyclic tests were conducted at four constant strain ranges with the strain rate of 0.1 percent/sec. Circular strain-path tests were also performed at three strain amplitudes $\epsilon_a = \Delta\dot{\epsilon}_{zz} = \Delta\gamma_{rz}/\sqrt{3}$ with the equivalent strain rate $(\dot{\epsilon}_{zz}^2 + \dot{\gamma}_{rz}^2/3)^{1/2}$ equal to 0.1 percent/sec. All tests were conducted at 550°C in air environment.

The main result is shown in Fig. 2, where the relation between stress and strain amplitudes at the stabilized cycle is plotted for both kinds of tests. The monotonic stress-strain relation is also given here. Significant enhancement of cyclic hardening by nonproportional loading can be clearly observed.

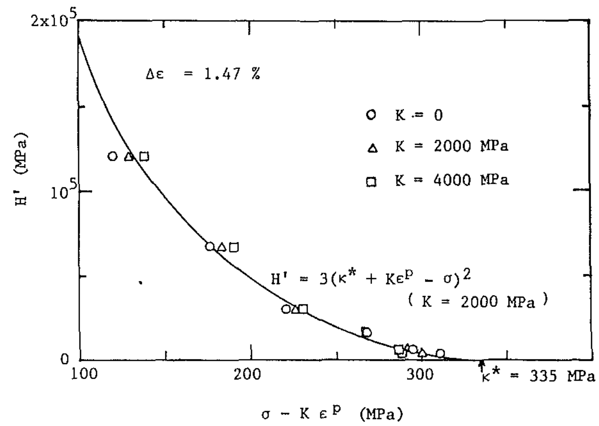


Fig. 3 Procedure for determining material constant K

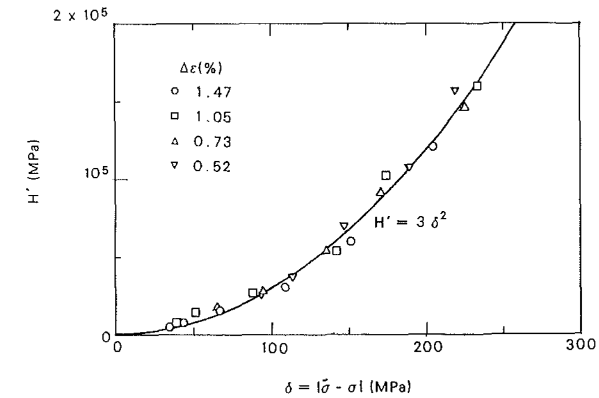


Fig. 4 Determination of material hardening function H'

5.2 Determination of Material Constants and Functions. Material constants and functions involved in the present constitutive model were determined according to the procedure briefly described in the previous section. Some details of the procedure will be given with the resulting values and functions in the following.

First, from the first cycle of uniaxial cyclic tests, κ_0 was determined as 100 MPa, where the Young's modulus was assumed as 157 GPa.

Secondly, from the tension-going portion of the stress-strain curve at the saturated condition in the uniaxial cyclic test with the largest strain range ($\Delta\epsilon = 1.47$ percent), the values of the plastic hardening modulus $H' = d\sigma/d\epsilon^p$ were plotted against $(\sigma - K\epsilon^p)$ for several values of K between 0 and the final slope of $\sigma - \epsilon^p$ relation. The result is shown in Fig. 3. Based on this figure, the value of K was determined in such a way that it resulted in the best fit of the data points, although it required human judgement. In this case, an assumption of $K = 4000$ MPa yielded an apparently nonsmooth behavior in the high-stress regime while a value of 2000 MPa allowed a smooth interpolation to be made. $K = 2000$ MPa was selected on the basis of this observation and κ^* in this condition was obtained as 335 MPa by an extrapolation to $H' = 0$. Tests with a larger strain range will make it possible to obtain a more precise value for K.

Next, similar relations were plotted for the results of the test with other three strain ranges, using the fixed value of K. This procedure gave us a set of almost parallel four curves, from which the optimal value of κ^* was determined for each strain range. Then, the relations between H' and $\delta = \sigma^* - \sigma = \kappa^* + K\epsilon^p - \sigma$ were replotted for all strain ranges, and the functional form of $H' = A(\delta)$ was finally determined. Fortunately, an excellent approximation was possible with the following very simple equation, as shown in Fig. 4.

Table 2 Result of analysis of circular strain path tests

Input Data			Result of Analysis								
ε (%)	σ(MPa)	ε_p(%)	ϕ(deg)	θ(deg)	x(deg)	α(MPa)	δ(MPa)	κ(MPa)	Q(MPa)	κ^*(MPa)	Q^*(MPa)
0.5	393	0.365	44.2	28.5	25.7	200	135	213	113	454	279
0.355	345	0.21	33.5	35.3	10.5	179	169	193	93	439	264
0.25	304	0.09	18.6	43.7	-9.8	170	251	174	74	478	303

Table 3 Summary of material constants and functions

κ_0 = 100 MPa
κ_0^* = 175 MPa
K = 2000 MPa
c = 0.01
C = 15
C^* = 15
$A(\delta) = 3.0\{\delta(\text{MPa})\}^{2.0}$ (MPa)
$Q(\rho,1) = 0$
$Q(\rho,0) = 610\{\rho(\text{mm/mm})\}^{0.30}$ (MPa)
$Q^*(\rho,1) = 390\{\rho(\text{mm/mm})\}^{0.17}$ (MPa)
$Q^*(\rho,0)$ = 280 MPa

$$H' = A(\delta) = 3.0\delta^{2.0} \ (H' \text{ and } \delta \text{ in MPa})$$

The cyclically stabilized behavior in these tests also showed that there was virtually no increase in the linear elastic range even after the large growth of the total stress range. $Q(\rho,1) = 0$ was assumed based on this observation.

As the next step, the value of c was determined as 0.01 from the approximate number of cycles to saturation of hardening observed in the uniaxial cyclic tests. Then, the values of κ^* and C^* were determined by several iterations of comparison between predictions and test data on the cyclic hardening process. It was found that the following constants generally yielded good agreement between the test data and simulation results for four test conditions.

$$\kappa_0^* = 175 \text{ MPa}, \ C^* = 12$$

At the same time, the function $Q^*(\rho,1)$ was determined as

$$Q^*(\rho,1) = 390\rho^{0.17} (Q^* \text{ in MPa and } \rho \text{ in mm/mm}).$$

The value of C was set equal to C^* for the sake of simplicity.

Finally, the functions $Q(\rho,0)$ and $Q^*(\rho,0)$ were determined using the results of three circular strain-path tests. The results of the analysis for three test cases are shown in Table 2. The calculated results for these tests indicated that $Q(\Delta\epsilon^p/2,0)$ showed a clear dependence on $\Delta\epsilon^p$ but that $Q^*(\Delta\epsilon^p/2,0)$ did not. The following expressions were employed based on this indication:

$$Q(\rho,0) = 610 \ \rho^{0.30} \ (Q \text{ in MPa and } \rho \text{ in mm/mm}),$$
$$Q^*(\rho,0) = 280 \text{ MPa}.$$

For convenience, the material constants and the functions determined by the above procedure are summarized in Table 3.

5.3 Result of Simulation. A small computer program was developed for performing the calculation based on the present constitutive model. A simple Euler scheme was used for time integration of the rate equations included in the model. In order to have the yield condition (equation (26)) satisfied at the end of each calculational step, the radial return method (Simo and Taylor, 1985) was applied. One hundred steps were used for the calculation of one cycle.

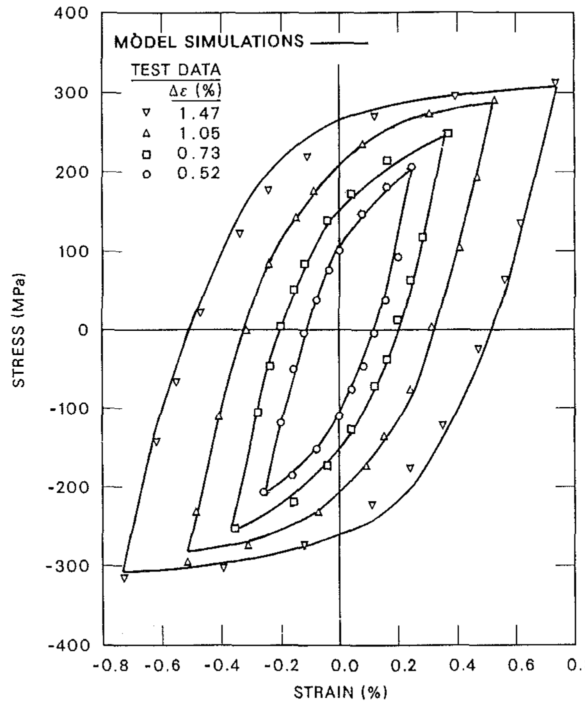


Fig. 5 Comparison of stabilized stress-strain relations in uniaxial cyclic loading condition

Figures 5–8 show a comparison of the experimental and theoretical results. Generally very good agreement was obtained in these comparisons. In particular, the stress-strain relations at the saturated state in both uniaxial and circular strain-path loadings were simulated with very small deviation from the experimental data. Not only the stress ranges but also the shapes of the stress-strain curves were described well. This comes from the employment of the precise procedure developed in the present study for determining the material constants and functions.

On the other hand, prediction of the cyclic hardening process in uniaxial loadings was not as excellent as the prediction of stabilized stress-strain behavior.

As can be seen in Fig. 7, the strain-range dependency of the number of cycles required for saturation of cyclic hardening was not simulated well by the model. It is expected that improved description can be obtained by employment of larger c and smaller C and C^* values than those used in the present simulation. But there is a possibility of imposing a too large memory effect with a larger value of c (e.g., maximum plastic strain-range criterion of Chaboche et al., 1979, for $c = 0.5$).

Very good agreement was obtained for the hardening rate under the circular strain-path condition was shown in Fig. 8. The model predicted more rapid attainment of the saturated stress range than the case of uniaxial tests. This was also observed in the test results. It is a very encouraging fact that the hardening rate under these two limit conditions can be well represented with the same value of C^* and C.

Figure 8 also includes the results predicted by the model not considering the additional hardening due to nonproportional loading. These results were obtained by simply replacing Q

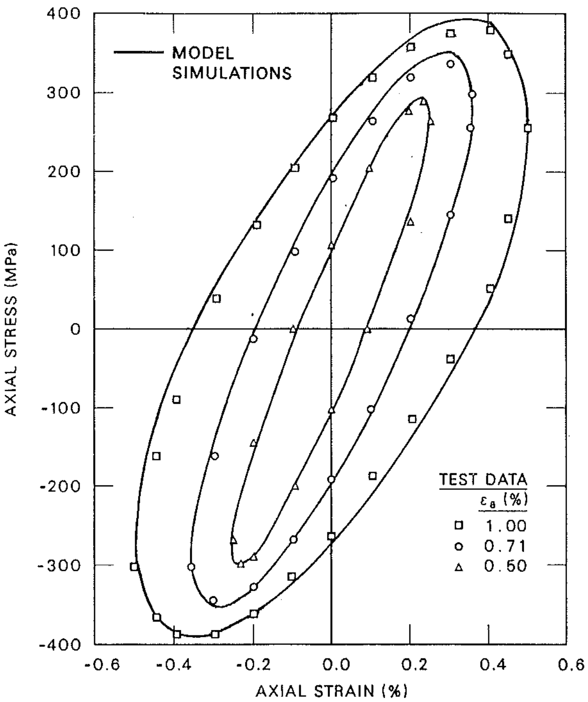


Fig. 6 Comparison of stabilized stress-strain relations in circular strain-path loading condition

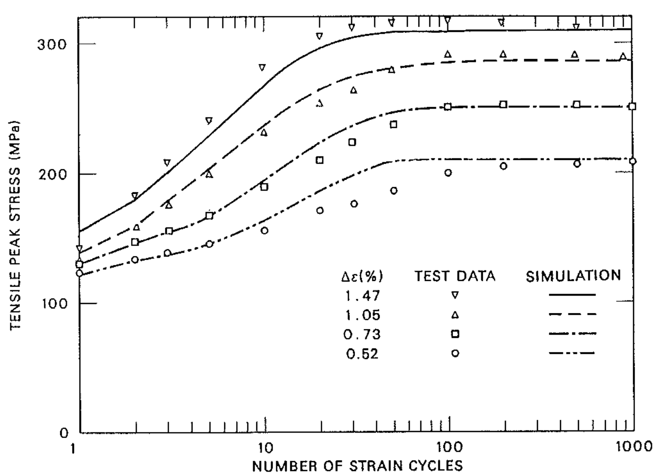


Fig. 7 Comparison of cyclic hardening process in uniaxial cyclic loading condition

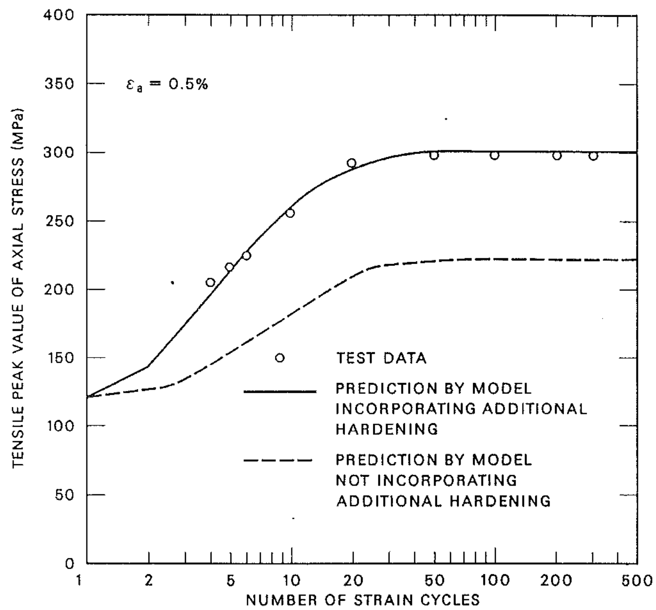


Fig. 8(*a*)

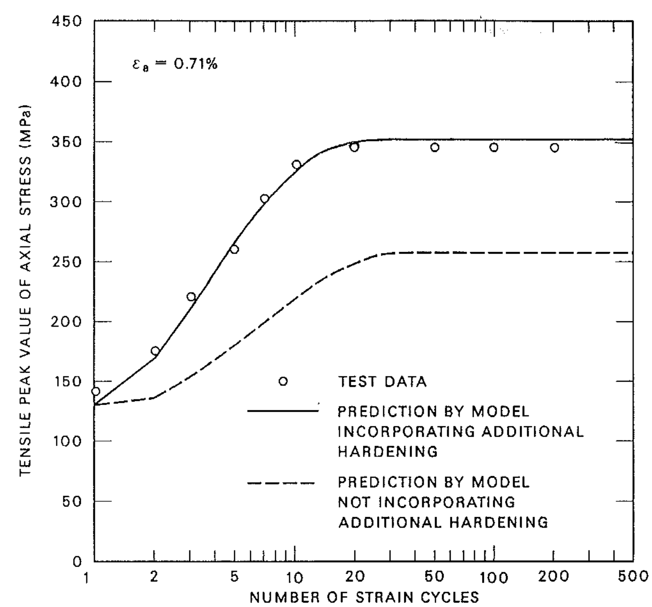


Fig. 8(*b*)

$(\rho,0)$ and $Q^*(\rho,0)$ by $Q(\rho,1)$ and $Q^*(\rho,1)$, respectively. These results considerably underestimated the stress range under circular strain-path loadings. This again clearly shows us the importance of considering the additional hardening in nonproportional cyclic loading.

6 Conclusion

In this study, a simple elastic-plastic constitutive model based on the two-surface theory was presented for description of the hardening behavior of austenitic stainless steels under nonproportional cyclic loadings. Expressions which can be used for the precise determination of the material parameters were derived based on analysis of the intervariable relations existing at the stabilized cycle of circular strain-path loading. It was shown that excellent descriptions of cyclic deformation behavior under uniaxial and circular strain-path conditions can be made using the material parameters determined by these expressions. The method of extension to general nonproportional loading—particularly the definition of q and the dependency of Q and Q^* on q—is an important subject for completed development of this model.

References

Benallal, A., and Marquis, D., 1987, "Constitutive Equations for Nonproportional Cyclic Elasto-Viscoplasticity," *Journal of Engineering Materials and Technology*, Vol. 109, pp. 326–336.

Cailletaud, G., Kaczmarek, H., and Policella, H., 1984, "Some Elements on Multiaxial Behaviour of 316 L Stainless Steel at Room Temperature," *Mechanics of Materials*, Vol. 3, pp. 333–347.

Chaboche, J. L., 1977, "Viscoplastic Constitutive Equations for the Description of Cyclic and Anisotropic Behavior of Metals," *Bull. de l'Acad. Polonaise des Sciences, Serie Sc. et Tech.*, Vol. 25, No. 1, pp. 33–42.

Chaboche, J. L., Dang Van, K., and Cordier, G., 1979, "Modelization of the Strain Memory Effect on the Cyclic Hardening of 316 Stainless Steel," *Transactions of the 5th International Conference on Structural Mechanics in Reactor Technology*, Paper No. L3/4, Vol. L.

Chaboche, J. L., and Rousselier, G., 1983, "On the Plastic and Visoplastic Constitutive Equations—Part I and Part II," *Journal of Pressure Vessel Technology*, Vol. 105, pp. 153–164.

Defalias, Y. F., and Popov, E. P., 1975, "A Model of Nonlinearly Hardening Materials for Complex Loading," *Acta Mechanica*, Vol. 21, pp. 173–192.

Krempl, E., and Lu, H., 1984, "The Hardening and Rate Dependent Behavior

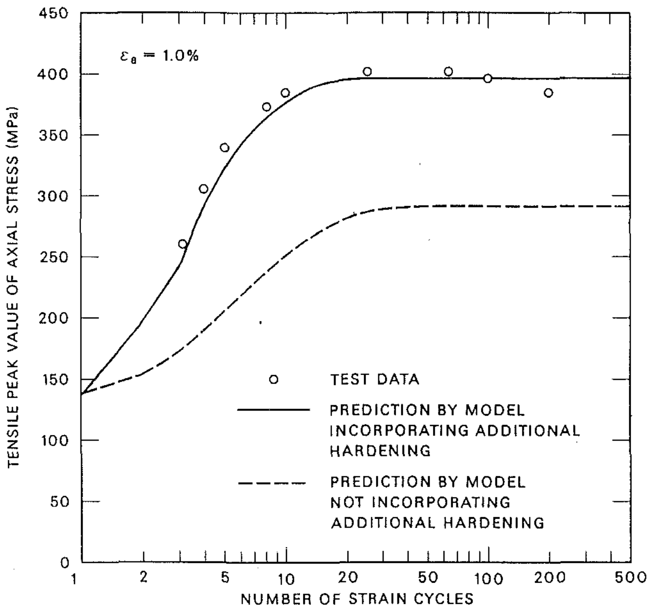


Fig. 8(c)

Fig. 8 Comparison of cyclic hardening process in circular strain-path loading condition

of Fully Annealed AISI type 304 Stainless Steel Under Biaxial in Phase and Out-of-Phase Strain Cycling at Room Temperature," *ASME Journal of Engineering Materials and Technology*, Vol. 106, pp. 376-382.

Krieg, R. D., 1975, "A Practical Two Surface Plasticity Theory," ASME Journal of Applied Mechanics, Vol. 42, pp. 641-646.

McDowell, D. L., 1985a, "An Experimental Study of the Structure of Constitutive Equations for Nonproportional Cyclic Plasticity," *ASME Journal of Engineering Materials and Technology*, Vol. 107, pp. 307-315.

McDowell, D. L., 1985b, "A Two-Surface Model for Transient Nonproportional Cyclic Plasticity: Part 1 and Part 2," ASME Journal of Applied Mechanics, Vol. 52, pp. 298-308.

Mroz, Z., 1967, "An Attempt to Describe the Behavior of Metals Under Cyclic Loads Using a More General Workhardening Model," *Acta Mechanica*, Vol. 7, pp. 199-212.

Murakami, S., Kawai, M., Aoki, K., and Ohmi, Y., "Temperature-Dependence of Multiaxial Non-Proportional Cyclic Hardening of Type 316 Stainless Steel," *Transaction of Japan Society of Mechanical Engineers*, No. 53-495, pp. 2077-2084 (in Japanese).

Nouailhas, D., 1987, "A Viscoplastic Modeling Applied to Stainless Steel Behavior," *Constitutive Laws for Engineering Materials: Theory and Applications*, C. S. Desai et al., eds., Elsevier Science Publishing Co., Inc., pp. 717-724.

Ohno, N., 1982, "A Constitutive Model of Cyclic Plasticity With a Nonhardening Strain Region," ASME Journal of Applied Mechanics, Vol. 49, pp. 721-727.

Ohno, N., and Kachi, Y., 1986, "A Constitutive Model of Cyclic Plasticity for Nonlinear Hardening Materials," ASME Journal of Applied Mechanics, Vol. 53, pp. 395-403.

Ohno, N., Takahashi, Y., and Kuwabara, K., 1989, "Constitutive Modeling of Anisothermal Cyclic Plasticity of 304 Stainless Steel," *ASME Journal of Engineering Materials and Technology*, Vol. 111, pp. 106-114.

Simo, J. C., and Taylor, R. L., 1985, "Consistent Tangent Operators for Rate-Independent Elastoplasticity," *Computer Methods in Applied Mechanics and Engineering*, Vol. 48, pp. 101-118.

Tanaka, E., Murakami, S., and Ooka, M., 1985, "Effects of Strain Path Shapes on Nonproportional Cyclic Plasticity," *Journal of Mechanics and Physics of Solids*, Vol. 33, No. 6, pp. 559-575.

Tanaka, E., Murakami, S., and Ooka, M., 1986, "A Constitutive Model of Cyclic Plasticity in Non-Proportional Loading Conditions (1st Report, Fundamental Notation and its Formulations)," *Transaction of Japan Society of Mechanical Engineers*, No. 52-480, pp. 2054-2061 (in Japanese).

Tanaka, E., Murakami, S., and Ooka, M., 1986, "A Constitutive Model of Cyclic Plasticity in Non-Proportional Loading Conditions (2nd Report, Identification of Material Constants and Material Functions, and Experimental Evaluation of the Model)," *Transaction of Japan Society of Mechanical Engineers*, No. 52-483, pp. 2520-2528 (in Japanese).

Void Growth in Plastically Deformed Free-Cutting Brass

M. J. Worswick[1]
Defence Research Establishment, Suffield, Ralston, Alberta T0J 2N0, Canada

R. J. Pick
University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

The void growth occurring during tensile testing of uniaxial and notched specimens of free-cutting brass has been determined experimentally. This material contains a globular lead phase which tears or bursts to nucleate voids during deformation. Using quantitative metallographic data from specimens whose deformation was interrupted prior to failure, histories of void volume fraction and void aspect ratio were determined. The measured stress-strain response from the tensile tests was shown to be close to predictions from a finite element model incorporating Gurson's constitutive model for a porous plastic solid. Predicted void growth rates agreed well with experiment for uniaxial specimens but were less than the measured growth rates in notched, high triaxiality specimens.

1 Introduction

Ductile fracture occurs in plastically deforming metals through the nucleation, growth, and coalescence of small internal voids or cavities (Rogers, 1960), the most common sites for void nucleation being hard second-phase particles or inclusions (Goods and Brown, 1979; Fisher and Gurland, 1981). The void growth stage of ductile fracture involves the stable expansion of voids within a material undergoing tensile plastic deformation. There are two aspects of this stage that must be considered: (i) the void expansion and change in shape during deformation; and (ii) the degradation in material load-carrying capacity due to the presence of the voids, referred to here as constitutive softening.

Previous analytical models (McClintock, 1968; Rice and Tracey, 1969; Budiansky et al., 1982) have considered the growth of an isolated void within an infinite medium. Models considering void growth and constitutive softening have also been developed for the case of finite porosity using analytical (Gurson, 1975, 1977) and numerical approaches (Needleman, 1972; Tvergaard, 1981; Koplik and Needleman, 1988; Worswick and Pick, 1989b). Experimental studies of void growth (Beremin, 1981; Marini et al., 1985; Bourcier et al., 1986; Becker et al., 1988; Spitzig et al., 1988) have shown qualitative agreement with these models. However, there appears to be quantitative differences between predicted and measured void growth rates, and between measured void growth rates from independent experimental studies.

In the present paper, measurements of void growth in uniaxial and notched tensile specimens of free-cutting brass (UNS C36000) are presented. This material contains a dispersion of globular lead particles (added to improve machinability) which burst or tear during plastic deformation to nucleate voids. Histories of damage versus applied deformation were constructed from a series of interrupted tensile tests in which the specimens were loaded to predetermined strain levels and then sectioned for quantitative metallographic examination.

The measured void growth rates were used to assess predictions by Rice and Tracey (1969) and Budiansky et al. (1982) based on isolated voids. In addition, the measured stress-strain and porosity histories were compared to those from a finite element model incorporating Gurson's (1975, 1977) porous continuum constitutive model.

In discussing the models considered in this paper, it is important to make a distinction between macroscopic and microscopic stresses and strains. Microscopic stresses and strains, σ_{ij} and ϵ_{ij}, respectively, refer to the detailed distributions around individual voids whereas the macroscopic stresses and strains, Σ_{ij} and ϵ_{ij}^{∞}, are average values such as would be measured in a tensile test.

2 Material

The tensile specimens were fabricated from free-cutting brass rod (UNS C36000) which has a nominal composition 61.5 percent Cu-35.5 percent Zn-3.0 percent Pb (weight percent). The material is three phase: alpha, beta, and lead. The lead phase, being essentially insoluble in Cu-Zn, forms small globular particles that are elongated along the rod axis due to the continuous casting and drawing process used in fabricating the rod.

The material was annealed at 800° C for two hours in order to promote spheroidization of the lead phase, as seen in Fig. 1. Using standard metallographic techniques (Underwood, 1970), the grain size after annealing was determined to be 55 μm. The volume fraction of lead was 0.025 and the mean lead particle aspect ratio in the plane of the section was close to

[1]Presently at the Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, Ont., K1S 5B6, Canada.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 15, 1989; final revision, Jan. 8, 1990.

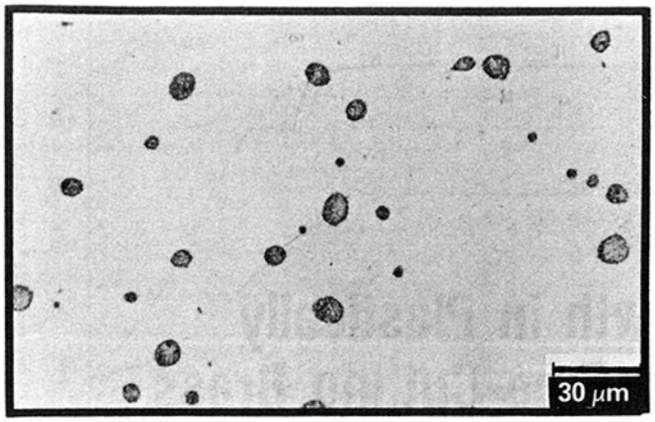


Fig. 1 Optical micrograph of UNS C36000 free-cutting brass showing spheroidized lead phase after annealing at 800°C for two hours (polished only)

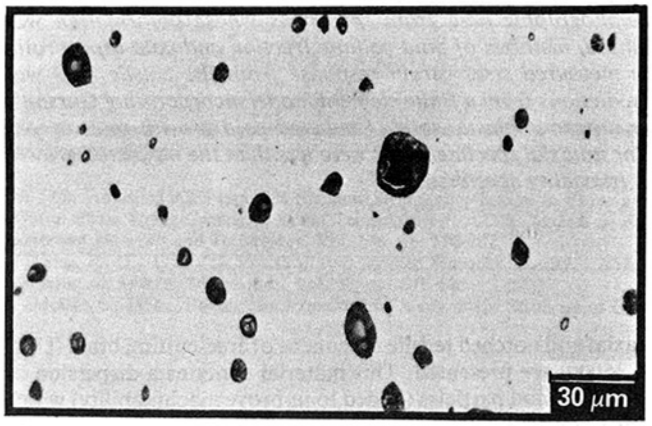


Fig. 2 Optical micrograph showing central tear or burst of lead particles after plastic deformation (polished only)

unity. The in-plane mean "nearest-neighbor" lead particle spacing was approximately 15 μm.

The presence of the globular lead phase was the primary reason for selecting free-cutting brass for this investigation. Pure lead has a tensile strength of approximately 20 MPa compared to roughly 400 MPa for half-hard free-cutting brass. Thus, the lead particles tear or burst during the early stages of tensile plastic deformation and nucleate voids (Fig. 2). In view of the disparity in strength between the lead particles and the brass matrix, the initial void size was taken as the size of the lead particle.

2.1 Material Properties. Values for Young's modulus and Poisson's ratio, equal to 97 GPA and 0.3, respectively, were adopted. The initial yield strength was 94.4 MPa and the uniaxial stress versus plastic strain behavior is described using a cubic polynomial of the form

$$\bar{\sigma} = 94.4 + 1594\bar{\epsilon}^{p} - 1693(\bar{\epsilon}^{p})^2 + 645\,(\bar{\epsilon}^{p})^3, \qquad (1)$$

in which $\bar{\sigma}$ is the material flow stress and $\bar{\epsilon}^{p}$ is the equivalent plastic strain. The constants in equation (1) were determined from linear regression of results from uniaxial tensile tests (Worswick, 1988). Bridgman corrections were not used in determining this equation.

3 Experimental Procedures

Six uniaxial specimens and ten notched specimens were pulled in displacement-controlled tests at speeds of 0.25 and 0.13 mm/min., respectively. Two notch profiles were used, corresponding to the A-notch and D-notch designations studied by Hancock and MacKenzie (1976). Figure 3 shows the notch detail of the D-notch specimen, for which the initial root radius, r^o, and notch radius, R^o, were equal to 3.8 mm and 1.27 mm, respectively. The initial root radius and notch radius of the A-notch specimen were both 3.8 mm. The uniaxial tensile specimen had an initial diameter and gauge length of 7.6 mm and 50 mm, respectively.

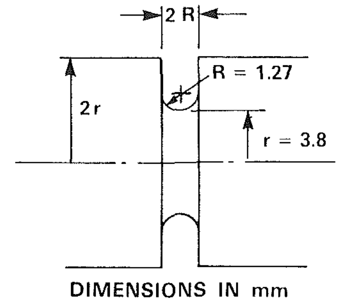


Fig. 3 Notch detail of the D-notch specimen. Tensile straining direction is in the horizontal plane.

Two uniaxial specimens and one of each notched specimen were tested to failure in order to establish the failure strain and stress-strain response for each specimen geometry. The remaining tests were interrupted at predetermined strains prior to fracture.

3.1 Metallography. After testing, the notch region of each specimen was cut out and sectioned longitudinally. The specimen surface was carefully ground and polished to ensure that the edges of the voids did not become rounded. The plane of the final polished section always lay within 0.1 mm of the specimen axis.

The polished specimens were examined under an optical microscope to determine the void volume fraction, f, and the average void aspect ratio, $\bar{Q}$. Quantitative assessment of the damage was performed using an on-line image processing system which measured the areal fraction of voids and the individual void aspect ratios, Q. The areal fraction of voids was taken as being equal to the void volume fraction, a statistically correct approach if a sufficient number of samples are taken (Underwood, 1970; VanderVoort, 1984). The average void aspect ratio, $\bar{Q}$, defined as the mean of the individual void aspect ratios in the plane of the section, was assumed to be representative of the aspect ratio of the embedded voids. Probabilistic approaches (DeHoff, 1964) for relating the mean aspect ratio of the embedded voids to that in the section plane were not considered.

Two approaches were used in sampling the sectioned specimens. In the first approach, micrographs were acquired at points along the specimen axis and along the radial direction across the minimum section. Three measurements were taken at each point, one directly on the axis and one on either side of the axis. The size of the sampled region in each micrograph was 192 μm × 144 μm. For the remaining specimens, the measurements were taken from a small region at the specimen center using a 9 × 12 grid pattern with 0.2 × 0.15 mm spacing. This pattern was used to measure f and $\bar{Q}$ at the specimen center and to ascertain the local statistical variation in these measured quantities while the axial and radial patterns serve to indicate the variation within the specimen.

4 Finite Element Model

The finite element meshes used to model the notched specimens were those used by Worswick (1988) and were similar to meshes used in a number of previous studies (Needleman and Tvergaard, 1984; Becker et al., 1988). The model of the uniaxial tensile specimen considered a "perfect" specimen, free

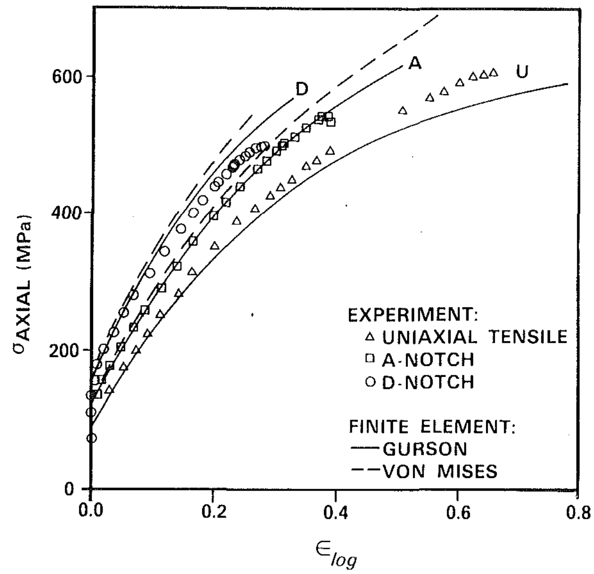


Fig. 4 Comparison between measured and predicted axial stress (σ_{axial}) versus logarithmic strain (ϵ_{log}) response. The letters U, A, and D indicate the predictions for the Uniaxial, A-notch, and D-notch specimens, respectively. Symbols indicate measured values.

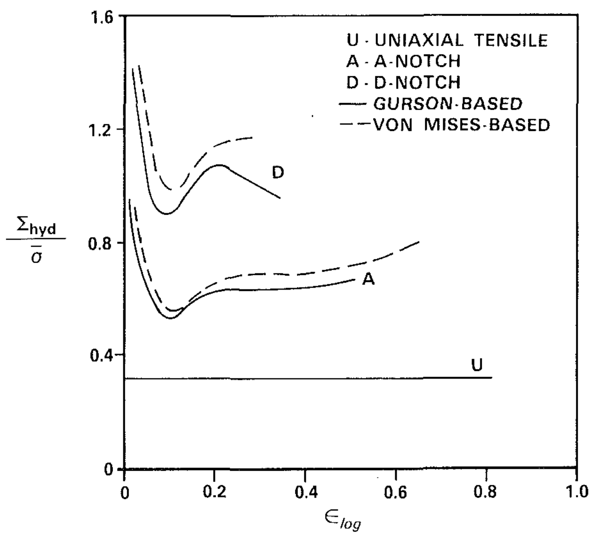


Fig. 5 Predicted stress triaxiality ($\Sigma_{hyd}/\bar{\sigma}$) versus logarithmic strain (ϵ_{log}) histories from the finite element models. The letters U, A, and D indicate the predictions for the Uniaxial, A-notch, and D-notch specimens, respectively.

of imperfections, and therefore would not account for the effects of geometric instabilities leading to necking. This approach provided predictions of void growth at constant stress triaxiality, $\Sigma_{hyd}/\bar{\sigma} = 1/3$, Σ_{hyd} being the hydrostatic stress. Note that in neglecting necking, errors in the predicted stress triaxiality and void growth rate result during the latter stages of deformation. The reader is referred to Tvergaard and Needlemen (1984) and Aravas (1987) for analyses of necking instability employing a Gurson constitutive model.

The finite element calculations were performed using the general purpose nonlinear finite element code ABAQUS (Hibbitt et al., 1984). A nonlinear geometric formulation, based on the updated Lagrangian approach due to McMeeking and Rice (1975), was used. The material model adopted was coded in a user-subroutine linked with the ABAQUS library which is called to integrate the constitutive model and calculate the material "stiffness" modulii.

The elastic response of the material was idealized as conforming to Hooke's Law for a linear elastic isotropic solid.

The macroscopic plastic response was modeled using the Gurson (1975, 1977) constitutive model describing porous continuum plastic solids. Central to this model is the use of the Gurson yield function to determine conditions to initiate or sustain plastic flow within a plastically dilating porous solid:

$$\phi = \left(\frac{\Sigma_{eq}}{\bar{\sigma}}\right)^2 + 2fq_1 \cosh\left(q_2 \frac{3\Sigma_{hyd}}{2\bar{\sigma}}\right) - 1 - q_3 f^2 = 0, \quad (2)$$

in which f is the void volume fraction and Σ_{eq} is the equivalent stress, defined as $\Sigma_{eq}^2 = 3/2\Sigma'_{ij}\Sigma'_{ij}$, Σ'_{ij} being the deviatoric components of Σ_{ij}. The coefficients q_1, q_2, and q_3 are "calibration" coefficients introduced by Tvergaard (1981, 1982) to better represent the effects of porosity in plastically deforming solids. The values adopted were $q_1 = 1.25$, $q_2 = 0.95$ and $q_3 = q_1^2$, given by Becker et al. (1988) and Worswick and Pick (1989a).

The detailed form of the constitutive equations used is given by Gurson (1975, 1977). A forward difference scheme was used to integrate these equations which required small time steps in order to avoid numerical instability. Future work will consider use of a backwards difference operator to integrate the Gurson constitutive model, as described by Aravas (1987). The initial void volume fraction was taken as being equal to 0.025, the volume fraction of the lead particles. Void growth was modeled as initiating immediately with the onset of tensile plastic deformation, a reasonable assumption if the void nucleation strain is low. The validity of this assumption will be addressed later in this paper.

5 Results

The axial stress-strain results from the specimens tested to failure are plotted in Fig. 4 (symbols). The axial stress, σ_{axial}, is the axial load on the specimen divided by the current cross-sectional area while the measure of strain is the logarithmic strain, $\epsilon_{log} = 2\ln(r^o/r)$. The figure serves to demonstrate the increase in axial stress and decrease in ductility with notch severity. Also plotted are the predicted curves obtained from the Gurson-based finite calculations (solid curves) and calculations using a Von Mises, fully dense, material model (dashed curves). The predicted curves using the Gurson model lie closer to the experimental data than the Von Mises results. The Gurson-based predictions of axial stress for the D-notch specimen exceed the measured value by roughly 5 percent at $\epsilon_{log} = 0.25$. Beyond this strain level, the measured stress drops off suggesting that void coalescence processes, not considered by the model, were beginning to occur. Similar behavior was seen in the A-notch experimental results near the failure point.

The increased constraint on plastic flow in the notched specimens led to the development of regions of high stress triaxiality at the center of the notch. The predicted histories of stress triaxiality from the finite element analyses are plotted in Fig. 5 and are similar to those obtained by Hancock and Brown (1983). After the initial "spike" associated with the onset of yielding, the triaxiality levels within the A and D-notch specimens ranged between approximately 0.6–0.7 and 0.9–1.1, respectively. The calculations using the Von Mises constitutive model (dashed curves) yielded somewhat higher values.

The predicted triaxiality within the uniaxial tensile specimen remained just below 1/3 for the duration of the analysis, since the model considered a "perfect" specimen without necking. Consequently, the finite element predictions of σ_{axial} for this specimen generally lay below the measured values (Fig. 4).

5.1 Initial Porosity and Void Aspect Ratio. The initial void volume fraction and void aspect ratio, summarized in Table 1, were measured using three undeformed samples taken from the same rod as the tensile specimens. Additional measurements were made on the threaded end of one tensile specimen. Note that the initial "voids" are actually the embedded lead particles which tear open during deformation. Shown in

Table 1 Initial void volume fraction and aspect ratio

Table 1(a): Void Volume Fraction

Specimen	N (micrographs)	f	Std. Dev.	CL(95%)	% RA
16	107	0.026	0.008	0.002	6.2
17	105	0.021	0.007	0.001	6.3
18	98	0.028	0.009	0.002	6.3
T*	88	0.025	0.008	0.002	6.4
Mean		0.025			

Table 1(b): Void Aspect Ratio

Specimen	N (voids)	$\bar{Q}$	Std. Dev.	CL(95%)	% RA
16	2222	1.006	0.233	0.010	1.0
17	1786	1.037	0.279	0.013	1.2
18	1603	1.013	0.234	0.011	1.1
T*	1468	0.985	0.229	0.012	1.2
Mean		1.010			

* at end of tensile specimen

Table 1 are the sample mean, standard deviation, 95 percent confidence interval, and percent relative accuracy. A total of 108 micrographs were acquired on 9 × 12 grids for the undeformed specimens. The number (N) reported was less than 108 since a number of micrographs were rejected due to the presence of surface flaws, such as polishing scratches.

The mean void areal fraction, f, varied between 0.021–0.028 with a relatively large sample standard deviation in the range 0.007–0.009. The mean void aspect ratio, $\bar{Q}$, was very close to unity, ranging between 0.98 and 1.04.

5.2 Porosity and Void Aspect Ratio Distributions. The measured radial distributions of void volume fraction (f) across the minimum sections of an A and D-notch specimen are plotted in Figs. 6(*a*) and (*b*). From examination of the figures, there is considerable scatter in f within the deformed specimens, attributed largely to the small sample size taken at each material point (3 micrographs) and the large variation in f between samples. No confidence intervals are shown in Fig. 6; however, an estimate of the 95 percent confidence interval on f, based on three micrographs, was obtained using the variance from the undeformed specimens to estimate the variance in the deformed material. This calculation yielded an error band of ± 0.009 on the mean value of f which was considered large. This error could be reduced by decreasing the magnification and thereby sampling larger areas per image; however, inaccurate measurement of the individual particle aspect ratios would result due to the limited resolution of the image acquisition system. Alternatively, the use of larger specimens would permit acquisition of more micrographs at each material "point," thereby reducing the error in f.

Distributions of void aspect ratio (Q) across the minimum sections are plotted in Fig. 7. The estimated error in $\bar{Q}$ from three micrographs (approximately 60 voids) was ± 0.06 and was considered acceptable.

The finite element predictions of f, for the corresponding values of $\epsilon_{\log}$, lie above the measured distributions suggesting that the predicted void growth rates exceed the measured rates. However, as will be discussed below, the lower measured increase in void volume fraction was due to a delay in the onset of void growth during the experiments.

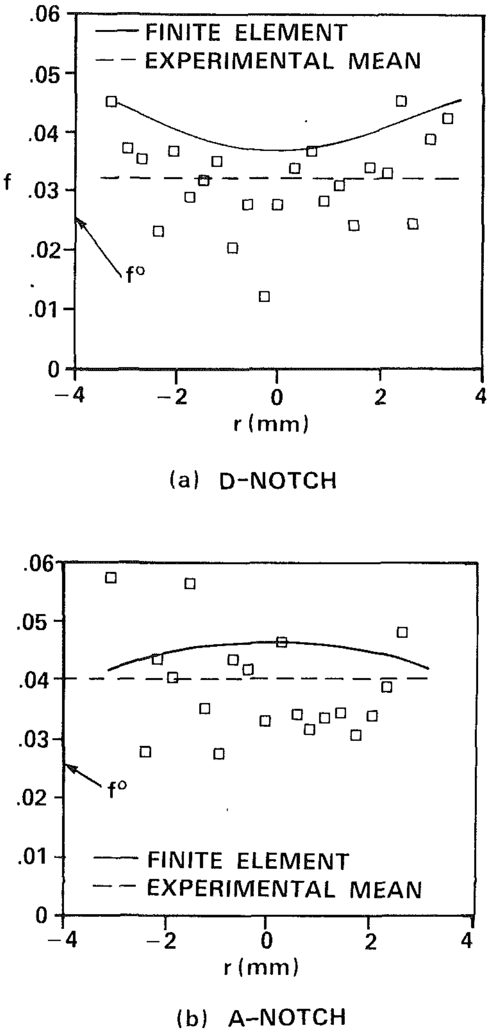


Fig. 6 Radial distributions of void volume fraction (f) across minimum section D and A-notch specimens strained to $\epsilon_{\log} = 0.2$ and 0.38, respectively. Symbols indicate measured values.

Examination of the predicted porosities in Fig. 6(*b*) for the A-notch specimen reveals higher porosity at the specimen center compared to the notch region. The increased porosity at the center of the specimen was due to the higher triaxiality in that region. The predicted distribution for the D-notch specimen shows the highest porosity at the notch. The D-notch is much sharper than the A-notch, causing a greater concentration of strain at the D-notch surface which led to higher porosities.

While the authors are unaware of a simple closed-form solution for the rate of void shape change (e.g., $\dot{Q}/Q$), results from Budiansky et al. (1982) and Worswick and Pick (1989b) have shown that $\dot{Q}/Q$ will decrease as triaxiality increases. As a result, larger void aspect ratios developed at the notch where the plastic strains were highest and the triaxiality was lowest (Fig. 7).

Axial distributions of void volume fraction and void aspect ratio along $r = 0$ are shown in Fig. 8 for the D-notch specimen tested to failure. The $z = 0$ coordinate in the figure corresponds to the fracture surface position. As with the radial distributions, there is a great deal of scatter in the distribution of f whereas the aspect ratio distribution is reasonably well defined.

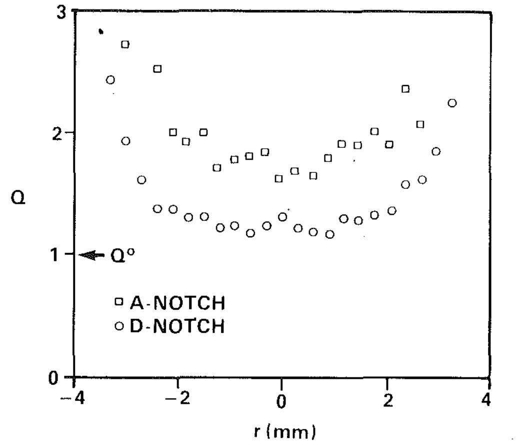


Fig. 7 Radial distributions of void aspect ratio (Q) across minimum section of D and A-notch specimens strained to $\epsilon_{log} = 0.2$ and 0.38, respectively

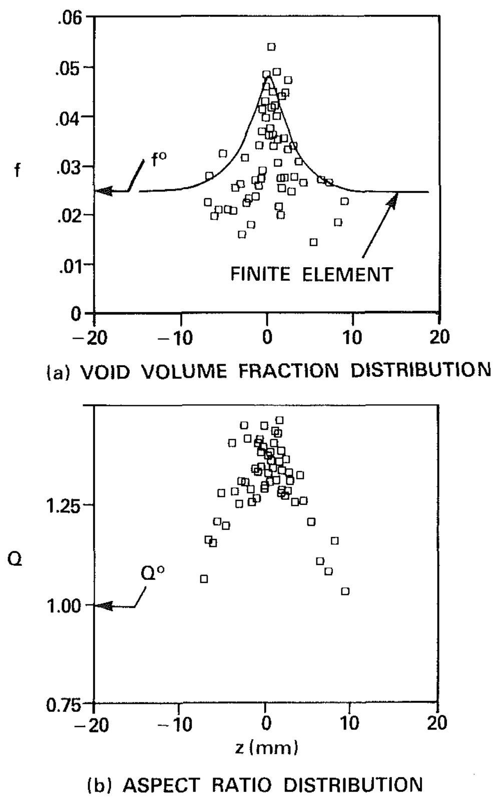


Fig. 8 Axial distributions of void volume fraction (f) and void aspect ratio (Q) along axis of the D-notch specimen tested to failure ($\epsilon_{log} = 0.3$)

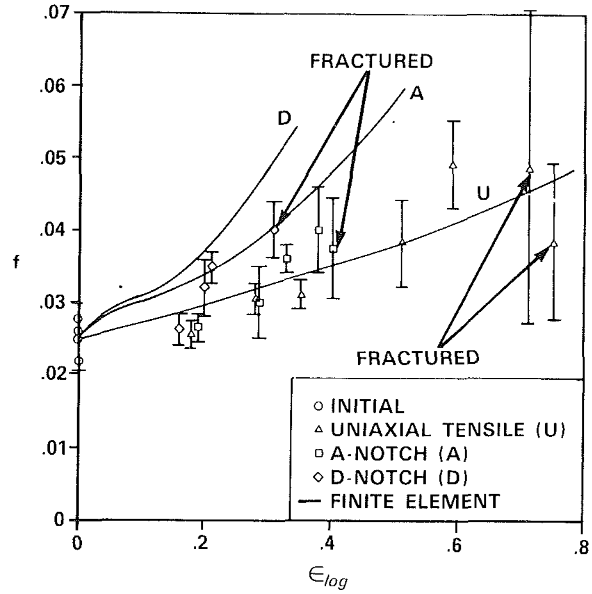


Fig. 9 Void volume fraction (f) versus logarithmic strain (ϵ_{log}) histories. Symbols indicate measured values while the solid lines correspond to the finite element predictions. The vertical bars indicate the 95 percent confidence intervals on the measured values. The letters U, A, and D indicate the results for the Uniaxial, A-notch, and D-notch specimens, respectively.

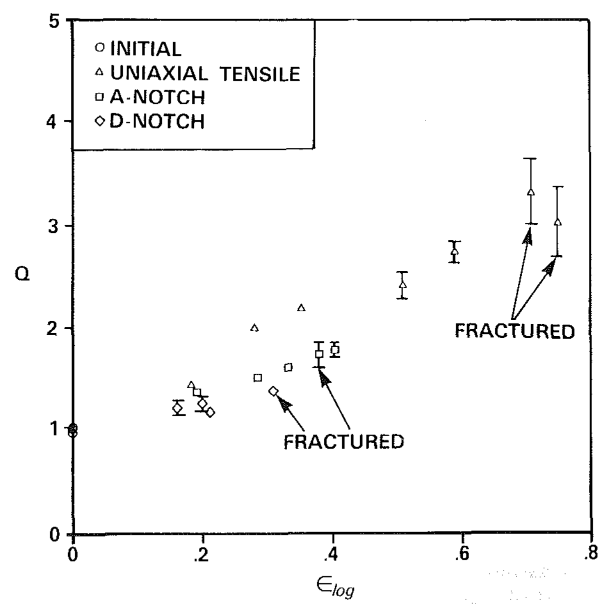


Fig. 10 Void aspect ratio (Q) versus logarithmic strain (ϵ_{log}) histories. Symbols indicate measured values while the solid lines correspond to the finite element predictions. The vertical bars indicate the 95 percent confidence intervals on the measured values. The letters U, A and D indicate the results for the Uniaxial, A-notch and D-notch specimens, respectively.

Only qualitative agreement with the finite element prediction can be suggested due to the scatter in the experimental results. Similar scatter was seen in the A-notch and uniaxial specimen distributions (not shown).

The sharp increase in f and Q, in Fig. 8, across the $z=0$ plane of symmetry prompts concern over the size of the D-notch region sampled using the grid pattern. The initial gauge length of the notch was 2.54 mm, while the grid pattern extends 1.8 mm along the axis; thus the actual values of f and Q at the specimen centre may exceed the reported mean values. This measurement inaccuracy could be reduced by using larger specimens.

5.3 Void Growth Rates. The measured void volume fraction and void aspect ratio at the specimen centers are plotted as functions of logarithmic strain in Figs. 9 and 10. The scatter bands indicate the 95 percent confidence intervals. The void volume fraction results from the fractured specimens have a high degree of uncertainty due to the scatter in the void volume fraction distributions as seen in Fig. 8 for the D-notch specimen.

The predicted void volume fraction histories from the finite element model are plotted in Fig. 9. The predicted curves echo the trends seen in the experimental data of increasing dilatational growth rate (the rate of void expansion) with triaxiality (notch severity). In Fig. 10, the extensional growth rate (the rate of void shape change) is highest for the low triaxiality uniaxial specimen.

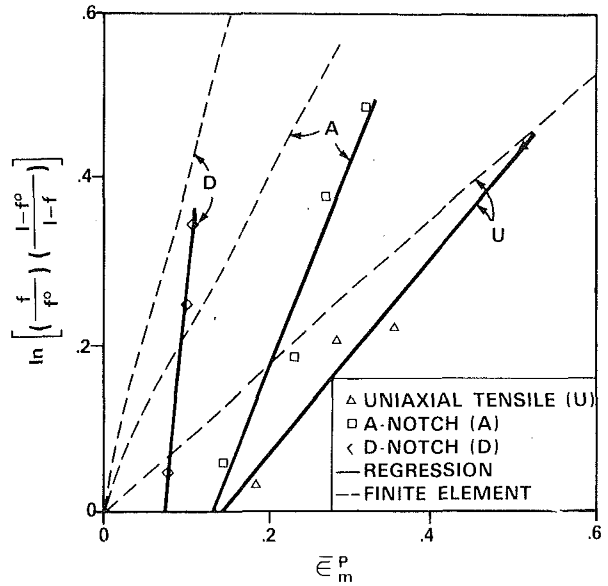


Fig. 11 Dilatational growth histories from the interrupted tensile specimens. Solid lines are the curve fits obtained using equation (5) while the dashed lines show the finite element predictions. The letters U, A, and D indicate the results for the Uniaxial, A-notch, and D-notch specimens, respectively.

Rice and Tracey (1969) introduced dilatational and extensional growth factors D and 1 + E, respectively, to characterize the rate of growth of an isolated void within an infinite media. For the case of finite porosity, assuming the voids are spherical and do not interact, the rate of dilatational growth can be characterized using

$$\frac{\dot{f}}{f(1-f)} = 3D\dot{\bar{\epsilon}}_m^p, \tag{3}$$

in which D relates $\dot{f}$ to the effective plastic strain rate in the matrix, $\dot{\bar{\epsilon}}_m^p$. Following Gurson (1975), the effective plastic strain rate within a porous material can be calculated using $\dot{\bar{\epsilon}}_m^p = \Sigma_{ij}\dot{\epsilon}_{ij}^{p\infty}/(1-f)\bar{\sigma}$, $\dot{\epsilon}_{ij}^{p\infty}$ being the plastic strain rate tensor. The rate of void shape change can be characterized using

$$\frac{\dot{Q}}{Q} = (1+E)(\dot{\epsilon}_{33}^{p\infty} - \dot{\epsilon}_{11}^{p\infty}), \tag{4}$$

in which 1 + E relates $\dot{Q}$ to the shape change inherent in the applied plastic strain field ($\dot{\epsilon}_{33}^{p\infty} - \dot{\epsilon}_{11}^{p\infty}$). In the tensile specimens, $\dot{\epsilon}_{33}^{p\infty}$ and $\dot{\epsilon}_{11}^{p\infty}$ correspond to the macroscopic plastic strain rates along the axial and radial directions, respectively.

By assuming constant values for D and 1 + E, equations (3) and (4) can be integrated to obtain

$$\ln\left[\left(\frac{f}{f^o}\right)\left(\frac{1-f^o}{1-f}\right)\right] = 3D(\bar{\epsilon}_m^p - (\bar{\epsilon}_m^p)_o) \tag{5}$$

and

$$\ln(Q/Q^o) = (1+E)[(\epsilon_{33}^{p\infty} - \epsilon_{11}^{p\infty}) - (\epsilon_{33}^{p\infty} - \epsilon_{11}^{p\infty})_o]. \tag{6}$$

The terms $(\bar{\epsilon}_m^p)_o$ and $(\epsilon_{33}^{p\infty} - \epsilon_{11}^{p\infty})_o$ are the plastic strains at which dilatational and extensional void growth common, respectively. Note that the integration of equations (5) and (6) in this manner implies that equations (3) and (4) can be extended to finite strains. Rice and Tracey (1969) make no provision for finite strains and equations (5) and (6) merely provide a means of estimating the effective void growth rates during the experiments.

5.3.1 Dilatational Growth Rate. Values of $\ln[(f/f^o)(1-f^o)/(1-f)]$ versus $\bar{\epsilon}_m^p$ at the center of the tensile specimens are plotted in Fig. 11 for the tensile tests interrupted prior to fracture. Note that while f is a measured quantity,

Table 2 Summary of least squares fits to dilational growth data

Specimen	N	D	$(\bar{\epsilon}_m^p)_o$	corr.^{2}
Uniaxial	4	0.397	0.142	0.971
A-notch	4	0.843	0.131	0.944
D-notch	3	3.224	0.072	0.991

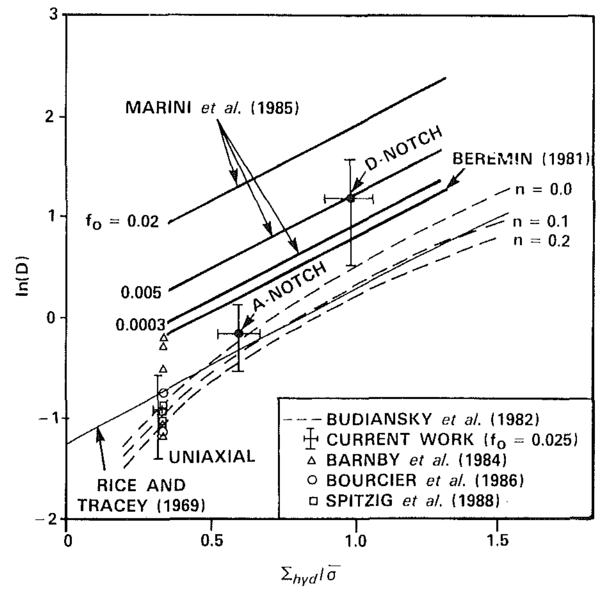


Fig. 12 Dilatational growth factor (D) from tensile test (crosses) as a function of stress triaxiality ($\Sigma_{hyd}/\bar{\sigma}$). The vertical span corresponds to the uncertainty in D while the horizontal span corresponds to the range in predicted stress triaxiality during the test.

$\bar{\epsilon}_m^p$ was determined from the finite element calculations. The linear trends seen in the data in Fig. 11 support the assumption of a constant value of D in integrating equation (3). There was also a noticeable delay in the onset of dilatational void growth as reflected by the finite $\bar{\epsilon}_m^p$-intercept. The dilatational growth factor, D and nucleation strain, $(\bar{\epsilon}_m^p)_o$, were estimated using least squares fits of equation (5) to the data in Fig. 11. The resulting values for D and $(\bar{\epsilon}_m^p)_o$ are summarized in Table 2 and the corresponding fits are plotted in Fig. 11. The correlation coefficients in the last column of Table 2 and the observed fit in Fig. 11 indicate that equation (5) provides a reasonable description of the experimental void growth behavior.

The finite element predictions of $\ln[(f/f^o))(1-f^o)/(1-f)]$ versus $\bar{\epsilon}_m^p$ are also plotted in Fig. 11. These curves pass through the origin due to the assumption of a zero void nucleation strain and, as a result, lie above the corresponding experimental data. The measured rates of void growth, however, were higher than the predicted rates as seen in the steeper slopes of the fits to the experimental data compared to the predicted curves. In spite of the lower predicted growth rates, the Gurson-based finite element model captured the trend of increased dilatational growth rate with triaxiality quite well.

5.3.2 Effect of Stress Triaxiality on D. The dependence of the measured dilatational void growth rate on stress triaxiality is examined in Fig. 12 (crosses). The natural logarithm

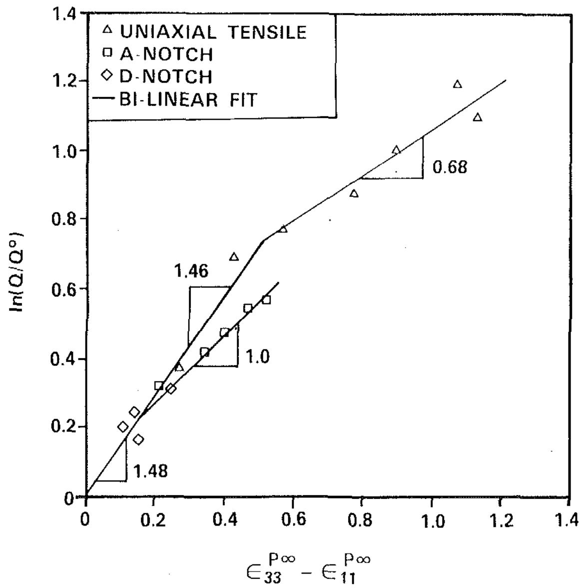


Fig. 13 Extensional growth histories from the interrupted tensile specimens. Solid lines are the curve fits obtained using equation (6).

of D was plotted in view of the exponential dependence of dilatational growth rate on triaxiality predicted by Rice and Tracey (1969). The vertical span of each cross indicates the uncertainty in D, calculated by Worswick (1988), whereas the horizontal span corresponds to the predicted range of $\Sigma_{hyd}/\bar{\sigma}$ during the test (Fig. 5). The figure shows the strong dependence of D on $\Sigma_{hyd}/\bar{\sigma}$.

The relatively large uncertainty in D, particularly at high triaxiality, was due to the limited number of specimens tested and the small plastic strains attained in the notched specimens. In spite of this uncertainty, the agreement between the present low triaxiality uniaxial results and those of Barnby et al. (1984), Bourcier et al. (1986) and Spitzig et al. (1988) was quite good, as seen in Fig. 12. At high triaxiality, the straight line fits to measurements by Beremin (1981) and Marini et al. (1985) exceeded the notched specimen growth rates in the current work. In contrast, the analytical predictions for D by Rice and Tracey (1969) and Budiansky et al. (1982) are less than the experimental values, particularly at high triaxiality levels.

5.3.3 Extensional Growth Rate. The relationship between the void aspect ratio and applied plastic strain suggested by equation (6) is examined in Fig. 13 using the measured values of $\bar{Q}$ and the predicted plastic strains at the specimen center from the finite element calculations. Examination of the figure reveals that the extensional growth rate, $1 + E$, is not a constant but decreases during deformation, as indicated by a decrease in the slope of the data at higher plastic strains. The figure also shows that extensional growth of the voids commenced immediately with the onset of plastic deformation. The early extensional growth occurs through elongation of the soft lead particles prior to their rupturing or tearing; thus the nucleation strain term, $(\epsilon_{33}^{p\infty} - \epsilon_{11}^{p\infty})_o$ in equation (6), would be zero.

The variation in $1 + E$ during the deformation history was estimated by fitting a bilinear form of equation (6) to the extensional growth histories shown in Fig. 13. The resulting slopes, shown in the figure, indicate an initial value of $1 + E$ in the range 1.45–1.5 for both the uniaxial and notched specimens, decreasing during the later stages of deformation to roughly 0.7 for the uniaxial and 1.0 for the notched specimens. These values compare well to uniaxial measurements for $1 + E$ in the range 1.3–1.6 by Leroy et al. (1981) and 0.8–1.1 determined using data published by Spitzig et al. (1988).

6 Discussion and Conclusion

No direct experimental evaluation of the degree of constitutive softening was possible since the stresses and stress triaxiality at the specimen center could not be measured. Instead, the measured axial stress-strain response of the test specimens was used to assess the performance of the Gurson and Von Mises constitutive models. The predicted stress-strain response using the Gurson-based model was in closer agreement with experiment than predictions using the Von Mises constitutive model for the stress triaxiality range $1/3 < \Sigma_{hyd}/\bar{\sigma} < 1.1$.

A question arises concerning the dependence of void nucleation within soft second-phase particles on the particle strength. The experimental results indicated a delay in the onset of dilatational void growth that was unexpected since the strength of the lead particles was low compared to the strength of the brass matrix. The majority of void nucleation studies have examined hard second-phase particles or inclusions which are brittle and nucleate voids once the local stresses acting on the particle exceed the particle strength or the strength of the particle-matrix interface (Goods and Brown, 1979; Fisher and Gurland; 1981). In the current study, the local shear stresses acting on the particles can be relieved by plastic flow of the ductile lead phase; thus plastic strains well in excess of the particle and matrix yield strains are attained prior to rupture of the particles. Void nucleation within the particles becomes highly dependent on the hydrostatic component of stress, as reflected in the decreased nucleation strain for the higher triaxiality notched specimens (Table 2). Direct verification of the reported void nucleation strains was not obtained since the lead phase smears during polishing, covering over cracks in the lead particles at small strains. However, experiments performed subsequent to those reported here[2] show no significant change in porosity within uniaxial specimens prior to strains of 20 percent, while the void aspect ratios increase steadily with applied plastic strain.

Once void expansion had begun, the dilatational void growth rates exceeded the predicted levels, in agreement with trends seen by Beremin (1981) and Marini et al. (1985). The decrease in the estimated extensional growth rate with stress triaxiality was in agreement with predictions by Budiansky et al. (1982) and Worswick and Pick (1989b) but was opposite to the trends predicted by Rice and Tracey (1969). Note that Rice and Tracey (1969) did obtain solutions in which $1 + E$ decreased with stress triaxiality; however, these solutions were discarded as being "nonintuitive."

In summary, quantitative comparisons between the Gurson-based finite element predictions and the measured void growth were hindered by the assumption of a zero void nucleation strain and by the variation in the measured void volume fraction within the specimens. In spite of these difficulties, the Gurson (1975, 1977) model was found to capture the specimen stress-strain response and void growth behavior reasonably well. The Gurson constitutive model represents a useful tool for describing the void growth stage of ductile fracture that can be readily incorporated into existing finite element programs.

7 Acknowledgment

The authors wish to thank Prof. P. Niessen and Mr. M. Kozdras of the University of Waterloo for their advice and assistance concerning metallurgical aspects of this work. Financial support for this work was provided by the Natural Sciences and Engineering Research Council of Canada.

[2]Work in progress.

References

Aravas, N., 1987, *Int. J. Num. Meth. Engng.*, Vol. 24, pp. 1395–1416.

Barnby, J. T., Shi, Y. W., and Nadkarni, A. S., 1984, *Int. J. Frac.*, Vol. 25, pp. 273–283.

Becker, R., Needleman, A., Richmond, O., and Tvergaard, V., 1988, *J. Mech. Phys. Sol.*, Vol. 36, pp. 317–351.

Beremin, F. M., 1981, *Advances in Fracture Research*, ICF5, Vol. 2, pp. 809–816.

Bourcier, R. J., Koss, D. A., Smelser, R. E., and Richmond, O., 1986, *Acta Metall.*, Vol. 34, pp. 2443–2453.

Budiansky, B., Hutchinson, J. W., and Slutsky, S., 1982, *Mechanics of Solids, The Rodney Hill 60th Anniversary Volume*, H. G. Hopkins and M. J. Sewell, eds., Pergamon Press, Oxford, U.K., pp. 13–45.

DeHoff, R. T., 1964, *AIME, Trans. Metall. Soc.*, Vol. 230, pp. 764–769.

Fisher, J. R., and Gurland, J., 1981, *Met. Sci.*, Vol. 15, pp. 185–202.

Goods, S. H., and Brown, L. M., 1979, *Acta Metall.*, Vol. 27, pp. 1–15.

Gurson, A L., 1975, Ph.D. Thesis, Brown University, Providence, R. I.

Gurson, A. L., 1977, *J. Engng. Mat. Tech.*, Vol. 99, pp. 2–15.

Hancock, J. W., and Brown, D. K., 1983, *J. Mech. Phys. Sol.*, Vol. 31, pp. 1–24.

Hancock, J. W., and MacKenzie, A. C., 1976, *J. Mech. Phys. Sol.*, Vol. 24, pp. 147–169.

Hibbitt, H. D., Karlsson, B. I., and Sorensen, E. P., 1984, *ABAQUS Theory Manual*, Hibbitt, Karlsson and Sorensen Inc., Providence, R. I.

Koplik, J., and Needleman, A., 1988, *Int. J. Sol. Struct.*, Vol. 24, pp. 835–853.

Leroy, G., Embury, J. D., Edward, G., and Ashby. M. F., 1981, *Acta Metall.*, Vol. 29, pp. 1509–1522.

Marini, B., Mudry, F., and Pineau, A., 1985, *Engng. Frac. Mech.*, Vol. 22, pp. 989–996.

McClintock, F. A., 1968, ASME Journal of Applied Mechanics, Vol. 35, pp. 363–371.

McMeeking, R. M., and Rice, J. R., 1975, *Int. J. Sol. Struct.*, Vol. 11, pp. 601–616.

Needleman, 1972, ASME Journal of Applied Mechanics, Vol. 39, pp. 964–970.

Needleman, A., and Tvergaard, V., 1984, *J. Mech. Phys. Sol.*, Vol. 32, pp. 461–490.

Rice, J. R., and Tracey, D. M., 1969, *J. Mech. Phys. Sol.*, Vol. 17, pp. 201–217.

Rogers, H. C., 1960, *Trans. Metall. Soc. AIME*, Vol. 218, pp. 498–506.

Spitzig, W. A:, Smelser, R. E., and Richmond, O., 1988, *Acta Metall.*, Vol. 36, pp. 1201–1211.

Tvergaard, V., 1981, *Int. J. Frac.*, Vol. 17, pp. 389–407.

Tvergaard, V., 1982, *Int. J. Frac.*, Vol. 18, pp. 237–252.

Tvergaard, V., and Needleman, A., *Acta Metall.*, Vol. 32, pp. 157–169.

Underwood, E. E., 1970, *Quantitative Stereology*, Addison-Wesley Publishing Co., Reading, Mass.

VanderVoort, G. F., 1984, *Metallography: Principles and Practice*, McGraw-Hill, New York.

Worswick, M. J., 1988, Ph.D. Thesis, University of Waterloo, Waterloo, Canada.

Worswick, M. J., and Pick, R. J., 1989a, "Gurson Yield Function Calibration Based on Three-Dimensional Unit Cell Calculations," *Proc. ABAQUS User's Conference*, Hibbitt, Karlsson and Sorensen Inc., Stressa, Italy, May 24–26.

Worswick, M. J., and Pick, R. J., 1989b, *J. Mech. Phys. Sol.*, accepted for publication.

A. E. Giannakopoulos
Department of Strength of Materials
and Solid Mechanics,
The Royal Institute of Technology,
S-100 44 Stockholm, Sweden

Dynamic Damage in Certain Monolithic Ceramic Materials

In the present work, the propagation of elasto-damage longitudinal stress waves in thin rods is investigated. The material behavior is characteristic to that of certain monolithic ceramics. The damage constitutive relation that characterizes this type of materials gives rise to certain dynamic behavior which is somewhat different from dynamic plastic behavior. Plastic and damage dynamic response are compared through an example.

1 Introduction

Ceramic materials at low temperatures, while not undergoing plastic deformations, can behave inelastically as a result of microcrack nucleation. In certain types of monolithic ceramic materials, such as polycrystalline and multiphase ceramics, microcracks develop mainly at grain facets as a result of residual stresses generated during cooling and of applied tensile stress (Fu, 1983). Microcracks tend to develop normal to the direction of maximum tension (normal microcracking). This gives rise to degradation of the elastic properties of the material. As the number of microcrack nucleation sites is exhausted a saturation stage ensues, during which the material sustains no further damage (Ortiz, 1987). We will refer to this type of materials as of damage type in order to distinguish them from those of plastic type.

Some important reasons to study the influence of dynamics in the mechanical behavior of ceramics are the tensile testings, measuring inelastic properties, and the loading resulting from compressive waves reflected on free surfaces of such materials. Although a lot has been done in the dynamic behavior of metals, very little has been done for ceramics. There is some analogy between ceramic and metallic materials. However, the treatment of the ceramics as classical plastic materials is not appropriate. As we will show in the following, the dynamic loading of ceramics exhibits some peculiarities due to the stress-strain concavity, whereas the case of dynamic unloading has a completely different mathematical modeling. In the present work we use a constitutive law suitable to describe microcracking. The equilibrium equations are then formulated and solved with the method of characteristics. A modified grapho-analytical method is then introduced, suitable for solving particular problems.

2 Dynamic Loading and Unloading of Ceramics

Figure 1(*a*) shows a typical stress-strain behavior of this type of ceramic, whereas Fig. 1(*b*) shows a trilinear idealization for this material behavior. In the tensile region of the stress-strain curve that will be investigated in the present work, we can distinguish three regions: (a) an elastic region with elastic modulus E_0, (b) a transition region with elastic tangent modulus E_T, and (c) a saturation region with elastic modulus E_S. Note that $E_0 > E_S > E_T$ and all the elastic moduli are positive.

We will consider waves propagating in a thin semi-infinite prismatic rod, whose axis coincides with the positive x-axis of space. We will assume the following kinematic conditions:

1 The cross-section of the rod remains plane and normal to the x-axis during deformation.

2 The deformations are small (small strains).

3 There is no influence from strain rates.

Furthermore, we will assume that the rod is initially at rest and at the end of the rod we impose some tensile stress history. Therefore, the initial conditions for the strain $\epsilon(x, t)$ and the velocity $v(x, t)$ will be

$$\epsilon(x \geq o,\ t=0)=0 \tag{1a}$$

$$v(x \geq 0,\ t=0)=0 \tag{1b}$$

and the boundary conditions for the stress $\sigma(x, t)$

$$\sigma(x=0,\ t \geq 0)=\sigma(0,\ t) \geq 0 \tag{2}$$

The equation of motion for the rod is

$$\frac{\partial \sigma}{\partial x}=\rho\,\frac{\partial^2 u}{\partial t^2}. \tag{3}$$

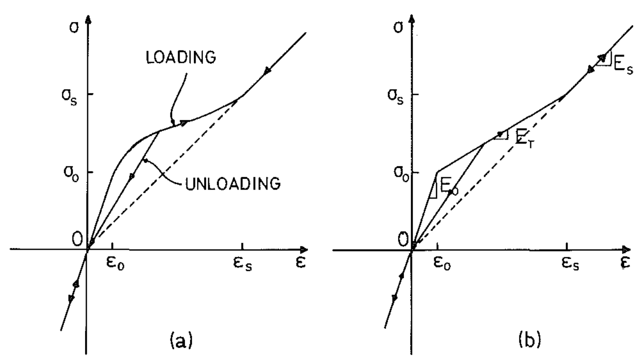


Fig. 1 (*a*) Stress-strain behavior of monolithic ceramics; (*b*) trilinear idealization

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 6, 1989; final revision, June 5, 1990.

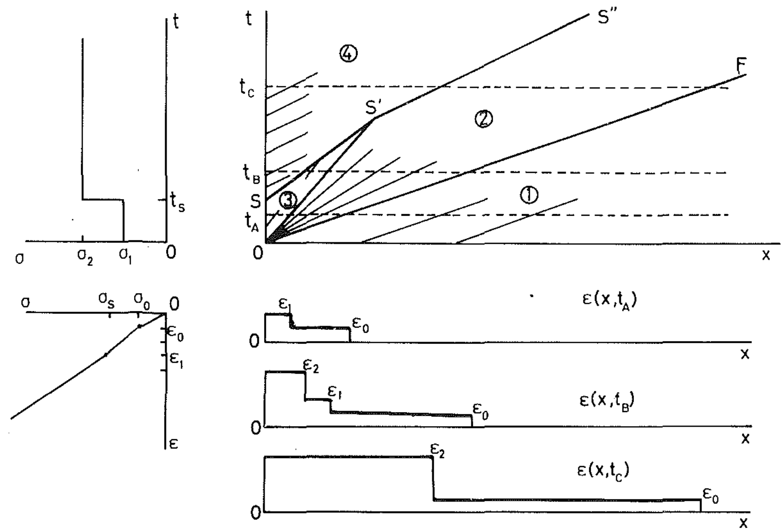


Fig. 2 Characteristic fields for successive impact loading and shock formation

and the compatibility equation for this one-dimensional problem is

$$\epsilon = \frac{\partial u}{\partial x} \tag{4}$$

where ρ is the density of the rod (assumed constant) and $u(x, t)$ is the displacement. We will try to find solutions for the strains and the displacements for positive space and time (x, $t \geq 0$).

We will assume that the loading function $\sigma(0, t)$ is an increasing function of time. Then, the finite constitutive equation for loading can be written as

$$\epsilon = \epsilon(\sigma) = \epsilon_e + \epsilon_d \tag{5}$$

where the elastic part of the strain is

$$\epsilon_e = \sigma / E_0 \tag{6a}$$

and the damage part of the strain is

$$\epsilon_d = \begin{cases} 0 & ; \quad \sigma \leq \sigma_0 \\ (\sigma - \sigma_0)\left(\dfrac{1}{E_T} - \dfrac{1}{E_0}\right) & ; \quad \sigma_0 \leq \sigma \leq \sigma_S \, \cdot \\ \left(\dfrac{1}{E_S} - \dfrac{1}{E_0}\right)\sigma & ; \quad \sigma_S \leq \sigma \end{cases} \tag{6b}$$

Denote by $c(\epsilon)$ the local velocity of propagation of disturbances, which is given by

$$c(\epsilon) = (1/\rho)^{1/2}(d\sigma(\epsilon)/d\epsilon)^{1/2}. \tag{7}$$

In this case $c(\epsilon)$ is

$$\begin{aligned} c_0 &= (E_0/\rho)^{1/2} \quad ; \sigma < \sigma_0 \\ c_T &= (E_T/\rho)^{1/2} \quad ; \sigma_0 < \sigma < \sigma_S \\ c_S &= (E_S/\rho)^{1/2} \quad ; \sigma_S < \sigma. \end{aligned} \tag{8}$$

Note that $c(\epsilon) > 0$ and $c_0 > c_S > c_T$ always.

The solution to equations (3), (4) can be stated in characteristic form as

$$dv = \pm c(\epsilon) d\epsilon \tag{9}$$

along the characteristic lines

$$dx = \pm c(\epsilon) dt. \tag{10}$$

In equations (9) and (10), the plus sign indicates a forward wave and the minus sign indicates a backward wave.

The solution given by equations (9) and (10) is the same as for the plastic loading behavior (Cristescu, 1967) for $\sigma(0, t) \leq \sigma_S$. Upon further increase of load we encounter a region where the stress-strain curve is concave toward the direction of increasing stress. Then the characteristic lines show a convergent bundle that will form a stress shock. This damage shock wave is formed because, due to the specific constitutive behavior of the material, the distance between the smooth wave fronts propagating in the rod decreases during propagation. Then, the waves carrying the largest strains tend to overtake all the others.

Suppose that at time t a given section x of the rod is reached by a shock front which moves with velocity c_D in the positive direction of the x-axis. The jump conditions across the shock front are the well-known Hugoniot relations, and relate the jumps in velocity $[v]$, strain $[\epsilon]$, and stress $[\sigma]$. These relations are

$$[v] = -c_D\,[\epsilon] \tag{11a}$$

$$c_D\,\rho[v] = -[\sigma]. \tag{11b}$$

Combining them, we obtain the velocity of propagation of the shock

$$c_D = ([\sigma]/([\epsilon]\rho))^{1/2}. \tag{12}$$

A particular type of loading history is shown in Fig. 2. In this particular example the shock line in the xt-plane is linear with constant velocity for the propagation of the shock front. In Fig. 2 the strains for different time levels t_A, t_B, and t_C are also shown. It should be noted here that this shock is due to the particular constitutive behavior, and if the material were simply plastic, the shock line would have been a characteristic line with $c_D = c_T$. The stresses can be easily computed from equation (6) and the displacements can be computed by integrating the strains along Ox-axis. The characteristic lines are also shown in Fig. 2. We can distinguish four regions and four lines that separate them (see Fig. 2):

In region 1: $dx/dt = c_0$
In region 2: $dx/dt = x/t$
In region 3: $dx/dt = c_T$
In region 4: $dx/dt = c_S$
Along line OF: $x = c_0\,t$
Along line OS': $x = c_T\,t$
Along line SS': $dx/dt = c_D = \{(\sigma_2 - \sigma_1)/((\epsilon_2 - \epsilon_1)\rho)\}^{1/2}$
Along line $S'S''$: $dx/dt = c_S$
At point S': $x_{S'} = c_D\,c_T\,t_S/(c_D + c_T)$, $t_{S'} = c_D\,t_S/(c_D + c_T)$

We may define the measure of cumulative damage by a scalar

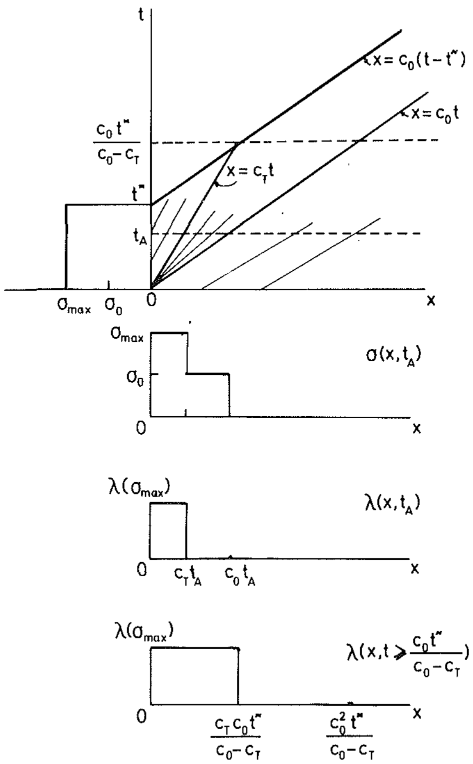


Fig. 3 Evolution of damage distribution due to a tensile impulse

λ, so that $\lambda^{\cdot}=(\epsilon/\sigma)^{\cdot}$ be the rate of change of the compliance of the material due to microcrack formation. The microcracks are formed parallel to the cross-section of the rod. Therefore, λ gives a measure of the microcrack density. The tensile stress is itself a function of the state of damage through the cumulative damage parameter λ. In the present model, in case of loading, λ can be given analytically by integrating $\lambda^{\cdot}=(\epsilon/\sigma)^{\cdot}$ with initial condition $\lambda(\sigma_0)=0$. (Ortiz and Giannakopoulos, submitted for publication):

$$\lambda(\sigma)=\left(1-\frac{\sigma_0}{\sigma}\right)\left(\frac{1}{E_T}-\frac{1}{E_0}\right)\ ;\ \sigma_0<\sigma<\sigma_S. \tag{13}$$

Upon unloading, the damage coefficient λ attains pointwise the maximum value suffered during the loading. Since the stress σ is a function of the position x along the rod at every instant t, then equation (13) can predict the damage distribution in the rod at every time t. Therefore, by solving for the stress in the loading region of the rod, we can compute the extent of damage in an initially undamaged rod. The damage evolution is described with the help of the following example. A tensile stress pulse of amplitude $\sigma_{\max}$ ($\sigma_0<\sigma_{\max}<\sigma_S$) and of duration t^* is imposed at the end of the rod (see Fig. 3). The rod is assumed to be initially at rest, unconstrained and intact. When the stress pulse starts propagating in the rod, microcracks are created at the wake of its front. The solution in characteristic form is shown in Fig. 3. In the same figure, the damage distribution $\lambda(x,t)$ is shown for $t=t_A$, $(0<t_A<t^*)$ and for $t>c_0t^*/(c_0-c_T)$. Therefore, the stress pulse creates a microcrack distribution which would have a density given by $\lambda(x,\ t>c_0t^*/(c_0-c_T))$, shown in Fig. 3.

The unloading ($\partial\sigma(0,t)/\partial t<0$) of the elasto-damage type of materials can be very different from the unloading of the elasto-plastic type of materials, as we may see from the following analysis. We will assume a perfectly damaged type of material where the unloading lines pass through the origin of the stress-strain curve (Fig. 4). For each section x of the bar, let $\sigma_m(x)$ and $\epsilon_m(x)$ be the maximum stress and maximum strain, correspondingly. Then, the constitutive equation for unloading becomes

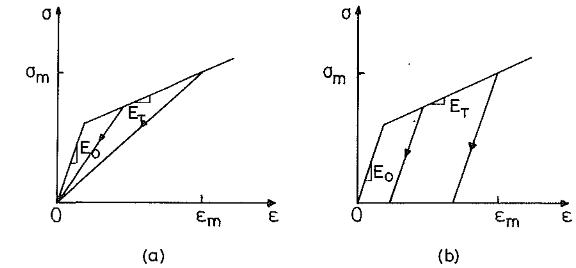


Fig. 4 (*a*) Perfect elasto-damage unloading; (*b*) Perfect elasto-plastic unloading

$$\sigma=\epsilon\,\frac{\sigma_m(x)}{\epsilon_m(x)}\ ;\ 0<\sigma_m<\sigma_S. \tag{14}$$

In the characteristic plane xOt, there will be two domains: a loading and an unloading one. The two domains are separated by a loading/unloading boundary $t=f(x)$. It is necessary to determine the solution in the unloading domain simultaneously with the solution in the loading domain (in order to find $\sigma_m(x)$ and $\epsilon_m(x)$).

The equations of motion for the unloading domain can be computed from equations (3) and (14), and can be written as a second-order partial differential equations in terms of the displacements u

$$\rho\ u_{tt}-\left(\frac{\sigma_m(x)}{\epsilon_m(x)}\right)u_{xx}-\frac{d}{dx}\left(\frac{\sigma_m(x)}{\epsilon_m(x)}\right)u_x=0. \tag{15}$$

Since $\sigma_m/\epsilon_m>0$ and $\rho>0$, equation (15) is of hyperbolic type. The characteristic lines for this equation are:

$$\frac{dx}{dt}=\pm\sqrt{\frac{\sigma_m(x)}{\rho\ \epsilon_m(x)}}. \tag{16}$$

Equation (15) can be written in canonical form by using new variables ξ and η, which are connected to x and t by

$$\xi=t-\rho^{1/2}\int(\epsilon_m/\sigma_m)^{1/2}dx \tag{17a}$$

$$\eta=t+\rho^{1/2}\int(\epsilon_m/\sigma_m)^{1/2}dx. \tag{17b}$$

The canonical form of equation (15) is a wave equation of the form

$$u_{\xi\eta}(\xi,\ \eta)=0. \tag{18}$$

The general solution of (15) is then given by

$$u(x,\ t)=F_1\left\{t-\rho^{1/2}\int(\epsilon_m/\sigma_m)^{1/2}dx\right\}+F_2\left\{t+\rho^{1/2}\int(\epsilon_m/\sigma_m)^{1/2}dx\right\} \tag{19}$$

where F_1 and F_2 are twice differentiable arbitrary functions to be determined by the boundary conditions at the end of the bar and the loading/unloading boundary $t=f(x)$. The differential relations to be satisfied along the characteristic lines are

$$dv=\pm\left(\frac{\epsilon_m(x)}{\rho\ \sigma_m(x)}\right)^{1/2}d\sigma. \tag{20}$$

On the other hand, in plastic-type unloading, the constitutive relation is

$$\sigma=\sigma_m(x)+E_0\{\epsilon-\epsilon_m(x)\}. \tag{21}$$

The dynamic equation for plastic unloading will then be

$$u_{tt}=c_0^2u_{xx}+\frac{d}{dx}\left\{\frac{\sigma_m(x)}{\rho}-c_0^2\ \epsilon_m(x)\right\}. \tag{22}$$

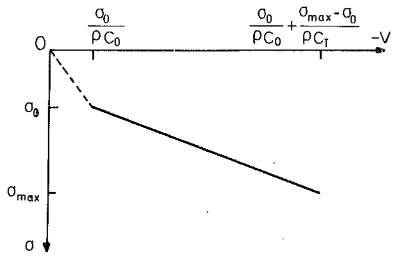


Fig. 5 Hodographic plane $vO\sigma$ associated with the characteristic field xOt

The characteristic lines for equation (22) are given by

$$\frac{dx}{dt} = \pm c_0. \tag{23}$$

The general solution of equation (22) is then of the form

$$u(x, t) = G_1(x + c_0 t) + G_2(x - c_0 t) - \frac{1}{E_0}\int \{\sigma_m(x) - E_0\,\epsilon_m(x)\}\,dx, \tag{24}$$

and the differential relations along the characteristic lines should be

$$dv = \pm\frac{1}{\rho\, c_0}\,d\sigma. \tag{25}$$

Obviously, there is a different dynamic behavior between damage and plastic type of unloading. In the plastic type of behavior, the characteristic lines in the unloading region are straight, whereas in the damage type of behavior, the characteristic lines are still parallel, but curved. The differences will become clearer when we will examine the loading/unloading boundary.

The shape of the loading/unloading boundary $t = f(x)$, can be determined by a grapho-analytical method similar to the one developed by Shapiro and Biderman (Cristescu, 1967). The method can be applied to long bars initially at rest and undeformed, and is described in the Appendix. It will be assumed that the stress at the end of the bar increases up to a maximum value $\sigma_{\max}$ and then decreases to zero ($\sigma_0 < \sigma_{\max} < \sigma_S$). The loading/unloading boundary emerges at the end of the bar, at the instant t_0, when the stress reaches its maximum. At this point ($x = 0$, $t = t_0$), by locally expanding the stress and velocity, we can compute the initial velocity of propagation of the loading/unloading boundary. The result is

$$\left.\frac{dx}{df(x)}\right|_{x=0} = \left(\frac{c_T^2\, c_u^2\,(k_1 - k_2)}{c_u^2\, k_1 - c_T^2\, k_2}\right)^{1/2} \tag{26}$$

where

$$k_1 = \frac{d\sigma}{dt}\,(x = 0,\ t = t_0 -) \geq 0$$

$$k_2 = \frac{d\sigma}{dt}\,(x = 0,\ t = t_0 +) \leq 0$$

$$k_1^2 + k_2^2 \neq 0 \tag{27}$$

and

$$c_u = \left(\frac{1}{\rho}\frac{\sigma_{\max}}{\epsilon_{\max}}\right)^{1/2} < c_0 \tag{28a}$$

$$\epsilon_{\max} = \frac{\sigma_{\max}}{E_0} + \frac{\sigma_{\max} - \sigma_0}{E_T}. \tag{28b}$$

In cases where $k_1 = k_2 = 0$ and $d^2\sigma(0, t)/dt^2$ is continuous at $t = t_0$, we have

$$\left.\frac{dx}{df(x)}\right|_{x=0} = c_u\left(\left(\frac{c_u^2}{c_T^2} + 3\right)^{1/2} - \frac{c_u}{c_T}\right). \tag{29}$$

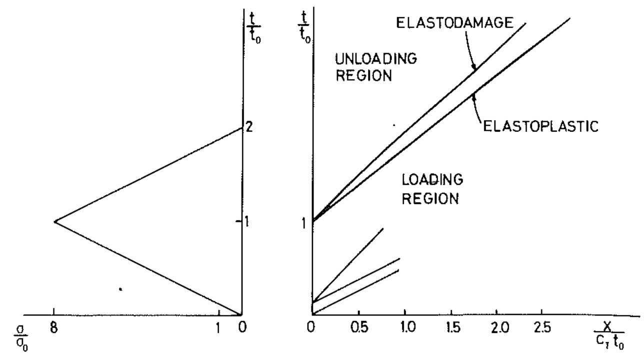


Fig. 6 The loading/unloading boundary for the plastically unloading and for the damage type of unloading materials

In plasticity, equations (26) and (29) also hold (Biderman's forms), but with c_0 in place of c_u. Since $c_u < c_0$, it can be easily shown that the initial velocity of propagation of the loading/unloading boundary is smaller for the damage type of behavior than that of the plastic type.

In order to construct the loading/unloading boundary, it is useful to have its image in the hodographic plane $vO\sigma$. In the hodographic plane, the loading constitutive equation can be stated as:

$$-v(\sigma) = \int_0^{\sigma} \frac{d\sigma}{\rho\, c(\sigma)}. \tag{30}$$

In our present model

$$-v = \frac{\sigma_0}{\rho\, c_0} + \frac{\sigma - \sigma_0}{\rho\, c_T}. \tag{31}$$

Therefore, the image of $t = f(x)$ in the $v\sigma$-plane is a straight line, shown in Fig. 5.

3 Example

In order to show the difference between the damage type and the plastic type of unloading, we will assume the same boundary and initial conditions and the same loading stress-strain behavior to apply in both cases. The only difference will be in the way the two types of material unload (Fig. 4).

The initial conditions will be $\epsilon = 0$, $\sigma = 0$. The boundary condition is a triangular pulse with $\sigma_{\max}/\sigma_0 = 8$, and is shown in Fig. 6. The material parameters used were $E_T/E_0 = 0.25$, $\rho = 1$, $\sigma_0 = 1$, $E_0/\sigma_0 = 4$. For the elastoplastic unloading we can use the grapho-analytical method proposed by Shapiro and Biderman (Cristescu, 1967). Note that the initial slope for the loading/unloading curve is $dx/dt = 0.6325\ c_0$. The loading/unloading curve for the elastoplastic response is shown in Fig. 6.

For determining the loading/unloading curve of the damage type of behavior, the aforementioned method can be used with the following modification: A net of characteristic curves given by equation (16) must also be constructed, emanating from the unloading part of the stress history boundary condition (see Appendix for details). The loading/unloading curve for the elastodamage response is shown in Fig. 6. The initial slope is $dx/dt = 0.5126\ c_0$. This indicates that the loading/unloading front propagates slower for the damage type of unloading than for the plastic type, as predicted from the previous analysis (equation (29)).

4 Conclusions

In case of dynamic loading up to the saturation level, we point out again the analogies between ceramic and metallic materials. These are expressed from equations (7), (8), (9), and

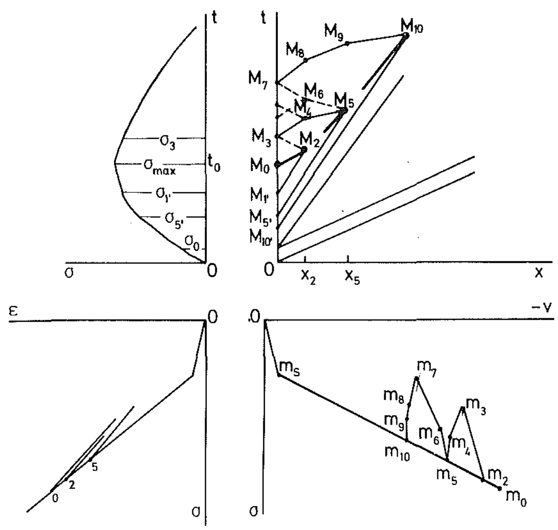


Fig. 7 The modified grapho-analytical method of Shapiro-Biderman used for the elasto-damage dynamic unloading

(10). The governing equation and its solution are similar to that of the elastoplastic materials and the characteristics are straight lines.

This is not true in case of unloading. The loading/unloading front propagates slower in the damage than in the plastic type of materials. For this reason the extent of the microcracked region is smaller than if it had to be computed from standard elastoplastic results. The governing equation (15) and its general solution (19) are very different from the ones that hold for the plastic type of materials (equations (22) and (24)). The characteristic lines in the unloading region are curved and are given by equations (17). The proposed grapho-analytical method described in Appendix is actually constructing these characteristics which depend on the specific dynamic boundary conditions. The method can be implemented numerically as well.

Pulse waves that do not exceed the saturation level, σ_s, can be used in dynamic tests to predict inelastic properties of ceramics. Loading beyond the saturation level requires confinement of microcracking. Shock conditions may then take place (as shown in Fig. 2) where the information from the saturation state travels faster and finally dominates over the transition region. Therefore, the extent of damage, as expressed by equation (13), is substantial when the duration of the pulse is sufficient. For this reason, we believe that uniaxial dynamic tests of ceramics must be performed at low stresses in cases where confinement of damage cannot be secured.

Acknowledgments

The author would like thank Profs. P. Gudmundson and B. Storakers for their valuable observations on this paper.

References

Cristescu, N., 1967, *Dynamic Plasticity*, North-Holland, Amsterdam.

Fu, Y., 1983, "Mechanics of Microcrack Toughening in Ceramics," Ph.D. Thesis, University of California, Berkeley, Calif.

Ortiz, M., 1987, "A Continuum Theory of Crack Shielding in Ceramics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 54–58.

Ortiz, M., and Giannakopoulos, A. E., 1991, "Crack Propagation in Monolithic Ceramics under Small Scale Mixed Mode Loading," submitted for publication.

APPENDIX

The grapho-analytical method used in the construction of the elasto-damage loading/unloading curve can be summarized as follows (see Fig. 7). First, we draw the initial slope of the loading/unloading boundary, given by equation (26) or equation (29). In this direction we choose a point M_2 close to the point $M_0(x=0,\ t=t_0)$. At point M_2, we can find the stress by drawing the characteristic line in the loading region $(dx/dt=c_T)$. Then, $\sigma_2=\sigma_{1'}$, where $\sigma_{1'}$ is the stress from the loading part of the curve $\sigma(0,\ t)$. Since we know σ_2, we can locate the point m_2 in the hodographic plane $vO\sigma$. The point m_2 must lie on the curve given by equation (31), since it belongs to the loading/unloading curve. Therefore, the slope of the characteristic lines in the unloading region for all $x=0$, will be given by equation (26) or equation (29).

We then draw the characteristic line m_2m_3, since we know the slope at the point M_0. From the point M_2 we draw the line M_2M_3 which corresponds to m_2m_3. The point m_3 can be located, since we know the stress σ_3 from the unloading branch of the curve $\sigma(0,\ t)$. Then, from point M_3, we draw a line M_3M_4 with the same slope as of M_0M_3. The point M_4 is located at the $x=x_2$-coordinate. The stress at the point M_4 can be found from the boundary conditions by writing equation (20) in a different form. Clearly, $\sigma_m(x_2)=\sigma_2$, and from the constitutive behavior we can find $\epsilon_m(x_2)=\epsilon_2$. Therefore, the slope of the characteristic lines for all points with $x=x_2$ in the unloading region would be given by $dx/dt=(\sigma_2/(\rho\ \epsilon_2))^{1/2}$. In the plane $vO\sigma$ we draw the line m_3m_4, which is the image of the line M_3M_4. The point m_4 can again be found, since we know σ_4. From the point M_4, we draw the line M_4M_5 which has a slope $(\sigma_2/(\rho\ \epsilon_2))^{1/2}$. We then draw the line m_4m_5 (which is the image of M_4M_5) until it intersects the line m_5m_0 in the $v\sigma$-plane. From the point m_5, we can find the stress σ_5 and then find the point $M_{5'}$ for which $\sigma_{5'}=\sigma_5$. From the point $M_{5'}$ we draw the loading characteristic line $(dx/dt=c_T)$.

The lines M_4M_5 and $M_{5'}M_5$ intersect at the point M_5, which belongs to the loading/unloading curve. The method continues as before for the point M_{10}, etc., and is shown schematically in Fig. 7.

P. Aravamadhu Balaji
Graduate Student.

T. Sundararajan
Assistant Professor.

G. K. Lal
Professor,
Mem. ASME

Department of Mechanical Engineering,
Indian Institute of Technology,
Kanpur 208016, India

Viscoplastic Deformation Analysis and Extrusion Die Design by FEM

A viscoplastic model for extrusion is discussed which simultaneously predicts the deformation field, optimal die geometry, and plastic boundaries. The die geometry and plastic boundaries are expressed in terms of chosen trial functions that satisfy certain geometrical and physical constraints. The variational power integral is minimized in the trial plastic domain using FEM technique to determine the deformation field and shape coefficients for the die contour and plastic boundaries. The proposed method is implemented for the optimal design of an axisymmetric streamlined die. The predicted values are in reasonable agreement with the experimental observations and are in conformity with the results published earlier.

Introduction

In forming processes such as extrusion, the geometry of the die constitutes an important aspect of die design. The die profile determines the extent of redundant work done during deformation. An optimum die profile minimizes the redundant work, thereby minimizing the extrusion power. In the past, several attempts have been made to obtain approximate solutions for the optimal shapes of straight or curved dies using the upper bound technique (Chen and Ling, 1968; Zeev Zimerman and Avitzur, 1970; Gunasekara and Hoshino, 1985; and Yang et al., 1985). These solutions have been derived by assuming kinematically admissible velocity fields which merely satisfy the incompressibility and velocity boundary conditions. A major drawback of the upper bound technique is that the velocity field does not satisfy stress equilibrium everywhere in the deformation zone. For this reason, the predicted optimal shape tends to be overly conservative (Aravamadhu Balaji, Sundararajan, and Lal, 1989). Another traditional approach used for the evaluation of optimal die shapes is the slip-line field technique (Sortais and Kobayashi, 1968; and Sowerby et al., 1968). However, this also provides an approximate solution only.

An extensively used tool to model metal-forming processes is the finite element method (FEM). Zienkiewicz et al. (1974, 1981) predicted the deformation field for extrusion by FEM, characterizing the material deformation as the flow of an incompressible viscoplastic material. A similar analysis was presented by Tayal and Natarajan (1981) for axisymmetric extrusion through conical dies. Altan and co-workers (1982) modeled the extrusion process using the rigid-viscoplastic finite element method. Although FEM provides an accurate description of the stresses in the deformation zone, no methodology appears to have been developed for utilizing the results so obtained for optimal die design.

In the present work, an attempt has been made to develop a generalized methodology for optimal die design using the viscoplastic formulation coupled with the finite element technique. The proposed method considers the plastic zone boundaries and the die geometry as additional variables to be determined apart from the deformation field. The shapes of the die profile and the plastic boundaries are expressed in terms of chosen trial functions along with some undetermined coefficients. The total power is then minimized with respect to all the undetermined coefficients and the deformation field, thus predicting the optimal die shape within the trial function space. This solution methodology has been illustrated for the optimal design of a streamlined extrusion die.

Viscoplastic Finite Element Formulation

In metal-forming operations such as extrusion, large progressive plastic deformation occurs. The elastic strains can therefore be neglected for the sake of simplicity in the analysis. Under such conditions, the deformation of a viscoplastic solid is analogous to the flow of an incompressible, non-Newtonian fluid (Zienkiewicz and Godbole, 1974). The constitutive law linking the stresses σ_{ij} and the current deformation rates $\dot{\epsilon}_{ij}$, can be expressed in the Eulerian frame as

$$\sigma_{ij}=\sigma'_{ij}+\frac{1}{3}\sigma_{kk}\delta_{ij}, \tag{1}$$

where σ_{kk} is the hydrostatic stress and δ is the kronecker delta.

The deviatoric stresses σ'_{ij} are given by the viscous flow relation

$$\sigma'_{ij}=2\mu\dot{\epsilon}_{ij}. \tag{2}$$

For a viscoplastic material (Zienkiewicz and Godbole, 1974; Altan et al., 1982; and Oh, 1982), the viscosity μ can be expressed as

$$\mu=\frac{\bar{\sigma}}{3\dot{\bar{\epsilon}}} \tag{3}$$

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, the Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Sept. 28, 1989; final revision, Mar. 14, 1990.

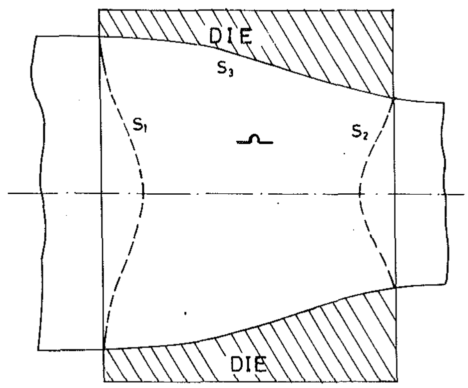


Fig. 1 Deforming zone in extrusion

where the generalized yield stress $\bar{\sigma}$ for a von Mises material is given by

$$\bar{\sigma}=\sqrt{\frac{3}{2}\sigma'_{ij}\sigma'_{ij}}=f(\bar{\epsilon},\ \dot{\bar{\epsilon}},\ T), \tag{4}$$

and the generalized strain rate $\dot{\bar{\epsilon}}$ is defined as

$$\dot{\bar{\epsilon}}=\sqrt{\frac{2}{3}\dot{\epsilon}_{ij}\dot{\epsilon}_{ij}}. \tag{5}$$

The form of function f in equation (4) will depend upon the deformation characteristics exhibited by the material such as ideal plastic behavior, strain rate sensitivity, work-hardening, and temperature dependence of properties.

Neglecting body and inertia forces, the stress equilibrium equation in the deformation zone Ω (Fig. 1) can be written using cartesian tensor notation, as

$$\sigma_{ij,j}=0. \tag{6}$$

From compatibility condition, the strain rate and velocity fields are related by

$$\dot{\epsilon}_{ij}=\frac{1}{2}(V_{i,j}+V_{j,i}). \tag{7}$$

The velocity field is required to satisfy the incompressibility condition

$$V_{i,i}=0. \tag{8}$$

Denoting the plastic boundaries and the die profile by surfaces S_1, S_2, and S_3 (Fig. 1) whose shapes are yet to be determined, the typical boundary conditions for the problem can be written as follows:

(i) On the die surface S_3, the normal component of velocity vanishes and the frictional shear stress is prescribed. These conditions can be written as

$$V_i n_i=0 \tag{9a}$$

and

$$n_i\sigma_{ij}t_j=\tau_s \text{ (the prescribed frictional shear stress)}, \tag{9b}$$

where **n** and **t** are the local unit vectors in the normal and tangential directions at the die surface.

(ii) On the plastic boundaries S_1 and S_2, either the components of the velocity vector or the boundary traction may be specified. Thus,

$$\sigma_{ij}n_i=F_j \tag{10a}$$

and

$$V_i=U_i, \tag{10b}$$

where F_j and U_i are the prescribed tractions and velocity components and **n** is the local unit normal vector on the plastic boundaries.

Equations (1 to 10) can be recast in an equivalent variational form for convenient treatment in the finite element solution approach. It can be shown that the functional ϕ corresponding to these equations is given by

$$\phi=\int_{\Omega^*}E(\dot{\epsilon}^*)dv+\int_{\Omega^*}\frac{1}{2}\lambda(\dot{\epsilon}^*_{kk})^2dv-\int_{S_3^*}\tau_s\Delta V_s ds$$
$$+\int_{S_1^*}\frac{\bar{\sigma}}{\sqrt{3}}|\Delta V_t|_{S_1}ds+\int_{S_2^*}\frac{\bar{\sigma}}{\sqrt{3}}|\Delta V_t|_{S_2}ds-\int_{S_2^*}F_jV_j^*ds, \tag{11}$$

where the work function E is of the form

$$E(\dot{\epsilon})=\int_0^{\dot{\epsilon}}\sigma_{ij}d\dot{\epsilon}_{ij}.$$

The sign of the frictional shear stress τ_s is opposite to that of the slip velocity ΔV_s on the die surface S_3. The tangential velocity discontinuities at S_1 and S_2 are denoted by $|\Delta V_t|_{S_1}$ and $|\Delta V_t|_{S_2}$, respectively. F_j is the prescribed tension at the exit end of the billet.

The incompressibility constraint on deformation field has been satisfied by penalizing it with a large positive penalty parameter λ in equation (11). The functional ϕ defined above represents the total extrusion power which is the sum of various power contributions. The terms in equation (11) in their order of occurrence represent the following:

(i) and (ii) the internal power required for deformation in Ω.
(iii) the power loss due to friction on S_3.
(iv) the shear losses due to tangential velocity discontinuities at S_1 and S_2.
(v) the power contribution when front tension is applied on S_2.

In equation (11), the asterisk indicates that the concerned variables are restricted in their respective trial spaces. The restrictions placed on the deformation field are that it must satisfy the prescribed velocity conditions on S_1, S_2, and S_3 and in addition, it must satisfy incompressibility everywhere in Ω. The die profile S_3 is constrained by inlet and exit cross-sections. Also, the slopes or curvatures of the profile may be prescribed at the end sections. In specific situations, the choice of S_3 may further be limited to a selected function space. For instance, in an axisymmetric extrusion problem the trial space for S_3 can be chosen as the set of all surfaces of revolution generated by third-order polynomials. The restrictions which apply for the shapes of the plastic boundaries S_1 and S_2 are that the normal velocity should be continuous across these surfaces while tangential velocity discontinuities may be permissible. As in the case of S_3, the surfaces S_1 and S_2 may also belong to some chosen function spaces. The shape of the deformation zone Ω is thus restricted by the individual restrictions placed on S_1, S_2, and S_3.

In the upper bound approach the velocity field, which satisfies only incompressibility, is substituted in the expression for ϕ. The shapes of S_1, S_2, and S_3 are then determined by minimizing ϕ (Chen and Ling, 1968; and Yang et al., 1985). Since the velocity fields do not satisfy the stress balance at all locations in Ω, the predicted optimal shapes are not accurate. On the other hand, viscoplastic models using FEM (Zienkiewicz and Godbole, 1974; and Tayal and Natarajan, 1981) have so far attempted to solve for the velocity field only for given shapes of S_1, S_2, and S_3. In most of the FEM analyses, the deforming zone is defined as the region bounded by the die profile and straight plastic boundaries at the inlet and exit sections. Discretizing the deformation zone into many elements, each having a prescribed number of nodes, the nodal velocities are calculated from the variational principle for ϕ.

In the present formulation the nodal velocities, as well as the solution domain Ω and its boundaries S_1, S_2, and S_3, are considered to be variables which determine the value of ϕ.

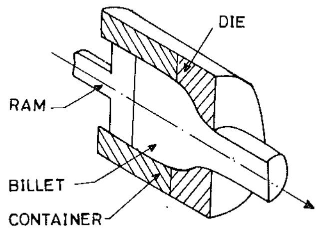


Fig. 2 A sectional view of an extruded cylindrical bar

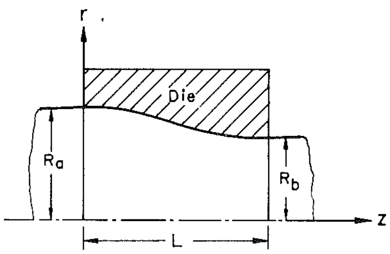


Fig. 3 Geometry of a streamlined die

With an appropriate choice of trial function spaces for S_1, S_2, and S_3, the shapes of these surfaces are obtained by minimizing ϕ. If V_i are the velocity unknowns and α_i, β_i, and γ_i are the undetermined coefficients which define the shapes of S_1, S_2, and S_3 in their respective trial spaces, the functional ϕ is minimized by setting

$$\frac{\partial \phi}{\partial V_i}=0 \text{ for } i=1, 2, \ldots\ldots N, \tag{12a}$$

$$\frac{\partial \phi}{\partial \alpha_i}=0 \text{ for } i=1, 2, \ldots\ldots M_1, \tag{12b}$$

$$\frac{\partial \phi}{\partial \beta_i}=0 \text{ for } i=1, 2, \ldots\ldots M_2, \tag{12c}$$

$$\frac{\partial \phi}{\partial \gamma_i}=0 \text{ for } i=1, 2, \ldots\ldots M_3, \tag{12d}$$

where N is the total number of velocity vector unknowns and M_1, M_2, M_3 are, respectively, the number of undetermined coefficients in the definitions of S_1, S_2, and S_3. Thus, the analysis simultaneously performs domain optimization for the plastic region along with the calculation of the deformation field solution.

Streamlined Die Design

In the present section, the method proposed above is illustrated for the optimal design of a derivative of sigmoidal die known as the streamlined die (Devenpeck and Richmond, 1965). The streamlined die possesses advantages such as uniform and homogeneous deformation characteristics, absence of an intense shear band, and low energy consumption (Sortais and Kobayashi, 1968; and Yang et al., 1985). Several researchers (Chang and Choi, 1971; Gunasekara and Hoshino, 1985; Yang et al., 1985) have obtained approximate upper-bound solutions for extrusion through streamlined dies. However, an accurate FEM-based determination of the optimal streamlined die shape as presented here has not so far been attempted.

The axisymmetric extrusion of an ideal plastic material for reducing the diameter of a cylindrical bar (Fig. 2) is considered. The extrusion process is assumed to take place at low ram speeds so that the process is free of rate effects. A constant shear friction condition is considered at the die-billet interface.

Since a streamlined die does not produce shear bands at the inlet and exit sections, it is reasonable to relax the constraints placed by the plastic boundaries on metal flow. Therefore, a deformation zone bounded by straight rigid plastic boundaries at die entry and exit can be assumed for analytical convenience. The upper-bound solutions obtained by earlier researchers (Chang and Choi, 1971; and Yang et al., 1985) using straight and arbitrarily shaped plastic boundaries indicate that there is little effect of the shapes of S_1 and S_2 on the overall solution. An additional advantage of using straight plastic boundaries is that this lays more emphasis on the optimization of the die shape.

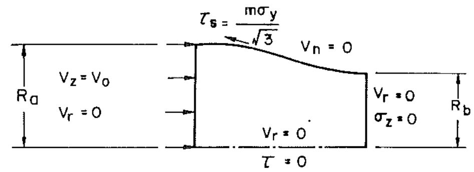


Fig. 4 Boundary conditions for extrusion

For a streamlined die profile, the slopes at the inlet and exit sections are prescribed as zero, in addition to the size reduction requirements. These geometrical constraints can be met by a third-order polynomial of the form

$$R(z)=a_0+a_1z+a_2z^2+a_3z^3. \tag{13}$$

Equation (13) defines the radius of a cross-section in terms of the axial distance, on surface S_3. Applying the conditions (Fig. 3)

$$\begin{pmatrix} R=R_a \text{ at } z=0 \\ R=R_b \text{ at } z=L \\ \dfrac{dR}{dz}=0 \text{ at } z=0 \\ \dfrac{dR}{dz}=0 \text{ at } z=L \end{pmatrix} \tag{14a-d}$$

it can be shown that the final form of equation (13) simplifies as

$$R(z)=\frac{R_a-R_b}{L^3}(2z^3-3z^2L)+R_a, \tag{15a}$$

where

$$R_b=R_a\sqrt{1-A_r}, \tag{15b}$$

with A_r denoting the area reduction ratio.

The streamlined die profile defined by equation (15a) represents a one-parameter family of surfaces for different values of the projected die length L. In the trial space of third-order polynomial die shapes, the optimum profile is found by minimizing ϕ with respect to L, by setting

$$\frac{d\phi}{dL}=0 \tag{16a}$$

and

$$\frac{d^2\phi}{dL^2}>0. \tag{16b}$$

For calculating the velocity field, the following boundary conditions (Fig. 4) are applied:

(a) At inlet the undeformed billet material approaches the die with a uniform velocity equal to that of the ram. Thus,

$$V_z=V_o \tag{17a}$$

and

$$V_r=0 \tag{17b}$$

at $z=0$.

The boundary condition (17b) is a consequence of the zero slope of streamlined die at the entry section.

(b) On the die surface S_3, the conditions are:

(i) The normal component of velocity vanishes at the die surface due to impenetrability. This leads to

$$V_i n_i = 0 \text{ at } r = R(z). \tag{18a}$$

(ii) A constant shear stress is assumed as the frictional condition at the die-billet interface (Zeev Zimerman and Avitzur, 1970). Thus,

$$\tau_s = \frac{m\sigma_y}{\sqrt{3}} \text{ at } r = R(z), \tag{18b}$$

where m is the friction factor and σ_y the yield stress of the material.

(c) At the exit section of the die:

(i) $V_r = 0$ (due to zero slope) at $z = L$. (19a)

(ii) In the absence of front tension, the exit condition for extrusion can be modeled as

$$\sigma_z = 0 \text{ at } z = L. \tag{19b}$$

(d) Along the axis, there is no motion normal to the axis and shear stress is zero due to axisymmetry. These conditions can be expressed mathematically as

$$V_r = 0 \tag{20a}$$

and

$$\tau = 0 \tag{20b}$$

at $r = 0$.

Since the entry and exit of the billet at the end sections are smooth, the shear losses can be taken to be zero. The simplified expression for ϕ for a streamlined die is therefore of the form

$$\phi = \int_{\Omega^*} E(\dot{\epsilon}^*)^* 2\pi r dr dz + \int_{\Omega^*} \frac{1}{2}\lambda(\dot{\epsilon}^*_{kk})^2 2\pi r dr dz$$

$$+ \int_{S_3^*} m \frac{\sigma_y}{\sqrt{3}} |\Delta V_s| 2\pi R(z) \sqrt{1 + \left(\frac{dR(z)}{dz}\right)^2} dz. \tag{21}$$

Expressing the total extrusion power ϕ as a function of velocity field and die length L, ϕ can be minimized with respect to these variables. The average extrusion pressure P can then be determined as

$$P = \frac{\phi}{\pi R_a^2 V_o}. \tag{22}$$

FEM Solution Procedure

The application of FEM to plastic deformation in metal forming processes has been described earlier by Zienkiewicz and co-workers (1974) and Tayal and Natarajan (1981). A similar procedure has been adopted in the present work for the velocity field solution with modifications incorporated into numerical scheme for evaluating the optimum die geometry. Eight-noded isoparametric elements have been employed whose boundaries, in general, are quadratic curves (Fig. 5). This gives quite an accurate representation of the curved die boundary. Starting with an initial guess for the die profile by choosing a value for L, an automatic mesh generation scheme is used to generate the FEM mesh of quadratic elements in the deformation zone. The value of L is updated iteratively by applying the minimum ϕ criterion in equation (21).

The axial and radial velocities in each element are expressed in terms of their nodal values using the interpolation expressions:

$$V_z = \sum_{m=1}^{8} N_m V_{zm} \tag{23}$$

and

$$V_r = \sum_{m=1}^{8} N_m V_{rm}, \tag{24}$$

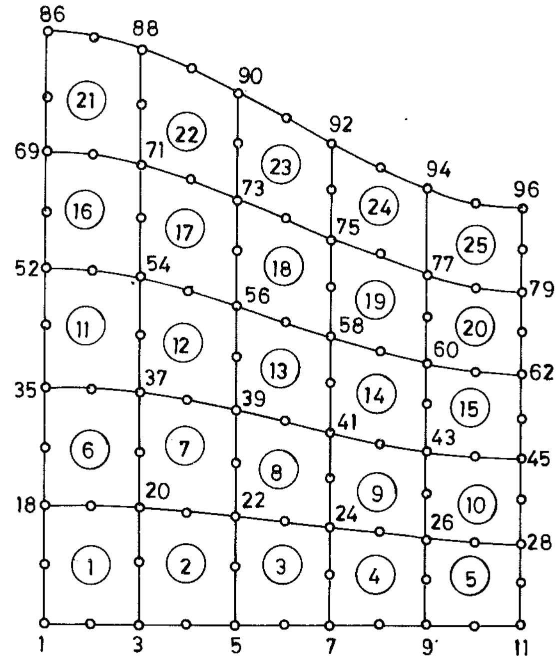


Fig. 5 A 25-element mesh with 8-noded isoparametric elements

where N_m are the quadratic shape functions for the eight-noded elements (Reddy, 1985) and V_{zm} and V_{rm} are the nodal velocity values. After substituting the compatibility relations for the strain rate components in equation (21), the minimization of ϕ with respect to the nodal velocity unknowns yields,

$$\int_{\Omega} \mu N_{m,k}(V_{k,i} + V_{i,k}) 2\pi r dr dz$$

$$+ \int_{\Omega} \lambda N_{m,i}(V_{k,k}) 2\pi r dr dz = \int_{S_3} N_m T_i 2\pi r ds, \tag{25}$$

where i and k are the directional subscripts, m is the subscript for the nodal number and T is the prescribed boundary traction and s is the arc length measured along S_3.

The value of the penalty parameter λ is obtained by trial and error, after varying it over a wide range of values until the velocity field solution becomes invariant with λ. An approximate value of λ can be found by taking $\lambda(\nabla.\mathbf{V}) \approx \sigma_y$ and estimating the smallest value of $(\nabla.\mathbf{V})$ from the word length of the computer used. In the present work, λ of the order of 10^6 was found to be suitable. Equation (25) can be written in the matrix form (after incorporating the boundary conditions) as

$$[K(\mu)] \begin{Bmatrix} V_{zm} \\ V_{rm} \end{Bmatrix} = \{F\}, \tag{26}$$

where K is the stiffness matrix which depends on the nonlinear viscosity function, $\begin{Bmatrix} V_{zm} \\ V_{rm} \end{Bmatrix}$ is the vector of the nodal velocity components and $\{F\}$ is the vector consisting of known velocity values or the contributions from known surface tractions. At nodes where the normal velocity is zero, the matrix K has been modified by substituting equation (18a) in place of the nodal equation for the radial direction.

The frontal solution technique has been used to solve the matrix equation (26) in order to reduce the core memory requirements. The stiffness matrix K is highly nonlinear, and hence an iterative solution procedure has been employed. For the first iteration, an initial guess for velocity field from the upper-bound solution (Aravamadhu Balaji, Sundararajan, and Lal, 1989) has been provided for evaluating μ. During the subsequent iterations, the value of μ is updated by using the

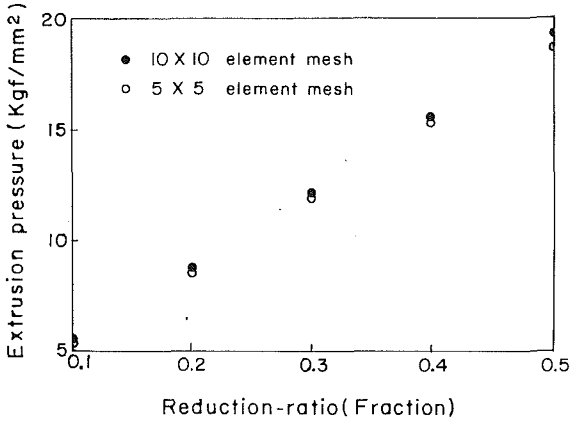


Fig. 6 Comparison of FEM results for two meshes

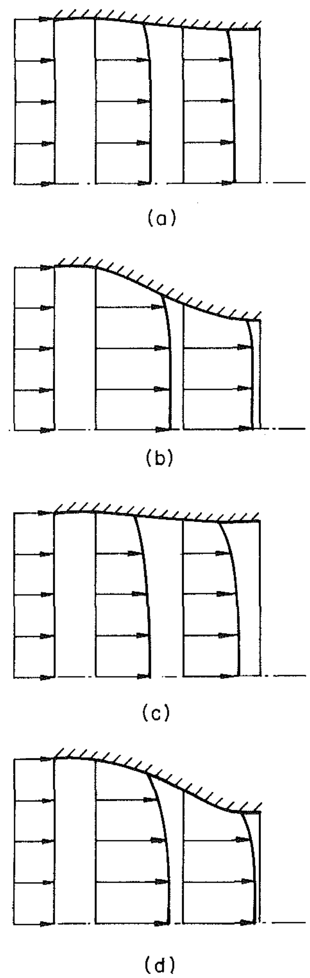


Fig. 7 Axial velocity distribution (*a*) $m=0.1, A_r=0.1$, (*b*) $m=0.1, A_r=0.5$, (*c*) $m=0.5, A_r=0.1$, and (*d*) $m=0.5, A_r=0.5$

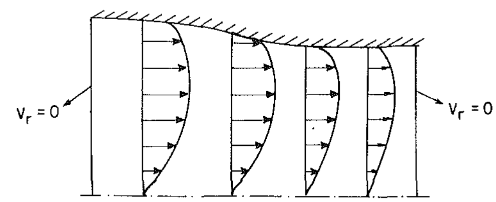


Fig. 8 Radial velocity distribution in the deforming zone

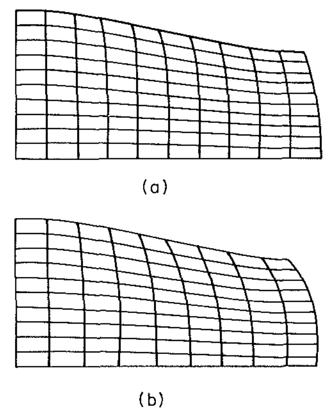


Fig. 9 Deformation patterns (*a*) $m=0.1, A_r=0.5$ and (*b*) $m=0.5, A_r=0.5$

previous iteration values of the velocity solution. In the course of numerical computations, the value of μ tends to become very large near the end sections of the die as $\bar{\dot{\epsilon}} \rightarrow 0$ at these locations. This difficulty is overcome by prescribing a suitable cutoff for $\bar{\dot{\epsilon}}$. To ensure stability, the velocity values are underrelaxed between iterations. The elemental integrals have been evaluated using 3×3 Gaussian quadrature, except for the incompressibility constraint for which a 2×2 Gaussian quadrature has been used. The frictional boundary integral at the die surface is computed by 3-point Gaussian quadrature considering the element surface as a one-dimensional isoparametric element (Oh, 1982). The iterations have been continued until the nodal velocity values converged within 0.1 percent between two successive iterations.

After converging the velocity unknowns for a given die shape, the optimal value of L has been obtained by the Newton-Raphson iterative procedure. It is possible to extend the present scheme to problems where the deformation zone Ω is bounded by multiparameter family of surfaces. Assuming initial guess shapes for S_1, S_2, and S_3 the velocity field solution can be obtained using a conventional FEM procedure and then the shape coefficients can be optimized using the Newton-Raphson procedure.

Results and Discussions

The die design procedure described above has been implemented for the evaluation of the optimum geometry of an axisymmetric streamlined die with a third-order polynomial profile. The billet material chosen for the present study is a soft low carbon steel with a typical composition of 0.09 percent C, 0.4 percent Mn, 0.015 percent P and 0.038 percent S whose yield stress value is equal to 21.1 kgf/mm^2 (Miner and Seastone, 1958). The material has been assumed to exhibit ideal plastic behavior obeying von Mises yield criterion. The following process parameters have been considered to be fixed for all the calculations made in the present study:

billet diameter – 25 mm
velocity of the ram – 0.1 mm/minute.

In order to estimate the accuracy of the predicted FEM results, the deformation field has been obtained using two different FEM meshes for a fixed value of the die length and die friction factor. A coarse 5×5 element mesh and a fine 10×10 element mesh have been employed and the sensitivity of the FEM results to the change in the element size has been shown in Fig. 6. The results of both the meshes compare very closely, implying that no further mesh refinement is necessary. Hence, for all the results presented here, calculations have been performed using a 10×10 element mesh.

In Fig. 7 (*a* to *d*), the axial velocity profiles in the deformation zone have been shown for different friction and reduction parameters. At every cross-section the axial velocity decreases in the radial direction due to the frictional retardation at the die surface. In the axial direction, however, die compression leads to an increase in the velocity. The velocity profile tends to be more uniform at low friction factor because of less retardation at the die surface. The nonuniformity in the velocity profile increases with reduction ratio. Higher reduction causes the radial velocity to increase which in turn leads to the lowering of the axial velocity on account of material incompressibility. Such a nonuniform velocity profile is expected to produce greater grid distortion.

The radial velocity profile variation in the axial direction (Fig. 8) corroborates the inference drawn in the discussion of Fig. 7 that the radial velocity increases with die compression. Indeed the maximum values are seen to occur in the location of large die curvature. A slight decrease in the value of radial velocity occurs at the die surface as compared to the interior, due to frictional retardation.

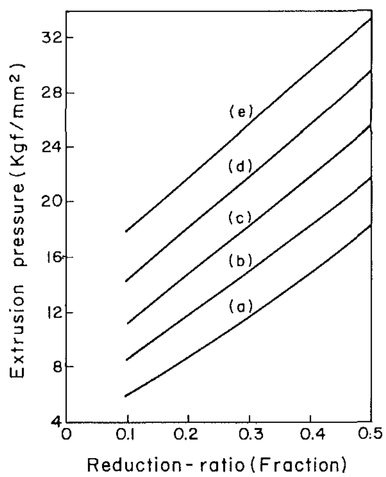


Fig. 10 Variation of extrusion pressure with reduction ratio (*a*) $m=0.1$, (*b*) $m=0.2$, (*c*) $m=0.3$, (*d*) $m=0.4$, and (*e*) $m=0.5$

Grid-deformation patterns for two different friction conditions are presented in Fig. 9. Distortion of the grid is larger at higher friction factor which is in agreement with the experimental results of Sortais and Kobayashi (1968). However, grid distortions predicted in the present study are not so severe as those obtained earlier (Sortais and Kobayashi, 1968; and Yang et al., 1985) because the billet material considered here is comparatively a very soft and nonwork-hardening material.

The extrusion pressure increases with reduction and friction factors, as presented in Fig. 10. This trend can be explained by equations (21) and (22) as a consequence of the increase in internal power of deformation and frictional dissipation at the die surface.

The plots of extrusion pressure variation with die length shown in Fig. 11 (*a* to *d*) indicate that a minimum value of extrusion pressure is reached at some die length. This value may be defined as the optimal die length. The optimality arises as a consequence of the opposing trends that the internal power of deformation decreases with the die length while the frictional power increases. The optimal die length reduces at higher friction factors in order to offset the tendency for increase in the frictional power with m. For a similar reason, the optimal length increases with reduction ratio by offsetting the increase in internal power of deformation. The value of the extrusion power, however, increases with both friction and reduction parameters. These trends are in agreement with those observed by Gunasekara and Hoshino (1985) and Yang, Han, and Lee (1985). However, the predicted optimal length values of the present study are smaller since a very soft nonwork-hardening billet material has been considered.

The shear stress σ_{rz}, the axial stress σ_{zz}, and the die pressure (normal stress exerted by the die) at the die-billet interface have been plotted against axial length in Fig. 12 (*a*, *b*, and *c*) for the optimal die geometry corresponding to $m=0.5$ and $A_r=0.1$. The shear stress is negative and increases in magnitude with z near the entry section. This implies that the retardation effect increases progressively and the die compression has not yet

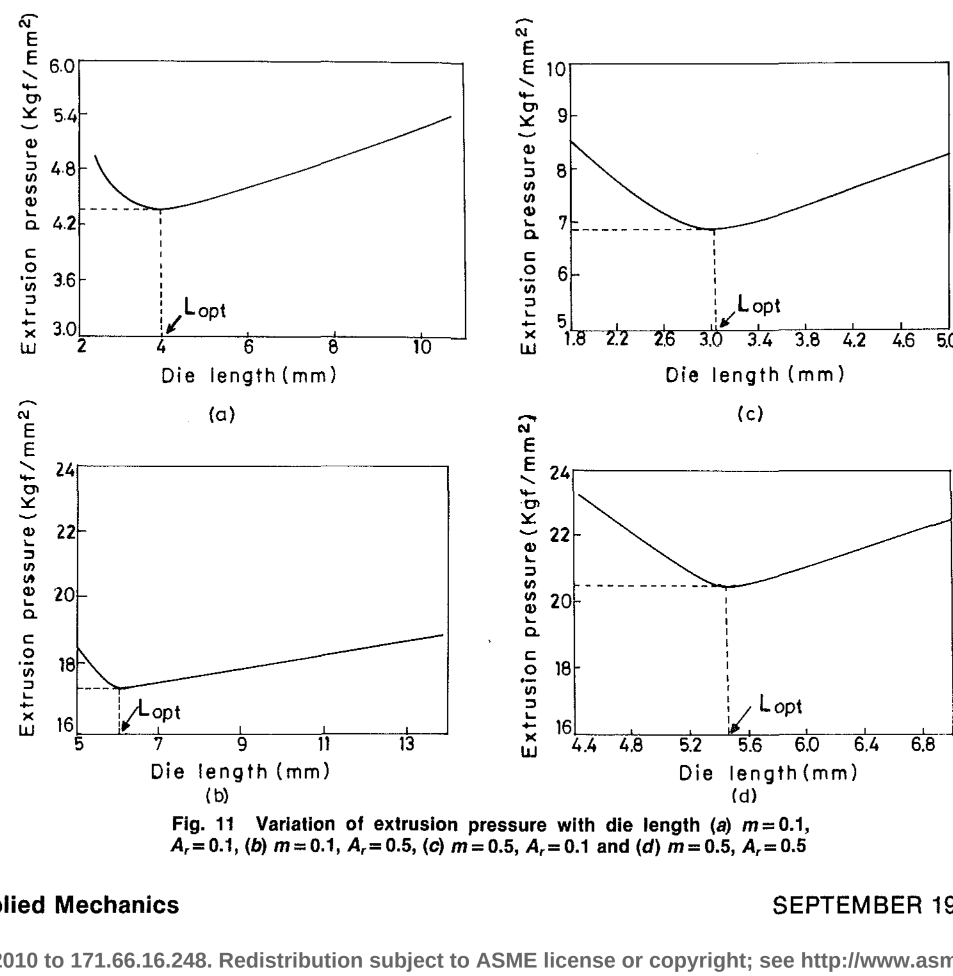


Fig. 11 Variation of extrusion pressure with die length (*a*) $m=0.1$, $A_r=0.1$, (*b*) $m=0.1$, $A_r=0.5$, (*c*) $m=0.5$, $A_r=0.1$ and (*d*) $m=0.5$, $A_r=0.5$

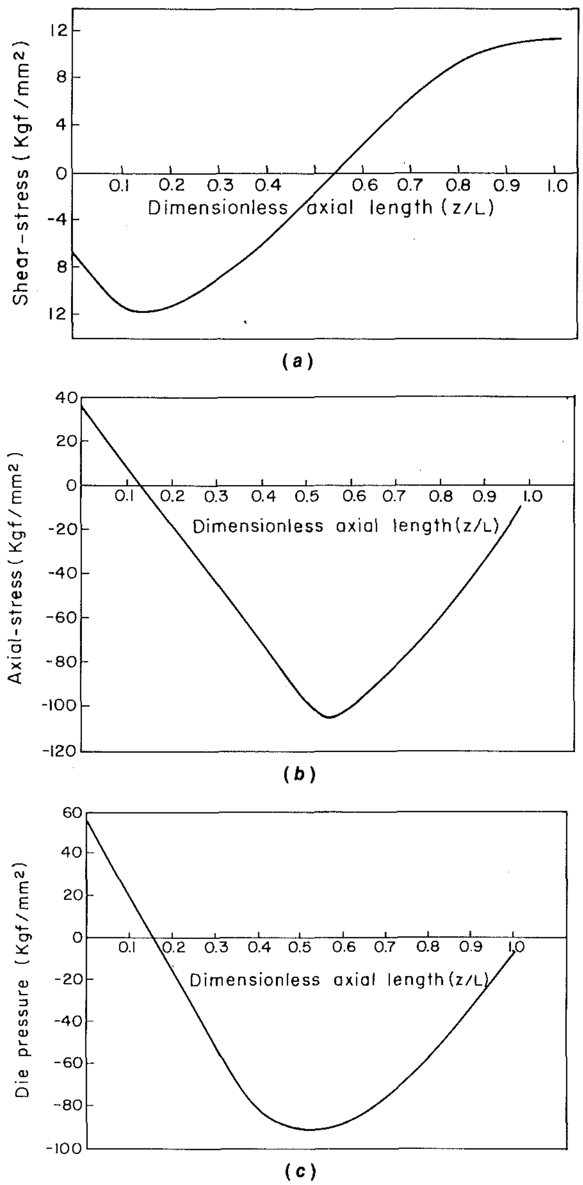


Fig. 12 Stress distribution on the die surface (*a*) shear stress, (*b*) axial stress, and (*c*) die pressure

begun. However, beyond a certain value of z, the compression effects come into picture and increase the value of the billet slip velocity at the die surface. This, in turn, decreases the magnitude of the shear stress. Around the mid-section where the die curvature is high, the shear stress becomes positive due to further increase in slip velocity. Near the die-exit as the die profile becomes flat and the compression effects vanish, the shear stress variation with z becomes negligible. Similar trends were observed by Lee Mallet and Yang (1977).

The axial and die pressure variations indicate a zone of tensile stress near the die-entry. Lee et al. (1977) also observed a similar zone in their study and suggested that this is because of the separation effects occurring at the die-billet interface. The billet contact with the die is reestablished at the zone where the tensile tractions are reduced to zero. The axial stress and die pressure clearly manifest the trend of increasing compression up to the length where die curvature is maximum and then a decrease in compression towards the die exit. The change in the sign of shear stress from negative to positive values appears to occur around the location of maximum compressive stresses.

In order to critically examine the accuracy of the FEM results the quadratic stress invariant $\bar{\sigma}$ was computed at all the nodes in the deformation zone, and was found to be equal to the material yield stress to a high degree of precision.

Conclusions

The viscoplastic deformation analysis by FEM has been successfully applied to predict the deformation field, die geometry, and plastic boundaries during metal extrusion. The predicted values and trends are in reasonable agreement with the experimental observations and are in conformity with the results published earlier.

The methodology proposed can be extended to analyze other metal-working operations where the domain of plastic deformation is to be predicted.

References

Altan, T., Gunasekara, J. S., Gegel, H. L., Malas, J. C., and Morgan, J. T., 1982, "Computer Aided Process Modelling of Hot Forging and Extrusion of Aluminum Alloys," *Annals of CIRP*, Vol. 31, pp. 131–135.

Aravamadhu Balaji, P., Sundararajan, T., and Lal, G. K., 1989, "Optimal Design of Curved Dies for Extrusion of Strain Hardening Materials," submitted for publication.

Chang, K. T., and Choi, J. C., 1971, "Upper Bound Solutions to Extrusion Problems through Curved Dies," *Proceedings of the 12th Midwestern Mechanics Conference*, University of Norte Dame, Ind., pp. 383–396.

Chen, C. T., and Ling, F. E., 1968, "Upper Bound Solutions to Axi-symmetric Extrusion Problems," *International Journal of Mechanical Sciences*, Vol. 10, pp. 863–879.

Devenpeck, M. L. and Richmond, O., 1965, "Strip-Drawing Experiments with a Sigmoidal Die Profile," *ASME Journal of Engineering for Industry*, Vol. 87, pp. 425–428.

Miner, D. I., and Seastone, J. B., 1958, *Handbook of Engineering Materials*, John Wiley and Sons, New York, pp. 3–146.

Gunasekara, J. S., and Hoshino, S., 1985, "Analysis of Extrusion of Polygonal Sections through Streamlined Dies," *ASME Journal of Engineering for Industry*, Vol. 107, pp. 229–233.

Lee, E. H., Mallet, R. L., and Yang, W. H., 1977, "Stress and Deformation Analysis of the Metal Extrusion Process," *Computer Methods in Applied Mechanics and Engineering*, pp. 339–353.

Oh, S. I., 1982, "Finite Element Analysis of Metal Forming Processes with Arbitrarily Shaped Dies," *International Journal of Mechanical Sciences*, Vol. 24, pp. 479–493.

Reddy, J. N., 1985, *An Introduction to the Finite Element Method*, McGraw-Hill, New York, pp. 242–254.

Sortais, H. C., and Kobayashi, S., 1968, "An Optimum Die Profile for Axisymmetric Extrusion," *International Journal of Machine Tool Design and Research*, Vol. 8, pp. 61–72.

Sowerby, R., Johnson, W., and Samantha, S. K., 1968, "Plane Strain Drawing and Extrusion of a Rigid-Perfectly Plastic Material through Concave Dies," *International Journal of Mechanical Sciences*, Vol. 10, pp. 231–238.

Tayal, A. K., and Natarajan, R., 1981, "Extrusion of Rate-Sensitive Materials using a Viscoplastic Constitutive Equation and the Finite Element Method," *International Journal of Mechanical Sciences*, Vol. 23, pp. 89–98.

Yang, D. Y., Han, C. H., and Lee, B. C., 1985, "The Use of Generalized Deformation Boundaries for the Analysis of Axisymmetric Extrusion through Curved Dies," *International Journal of Mechanical Sciences*, Vol. 27, pp. 653–663.

Zimerman, Z., and Avitzur, B., 1970, "Metal Flow through Conical Converging Dies—A Lower Upper Bound Approach using Generalized Boundaries of the Plastic Zone," ASME *Journal of Engineering for Industry*, Vol. 92, pp. 119–129.

Zienkiewicz, O. C., and Godbole, P. N., 1974, "Flow of Plastic and Visco Plastic Solids with Special Reference to Extrusion and Forming Processes," *International Journal for Numerical Methods in Engineering*, Vol. 8, pp. 3–16.

Zienkiewicz, O. C., Onate, E., and Heinrich, I. C., 1981, "A General Formulation for Coupled Thermal Flow of Metals using Finite Elements," *International Journal for Numerical Methods in Engineering*, Vol. 17, pp. 1497–1514.

Analytical Solutions for the Plane-Strain Bifurcation of Compressible Solids

J. P. Bardet
Civil Engineering Department,
University of Southern California,
Los Angeles, CA 90089-0242

Analytical solutions are presented for the diffuse and localized bifurcations of compressible solids subjected to plane-strain loadings. The solutions generalize the works for incompressible solids of Biot (1965) and Hill and Hutchinson (1975). They are verified by comparing them to results previously established for incompressible solids and elastoplastic Mohr-Coulomb materials.

1 Introduction

Biot (1965) and Hill and Hutchinson (1975) analyzed the bifurcation of incompressible solids in plane-strain tension. Young (1976) applied their method to plane-strain compression and Needleman (1979) adapted it for elastoplastic solids with nonassociative flow rule. Vardoulakis (1981) introduced the effect of compressibility by treating it as an internal constraint for a particular elastoplastic Mohr-Coulomb materials. Chau and Rudnicki (1989) generalized the analysis of Needleman and Vardoulakis by considering a class of compressible solids with five parameters. Bazant (1971) reviewed several incremental formulations for stability and compared their results in the case of buckling of free surfaces and columns.

The bifurcation analyses of solid mechanics have migrated to the field of finite element methods in order to be generalized to realistic but complicated boundary value problems (Bardet, 1990; deBorst, 1987; Needleman and Tvergaard, 1984). However, the rational and methodic development of numerical techniques for bifurcation is still impeded by the scarcity of the analytical solutions that are available to calibrate numerical results. The analysis of Chau and Rudnicki (1989) is restrained to a class of compressible solids, the incremental response of which is identified by five parameters.

This paper presents general analytical solutions that are useful to assess the role of material compressibility on plane-strain bifurcations and to calibrate the finite element analysis of bifurcations within compressible solids. A similar approach was adopted by Bardet (1990) to support numerical results on surface instability.

Following the Introduction, the second section reviews the concepts of continuum mechanics relevant to material instability, which include stress rates and constitutive models. The third section poses and analytically solves the plane-strain problem of diffuse bifurcations and relates it to localized bifurcations. The fourth section attempts to validate the analytical solutions by comparing them with existing results on incompressible solids and elastoplastic Mohr-Coulomb materials.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Sept. 21, 1989; final revision, Apr. 9, 1990.

2 Definition of Stress Rates and Constitutive Models

2.1 Stress Tensors and Rates. By definition, the contact force vector $\mathbf{t}$ acting on the deformed surface, with area dS_t and unit normal vector $\mathbf{n}$, is related to the Cauchy stress tensor σ and the nominal (Piola-Kirchhoff) stress tensor Σ through:

$$\mathbf{t}=\mathbf{n}\cdot\sigma dS_t=\mathbf{N}\cdot\Sigma dS_o \tag{1}$$

where $\mathbf{N}$ and dS_o are the unit normal vector and area, respectively, of the undeformed surface. Nominal and Cauchy stresses are related through:

$$\Sigma=\det(\mathbf{F})\mathbf{F}^{-1}\cdot\sigma \tag{2}$$

where $\mathbf{F}^{-1}$ is the inverse transformation of the deformation gradient $\mathbf{F}$. By definition, the Kirchhoff stress tensor τ is related to the Cauchy stress tensor σ:

$$\tau=\det(\mathbf{F})\sigma. \tag{3}$$

Hereafter, the derivative with respect to time are designated by the superscript "°," and the partial differentiation with respect to the coordinate x_j, are denoted by ",j." The rate of $\mathbf{t}$ is:

$$\mathring{\mathbf{t}}=\mathbf{N}\cdot\mathring{\Sigma}dS_o \tag{4}$$

and the rate of Σ is:

$$\mathring{\Sigma}=\det(\mathbf{F})\mathbf{F}^{-1}\cdot(\mathring{\sigma}-\mathbf{L}\cdot\sigma+\sigma\ \mathrm{trace}(\mathbf{L})) \tag{5}$$

where $\mathbf{L}$ is the gradient of the velocity $\mathbf{v}$ with respect to the deformed position $\mathbf{x}$:

$$L_{ij}=\frac{\partial v_i}{\partial x_j}=v_{i,j} \tag{6}$$

2.2 Rate-Type Constitutive Models. In the present analysis, the material behavior is modeled with rate-type equations (Truesdell and Noll, 1965)

$$\hat{\tau}_{ij} = C_{ijkl} D_{kl} \tag{7}$$

where $\hat{\tau}$ is the Jaumann rate of Kirchhoff stress τ, and $\mathbf{D}$ the rate of deformation. In general, C_{ijkl} is homogeneous of degree zero in $\mathbf{D}$ and depends on the states of stress and strain. The Jaumann rate of Kirchhoff stress is:

$$\hat{\tau} = \overset{\circ}{\tau} - \mathbf{W} \cdot \tau + \tau \cdot \mathbf{W}. \tag{8}$$

The rate of deformation $\mathbf{D}$ and spin tensor $\mathbf{W}$ are

$$D_{ij} = \frac{1}{2}(L_{ij} + L_{ji}) \quad W_{ij} = \frac{1}{2}(L_{ij} - L_{ji}). \tag{9}$$

2.3 Relations Between Material and Objective Stress Rates. The Jaumann rate of Cauchy stress $\hat{\sigma}$ is defined as $\hat{\tau}$ by using Eq. (8). It is related to $\hat{\tau}$ through

$$\hat{\tau} = \det(\mathbf{F})(\hat{\sigma} + \sigma v_{k,k}). \tag{10}$$

If the present configuration at time t is chosen as reference and the configuration at a later time $t + dt$ as the deformed configuration, then the deformation gradient is approximately equal to the unity transformation $\mathbf{1}$:

$$\mathbf{F} = \mathbf{F}^{-1} \approx \mathbf{1} \quad \text{and} \quad \det(\mathbf{F}) \approx 1. \tag{11}$$

In this condition, the Piola-Kirchhoff, Cauchy, and Kirchhoff stress tensors are identical:

$$\Sigma = \sigma = \tau, \tag{12}$$

and their rates are related through

$$\hat{\tau} = \hat{\sigma} + \sigma D_{kk} \tag{13}$$

$$\overset{\circ}{\Sigma} = \hat{\tau} - \sigma \cdot \mathbf{W} - \mathbf{D} \cdot \sigma. \tag{14}$$

By using equation (14), equation (7) becomes:

$$\overset{\circ}{\Sigma}_{ij} = (C_{ijkl} + \tilde{C}_{ijkl}) v_{k,l}, \tag{15}$$

where the additional terms $\tilde{C}_{ijkl}$ are only stress dependent

$$\tilde{C}_{ijkl} = \frac{1}{2}(\sigma_{il}\delta_{jk} - \sigma_{jl}\delta_{ik} - \sigma_{jk}\delta_{il} - \sigma_{ik}\delta_{jl}). \tag{16}$$

δ_{ij} is the Kroneker symbol:

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \quad i, j = 1,2,3. \tag{17}$$

The choice of the Jaumann rate of Kirchhoff stress does not affect the generality of the present analysis. Equation (7) is simply related to the constitutive equations that are formulated in terms of other types of objective stress rates. For instance, when the constitutive equation is formulated in terms of the Jaumann rate of Cauchy stress

$$\hat{\sigma}_{ij} = C^{\sigma}_{ijkl} D_{kl}, \tag{18}$$

the constitutive matrix of equation (7) is:

$$C_{ijkl} = C^{\sigma}_{ijkl} + \sigma_{ij}\delta_{kl}. \tag{19}$$

3 Analytical Solutions for Plane-Strain Bifurcation

3.1 Formulation of Problem for Diffuse Bifurcations. In the case of plain-strain loadings (Hill, 1959), the stress-rate equilibrium equations are:

$$\begin{cases} \overset{\circ}{\Sigma}_{11,1} + \overset{\circ}{\Sigma}_{21,2} = 0 \\ \overset{\circ}{\Sigma}_{12,1} + \overset{\circ}{\Sigma}_{22,2} = 0. \end{cases} \tag{20}$$

Equation (20) is solved for the boundary value problem of Fig. 1. On the top and bottom edges where $x_2 = \pm H$, the velocity is prescribed in the x_2-direction without causing shear traction. A constant stress σ_{11} is applied on the lateral surfaces at $x_1 = \pm L$. The boundary conditions are:

$$\begin{cases} \overset{\circ}{\Sigma}_{21} = 0 \\ v_2 = 0 \end{cases}$$

$$\text{for } x_2 = \pm H \text{ and } x_1 \in [-L, +L] (N_1 = 0 \text{ and } N_2 = \pm 1) \tag{21}$$

$$\begin{cases} \overset{\circ}{\Sigma}_{11} = 0 \\ \overset{\circ}{\Sigma}_{12} = 0 \end{cases}$$

$$\text{for } x_1 = \pm L \text{ and } x_2 \in [-H, +H] (N_1 = \pm 1 \text{ and } N_2 = 0). \tag{22}$$

Assuming that equation (7) retains an orthotropic symmetry at any stage of straining up to bifurcation, it is appropriate to introduce the following coefficients:

$$\begin{cases} d_1 = C_{1111} - \sigma_{11} \\ d_2 = C_{2222} - \sigma_{22} \\ d_3 = C_{1212} - \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22} \\ d_4 = C_{1212} - \frac{1}{2}\sigma_{11} - \frac{1}{2}\sigma_{22} \\ d_5 = C_{1212} + \frac{1}{2}\sigma_{11} - \frac{1}{2}\sigma_{22} \\ d_6 = C_{1212} + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22} \\ d_7 = C_{1122} \\ d_8 = C_{2211}. \end{cases} \tag{23}$$

Equation (20) becomes:

$$\begin{cases} d_1 v_{1,11} + d_3 v_{1,22} + (d_4 + d_7) v_{2,12} = 0 \\ d_5 v_{2,11} + d_2 v_{2,22} + (d_4 + d_8) v_{1,12} = 0 \end{cases} \tag{24}$$

while equations (21) and (22) become:

$$\overset{\circ}{\Sigma}_{21} = d_3 v_{1,2} + d_4 v_{2,1} = 0 \quad \text{for } x_2 = \pm H \text{ and } x_1 \in [-L, +L] \tag{25}$$

$$\begin{cases} \overset{\circ}{\Sigma}_{11} = d_1 v_{1,1} + d_7 v_{2,2} = 0 \\ \overset{\circ}{\Sigma}_{12} = d_4 v_{1,2} + d_5 v_{2,1} = 0 \end{cases} \quad \text{for } x_1 = \pm L \text{ and } x_2 \in [-H, +H]. \tag{26}$$

The trivial solution of the boundary value problem posed in equations (24), (25), and (26) is a homogeneous stress and displacement field. Following the method of Hill and Hutchinson (1975), the existence of a bifurcated velocity field is investigated. This velocity field is either x_2-symmetric:

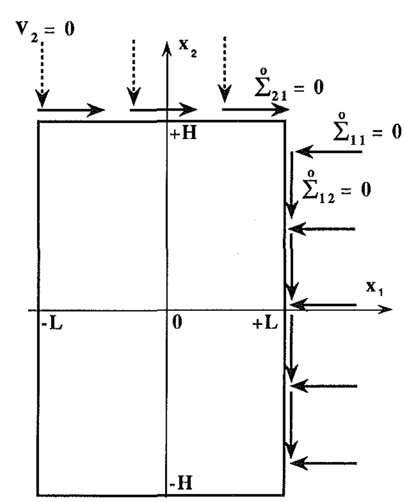


Fig. 1 Geometry and boundary conditions of block subjected to plane-strain compression

$$\begin{cases} v_1(x_1,x_2) = V_1(x_1)\cos\beta x_2 \\ v_2(x_1,x_2) = V_2(x_1)\sin\beta x_2 \end{cases} \tag{27}$$

or x_2-antisymmetric:

$$\begin{cases} v_1(x_1,x_2) = V_1(x_1)\sin\beta x_2 \\ v_2(x_1,x_2) = V_2(x_1)\cos\beta x_2 \end{cases}. \tag{28}$$

The coefficient β is selected to satisfy equation (21) for $x_2 = \pm H$ and $x_1 \in [-L,+L]$:

$$\beta H = m\frac{\pi}{2},\ m = 1,2,3,\text{etc.} \tag{29}$$

The velocity field is x_2-symmetric for even values of m and x_2-antisymmetric for odd values of m. By substituting equation (27) into equation (24), the following system of ordinary differential equations is obtained:

$$\begin{cases} d_1\dfrac{d^2V_1}{dx_1^2} - d_3\beta^2V_1 + (d_4+d_7)\beta\dfrac{dV_2}{dx_1} = 0 \\ d_5\dfrac{d^2V_2}{dx_1^2} - d_2\beta^2V_2 - (d_4+d_8)\beta\dfrac{dV_1}{dx_1} = 0. \end{cases} \tag{30}$$

The solutions of equation (30) must obey the following boundary conditions, which are obtained by substituting equation (27) into equation (26):

$$\begin{cases} d_1\dfrac{dV_1}{dx_1}(\pm L) + d_7\beta V_2(\pm L) = 0 \\ -d_4\beta V_1(\pm L) + d_5\dfrac{dV_2}{dx_1}(\pm L) = 0. \end{cases} \tag{31}$$

Equations (30) and (31) apply to x_2-symmetric solutions. For x_2-antisymmetric solutions, β is replaced by $-\beta$ in equations (30) and (31). The velocity solutions have the following type:

$$\begin{cases} V_1(x_1) = Ae^{i\alpha x_1} \\ V_2(x_1) = Be^{i\alpha x_1} \end{cases} \tag{32}$$

where

$$i = \sqrt{-1}. \tag{33}$$

After introducing the variable $Z = \alpha/\beta$, equation (30) becomes:

$$\begin{cases} (d_1Z^2 - d_3)A + (d_4+d_7)iZB = 0 \\ -(d_4+d_8)iZA + (d_5Z^2 - d_2)B = 0. \end{cases} \tag{34}$$

In order to obtain nontrivial solutions for A and B, the determinant of equation (34) is equal to zero, which gives the following characteristic equation:

$$aZ^4 + bZ^2 + c = 0 \tag{35}$$

where the coefficients a, b, and c are:

$$\begin{cases} a = d_1d_5 \\ b = d_1d_2 + d_3d_5 - (d_4+d_7)(d_4+d_8) \\ c = d_2d_3. \end{cases} \tag{36}$$

Depending on the values of a, b, and c, equation (35) has four different types of solution in Z:

(EI) elliptic imaginary when it has four imaginary roots,
(EC) elliptic complex when it has four complex roots,
(P) parabolic when it has two real and two purely imaginary roots, and
(H) hyperbolic when it has four real roots.

Most of the mathematical derivations are omitted hereafter. The following section presents each type of solution and divides them into x_1-symmetric and x_1-antisymmetric solutions as in Hill and Hutchinson (1975).

3.2 Elliptic Imaginary Solutions. Elliptic imaginary solutions are found when

$$\delta = b^2 - 4ac \geq 0,\ -\frac{b}{a} < 0 \text{ and } \frac{c}{a} \geq 0. \tag{37}$$

After defining the following real quantities

$$Z_1 = \sqrt{\frac{b-\sqrt{\delta}}{2a}} \quad Z_2 = \sqrt{\frac{b+\sqrt{\delta}}{2a}} \tag{38}$$

the general form of the x_1-symmetric bifurcating velocity field is

$$\begin{cases} V_1(x_1) = A_1\sinh(\beta Z_1x_1) + B_1\sinh(\beta Z_2x_1) \\ V_2(x_1) = A_2\cosh(\beta Z_1x_1) + B_2\cosh(\beta Z_2x_1). \end{cases} \tag{39}$$

By substituting the velocity field of equation (39) into equation (30) and enforcing the boundary conditions of equation (31), it can be shown that the velocity field of equation (39) emerges from the homogeneous velocity field when the following bifurcation condition is met

$$\frac{aZ_1^2 - f}{gZ_1^2 - h}Z_1\tanh(\beta Z_1L) - \frac{aZ_2^2 - f}{gZ_2^2 - h}Z_2\tanh(\beta Z_2L) = 0 \tag{40}$$

where the coefficients f, g, and h are

$$\begin{cases} f = d_3d_5 - d_4(d_4+d_7) \\ g = -d_1d_4 \\ h = d_3d_7. \end{cases} \tag{41}$$

The x_1-antisymmetric solution has the following general form:

$$\begin{cases} V_1(x_1) = A_1\cosh(\beta Z_1x_1) + B_1\cosh(\beta Z_2x_1) \\ V_2(x_1) = A_2\sinh(\beta Z_1x_1) + B_2\sinh(\beta Z_2x_1). \end{cases} \tag{42}$$

By using the same approach as for symmetric modes, the condition for the emergence of an x_1-antisymmetric velocity field is

$$\frac{aZ_1^2 - f}{gZ_1^2 - h}Z_1\operatorname{cotanh}(\beta Z_1L) - \frac{aZ_2^2 - f}{gZ_2^2 - h}Z_2\operatorname{cotanh}(\beta Z_2L) = 0. \tag{43}$$

3.3 Elliptic Complex Solutions. Elliptic complex solutions emerge when

$$\delta < 0. \tag{44}$$

By defining the following real quantities p and q such as:

$$p = \Im\left(\sqrt{\frac{-b+\sqrt{\delta}}{2a}}\right) \quad q = \Re\left(\sqrt{\frac{-b+\sqrt{\delta}}{2a}}\right), \tag{45}$$

the general form of the x_1-symmetric velocity field is:

$$\begin{cases} V_1 = A_1\sinh(\beta px_1)\cos(\beta qx_1) + B_1\cosh(\beta px_1)\sin(\beta qx_1) \\ V_2 = A_2\cosh(\beta px_1)\cos(\beta qx_1) + B_2\sinh(\beta px_1)\sin(\beta qx_1). \end{cases} \tag{46}$$

It can be shown that the x_1-symmetric velocity field emerges when:

$$[q\sinh(2\beta pH) + p\sin(2\beta qH)](ag(p^2+q^2)^2 - 2ah(p^2-q^2) + hf) + [q\sinh(2\beta pH) - p\sin(2\beta qH)](p^2+q^2)(gf-ah) = 0 \tag{47}$$

where the coefficients f, g, and h are given in equation (41).

The x_1-antisymmetric velocity field

$$\begin{cases} V_1 = A_1\cosh(\beta px_1)\cos(\beta qx_1) + B_1\sinh(\beta px_1)\sin(\beta qx_1) \\ V_2 = A_2\sinh(\beta px_1)\cos(\beta qx_1) + B_2\cosh(\beta px_1)\sin(\beta qx_1) \end{cases} \tag{48}$$

emerges when

$$[q\sinh(2\beta pH) - p\sin(2\beta qH)](ag(p^2+q^2)^2 - 2ah(p^2-q^2) + hf) + [q\sinh(2\beta pH) + p\sin(2\beta qH)](p^2+q^2)(gf-ah) = 0. \tag{49}$$

3.4 Parabolic Solutions. Parabolic solutions are found when

$$\frac{c}{a}<0. \tag{50}$$

By defining the following real quantities

$$Z_1=\sqrt{-\min\left(\frac{-b-\sqrt{\delta}}{2a},\ \frac{-b+\sqrt{\delta}}{2a}\right)}$$

$$Z_2=\sqrt{\max\left(\frac{-b-\sqrt{\delta}}{2a},\ \frac{-b+\sqrt{\delta}}{2a}\right)}, \tag{51}$$

the general form of the x_1-symmetric velocity field is

$$\begin{cases} V_1=A_1\sinh(\beta Z_1 x_1)+B_1\sin(\beta Z_2 x_1) \\ V_2=A_2\cosh(\beta Z_1 x_1)+B_2\cos(\beta Z_2 x_1). \end{cases} \tag{52}$$

It emerges when

$$\frac{aZ_1^2-f}{gZ_1^2-h}\,Z_1\tanh(\beta Z_1 L)+\frac{aZ_2^2+f}{gZ_2^2+h}\,Z_2\tan(\beta Z_2 L)=0. \tag{53}$$

The x_1-antisymmetric velocity field

$$\begin{cases} V_1=A_1\cosh(\beta Z_1 x_1)+B_1\cos(\beta Z_2 x_1) \\ V_2=A_2\sinh(\beta Z_1 x_1)+B_2\sin(\beta Z_2 x_1) \end{cases} \tag{54}$$

emerges when

$$\frac{aZ_1^2-f}{gZ_1^2-h}\,Z_1\text{cotanh}(\beta Z_1 L)+\frac{aZ_2^2+f}{gZ_2^2+h}\,Z_2\text{cotan}(\beta Z_2 L)=0. \tag{55}$$

3.5 Hyperbolic Solutions. Hyperbolic solutions are found when

$$\delta\geq 0,\ -\frac{b}{a}\geq 0 \text{ and } \frac{c}{a}\geq 0. \tag{56}$$

By defining the solutions as

$$Z_1=\sqrt{\frac{-b-\sqrt{\delta}}{2a}}\quad Z_2=\sqrt{\frac{-b+\sqrt{\delta}}{2a}}, \tag{57}$$

the general form of the x_1-symmetric velocity field is

$$\begin{cases} V_1=A_1\sin(\beta Z_1 x_1)+B_1\sin(\beta Z_2 x_1) \\ V_2=A_2\cos(\beta Z_1 x_1)+B_2\cos(\beta Z_2 x_1)^{\cdot} \end{cases} \tag{58}$$

It emerges when

$$\frac{aZ_1^2+f}{gZ_1^2+h}\,Z_1\tan(\beta Z_1 L)-\frac{aZ_2^2+f}{gZ_2^2+h}\,Z_2\tan(\beta Z_2 L)=0. \tag{59}$$

The x_1-antisymmetric velocity field

$$\begin{cases} V_1=A_1\cos(\beta Z_1 x_1)+B_1\cos(\beta Z_2 x_1) \\ V_2=A_2\sin(\beta Z_1 x_1)+B_2\sin(\beta Z_2 x_1) \end{cases} \tag{60}$$

emerges when

$$\frac{aZ_1^2+f}{gZ_1^2+h}\,Z_1\text{cotan}(\beta Z_1 L)-\frac{aZ_2^2+f}{gZ_2^2+h}\,Z_2\text{cotan}(\beta Z_2 L)=0. \tag{61}$$

3.6 Analytical Solution for Localized Bifurcation. According to Rice (1976), strain may bifurcate in localized modes instead of diffuse modes. Strain localizes when a velocity field **v** different from the homogeneous fields emerges in a planar region, which is referred to as shear band. The difference between the gradients of the velocity field inside and outside the shear band is

$$\Delta v_{i,j}=g_i n_j \tag{62}$$

where n_j represents the normal to the shear band and g_i depends only on the distance across the band and vanishes outside the band. Instead of being enforced uniformly through partial differential equations (equation (26)), the equilibrium of stress rate is enforced only in the direction normal to the shear band

$$n_j\Delta\mathring{\Sigma}_{ji}=0 \tag{63}$$

where $\Delta\mathring{\Sigma}_{ji}$ is the difference between the nominal stress rate inside and outside the band. By substituting the constitutive equation (15) into (63), the following system of two linear equations with two unknowns g_1 and g_2 is obtained in the case of plane-strain deformation:

$$\begin{cases} g_1(d_1 n_1^2+d_3 n_2^2)+g_2(d_4+d_7)n_1 n_2=0 \\ g_1(d_4+d_8)n_1 n_2+g_2(d_5 n_1^2+d_2 n_2^2)=0. \end{cases} \tag{64}$$

Equation (64) has nontrivial solutions when its determinant is equal to zero, i.e., when the variable Z,

$$Z=\frac{n_2}{n_1}, \tag{65}$$

obeys the characteristic equation for diffuse bifurcation, equation (35). Therefore, (35) also gives the inclination of the shear band. However, since localized modes are not required to obey boundary conditions they systematically appear in the parabolic and hyperbolic regimes.

4 Applications

The general solutions derived in the previous sections are verified by comparing them to existing results for particular materials.

4.1 Nearly Incompressible Solids. The present solutions generalize the results of Biot (1965), Hill and Hutchinson (1975), and Young (1976) by defining the following linearized solids:

$$\begin{cases} C_{1111}=2\mu_1+\lambda \\ C_{2222}=2\mu_2+\lambda \\ C_{1212}=2\mu \\ C_{1122}=C_{2211}=\lambda \end{cases} \tag{66}$$

where μ_1 and μ_2 are the shear moduli in the x_1 and x_2-directions, respectively, μ is the shear modulus, and λ is Lame's modulus. Shear modulus and Lame's modulus are related through Poisson's ratio ν

$$\lambda=\frac{2\mu^*\nu}{1-2\nu}. \tag{67}$$

μ^* is the average shear modulus

$$\mu^*=\frac{1}{2}(\mu_1+\mu_2). \tag{68}$$

The constitutive equation of Hill and Hutchinson (1975),

$$\begin{cases} \hat{\tau}_{11}-\hat{\tau}_{22}=2\mu^*(D_{11}-D_{22}) \\ \hat{\tau}_{12}=2\mu D_{12} \\ D_{11}+D_{22}=0, \end{cases} \tag{69}$$

may be approximately described by equation (66) by using nearly incompressible materials, e.g., ν = 0.4999. Figure 2 represents the bifurcation regimes of nearly incompressible materials as a function of the dimensionless stress σ_{22}/μ^* and moduli ratio μ/μ^* when the lateral stress σ_{11} is equal to zero. As in Hill and Hutchinson (1975), Fig. 2 shows that the bifurcation domains coincide for compressive and tensile stress. Figure 3 represents the buckling stress σ_{22}/μ^* versus the wavelength of the diffuse mode of bifurcation $m(\pi L)/(2H)$ in the EC, H, and P domains for the particular case μ/μ^* = 5. For slender specimens (L/H < 1) subjected to compression, the bifurcation stress of the first antisymmetric mode coincides with the buckling stress of a Euler column that is axially loaded while its extremities are prevented from rotating (Timoshenko and Gere, 1961):

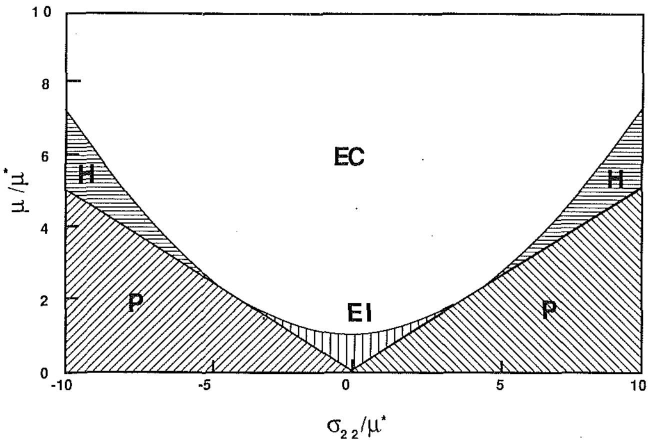


Fig. 2 Bifurcation regimes of incompressible hypoelastic materials (ν = 0.4999 and σ_{11} = 0)

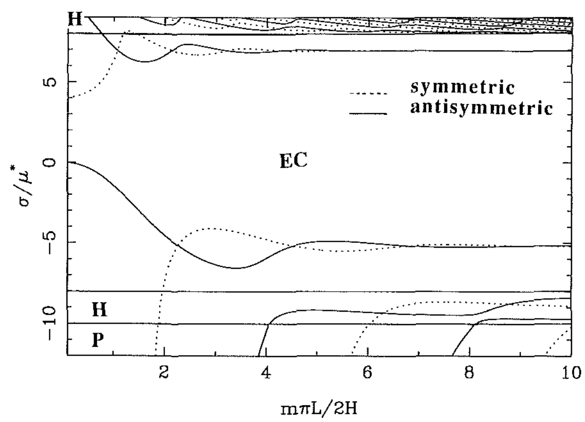


Fig. 3 Buckling stress versus wavelength of the elliptic (E), parabolic (P), and hyperbolic (H) bifurcation modes of nearly compressible hypoelastic material (μ/μ^* = 5, ν = 0.4999, and σ_{11} = 0)

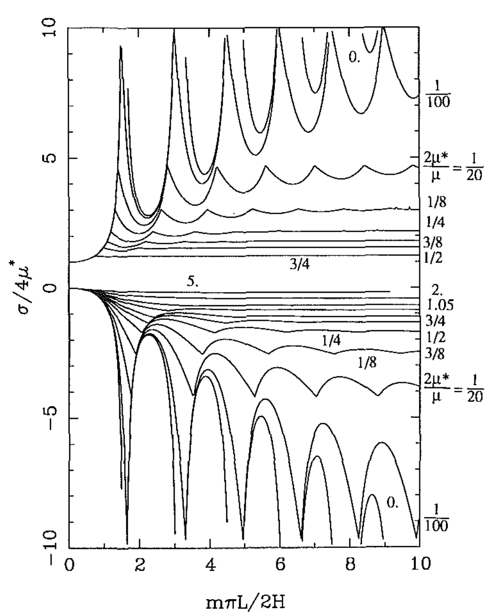


Fig. 4 Lowest bifurcation stress of symmetric and antisymmetric modes during plane-strain compression and tension of nearly incompressible solids (ν = 0.4999)

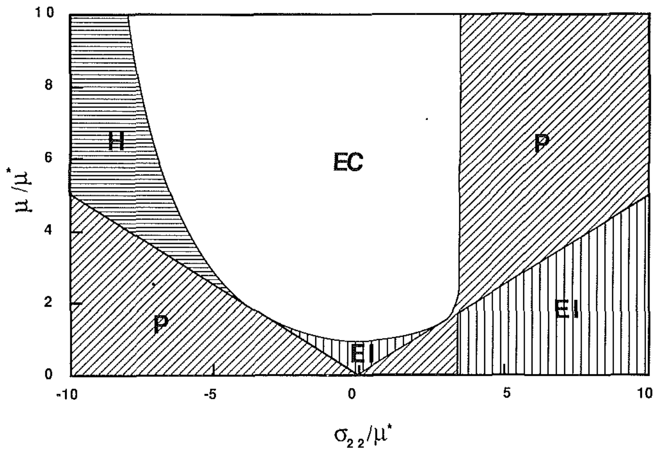


Fig. 5 Bifurcation regimes of compressible hypoelastic materials (ν = 0.3)

$$\frac{\sigma_{22}}{\mu^*}=\frac{(1+\nu)\pi^2}{6}\left(\frac{L}{H}\right)^2. \tag{70}$$

For bulky specimens $\left(\frac{L}{H}\to\infty\right)$ and very short wavelengths ($m\to\infty$), the symmetric and antisymmetric modes emerge for the same buckling stress independently of the wavelength mL/H, which is a characteristic result for surface instability (Biot, 1965). For intermediate wavelengths, symmetric and antisymmetric bifurcations alternate to give the lowest buckling stress. Figure 3 also displays the P and H solutions in addition to the EC solutions. The buckling stress varies continuously through the EC-H and H-P transitions. The symmetric and antisymmetric modes alternate orderly in the P domain, while they interlace in wavily fashion in the H regimes. As pointed out by Hill and Hutchinson (1975), the lowest buckling stress in the P (or H) regime coincides with the E-P (or E-H) transition for bulky specimens or very short wavelengths.

Figure 4 shows the lowest bifurcation stress as a function of the wavelength $\gamma = m(\pi L)/(2H)$ obtained for symmetric and antisymmetric elliptic modes of bifurcation corresponding to selected values of μ/μ^*. Since Fig. 4 duplicate Young's results, it is concluded that our analytical solutions apply to incompressible solids.

4.2 Compressible Solids. Figure 5 shows the bifurcation diagram of compressible materials (equation (66) and ν = 0.3). The E, P, and H domains are different for compressive and tensile stress, which contrasts to the symmetric domains of Fig. 2. As shown for tensile stress in Fig. 5, the P domain moves to partially occupy the EC and H domains of Fig. 2 and to give its place to EI solutions. Strain localization and internal buckling with very short wavelength emerges when the tensile stress σ_{22} reaches about $-3.5\mu^*$. As in Fig. 4, Fig. 6 represents the compressive and tensile buckling stress σ_{22}/μ^* versus the wavelength $m(\pi L)/(2H)$ of the diffuse bifurcation for the particular case μ/μ^* = 5. The first antisymmetric mode emerges in the P regime in tension and in the EC regime for compression. Diffuse elliptic bifurcation is not obtained during tension of bulky specimens ($L/H > 1$).

4.3 Elastoplastic Mohr-Coulomb Materials. The incremental stress-strain relationship of the flow theory of plasticity (Hill, 1950) has the following coefficients:

$$C_{ijkl}=\mu(\delta_{ik}\delta_{jl}+\delta_{ik}\delta_{jl}) + \lambda\delta_{ij}\delta_{kl}-\frac{[\mu P_{ij}+\lambda(P_{rr})\delta_{ij}][\mu Q_{kl}+\lambda(Q_{ss})\delta_{kl}]}{H+\lambda(P_{aa})(Q_{bb})+2\mu P_{cd}Q_{cd}} \tag{71}$$

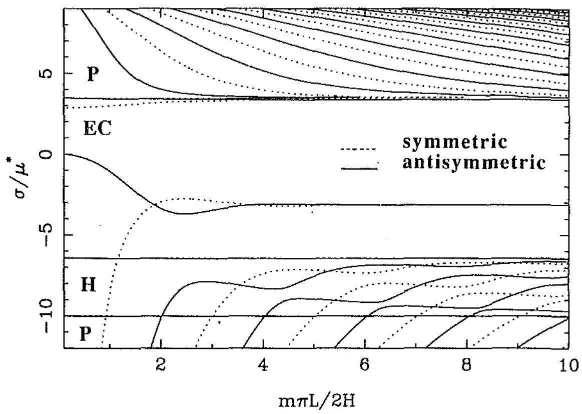


Fig. 6 Buckling stress versus wavelength of the elliptic, parabolic, and hyperbolic bifurcation modes of compressible hypoelastic material ($\mu/\mu^* = 5$, $\nu = 0.3$ and $\sigma_{11} = 0$)

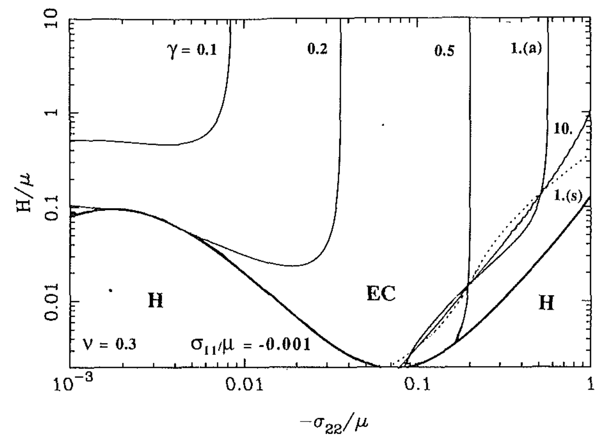


Fig. 8 Bifurcation regimes and elliptic symmetric (dotted line) and antisymmetric modes for non-associative compressible elastoplastic solids (ϕ_μ = 30 deg)

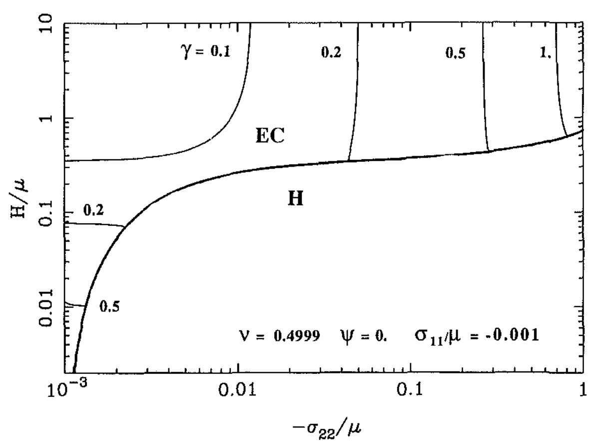


Fig. 7 Bifurcation regimes and elliptic bifurcation modes for nearly incompressible elastoplastic solids (ψ = 0 deg)

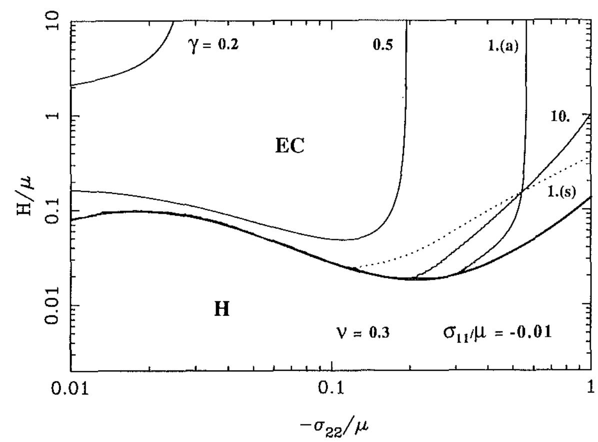


Fig. 9 Bifurcation regimes and elliptic symmetric (dotted line) and antisymmetric modes for nonassociative compressible elastoplastic solids (ϕ_μ = 30 deg)

where H is the plastic modulus, μ the elastic shear modulus, λ the elastic Lame's modulus, and P_{ij} and Q_{kl} the plastic flow and yield directions, respectively. The flow direction P_{ij} of elastoplastic Mohr-Coulomb solids is

$$\left\{\begin{array}{l} P_{11}=\dfrac{1-\sin\psi}{\sqrt{2(1+\sin^2\psi)}} \\ P_{22}=-\dfrac{1+\sin\psi}{\sqrt{2(1+\sin^2\psi)}} \\ P_{33}=0 \end{array}\right. \tag{72}$$

where the dilatancy angle ψ (Rowe, 1971) is

$$\sin\psi=-\frac{P_{11}+P_{22}}{P_{11}-P_{22}}=-\frac{d\epsilon_{11}^p+d\epsilon_{22}^p}{d\epsilon_{11}^p-d\epsilon_{22}^p}. \tag{73}$$

The yield direction Q_{ij} is

$$\left\{\begin{array}{l} Q_{11}=\dfrac{1-\sin\phi}{\sqrt{2(1+\sin^2\phi)}} \\ Q_{22}=-\dfrac{1+\sin\phi}{\sqrt{2(1+\sin^2\phi)}} \\ Q_{33}=0 \end{array}\right. \tag{74}$$

where ϕ is the tangential friction angle. In the case of cohesionless materials, ϕ is approximately equal to the mobilized friction angle ϕ_m,

$$\sin\phi \approx \sin\phi_m = \frac{\sigma_1-\sigma_3}{\sigma_1+\sigma_3}, \tag{75}$$

where σ_1 and σ_3 are the major and minor principal stresses, respectively.

4.3.1 Nearly Incompressible Elastoplastic Mohr-Coulomb Solids. By selecting $\psi = 0$ and $\nu = 0.4999$, the present analysis generalizes the work of Needleman (1979) for incompressible elastoplastic solids with nonassociative flow rule.

Figure 7 represents the various regimes of bifurcation as a function of $-\sigma_{22}/\mu$ and H/μ. The logarithmic representation of Fig. 7 allows us to include stresses σ_{22} that are small compared to the elastic shear modulus. σ_{11}/μ is assumed to be equal to 10^{-3}, which is a reasonable value for sands. Lade and Nelson (1987) established experimentally that μ for sand is typically between 100 to 1000 times larger than the current stress level. If the material is isotropically stressed ($\sigma_{11} = \sigma_{22}$) and behaves elastically ($H/\mu = +\infty$) at the beginning of the plane-strain compression, the point $(-\sigma_{22}/\mu, H/\mu)$ lies in the upper left corner of Fig. 7. As the compression proceeds, the point $(-\sigma_{22}/\mu, H/\mu)$ moves gradually downward and toward a failure point on the σ_{22}/μ-axis. Therefore, diffuse bifurcation is first possible in the EC domain. As the plastic modulus

H/μ decreases further, localized bifurcations may occur first at the EC-H transition and later in the H domain.

Figure 7 also represents the location of the emergence of the antisymmetric elliptic modes corresponding to the wavelengths $\gamma = m(\pi L)/(2H) = 0.1$, 0.2, 0.5, and 1. Symmetric modes do not emerge in the EC domain for the selected range of parameters. It is concluded that slender specimens (e.g., $L/H < 0.1$) are more likely to bifurcate than bulky specimens ($L/H > 1$) in antisymmetric modes.

4.3.2 Compressible Elastoplastic Mohr-Coulomb Solids. Needleman's analysis can be extended to compressible elastoplastic materials. As in Rowe (1971) and Vardoulakis (1981), the dilatancy angle varies as a function of the mobilized friction angle ϕ_m,

$$\sin\psi = -\frac{\sin\phi_m - \sin\phi_\mu}{1-\sin\phi_m \sin\phi_\mu}, \tag{76}$$

where ϕ_μ is the friction angle between particles. Equation (76) implies that the material is compacting for $\phi_m < \phi_\mu$, incompressible for $\phi_m = \phi_\mu$, and dilatant for $\phi_m > \phi_\mu$.

As in Fig. 7, Figs. 8 and 9 shows the various bifurcation regimes and the elliptic modes for compressible elastoplastic materials. The typical value of 30 deg is selected for ϕ_μ. The lateral stress σ_{11} is arbitrarily fixed to $\sigma_{11}/\mu = -0.001$ and $\sigma_{11}/\mu = -0.01$. The axial stress σ_{22} varies, but remains smaller than μ. As shown in Figs. 8 and 9, the symmetric and antisymmetric EC modes appear simultaneously for $\gamma = 10$. This case corresponds to surface instability ($L/H \to \infty$) and diffuse bifurcation with very short wavelength ($m \to \infty$). Antisymmetric and symmetric modes also appear almost simultaneously for $\gamma = 1$. However, in contrast to antisymmetric modes, symmetric modes do not appear for $\gamma < 0.5$. Figures 8 and 9 show that diffuse bifurcations may occur when the applied stress σ_{22} is on the order of μ, which is unrealistic for cohesionless materials. The comparison of Fig. 7 and 8 indicates that the flow rule (equation (76)) influences the EC-H transition (i.e., strain localization) more than it affects diffuse elliptic bifurcation. Based on the scarcity of diffuse bifurcation solutions for the parameters characteristic of sands, it is concluded with Vardoulakis (1981) that sand specimens may deform under plane-strain rectilinear deformation without experiencing diffuse bifurcation.

5 Conclusion

Analytical solutions have been derived for the diffuse bifurcations of compressible solids subjected to plane-strain loading. These solutions generalize the works for incompressible solids of Biot (1965) and Hill and Hutchinson (1975). They describe diffuse symmetric and antisymmetric bifurcations and localized bifurcations. They have been validated by applying them to the incompressible solids of Hill and Hutchinson (1975), Young (1976), and Needleman (1979) and to the elastoplastic Mohr-Coulomb materials of Vardoulakis (1981).

The present analytical solutions are capable of reproducing the lowest bifurcation stresses for incompressible solids calculated by Hill and Hutchinson (1975). In contrast to incompressible solids that systematically experience diffuse elliptic bifurcations before any other types of bifurcation, compressible materials are found to reach the elliptic-hyperbolic boundary (i.e., to localize) prior to encountering diffuse elliptic bifurcation. When applied to elastoplastic Mohr-Coulomb materials, the analysis corroborates the conclusions of Needleman (1979) and Vardoulakis (1981). It emphasizes that the elastic shear modulus influences diffuse bifurcation much more than the flow rule.

The proposed analytical solutions are useful to assess the role of material compressibility on plane-strain bifurcations, to develop plane-strain bifurcation analysis of specific solids and to calibrate the numerical bifurcation analysis performed with finite elements.

Acknowledgment

The financial support of the National Science Foundation is acknowledged (grant MSM 8657999).

References

Bardet, J. P., 1990, "Finite Element Analysis of Surface Instability," *J. of Computer Methods in Applied Mechanics and Engineering*, Vol. 78, pp. 273–296.

Computer Methods in Applied Mechanics and Engineering, in press.

Bazant, Z. P., 1971, "A Correlation Study of Formulations of Incremental Deformation and Stability of Continuous Bodies," ASME Journal of Applied Mechanics, pp. 919–928.

Biot, M., 1965, *Mechanics of Incremental Deformations*, John Wiley and Sons, New York.

deBorst, R., 1987, "Computation of Post-Bifurcation and Post-Failure Behavior of Strain Softening Solids," *Computers and Structures*, Vol. 25, pp. 211–224.

Chau, K. T., and Rudnicki, J. W., 1989, "Bifurcation of Compressible Pressure-Sensitive Materials in Plane Strain Tension and Compression," submitted to *J. Mech. Phys. Solids*.

Hill, R., 1950, *The Mathematical Theory of Plasticity*, Oxford University Press, London.

Hill, R., 1959, "Some Basic Principles in the Mechanics of Solids Without a Natural Time," *J. Mech. Phys. Solids*, Vol. 7, p. 209.

Hill, R., and Hutchinson, J. W., 1975, "Bifurcation Phenomena in the Plane Tension Test," *J. Mech. Phys. Solids*, Vol. 23, pp. 239–264.

Lade, P. V., and Nelson, R. B., 1987, "Modelling the Elastic Behavior of Granular Materials," *Int. Jnl. for Numerical and Analytical Methods in Geomechanics*, Vol. 11, pp. 521–542.

McMeeking, R. M., and Rice, J. R., 1975, "Finite Element Formulation for Problems of Large Elastic-Plastic Deformation," *Int. Jnl. Solids Structures*, Vol. 11, pp. 601–616.

Needleman, A., 1979, "Non-normality and Bifurcation in Plane Strain Tension and Compression," *J. Mech. Phys. Solids*, Vol. 27, pp. 231–254.

Needleman, A., and Tvergaard, V., 1984, "Finite Element Analysis of Localization in Plasticity," *Finite Elements: Special Problems in Solid Mechanics*, Vol. 5, J. T. Oden and G. F. Carey, eds., Prentice-Hall, New York, pp. 94–157.

Rowe, P. W., 1971, "Theoretical Meaning and Observed Values of Deformation Parameters for Soil," *Proceedings of the Roscoe Memorial Symposium*, R. H. G. Perry, ed., Cambridge University Press, pp. 143–194.

Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw-Hill, New York, p. 49.

Truesdell, C., and Noll, W., 1965, *Nonlinear Field Theory of Mechanics* (Handbuch of Physics III/3), Springer-Verlag, Berlin.

Vardoulakis, I., 1981, "Bifurcation Analysis of the Plane Rectilinear Deformation on Dry Sand Samples," *Int. J. Solids and Structures*, Vol. 17, No. 11, pp. 1085–1101.

Young, N. J. B., 1976, "Bifurcation Phenomena in Plane Compression Test," *J. Mech. Phys. Solids*, Vol. 24, pp. 77–91.

T. Belytschko
Fellow ASME

B. Moran
Assoc. Mem. ASME

M. Kulkarni

Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208

On the Crucial Role of Imperfections in Quasi-static Viscoplastic Solutions

The stability and structure of shear bands and how they relate to initial imperfections is studied within the framework of a one-dimensional boundary value problem. It is shown that in strain-softening viscoplasticity the structure of the band depends on the structure of the imperfection. A Fourier analysis shows that the width of the shear band depends directly on the width of the imperfection, suggesting that the imperfection scales the response of the viscoplastic material. For continuously differentiable imperfections, the shear band is continuously differentiable, whereas when the imperfection is C^0 at the maximum, the shear band is C^0, and cusp-shaped. For step function imperfections, the shear band is shown to be a step function, but it is shown that this solution is unstable.

1 Introduction

The scaling or size of a shear band is an important physical feature which does not emanate from the governing differential equations for the quasi-static response of a viscoplastic material, which does not possess a parameter of the dimension of length. A length scale does emerge in a dynamic or coupled thermal boundary value problem. However, in a Hopkinson torsional experiment, the rise time is very large compared to the traversal time of a torsional pulse across the shear band and yet too small to engender significant thermal conduction. Therefore, it can be argued, that these scale factors do not explain the scatter in size of shear band widths as reported by experimentalists such as Duffy (1984) and Marchand and Duffy (1988).

In this paper, we propose that the shear bands are scaled primarily by the imperfections in the specimen. For this purpose, we will examine the effect of various initial imperfections on the structure of the shear band. For this purpose, a closed-form solution of a traction controlled boundary value problem, and some numerical solutions of a velocity-controlled boundary value problem, are used.

Considerable progress has been achieved in the past decade in the understanding of shear banding. Strain-softening viscoplastic models have proven to be effective in capturing some important characteristics of shear banding. Molinari and Clifton (1987) have shown that viscoplastic material models exhibit what they call L_∞ localization. Pan (1983) has shown that viscoplastic material models are unstable if the strain field is perturbed as the stress passes a maximum. It has been shown that with strain gradient regularization, perturbations with wavelengths above a threshold, which depends on the localization parameter, do not grow unboundedly (Lasry and Belytschko, 1988). Wu and Freund (1988) have shown that for dynamic problems, rate-dependent softening models exhibit a phenomenon which they term "deformation-trapping" as a consequence of imaginary wave speeds. They then showed that they could obtain a solution numerically by adding rate dependence (viscoplasticity) and coupled heat transfer. Bazant and Belytschko (1985) obtained a closed-form solution for rate-independent strain softening. They showed that the strain softening was limited to a set of measure zero. Needleman (1988) showed that the ill-posedness of the rate-independent model could be eliminated by the addition of rate dependence in the form of a viscoplastic model.

Shawki and Clifton (1989) have presented a closed-form solution for the strain field subjected to a temperature perturbation, neglecting the effects of elasticity. Tzavaras (1986) has addressed the issues of the existence of classical solutions and stability of uniform shearing, using a rigorous mathematical analysis. Wright and Batra (1985) and Wright and Walter (1987) have reported numerical solutions for viscoplastic dynamic problems coupled with heat transfer. Wright and Walter show that the shear strain distribution at the maximum is actually flat over a very small distance; however, extremely fine meshes with a logarithmic distribution about the center of the band and on the order of 10^3 mesh points were needed to achieve such solutions.

In most of these solutions, the role of the initial imperfections, or perturbations of the initial data, was given little attention. Most of the authors cited have used triangular imperfections or step imperfections. In this paper, a closed-form solution is developed for a traction-controlled boundary value problem which is essentially one of constitutive response.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 1, 1989; final revision, May 21, 1990.

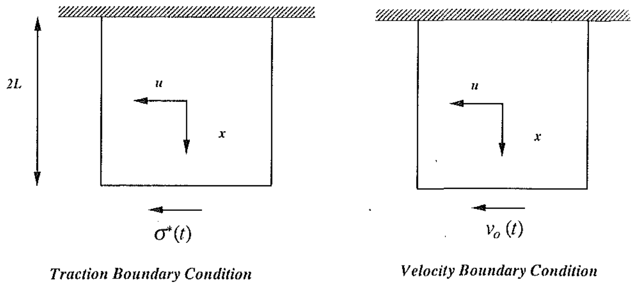


Fig. 1 A slab of viscoplastic material subjected to simple shear

This simple solution facilitates the demonstration of the crucial role of imperfections in the evolution of the strain field. It is shown via one-dimensional finite element studies that similar behavior is also seen in a velocity-controlled problem. In both cases for a quasi-static analysis, the strain field in the shear band is governed by the initial imperfection.

We will adopt the classical mathematical viewpoint of dynamical systems theory in that an evolution solution is considered stable if a perturbation results in a solution which is in a finite "hypersphere" about the unperturbed solution. By using this framework, we show that the homogeneous solution for a strain-softening viscoplastic material is unstable. Solutions which are step functions with one subdomain in a strain-softening regime, the other in a nonsoftening viscoplastic regime, are also shown to be unstable solutions. These solutions cannot be achieved physically and are meaningless.

The morphology of the shear band is shown to depend strongly on the structure of the imperfection. If the imperfection is smooth (C^1), then the shear band has the same character with a continuous derivative at its maximum. If the imperfection is C^0 with a discontinuity in its derivative at the maximum, then the shear band has a cusped structure with discontinuous derivatives at its maximum. Furthermore, we show that imperfections introduce a length scale into the evolving shear band. In particular, the spectrum of the shear band is related to the spectrum of the imperfection. It is possible that the viscoplastic model, when used correctly with imperfections that are representative of those which occur in nature, will reproduce the strain fields observed in experiments.

The behavior of the strain field at the point of maximum strain is examined using a Taylor series expansion, and it is shown that though the viscoplastic model is well posed, and leads to a unique solution at a given time, the shear band narrows with increasing deformation. This feature has also been observed experimentally.

This paper is organized as follows. In Section 2, the relevant governing equations are outlined, and in Section 3, a closed-form solution of the governing equations for a traction boundary condition is presented. In Section 4, 5, and 6, the issues of stability, scaling, and the narrowing of the shear band are addressed using the closed-form solution as a vehicle. In Section 7, the conclusions reached in Sections 4, 5, and 6 are verified for a velocity boundary condition by one-dimensional linear finite element studies.

2 Governing Differential Equations for Shear Banding

Consider a viscoplastic slab of length $2L$ subjected to pure shear as shown in Fig. 1. The slab is fixed at $x = -L$: the edge $x = +L$, may be subjected to either a velocity v_o or a traction σ^*. The relevant governing equations are the momentum equation,

$$\sigma_{,x} = \rho\, \ddot{u}, \tag{1}$$

the strain displacement equation,

$$\epsilon = u_{,x}, \tag{2}$$

the viscoplastic constitutive equations,

$$\dot{\sigma} = G(\dot{\epsilon} - \dot{\epsilon}^{vp}) \tag{3a}$$

$$\dot{\epsilon}^{vp} = \dot{\epsilon}^{vp}(\sigma, \epsilon^{vp}, \theta), \tag{3b}$$

and the energy balance equation,

$$\rho C_p \dot{\theta} = k\theta_{,xx} + \kappa\sigma\dot{\epsilon}^{vp}. \tag{4}$$

In the above, σ is the shear stress, ϵ is the corresponding shear strain, G is the elastic shear modulus, ρ is the density, C_p is the specific heat at constant pressure, k is the thermal diffusivity, κ is the Taylor-Quinney coefficient, and θ is the temperature. Superposed dots denote material time derivatives and commas denote derivatives with respect to the variable which follows.

The boundary conditions are

(1) Traction controlled:

$$\sigma(L,t) = \sigma^*(t) \tag{5a}$$

$$u(-L) = 0. \tag{5b}$$

(2) Velocity controlled:

$$\dot{u}(L,t) = v_0(t) \tag{5c}$$

$$u(-L) = 0. \tag{5d}$$

In this paper attention is restricted to the quasi-static, isothermal case, i.e., with $\rho\ddot{u}$, k, and $\rho C_p \dot{\theta}$ assumed to be negligible.

3 Closed-Form Solution for a Traction Boundary Condition

We consider a slab of length $2L$ subjected to a shear traction which rises instantaneously to a value σ^* and then remains constant (see Fig. 1). The boundary conditions are

$$\sigma(L,t) = \sigma^* H(t) \tag{6a}$$

$$u(-L) = 0 \tag{6b}$$

where $H(t)$ is the Heaviside step function.

The viscoplastic constitutive function ($3b$) is chosen to be

$$\dot{\epsilon}^{vp} = a_0 \left(\frac{\sigma}{g(\epsilon^{vp})}\right)^{\frac{1}{m}}. \tag{7}$$

Here, a_0 and m are material data, $0 < m \leq 1$ and $m \to 0$ represents the rate independent limit. Equilibrium in the quasi-static case implies

$$\dot{\sigma} = 0 \text{ for } x \in [-L, +L], t > 0. \tag{8a}$$

Hence,

$$\sigma = \sigma^*,\ t > 0. \tag{8b}$$

As a consequence of the loading condition, the elastic strain remains constant and is given by

$$\epsilon^* = \frac{\sigma^*}{G}. \tag{9a}$$

Therefore,

$$\dot{\epsilon} = \dot{\epsilon}^{vp}. \tag{9b}$$

The plastic response function for the softening material is taken to be

$$g(\epsilon^{vp}) = \sigma_Y \left(\frac{\epsilon_0 + \epsilon^{vp}}{\epsilon_0}\right)^{-p} \tag{10}$$

so that when $\epsilon^{vp} = 0$, $g(\epsilon^{vp}) = \sigma_Y$, where $p > 0$ is a material constant, σ_Y is a reference stress that corresponds to the yield stress and $\epsilon_o = \dfrac{\sigma_Y}{G}$.

Let

$$e = \epsilon_o + \epsilon - \epsilon^*. \tag{11}$$

Using equation (10), and since the elastic strain rate $\dot{\epsilon}^e = 0$, for $t > 0$, the viscoplastic response function may be written as

$$g(e) = \sigma_Y \left(\frac{e}{\epsilon_o}\right)^{-p}. \tag{12}$$

This, with equation (7) and equation (8*b*), gives

$$\sigma^* = g(e)\left(\frac{\dot{e}}{a_o}\right)^m. \tag{13}$$

To represent initial imperfections, the yield stress σ_Y is perturbed as follows:

$$\sigma_Y = \sigma_o(1 - \mu f(x)) \tag{14}$$

where μ is a small parameter representing the amplitude of the strength imperfection. Combining equation (13) and equation (14) yields

$$\frac{\sigma}{\sigma_o \alpha(x)} = \left(\frac{e}{\epsilon_o}\right)^{-p}\left(\frac{\dot{e}}{a_o}\right)^m \tag{15}$$

where $\alpha(x) = (1 - \mu f(x))$.

Introduce the following dimensionless variables

$$\bar{t} = a_o t \tag{17a}$$

$$\gamma = \frac{e}{\epsilon_o} \tag{17b}$$

$$\tau^* = \frac{\sigma^*}{\sigma_o \epsilon_o{}^m}. \tag{17c}$$

Equation (9*a*) may now be written in dimensionless form as

$$\tau^* = \alpha(x)\gamma^{-p}\left(\frac{\partial \gamma}{\partial \bar{t}}\right)^m. \tag{18}$$

Integrating in time yields

$$\int_0^{\bar{t}} \left\{\frac{\tau^*}{\alpha(x)}\right\}^{\frac{1}{m}} d\zeta = \int_1^{\gamma} \eta^{-\frac{p}{m}}\left(\frac{\partial \eta}{\partial \bar{t}}\right) d\bar{t} \tag{19}$$

where η and ζ are dummy variables. After integration, the following expression is obtained for the strain field

$$\gamma(x, \bar{t}) = \left[\left(\frac{m-p}{m}\right)\left\{\frac{\tau^*}{\alpha(x)}\right\}^{\frac{1}{m}1/m}(\bar{t} + 1)\right]^{\frac{m}{m-p}}. \tag{20}$$

For simplicity of notation, this expression is recast in the following form and is used as such henceforth.

$$\gamma(x, \bar{t}) = \left[1 + \xi \bar{t} \alpha^{-\frac{1}{m}}\right]^n \tag{20a}$$

where

$$\frac{m}{m-p} = n \tag{20b}$$

$$\left(\frac{m-p}{m}\right)\tau^{*\frac{1}{m}} = \xi. \tag{20c}$$

Unless stated otherwise all results presented subsequently have been obtained using $\tau^* = 1$, $m = 0.01$, $p = 0.02$, and all spatial profiles of field variables have been evaluated at $\bar{t} = 0.9$. Shawki and Clifton (1989) have demonstrated via a closed-form solution which neglects elastic effects that the behavior of the strain field depends strongly on the sign of the exponent. They also showed that the most interesting case for modelling shear band development is when the exponent is negative, i.e., for a strain-softening case.

4 Stability of the Strain Field Solution

For a perfect slab, i.e., $\alpha(x) = 1.0$, equation (20) yields the homogeneous solution

$$\gamma(x, \bar{t}) = [1 + \xi \bar{t}]^n. \tag{21}$$

However, this solution is unstable when the exponent is negative. Stability requires that if the initial data is perturbed by a small amount, then the solutions must differ by a small amount. To be specific, if solutions with initial data $\alpha_1(x)$ and $\alpha_2(x)$ are $u_1(x, \bar{t})$ and $u_2(x, \bar{t})$, then if

$$\|\alpha_1(x) - \alpha_2(x)\| < \mu, \tag{22}$$

then $u_1(x)$ is a stable solution is

$$\|\alpha_1(x, \bar{t}) - u_2(x, \bar{t})\| < C\mu \quad \forall \bar{t}, \tag{23}$$

where C is a constant, μ is a small constant, and $\|\cdot\|$ indicates a norm. If

$$\alpha_1(x) = 1.0 \tag{24}$$

$$\alpha_2(x) = 1 - \mu \cos\frac{\pi x}{L} \tag{25}$$

then $\gamma_1(x, \bar{t})$ is a homogeneous solution given by equation (21), whereas $\gamma_2(x, \bar{t})$ and the numerically integrated displacement fields $u_1(x, \bar{t})$ and $u_2(x, \bar{t})$ are shown in Fig. 3. It can be seen that the two displacement fields do not satisfy equation (23), so the homogeneous solution $u_1(x, \bar{t})$ is unstable.

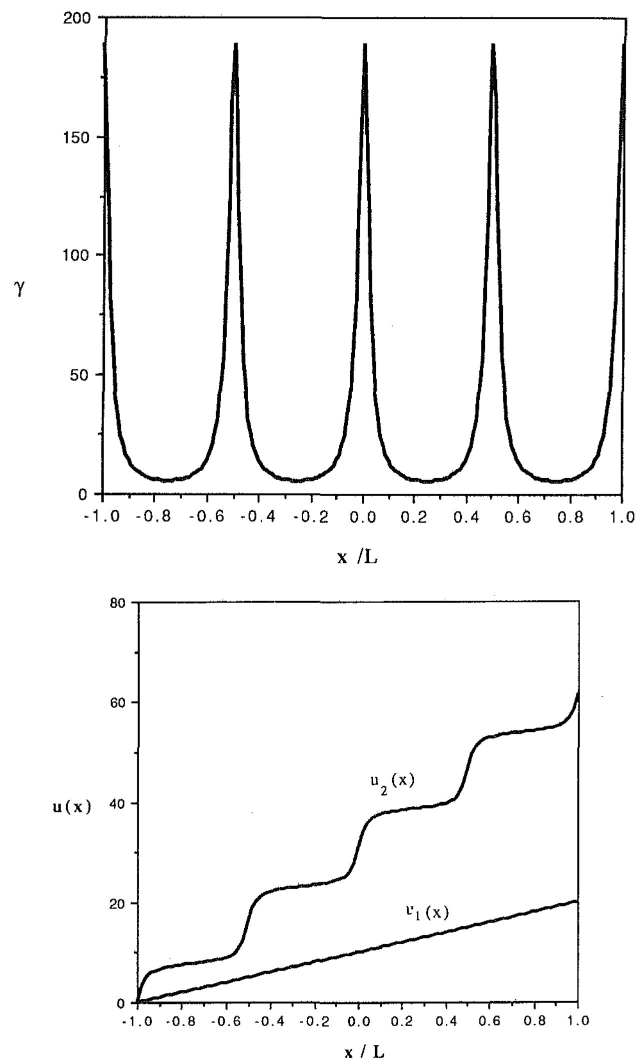


Fig. 2 Strain and displacement field response to the imperfection represented by equation (25)

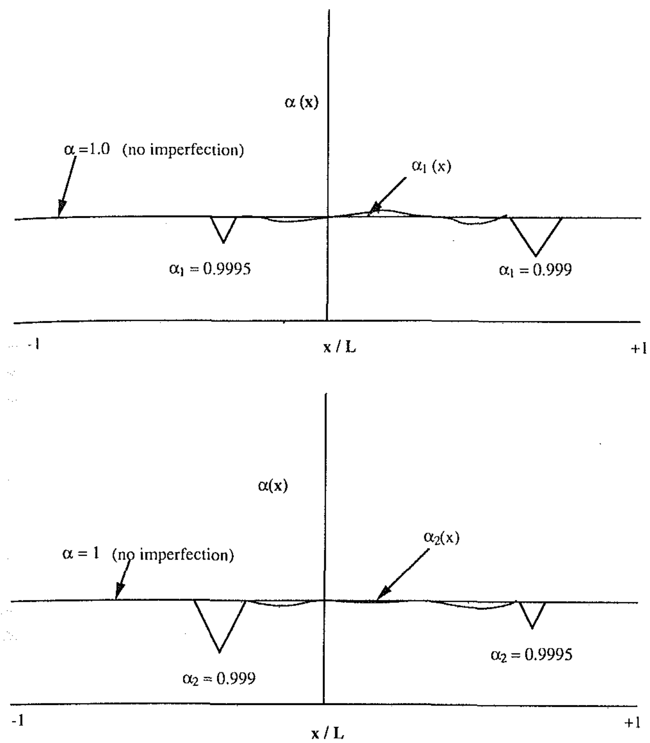


Fig. 3 Imperfections used to illustrate the random location of shear bands. Note that these imperfections are different.

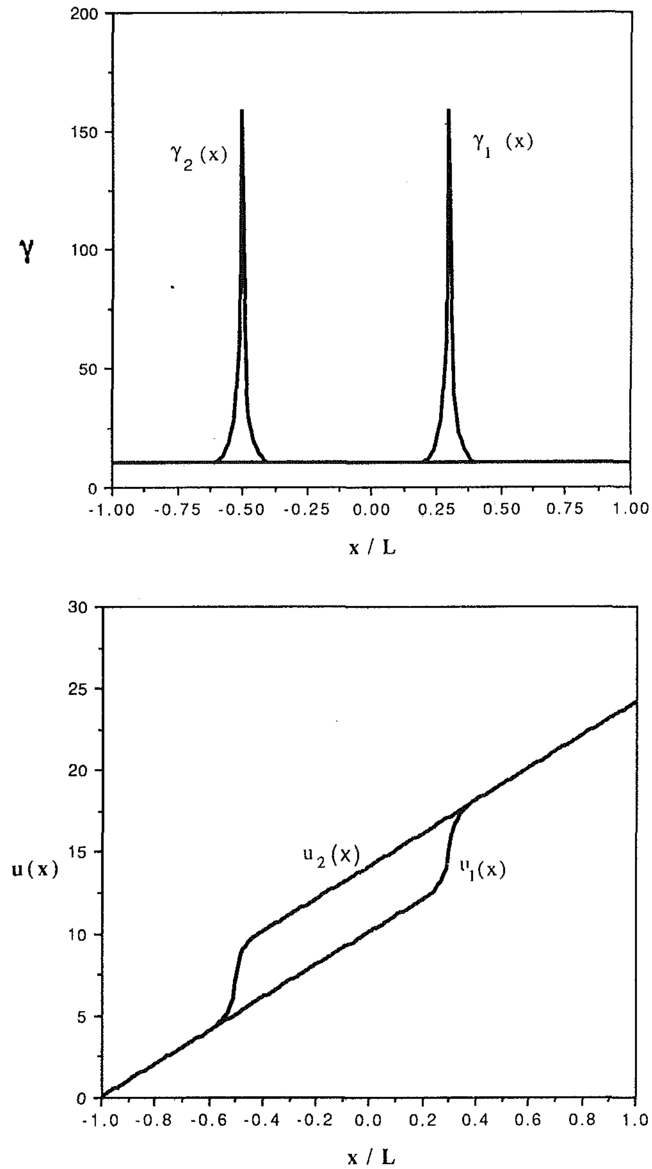


Fig. 4 Strain and displacement fields in response to the imperfections shown in Fig. 3 for a traction boundary condition

It is interesting that even for two perturbations which satisfy equation (22), the displacement fields may vary markedly. Thus, if we consider $\alpha_1(x)$ and $\alpha_2(x)$ as shown in Fig. 3, these perturbations satisfy equation (22) but the solutions, shown in Fig. 4, do not satisfy the stability condition (23). It appears that stability in solutions with these materials is a subtle problem; perhaps once a perturbation is defined, its stability should be examined by considering small perturbations of this initial perturbation, i.e., by considering $\alpha_3(x)$ where

$$\|\alpha_2(x) - \alpha_3(x)\| < \mu^2. \tag{26}$$

Another approach to obtaining a less sensitive measure of stability is to use a response function approach to stability, described subsequently.

A commonly used perturbation in the literature is the step function

$$\alpha(x) = 1.0 - \mu(H(x-b) - H(x+b)). \tag{27}$$

The solution to the above perturbation has the interesting property that even under the restriction (26), it is unstable. To illustrate, we consider the perturbation $\alpha_3(x)$ shown in Fig. 6(*a*). The solutions $u_2(x, \bar{t})$ and $u_3(x, \bar{t})$ do not satisfy (26) (see Fig. 6(*b*)), so $u_2(x, \bar{t})$ is an unstable solution. Step-function imperfections are tacitly used in most finite element solutions. This analysis shows that the resulting solution is not stable or physically meaningful for this constitutive model.

4.1 Stability in Terms of a Response Function. A feature of the viscoplastic constitutive model under consideration here is that localization is triggered at the point at which $\alpha_1(x)$ is a global minimum. As shown in Fig. 4, the shear band occurs at the location of the minimum of $\alpha_2(x)$, which is quite random in a real body. Thus, for an arbitrary, small imperfection, the location of the band may differ markedly. This randomness is in accord with experimental observations; unless notches or other tailored geometric imperfections are used to trigger localization, the location of the shear band is often quite unpredictable. However, the overall response only depends on the occurrence of the band, not its location. Thus, for engineering purposes, it appears worthwhile to consider an alternative definition of stability in terms of a response function, the appropriate choice of which will depend on the problem.

This definition of stability is stated as follows. Given two imperfections $\alpha_1(x)$ and $\alpha_2(x)$ satisfying (22), then if

$$\|r_1(t) - r_2(t)\| \le C_0\mu \tag{28}$$

where $r(t)$ is the response function, the solution is stable. For the particular problem considered here, the response function may be taken to be

$$r(t) = u(L,t),$$

so stability requires that

$$\|r_1(t) - r_2(t)\| = \left\{\int_0^t [u_1(L,t) - u_2(L,t)]^2 dt\right\}^{\frac{1}{2}} < C_0\mu. \tag{29}$$

Here, t may be any suitable measure of time.

Note that neither the homogeneous solution (20) nor the solution with the step-function imperfection (27) meets this stability criterion. On the other hand, the solutions corresponding to the perturbations $\alpha_1(x)$ and $\alpha_2(x)$ as defined in Fig. 3 meet this criterion.

The appropriate response function depends on the problem. For engineering purposes, a solution which is stable in terms of a response function should be an adequate solution. In a problem with a prescribed velocity, the response function would be the resulting load or average stress.

The relevance of this definition of stability for a velocity boundary condition is demonstrated in Fig. 5. It is seen that though the resulting strain fields are such that criterion (23) is violated, these solutions satisfy (28). For a velocity boundary condition, localized deformation is accompanied by a sharp drop in the stress/nominal strain curve, this has been observed experimentally and is in fact used to determine the onset of localization (Marchand and Duffy (1988)). This sharp drop represents a decrease in the load-carrying capacity of the material and maybe construed as failure in any sense of the term. Thus, the response function concept provides a physically meaningful framework to examine localization failure.

5 Imperfections and Scale in Shear Banding

The issue of length scales has particular significance in the shear band problem. A dimensional analysis of the governing equations (1)-(4) reveals two length scales, an inertial length scale $L_I \equiv (1/a_o)\ (G/\rho)^{1/2}$, and a thermal length scale $L_T \equiv (k/\rho C_p a_o)^{1/2}$. Needleman (1988) showed that the inertial length scale does not alleviate the problem of an ever increasing strain gradient. In the isothermal quasi-static case only a boundary condition-dependent length scale $L_B \equiv v_o/a_o$ is present, which only scales the nominal strain rate.

For a step function perturbation, the width of the region of intense straining is set by the width of the imperfection, i.e., the width of the shear band is equal to the width of the imperfection. But, as shown previously, this is an unstable solution which is not physically meaningful. Given the absence of a length scale in viscoplastic materials models for quasi-static loading conditions, it has been of interest as to what determines the width of shear bands. To better understand the effect of C^1 imperfections, the strain field given in equation (20) is expanded in a Fourier series, $c_o + c_m \cos(mx/L)$, where a repeated index represents a summation. The effect of the width of the perturbation on the spectrum is then studied where the width of the imperfection implies the base of the C^1 imperfection. For the case of a parabolic imperfection it was found that the amplitude of the dominant mode, the coefficient c_1, was directly proportional to the width of the initial perturbation. From Fig. 7 it can be seen that the width of the initial imperfection scales the amplitude of the dominant mode of the Fourier spectrum of the strain field.

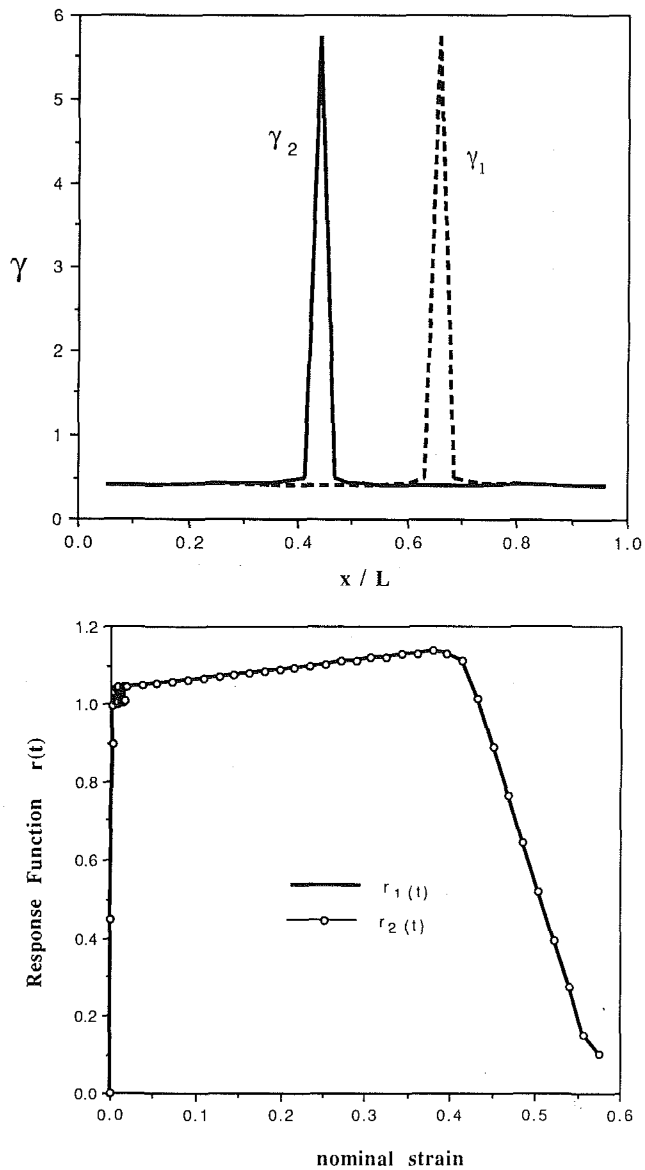


Fig. 5 Strain field and stress/nominal strain curve in response to the imperfection shown in Fig. 3 for a velocity boundary condition

6 The Spatial Distribution of the Strain Field

In this section we characterize the effect of the initial imperfections on the spatial distribution of the strain field. We first focus attention on the traction-controlled boundary condition and present results from the numerical solutions of the velocity boundary condition.

It can be seen from equation (20) that if $\alpha(x)$ is a C^1 function, then $\gamma(x)$ is also a C^1 function. Thus, for $\alpha(x) \in C^1$, the maximum of $\gamma(x, \bar{t})$ occurs where $\frac{\partial\gamma}{\partial x} = 0$. The first and second spatial derivatives of the strain field are given by

$$\frac{\partial\gamma}{\partial x} = -\left(\frac{\xi n \bar{t}}{m}\left(1 + \xi\bar{t}\alpha^{\frac{1}{m}}\right)^{n-1}\alpha^{\frac{1+m}{m}}\right)\frac{\partial\alpha}{\alpha x} \tag{30}$$

$$\frac{\partial^2\gamma}{\partial x^2} = \left(-\frac{\xi n\bar{t}}{m}\right)\beta\left(\frac{\partial^2\alpha}{\partial x^2}\right) + \left(-\frac{\xi n\bar{t}}{m}\right)\left(\frac{\partial\beta}{\partial x}\right)\left(\frac{\partial\alpha}{\partial x}\right) \tag{31}$$

where

$$\beta = \left\{1 + \xi\bar{t}\alpha^{\frac{1}{m}}\right\}^{n-1}\alpha^{-\frac{1+m}{m}}.$$

From the foregoing expressions it is clear that the derivative of the strain field vanishes only when the derivative of the imperfection function vanishes, i.e., $\frac{\partial\gamma}{\partial x} \Leftrightarrow \frac{\partial\alpha}{\partial x} = 0$. If $\alpha(x)$ is a C^0 or C^1 function, then $\gamma(x)$ is a C^0 or C^1 function, respectively, and $\gamma(x, \bar{t})$ is a maximum where $\alpha(x)$ is minimum (see Fig. 8). If $\alpha(x)$ is a C^0 function, with a cusp-shaped maximum at x_0, then $\frac{d\alpha}{dx}(x_0 - \Delta x) < 0$ and $\frac{d\alpha}{dx}(x_0 + \Delta x) > 0$ for $\Delta x > 0$, i.e., the left-hand derivative is negative and the right-hand derivative is positive. Equation (24) then implies that the right spatial derivative of $\gamma(x, \bar{t})$ is positive and the left spatial derivative is negative since the quantities in the parentheses are positive. Thus, the cusp-shaped spatial distribution for $\gamma(x, \bar{t})$ occurs only when $\alpha(x)$ is a C^0 function.

6.1 Behavior of the Strain Field Near the Point of Maximum Strain.

To examine the behavior of the strain field near the point of maximum strain, the strain field is expanded in a Taylor series about the point of maximum strain $\bar{x}$, and the

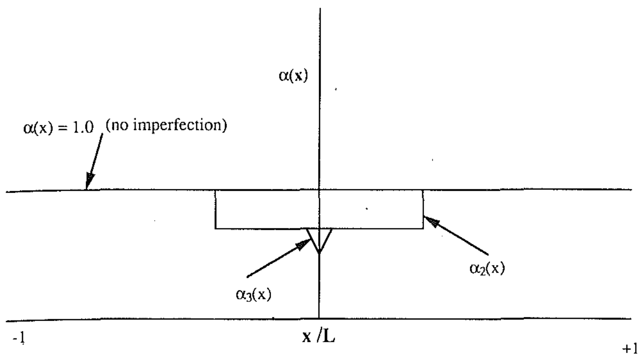


Fig. 6(*a*) Imperfection to illustrate the instability of a step-function perturbation

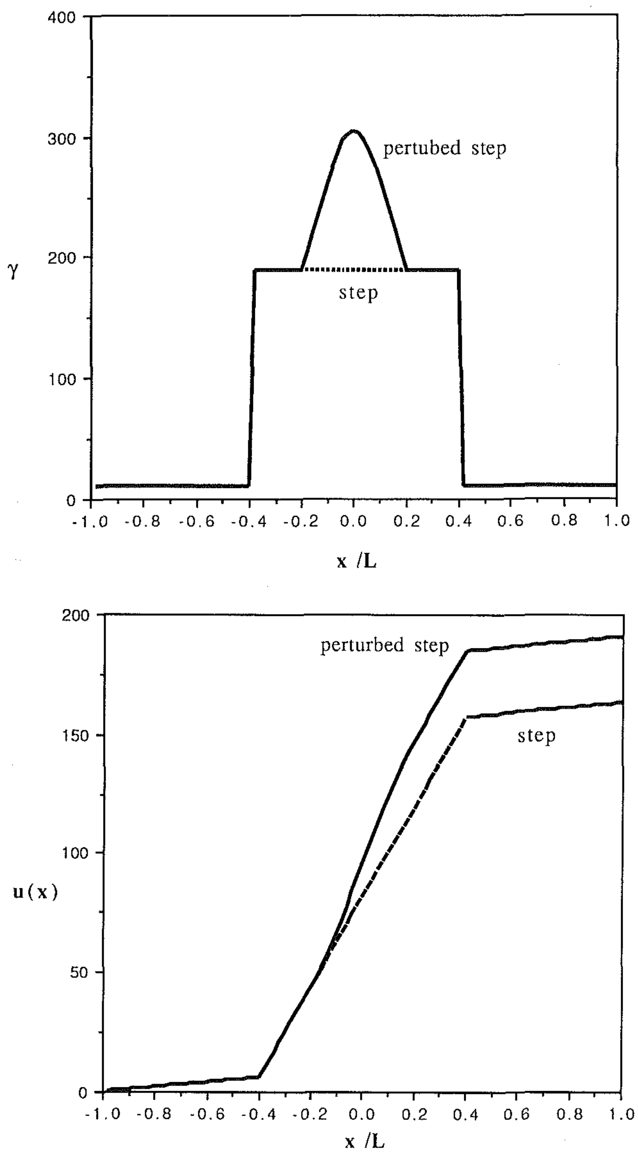


Fig. 6(*b*) Strain and displacement fields in response to the imperfection in Fig. 6(*a*)

behavior of $f(t) = \dfrac{\gamma(\bar{x}, \bar{t})}{\gamma(\bar{x}+\Delta x, \bar{t})}$ is examined:

$$\gamma(\bar{x}+\Delta x, \bar{t}) = \gamma(\bar{x}+\bar{t}) + \frac{1}{2}\frac{\partial \gamma(\bar{x}, \bar{t})}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 \gamma(\bar{x}, \bar{t})}{\partial x^2}\Delta x^2 + \dots \quad (32)$$

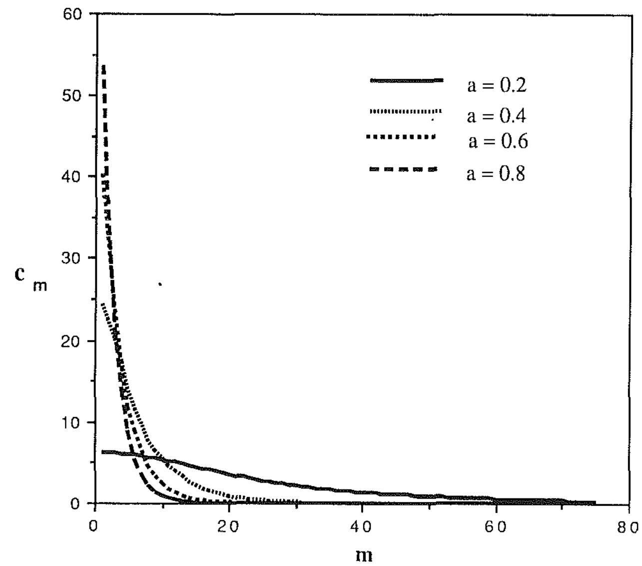


Fig. 7(*a*) Effect of the width of the imperfection on the components of the Fourier series c_m. Here $a = w/2L$, where w is the width of the imperfection.

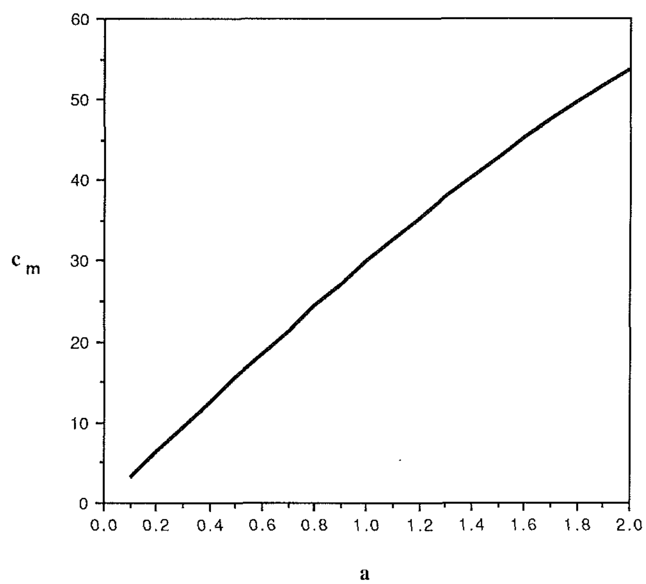


Fig. 7(*b*) Effect of the width of the initial imperfection on the dominant mode ($m = 1$) of the Fourier series

If the imperfection function is assumed to be C^1 then

$$\frac{\partial \alpha(\bar{x})}{\partial x} = 0, \; \frac{\partial \gamma(\bar{x})}{\partial x} = 0, \; \frac{\partial^2 \alpha(\bar{x})}{\partial x^2} > 0,$$

$$\frac{\partial^2 \gamma(\bar{x})}{\partial x^2} = \left(-\frac{n\xi \bar{t}}{m}\right)\beta \frac{\partial^2 \alpha(\bar{x})}{\partial x^2}$$

$$\gamma(\bar{x}+\Delta x, \bar{t}) = \gamma(\bar{x}, \bar{t}) + \frac{1}{2}\left(-\frac{\xi n \bar{t}}{m}\right)\left\{1+\xi \bar{t}\alpha^{-\frac{1}{m}}\right\}^{n-1}\alpha^{-\frac{1+m}{m}}\frac{\partial^2 \alpha(\bar{x})}{\partial x^2}\Delta x^2. \quad (33)$$

Dividing the above expansion for $\gamma(\bar{x} + \Delta x, \bar{t})$ by

$$\gamma(x, \bar{t}) = \left[1+\xi \bar{t}\alpha^{-\frac{1}{m}}\right]^n$$

yields the following

$$\frac{\gamma(\bar{x}+\Delta x,\bar{t})}{\gamma(\bar{x},\bar{t})} 1-\frac{\phi(\Delta x)\,\bar{t}}{\left(1+\xi\,\bar{t}\,\alpha^{-\frac{1}{m}}\right)} \tag{34}$$

where

$$\phi(\Delta x)=\frac{1}{2}\left(\frac{\xi n}{m}\,\alpha(\bar{x})^{-\frac{1+m}{m}}\,\frac{\partial^2\alpha}{\partial x^2}(\bar{x})\right)\Delta x^2. \tag{35a}$$

Let

$$f(\Delta x,\bar{t})=\frac{\gamma(\bar{x},\bar{t})}{\gamma(\bar{x}+\Delta x,\bar{t})}. \tag{35b}$$

It may be noted that $\xi n > 0$, $\frac{\partial^2\alpha(\bar{x})}{\partial x^2} > 0$, $\alpha(\bar{x}) > 0$, $\Delta x^2 > 0$, $\Rightarrow \phi(\Delta x) > 0$. Equation (41) may then be written as

$$\frac{1}{f(\Delta x,\bar{t})}=1-\frac{\phi(\Delta x)\,\bar{t}}{\left(1+\xi\,\bar{t}\,\alpha^{-\frac{1}{m}}\right)} \tag{35c}$$

$$\frac{\partial f}{\partial \bar{t}}=f^2\,\frac{\phi(\Delta x)}{\left(1+\xi\,\bar{t}\,\alpha^{-\frac{1}{m}}\right)^2}. \tag{35d}$$

Irrespective of the sign of ξ, $\frac{\partial f}{\partial \bar{t}} > 0 \Rightarrow f(\Delta x,\bar{t})$ is a monotonically increasing function of time. For $\xi > 0$, $\frac{\partial f}{\partial \bar{t}}$ is monotonically decreasing with time for all Δx. For $\xi < 0$, $\frac{\partial f}{\partial \bar{t}}$ is monotonically increasing with time for all Δx, i.e., $f(\Delta x,\bar{t})$ is a convex upward function of time. Thus, for $\xi < 0$, i.e., $m < p$, the ratio of the strain at the point of maximum strain to the strain at a point very close to it grows at an ever increasing rate and results in an ever narrowing region of intense straining, this feature has been observed experimentally by Marchand and Duffy (1988).

7 Numerical Analyses for a Velocity Boundary Condition

In this section, we consider a viscoplastic slab subjected to velocity boundary conditions (5*c*) and (5*d*). The problem geometry is defined in Fig. 1. The results are obtained by one-dimensional constant strain finite elements. The slab is modeled as elasto-viscoplastic. The viscoplastic function is chosen to be the power-law model described by equation (7). The yield function is chosen to be bilinear. The following dimensionless parameters were used in the analysis:

$$\sigma_0/G=0.002$$

$$v_o/(L\,a_o)=100.$$

The value of m in equation (7) was taken to be 0.01, a typical value for structural steels. The response of the slab to C^0, imperfections was studied and the resulting strain profile is shown in Fig. 9. The imperfections are modeled by assuming a variation in the yield strength in the appropriate elements. It is seen that for C^0 imperfections, the strain fields have a maximum here the imperfection function has a global minimum.

The different boundary conditions lead to some differences, the most noteworthy of these are:

(1) for the velocity boundary condition the strain field is always bounded for finite end displacement and localization in the L_∞ sense of Molinari and Clifton (1987) is accompanied by a sharp drop in the stress/nominal strain curve;

(2) for the traction boundary condition, localization in the L_∞ sense occurs at *finite* end displacement.

However, there are some features that are common to both situations:

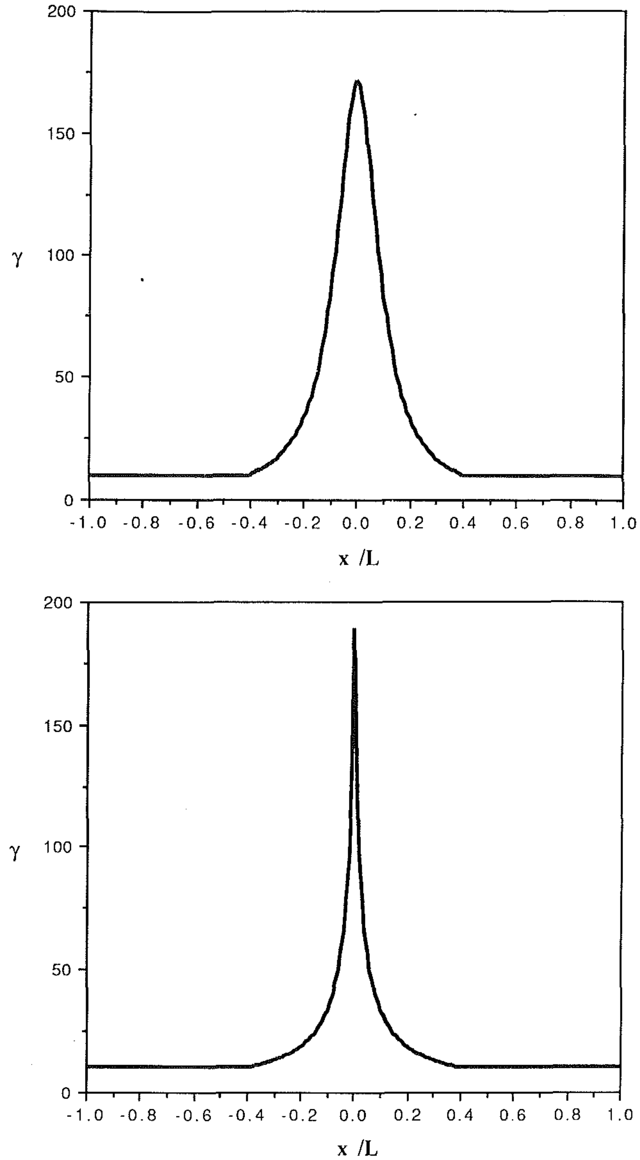


Fig. 8 Strain field response to C^0 and C^1 imperfections, respectively, for a traction boundary condition

(1) The strain fields in both cases are almost entirely determined by the initial imperfections. In particular:

(a) the relative maximum of the strain field occurs where the yield strength is a global minimum;

(b) wherever the imperfection function has a constant magnitude, the strain field is constant; this further implies that if the imperfection function is a step function, then the strain field also contains a step function (see Fig. 8).

(2) The response of the material to initial imperfections is initially quasi-homogeneous in the softening regime followed by localization of the deformation in a narrow region.

8 Conclusions

It has been shown that the scale or size of the shear band in a viscoplastic material is governed by the scale of the imperfection. A closed-form solution for a traction-controlled quasi-static shear banding in a viscoplastic, one-dimensional problem and numerical results for a velocity-controlled boundary condition have been presented. The solutions exhibit a strong dependence on the morphology of the initial imperfection. In particular:

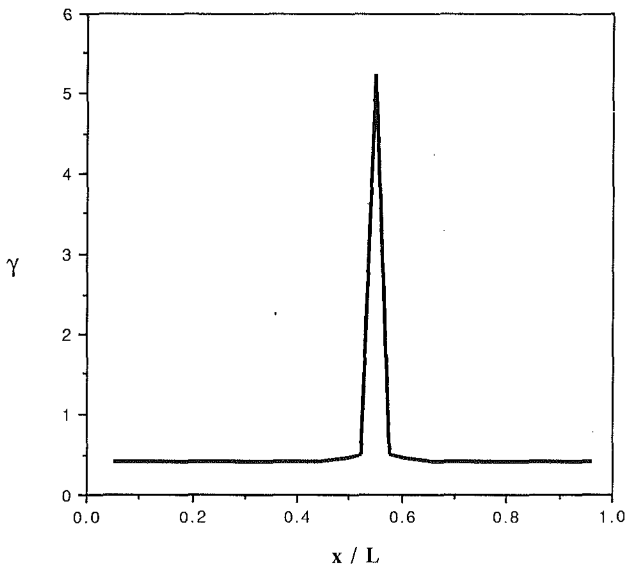


Fig. 9 Strain field for C^0 imperfections for a velocity boundary condition

(1) in a homogeneous stress field, the shear band occurs at an extremum of the imperfection, so the position of the shear band is quite random;

(2) the continuity of the strain field is identical to the continuity of the imperfection, i.e., for a C^1 imperfection, the shear strain field is C^1.

(3) when the imperfection is C^{-1}, i.e., a gate function, the strain field is also C^{-1}.

It has been shown that the solution with a gate function imperfection is unstable. This finding is of relevance in finite element solutions with constant strain elements, where the imperfections are gate functions, when constant strain elements are used, and the imperfection is a perturbation of one element. Furthermore, when the imperfection is a gate function, the morphology of the strain field differs significantly from that which arises from more realistic imperfections. Thus, solutions based on step function imperfections cannot reveal the details of the strain field which are needed to understand material behavior at high strains.

The definition of a stable solution for a viscoplastic material presents an interesting problem. Even two imperfections which are close to each other in the sense of a norm can lead to strikingly different solutions. Therefore, a definition of stability in terms of a response function has been proposed. This definition is useful for engineering purposes since it overcomes difficulties which arise from random locations of shear bands.

It has also been shown that the ratio of the strain at the point very close to it is ever increasing resulting in an ever narrowing shear band. This feature has also been observed experimentally:

Acknowledgment

The support of the Office of Naval Research under Award N00014-89-J-3066 to Northwestern University is gratefully acknowledged.

References

Bazant, Z. P., and Belytschko, T., 1985, "Wave Propagation In a Strain Softening Bar: Exact Solution," *J. Engrg. Mech.*, Vol. 111, pp. 381-389.

Duffy, J., 1984, "Mechanics of Material Behavior," *The Daniel C. Drucker Anniversary Volume*, G. J. Dvorak and R. T. Shield, eds. Elsevier, New York, p. 75.

Lasry, D., and Belytschko, T., 1988, "Localization Limiters in Transient Problems," *Int. J. Solids and Structures*, Vol. 24, pp. 581-597.

Marchand, A., and Duffy, J., 1988, "An Experimental Study of the Formation Process of Adiabatic Shear Bands in a Structural Steel," *J. Mech. Phys. Solids*, Vol. 36, pp. 251-283.

Molinari, A., and Clifton, R. J., 1987, "Analytical Characterization of Shear Localization in Thermoviscoplastic Materials," *J. App. Mech.*, Vol. 54, pp. 806-812.

Needleman, A., 1988, "Material Rate Dependence and Mesh Sensitivity in Localization Problems," *Comp. Meth. App. Mech. Engrg.*, Vol. 67, pp. 61-85.

Pan, J., 1983, "Perturbation Analysis of Shear Strain Localization in Rate Sensitive Materials," *Int. J. Solids and Structures*, Vol. 19, pp. 153-164.

Shawki, T. G., and Clifton, R. J., 1989, "Shear Band Formation in Thermal Viscoplastic Materials," *Mechanics of Materials*, Vol. 8, pp. 13-43.

Tzavaras, A. E., 1986, "Plastic Shearing of Materials Exhibiting Strain Hardening or Strain Softening," *Arch. Rat. Mech. Anal.*, Vol. 94, p. 39.

Wright, T. W., and Batra, R. C., 1985, "The Initiation and Growth of Adiabatic Shear Bands," *Int. J. Plasticity*, Vol. 1, pp. 205-212.

Wright, T. W., and Walter, J. W., 1987, 'On Stress Collapse in Adiabatic Shear Bands," *J. Mech. Phys. Solids*, Vol. 35, pp. 205-212.

Wu, F. H., and Freund, L. B., 1988, "Deformation Trapping due to Thermoplastic Instability in One-Dimensional Wave Propagation," *J. Mech. Phys. Solids*, Vol. 32, pp. 119-132.

J. J. Kalker
Professor,
Faculty of Technical Mathematics and Informatics,
Delft University of Technology,
Delft, The Netherlands

Viscoelastic Multilayered Cylinders Rolling With Dry Friction

Two circular cylinders consisting of a rigid core which is covered by an arbitrary number of homogeneous, isotropic, viscoelastic coats of arbitrary, but uniform thickness are pressed together so that a contact area in the form of a strip forms between them, and subsequently rolled in the presence of dry friction. A Maxwell model of viscoelasticity is employed; the friction is finite and modeled by Coulomb's law; partial slip in the contact area is allowed. It is required to find the viscoelastic field in the cylinder, notably in the contact strip, when the compressive force and the creepage in rolling direction are specified. The proposed method works almost equally fast in the case of pure elasticity and of viscoelasticity. It is akin to the method of Bentall et al. (1968), but automated, modernized, and extended.

Introduction

Consider two infinite rigid cylinders with parallel axes which are covered with a number of homogeneous, isotropic, linearly elastic or viscoelastic layers that are completely bonded to each other and to the cylinder they cover. They are pressed together and subsequently rolled over each other until a steady state sets in. The circumferential velocities of the cylinder-cum-layers systems may differ, so that partial or complete slip occurs in the interface. Friction is present in the interface; it is assumed to behave according to Coulomb's law with a constant friction coefficient.

It is required to find the displacement and the stress in the layers with respect to the Eulerian coordinate system that is attached to the axes of both cylinders and which is, consequently, fixed to the contact area. In particular, one is interested in the displacements and loads present in the contact area.

The analysis is linear and two-dimensional. For the calculation of the (visco)elastic field the cylinders are approximated by layered (visco)elastic half spaces in contact. The surface load is approximated by a function which is constant in adjoining, equally long intervals with the aid of algorithms that have proved their utility before in the programs CONTACT and LAAGROL and that have been established rigorously by Kalker (1988) in the elastic case.

In order to use these algorithms the (visco)elastic field due to a typical normal surface load, and that due to a typical shearing surface load, are required. To find these fields, the displacements and stresses are expressed in Airy's stress function which obeys a two-dimensional bi-potential PDE. This equation is attacked by applying a complex Fourier transform in the tangential direction and by analytically solving the resulting fourth-order ODE in the normal coordinate. The solution, a Fourier transform, is inverted numerically by a method which guarantees a prescribed accuracy.

The method presented is akin to that of Bentall and Johnson (1968) but extended, automated, and modernized.

1 Elasticity Theory

The cylinders are numbered 1 and 2 (see Fig. 1). We introduce a Cartesian coordinate system with the plane $z = 0$ in the mutual tangent plane of the cylinders, the y-axis in the axial direction, and the x-axis in the tangential direction; the z-axis points into body 1. We denote the normal stresses σ_x, σ_y, σ_z, and the shear stresses as τ_{xy}, τ_{yz}, τ_{zx}. All quantities (displacements, stresses, strains) can be provided with a subscript $\alpha =$ 1, 2, signifying cylinder α. The displacements are $u(u, v, w)$; the linearized strain e_x, e_y, e_z, e_{xy}, e_{yz}, e_{zx}. We assume a two-dimensional situation in which the dependence on y of all quantities disappears and in which the y-component of the displacement $v = 0$ (plane strain),

$$\frac{\partial}{\partial y} = 0,\ v = 0. \tag{1}$$

An Airy function approach yields (H is the Airy function):

$$u = \frac{1-\nu^2}{E} H_{,zzz} - \frac{\nu(1+\nu)}{E} H_{,xxz} \tag{2}$$

$$w = \frac{1-\nu^2}{E} H_{,xxx} - \frac{\nu(1+\nu)}{E} H_{,xzz} \tag{3}$$

$$\sigma_x = H_{,zzzx},\ \sigma_z = H_{,xxxz},$$

$$\tau_{xz} = -H_{,xxzz},\ \sigma_y = \nu(\sigma_x + \sigma_z) \tag{4}$$

$$H_{,zzzz} + 2H_{,xxzz} + H_{,xxxx} = 0. \tag{5}$$

$$_{,x} = \partial/\partial x,\ _{,z} = \partial/\partial z \tag{6}$$

E: Young's modulus, ν: Poisson's ratio (7)

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 8, 1989; final revision, Mar. 27, 1990.

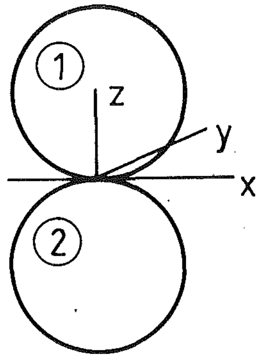


Fig. 1 Two cylinders in contact

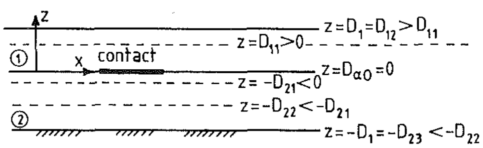


Fig. 2 The layers and the cylinders

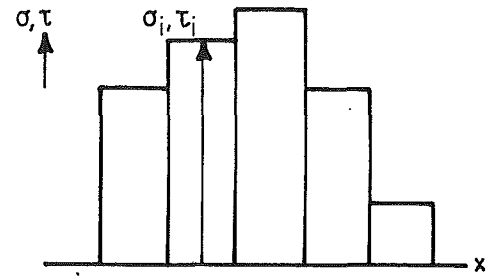


Fig. 3 Piecewise constant traction

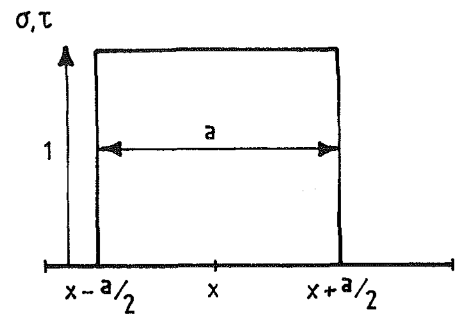


Fig. 4 An element

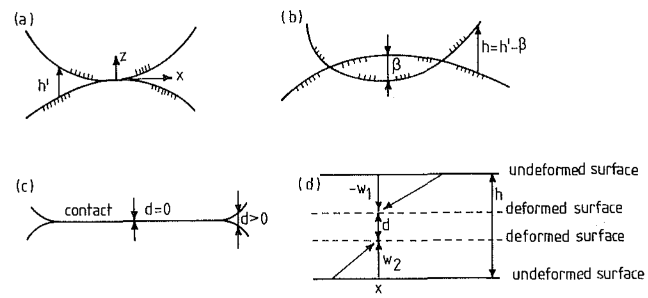


Fig. 5 The cylinders touch in Fig. 5(*a*). The layers are not shown. They approach each other over a certain distance in Fig. 5(*b*). In Fig. 5(*c*) a deformation occurs, which cancels the penetration in Fig. 5(*b*). In Fig. 5(*d*) the construction of the deformed distance is shown.

We note that equations (2)-(7) hold for any two-dimensional, homogeneous isotropic elastic body. When we consider a layered medium with each layer homogeneous and isotropic, equations (2)-(7) hold for each layer separately.

1.1 A Fundamental Boundary Value Problem. We assume that both layer and cylinder can be taken as plane as far as elasticity calculations are concerned. For the boundary values we retain the real geometry. The situation is shown in Fig. 2. The layers consist of several sublayers; the interface between sublayer i and sublayer $i+1$ has the equation

$$z=(-1)^{\alpha-1}D_{\alpha i},\ \alpha=1,2;\ i=0,\ldots,m_{\alpha i},\ D_{\alpha i}\in I\!R, \tag{8}$$

$$m_{\alpha i}:\ \text{number of sublayers of layer } \alpha. \tag{9}$$

$$D_{\alpha 0}=0<D_{\alpha 1}<D_{\alpha 2}\cdots<D_{\alpha m_\alpha}\underset{=}{def}\,D_\alpha. \tag{10}$$

The situation is shown in Fig. 2.

The sublayers are completely bonded together, so that we have that

$$u,\ w,\ \sigma_z,\ \tau_{xz} \tag{11}$$

are continuous across the interface of two sublayers. The subscript α is omitted as long as this causes no confusion. The layers are rigidly bonded to the cylinder. So

$$u(x,(-1)^{\alpha-1}D_\alpha)\ =\ w(x,(-1)^{\alpha-1}D_\alpha)=0. \tag{12}$$

Moreover, it lies at hand to suppose that u, w, σ_x, σ_z, $\tau_{xz}\rightarrow 0$, far from the contact area. Therefore, we assume

$$H\rightarrow 0 \text{ with all derivatives if } |x|\rightarrow\infty. \tag{13}$$

Finally, we assume that at the surface $z=0$,

$$\sigma_x(x,0)\ \underset{=}{def}\sigma(x)=\ \text{prescribed} \tag{14}$$

$$\tau_{xz}(x,0)\ \underset{=}{def}\tau(x)=\ \text{prescribed}. \tag{15}$$

We shall use solutions of this type to construct the more complicated solutions we need. In particular, we consider the case when $\sigma(x)$ and $\tau(x)$ are piecewise constant, (see Fig. 3). These traction distributions can be considered as built from elements, see Fig. 4. We are interested in the influence numbers, i.e., the displacements in y due to the element with height 1, width a, and center x. The traction distribution, hence the displacement, is determined by the numbers σ_i, τ_i, heights of the ith element.

2 Contact Formation

Consider two cylinders (see Fig. 5). The boundaries between rubber and steel are not shown.

In Fig. 5(*a*) the cylinders touch; their vertical distance at the position x is $h'(x)$. $h'(x)$ is given by

$$h'(x)=\frac{1}{2}\left(\frac{1}{R_1}+\frac{1}{R_2}\right)x^2, \tag{16}$$

R_α:radius of cylinder α.

Subsequently the cylinders are pressed together, causing their centers to approach each other over a distance β. If we omit the elastic deformation, their distance is now

$$h(x)\ =\ h'(x)-\beta\ =\ \left(\frac{1}{2R_1}+\frac{1}{2R_2}\right)x^2-\beta. \tag{17}$$

Now the bodies overlap (see Fig. 5(*b*)); this is cancelled by the elastic deformation shown in Fig. 5(*d*), and the bodies look as shown in Fig. 5(*c*). It holds that the distance after deformation is $d(x)$, and from Fig. 5(*d*) we can tell

$$d(x)\ =\ h(x)\ -\ w_2(x)\ +\ w_1(x)\ \geq\ 0$$

$$\text{as the bodies cannot overlap} \tag{18}$$

$d(x)=0$: contact.

We assume that the bodies cannot exert tractive forces on each other, or $\sigma(x) \le 0$; and that tractions can occur in the contact only, or $\sigma(x)d(x) = 0$, while the normal force is continuous, in the sense that $\sigma_1(x) = \sigma_2(x) + \sigma(x)$. Summarizing, we obtain

$$\left.\begin{aligned} &h(x)=\left(\frac{1}{2R_1}+\frac{1}{2R_2}\right)x^2-\beta,\\ &d(x)=h(x)-w_2(x)+w_1(x)\ge 0,\ \sigma_1(x)=\sigma_2(x)=\sigma(x)\le 0,\ \sigma(x)d(x)=0. \end{aligned}\right\} \quad (19)$$

We assume an element with length a; and we want to know the traction/displacement in the interval $(-na, na)$ of the x-axis. This is called the potential contact; n can be chosen freely as long as the entire contact area is within the potential contact. We cover this interval with the elements of Figs. 3 and 4. These elements are numbered $i = 1, n$ and they are determined by the tractions in their centers, σ_i, and $\tau_{\alpha i}$. Their centers are in $x_i = \left(i-\frac{n-1}{2}\right) a$. $w_{\alpha i}$ is also sampled in these points; it holds the following connection between $w_{\alpha i}$ and the tractions σ_j and the tractions $\tau_{\alpha j}$, the latter of which we consider given. Then we want to solve the following problem. Find all σ_j:

$$\sigma_j \le 0;\ d_j = d(x_j) \ge 0;\ \sigma_j d_j = 0;\ \tau_j \text{ given.} \quad (20)$$

The following algorithm solves the so-called complementarity system (20), if approach β is given:

Algorithm N, Kalker (1983, 1988).

For all i *within the potential contact:*

Step 0 Suppose $\sigma_i = 0$, $1 \le i \le 2n + 1$; we assume τ_i given.

Step 1 If $d_i \le 0$, then i is placed in index set (= set of indices) $\boldsymbol{K}$. If $d_i > 0$, then i is placed in index set $\boldsymbol{B}$ ($\boldsymbol{K}$: contact, $\boldsymbol{B}$: surface of the half space outside contact ("exterior")).

Step 2 We set $d_i = 0$ if i is in $\boldsymbol{K}$; we set $\sigma_i = 0$ if i is in $\boldsymbol{B}$. These are $2n + 1$ linear equations for the $2n + 1$ unknowns σ_i; solve them.

Step 3 If i lies in $\boldsymbol{K}$ and the just-found $\sigma_i \le 0$, i will remain in $\boldsymbol{K}$. If i lies in $\boldsymbol{K}$ and the just-found $\sigma_i > 0$, one such i will go to $\boldsymbol{B}$. If i lies in $\boldsymbol{B}$, i will remain in $\boldsymbol{B}$.

Step 4 If $\boldsymbol{K}$ is changed in Step 3, then go to Step 2.

Step 5 Now $\sigma_i \le 0$ in $\boldsymbol{K}$; $\sigma_i = 0$ in $\boldsymbol{B}$, $d_i = 0$ in $\boldsymbol{K}$. We verify whether $d_i \ge 0$ in $\boldsymbol{B}$. If i is in $\boldsymbol{B}$ and $d_i < 0$, one such i will go to $\boldsymbol{K}$; otherwise $\boldsymbol{K}$ and $\boldsymbol{B}$ remain unchanged.

Step 6 If $\boldsymbol{K}$ is changed in Step 5, then go to Step 2.

Step 7 Now $\sigma_i \le 0$ in $\boldsymbol{K}$, $\sigma_i = 0$ in $\boldsymbol{B}$, $d_i = 0$ in $\boldsymbol{K}$, $d_i \ge 0$ in $\boldsymbol{B}$: We are ready.

Remark. This algorithm can be proved rigorously to converge towards the unique solution of equation (20) in a finite number of steps (see Kalker, 1983, 1988). The number of iterations, though finite, may be large; actually it will be of the order of $4n$, n: the number of elements in the contact area.

Two methods may be used to accelerate the process:

(1) Only one element of $\boldsymbol{K}$ goes to $\boldsymbol{B}$, and vice versa. Hence, the system of linear equations of Step 2 changes only little, so little that the solution may be updated by a simple update formula which requires only $O(n^2)$ elementary operations.

(2) Alternatively one may modify $\boldsymbol{K}$ and $\boldsymbol{B}$ in Step 3 by allowing *all i* to go to $\boldsymbol{B}$ for which the just-found $\sigma_i > 0$; and by allowing *all i* to go to $\boldsymbol{K}$ in Step 5 for which $d_i < 0$.

The first alternative is attractive in that the proof of the algorithm remains unchanged. The second alternative is simpler, but our proof of the algorithm breaks down. Yet the algorithm is so robust that even with the modification it has never failed so far.

This, added to the fact that the second alternative allows easy generalization to the three-dimensional case (Kalker, 1988) leads us to adopt the second alternative. If instead of β the total force $Q_z = \int_{-\infty}^{\infty} (\sigma(x))\, dx$ is given, we use the discretized version $Q_z = -\sum_i a\,\sigma_i$ as an auxiliary condition. The approach β will be considered as Lagrange multiplier (an unknown) of this auxiliary condition, and if we define

$$\left.\begin{aligned} &w_1(x_i) - w_2(x_i) = \sum_j (A_{ij}\sigma_j + B_{ij}\tau_j)\\ &(A_{ij},\ B_{ij}\text{: influence numbers})\\ &\text{then the tableau is:}\\ &\sum_j A_{ij}\sigma_j - \beta = -h_{ti}\ (i \text{ and } j \text{ in } \boldsymbol{K})\\ &-\sum_j a\,\sigma_j = N\\ &\text{with } h_{ti} = h'(x_i) - \sum_j B_{ij}\tau_j. \end{aligned}\right\} \quad (21)$$

The algorithm runs as above, with β as an extra variable that can take any value.

3 Friction

Let Vv be the slip of body 2 over body 1 when 1 and 2 are considered rigid. In the time interval $(T - t, T)$ the relative displacement will be Vvt. As a result of the elastic deformation of the layers on the cylinders, the particle that was in x at the time T has undergone a tangential displacement of $u(x)$ with respect to the cylinder. At the time $T - t$ the particle was in $x + Vt$ (see Fig. 6). The tangential displacement of the particle with respect to the cylinder at the time $T - t$ is $u(x + Vt)$. The net tangential displacement of the particle with respect to the cylinder in the time interval $(T-t, T)$ will be $u(x) - u(x + Vt)$.

Therefore, the total tangential displacement of the cylinder (2) with respect to (1) is:

$$v_1(x) = Vvt + (u_1(x+Vt) - u_2(x+Vt)) - (u_1(x) - u_2(x)). \quad (22)$$

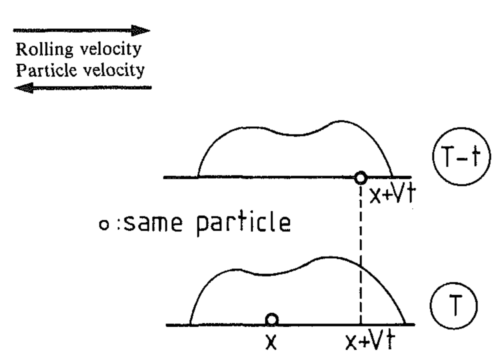


Fig. 6 The movement of a particle through the contact plane

If we choose $Vt = a$, where a is the width of an element, then we get

$$v_i = av + u_{i+1} - u_i;\ Vt = a \quad (u = u_1 - u_2). \tag{23}$$

With the aid of the influence numbers v_j is expressed linearly in the σ_i and the τ_i. Concerning the forces we observe that the stress is continuous across the interface of the bodies, so that

$$\tau_1(x) = \tau_2(x),\ \sigma_1(x) = \sigma_2(x). \tag{24}$$

We assume that the friction can be described by the law of Coulomb-d'Amontons, with a friction coefficient f. Discretized, this reads:

$$|\tau_i| \le -f\sigma_i\ (\sigma_i < 0).$$

If the slip at element i, $v_i \neq 0$ then we have

$$\tau_i = f\sigma_i \operatorname{sign}(v_i) \text{ in } \boldsymbol{K}. \tag{25}$$

The problem is now solved by the following algorithm.

Algorithm F (Kalker (1983, 1988)).

Step 0 Initiate with $\tau_i = 0$. We work only within the contact area $\boldsymbol{K}$, i in $\boldsymbol{K}$, σ_i is given.

Step 1 If $|\tau_i| > -f\sigma_i$ ($\sigma_i < 0_i$), then i is placed in $\boldsymbol{S}$ (index set: area of slip). If $|\tau_i| \le -f\sigma_i$, then i is placed in $\boldsymbol{A}$ (index set: area of adhesion).

Step 2 If i is in $\boldsymbol{S}$, τ_i is set equal to $-f\sigma_i \operatorname{sign}(\tau_i)$ which means that $|\tau_i|$ is reduced to $-f\sigma_i$, while τ_i retains its sign. If i is in $\boldsymbol{A}$ then v_i is set equal to 0. These are linear equations. Solve them.

Step 3 If i is in $\boldsymbol{A}$, as well as the just-found $|\tau_i| > -f\sigma_i$, then i is placed in the area of slip. No further mutations; the area of adhesion can only decrease.

Step 4 If $\boldsymbol{A}$ is changed in Step 3, then go to Step 2.

Step 5 Now it holds in $\boldsymbol{A}$: $|\tau_i| \le -f\sigma_i$ and $v_i = 0$, and in $\boldsymbol{S}$: $|\tau_i| = -f\sigma_i$. If i is in $\boldsymbol{S}$ and τ_i and v_i have the wrong sign with respect to each other, see equation (25), then i is placed in the area of adhesion $\boldsymbol{A}$, see equation (25).

Step 6 If $\boldsymbol{S}$ is changed in Step 5, then go to Step 2.

Step 7 Now it holds in $\boldsymbol{A}$: $|\tau_i| \le -f\sigma_i$, and $v_i = 0$: area of adhesion, and in $\boldsymbol{S}$: $\tau_i = f\sigma_i\, v_i\, /\, |v_i|$: area of slip. We are ready.

Remark. When in step 3 *only one i* with $|\tau_i| > -f\sigma_i$ is allowed to pass from the area of adhesion $\boldsymbol{A}$ to the area of slip $\boldsymbol{S}$, and when in Step 5 *only one i* for which τ_i and v_i have the same sign is allowed from $\boldsymbol{S}$ to $\boldsymbol{A}$, the algorithm F can be proved rigorously, in the same manner as N. Then, as in N (see the Remark after it) there are the same two alternatives by which F may be accelerated. We have adopted the second alternative, which has actually been given in our presentation of F. As in N, no failures have been observed.

3.1 The Panagiotopoulos Process. Because the algorithm F assuming a certain σ_i will change the τ_i with respect to their original values, the algorithm N, which uses these values as data, will change the σ_i as well, thus causing a discrepancy.

We solve this by repeating the algorithms N and F: $NF\,NF\,NF$. . . until a convergence of σ_i and τ_i occurs. This process is called the Panagiotopoulos (1975) process. If it is performed once: NF, it is called the *Johnson process* (Bentall et al., 1967). There is no guarantee that the Panagiotopoulos process converges, and if it converges, whether the solution found makes sense. In my examples, treated by the LAAGROL program, no complication occurred, and a most convincing convergence was reached after $4 \times NF$.

4 Elasticity and Viscoelasticity

In Appendix E it is shown that for a certain class of viscoelastic materials,

$$e^f_{ij}(x_h,r) = \frac{1+\nu}{E(r)}\sigma^f_{ij} - \frac{\nu}{E(r)}\sigma^f_{hh}\delta_{ij} \tag{E14}$$

with $E(r)$ as in Appendix E, equation (E15). The superscript f indicates a complex Fourier transform with respect to time and parameter r (see Appendix E, equation (E7, sqq.). It is shown in Appendix E that e^f_{ij}, σ^f_{ij}, u^f_i form a purely elastic field for any value of the parameter r. Hence, we may introduce the Airy stress function H as in equations (2)–(7) (we omit the parameter and the superscript f):

$$u = \frac{1-\nu^2}{E}H_{,zzz} - \frac{\nu(\nu+1)}{E}H_{,xxz} \tag{26a}$$

$$w = \frac{1-\nu^2}{E}H_{,xxx} - \frac{\nu(\nu+1)}{E}H_{,xzz} \tag{26b}$$

$$H_{,zzzz} + 2H_{,xxzz} + H_{,xxxx} = 0,\ H \to 0 \text{ if } |x| \to \infty \tag{26c}$$

$$\sigma_z = H_{,xxxz},\ \tau_{xz} = -H_{,xxzz}, \text{ given at } z = 0 \tag{26d}$$

$$(u,w) = (0,0) \text{ if } z = (-1)^{\alpha-1}D_\alpha. \tag{27}$$

Equation (E15a) specializes ν to be constant, and E to

$$E(r) = (1 - jqr)/(K - jqQr) \tag{E15a}$$

with K, Q, q viscoelastic constants which are interpreted in equation (E15). When $K = Q$, one regains elasticity with $E = 1/K$. Consequently, equations (26a,b) become

$$(1 - jqr)u^f = (K - jqQr)\{(1-\nu^2)H^f_{,zzz} - \nu(1+\nu)H^f_{,xxz}\}. \tag{28a}$$

$$(1 - jqr)w^f = (K - jqQr)\{(1-\nu^2)H^f_{,xxx} - \nu(1+\nu)H^f_{,xzz}\}. \tag{28b}$$

We transform back, (cf. equation (E16)):

$$(1 + q\ d/dt)u = (K + qQd/dt)\{(1-\nu^2)H_{,zzz} - \nu(1+\nu)H_{,xxz}\} \tag{29a}$$

$$(1 + q\ d/dt)(w = (K + qQd/dt)\{(1-\nu^2)H_{,xxx} - \nu(1+\nu)H_{,xzz}\}. \tag{29b}$$

In steady-state rolling in the positive x-direction with velocity $V > 0$, $d/dt = -V\partial/\partial x$, hence

$$(1 - qV\,\partial/\partial x)u = (K - qQV\,\partial/\partial x)\ \{(1-\nu^2)H_{,zzz} - \nu(1+\nu)H_{,xxz}\} \tag{30a}$$

$$(1 - qV\,\partial/\partial x)w = (K - qQV\,\partial/\partial x)\ \{(1-\nu^2)H_{,xxx} - \nu(1+\nu)H_{,xzz}\}. \tag{30b}$$

We introduce a complex Fourier transform with respect to x, with k as parameter and a hat ($\hat{}$) as transform indicator. The complex Fourier transform with respect to time is described in Appendix E, Section E4; the transform with respect to position is analogous. We obtain in a multilayer

$$\frac{E_i}{1+\nu_i}\hat{u}(k,z) = (1-\nu_i)\hat{H}_{,zzz} + \nu_i k^2 \hat{H}_{,z}, \tag{31}$$

$i = 1. . . ,m_\alpha$, m_α: number of layers; $E_i = \dfrac{1 + jq_iVk}{K_i + jq_iQ_iVk}$ (32)

$$\frac{E_i}{1+\nu_i}\hat{w}(k,z) = jk^3(1-\nu_i)\hat{H} + jk\nu_i\hat{H}_{,zz} \tag{33}$$

$$\hat{\sigma}_z = jk^3\hat{H}_{,z},\ \hat{\tau}_{xz} = k^2\hat{H}_{,zz} \tag{34}$$

with

$$\hat{H}_{,zzzz} - 2k^2\hat{H}_{,zz} + k^4\hat{H} = 0. \tag{35}$$

We confine ourselves to body 1; the layer of body 2 is treated similarly.

We solve the differential equations (35); they are ODE in z; for $\hat{H}$. We find in layer L_i

$$L_i = \{(x,z) \mid D_{i-1} \le z \le D_i\},\ i=1,\ldots,m,$$
$$D_0 = 0,\ D_m = D \tag{36}$$

that

$$\hat{H}(k,z) = \{A_i(k) + zB_i(k)\}e^{k} + \{C_i(k) + zG_i(k)\}e^{-kz}, \tag{37}$$

A_i, B_i, C_i, G_i: integration constants depending on k.

Consequently, the transformed field quantities become, in the transformed layer $\hat{L}_i$,

$$\hat{L}_i = \{(k,z) \mid D_{i-1} \le z \le D_i\},\ i=1,\ldots,m \tag{38}$$

$$\hat{u}_i(k,z) = \frac{1+\nu_i}{E_i}[\{k^3(A_i + zB_i) + (3-2\nu_i)k^2B_i\}e^{kz} + \{-k^3(C_i+zG_i) + (3-2\nu_i)k^2G_i\}e^{-kz}] \tag{39a}$$

$$\hat{w}_i(k,z) = \frac{j(1+\nu_i)}{E_i)}[\{k^3(A_i+zB_i) + 2\nu_i k^2B_i\}e^{kz} + \{k^3(C_i+zG_i - 2\nu_i k^2 G_i\}e^{-kz}] \tag{39b}$$

$$\hat{\sigma}_{zi}(k,z) = jk[\{k^3(A_i+zB_i) + k^2B_i\}e^{kz} + \{-k^3(C_i+zG_i) + k^2G_i\}e^{-kz}] \tag{39c}$$

$$\hat{\tau}_{xzi}(k,z) = k[\{k^3(A_i+zB_i) + 2k^2B_i\}e^{kz} + \{k^3(C_i+zG_i) - 2k^2G_i\}e^{-kz}] \tag{39d}$$

$$\hat{\sigma}_{xi}(k,z) = -jk[\{k^3(A_i+zB_i) + 3k^2B_i\}e^{kz} + \{-k^3(C_i+zG_i) + 3k^2G_i\}e^{-kz}]\ (k,z\epsilon\hat{L}_i. \tag{39e}$$

We write

$$A_i' = k^3A_i,\ B_i' = k^2B_i,\ C_i' = k^3C_i,\ G_i' = k^2G_i \tag{40}$$

and omit the primes again. In terms of the new definition (40) of A_i, B_i, C_i, G_i, we obtain

$$\hat{u}_i(k,z) = \frac{1+\nu_i}{E_i}[\{A_i + (3-2\nu_i+kz)B_i\}e^{kz} + \{-C_i + (3-2\nu_i-kz)G_i\}e^{-kz}] \tag{41a}$$

$$\hat{w}_i(k,z) = \frac{j(1+\nu_i)}{E_i}[\{A_i + (2\nu_i+kz)B_i\}e^{kz} + \{C_i + (kz-2\nu_i)G_i\}e^{-kz}] \tag{41b}$$

$$\hat{\sigma}_{zi}(k,z) = jk[\{A_i + (1+kz)B_i\}e^{kz} + \{-C_i + (1-kz)G_i\}e^{-kz}] \tag{41c}$$

$$\hat{\tau}_{xzi}(k,z) = k[\{A_i + (2+kz)B_i\}e^{kz} + \{C_i + (kz-2)G_i\}e^{-kz}] \tag{41d}$$

$$\hat{\sigma}_{xi}(k,z) = -jk[\{A_i + (3+kz)B_i\}e^{kz} + \{-C_i + (3-kz)G_i\}e^{-kz}],\ (k,z)\epsilon\hat{L}_i. \tag{41e}$$

We consider the boundary conditions:

(a) $\hat{u}_m(k,D) = \hat{w}_m(k,D) = 0$, $D = D_m$. Perfect adhesion between the layer system and the rigid substrate. (42a)

(b) $\hat{\sigma}_z(k,0) \underset{=}{def}\ \hat{\sigma}(k) =$ prescribed, $\hat{\tau}_{xzl}(k,0) \underset{=}{def}\ \hat{\tau}(k) =$ prescribed. Surface loads prescribed. (42b)

(c) $\hat{u}_i(k,D_i) = \hat{u}_{i+1}(k,D_i)$; $\hat{w}_i(k,D_i) = \hat{w}_{i+1}(k,D_i)$; $\hat{\sigma}_{zi}(k,D_i) = \hat{\sigma}_{z,i+1}(k,D_i)$; $\hat{\tau}_{xzi}(k,D_i) = \hat{\tau}_{xz,i+1}(k,D_i)$, $i=1,\ldots m-1$. Continuity of u, w, σ_z, τ_{xz} on the inner boundaries between the constituent layers. (42c)

Equations (42) constitute $4m$ linear equations; for the $4m$ integration constants A_i, B_i, C_i, G_i, $i=1,\ldots,m$. The coefficients of these equations and the integration constants are functions of the transform parameter k.

Let

$$\mathbf{A}_i = (A_i, B_i, C_i, G_i)^T,\ \mathbf{O}_i = (0,0,0,0)^T,\ i=1\ldots,m. \tag{43}$$

Then the boundary conditions (42) become, in symbolic form

$$N\begin{bmatrix} A_1 \\ B_1 \\ C_1 \\ G_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} = \begin{bmatrix} \hat{\sigma}(k)/k \\ \hat{\tau}(k)/k \\ 0 \\ 0 \\ \mathbf{O}_2 \\ \vdots \\ \mathbf{O}_m \end{bmatrix} \Rightarrow \begin{bmatrix} A_1 \\ B_1 \\ C_1 \\ G_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} = N^{-1}\begin{bmatrix} \hat{\sigma}(k)/k \\ \hat{\tau}(k)/k \\ 0 \\ 0 \\ \mathbf{O}_2 \\ \vdots \\ \mathbf{O}_m \end{bmatrix} \tag{44}$$

where N is a $4m \times 4m$ complex matrix function of k which is regular and continuous everywhere.

Define $\mathbf{S}(k)$ as the first column of N^{-1}, and $\mathbf{T}(k)$ as the second column of N^{-1}. Then we have, if $\mathbf{S}$ and $\mathbf{T}$ are partitioned:

$$\begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} = \begin{bmatrix} \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_m \end{bmatrix}\hat{\sigma}/k + \begin{bmatrix} \mathbf{T}_1 \\ \vdots \\ \mathbf{T}_m \end{bmatrix}\hat{\tau}/k$$

$\mathbf{S}_i$, $\mathbf{T}_i$ are complex 4-vectors. (45)

We are interested in the response of the layered medium to the normal surface load

$$\sigma_0(x) = I(x) = 0 \qquad |x| > a$$
$$= 1 \qquad |x| < a \tag{46a}$$
$$\hat{I}(k) = \{2\sin(ka)\}/k$$

and also to the tangential surface load

$$\tau_0(x) = I(x). \tag{46b}$$

We will approximate the true surface stress distribution by the following piecewise constant one:

$$\sigma(x) = \sum_{h=-n}^{n} \sigma_h I(x-2ha)$$
$$\tau(x) = \sum_{h=-n}^{n} \tau_h I(x-2ha) \tag{47}$$

where $\{\sigma_h, \tau_h\}$ are constants to be determined by the algorithms N and F, and the Panagiotopoulos process.

The calculation consists accordingly of two parts:

(1) We must know the elastic field $\mathbf{F}^n(x,z) = (u^n(x,z), w^n(x,z), \sigma_z^n(x,z), \tau_{xz}^n(x,z), \sigma_x^n(x,z))^T$ due to the surface load

$$\sigma(x) = I(x),\ \tau(x) = 0 \tag{48a}$$

and the similarly defined elastic field $\mathbf{F}^t(x,z) = (u^t(x,z),\ldots,\ldots,\ldots,\ldots)^T$ due to the surface load

$$\sigma(x) = 0,\ \tau(x) = I(x). \tag{48b}$$

We call $\mathbf{F}^n$ and $\mathbf{F}^t$ the normal and tangential *influence functions* of the problem.

(2) Once we have the influence functions, we determine the weight factors $\{\sigma_n, \tau_h\}$ by the algorithms N and F, and the Panagiotopoulos process. Then the resulting field is

$$\mathbf{F}(x,z) = \sum_{h=-n}^{n} \{\sigma_h \mathbf{F}^n(x-2ha,z) + \tau_h \mathbf{F}^t(x-2ha,z)\}. \quad (49)$$

For the determination of $\{\sigma_h, \tau_h\}$ we need, apart from the surface load, the surface displacement:

$$u(x) \underline{\underline{def}} u_1(x,0) = \sum_{h=-n}^{n} \{\sigma_h u^n(x-2ha) + \tau_h u^t(x-2ha)\} \quad (50a)$$

$$w(x) \underline{\underline{def}}\ w_1(x,0) = \sum_{h=-n}^{n} \{\sigma_h w^n(x-2ha) + \tau_h w^t(x-2ha)\}. \quad (50b)$$

4.1 The Fourier Transform of the Influence Functions. We transform σ_0 (x) and $\tau_0(x)$; we denote the result by $\hat{\sigma}_0(k)$, $\hat{\tau}_0(k)$, and we drop the subscript zero. For a purely normal or a purely tangential load we have

$$\hat{\sigma}^n(k) = (\sin ka)/k, \quad \hat{\tau}^n(k) = 0 \quad (51a)$$

$$\hat{\sigma}^t(t) = 0, \qquad \hat{\tau}^t(k) = (\sin ka)/k. \quad (51b)$$

Usually, our considerations hold for a purely normal and for a purely tangential load, and then we will drop the superscripts "n" or "t". Occasionally, however, it is essential to distinguish between the two types of loading, and then we will use the superscripts.

We recall that the connection between the elastic field $\hat{\mathbf{F}}$ and the integration constants $\mathbf{A}_i$ was given in equation (41). In symbolic form we have

$$\hat{\mathbf{F}} = M_i \mathbf{A}_i = M_i(\mathbf{S}_i\ \hat{\sigma}/k + \mathbf{T}_i \hat{\tau}/k) = \mathbf{S}_i^* \hat{\sigma}/k + \mathbf{T}_i^* \hat{\tau}/k, \quad (52a)$$

with M_i a complex 5×4 matrix depending on k and z, defined by equation (41); (52b)

$\mathbf{S}_i$, $\mathbf{T}_i$: see equations (42), (44), (45); they are complex 4-vectors, functions of k and z. (52c)

So we have obtained the elements of $\hat{\mathbf{F}}$; they are complex. However $\mathbf{F}$ itself is real, so that

$$\hat{\mathbf{F}}(k) = \int_{-\infty}^{\infty} e^{jkx}\, \mathbf{F}(x)\, dx = \int_{-\infty}^{\infty} \mathbf{F}(x) \cos(kx)\, dx + j\int_{-\infty}^{\infty} \mathbf{F}(x) \sin(kx)\, dx$$

so that

$$\mathrm{Re}\,(\hat{\mathbf{F}}(k)) = \int_{-\infty}^{\infty} \mathbf{F}(x) \cos(kx)\, dx\text{: even in } k. \quad (53a)$$

$$\mathrm{Im}\,(\hat{\mathbf{F}}(k)) = \int_{-\infty}^{\infty} \mathbf{F}(x) \sin(kx)\, dx\text{: odd in } k. \quad (53b)$$

On the other hand,

$$\mathbf{F}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\mathbf{F}}(k) e^{-jkx} dk = $$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Re}\,(\hat{\mathbf{F}}(k)) e^{-jkx} dk + \frac{j}{2\pi}\int_{-\infty}^{\infty} \mathrm{Im}(\hat{\mathbf{F}}(k)) e^{-jkx} dk$$

so that, since $\mathbf{F}(x)$ is real,

$$\mathbf{F}(x) = \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}(\hat{\mathbf{F}}(k)) \cos kx\, dk + \frac{1}{\pi}\int_0^{\infty} \mathrm{Im}\,(\hat{\mathbf{F}}(k)) \sin(kx)\, dk \quad (54)$$

which are real integrals, one a cosine, one a sine transform. We will show in Appendix C how these integrals can be calculated numerically, fast, and with a prescribed accuracy.

When we consider elasticity rather than viscoelasticity, $K = Q$, and some gain in calculation speed can be obtained by keeping track of the real and imaginary quantities. Then equation (52) can be formulated in terms of purely real and purely imaginary components of $\hat{\mathbf{F}}$, while there is hardly any need of complex arithmetic. Either one or the other integral appears alone in equation (54). This yields a reduction in calculating speed of roughly a factor 4, which is due to the circumstance that a complex multiplication results in 4 real products instead of in 1. On the other hand, the algorithms N and F are equally fast in viscoelasticity as in elasticity, so that once the influence functions are known, the viscoelastic and elastic calculations are equally fast.

5 Conclusion

A fast method has been presented for the calculation of the elastic field on and inside a viscoelastic or elastic multilayered cylinder. It is found that the calculation times for a viscoelastic multilayer differs not too much from its elastic counterpart. Details of the calculation are given in the Appendices of this paper, as well as a dimensional analysis.

References

Bentall, R. H., and Johnson, K. L., 1967, "Slip in the Rolling Contact of Two Dissimilar Elastic Rollers," *International Journal of Mechanical Science*, Vol. 9, pp. 389–404.

Bentall, R. H., and Johnson, K. L., 1968, "An Elastic Strip in Rolling Contact," *International Journal of Mechanical Science*, Vol. 10, pp. 637–663.

Kalker, J. J., 1983, "Two Algorithms for the Contact Problem in Elastostatics," *Contact Mechanics and Wear of Rail/Wheel Systems*, J. Kalousek, et al., eds., University of Waterloo Press, pp. 101–120.

Kalker, J. J., 1988, "Contact Mechanical Algorithms," *Communications in Applied Numerical Methods*, Vol. 4, pp. 25–32.

APPENDIX A

A Second-Order Approximation for the Deformed Distance

In some applications it is found that in equation (17)

$$0 = d(x) \simeq d_1(x) \underline{\underline{def}}\ h(x) + w_1(x) - w_2(x)$$

$$h(x) = z_1(x) - z_2(x) \quad (17)(\mathrm{A}1)$$

is not satisfied at all points of the contact area of one specifies that $d_1(x) = 0$ in contact. Apparently, the deformed distance of equation (A1), which has a second-order error, is not accurate enough. In this Appendix we derive an expression $d_2(x)$ of $d(x)$, with third-order error. To that end we reanalyze the deformed distance (see Fig. A1).

From Fig. A1(a) it is clear that

$$d(x) = z_1^* - z_2^*. \quad (\mathrm{A}2)$$

So we must calculate z_a^*, $a = 1, 2$ (see Fig. 1A(b)). We note that $\mathbf{u}_a$, $a = 1, 2$, is the displacement at the surface of the ath body, at x-coordinate x_a (XZ in Fig. A1(b)). We calculate u_a, the displacement component tangential to the undeformed surface (XV in Fig. A1 (b)), and w_a, the displacement component normal to it (VZ in Fig. A1(b)). They are assumed to coincide with the x, z displacement components in the half-space approximation. The auxiliary quantities $w_a^* = w_a/\cos(z_{a,x}) = w_a$ + third-order terms. So we have that

$$\begin{aligned} z_a^* &= z_a(x) + w_a^* = z_a(x) + w_a(x_a) + \text{third-order terms} \\ &= z_a(x) + w_a(x - u_a(x_1) + w_a(x) z_{a,x}) + \text{third-order-terms} \\ &= z_a(x) + w_a(x - u_a(x)) + \text{third-order terms} \quad (\mathrm{A}3a) \\ &= z_a(x) + w_a(x) - u_a(x) w_{a,x}(x) + \text{third-order terms}. \quad (\mathrm{A}3b) \end{aligned}$$

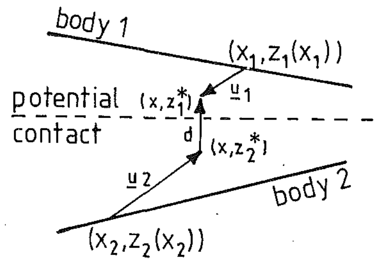


Fig. A1(*a*)

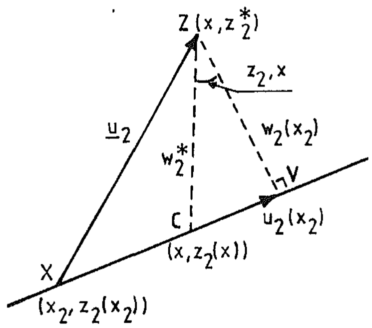


Fig. A1(*b*)

Fig. A1 (*a*) The deformed distance *d* (*b*) construction of (x, z_2^*)

Therefore,

$$d(x) = d_2(x) + \text{third-order terms,}$$

$$d_2(x) = z_1(x) - z_2(x) + w_1(x) - w_2(x) - w_{1,x}(x)\, u_1(x) + w_{2,x}(x) u_2(x) \qquad \text{(A4}a\text{)}$$

$$= d_1(x) + w_{2,x}(x) u_2(x) - w_{1,x}(x)\, u_1(x) \qquad \text{(A4}b\text{)}$$

$$= z_1(x) - z_2(x) + w_1(x - u_1(x)) - w_2(x - u_2(x)) \qquad \text{(A4}c\text{)}$$

$$= d_1(x) + w_1(x - u_1(x)) - w_1(x) - w_2(x - u_2(x)) + w_2(x). \qquad \text{(A4}d\text{)}$$

$d_2(x)$ is the promised expression for $d(x)$ with third-order error.

If we calculate so that $d_1(x) = 0$, we are left with a deformed distance d_{error} which equals

$$d_{\text{error}}(x) = w_{2,x}(x) u_2(x) - w_{1,x}(x) u_1(x) + \text{third-order terms,} \qquad \text{(A5}a\text{)}$$

$$= w_1(x - u_1(x)) - w_1(x) - w_2(x - u_2(x)) + w_2(x) + \text{third-order terms.} \qquad \text{(A}b\text{)}$$

In the correction terms of the constituent $w_{a,x}$ and $u_a(x)$ need not be calculated very precisely.

APPENDIX B

Dimension Analysis

In this Appendix we consider the interdependencies of the parameters of the problem. This will result in the construction of a number of parameters by which the solution, viz. the Airy function H, will be nondimensionalized.

The problem is governed by the following equations and inequalities.

$$(1 - q_{j\alpha} V \partial/\partial x) u = (K_{j\alpha} - q_{j\alpha} Q_{j\alpha} V\, \partial/\partial x) \times \{(1 - \nu_{j\alpha}^2) H_{,zzz} - \nu_{j\alpha}(1 + \nu_{j\alpha}) H_{,xxz}\} \qquad (30a)(\text{B}1a)$$

$$(1 - q_{j\alpha} V \partial/\partial x) w = (K_{j\alpha} - q_{j\alpha} Q_{j\alpha} V\, \partial/\partial x) \times \{(1 - \nu_{j\alpha}^2) H_{,xxx} - \nu_{j\alpha}(1 + \nu_{j\alpha}) H_{,xzz}\} \qquad (30b)(\text{B}1b)$$

$$\sigma_z = H_{,xxxz},\ \tau_{xz} = -H_{,xxzz},\ \sigma_x = H_{,xzzz} \qquad (26d)(\text{B}2a)$$

$$H_{,xxxx} + 2H_{,xxzz} + H_{,zzzz} = 0. \qquad (26e)(\text{B}2b)$$

Subscript j: sublayer number, α: cylinder number.

Boundary conditions are:
The field quantities u, w, σ_z, τ_{xz} are continuous at the interfaces $(-1)^{\alpha-1} D_{j\alpha}$, $j = 1, \ldots m^\alpha - 1$; $u = w = 0$ at $z = (-1)^{\alpha-1} D_\alpha$, where the interface is with the rigid substrate, see equation (13). (B3)

Further, the contact formation conditions:

$$\left.\begin{aligned} &d(x) = h(x) + w(x) \geq 0, = 0 \text{ in contact;} \\ &w(x) = w_1(x,0) - w_2(x,0); \\ &\sigma(x) = \sigma_z(x,0) \leq 0 \text{ in contact, } = 0 \text{ outside contact;} \end{aligned}\right\} \qquad (19)(\text{B}4a)$$

$$h(x) = \left(\frac{1}{2R_1} + \frac{1}{2R_2}\right) x^2 - \beta = Ax^2 - \beta \qquad (17)(\text{B}4b)$$

and the frictional conditions are:

$$\left.\begin{aligned} &|\tau| \leq -f\sigma,\ \tau(x) = \tau_{xz}(x,0) && \text{on the entire surface} \\ &\text{if } v \neq 0 \text{ then } \tau = -f\sigma \text{ sign } (v) && (v\text{: slip of body 2 over body 1}) \end{aligned}\right\} \qquad (25)(\text{B}5a)$$

where it can be derived from equation (22) that

$$\left.\begin{aligned} s = v/Vt = \nu + \partial u/\partial x;\ &t\text{:a time increment} \\ &V\text{: rolling velocity} \end{aligned}\right\} Vt = a, \text{ element length.} \qquad (\text{B}5b)$$

$$u(x) = u_1(x,0) - u_2(x,0);$$

As $Vt > 0$ we can use $s = v/Vt$ instead of v in equation (25).

It is seen, by enumeration, that the problem is uniquely defined by the parameters

$$P = (q_{j\alpha} V, K_{j\alpha}, Q_{j\alpha}, \nu_{j\alpha}, f, \upsilon, \beta, A, D_{j\alpha}) \qquad (\text{B}6)$$

and that the variables are (x, z). Furthermore, the problem is described by the Airy function H

$$H = H(P; x,z). \qquad (\text{B}7)$$

Quantities of interest are:

$$u, w, \sigma_x, \tau_{xz}, \sigma_z, b_1 < b_2$$

(the endpoints of contact), $\mathbf{Q}$ (total force). (B8)

Let λ, μ, η > 0, 0, 0 be real, positive scale factors, and define

$$\left.\begin{aligned} &\overline{q_{j\alpha} V} = \lambda q_{j\alpha} V, && \bar{K}_{j\alpha} = K_{j\alpha}/\mu && \bar{Q}_{j\alpha} = Q_{j\alpha}/\mu \\ &\bar{\nu}_{j\alpha} = \nu_{j\alpha}, && \bar{f} = f, && \bar{\nu} = \eta\nu \\ &\bar{\beta} = \lambda\eta\beta, && \bar{A} = \eta A/\lambda, && \bar{D}_{j\alpha} = D_{j\alpha} \\ &\bar{x} = \lambda x, && \bar{z} = \lambda z && \end{aligned}\right\} \qquad (\text{B}9)$$

and let

$$\bar{P}=(\overline{q_{j\alpha}V},\ \bar{K}_{j\alpha},\ \bar{Q}_{j\alpha},\ \bar{\nu}_{j\alpha},\ \bar{f},\ \bar{\upsilon},\ \bar{\beta},\ \bar{A},\ \bar{D}_{j\alpha}). \quad \text{(B10)}$$

The Airy function with parameters P, variables (x, z), which satisfies the biopotential equation, and has displacements (u, v), stresses (σ,τ,σ_x), deformed distance d and slip s, is given by equation (B7). The Airy function of the contact problem with parameter $\bar{P}$, variables $(\bar{x},\bar{z})$ is given by $H(\bar{P};\bar{x},\bar{z})$; it satisfies the bipotential equation in the variables $(\bar{x},\bar{z})$. *It will be shown that*

$$H(\bar{P};\ \bar{x},\ \bar{z})=\lambda^4\mu\eta\ H(P;\ x,\ z). \quad \text{(B11)}$$

To that end we define

$$H_0(\bar{P};\ \bar{x},\ \bar{z})\ \underline{\underline{\text{def}}}\lambda^4\mu\eta\ H(P;\ x,z).$$

The we have:

$$H_{0,\bar{x}}=\lambda^{-1}H_{0,x}=\lambda^3\mu\eta\ H_{,x}(P;x,z)$$
$$H_{0,\bar{z}}=\lambda^{-1}H_{0,z}=\lambda^3\mu\eta\ H_{,z}(P;x,z).$$

Higher derivatives satisfy similar formulae. Clearly, then, $H_0(\bar{P};\ \bar{x},\ \bar{z})$ satisfies the bipotential equation in the variables $\bar{x},\ \bar{z}$. So, H_0 is a feasible Airy function in the variables $\bar{x},\ \bar{z}$, and field quantities $(\bar{u},\ \bar{w},\ \bar{\sigma},\ \bar{\tau},\ \bar{\sigma}_x,\ \bar{d},\ \bar{s})$. We calculate these field quantities. Consider $\bar{u}$:

$$(1-\overline{q_{j\alpha}V}\ \partial/\partial\bar{x})\bar{u}=(1-q_{j\alpha}V\ \partial/\partial x)\ \bar{u}=$$
$$=(\bar{K}_{j\alpha}-\overline{q_{j\alpha}V}\bar{Q}_{j\alpha}\ \partial/\partial\bar{x})$$
$$\times\{(1-\bar{\nu}^2_{j\alpha})\ H_{0},\bar{z}\bar{z}\bar{z}-\bar{\nu}_{j\alpha}(1+\nu_{j\alpha})H_{0},\bar{x}\bar{x}\bar{z}\}=$$
$$=\lambda\eta(K_{j\alpha}-q_{j\alpha}VQ_{j\alpha}\partial/\partial x)$$
$$\times\{(1-\nu^2_{j\alpha}\ H_{,zzz}(P;x,z)-\nu_{j\alpha}(1+\nu_{j\alpha})H_{,xxz}(P;x,z)\}$$
$$\Rightarrow\ \bar{u}=\lambda\eta u.$$

Similarly, by equation (B1*b*), $\bar{w}=\lambda\eta w$; by equation (B2*a*), $\bar{\sigma}=\mu\eta\sigma$, $\bar{\tau}=\mu\eta\tau$, $\bar{\sigma}_x=\mu\eta\sigma_x$. Therefore, the field quantities $\bar{u}$, $\bar{w}$, $\bar{\sigma}$, $\bar{\tau}$ are continuous throughout the layer, since u, w, σ, τ are, and λ, η, μ are constants. The undeformed distance $\bar{h}(\bar{x})$ corresponding to $H_0(\bar{P};\ \bar{x},\ \bar{z})$ is (see equation (B4*b*),

$\bar{h}(\bar{x})=\lambda\eta h(x)\Rightarrow$ equation (B4*a*)

is satisfied in the barred system

and the frictional conditions (B5) are likewise satisfied with $\bar{s}=\eta s$.

This establishes that H_0 is the Airy function of the contact problem for the parameters $\bar{P}$ and the variables $\bar{x},\ \bar{z}$, so that it can be identified with $H(\bar{P};\ \bar{x},\ \bar{z})$, as we set out to prove.

We can derive expressions for the modified contact end points and the total force:

$$\bar{b}_1=\lambda b_1,\ \bar{b}_2=\lambda b_2,\ \bar{\mathbf{Q}}=-\int_{\bar{b}_1}^{\bar{b}_2}(\bar{\tau},\ \bar{\sigma})\ d\bar{x}=\lambda\mu\eta\mathbf{Q}. \quad \text{(B12)}$$

We can use equations (B9)–(B11) to reduce the number of parameters needed to specify a problem. Set

$$\left.\begin{array}{l}\lambda=A^{1/2}\beta^{-1/2},\ \eta=A^{-1/2}\beta^{-1/2},\ \mu=K\\ \text{with } K \text{ some representative elastic constant, of dimension m}^2/N\end{array}\right\}. \quad \text{(B13}a\text{)}$$

Then,

$$\left.\begin{array}{lll}H(P;x,z)=\lambda^{-4}\mu^{-1}\eta^{-1}\ H(\bar{P}\ \bar{x},\bar{z}) & & \\ \lambda^{-4}\mu^{-1}\eta^{-1}=A^{-3/2}\beta^{5/2}K^{-1}, & \bar{P}\text{ given by equation (B10);} & \\ \text{(B9) is:} & & \\ \overline{q_{j\alpha}V}=A^{1/2}\beta^{-1/2}q_{j\alpha}V, & \bar{K}_{j\alpha}=K_{j\alpha}/K, & \bar{Q}_{j\alpha}=Q_{j\alpha}/K \\ \bar{\nu}_{j\alpha}=\nu_{j\alpha}, & \bar{f}=f & \bar{\nu}=A^{-1/2}\beta^{-1/2}\nu \\ \boxed{\bar{\beta}=1} & \boxed{\bar{A}=1} & \bar{D}_{j\alpha}=A^{1/2}\beta^{-1/2}D_{j\alpha} \\ \bar{x}=A^{1/2}\beta^{-1/2}x, & \bar{z}=A^{1/2}\beta^{-1/2}z & \end{array}\right\}. \quad \text{(B13}b\text{)}$$

As we see two parameters drop out, as $\bar{\beta}=1$ and $\bar{A}=1$, while one of the constants $\bar{K}_{j\alpha}$, $\bar{Q}_{j\alpha}$ may be likewise set equal to 1, by a proper choice of K.

An important special case is *the case of symmetry* about the plane $z=0$. Then we have that

$$\left.\begin{array}{ll}D_j=D_{j1}=D_{j2};R_1=R_2=R\,(A=1/R) & \text{geometric symmetry}\\ \nu_j=\nu_{j1}=\nu_{j2};q_j=q_{j1}=q_{j2}; & \\ K_j=K_{j1}=K_{j2};Q_j=Q_{j1}=Q_{j2} & \text{material symmetry}\end{array}\right\} \quad \text{(B14)}$$

and in equation (B13*b*) one may omit the subscript α.

In this case of symmetry we decompose $H(P;\ x,\ z)$ into two sub-Airy functions $H_n(x,\ z)$ and $H_t(x,z)$. Quantities pertaining to H_n are given a subscript n; similarly, quantities pertaining to H_t are distinguished by a subscript t. In addition, quantities of the upper cylinder ($z\geq 0$) carry a subscript 1, the quantities of the lower cylinder ($z\leq 0$) carry a a subscript 2. All Airy functions are defined in the entire space, and satisfy equations (B1), (B2), (B3);

$$\sigma_n,\ \sigma_t,\ \sigma \text{ and } \tau_n,\ \tau_t,\ \tau \text{ are continuous at } z=0. \quad \text{(B15)}$$

In addition,

$$H(P;\ x,z)=H_n(x,\ z)+H_t(x,\ z); \quad \text{(B16}a\text{)}$$

$H_n(x\ ,z)$ is odd in z

$$\Rightarrow H_{n,z}\ (x,\ z) \text{ is even in } z \quad \text{(B16}b\text{)}$$

$H_t(x\ ,z)$ is even in z

$$\Rightarrow H_{t,z}\ (x,\ z) \text{ is odd in } z. \quad \text{(B16}c\text{)}$$

Simultaneous satisfaction of equations (B15) and (B16) is possible owing to the symmetry of the problem.

Inspection of equation (B1*a*) teaches us that in view of equation (B16*b*) u_n is even in z. Similarly, we find by inspection of equations (B1) and (B2*a*):

$$u_n(x,\ z),\ w_t(x,\ z),\ \sigma_{nz}\ (x,\ z),$$
$$\tau_{txz}(x,z),\ \sigma_{nx}(x,\ z) \text{ are even in } z; \quad \text{(B17}a\text{)}$$

$$u_t(x,\ z),\ w_n(x,\ z),\ \sigma_{tz}\ (x,\ z),$$
$$\tau_{nxz}(x,\ z),\ \sigma_{tx}(x,z) \text{ are odd in } z. \quad \text{(B17}b\text{)}$$

Consider the contact formation conditions (B4). We have:

$$w(x)=w_1(x,0)-w_2(x,0)$$
$$=w(x,0+)-w(x,0-) \text{ is discontinuous}; \qquad \text{(B18}a\text{)}$$
$$=2w_n(x,0+).$$

Since $\sigma_{tz}(x, z)$ is odd and continuous at $z = 0$, $\sigma_{tz}(x, 0) = 0$, and

$$\sigma(x)=\sigma_{nz}(x,0)+\sigma_{tz}(x,0)=\sigma_{nz}(x,0). \qquad \text{(B18}b\text{)}$$

So the tangential (t) quantities do not affect the contact formation which is governed by equation (B4), and we determine $H_n(x, z)$ without being influenced by the tangential quantities (t). Consider the frictional conditions (B5)

$$u(x)=u_1(x,0)-u_2(x,0)$$
$$=u(x,0+)-u(x,0-) \text{ is discontinuous}; \qquad \text{(B18}b\text{)}$$
$$=2u_t(x,0+).$$

Since $\tau_{nxz}(x, z)$ is odd and continuous at $z = 0$, $\tau_n(x, 0) = 0$, and

$$\tau(x)=\tau_{nxz}(x,0)+\tau_{txz}(x,\ 0)=\tau_{txz}(x,0). \qquad \text{(B19}a\text{)}$$

So the normal (n) quantities affect the tangential problem only through the traction bound ($-f\sigma$), which is determined without reference to the tangential quantities (t). Apart from that, the normal quantities (n) do not affect the calculation of H_t. The parameters P_n which determine the Airy function H_n of the normal contact problem are the same as of equation (B13b), but the friction coefficient $\bar{f}$ and the creepage $\bar{\nu}$ may be omitted. The parameters are

$$A^{1/2}\beta^{-1/2}q_jV,\ K_j/K,\ Q_j/K,\ \nu_j,$$
$$A^{1/2}\beta^{-1/2}\ D_j \text{ (all dimensionless).} \qquad \text{(B20}a\text{)}$$

In case there is only one homogeneous elastic layer on each cylinder, we can take

$$K_j=Q_j=K,\ q_j=0,\ D_j=D,$$
$$\text{and the parameters are only } \nu_1,\ A^{1/2}\beta^{-1/2}D. \qquad \text{(B20}b\text{)}$$

The parameters P_t which determine the Airy function H_t of the tangential problem are (B13b),

$$A^{1/2}\beta^{-1/2}q_jV,\ K_j/K,\ Q_j/K,\ \nu_j,\ f,$$
$$A^{-1/2}\beta^{-1/2}\nu,\ A^{1/2}\beta^{-1/2}D_j$$

ν and f occur only in the slip, and in the traction bound of Coulomb's law (see equation (B5)).

Suppose we divide τ by the friction coefficient f in equation (B5a). The easiest way to accomplish this is to divide H_t by f. This does not affect the normal stress σ, as H_n, and hence the normal quantities are not affected. u_t, however, is affected; it is divided by f. To retain the slip equation (B5b) intact, ν should be divided by f. The direction of s is not affected, as $f > 0$. So the governing parameters of the tangential problem become:

$$A^{1/2}\beta^{-1/2}q_jV,\ K_j/K,\ Q_j/K,\ \nu_j,$$
$$A^{-1/2}\beta^{-1/2}f^{-1}\nu,\ A^{1/2}\beta^{-1/2}D_j. \qquad \text{(B21(}a\text{))}$$

In case that there is only one homogeneous elastic layer on each cylinder, we can take, as in the normal problem, $K = K_j = Q_j$, $q_j = 0$, $D_j = D$, and then the parameters are only

$$\nu_1,\ A^{1/2}\beta^{-1/2}D,\ A^{-1/2}\beta^{-1/2}f^{-1}\nu, \qquad \text{(B21}b\text{)}$$

all three of which are dimensionless.

In the case that there is no symmetry, but that each cylinder is covered by a homogeneous, elastic layer, there are seven dimensionless parameters, viz.

$$K_1/K_2,\ \nu_1,\ \nu_2,\ f,\ A^{-1/2}\beta^{-1/2}\nu,\ A^{1/2}\beta^{-1/2}D_1,A^{1/2}\beta^{-1/2}D_2. \qquad \text{(B22)}$$

APPENDIX C

Numerical Analysis

In this Appendix we consider the numerical analysis of the inversion of the field quantities $\hat{\mathbf{F}}(k;\ z)$. To that end we investigate

(1) the behavior of $\hat{\mathbf{F}}(k,\ z)$ near $k = 0$,
(2) the behavior of $\hat{\mathbf{F}}(k,\ z)$ as $k \to \infty$, and
(3) the numerical inversion of the Fourier integrals.

C1 The behavior of $\hat{\mathbf{F}}(k,\ z)$ near $k = 0$. As we see from equation (52), a factor k occurs in the denominator of the transform $\hat{\mathbf{F}}$. It is therefore interesting to investigate what happens when $k \to 0$. To that end, we expand $\hat{\mathbf{F}}$ in a Laurent series about $k = 0$:

$$\hat{\mathbf{F}}(k)=\hat{\mathbf{F}}_{-1}/k+\hat{\mathbf{F}}_0+\hat{\mathbf{F}}_1k+\hat{\mathbf{F}}_2k^2+\ldots \qquad \text{(C1)}$$

where $\hat{\mathbf{F}}_\ell \in \mathbb{C}$

Since Re $(\hat{\mathbf{F}}(k))$ is even in k, and Im $(\hat{\mathbf{F}}(k))$ odd, $\hat{\mathbf{F}}_{2\ell}$ is real, and $\hat{\mathbf{F}}_{2\ell-1}$ is purely imaginary.

It is of interest numerically to know $\hat{\mathrm{F}}_0$ and $\hat{\mathbf{F}}_{-1}$.

C1.1 $\hat{\mathbf{F}}_0$. We have:

$$\hat{\mathbf{F}}_0=\frac{d}{dk}\,(k\,\hat{\mathbf{F}}_k)=\frac{d}{dk}\,\{M_i\ (\mathbf{S}_i\hat{\sigma}+\mathbf{T}_i\hat{\tau})\}, \qquad \text{(C2)}$$
$$\mathbf{S}_i,\ \mathbf{T}_i,\ M_i \text{ regular in } k=0.$$

By equations (48) and (46a) we also have

$$\hat{\sigma}(k)=\theta\ \sin\ (ka)/k=\theta a\left(1-\frac{1}{6}\,k^2a^2\right),\ \theta=0, \text{ or } 1$$
$$\hat{\tau}(k)=(1-\theta)\ \sin\ (ka)/k=(1-\theta)\ a\ \left(1-\frac{1}{6}\,k^2a^2\right) \qquad \text{(C3)}$$

so that

$$\hat{\sigma}(0)=\theta a,\ \hat{\sigma}'(0)=0;\ '\underline{\underline{def}}\ d/dk$$
$$\hat{\tau}(0)=(1-\theta)a,\ \ \hat{\tau}'(0)=0. \qquad \text{(C4)}$$

Then we have

$$\hat{\mathbf{F}}_0=a\mathbf{M}_i'(0)\ (\theta\mathbf{S}_i(0)\hat{\sigma}+(1-\theta)\mathbf{T}_i(0)\hat{\tau})$$
$$+\,a\mathbf{M}_i(0)(\theta\mathbf{S}_i'(0)\hat{\sigma}+(1-\theta)\ \mathbf{T}_i'(0)\tau). \qquad \text{(C5)}$$

$\mathbf{M}_i'$ follows straight from equation (41). The $\mathbf{S}_i$ and $\mathbf{T}_i$ constitute the first two columns $\mathbf{S}$ and $\mathbf{T}$ of the square, regular matrix N^{-1} of equations (42) and (44), independent of z, as follows:

$$\mathbf{S}=(\mathbf{S}_1^T,\ldots,\mathbf{S}_m^T)^T,\ \mathbf{T}=(\mathbf{T}_1^T,\ldots,\mathbf{T}_m^T)^T. \qquad \text{(C6)}$$

The derivative of the inverse of N reads

$$\{N(k)^{-1}\}'=-N(k)^{-1}\{N(k)\}'N(k)^{-1} \qquad \text{(C7}a\text{)}$$

so that

$$\mathbf{S}'(k)=-N(k)^{-1}\{N(k)\}'\mathbf{S}(k)$$
$$\mathbf{T}'(k)=-N(k)^{-1}\{N(k)\}'\mathbf{T}(k) \qquad \text{(C7}b\text{)}$$

all factors of which can be determined straightforwardly.

C1.2 $\hat{\mathbf{F}}_{-1}$. It will appear useful to separate the term $\hat{\mathbf{F}}_{-1}/k$ (see equations (C1), (52)) from the rest:

$$\hat{\mathbf{F}}(k)=\{\hat{\mathbf{F}}(k)-\hat{\mathbf{F}}_{-1}/k\}+\hat{\mathbf{F}}_{-1}/k$$
$$\hat{\mathbf{F}}_{-1}=\lim_{k\to\infty}k\hat{\mathbf{F}}_k=a\{\theta\mathbf{S}_i^*(0,z)+(1-\theta)\mathbf{T}_i^*(0,z)\} \qquad \text{(C8)}$$
$$\hat{\mathbf{F}}(k)-\hat{\mathbf{F}}_{-1}/k=O(1) \text{ as } k\to\infty. \qquad \text{(C9)}$$

C2 The behavior of $\hat{\mathbf{F}}$ (k, z) as $k \to \infty$. Also important is the behavior of $\hat{\mathbf{F}}(k, z)$ as $k \to \pm\infty$. In view of equation (54) it is sufficient to consider only the case that $k \to \infty$. We must consider the equations (42)–(41), which we will write out in abbreviated form in Table 1. For the abbreviation we use the following convention:

$+i$	means $(a+bkz)e^{+D_ik}$	a,b: arbitrary functions of k	
$-i$	means $(a+bkz)e^{-D_ik}$	$i=1,2,3,\ldots$	
$+$	means $(a+bkz)$		(C10)
0	means 0		
X	means "a known function of k."		

In our order-of-magnitude considerations we confine ourselves to the most important terms, viz. the exponentials.

We write out the equations for a three-layered medium.

We start from the supposition that a positive exponential behavior of $\hat{\mathbf{F}}$ does not occur, since otherwise a Fourier inversion is impossible. At the first interface, A_1 and B_1 are multiplied by e^{D_1k}, and at the top by e^{0k}. So A_1 and B_1 are at most of magnitude $O(e^{-D_1k})$ as $k \to +\infty$, and the contribution of A_1 and B_1 at the top will be at most $O(e^{-D_1k})$, which will turn out to be of secondary importance. We encircle the elements in the first interface corresponding to A_1 and B_1, as the coefficients that determine the behavior of A_1, B_1.

We turn to C_1, G_1. At the top, they are multiplied by e^{0k}. At the first interface, they are multiplied by e^{-D_1k}. So they are at most $O(1)$: then they remain bounded in the first layer, and on the first interface they contribute $O(e^{-D_1k})$ *to* $\hat{\mathbf{F}}$, well below the $O(1)$ contributed by A_1 and B_1. We encircle the elements of C_1 and G_1 at the top, as the coefficients that determine the behavior of C_1 and G_1. Now we see that C_1 and G_1 are actually of the order of magnitude of the knowns. A_1 and B_1 are not determined by the first two equations, and $A_2, \ldots$ and $A_3, \ldots$ are not involved at all in the first two equations.

We turn to A_2 and B_2, and try to find the significant coefficients.

First Interface: coeff $= O(e^{kD_1})$

Second Interface: coeff $= O(e^{kD_2})$, $D_2 > D_1$; $\Rightarrow A_2, B_2 = O(e^{-kD_2})$.

From this we see that the coefficients of A_2, B_2 in the second interface are the significant ones. We encircle them. Similarly, all other encirclements are placed.

When $k \to \infty$, the nonencircled coefficients give no contribution. Concentrating on the encircled elements alone, we see that the top layer completely determines C_1 and G_1: they are $O(X)$. The first interface makes A_1, B_1, C_2, G_2 vanish asymptotically. The second interface makes A_2, B_2, C_3, G_3 vanish. The bottom makes A_3, B_3 vanish, and we have solved the equations in first approximation, valid for $k \to \infty$. It may be found that

$$\begin{aligned} jk(-C_1+G_1)&=\hat{\sigma}(k)\\ k(C_1-2G_1)&=\hat{\tau}(k) \end{aligned} \Rightarrow \begin{pmatrix}-1 & 1\\ 1 & -2\end{pmatrix}\begin{pmatrix}C_1\\ G_1\end{pmatrix} = \begin{pmatrix}\hat{\sigma}/jk\\ \hat{\tau}/k\end{pmatrix} \Longleftrightarrow \begin{pmatrix}-2 & -1\\ -1 & -1\end{pmatrix}\begin{pmatrix}\hat{\sigma}/jk\\ \hat{\tau}/k\end{pmatrix} = \begin{pmatrix}C_1\\ G_1\end{pmatrix}. \tag{C11}$$

Note that *all* field quantities $\to 0$ as $k \to +\infty$ when $z > 0$.

One can obtain a better approximation of the behavior of the A_i by using the value of C_1, G_1 we just obtained, viz. $O(X)$. Enter these into the equations of the first interface; they are of magnitude $O(Xe^{-kD_1})$. Next, one regards them as right-hand sides. We see:

$$A_1, B_1 = O(Xe^{-2D_1k})$$
$$C_2, G_2 = O(X).$$

We enter C_2 and G_2 into the second interface equation, and find

$$A_2, B_2 = O(Xe^{-2D_2k})$$
$$C_3, G_3 = O(X).$$

Finally, we enter C_3, G_3 into the bottom equation, and find

$$A_3, B_3 = O(Xe^{-2D_3k}).$$

Generally, it may be stated:

$$C_i,\ G_i=O(X),\ A_i,\ B_i=(Xe^{-2D_ik}),\ i=1,\ 2,\ 3,\ldots \tag{C12}$$

and

$$\begin{aligned}\hat{\mathbf{F}}(k,\ z)&=O(Xe^{k(z-2D_i)})+O(Xe^{-kz})\\ &=O(Xe^{-kz}),\ D_{i-1}\le z\le D_i,\ k\to\infty.\end{aligned} \tag{C13}$$

We emphasize once more that only the exponential behavior in k has been taken into account.

Table 1 The equations for a three layered medium with the conventions (C10)

Top	A_1	B_1	C_1	G_1		A_2	B_2	C_2	G_2		A_3	B_3	C_3	G_3		*knowns*
$\hat{\sigma}(k)$	+	+	+	+		0	0	0	0		0	0	0	0	=	X
$\hat{\tau}(k)$	+	+	+	+		0	0	0	0		0	0	0	0	=	X
1st interface																
$\hat{u}_1$	+1	+1	−1	−1	$\hat{u}_2$	+1	+1	−1	−1		0	0	0	0	=	0
$\hat{w}_1$	+1	+1	−1	−1	$\hat{w}_2$	+1	+1	−1	−1		0	0	0	0	=	0
$\hat{\sigma}_{z1}$	+1	+1	−1	−1	$\hat{\sigma}_{z2}$	+1	+1	−1	−1		0	0	0	0	=	0
$\hat{\tau}_{xz1}$	+1	+1	−1	−1	$\hat{\tau}_{xz2}$	+1	+1	−1	−1		0	0	0	0	=	0
2nd interface																
	0	0	0	0	$\hat{u}_2$	+2	+2	−2	−2	$\hat{u}_3$	+2	+2	−2	−2	=	0
	0	0	0	0	$\hat{w}_2$	+2	+2	−2	−2	$\hat{w}_3$	+2	+2	−2	−2	=	0
	0	0	0	0	$\hat{\sigma}_{z2}$	+2	+2	−2	−2	$\hat{\sigma}_{z3}$	+2	+2	−2	−2	=	0
	0	0	0	0	$\hat{\tau}_{xz2}$	+2	+2	−2	−2	$\hat{\tau}_{xz3}$	+2	+2	−2	−2	=	0
Bottom																
	0	0	0	0		0	0	0	0	$\hat{u}_3$	+3	+3	−3	−3	=	0
	0	0	0	0		0	0	0	0	$\hat{w}_3$	+3	+3	−3	−3	=	0

Remark. It may very well be that the boundary conditions possess no "bottom" equations. Under those circumstances, "$A_3 = B_3 = 0$" in Table 1 replaces the bottom equations. This will occur when the elastic layers are bonded to an elastic half space as substrate.

C3 The Numerical Inversion of the Fourier Integrals. We saw in equation (54) that for the inversion the following types of integral must be evaluated.

(54*a*) $\hat{f}(k)$ is even in $k \Rightarrow f(x)$

$$=\frac{1}{\pi}\int_0^\infty \hat{f}_R(k)\cos(kx)\,dk = \text{even in } x.$$

(54*b*) $\hat{f}(k)$ is odd in $k \Rightarrow f(x)$

$$=\frac{1}{\pi}\int_0^\infty \hat{f}_I(k)\cos(kx)\,dk = \text{odd in } x.$$

In both cases we split the interval of integration into two parts, viz. $[0,b]$ and $[b,\infty]$, where b will be chosen presently. It depends on f.

C3.1 The Integration Over $[0,b]$. We denote $\hat{f}_R$ and $\hat{f}_I$ both by g. We approximate g by a function g_a which is so that $\int_0^b g_a(k)(\cos(kx))\,dk$ can be calculated exactly. Examples of such functions are

— piecewise constant functions;
— piecewise linear, continuous functions;
— splines, i.e., piecewise cubic functions that are continuously twice differentiable.

The simplest are the piecewise constant functions. Let g_a be constant in the interval $((h-1)c, hc)$, with c constant, > 0, and value g_h, $h = 1, 2, 3, \ldots, q$, $qc = b$. Then

$$\int_{(h-1)c}^{hc} g_a(k)\cos(kx)\,dk$$

$$=\frac{1}{x}g_h\,\{\sin[hcx]-\sin[(h-1)cx]\}=$$

$$=\{(2g_h/x)\sin(cx/2)\}\cos\left[\left(h-\frac{1}{2}\right)cx\right] \qquad (C14a)$$

$$\int_{(h-1)c}^{hc} g_a(k)\sin(kx)\,dx$$

$$=\frac{1}{x}g_h\,\{\cos[(h-1)cx]-\cos[hcx]\}=$$

$$=\{(2g_h/x)\sin(cx/2)\}\sin\left[\left(h-\frac{1}{2}\right)cx\right]. \qquad (C14b)$$

Note that the (finite) first factor is common to all terms, while the second can be determined recursively:

$$\cos\left[\left(h-\frac{1}{2}\right)cx\right]=\cos\left[\left(h-\frac{3}{2}\right)cx\right]\cos(cx) - \sin\left[\left(h-\frac{3}{2}\right)cx\right]\sin(cx)$$

$$\sin\left[\left(h-\frac{1}{2}\right)cx\right]=\cos\left[\left(h-\frac{3}{2}\right)cx\right]\sin(cx) + \sin\left[\left(h-\frac{3}{2}\right)cx\right]\cos(cx). \qquad (C14c)$$

Consequently,

$$\int_0^{b=qh} g_a(k)\cos(kx)\,dk = \left\{\frac{2\sin(cx/2)}{x}\right\}\sum_{h=1}^{q} g_h\cos\left[\left(h-\frac{1}{2}\right)cx\right]$$

$$\int_0^{b=qh} g_a(k)\sin(kx)\,dk = \left\{\frac{2\sin(cx/2)}{x}\right\}\sum_{h=1}^{q} g_h\sin\left[\left(h-\frac{1}{2}\right)cx\right]$$

$\cos\left[\left(h-\frac{1}{2}\right)cx\right]$, $\sin\left[\left(h-\frac{1}{2}\right)cx\right]$ determined recursively (see equation (C14*c*)). (C15)

We can easily approximate a continuous g by a piecewise constant $g_a(k)$, by setting

$$g_h = g\left[\left(h-\frac{1}{2}\right)c.\right] \qquad (C16)$$

Call

$$\epsilon_h = \max_{k\in[(h-1)c,hc]} |g(k)-g_h|,\ \epsilon = b\max_h \epsilon_h \qquad (C17)$$

then the error of the inversion integral due to the integration over the finite interval $[0, b]$ is

$$\text{error}_{\text{finite part}} = \left|\int_0^b [g(k)-g_a(k)]\begin{pmatrix}\sin kx\\ \cos kx\end{pmatrix}dx\right| \le \left|\int_0^b [g(k)-g_a(k)]dk\right| \le \epsilon. \qquad (C18)$$

We observe that we have that $\hat{f}_R$ is continuous in $[0,\infty]$, as is $\{\hat{f}_I - (\hat{f}_{-1}/k)\}$, see equation (C9), with $\hat{f}_{-1}$ constant. The term $(\hat{f}_{-1}/k)$ which must be added to the latter function to obtain $\hat{f}_I$, can be inverted analytically:

$$\frac{1}{\pi}\int_0^\infty f_{-1}\sin(kx)\frac{dk}{k} = \frac{\text{sign}(x)}{\pi}\int_0^\infty \hat{f}_{-1}\sin(k|x|)\frac{dk}{k} = \frac{1}{2}\hat{f}_{-1}\,\text{sign}(x) \qquad (C19)$$

by Abramowitz-Stegun (1965), equations (5.2.1) and (5.2.5), so that

$$\int_0^\infty \hat{f}_I(k)\sin(kx)\,dx = \int_0^\infty \{\hat{f}_I - (\hat{f}_{-1}/k)\}\sin(kx)\,dk + \frac{1}{2}\hat{f}_{-1}\,\text{sign}(x). \qquad (C20)$$

C.3.2. The Integration Over $[b,\infty)$, *When* $z > 0$. We recall that by equation (C13) the Fourier transform of all field quantities is

$$\hat{\mathbf{f}}(k,z) = O(e^{-kz}) \Rightarrow f(k) = O(e^{-kz}) \text{as } k\to\infty \qquad (C21)$$

where we have omitted the factor X. We assume that we want the field quantities a fixed distance $z > 0$ below the surface.Then, if b is large enough, the behavior of

$$f(k) \approx f(b)\,e^{(b-k)z},\ b \le k \to \infty. \qquad (C22)$$

We given an estimate of the contribution to the Fourier integral by the tail of he integration, i.e.,

$$\int_b^\infty f(k)\begin{pmatrix}\sin(kx)\\ \cos(kx)\end{pmatrix}dk.$$

We have

$$|\int_b^\infty f(k)\begin{pmatrix}\sin kx\\ \cos kx\end{pmatrix}dk| \le \int_b^\infty |f(k)|dk$$

$$\approx \int_b^\infty |f(b)|e^{(b-k)z}dk = \frac{1}{z}|f(b)|.$$

As $f(k) = O(e^{-kz})$ as $k \to \infty$,

$$|\int_b^\infty f(k)\begin{pmatrix}\sin kx\\ \cos kx\end{pmatrix}dk| = O\left(\frac{1}{z}e^{-bz}\right). \tag{C23}$$

This goes to zero rather fast, so the procedure is:

(1) Choose ϵ;

(2) Determine b so that $\frac{1}{z}e^{-bz} = O(\epsilon)$;

(3) Determine c so that

$$\epsilon_h = \max_{k\in[(h-1)c,hc]} |f(k)-f_h| < \epsilon/b$$

by inspection of $f(k)$;

(4) Perform equation (C15). (C24)

C.3.3. The Integration Over $[b,\infty)$, *When* $z = 0$. According to Section C2, A_1 and B_1 are $O(e^{-D_1k})$, $k \to \infty$, and therefore the tail of the integration may be neglected when b is large enough. Also, it was shown that

$$\begin{pmatrix}C_1\\ G_1\end{pmatrix} = -\begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}\hat\sigma(k)/jk\\ \hat\tau(k)/k\end{pmatrix} \quad (|k|\to\infty),$$

$$\text{with} \quad \begin{matrix}\hat\sigma^n(k) = \sin(ka)/k\\ \hat\tau^t(k) = \sin(ka)/k.\end{matrix} \tag{C25}$$

We go a step further, and determine the three as yet unknown surface field quantities, viz. $\hat u(k)$, $\hat w(k)$, $\hat\sigma_x(k)$. Indeed we have, by equations (41*a*), (41*b*), (41*c*)

$$\begin{matrix}\nu=\nu_1\\ E=E_1\\ C=C_1,\ G=G_1\end{matrix} \quad \begin{pmatrix}\hat u(k)\\ \hat w(k)\\ \hat\sigma_x(k)\end{pmatrix} = \begin{pmatrix}[-C+(3-2\nu)G](1+\nu)/E\\ j[C-2\nu G](1+\nu)/E\\ -jk[-C+3G]\end{pmatrix}. \tag{C26}$$

From a comparison of equations (C25) and (C26) we find

$$\begin{pmatrix}\hat u(k)\\ \hat w(k)\\ \hat\sigma_x(k)\end{pmatrix} = \begin{pmatrix}-H(1-2\nu) & -2H(1-\nu)\\ -2jH(1-\nu) & -jH(1-2\nu)\\ jk & 2jk\end{pmatrix}\begin{pmatrix}\hat\sigma(k)/jk\\ \\ \hat\tau(k/k\end{pmatrix} \quad H \underset{=}{def} \frac{1+\nu}{E}, \tag{C27}$$

so that we see, since $\hat\sigma(k) = \theta\sin(ka)/k$, $\hat\tau(k) = (1-\theta)\sin(ka)/k$, $\theta = 0$ or 1, that

$$|\hat u(k)| \le 2H(1-\nu)/k^2, \ |\hat w(k)| \le 2(-\nu)/k^2,$$

$$|\hat\sigma_x(k)| \le 2/k \text{ as } k\to\infty. \tag{C28}$$

As to $\hat u(k)$ and $\hat w(k)$, we propose to neglect the tail. The procedure is as in equation (C24), where (C24.2) is replaced by

2′) Determine b so that $2H(1-\nu)/b < \epsilon$. (C29)

The tail of the inversion of $\hat\sigma_x(k)$ can be expressed in sine and cosine integrals (see A-S) (= Abramowitz-Stegun, 1965) Chapter 5. Indeed,

$$\text{Tail}\begin{pmatrix}\mathrm{Re}\hat\sigma_x(k)\\ \mathrm{Im}\hat\sigma_x(k)\end{pmatrix} \simeq \frac{1}{\pi}\int_b^\infty \frac{\sin(ka)}{k}\begin{pmatrix}\cos kx\\ \sin kx\end{pmatrix}dx =$$

$$= \frac{1}{2\pi}\int_b^\infty \frac{1}{k}\begin{pmatrix}\sin[k(a+x)]+\sin[k(a-x)]\\ \cos[k(a+x)]-\cos[k(a-x)]\end{pmatrix}dk =$$

$$= -\frac{1}{2\pi}\begin{pmatrix}Si[b(a+x)]+\mathrm{sgn}(a-x)Si[b|a-x|]\\ Ci[b|(a-x|]-Ci(b|a+x|)\end{pmatrix},$$

$$\begin{matrix}\mathrm{sgn}(t) = 1 \text{ if } t>0\\ = -1 \text{ if } t<0\end{matrix} \tag{C30}$$

where

$$Si(z) \underset{=}{def}\int_z^\infty \frac{\sin t}{t}dt = \frac{\pi}{2} - \sum_{n=0}^\infty \frac{(-1)^n z^{2n+1}}{(2n+1)!(2n+1)}$$

(A-S 5.2.14)

$$= f(z)\cos z + g(z)\sin z;\ f,\ g \text{ (see below)}$$

(A-S 5.2.8)

$$Ci(z) \underset{=}{def} -\int_z^\infty \frac{\cos t}{t}dt = +\gamma + \ln(z) + \sum_{n=1}^\infty \frac{(-1)^n z^{2n}}{(2n)!(2n)}$$

$\gamma = 0.57722\ldots$ Euler's Constant

$$= f(z)\sin(z) - g(z)\cos(z)$$ (A-S 5.2.9)

with

$$f(z) = z^{-1}(z^4+a_1z^2+a_2)/(z^4+b_1z^2+b_2)+\eta(z),$$

$1 \le z < \infty$; $|\eta(z)| < 2\times 10^{-4}$

$$g(z) = z^{-2}(z^4+a_1'z^2+a_2')/(z^4+b_1'z^2+b_2')+\eta'(z),$$

$1 \le z < \infty$ $|\eta'(z)| < 10^{-4}$

$a_1 = 7.241163$, $a_2 = 2.463936$; $b_1 = 9.068580$, $b_2 = 7.157433$ (A-S 5.2.36)

$a_1' = 7.547478$, $a_2' = 1.564072$, $b_1' = 15.723606$, $b_2' = 12.723684$. (A-S 5.2.3.7)

The procedure is as in (C24), but (C24.2) is replaced by 2″. Determine b so that A_1 and B_1 are negligible, that is

$$e^{-D_1b} = O(\epsilon). \tag{C31}$$

Reference

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

APPENDIX D

Determination of the Position of the Ends of the Contact Area

In many applications it is of interest to know the coordinate of the left and right ends of the contact area. A crude estimate of these coordinates is found by equating them to the outermost boundaries of the elements in contact. However, a better estimate is possible. To find it, consider the normal traction $\sigma(x)$. It has a vertical tangent near the ends of the contact area, and in fact $\sigma(x)$ behaves like $p|x-z|^{1/2}$ inside the contact area, where z denotes the x-coordinate of the end-point under consideration. Here, p is a slowly-varying function of x inside the contact, and near z.

To find z, we propose to take p constant. Consider the two abscissa x_1 and x_2 where the traction σ is known: σ_1 and σ_2; we assume $\sigma_1 > 0$, $\sigma_2 > 0$, $\sigma_1 \neq \sigma_2$. x_1 and x_2 are best taken closest to z, a crude estimate of which is given above. For best signficance, x_1 and x_2 are each identified with the center of an element. Then we have:

$$\sigma_1 = p|x_1 - z|^{1/2},\ \sigma_2 = p|x_2 - z|^{1/2}$$

$$\Rightarrow \quad \sigma_1^2/\sigma_2^2 = (x_1 - z)/(x_2 - z)$$

$$\Rightarrow \quad z = (\sigma_1^2 x_2 - \sigma_2^2 x_1)/\sigma_1^2 - \sigma_1^2). \tag{D1}$$

Example: $x_1 = 1,\ x_2 = 4;\ \sigma_1 = 1,\ \sigma_2 = 2 = \Rightarrow z = 0.$

APPENDIX E

A Note on Viscoelasticity

E1 Elastostatics and viscoelastics are governed by

$$e_{ij}(x_\ell) = \frac{1}{2}\{u_{i,j}(x_\ell) + u_{j,i}(x_\ell)\}: \text{ compatibility equations} \tag{E1a}$$

x_ℓ:Cartesian coordinates, $i, j, k, h, \ell = 1, 2, 3$ (E1b)

e_{ij} = linearized strain $= e_{jj}$ (E1c)

u_j = displacement component (E1d)

$_{,j} = \partial/\partial x_j$; summation convention (E1e)

and by

$$\left.\begin{aligned} &Q^f(r) = \int_{-\infty}^{\infty} Q(t)\, e^{jrt} dt,\ R^f(r) = \int_{-\infty}^{\infty} R(t)\, e^{jrt} dt \\ &\text{and} \\ &S^f(r) = Q^f(r)\, R^f(r), \qquad S^f(r) = \int_{-\infty}^{\infty} S(t)\, e^{jrt} dt, \\ &\text{then} \\ &S(t) = \int_{-\infty}^{\infty} Q(t-\tau)\, R(\tau) d\tau = \int_{-\infty}^{\infty} Q(\tau)\, R(t-\tau)\, d\tau \ \underset{=}{\text{notation}} Q^*R = Q. \end{aligned}\right\} \tag{E10}$$

$\sigma_{ij,j} = 0$; equilibrium without inertia and body force (E2a)

σ_{ij}: stress component, $\sigma_{ij} = \sigma_{ji}$. (E2$b$)

E2 Moreover, we have constitutive relations; for elasticity they are:

$$\sigma_{ij} = E_{ijhk} e_{hk},\ E_{ijhk} = E_{jihk} = E_{hkij}. \text{ Hooke's law} \tag{E3a}$$

$$e_{hk} = S_{hkij}\sigma_{ij},\ S_{hkij} = S_{khij} = S_{ijhk}. \tag{E3b}$$

The symmetries (E1c), (E2b), (E3a), (E3b) entail that there are only six independent stress and strain components

$$\sigma_I(x_i),\ e_I(x_i),\ I = 1, \ldots, 6 \tag{E4a}$$

(3a) and (3b) become

$$\sigma_I = E_{IJ} e_J,\ e_J = S_{IJ}\,\sigma_J,\ (E_{IJ}),\ (S_{IJ}) > 0.0. \tag{E4b}$$

In case of isotropy, we have

$$e_{ij} = \frac{1+\nu}{E}\sigma_{ij} - \frac{\nu}{E}\sigma_{hk}\,\delta_{ij} \tag{E5a}$$

$$\delta_{ij} = 0 \text{ if } i \neq j,\ \delta_{ij} = 1 \text{ if } i = j \tag{E5b}$$

E: Young's modulus, ν: Poisson ratio (E5c)

E3 In viscoelasticity, the constitutive relations are time dependent, as well as the viscoelastic field; Hooke's law reads:

$$\sigma_I + S_{IJ}^{(1)}\dot{\sigma}_J + \ldots + S_{IJ}^{(n)} d^n\sigma_J/dt^n = \\ = E_{IJ}^{(0)}\dot{e}_J + E_{IJ}^{(1)}\dot{e}_J + \ldots + E_{IJ}^{(m)} d^m e_J/dt^m \tag{E6a}$$

$m,n \in \{0,1,2, \ldots\}$; $(\bullet) = d/dt$, material derivative with time t. (E6b)

$S_{IJ}^{(k)}$, $E_{IJ}^{(k)}$ are time independent parameters. (E6c)

In case of homogeneity, all (visco) elastic parameters are independent of the x_i.

E4 We can reduce the viscoelastic field to an elastic field by applying a complex Fourier transform (FT).

The FT of a function $H(x_i,t)$ is given by

$$H^f(x_i,r) = \int_{-\infty}^{\infty} H(x_i,\, t)\, e^{jrt} dt,$$

$$j: \text{ imaginary unit, } j^2 = -1. \tag{E7}$$

The following inversion formula holds:

$$H(x_i,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H^f(x_i,\, r)\, e^{-jrt}\, dr \tag{E8}$$

while

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial t} H(x_i,\, t)\, e^{jrt} dt = -jr\, H^f(x_i,r). \tag{E9}$$

The convolution theorem holds:

Q^*R is the *convolution* of Q and R.

E5 We apply the *FT* to (E6a):

$$\{\delta_{IJ} + (-jr)\, S_{IJ}^{(1)} + \ldots + (-jr)^n\, S_{IJ}^{(m)}\}\sigma_J^f(x_\ell,r) = \\ = \{E_{IJ}^{(0)} + (-jr)\, E_{IJ}^{(1)} + \ldots + (-jr)^m\, E_{IJ}^{(m)}\} e_J^f(x_\ell,r). \tag{E11}$$

This can be written as

$$e_J^f(x_\ell,r) = (E_{IJ}^{(0)} + (-jr)\, E_{IJ}^{(1)} + \ldots + (-jr)^m\, E_{IJ}^{(m)})^{-1} \times \\ \times (\delta_{IK} + (-jr) S_{IK}^{(1)} + \ldots + (-jr)^n S_{IK}^{(n)} \sigma_K^f)\sigma_K^f(x_\ell,r) \\ \underset{=}{def}\ S_{JK}^f(r)\, \sigma_K^f(x_\ell,r) \tag{E12a}$$

or

$$e^f_{ij}(x_\ell,r)=S^f_{ijhk}(r)\sigma^f_{hk}(x_\ell,r);\; e^f_{ij}=e^f_{ji};\; \sigma^f_{hk}=\sigma^f_{kh}. \quad \text{(E12}b\text{)}$$

We also have, by (E1*a*) and (E2*a*),

$$e^f_{ij}=\frac{1}{2}\,(u^f_{j,i}+u^f_{i,j}) \quad \text{(E12}c\text{)}$$

$$\sigma^f_{ij,j}=0. \quad \text{(E12}d\text{)}$$

Clearly, $u^f_i\,(x_\ell,\,r)$; $e^f_{ij}\,(x_\ell,\,r)$; $\sigma^f_{ij}\,(x_\ell,\,r)$ form a compatible, equilibrium Hookean elastic field for every value for the Fourier parameter r. Therefore, this transformed elastic field obeys all elastic laws, so that when the elastic solution of a problem may be found, an inverse transform yields the corresponding viscoelastic solution. This principle is called the *correspondence principle of linear viscoelasticity*. It is a well-known principle.

E6 An application will be given of the correspondence principle. If the material is homogeneous and isotropic, the stress-strain relations (E12*b*) become

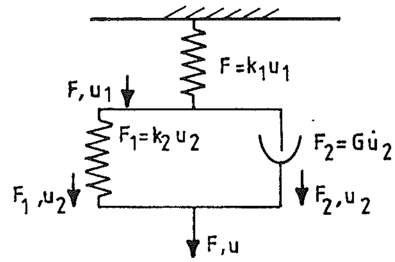
$$e^f_{ij}(x_\ell,r)=\frac{1+\nu(r)}{E(r)}\,\sigma^f_{ij}-\frac{\nu(r)}{E(r)}\,\sigma^f_{hk}\,\delta_{ij}. \quad \text{(E13)}$$

The material is incompressible if and only if $\nu\,=\,0.5$. So, often one takes $\nu(r)\,=$ constant, albeit not necessarily 0.5. Then equation (E13) becomes

$$e^f_{ij}(x_\ell,r)=\{(1+\nu)\,\sigma^f_{ij}-\nu\,\sigma^f_{hk}\,\delta_{ij}\}/E(r). \quad \text{(E14)}$$

Assume now

$$E(r)=(1-jqr)/(K-jqQr) \quad \text{(E15}a\text{)}$$



F: force, u_1, u_2, u: displacement $u=u_1+u_2$;
K_1: spring constant, $1/Q$, see equation (E15*b*)
K_2: spring constant, $1/(K-Q)$, see equation (E15*c*)
G: damper constant, $q/(K-Q)$, see equation (E15*d*)
$D.E.$: $u+q\dot{u}\;=\;KF\;+\;qQ\dot{F}\;u \Longleftrightarrow e_{ij},\,F \Longleftrightarrow \sigma_{ij}$.

Fig. E1 Two-spring, one-damper model of a viscoelastic solid

$$Q\text{: ``initial compliance,'' } Q>0 \quad \text{(E15}b\text{)}$$

$$K\text{: ``final compliance,'' } K\geq Q>0 \quad \text{(E15}c\text{)}$$

$$q\text{: relaxation time, } q>0. \quad \text{(E15}d\text{)}$$

This form is based on the two-spring, one-damper model of a viscoelastic one-dimensional solid (see Fig. E1).

So we may write:

$$(1-jqr)\,e^f_{ij}(x_\ell,r)=(K-jQqr)\{(1+\nu)\sigma^f_{ij}-\nu\sigma^f_{hk}\delta_{ij}\} \quad \text{(E16}a\text{)}$$

which we may transform back to obtain the constitutive equations of homogeneous, isotropic viscoelasticity of the type - Fig. E1:

$$e_{ij}+q\;\dot{e}_{ij}=(K+qQ\;d/dt)\{(1+\nu)\sigma_{ij}-\nu\;\sigma_{hk}\delta_{ij}\}. \quad \text{(E16}b\text{)}$$

K. M. Liechti
Associate Professor,
Mem. ASME.

Y.-S. Chai[1]

Department of Aerospace Engineering
and Engineering Mechanics,
The University of Texas at Austin,
Austin, TX 78712

Biaxial Loading Experiments for Determining Interfacial Fracture Toughness

The paper establishes the range of in-plane fracture mode mixtures and contact zone sizes that can be obtained from an edge-cracked bimaterial strip under biaxial applied displacements. The development of a suitable loading device for and the application of crack opening interferometry to interfacial crack initiation experiments is described. The crack initiation process under bond-normal loading is examined in detail for a glass/epoxy interface in order to establish a hybrid optical interference/finite element analysis technique for extracting mixed-mode fracture parameters.

1 Introduction

It has become increasingly clear that the fracture resistance of composite materials can be strongly affected by the toughness of the interface between constituents. The reliability of microelectronic devices, which may contain a large number of different interfaces, may also be compromised by their toughness. The same may also be true of structural adhesively bonded joints and coatings subjected to hostile environments. If a crack is constrained to grow along the interface, then the growth is inherently mixed mode in nature and a suitable parameter must be found that characterizes critical and subcritical growth over a range of mode mixes. The purpose of this paper is to describe the examination and development of a method for providing mixtures of mode I and II over a wide range of mode mixes.

Although any interfacial fracture test will, in general, involve some mode mix, a series of specimens loaded in different ways, a single specimen under biaxial load or a change in delamination shape will usually be required to determine interfacial toughness over a range of mode mixes. The first strategy was recognized early by Malyshev and Salganik (1965) and Gent and Kinloch (1971) and later by Takashi et al. (1978), but fracture mode mixes were not explicitly extracted. Trantina (1972) using scarf joints, Anderson, DeVries, and Williams (1974) using cone, peel, and blister specimens, Liechti and Hanson (1988) using blister specimens, Cao and Evans (1988) using symmetric and asymmetric double cantilever beams, four-point flexure (Charalambides et al., 1989a) and composite cylinder (Charalambides and Evans, 1989b), and Rosenfeld et al. (1990) introducing the microindentation test all used finite element analyses to extract fracture mode mixtures. Analytical stress intensity factor solutions were obtained for blister specimens (Arin and Erdogan, 1971), sandwich specimens (Suo and Hutchinson, 1989), and brazil nut sandwiches (Wang and Suo, 1990). Single specimens under multiaxial loads were employed by Mulville et al. (1978) and Liechti and Knauss (1982a,b) and suggested by Suo and Hutchinson (1990). Finally, in the realm of thin coatings, a clever use of residual stresses has been made in determining the effect of mode III on interfacial toughness by examining the shape of the delamination emanating from a straight cut made through the coating to the substrate interface (Jensen et al., 1990). A simplified analysis for extracting three-dimensional mode mixes from curved delamination fronts in thin films has recently been presented by Chai (1989).

The approach that was chosen here for obtaining a wide range of mode mixes was to use a single specimen in conjunction with a biaxial loading device. The stress analysis of the specimen and loading is considered first in order to establish the potential mode mix range and crack-face contact effects. The development of the biaxial loading device and the measurement of normal crack opening displacements (NCOD) is then described. A hybrid procedure for extracting stress intensity factors based on the measured NCOD and complementary finite element analyses is then discussed with reference to crack initiation under some initial experiments bond-normal loading. The results of a series of experiments over a wide range of mode mixes are presented in a companion paper (Liechti and Chai, 1989).

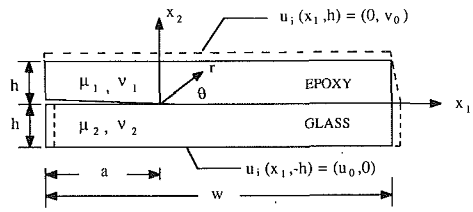
2 Specimen Geometry and Analysis

The choice of specimen geometry was motivated by a number of factors. First, it was desirable to have a specimen that gave rise to crack-length independence of fracture parameter and mode mix. This feature simplifies data reduction, particularly for crack propagation studies and allows cracks to be initiated and arrested by suitable control of the loading. The use of a single specimen minimizes variations in surface preparations

[1] Currently Assistant Professor, Yeungnam University, Seoul, South Korea.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Nov. 10, 1989; final revision, Apr. 4, 1990.



w = 17.78 cm   h = 1.27 cm

Fig. 1 The edge-cracked bimaterial strip specimen

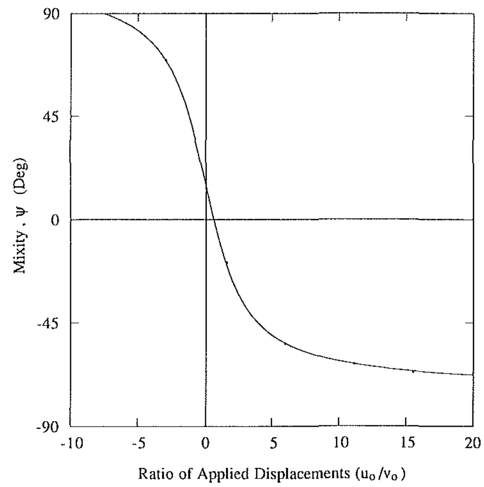


Fig. 2 The range of mixity available under positive bond-normal displacements

which affect the intrinsic adhesion or toughness which, in turn, control the overall toughness (Argon et al., 1989). It also means that the specimen should be amenable to biaxial loading. Continuing the desire the make measurements of NCOD near the crack front in order to assess the importance of nonlinear, three-dimensional, and crack-face contact effects (Liechti and Knauss, 1982; Liechti and Hanson, 1988) required that at least one material be transparent. Glass was chosen for this work, but transparency need not be limited to the visible spectrum. In view of these considerations, the specimen geometry and loading that was adopted was the edge-cracked bimaterial strip shown in Fig. 1.

The homogeneous strip geometry is well known for its linear compliance versus crack-length relation for sufficiently long cracks (Knauss, 1966; Rice, 1967). The extension to the bimaterial case under bond-normal loading has been made by Atkinson (1977) and energy arguments yield the bond-tangential contribution so that

$$G=\frac{v_0^2}{h}\left[\frac{(1-2\nu_1)}{\mu_1(1-\nu_1)}+\frac{(1-2\nu_2)}{\mu_2(1-\nu_2)}\right]^{-1}+\frac{u_0^2}{2h}\left[\frac{1}{\mu_1}+\frac{1}{\mu_2}\right]^{-1} \quad (1)$$

in the notation of Fig. 1.

From the analysis by Knauss (1966) we expect the steady-state solution (1) to be valid for $a/h>2$. This expectation was verified by finite element analysis (Chai, 1990). However, a more important contribution of the finite element analysis was in the extraction of the mode mix associated with any particular combination of materials and u_0 and v_0. The definitions of complex interfacial stress intensity factor K, bimaterial constant, ϵ, etc., that was used in this work follow those given by Rice (1988).

The mode mix or mixity, ψ, was taken to be

$$\psi=\tan^{-1}\left\{\frac{\mathrm{Im}[Kh^{i\epsilon}]}{\mathrm{Re}[Kh^{i\epsilon}]}\right\}. \quad (2)$$

Table 1 Material properties

MATERIAL TYPE	YOUNG'S MODULUS, $E(GPa)$	POISSON'S RATIO, ν	σ_0 (MPa)	HARDENING EXPONENT (n)
Epoxy	2.07	.37	34.5	5
Glass	68.9	.20	-	-

Epoxy-glass Dundurs' parameter $\alpha = -0.935$
$\beta = -0.188$
Bimaterial constant $\varepsilon = +0.0604$

Ramberg Osgood Representation:

$$\varepsilon=\sigma/E+\frac{3}{7}\left(\frac{\sigma}{\sigma_0}\right)^{n-1}$$

Table 2 Energy release rates under bond-normal and tangential displacements

u_0 (μm)	v_0 (μm)	$G(J/m^2)$ NUMERICAL	$G(J/m^2)$ ANALYTICAL	ψ (DEG)
1.27	0	0.04669	0.04673	-74.03
0	1.27	0.22160	0.22160	16.00

The energy release rate can also be obtained from K and its complex conjugate $\bar{K}$ through

$$G=\left[\frac{(1-\nu_1)}{\mu_1}+\frac{(1-\nu_2)}{\mu_2}\right]\frac{K\bar{K}}{4\cosh^2\pi\epsilon}. \quad (3)$$

Following a comparison (Ginsburg, 1987) of techniques for extracting mixed-mode interfacial fracture parameters based on crack opening displacements (Smelser, 1979), virtual crack closure (Raju, 1986) and a conservation integral approach (Yau and Wang, 1984), the latter was found to be most satisfactory and was incorporated as a post-processing routine in the finite element code VISTA (Becker et al., 1984). The auxiliary solutions required for the technique were taken from the paper by Smelser (1979), taking into account the stress intensity factor definition in Rice (1988). The invariance of energy release rate and mixity over the range of crack lengths used in the experiments was established for unit applied displacements normal and tangential to the interface. For the same displacement level (Table 2), the bond-normal displacements give rise to an energy release rate that is approximately four times higher than that produced by tangential displacements. This can also be seen from equation (1) which differs from the finite element solutions shown by less than 1 percent.

Under some general combination of applied bond-normal and bond-tangential displacements, the real and imaginary parts of the complex stress intensity factor can be written as

$$K_1=aK_1^{(u0)}+bK_1^{(v0)} \quad (4)$$

$$K_2=aK_2^{(u0)}+bK_2^{(v0)} \quad (5)$$

where $K_i^{(u0)}$ and $K_i^{(v0)}$, $i=1,2$ are the base stress intensity factors due to unit applied displacements tangential and normal to the interface, respectively, and the coefficients a and b are load factors. In view of the crack length invariance of K and (4) and (5), only two finite element analyses are required in order to map out the spectrum of mixities that can be obtained from the geometry and loading shown in Fig. 1.

The range of mixities that can be obtained for $v_0>0$ are shown in Fig. 2. Pure bond-normal displacements ($u_0=0$) give rise to a mixity of 16 deg, bringing out the mismatch between the glass and epoxy elastic properties (Table 1). A 1:1 ratio of bond-tangential to bond-normal displacements is required to produce $\psi=0$ deg, whereas a -7:1 ratio gives rise to $\psi=90$ deg. The mixity does not drop much below -60 deg for $u_0/v_0>20$. Thus, for positive bond-normal displacements, the range of mixities is essentially -60 deg$<\psi<90$ deg.

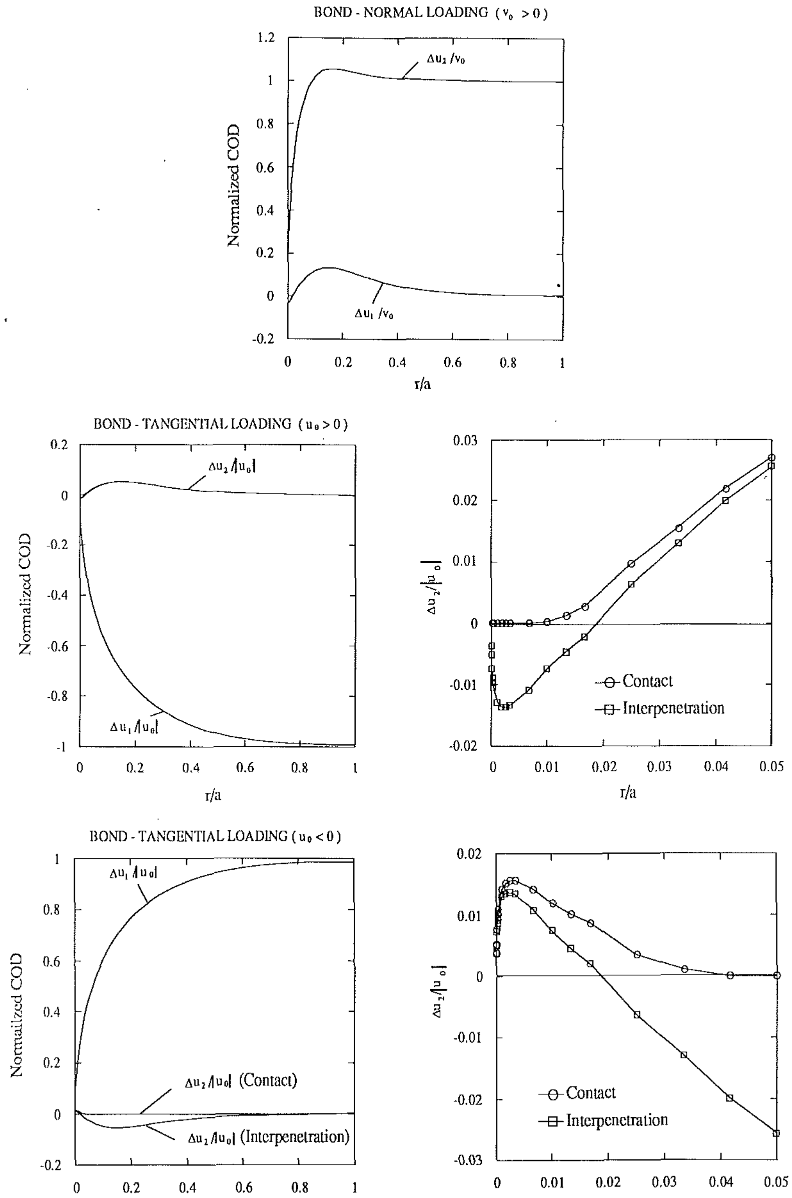


Fig. 3 Crack opening displacements under various loadings

Another interesting aspect of the proposed specimen geometry and loading is the extent of crack-face contact. For two semi-finite bodies with a central interface crack, Comninou (1978) found that, under a shear load, frictionless contact could occur over as much as 33 percent of the crack length. Even larger contact zones are possible for compression and shear, although complete contact can never occur. Experimental evidence of these trends has also been provided (Liechti and Knauss, 1982).

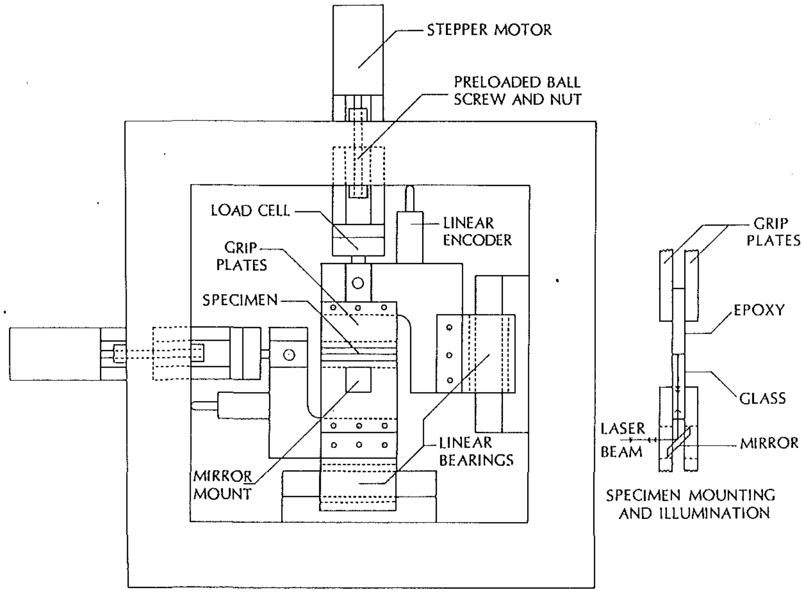
The stress analysis for this portion of the work was conducted with the ABAQUS finite element code[3], making use of special gap elements to eliminate interpenetration of crack faces. The response of the glass and epoxy was considered to be linearly elastic using the properties noted in Table 1. The size of the smallest elements surrounding the crack tip was $2\times10^{-4}h$. The components, Δu_i, of the displacement jump across the crack faces were taken to be

$$\Delta u_i = u_i^{(1)} - u_i^{(2)} \tag{6}$$

[3]The permission to use ABAQUS under academic license, granted by Hibbit, Karlsson, and Sorensen, Inc., is gratefully acknowledged.

where the superscripts (1) and (2) refer to the epoxy and glass, respectively. Under bond-normal loading (Fig. 3), Δu_2 was always positive, implying no crack-face contact within the resolution of the mesh. However, the tangential crack opening was negative over a small region ($r/a<0.02$). Positive bond-tangential displacements gave rise to some contact (Fig. 3) near the crack tip and mouth ($r/a=1$). The near-tip contact zone was $0.007r/a$. When the constraint was removed to allow interpenetration of crack faces, the near-tip interpenetration region was more than double the contact zone. Comninou (1978) also noticed that contact zones were smaller than interpenetration zones. For negative bond-tangential loading, Fig. 3 and its insert indicate that there was some opening at the crack tip but the crack faces were in contact over most of the crack length (96 percent). When interpenetration was allowed, the open region was again crack length (96 percent). When interpenetration was allowed, the open region was again smaller ($r/a<0.02$).

The noted differences between sizes of the contact zones and interpenetration regions did not give rise to any variations in energy release rate values. This is probably due to the as-



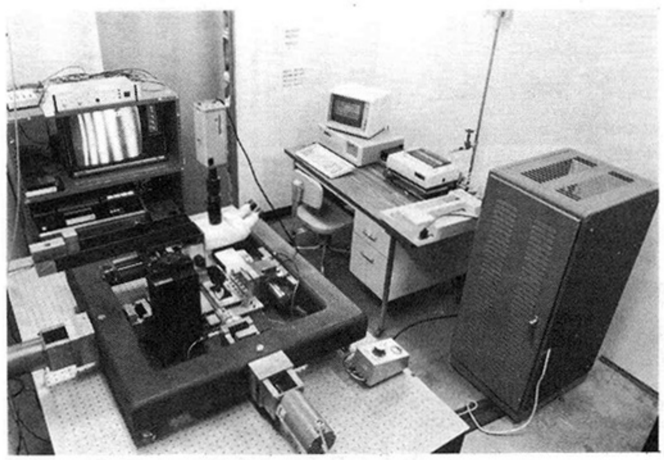


Fig. 4 Biaxial loading device

sumption of frictionless contact, although an analysis of cohesive mode II cracks in an adhesive layer did not reveal much change in energy release rate when frictional contact was allowed (Liechti and Freda, 1989). Moreover, energy release rate values calculated using the conservation integral approach and crack-opening displacements (VISTA) and virtual crack extension (ABAQUS) were all within 1 percent of the values obtained from (1). A positive bond-normal applied displacement yielded positive K_1 and K_2 values. Surprisingly, the conservation integral calculation indicated that $K_1 > 0$ for positive bond-tangential displacements, in spite of the local crack-tip closure (Fig. 3). The positive K_1 value may have been due to the fact that the contours were evaluated in regions where the crack was opening. The K_2 value was negative under $u_0 > 0$, which seems reasonable. All crack initiation experiments described later involved combinations of u_0 and v_0 that gave rise to a total $K_1(4)$ that was positive at initiation. Under combined tension and shear, Comninou and Schmueser (1979) found that K_2 was a nonlinear function of load ratio due to variations in contact lengths. As a result, one would think that the superpositions in (4) and (5) are invalid. However, the nonlinearity did not appear in the edge-cracked bimaterial strip in the sense that the energy release rate was linearly proportional to $(u_0^2 + v_0^2)$ even when crack contact was allowed.

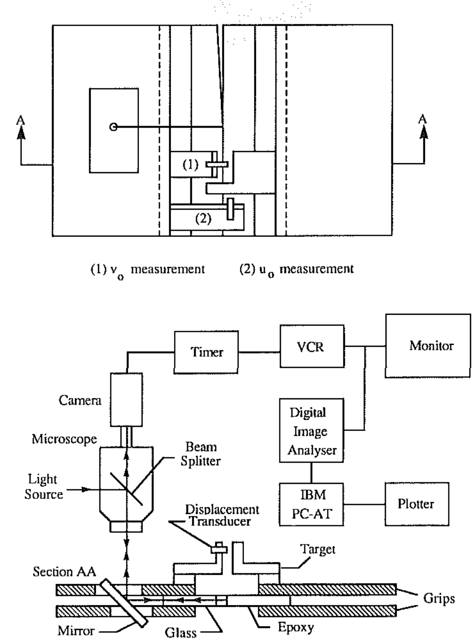


Fig. 5 Measurements of applied and normal crack opening displacements

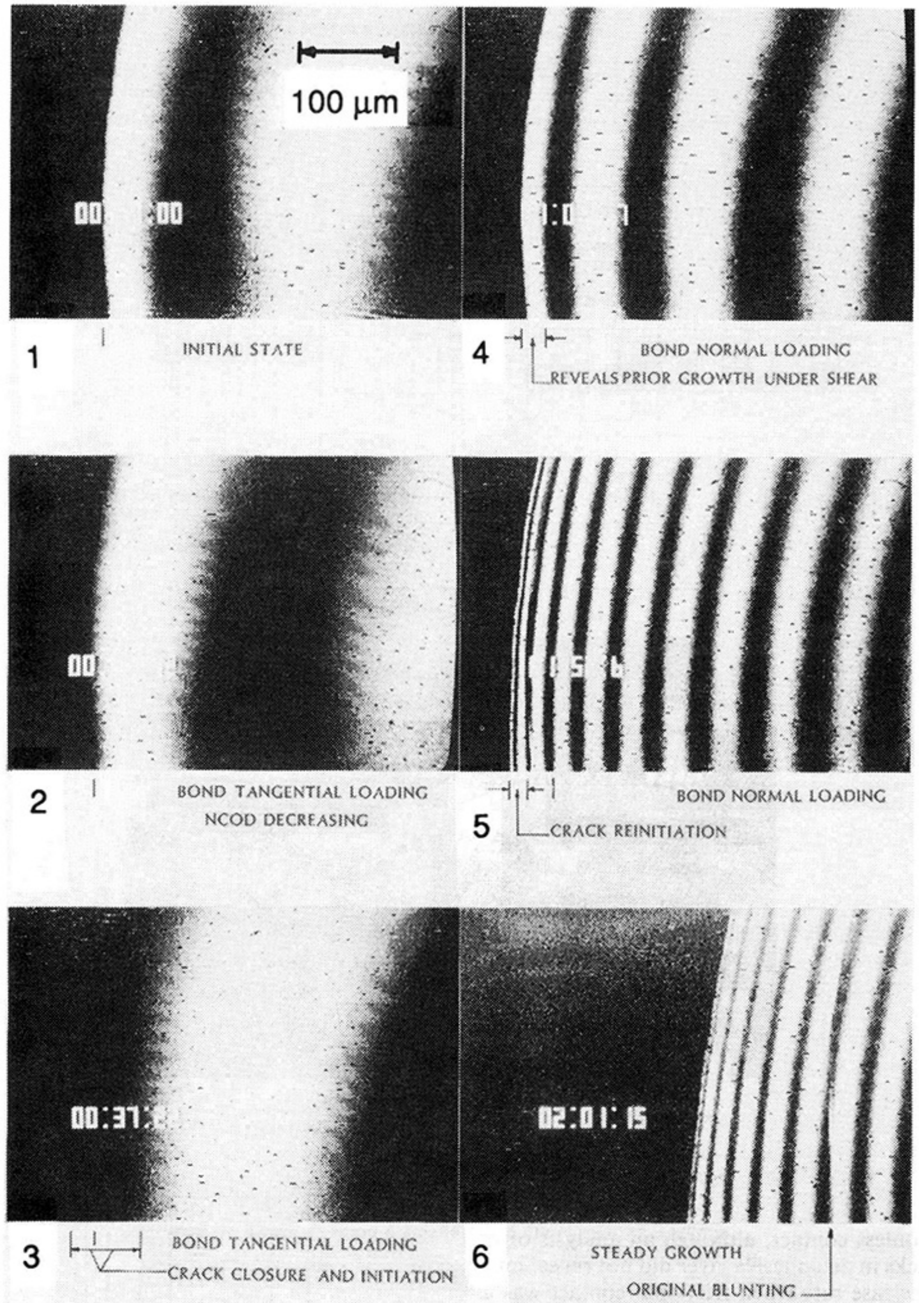


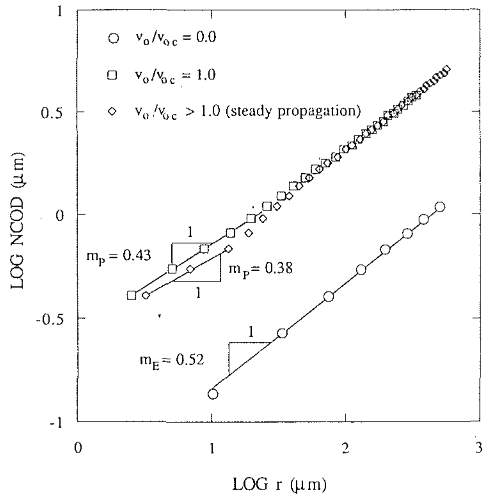
Fig. 6 Crack-face contact and extension under a sequential loading: bond-tangential displacements followed by bond-normal displacements

3 Experimental Aspects

From the preceding analysis it is clear that the development of a special biaxial loading device was required. The relatively low toughness of interfaces led to the additional requirement that the applied displacements be controlled with high resolution. Furthermore, since with slightest addition of bond-normal displacements removes the contact zones that arise under bond-tangential displacements, it was important to minimize any interaction (crosstalk) between the two loading modes. The stiffness of the loading device had to be high enough that the crack initiation process (slow extension) and steady growth could be examined. Finally, the desire to measure normal crack opening displacements (NCOD) in order to examine nonlinear and three-dimensional effects in the crack-tip region meant that microscope access be provided.

A schematic of the loading device is shown in Fig. 4 and a more detailed description is given in Chai (1990). The stepper motors, ball screws, and optical encoders were used in a computer-controlled feedback loop to provide independent displacement control to a resolution of 1.27μm. Stiff load cells measured the reactions normal and tangential to the interface. The relative displacements of the clamped boundaries of the specimens were also measured in the two directions using miniature capacitative displacement transducers having a range of 0.5 mm and resolution of 0.5 μm (Fig. 6). All transducers were calibrated in accordance with manufacturer's specifications and no drift was observed over periods much longer than the time scale of the experiments.

In order to examine near-tip asymptotics, crack-face contact, and three-dimensional effects, it was necessary to make measurements in the crack-front region in addition to globally applied displacements and their associated reactions. Crack opening interferometry, which has revealed interesting non-



(a) Crack Tip NCOD Asymptotics

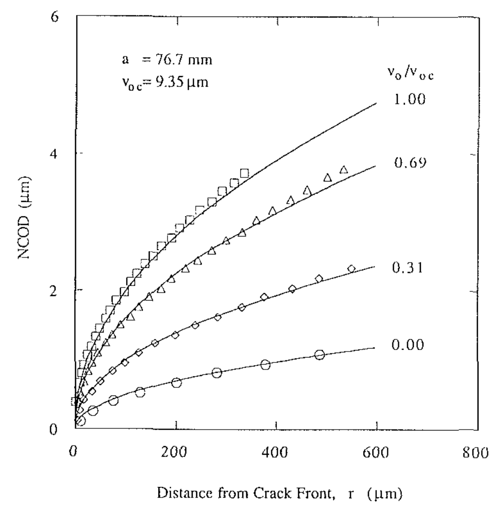


Fig. 7 Crack-tip asymptotics and comparisons with finite element analysis

linear and three-dimensional effects in the past (Liechti and Knauss, 1981, 1982b; Liechti and Hanson, 1988), was therefore employed in this study.

A 45 deg mirror was mounted in one of the grips to introduce a beam of monochromatic light through the glass for reflection by the crack faces. The reflected beams interfere and the resulting fringe patterns were resolved by a microscope with a 100-mm working distance. For the normal incidence and air-filled crack used here, a dark fringe of order m is a contour of NCOD, $\Delta\mu_2$, given by

$$\Delta\mu_2 = m\lambda/2. \tag{7}$$

The wavelength λ was 546-nm, yielding a resolution per half fringe (bright to dark) of 0.137 μm. The field of view was approximately 0.5 mm and fringes could be located to within 5 μm. The fringe patterns were recorded through a video camera and timer onto a high resolution video cassette recorder. The recordings were later analyzed using a digital image analysis system to obtain light intensity profiles along the center of the specimen. Because the fringe patterns were recorded for all times, there was no possibility of ambiguity in assigning fringe numbers and signs. The intensity profiles were filtered prior to thresholding in order to determine fringe locations. The whole procedure and fringe counting was automated so that NCOD profiles could be obtained every 1/30 of a second if necessary. The video timer allowed the NCOD measurements to be synchronized with those of the applied displacements and reactions.

A series of video frames is shown in Fig. 6 to illustrate crack-face contact and propagation under sequential loading in which positive bond-tangential displacements were applied first, followed by bond-normal applied displacements. The first frame, taken at a slight preload, indicated some initial opening. The crack front was convex in the direction of crack growth and, with proper specimen alignment, was symmetric with respect to the specimen midthickness. In the second frame, the fringe density decreased (indicating decreasing NCOD) under positive bond-tangential loading. However, crack closure had not yet occurred, due to the slight preload. Crack closure can be seen in the third frame which corresponds to the maximum level of bond-tangential displacements that were applied. Although it is not possible to discern crack extension under contact, the fourth frame, taken just after bond-normal displacements were initiated, indicates that such crack extension had indeed occurred. This interpretation of events is based on the fact that the crack faces immediately opened up to the new crack length upon application of bond-normal and further crack extension did not occur until some time later (frame 5). A slight discontinuity in the fringe pattern reveals some blunting at the original crack front. The sixth frame was taken during steady crack growth. Although detailed analyses have yet to be conducted, it appears that steady growth at a given load combination is characterized by a constant fringe spacing or NCOD profile.

Although quantitative comparisons between measured and predicted contact zone sizes will follow in a companion paper, the contact zone shown in Fig. 6 was indeed small as was predicted in Fig. 3(b). Chiang et al. (1988, 1989) did not observe any contact zones in their experiments, even though the contact zones should have been much larger for their specimen and loading. In view of our experience here, a number of possibilities come to mind. First of all, the cracks used for their work were formed by Teflon inserts which give rise to relatively large initial gaps. Secondly, the displacement measurements (1988) were made at the specimen edge and the stress measurements (1989) were averaged through the thickness. In both cases three-dimensional effects could have contributed to the lack of contact. Finally, our observations indicate that very small bond-normal displacements eliminate contact and great care must be taken in loading device design and specimen alignment as was outlined in Section 2.

4 Analysis of Crack Initiation

The measured NCOD, by themselves, do not provide sufficient information with which to extract mixed-mode fracture parameters of interest, particularly during crack initiation. A previously developed hybrid experimental/finite element analysis procedure (Liechti et al., 1987) was again implemented. The procedure consists of matching the measured NCOD and finite element solutions in a region of linear elastic response and then using the matched finite element solutions to extract the mixed-mode fracture parameters. The validity of the procedure is established here and applied to a detailed analysis of crack initiation under bond-normal loading.

A sequence of NCOD profiles was taken from the center of a series of interference patterns, corresponding to the center of the specimen, far removed from any edge effects. Crack-tip asymptotics were examined through logarithmic plots of NCOD versus distance from the crack front (Fig. 7(a)). At low load levels, the data fall on one straight line which, in these experiments, had a slope of 0.52. The expected value is, of course, 0.5 and the slightly higher value may indicate that higher-order terms are having some effect. As the load level was increased, a point was reached where the plots took on a bilinear form. The particular examples shown here are at the

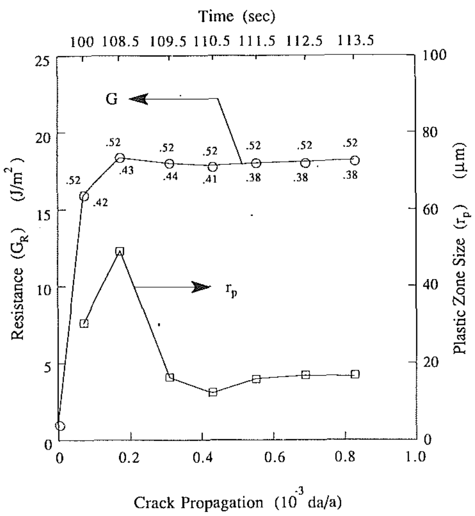


Fig. 8 Crack extension under bond-normal displacements

critical value of applied displacement (defined later) and somewhat later as the crack propagated steadily. In both cases, the original slope of 0.52 was retained well away from the crack front. Near the crack front the slopes reduced to 0.4 and 0.38 at the critical and post critical applied displacements, respectively. These lower slopes are indicative of some inelastic response but do not yield the value of $\frac{1}{n+1} = 0.167$ that would be expected from the power-law hardening exponent of $n=5$ for the epoxy (Table 1) and HRR singular fields. However, some crack extension had occurred at the times that these analyses were conducted and the singularities are really those of a growing, rather than stationary crack. Shih and Asaro (1988) recently showed that the asymptotic fields of a stationary crack between a power-law hardening and a rigid one are nearly similar to the HRR fields that arise in a cracked, homogeneous, power-law hardening material under mixed-mode loading. On the other hand, experimental analyses (Epstein, 1989) have not revealed HRR fields on the specimen surface near the tip of an interface crack.

The extent of the plastic zone behind the crack front was taken to be at the intersection of the lines representing the elastic and inelastic response in Fig. 7(*a*). The plastic zone size at crack initiation ($v_0 = v_{0_c}$) was therefore found to be 49.3 μm. Considering that the specimen thickness was 5.97 mm, the yielding was small scale in nature, thus permitting a previously employed hybrid approach (Liechti et al., 1987) for extracting mixed-mode fracture parameters from NCOD measurements to be considered here. The basis for the approach is the comparison between measured NCOD and linear elastic finite element solutions of the corresponding geometry and loadings, an example of which is shown in Fig. 7(*b*). The initial profile was matched by applying a suitable bond-normal displacement in the finite element analysis. The subsequently applied displacements were then added to the initial displacements so that measured and predicted NCOD could be properly compared. The experimental and numerical results for various applied bond normal displacements levels up to the critical one were in close agreement, thus permitting the finite element solution to be used for extracting mixed-mode fracture parameters. The favorable comparison also indicates that plane-strain conditions prevail at the center of the specimen. All values of fracture parameters subsequently reported were obtained by matching NCOD well outside any regions of inelastic response.

The relatively high degree of magnification that was used to resolve the interference fringe patterns meant that very small amounts of crack extension ($\Delta a/a \simeq 2\times10^{-5}$) could be resolved. The question arose as to what degree of crack extension constituted "initiation." The procedure that was adopted is now described with reference to Fig. 8, where a number of parameters are presented as a function of crack extension. First, it can be seen that the energy release rate increased with load level and crack extension until the crack attained a steady velocity. The elapsed times from load initiation are noted for various values of crack extension and indicate the energy release rate peaked just prior to dropping off slightly to a constant value as steady propagation occurred. If the load was held constant during the time when the energy release rate was increasing, then crack extension would stop. Thus, on this scale, the relatively brittle crack initiation process under bond-normal displacements as judged by the maximum G value of 18.4 J/m^2 displays a response that is reminiscent of stable crack initiation in very tough materials. The critical value of energy release rate was taken to be the constant value associated with steady crack extension and all quoted values of critical applied displacements were likewise associated with the attainment of constant crack velocity. Since the G values in Fig. 8 were essentially derived from NCOD profiles, the results indicate that steady crack propagation is associated with a fixed NCOD profile.

The other parameters noted in Fig. 8 were derived from logarithmic plots of the type shown in Fig. 7(*a*). The values noted above and below the resistance curve at various degrees of crack extension correspond respectively to the slopes of the lines in the regions of elastic and inelastic response. Thus, it can be seen that the exponent in the elastic region was consistently the 0.52 value noted in Fig. 7(*a*) and that exponent variations occurred in the inelastic region, depending on the degree of crack extension, until steady crack propagation occurred. The plastic zone sizes, r_p, were also recorded as a function of crack extension. An increase in r_p was noted during stable crack extension but it was then followed by a sharp decrease to a constant value which was associated with steady crack growth. The smaller plastic zone size during steady crack extension is presumably due to rate effects. The synchronization of changes in energy release rate values, inelastic exponents and plastic zone sizes were all very consistent and give a picture of blunting (on a very small scale) prior to steady propagation.

A series of experiments under bond-normal applied displacements were conducted on a single specimen by unloading very quickly once steady crack propagation was well established. The arrested crack became the starter crack for the next experiment. This procedure gave rise to starter cracks that were sharp and not influenced by the previous experiment. The same crack extension behavior noted above was observed in all experiments and the critical value of energy release rate associated with steady extension was found to be 17 J/m^2 with a coefficient of variation of 8.3 percent, indicating reasonable reproducibility within one specimen.

5 Conclusions

The paper has described the analysis of a single specimen which, when used with a specially developed biaxial loading device, should be capable of providing a wide range of mixtures of mode I and mode II. A stress analysis revealed that, for positive bond-normal applied displacements, the mixity ranged from -60 deg to 90 deg for ratios of applied bond-tangential displacement to bond-normal displacements of 10 to -7.5, respectively. The degree of crack face contact near the crack tip was relatively small ($< h/100$) under positive bond-tangential displacements and nonexistent (within the resolution of the finite element mesh) for negative bond tangential and positive bond-normal applied displacements.

The biaxial loading device was capable of producing steady crack propagation under all loading directions. Crack extension accompanied by near-tip crack face contact was observed using optical interferometry to measure NCOD. In a series of experiments under bond-normal applied displacements, the measured NCOD revealed that, for the relatively weak bond between epoxy and very smooth glass, crack initiation was accompanied by small-scale blunting whose extent was traced as a function of crack velocity. Energy release rates were extracted from linear elastic finite element solutions that matched the measured NCOD in regions of elastic response. Due to the high resolution in crack extension measurements, the energy release rates were found to increase with increasing crack extension until steady propagation occurred. The constant G value corresponding to steady propagation was taken to be the critical value and was found to be 17 J/m^2 for $\psi = 16$ deg. The extension of the analyses and the procedures developed here to determine and examine the increase in G_c with positive and negative mixities is described in an accompanying paper (Liechti and Chai, 1989).

Acknowledgments

The authors would like to acknowledge the support of the National Science Foundation through Grant Number MSM-8813822. The digital image analysis system and some of the loading device components were provided under the University Research Instrumentation Program through the Office of Naval Research (Grant Number N00014-84-G-0175). We would also like to thank Jan Shrode for timely preparation of the manuscript.

References

Anderson, G. P., DeVries, K. L., and Williams, M. L., 1974, "Mixed Mode Stress Field Effect in Adhesive Fracture," *International Journal of Fracture*, Vol. 10, pp. 565–583.

Argon, S. A., Gupta, V., Landis, H. S., and Cornie, J. A., 1989, "Instrinsic Toughness of Interfaces Between Si C Coatings and Substrates of Si or C Fiber," *Journal of Materials Science*, in press.

Arin, K., and Erdogan, F., 1971, "Penny-Shaped Crack in an Elastic Layer Bonded to Dissimilar Half Spaces," *International Journal of Engineering Science*, Vol. 9, pp. 213–232.

Atkinson, C., 1977, "On Quasistatic Problems of Cracks in a Non-Homogeneous Elastic Layer," *Acta Mechanica*, Vol. 26, pp. 103–113.

Becker, E. B., Chambers, R. S., Collins, L. R., Liechti, K. M., Knauss, W. G., and Romanko, J., 1984, "Viscoelastic Stress Analysis Including Moisture Diffusion for Adhesively Bonded Joints," AFWAL-TR-84-4057.

Cao, H. C., and Evans, A. G., 1989, " An Experimental Study of the Fracture Resistance of Bimaterial Interfaces," *Mechanics of Materials*, Vol. 7, pp. 295–304.

Chai, H., 1989, "Three-Dimensional Fracture Analysis of Thin Film Debonding," *International Journal of Fracture*, in review.

Chai, Y.-S., 1990, "Interfacial Crack Initiation Under Biaxial Loads," Engineeering Mechanics Research Laboratory Report #EMRL 90-1, Ph.D. Dissertation, The University of Texas at Austin, Austin, Tex.

Charalambides, P. G., Lund, J., Evans, A. G., McMeeking, R. M., 1989a, "A Test Specimen for Determining the Fracture Resistance of Bimaterial Interfaces," ASME Journal of Applied Mechanics, Vol. 56, pp. 77–82.

Charalambides, P. G., and Evans, A. G., 1989b, "Debonding Properties of Residually Stressed Brittle Matrix Composites," *Journal of the American Ceramics Society*, Vol. 72, pp. 746–753.

Chiang, F. P., and Hua, L., 1988, "Measurement of Displacement Fields Around an Interface Crack in a Bimaterial," Technical Report #520, College of Engineering and Applied Sciences, State University of New York, Stony Brook, N.Y.

Chiang, F. P., Hua, L., and Yan, X. T., 1989, "Photoelastic Analysis of a Crack at a Bimaterial Interface," *Advances in Fracture Research*, Salama, Ravi-Chandar Taplin, and Rama Rao, eds., Vol. 4, pp. 3063–3072.

Comninou, M., 1978, "The Interface Crack in a Shear Field," ASME Journal of Applied Mechanics, Vol. 45, pp. 287–290.

Comninou, M., and Schmueser, D., 1979, "The Interface Crack in a Combined Tension-Compression and Shear Field," ASME Journal of Applied Mechanics, Vol. 46, pp. 345–348.

Epstein, J. S., Graham, S. M., Lloyd, W. R., and Reuter, W. G., 1989, "Elastic Plastic Displacement Fields for Diffusion Bondline Through Cracks Under Small and Large Scale Yield," Proceedings 1989 SEM Spring Meeting, pp. 411–413.

Gent, A. N., and Kinloch, A. J., 1971, "Adhesion of Viscoelastic Materials to Rigid Substrates III. Energy Criterion for Failure," *Journal of Polymer Science*, Part A2, Vol. 9, pp. 659–668.

Ginsburg, D., 1987, "Calculation of Fracture Parameters for Interface Cracks with Application to Mixed-Mode Crack Initiation," Engineering Mechanics Research Laboratory Report #EMRL 87/2, MS Thesis, The University of Texas at Austin, Austin, Tex.

Jensen, H. M., Hutchinson, J. W., and Kim, K.-S., 1990, "Decohesion of a Cut Prestressed Film on a Substrate," *International Journal of Solids and Structures*, Vol. 26, pp. 1099–1114.

Knauss, W. G., 1966, "Stresses in an Infinite Strip Containing a Semi-Infinite Crack," ASME Journal of Applied Mechanics, Vol. 88, pp. 356–362.

Liechti, K. M., and Knauss, W. G., 1981, "The Use of Crack Profile Measurements to Determine Mode Interactions of Propagating Cracks at Material Interfaces," *Advances in Aerospace Structures and Materials*, Wang and Renton, eds., No. AD-01, ASME, New York, pp. 51–60.

Liechti, K. M., and Knauss, W. G., 1982a, "Crack Propagation at Material Interfaces: I. Experimental Technique to Determine Crack Profiles," *Experimental Mechanics*, Vol. 22, pp. 262–269.

Liechti, K. M., and Knauss, W. G., 1982b, "Crack Propagation at Material Interfaces: II. Experiments on Mode Interaction," *Experimental Mechanics*, Vol. 22, pp. 383–391.

Liechti, K. M., Ginsburg, D., and Hanson, E. C., 1987, "A Comparison Between Measurements and Finite Element Predictions of Crack Opening Displacements Near the Front of an Interface Crack," *Journal of Adhesion*, Vol. 23, pp. 123–146.

Liechti, K. M., and Hanson, E. C., 1988, "Nonlinear Effects in Mixed-Mode Interfacial Delaminations," *International Journal of Fracture*, Vol. 36, pp. 199–217.

Liechti, K. M., and Freda, T., 1989, "On the Use of Laminated Beams for the Determination of Pure and Mixed-Mode Fracture Properties of Structural Adhesives," *Journal of Adhesion*, Vol. 28, pp. 145–169.

Liechti, K. M., and Chai, Y.-S., 1989, "Asymmetric Shielding in Interfacial Fracture Under In-Plane Shear," Engineering Mechanics Research Laboratory Report No. EMRL 89/4, The University of Texas at Austin, Austin, Tex.

Malyshev, B. M., Salganik, R. L., 1965, "The Strength of Adhesive Joints Using the Theory of Cracks," *International Journal of Fracture*, Vol. 1, pp. 114–128.

Mulville, D. R., Hunston, D. L., and Mast, P. W., 1978, "Developing Failure Criteria for Adhesive Joints under Complex Loading," *J. Eng. Mat. Tech.*, Vol. 100, pp. 25–31.

Raju, I. S., 1986, "Simple Formulas for Strain Energy Release Rates with Higher Order and Singular Elements," NASA Contractor Report 178196.

Rice, J. R., 1967, "Stresses in a Semi-Infinite Strip Containing a Semi-Infinite Crack," ASME Journal of Applied Mechanics, Vol. 89, pp. 248–249.

Rice, J. R., 1988, "Elastic Fracture Mechanics Concepts for Interfacial Cracks," ASME Journal of Applied Mechanics, Vol. 55, pp. 98–103.

Rosenfeld, L. G., Ritter, J. E., Lardner, T. J., and Lin, M. R., 1990, "Use of Microindentation Technique for Determining Interfacial Fracture Energy," *Journal of Applied Physics*, to appear.

Shih, C. F., and Asaro, R. J., 1988, "Elastic-Plastic Analysis of Cracks on Bimaterial Interfaces: Part I—Small-Scale Yielding," ASME Journal of Applied Mechanics, Vol. 55, pp. 299–316.

Smelser, R. E., 1979, "Evaluation of Stress Intensity Factors for Bimaterial Bodies Using Numerical Crack Flank Displacement Data," *Int. J. Fracture*, Vol. 15, pp. 135–143.

Suo, Z., and Hutchinson, J. W., 1989, "On Sandwich Test Specimens for Measuring Interface Crack Toughness," *Materials Science and Engineering*, Vol. A107, pp. 135–143.

Suo, Z., and Hutchinson, J. W., 1990, "Interface Crack Between Two Elastic Layers," *International Journal of Fracture*, Vol. 43, pp. 1–18.

Takashi, N., Yamazaki, K., Natsume, T., and Takebe, T., 1978, "Debonding of an Al-Epoxy System by Blister Method," *21st Japan Congress on Materials Research—Non-Metallic Materials*, pp. 260–264.

Trantina, G. C., 1972, "Combined Mode Crack Extension in Adhesive Joints," *Journal of Composite Materials*, Vol. 6, pp. 271–385.

Wang, J.-S., and Suo, Z., 1990, "Experimental Determination of Interfacial Toughness Curves Using Brazil-Nut Sandwiches," *Acta Metall. Mater.*, to appear.

Yau, J. F., and Wang, S. S., 1984, "An Analysis of Interface Cracks Between Dissimilar Isotropic Materials Using Conservation Integrals in Elasticity," *Engineering Fracture Mechanics*, Vol. 20, pp. 423–432.

Reflection and Transmission of Rayleigh Surface Waves by a Material Interphase

Z. L. Li

J. D. Achenbach
Fellow ASME

Center for Quality Engineering and Failure Prevention, Northwestern University, Evanston, IL 60208

Reflection and transmission of Rayleigh surface waves by a juncture normal to the free surface, between identical or different materials, has been investigated. The juncture, which may be an interface containing defects or a thin layer, is represented by a layer of extensional and shear springs. The mathematical statement of the problem is reduced to a system of singular integral equations for the displacements on the free surface and the tractions and the displacements across the juncture. Numerical solutions of this system have been computed by the use of the boundary element method. Expressions for the reflection and transmission coefficients have subsequently been obtained by the use of half-plane Green's functions in conjunction with an elastodynamic representation integral. Results are presented for selected values of the elastic constants of the joined bodies and the stiffness parameters of the juncture.

1 Introduction

Interfaces between adjoining materials of similar or dissimilar properties often are of a more complex nature than can be represented by the conditions of continuity of tractions and displacements across a surface of perfect contact. This is obvious when the interface is actually an interphase, i.e., a very thin layer of different mechanical properties, or when the interface contains defects such as cracks, voids, or small inclusions.

One of the ways of obtaining information on interfaces is from the reflection and transmission of ultrasonic waves. The reflection of plane harmonic waves by a perfect interface of by a thin layer in between two half-spaces is a well-known and relatively simple problem. Reflection and transmission coefficients for a planar distribution of cracks in an otherwise homogeneous unbounded solid have also been calculated by Angel and Achenbach (1985) and Sotiropoulos and Achenbach (1988). The analogous results for a planar distribution of spherical cavities have been obtained by Achenbach and Kitahara (1986). In these studies it was found that the reflection and transmission coefficients for a thin layer and a planar distribution of defects depend on the frequency, while the corresponding coefficients for the perfect interface (as well as for the still simpler case of frictionless "sliding" contact) are independent of the frequency.

A convenient way of modeling the reflection and transmission properties of an interface or interphase is by a distribution of linear springs. This model has previously been considered by a number of authors, most recently by Mal et al. (1989), Mal and Xu (1989), and Olsson et al. (1990). Various approximations for interphase behavior have recently been reviewed by Martin (1990). For an interface containing a distribution of cracks, the spring model was proposed by Thompson and Fiedler (1984). For a periodic planar array of cracks, the validity of the model over a specified range of frequencies was verified by Angel and Achenbach (1985) who also presented explicit expressions for the spring constants in terms of the crack widths and the crack spacings. For an interphase, i.e., a thin interface layer, it has been pointed out by Datta et al. (1988) and Olsson et al. (1990) that a spring layer, whose constants are expressed in terms of the elastic constants of the interphase only, generally is not entirely adequate, since the inertia effect of the interphase cannot be ignored. From some simple calculations the present authors have, however, concluded that if the ratio of the thickness of the interphase to the wavelength is less than 0.2, the interphase can in fact be modeled by a spring layer, but one whose spring constants depend not only on the elastic properties of the interphase, but also on its mass density, as well as on the angle of incidence. Hence, the use of a spring layer can be justified, certainly for an interface containing defects, but also for an interphase, provided that in the latter case the spring constants are appropriately selected.

Of frequent interest are the mechanical properties of an interphase near the point of intersection with a free boundary. This configuration is considered in this paper, specifically, the case when the interphase is normal to the free boundary. The spring layer model is used to model the interphase properties.

The presence of the free surface suggests the use of Rayleigh surface waves for the interrogation of the interphase, in the

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Aug. 24, 1989; final revision, Mar. 12, 1990.

same manner as this has been done for surface breaking cracks (see, for example, the paper by Vu and Kinra (1985)). On a smaller scale such an interrogation could be carried out by a line focus acoustic microscope, similarly to the investigations reported by Kushibiki and Chubachi (1985).

The specific configuration is shown in Fig. 1. It is shown in this paper that the mathematical statement of the problem of reflection and transmission of Rayleigh surface waves by a spring layer interphase can be reduced to a system of singular integral equations over the boundaries. In our procedure full-plane Green's functions have been used. As a consequence the system of singular integral equations consists of equations over the traction-free boundaries of the quarter-planes as well as over the interphase. The boundary element method has been used to solve the system of integral equations, but for practical purposes it was necessary to truncate the infinite integrals. To reduce the error related to the truncation, the omitted integration paths over the traction-free boundaries have been replaced by infinite elements which can accommodate the outgoing Rayleigh surface waves reflected and transmitted by the interphase. Once the boundary integral equations have been solved for the interphase fields, integral representations using half-plane Green's functions have been used to write expressions for the reflection and transmission coefficients. Numerical results are presented for selected ratios of the mechanical properties of the quarter-planes and the spring layer constants.

The results presented in this paper also have relevance to earlier generally analytical results for the reflection and transmission of surface waves in a quarter-plane. The many analytical efforts devoted to this problem have been reviewed by Knopoff (1969) and Miklowitz (1987). The case of joined quarter-planes was investigated by Viswanathan (1966) who employed an iteration method. The quarter-plane problem was recently reconsidered by Gautesen (1985). It is shown that for the appropriate limit case the results of this paper agree with those obtained for the quarter-plane.

2 Statement of the Problem

A half-plane is composed of two homogeneous, isotropic, linearly elastic quarter-planes, A and B, which are joined by an interphase. The materials of A and B may be the same or they may be different. The half-plane occupies the domain $x_2 \geq 0$, and the interphase is located along $x_1 = 0$. The two-dimensional geometry is shown in Fig. 1.

The motion of the half-plane is time harmonic with angular frequency ω. In the following analysis, the time-harmonic factor $\exp(-i\omega t)$ will, however, be omitted.

The incident wave is a Rayleigh surface wave which propagates from $x_1 = -\infty$ to $x_1 = 0$ in the region A. The displacement components of the incident wave may be written as

$$u_1^{in}(\mathbf{x}) = U \frac{k_R^A}{k_L^A} \left\{ e^{-\gamma_L^A x_2} + \frac{2\gamma_L^A \gamma_T^A}{(k_T^A)^2 - 2(k_R^A)^2}\, e^{-\gamma_T^A x_2} \right\} e^{ik_R^A x_1}, \qquad (1a)$$

$$u_2^{in}(\mathbf{x}) = iU \frac{\gamma_L^A}{k_L^A} \left\{ e^{-\gamma_L^A x_2} + \frac{2(k_R^A)^2}{(k_T^A)^2 - 2(k_R^A)^2}\, e^{-\gamma_T^A x_2} \right\} e^{ik_R^A x_1}, \qquad (1b)$$

where k_L^A, k_T^A, and k_R^A are the wave numbers of longitudinal, transverse, and Rayleigh waves, respectively,

$$k_L^A = \omega / c_L^A\ , \qquad c_L^A = [(\lambda^A + 2\mu^A)/\rho^A]^{1/2}\ , \qquad (2a,b)$$

$$k_T^A = \omega / c_T^A\ , \qquad c_T^A = (\mu^A/\rho^A)^{1/2}\ , \qquad (3a,b)$$

$$k_R^A = \omega / c_R^A\ . \qquad (4)$$

Here, c_L^A, c_T^A, c_R^A are the longitudinal, transverse, and Rayleigh wave velocities, λ^A and μ^A are the Lamé elastic constants, and ρ^A is the mass density of A. Also,

$$\gamma_L^A = [(k_R^A)^2 - (k_L^A)^2]^{1/2}\ , \ \gamma_T^A = [(k_R^A)^2 - (k_T^A)^2]^{1/2}\ . \qquad (5a,b)$$

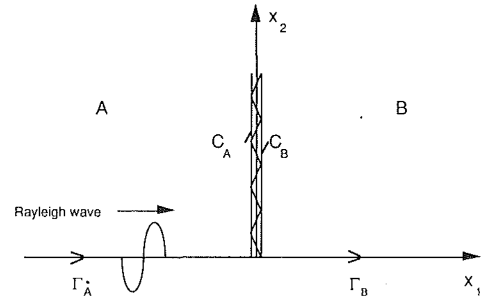


Fig. 1 Two quarter-planes joined by a spring layer

In region A, the components of the total displacement field, u_j^A (in this paper the lower case Latin subscripts range over 1 and 2), are the sums of the corresponding displacement components, of the incident wave, u_j^{in}, and the back-scattered wave, u_j^s:

$$u_j^A = u_j^{in} + u_j^s\ ,\ j = 1,2. \qquad (6)$$

The components of the transmitted displacement field in region B are denoted by u_j^B. According to Hooke's law, the corresponding stress fields are

$$\sigma_{ij}^A = \lambda^A \delta_{ij} u_{k,k}^A + \mu^A (u_{i,j}^A + u_{j,i}^A)\ , \qquad (7)$$

$$\sigma_{ij}^B = \lambda^B \delta_{ij} u_{k,k}^B + \mu^B (u_{i,j}^B + u_{j,i}^B)\ , \qquad (8)$$

where λ^B and μ^B are the Lamé elastic constants of the material in region B. On a surface with unit outward normal vectors $\mathbf{n}^A$ or $\mathbf{n}^B$, the tractions are

$$f_i^A = \sigma_{ij}^A n_j^A\ \ ,\ f_i^B = \sigma_{ij}^B n_j^B\ . \qquad (9a,b)$$

A spring layer with appropriate spring constants may often be used to model the interphase, and the problem can then be simplified significantly. Let S_L and S_T denote the extensional and shear constants of the spring layer. The conditions of continuity along $x_1 = 0$ can then be expressed as

$$f_j^A = -f_j^B\ , \qquad (10)$$

$$f_1^A = S_L(u_1^B - u_1^A) = S_L(u_1^B - u_1^{in} - u_1^s)\ , \qquad (11a)$$

$$f_2^A = S_T(u_2^B - u_2^A) = S_T(u_2^B - u_2^{in} - u_2^s)\ . \qquad (11b)$$

On the free surface, $x_2 = 0$, the tractions vanish and thus

$$f_j^A(\mathbf{x}) = f_j^s(\mathbf{x}) + f_j^{in}(\mathbf{x}) = 0, \quad \mathbf{x} \text{ on } \Gamma_A \qquad (12a)$$

$$f_j^B(\mathbf{x}) = 0, \qquad \mathbf{x} \text{ on } \Gamma_B \qquad (12b)$$

where f_j^s and f_j^{in} are the tractions related to the back-scattered and incident waves, respectively. Since $f_j^{in}(\mathbf{x}) = 0$, for $\mathbf{x}$ on Γ_A, we have

$$f_j^s(\mathbf{x}) = 0, \quad \mathbf{x} \text{ on } \Gamma_A\ . \qquad (13)$$

3 Boundary Integral Equations

To derive the boundary integral equations which can be used in conjunction with the boundary element method to obtain numerical results, we start with an integral representation of the elastodynamic solution for the back-scattered and transmitted fields. For the present problem two integral representations can be used: one based on the half-plane Green's functions for materials A and B, and the other based on the full-plane Green's functions. The integral representations of the scattered fields in terms of the half-plane Green's functions (which satisfy traction-free conditions on the surface of the half-plane) include only integrals along the interface of the two quarter planes. Unfortunately, the use of the half-plane Green's functions increases the computing effort due to the high complexity of these functions. The integral representations of the scattered fields in terms of the full-plane Green's functions (the fundamental solutions) on the other hand include not only the integral along the interface of the quarter-planes but also along the traction-free surfaces of the quarter planes. The full-plane Green's functions are, however, of simpler analytical forms, and they make the boundary element computations more efficient. However, as we will see in the next section, once the displacement and traction fields on the interface of the two quarter-planes are known, it is easier and more accurate to calculate the solutions in the far-fields by the use of the half-plane Green's functions.

In the sequel $u_{ik}^G(\mathbf{x};\mathbf{x}_p)$ and $\sigma_{ijk}^G(\mathbf{x};\mathbf{x}_p)$ denote the full-plane fundamental solution and the second fundamental solution, respectively. The physical meaning of $u_{ik}^G(\mathbf{x};\mathbf{x}_p)$ is that these components denote the displacement in the direction x_i at point $\mathbf{x}$, due to a concentrated load at $\mathbf{x}_p$, applied in the direction x_k. The components $\sigma_{ijk}^G(\mathbf{x};\mathbf{x}_p)$ are the corresponding stresses.

The integral representation for the back-scattered field in region A can be written as

$$\epsilon_A(\mathbf{x}_p)u_k^s(\mathbf{x}_p)=\int_{\Gamma_A+C_A}[u_{ik}^{GA}(\mathbf{x};\mathbf{x}_p)f_i^s(\mathbf{x}) - \sigma_{ijk}^{GA}(\mathbf{x};\mathbf{x}_p)u_i^s(\mathbf{x})n_j^A(\mathbf{x})]ds_x\ , \quad (14)$$

where

$$\epsilon_A(\mathbf{x}_p)=\begin{cases}1\ ,\ \mathbf{x}_p\in A\\ 0\ ,\ \mathbf{x}_p\notin A\ ,\end{cases}$$

and the contours Γ_A and C_A are indicated in Fig. 1. The analogous integral representation for the displacement field in region B is

$$\epsilon_B(\mathbf{x}_p)u_k^s(\mathbf{x}_p)=\int_{C_B+\Gamma_B}[u_{ik}^{GB}(\mathbf{x};\mathbf{x}_p)f_i^B(\mathbf{x}) - \sigma_{ijk}^{GB}(\mathbf{x};\mathbf{x}_p)u_i^B(\mathbf{x})n_j^B(\mathbf{x})]ds_x \quad (15)$$

where

$$\epsilon_B(\mathbf{x}_p)=\begin{cases}1\ ,\ \mathbf{x}_p\in B\ ,\\ 0\ ,\ \mathbf{x}_p\notin B\ ,\end{cases}$$

and again Γ_B and C_B are indicated in Fig. 1. Explicit expressions for the fundamental solution and the second fundamental solution may be found, for example, in the work by Kobayashi (1987). For reference purposes the expressions have been stated in the Appendix.

For $\mathbf{x}_p\notin A$, substitution of (6) and (10)–(12a) into (14) yields

$$\int_{\Gamma_A}\sigma_{ijk}^{GA}u_i^s n_j^A ds_x+\int_{C_A}[\sigma_{ijk}^{GA}u_i^B n_j^A+(u_{ik}^{GA}+\sigma_{1jk}^{GA}n_j^A/S_L)f_1^B + (u_{2k}^{GA}+\sigma_{2jk}^{GA}n_j^A/S_T)f_2^B]ds_x$$
$$\int_{C_A}(\sigma_{ijk}^{GA}u_i^{in}n_j^A-u_{ik}^{GA}f_i^{in})ds_x\ ,\ \mathbf{x}_p\notin A\ . \quad (16)$$

For the $\mathbf{x}_p\notin B$, (15) becomes, by the use of (12b),

$$\int_{C_B+\Gamma_B}\sigma_{ijk}^{GB}u_i^B n_j^B ds_x-\int_{C_b}u_{ik}^{GB}f_i^B ds_x=0\ ,\ \mathbf{x}_p\notin B\ . \quad (17)$$

Taking the limit of $\mathbf{x}_p\notin A\rightarrow\mathbf{x}_p\in\Gamma_A$ or C_A, and $\mathbf{x}_p\notin B\rightarrow\mathbf{x}_p\in\Gamma_B$ or C_B, equations (16) and (17) become a set of boundary integral equations for the following unknowns:

$$u_i^s(\mathbf{x}_p)\ ,\ \mathbf{x}_p \text{ on } \Gamma_A\ ,$$
$$u_i^B(\mathbf{x}_p)\ ,\ \mathbf{x}_p \text{ on } C_A\ ,$$
$$f_i^B(\mathbf{x}_p)\ ,\ \mathbf{x}_p \text{ on } C_B\ ,$$
$$u_i^B(\mathbf{x}_p)\ ,\ \mathbf{x}_p \text{ on } \Gamma_B\ .$$

After the boundary integral equations have been solved, we can calculate the displacement field of back-scattered and transmitted waves at any point in the half-space from (14) and (15).

4 Far-Field Behavior of Back-Scattered and Transmitted Waves

Once the integral equations (16) and (17) have been solved, the fields in the quarter-planes can conveniently be obtained by the use of an integral representation in terms of the half-plane Green's function. This integral representation may be written as

$$u_k^s(\mathbf{x}_p)=\int_{C_A}[U_{ik}^{GA}(\mathbf{x};\mathbf{x}_p)\sigma_{ij}^s(\mathbf{x}) - \tau_{ijk}^{GA}(\mathbf{x};\mathbf{x}_p)u_i^s(\mathbf{x})]n_j^A(\mathbf{x})ds_x\ \mathbf{x}_p\in A\ , \quad (18)$$

where U_{ik}^{GA} and τ_{ijk}^{GA} are half-plane Green's displacement function and Green's stress function for material A.

According to Neerhoff, by expanding the half-plane Green's displacement and Green's stress functions for $|x_{p1}|\rightarrow\infty$, we obtain

$$u_k^s(\mathbf{x}_p)\approx R\ u_k^{SR}(\mathbf{x}_p)\ , \quad (19)$$

where R is defined as the reflection coefficient and u_k^{SR} is a Rayleigh wave propagating in the negative x_1 direction,

$$u_1^{SR}(\mathbf{x})=U\frac{k_R^A}{k_L^A}\left[e^{-\gamma_L^A x_2}+\frac{2\gamma_L^A\gamma_T^A}{(k_T^A)^2-2(k_R^A)^2}e^{-\gamma_T^A x_2}\right]e^{-ik_R^A x_1}\ , \quad (20a)$$

$$u_2^{SR}(\mathbf{x})=-iU\frac{\gamma_L^A}{k_L^A}\left[e^{-\gamma_L^A x_2}+\frac{2(k_R^A)^2}{(k_T^A)^2-2(k_R^A)^2}e^{-\gamma_T^A x_2}\right]e^{-ik_R^A x_1}\ . \quad (20b)$$

We have

$$R=\int_{C_A}[Q_i^A(\mathbf{x})\sigma_{ij}^s(\mathbf{x})-S_{ij}^A(\mathbf{x})u_i^s(\mathbf{x})]n_j^A(\mathbf{x})ds_x\ , \quad (21)$$

where

$$Q_i^A(\mathbf{x}) =\frac{i}{2(\lambda^A+2\mu^A)\gamma_L^A D^A}\left[N_{LL}^A k_i^{LA}e^{ik_L^A k_j^{LA}x_j}+N_{TL}^A\epsilon_{ki}k_i^{TA}e^{ik_T^A k_j^{TA}x_j}\right]$$

$$S_{ij}^A(\mathbf{x})=\frac{-1}{2(\lambda^A+2\mu^A)\gamma_L^A D^A}\left[N_{LL}^A k_L^A(\lambda^A\delta_{ij} +2\mu^A k_i^{LA}k_j^{LA})e^{ik_L^A k_l^{LA}x_l} + N_{TL}^A k_T^A\mu^A(\epsilon_{lj}k_i^{TA}+\epsilon_{li}k_j^{TA})k_l^{TA}e^{ik_T^A k_l^{TA}x_l}\right]\ . \quad (22)$$

In these expressions, ϵ_{ij} is the two-dimensional permutation tensor, and

$$N_{LL}^A=-2[(k_T^A)^2-2(k_R^A)^2]^2\ , \quad (23)$$

$$N_{TL}^A=\frac{4i}{k_L^A}k_T^A k_R^A k_R^A\gamma_L^A[(k_T^A)^2-2(k_R^A)^2]\ , \quad (24)$$

$$D^A=8k_R^A\gamma_L^A\gamma_T^A+4(k_R^A)^3(\gamma_L^A/\gamma_T^A+\gamma_T^A/\gamma_L^A) +8k_R^A[(k_T^A)^2-2(k_R^A)^2] \quad (25)$$

$$k_j^{LA}=\frac{1}{k_L^A}(-k_R^A,\ i\gamma_L^A)\ , \quad (26)$$

$$k_j^{TA}=\frac{1}{k_T^A}(-k_R^A,\ i\gamma_T^A)\ . \quad (27)$$

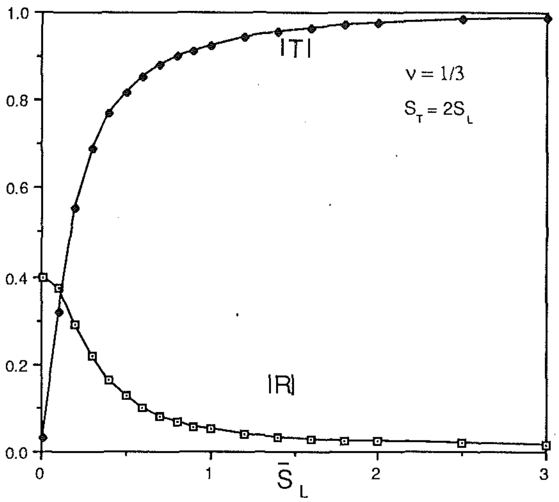


Fig. 2 Absolute values of the reflection and transmission coefficients for a spring layer between identical materials

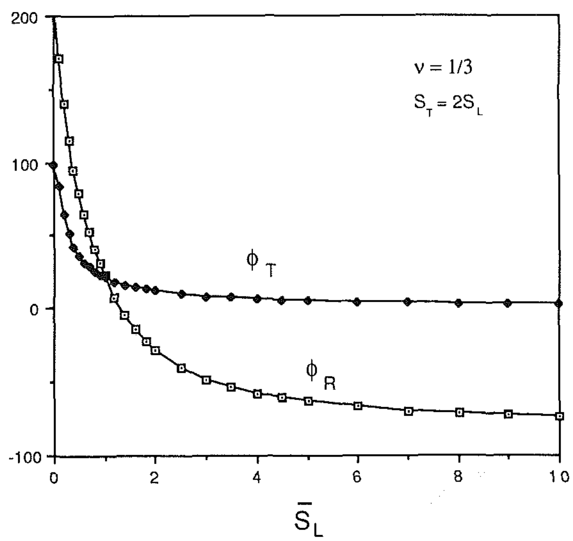


Fig. 3 Phase angles of the reflection and transmission coefficients for a spring layer between identical materials

Similarly, when $x_1 \rightarrow \infty$, for the transmitted wave, we have

$$u_k^B(\mathbf{x}_p) \approx T\, u_k^{BR}(\mathbf{x}_p)\ , \tag{28}$$

where u_k^{BR} is a Rayleigh wave propagating in the positive x_1 direction, and T is the transmission coefficient:

$$u_1^{BR}(\mathbf{x}) = U \frac{k_R^B}{k_L^B}\left[e^{-\gamma_L^B x_2} + \frac{2\gamma_L^B\gamma_T^B}{(k_T^B)^2-2(k_R^B)^2}\, e^{-\gamma_T^B x_2}\right] e^{ik_R^B x_1}, \tag{29a}$$

$$u_2^{BR}(\mathbf{x}) = iU \frac{\gamma_L^B}{k_L^B}\left[e^{-\gamma_L^B x_2} + \frac{2(k_R^B)^2}{(k_T^B)^2-2(k_R^B)^2}\, e^{-\gamma_T^B x_2}\right] e^{ik_R^B x_1}, \tag{29b}$$

$$T = \int_{C_B} [Q_i^B(\mathbf{x})\sigma_{ij}^B(\mathbf{x}) - S_{ij}^B(\mathbf{x})u_i^B(\mathbf{x})]n_j^B(\mathbf{x})ds_x\ . \tag{30}$$

Replacing all A's by B's in equations (22)–(23) and (25) and changing (24), (26) and (27) to

$$N_{TL}^B = -4i\frac{k_T^B}{k_L^B}\, k_R^B\gamma_L^B[(k_T^B)^2 - 2(k_R^B)^2]\ , \tag{31a}$$

$$k_j^{LB} = \frac{1}{k_L^B}\,(k_R^B, i\gamma_L^B), \quad k_j^{TB} = \frac{1}{k_T^B}\,(k_R^B, i\gamma_T^B) \tag{31b,c}$$

we obtain the expressions for Q_i^B and S_{ij}^B in the integrand of (30). Thus, if the displacement and stress fields along C_A and C_B are known, reflection and transmission coefficients R and T can be obtained from (21) and (30).

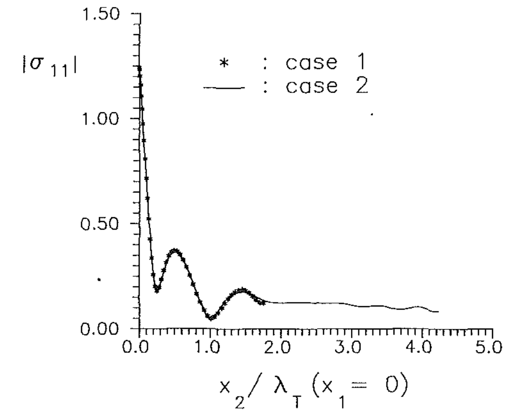


Fig. 4 Absolute values of the total stress component σ_{11} along the interface

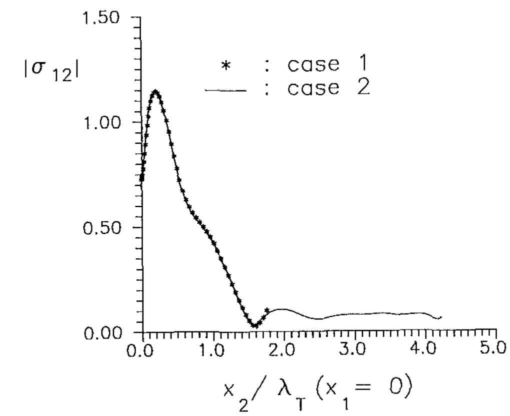


Fig. 5 Absolute values of the total stress component σ_{12} along the interface

5 Boundary Element Method

In this section, we will obtain numerical solutions to the boundary integral equations (16) and (17) by using the boundary element method. In (16) and (17), the integral paths along Γ_A, C_A, C_B and Γ_B extend, however, to infinity, which is not suitable for the numerical procedure.

The wave system generated by interaction of the incident surface wave with the interface consists of reflected and transmitted surface waves, diffracted body waves, and possibly interface waves along C_A and C_B. Interface waves along a spring-layer juncture have been discussed by Jones and Whittier (1967). Along the interface the free surface waves and the body waves decay as x_2 increases. If in addition it may be assumed that sufficiently far from the free surface, interface waves along C_A and C_B yield negligible contributions to R and T, the integral paths C_A and C_B may be truncated to the extent that a desired accuracy is achieved. The integrals along C_A and C_B can then be replaced by integrals along finite paths, say C_{A1} and C_{B1}. For the integrals along Γ_A and Γ_B, on the other hand, the reflected and transmitted waves do not decay as $|x_1| \rightarrow \infty$, since the main parts of these waves are outgoing Rayleigh surface waves. The relative amplitudes of the outgoing Rayleigh waves (T and R) can, however, be computed by truncating Γ_A and Γ_B at some finite length, as was done by Zhang and Achenbach (1988). The accuracy obtained by such a truncation is not very clear. Indeed, if the displacement fields on the free surface must be determined at values of $|x_1|$ which are not large enough, the truncations on Γ_A and Γ_B may produce big errors.

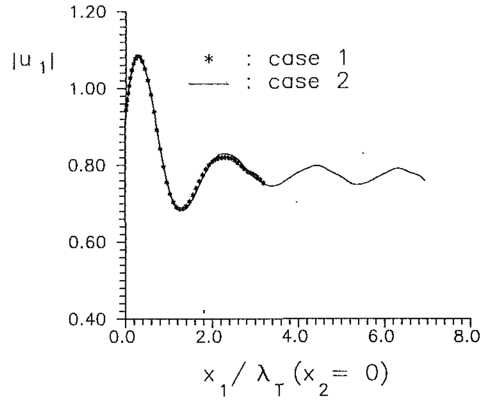


Fig. 6 Absolute values of the total displacement component u_1 along the free surface for $x_1>0$

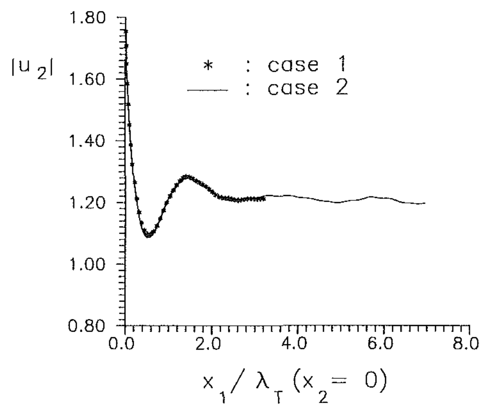


Fig. 7 Absolute values of the total displacement component u_2 along the free surface for $x_1>0$

In order to reduce the error related to the truncations of Γ_A and Γ_B we will add additional terms. For that purpose we split Γ_A at $x_1 = x_{1A}$ in Γ_{A0} and Γ_{A1}, where

$$\Gamma_{A0}:\quad x_2=0,\ x_1<x_{1A}\ ,$$

$$\Gamma_{A1}:\quad x_2=0,\ 0>x_1>x_{1A}\ .$$

Here, $|x_{1A}|$ is large enough so that for $x_1<x_{1A}$, the back-scattered displacement fields can be expressed by (19). If f_j^{SR} denote the traction components corresponding to u_j^{SR}, we can write according to equation (14)

$$\int_{\Gamma_{AO}+\Gamma_{A1}+C_{A1}} [u_{ik}^{GA}(\mathbf{x};\mathbf{x}_p)f_i^{SR}(\mathbf{x}) - \sigma_{ijk}^{GA}(\mathbf{x};\mathbf{x}_p)u_i^{SR}(\mathbf{x})n_j^A(\mathbf{x})]ds_x=0 \quad (32)$$

for $\mathbf{x}_p \notin A$. By the use of the traction-free conditions on Γ_A, equation (32) becomes

$$\int_{\Gamma_{A0}} \sigma_{ijk}^{GA}u_i^{SR}n_j^A ds_x = -\int_{\Gamma_{A1}} \sigma_{ijk}^{GA}u_i^{SR}n_j^A ds_x + \int_{C_{A1}} (u_{ik}^{GA}f_i^{SR} - \sigma_{ijk}^{GA}u_i^{SR}n_j^A)ds_x\ . \quad (33)$$

Now, substituting (19), into (16) and using (33), we obtain

$$\int_{\Gamma_{A1}} \sigma_{ijk}^{GA}u_i^S n_j^A ds_x + \int_{\Gamma_{A1}} [\sigma_{ijk}^{GA}u_i^B n_j^A + (u_{9k}^{GA} + \sigma_{1jk}^{GA}n_j^A/S_L)f_1^B + (u_{2k}^A + \sigma_{2jk}^{GA}n_j^A/S_T)f_2^B]ds_x - R\left(\int_{\Gamma_{A1}+C_{A1}} \sigma_{ijk}^{GA}u_i^{SR}n_j^A ds_x - \int_{C_{A1}} u_{ik}^{GA}f_i^{SR}ds_x\right)$$

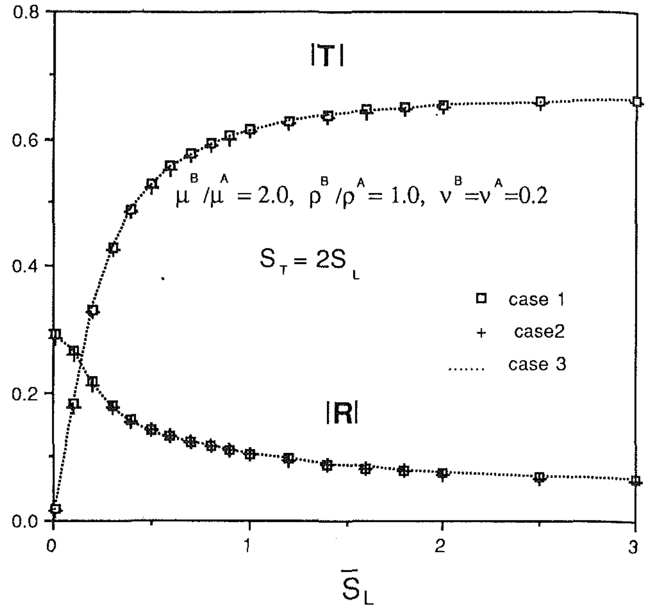


Fig. 8 Absolute values of the reflection and transmission coefficients for a spring layer between different materials

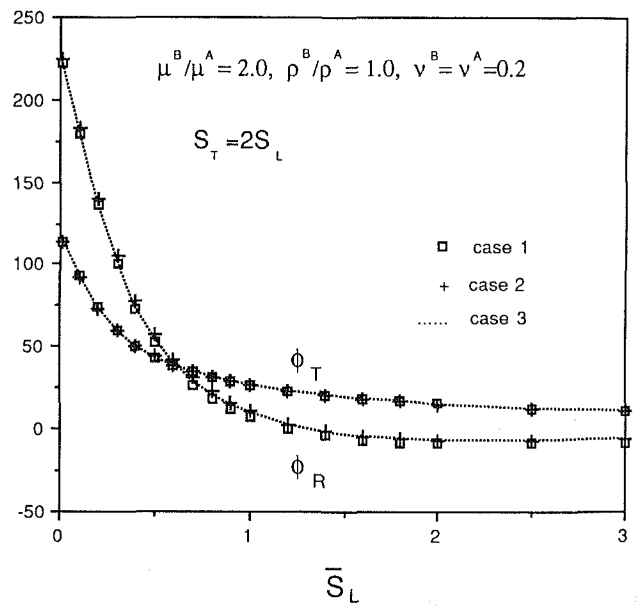


Fig. 9 Phase angles of the reflection and transmission coefficients for a spring layer between different materials

$$= \int_{C_{A1}} (\sigma_{ijk}^{GA}u_i^{in}n_j^A - u_{ik}^{GA}f_i^{in})ds_x\ . \quad (34)$$

Using a similar approach, for the transmitted waves, we have from (17)

$$\int_{C_{B1}+\Gamma_{B1}} \sigma_{ijk}^{GB}u_i^B n_j^B ds_x - \int_{C_{B1}} u_{ik}^B f_i^B ds_x - T\left(\int_{C_{B1}\Gamma_{B1}} \sigma_{ijk}^{GB}u_i^{BR}n_j^B ds_x - \int_{C_{A1}} u_{ik}^{GB}f_i^{BR}ds_x\right) = 0 \quad (35)$$

where Γ_{B1} is the part of the boundary $x_2=0$ defined by $0<x_1<x_{1B}$, and x_{1B} is large enough such that for $x_1>x_{1B}$, the transmitted waves can be expressed approximately by (28). Also, f_j^{BR} denotes the tractions corresponding to u_j^{BR}, where the latter are defined by (29).

Equations (34) and (35) are boundary integral equations

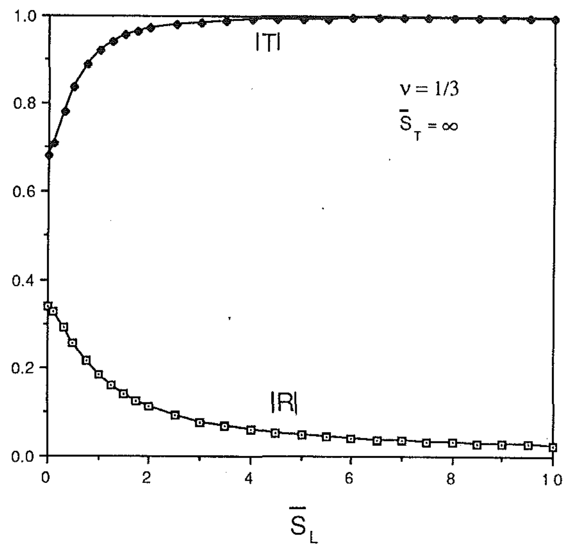


Fig. 10 Absolute values of the reflection and transmission coefficients for an extensional spring layer between identical materials; $S_T = \infty$

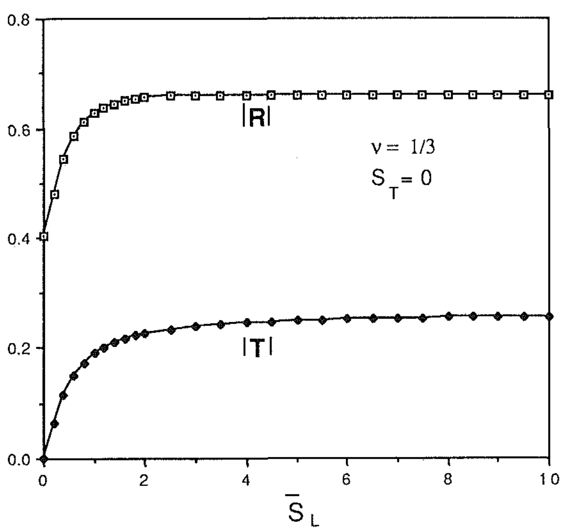


Fig. 11 Absolute values of the reflection and transmission coefficients for an extensional spring layer between identical materials; $S_T = 0$

which can be solved in the usual manner by the boundary element method.

It is useful to examine equations (34) and (35) for the special case when the two materials in A and B are the same and $S_T = S_L \to \infty$. This case corresponds to a homogeneous half-plane problem, and therefore the back-scattered fields must be exactly zero and the transmitted wave is just the same as the incident Rayleigh surface wave. Indeed, $u_j^s = 0$, $u_j^B = u_j^{in}$, $f_j^B = -f_j^{in}$, $R = 0$, $T = 1$ satisfy the boundary integral equations (34) and (35) exactly. To check the effect of truncation for this case, the terms with R and T in equations (34) and (35) can be omitted. Numerical results for neglecting the terms with T and R of (34) and (35) have been computed. The error of the displacement fields of the transmitted waves on the free surface Γ_{B1}, is not negligible, but both the stresses and displacements on the interface C_{A1} and C_{B1}, show excellent agreement with the exact results. If T and R are then calculated by substituting the results obtained on C_{A1} and C_{B1} into (30) and (21), we obtain accurate results, namely: $T = 0.9994$, $R = 0.0000$.

6 Numerical Results

For the same materials in quarter-planes A and B and Poisson's ratio $\nu = 1/3$, the absolute values of the reflection coefficient R and the transmission coefficient T are shown in Fig. 2 as functions of the dimensionless spring constant

$$\bar{S}_L = \frac{1}{\mu^A k_T^A} S_L . \tag{36}$$

For these results we have also taken $\bar{S}_T = S_T/\mu^A k_T^A = 2\bar{S}_L$. The corresponding phase angles have been shown in Fig. 3. In the numerical calculation the lengths of Γ_{A1} and Γ_{B1}, are $4\lambda_T$, the lengths of C_{A1} and C_{B1} are $3\lambda_T$, where $\lambda_T = 2\pi/k_T^A$ is the wavelength of transverse waves in region A, and Γ_{A1}, Γ_{B1}, C_{A1}, C_{B1} are divided into 50 elements. The results show that T increases monotonically from zero to unity as $\bar{S}_L$ increases, while R decreases monotonically from a value smaller than unity to zero as $\bar{S}_L$ increases. In the limit $\bar{S}_L \to 0$, the results should reduce to the ones for the quarter-plane. For the quarter-plane case and $\nu = 0.33$, the reflection coefficient was obtained as 0.40 by Achenbach et al. (1980), and 0.39 by Gautesen (1985). In the present work the smallest value of $\bar{S}_L$ was chosen to be 0.01 and the reflection coefficient was obtained as 0.3997, which shows a very satisfactory agreement, with the above-mentioned results.

The effects of truncation of the integral along the interface have been investigated for the case of two identical quarter planes which are connected by an interface defined by $\bar{S}_L = 0.5$ and $\bar{S}_T = 1$. Calculations were carried out for two lengths of $C_{A1} = C_{B1}$, namely $2\lambda_T$ and $4\lambda_T$. For these cases the number of elements along the interface was 50 and 100, respectively. Figures 4 and 5 show the absolute values of the total stresses along the interface. Similar plots have been obtained for the displacements, but they are not shown here. It is noted that the stresses decay rapidly. It is also noted that there is no significant difference between the two cases over the range where both are obtained. Further calculations have shown that the additional information that is obtained for the longer length does not affect the numerical results for R and T.

To check the effect of the integration lengths along the free surface, Γ_{A1} and Γ_{B1}, calculations were carried out for two lengths, namely, $4\lambda_T$ and $8\lambda_T$, for the case of identical quarter-planes and $\bar{S}_L = 0.5$, $\bar{S}_T = 1$. For these cases the absolute values of the total displacments were obtained along the free surface, for $x_1 > 0$. The results are shown in Figs. 6 and 7. Only minor discrepancies are noted for $|u_1|$.

The results of Figs. 4–5 and 6–7 show that truncation of the interface and the use of a finite integration length with an additional term along the free surface yield accurate results.

For different materials in quarter-planes A and B, ($\mu^B/\mu^A = 2.0$, $\rho^B/\rho^A = 1.0$, $\nu^B = \nu^A = 0.20$) the absolute values of T and R versus $\bar{S}_L$, ($\bar{S}_T = 2\bar{S}_L$), and the phase angles ϕ_T and ϕ_R are shown versus $\bar{S}_L$ in Figs. 8 and 9. As $\bar{S}_L$ increases, T increases, but the limit value of T as $\bar{S} \to \infty$ in not unity. When $\bar{S}_L \to 0$ the problem again reduces to a quarter-plane problem. For $\nu = 0.2$, Gautesen (1985) obtained $|R| \approx 0.3$, and for $\bar{S}_L = 0.01$, we obtain $|R| \approx 0.2837$.

Figures 8 and 9 also show further checks on the accuracy of the results. Case 1 corresponds to $C_{A1} = C_{B1} = 1.7\lambda_T$, $\Gamma_{A1} = \Gamma_{B1} = 2.3\lambda_T$, while C_{A1}, C_{B1}, Γ_{A1}, Γ_{B1} are divided into 50 elements each. Similarly, case 2 corresponds to $C_{A1} = C_{B1} = 2.1\lambda_T$, $\Gamma_{A1} = \Gamma_{B1} = 3.45\lambda_t$ and each segment is divided into 100 elements. For case 3, $C_{A1} = C_{B1} = 4.2\lambda_T$, $\Gamma_{A1} = \Gamma_{B1} = 6.9\lambda_T$, and 100 elements were used on each segment. It is noted that there are no significant differences between the results for these three cases.

Finally, for the special cases of $\bar{S}_T \to \infty$ and $\bar{S}_T \to 0$, (only extensional springs exist), T and R versus $\bar{S}_L$ are shown in Fig. 10 and Fig. 11 for identical quarter-planes. Very different variations of $|R|$ and $|T|$ with $\bar{S}_L$ are obtained.

Acknowledgment

This work was carried out in the course of research supported

by the Mechanics Division of the Office of Naval Research under Contract N00014-88-K-0452 with Northwestern University.

References

Achenbach, J. D., Gautesen, A. K., and Mendelsohn, D. A., 1980, "Ray Analysis of Surface-Wave Interaction with an Edge Crack," *IEEE Transactions of Sonics and Ultrasonics*, Vol. SU-27, pp. 124–129.

Achenbach, J. D., and Kitahara, M., 1986, "Reflection and Transmission of an Obliquely Incident Wave by an Array of Spherical Cavities," *Journal of the Acoustical Society of America*, Vol. 80, pp. 1209–1214.

Angel, Y. C., and Achenbach, J. D., 1985, "Reflection and Transmission of Elastic Waves by a Periodic Array of Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 33–41.

Datta, S. K., Olsson, P., and Boström, A., 1988, "Elastodynamic Scattering from Inclusions with Thin Interface Layers," *Wave Propagation in Structural Composites*, AMD-Vol. 90, A. K. Mal and T. C. T. Ting, eds., pp. 109–116.

Gautesen, A. K., 1985, "Scattering of a Rayleigh Wave by an Elastic Quarter Space," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 664–668.

Jones, J. P., and Whittier, J. S., 1967, "Waves at a Flexibly Bonded Interface," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 905–909.

Knopoff, L., 1969, "Elastic Wave Propagation in a Wedge," *Wave Propagation in Solids*, J. Miklowitz, ed., ASME, New York, pp. 3–42.

Kobayashi, S., 1987, "Elastodynamics," *Boundary Element Methods in Mechanics*, D. E. Beskos, ed., North-Holland, Amsterdam, pp. 191–255.

Kushibiki, J., and Chubachi, N., 1985, "Material Characterization by Line-Focus-Beam Acoustic Microscope," *IEEE Transactions on Sonics and Ultrasonics*, Vol. SU-32, pp. 189–212.

Mal, A. K., Xu, P.-C., and Bar-Cohen, Y., 1989, "Analysis of Leaky Lamb Waves in Bonded Plates," *International Journal of Engineering Science*.

Mal, A. K., and Xu, P.-C., 1989, "Elastic Waves in Layered Media," in *Elastic Wave Propagations*, M. F. McCarthy and M. A. Hayes, eds., Elsevier Publishers B.V., New York, pp. 67–73.

Martin, P. A., 1990, "Thin Interface Layers: Adhesives, Approximations and Analysis," *Elastic Wave Propagation and Ultrasonic Non-Destructive Evaluation*, S. K. Datta, J. D. Achenbach, and Y. Rajapakse, eds., to appear.

Miklowitz, J., 1987, "Modern Corner, Edge, and Crack Problems in Linear Elastodynamics Involving Transient Waves," *Advances in Applied Mechanics*, Vol. 25, T. Y., Wu, and J. W. Hutchinson., eds., Academic Press, New York, pp. 47–181.

Neerhoff, F. L., 1980, "Reciprocity and Power-Flow Theorems for the Scattering of Plane Elastic Waves in a Half Space," *Wave Motion*, Vol. 2, pp. 277–292.

Olsson, P., Datta, S. K., and Boström, A., 1990, "Elastodynamic Scattering from Inclusions Surrounded by Thin Interface Layers," ASME JOURNAL OF APPLIED MECHANICS, in press.

Sotiropoulos, D. A., and Achenbach, J. D., 1988, "Reflection of Elastic Waves by Distribution of Coplanar Cracks," *Journal of the Acoustical Society of America*, Vol. 84, pp. 752–759.

Thompson, R. B., and Fiedler, C. J., 1984, "The Effect of Crack Closure on Ultrasonic Scattering Measurements," *Review of Progress in QNDE*, Vol. 3A, D. O. Thompson and D. E. Chimenti, eds., Plenum, New York, pp. 207–215.

Viswanathan, K., 1966, "Wave Propagation in Welded Quarter-Spaces," *Geophysical Journal of the Royal Astronomical Society*, Vol. 11, pp. 293–322.

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.

Zhang, C., and Achenbach, J. D., 1988, "Numerical Analysis of Surface Wave Scattering by the Boundary Element Method," *Wave Motion*, Vol. 10, pp. 365–374.

APPENDIX

Two-Dimensional Time-Harmonic Elastodynamic Fundamental Solutions

The expressions for the displacements are

$$u_{ik}^G(\mathbf{x};\mathbf{x}_p)=\frac{i}{4\mu}\left\{U_1(r)\delta_{ik}-U_2(r)\,\frac{\partial r}{\partial x_i}\,\frac{\partial r}{\partial x_k}\right\}, \qquad \text{(A1)}$$

where $r=|\mathbf{x}-\mathbf{x}_p|$, and

$$U_1(r)=H_0^{(1)}(k_Tr)-\frac{1}{k_Tr}\left\{H_1^{(1)}(k_Tr)-\frac{k_L}{k_T}H_1^{(1)}(k_Lr)\right\}, \qquad \text{(A2)}$$

$$U_2(r)=H_0^{(1)}(k_Tr)-\left(\frac{k_L}{k_T}\right)^2 H_0^{(1)}(k_Lr)$$
$$-\frac{2}{k_Tr}\left\{H_1^{(1)}(k_Tr)-\frac{k_L}{k_T}H_1^{(1)}\ (k_Lr)\right\}, \qquad \text{(A3)}$$

here $H_n^{(1)}(\ )$ is the Hankel function of the nth order of the first kind. The corresponding stresses follow from Hooke's law:

$$\sigma_{ijk}^G(\mathbf{x};\mathbf{x}_p)=\lambda\,\frac{u_{mk}^G}{\partial x_m}\,\delta_{ij}+\mu\left(\frac{\partial u_{ik}^G}{\partial x_j}+\frac{\partial u_{jk}^G}{\partial x_i}\right). \qquad \text{(A4)}$$

T. Kundu
Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, AZ 85721
Mem. ASME

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

Axisymmetric Scattering of a Plane Longitudinal Wave by a Circular Crack in a Transversely Isotropic Solid

The scattering of elastic waves by a circular crack situated in a transversely isotropic solid is studied here. The axis of material symmetry and the axis of the crack coincides. The incident wave is taken as a plane longitudinal wave propagating perpendicular to the crack surface. A Hankel transform representation of the scattered field is used, and after some manipulations using the boundary conditions this leads to an integral equation over the crack for the displacement jump across the crack. This jump is expanded in a series of Legendre polynomials which fulfill the correct edge condition and the integral equation is projected on the same set of Legendre polynomials. The far field is computed by the stationary phase method. A few numerical computations are carried out for both isotropic and anisotropic solids. Results for the isotropic solid compare favorably with those available in the literature.

1 Introduction

Scattering of elastic waves by an internal crack is a problem of considerable importance in the field of quantitative nondestructive evaluation of materials. The circular crack has been extensively studied analytically, because compared to other shapes the circular crack is relatively simple to analyze. Representative examples of previous work can be found in Robertson (1967), Mal (1968a, 1968b, 1968c, 1970), Sih and Loeber (1968, 1969), Martin (1981), Krenk and Schmidt (1982), Kristensson and Waterman (1982), Martin and Wickham (1983), Keogh (1986), Niwa and Hirose (1987), Nishimura and Kobayashi (1988), and others. However, in most of the previous studies the crack is located in an isotropic solid. Some investigations have also been carried out on elastic wave scattering by a circular crack at the interface between two isotropic solids (Srivastava, Palaiya and Gupta, 1979; Boström and Peterson, 1989a) or by a soundhard circular disk at the interface between two fluids (Boström and Peterson, 1989b). No analytical study has yet been carried out to investigate the scattering of elastic waves by a circular crack in an anisotropic solid. For detection and quantitative measurement of internal cracks in a fiber-reinforced composite solid, one needs to study the crack scattering problem in an anisotropic solid. With this goal in mind, the scattering of elastic waves by a circular crack in an anisotropic solid has been studied in this paper.

As a first step, an axisymmetric problem is solved. The

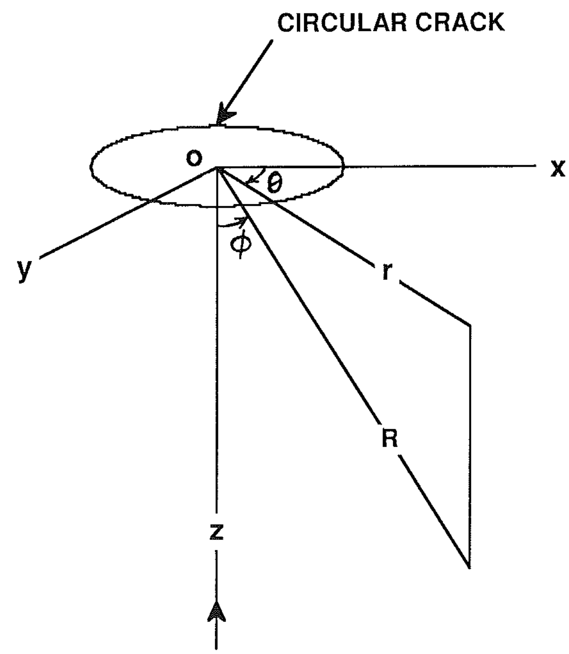


Fig. 1 Schematic diagram of the problem geometry

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Oct. 11, 1989; final revision, June 5, 1990.

composite solid is made of unidirectionally oriented fibers. The fiber diameter and fiber spacing are assumed to be much smaller than the wavelengths of the elastic waves in the material, so that the fiber-reinforced composite solid can be modeled as a transversely isotropic solid with the fiber direction as the axis of symmetry. The circular crack is located perpendicular to the fiber direction. This orientation of the crack naturally appears when it is generated due to fiber breakage. A plane longitudinal wave propagating along the direction of the fibers strikes the crack. The crack opening displacement (COD) and the far-field scattered displacements are studied in this paper. COD is an important parameter to study, since the critical COD determines when a crack starts to propagate; thus the study of COD gives one some insight about the crack propagation. Investigation of the far-field scattered displacements is necessary for internal crack detection by nondestructive testing techniques.

2 Problem Statement

Consider a cylindrical polar coordinate system r, θ, z with the origin at the center of the crack, so that the crack is located at $z = 0$, $0 \le r \le a$, see Fig. 1. Let ω be the circular frequency of the incident waves. In what follows the time dependence of all the field quantities, assumed to be of the form $e^{-i\omega t}$, will be suppressed.

The material is assumed to be a unidirectionally fiber-reinforced composite solid whose fiber diameter is small compared to the wavelength so that one can consider the material as a transversely isotropic solid. The fiber direction is parallel to the z-axis. The stress-strain relation for this material has the following form [Vinson and Sierakowski (1986)]

$$\begin{Bmatrix} \sigma_{rr} \\ \sigma_{\theta\theta} \\ \sigma_{zz} \\ \tau_{\theta z} \\ \tau_{rz} \\ \tau_{r\theta} \end{Bmatrix} = \begin{bmatrix} c_1 & c_2 & c_3 & 0 & 0 & 0 \\ c_2 & c_1 & c_3 & 0 & 0 & 0 \\ c_3 & c_3 & c_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & c_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_6 \end{bmatrix} \begin{Bmatrix} \epsilon_{rr} \\ \epsilon_{\theta\theta} \\ \epsilon_{zz} \\ \gamma_{\theta z} \\ \gamma_{rz} \\ \gamma_{r\theta} \end{Bmatrix} \tag{1}$$

Only five of the above six material constants are linearly independent. c_6 can be expressed in terms of c_1 and c_2:

$$c_6 = \frac{1}{2}(c_1 - c_2) \tag{2}$$

Let a longitudinal wave propagate in the negative z direction as shown in Fig. 1. Since the geometry, material properties, and the loading are symmetric about the z-axis, the response must be axisymmetric. For such an axisymmetric problem the stresses can be expressed in terms of the displacements in the following form

$$\sigma_{rr} = c_1 \frac{\partial u_r}{\partial r} + c_2 \frac{u_r}{r} + c_3 \frac{\partial u_z}{\partial z}$$

$$\sigma_{\theta\theta} = c_2 \frac{\partial u_r}{\partial r} + c_1 \frac{u_r}{r} + c_3 \frac{\partial u_z}{\partial z}$$

$$\sigma_{zz} = c_3 \left(\frac{\partial u_r}{\partial r} + \frac{u_r}{r}\right) + c_4 \frac{\partial u_z}{\partial z}$$

$$\tau_{rz} = c_5 \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right)$$

$$\tau_{r\theta} = \tau_{z\theta} = 0 \tag{3}$$

where u_r and u_z are the radial and vertical displacement components, which are functions of r and z only.

Equations of motion for the problem are

$$c_1\left(\frac{\partial^2 u_r}{\partial r^2} + \frac{1}{r}\frac{\partial u_r}{\partial r} - \frac{u_r}{r^2}\right) + c_3 \frac{\partial^2 u_z}{\partial r \partial z} + c_5\left(\frac{\partial^2 u_z}{\partial r \partial z} + \frac{\partial^2 u_r}{\partial z^2}\right) + \rho\omega^2 u_r = 0$$

$$(c_3 + c_5)\left(\frac{\partial^2 u_r}{\partial r \partial z} + \frac{1}{r}\frac{\partial u_r}{\partial z}\right) + c_5\left(\frac{\partial^2 u_z}{\partial r^2} + \frac{1}{r}\frac{\partial u_z}{\partial r}\right) + c_4 \frac{\partial^2 u_z}{\partial z^2} + \rho\omega^2 u_z = 0 \tag{4}$$

where ρ is the material density. The solution of this system of equations can be assumed to have the following form

$$u_z = \int_0^\infty A(k) e^{ipz} J_0(kr) k dk$$

$$u_r = \int_0^\infty s(k) A(k) e^{ipz} J_1(kr) k dk \tag{5}$$

where J_0 and J_1 are Bessel functions of first kind of orders zero and one, respectively.

Substitution of equation (5) into equation (4) gives after some manipulation

$$s(k) = \frac{c_5 k^2 + c_4 p^2 - \rho\omega^2}{ipk(c_3 + c_5)} \tag{6}$$

and

$$c_4 c_5 p^4 + \{(c_1 c_4 - c_3^2 - 2c_3 c_5)k^2 - \rho\omega^2(c_4 + c_5)\}p^2 + \{c_1 c_5 k^4 - \rho\omega^2(c_1 + c_5)k^2 + \rho^2\omega^4\} = 0. \tag{7}$$

Equation (7) is a quadratic equation in p^2 so it has two roots p_1^2 and p_2^2 which are given by

$$p_j^2 = \frac{1}{2c_4 c_5}\{b_2 - b_1 k^2 + (-1)^j (B_1 k^4 + B_2 k^2 + B_3)^{1/2}\} \tag{8}$$

where

$$b_1 = c_1 c_4 - c_3^2 - 2c_3 c_5$$
$$b_2 = \rho\omega^2(c_4 + c_5)$$
$$B_1 = b_1^2 - 4c_1 c_4 c_5^2$$
$$B_2 = 4\rho\omega^2 c_4 c_5 (c_1 + c_5) - 2b_1 b_2$$
$$B_3 = \rho^2\omega^4(c_4 - c_5)^2.$$

From equation (8) it can be shown that $p_j (j = 1, 2)$ is equal to zero at $k = k_j$ where,

$$k_1 = \omega\sqrt{\frac{\rho}{c_1}}$$

$$k_2 = \omega\sqrt{\frac{\rho}{c_5}}. \tag{9}$$

For $k < k_j$, p_j^2 is positive and p_j is defined to be a positive real number, and for $k > k_j$, p_j^2 is negative and p_j is defined to be a positive imaginary number.

The general axisymmetric solution, hence, have the following form

$$u_z = \pm \int_0^\infty \sum_{j=1}^{2} A_j(k) e^{\pm ip_j z} J_0(kr) k dk$$

$$u_r = \int_0^\infty \sum_{j=1}^{2} s(k) A_j(k) e^{\pm ip_j z} J_1(kr) k dk \tag{10}$$

with $\pm$ for waves (including evanescent ones) in the positive and negative z directions.

For a vertically propagating longitudinal plane wave, there is no r dependence. With $k = 0$ in equation (7), the two roots are $\omega\sqrt{\rho/c_4}$ and $\omega\sqrt{\rho/c_5}$. These two roots correspond to the longitudinal and transverse wave numbers, respectively, in the fiber direction. Thus, the displacement and stress fields corresponding to an incident vertically propagating longitudinal plane wave are

$$u_z^i = e^{-i\omega z\sqrt{\rho/c_4}}$$

$$u_r^i = 0$$

$$\sigma_{zz}^i = -i\omega\sqrt{\rho c_4}\, e^{-i\omega z\sqrt{\rho/c_4}}. \tag{11}$$

3 Integral Equation Formulation

The scattered field is defined such that when it is added to the incident field then the total field is obtained. It can be clearly seen that, since the crack surface is traction-free, the scattered field should produce a traction which is equal but opposite in sign to the incident field traction at the crack position:

$$\sigma_{zz}^s = i\omega\sqrt{\rho c_4} \text{ at } z=0,\ r<a. \tag{12}$$

The scattered field should also satisfy the same governing equations (4), so it must have the form given in equation (10). Using the continuity of σ_{zz}^s for $z = 0$ and all r one can set

$$u_z^s = (\mathrm{sgn} z)\int_0^\infty \sum_{j=1}^{2} A_j(k)e^{ip_j z(\mathrm{sgn}\, z)}J_0(kr)kdk$$

$$u_r^s = \int_0^\infty \sum_{j=1}^{2} s_j(k)A_j(k)e^{ip_j z(\mathrm{sgn}\, z)}J_1(kr)kdk \tag{13}$$

where

$$\mathrm{sgn}\, z = 1,\ z>0$$
$$= -1,\ z<0. \tag{14}$$

It should be noted here that u_r^s becomes automatically continuous across the $z = 0$ plane by this solution. The vertical displacement jump across the $z = 0$ plane is equal to

$$u_z^{s+} - u_z^{s-} = 2\int_0^\infty \sum_{j=1}^{2} A_j(k)J_0(kr)kdk = 0 \qquad r>a$$

$$= \Delta u(r) \qquad r<a \tag{15}$$

where $\Delta u(r)$ is the crack opening displacement (COD). Inversion of the Hankel transform gives

$$A_1(k) + A_2(k) = \frac{1}{2}\int_0^a \Delta u(r)J_0(kr)rdr. \tag{16}$$

From the continuity of τ_{rz}^s across the $z = 0$ plane follows

$$A_2(k) = -\frac{k - ip_1 s_1}{k - ip_2 s_2}A_1(k). \tag{17}$$

Substitution of equation (17) into equation (16) gives

$$A_1(k) = \frac{k - ip_2 s_2}{2i(p_1 s_1 - p_2 s_2)}\int_0^a \Delta u(r)J_0(kr)rdr. \tag{18}$$

After some algebraic manipulations, the boundary condition equation (12) becomes

$$\int_0^\infty A_1(k)\{(c_3 s_1 k + ic_4 p_1) - \frac{k - ip_1 s_1}{k - ip_2 s_2}(c_3 s_2 k + ic_4 p_2)\}J_0(kr)kdk = i\omega\sqrt{\rho c_4}. \tag{19}$$

Finally, equation (18) is substituted into equation (19) to give

$$\int_0^\infty \left(\int_0^a \Delta u(r)J_0(kr)rdr\right)\{(k - ip_1 s_1)(c_3 s_2 k + ic_4 p_2) - (k - ip_2 s_2)(c_3 s_1 k + c_4 p_1)\}\frac{J_0(kr)k}{2(p_1 s_1 - p_2 s_2)}dk = \omega\sqrt{\rho c_4}. \tag{20}$$

This is an integral equation over the crack for the COD $\Delta u(r)$.

The COD $\Delta u(r)$ can be expressed in terms of Legendre polynomials (Krenk and Schmidt, 1982).

$$\Delta u(r) = \sum_{m=0}^{\infty} \alpha_m (-1)^m \frac{P_{2m+1}^0(\sqrt{1-r^2/a^2})}{P_{2m+1}^1(0)} \tag{21}$$

where α_m is an unknown expansion coefficient and P_{2m+1}^j is the associated Legendre function. This expansion incorporates the correct square root behavior of the crack edge (see, e.g., Kuo, 1984).

Using the relation (Krenk, 1982),

$$\int_0^a P_{2m+1}^0(\sqrt{1-r^2/a^2})J_0(kr)rdr = (-1)^m P_{2m+1}^1(0)j_{2m+1}(ka)\frac{a}{k} \tag{22}$$

one can evaluate

$$\int_0^a \Delta u(r)J_0(kr)rdr = \frac{a}{k}\sum_{m=0}^{\infty}\alpha_m j_{2m+1}(ka) \tag{23}$$

where $j_{2m+1}(ka)$ is the spherical Bessel function of first kind of order $2m + 1$ (Abramowitz and Stegun, 1972).

Equation (23) is then substituted into equation (20) to give

$$\frac{a}{2}\int_0^\infty f(k)\frac{J_0(kr)}{p_1 s_1 - p_2 s_2}\sum_{m=0}^{\infty}\alpha_m j_{2m+1}(ka)dk = -\omega\sqrt{\rho c_4} \tag{24}$$

where

$$f(k) = (s_1 - s_2)(c_3 k^2 - p_1 p_2 c_4) + ik(p_1 - p_2)(c_4 + c_3 s_1 s_2). \tag{25}$$

Multiplication of both sides of equation (24) by $rP_{2j+1}^0(\sqrt{1-r^2/a^2})$ and integration over r from 0 to a yields

$$\frac{a}{2}\sum_{m=0}^{\infty}\alpha_m\int_0^\infty \frac{f(k)}{k(p_1 s_1 - p_2 s_2)}j_{2j+1}(ka)j_{2m+1}(ka)dk = -\delta_{j0}\omega\sqrt{\rho c_4}\frac{a}{3} \tag{26}$$

where δ_{j0} is the Kronecker delta.

Thus, equation (26) can be rewritten in the following form:

$$\sum_{m=0}^{\infty} Q_{jm}\alpha_m = -\frac{2}{3}\omega\sqrt{\rho c_4}\delta_{j0} \tag{27}$$

where

$$Q_{jm} = \int_0^\infty \frac{f(k)}{k(p_1 s_1 - p_2 s_2)}j_{2j+1}(ka)j_{2m+1}(ka)dk. \tag{28}$$

Equation (27) can now be solved for α_m; then equations (21), (18), (17), and (13) give the scattered field components u_z^s and u_r^s.

The semi-infinite integral of equation (28) can be converted to finite integrals by considering proper contours as suggested by Krenk and Schmidt (1982). The details are omitted here and only the final result is given:

$$Q_{jm} = \int_0^{k_1}\left[f(k)\frac{j_{2j+1}(ka)h_{2m+1}^{(1)}(ka)}{p_1 s_1 - p_2 s_2} - \delta_{jm}\frac{i\omega\sqrt{\rho c_4}}{a(4j+3)}\frac{1}{k}\right]\frac{dk}{k} + \int_{k_1}^{k_2} g(k)\frac{j_{2j+1}(ka)h_{2m+1}^{(1)}(ka)}{p_1 s_1 - p_2 s_2}\frac{dk}{k} - \delta_{jm}\frac{i\omega\sqrt{\rho c_4}}{k_1 a(4j+3)} \tag{29}$$

where

$$g(k) = ik(c_3 p_1 s_1 s_2 - p_2 c_4) - (s_2 c_3 k^2 + p_1 p_2 s_1 c_4) \tag{30}$$

and $h_{2m+1}^{(1)}(ka)$ is the spherical Hankel function of first kind of order $2m + 1$. Equation (29) can be used to evaluate Q_{jm} for $j \geq m$ and the symmetry $Q_{mj} = Q_{jm}$ then gives the remaining elements.

Two comments are in order concerning the obtained solu-

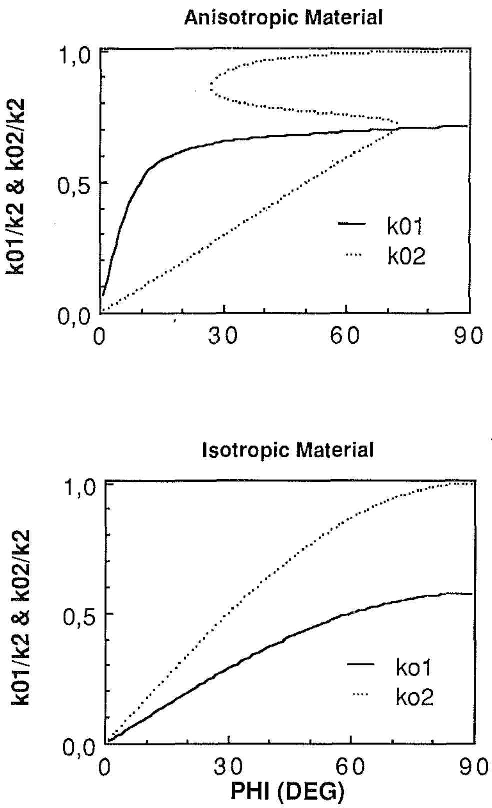


Fig. 2 Variations of stationary points (solutions of equation (35)) for an anisotropic material and an isotropic material. Material properties are given below equation (35).

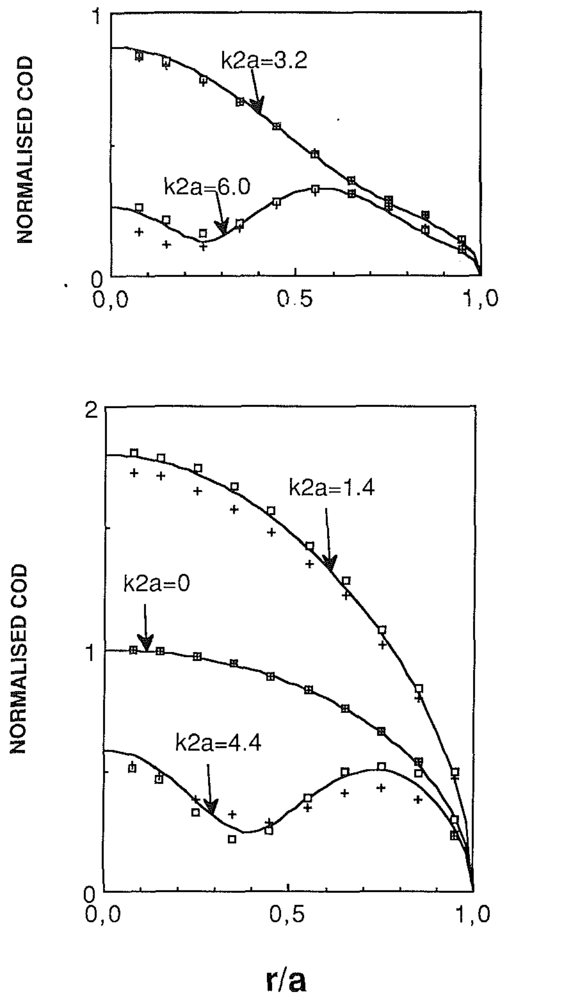


Fig. 3 Normalized COD as a function of r/a in the isotropic solid, for k_2a = 0, 1.4, 3.2, 4.4, and 6.0. Solid lines are from the present analysis, crosses have been obtained by Mal (1970), and white squares are from Budreck and Achenbach (1988).

tion. For the case of an isotropic solid, the system of equations (27) and (28) reduces to the solution given by Krenk and Schmidt (1982) for their symmetric part with $m = 0$ (rotational symmetry), although the normalizations are somewhat different. In the static limit only the last term in the Q matrix in equation (29) survives, and this matrix thus becomes diagonal. The system of equations (27) can thus be solved and this gives

$$\alpha_0 = -2ik_1a$$
$$\alpha_m = 0,\ m = 1,2,\ldots. \quad (31)$$

As in the case of an isotropic solid, the COD in equation (21) in the static limit thus becomes a simple square-root function.

4 Far-Field Calculation

In the near field (small z and r), the integrals in equation (13) can be evaluated numerically to compute u_z^s and u_r^s. However, in the far field, the numerical evaluation of these integrals is awkward because of the rapid oscillations introduced in the integrand by the Bessel functions and the trigonometric functions.

Keeping only the first term in the asymptotic expansion of the Bessel function

$$J_m(kr) \simeq \frac{1}{\sqrt{2\pi kr}} \left\{ e^{ikr - i(2m+1)\frac{\pi}{4}} + e^{-ikr + i(2m+1)\frac{\pi}{4}} \right\}$$

in equation (13) gives, after some algebraic manipulations,

$$u_z^s \simeq (\text{sgn } z)\frac{ia}{\sqrt{8\pi r}} \sum_{j=0}^{\infty} \alpha_j \int_0^{\infty} \frac{j_{2j+1}(ka)}{\sqrt{k}(p_1s_1 - p_2s_2)} \Big\{ (k - ip_1s_1)$$
$$\times e^{ip_2z(\text{sgn } z)} - (k - ip_2s_2)e^{ip_1z(\text{sgn } z)} \Big\} \left\{ e^{ikr - i\frac{\pi}{4}} + e^{-ikr + i\frac{\pi}{4}} \right\} dk$$

$$u_r^s \simeq \frac{ia}{\sqrt{8\pi r}} \sum_{j=0}^{\infty} \alpha_j \int_0^{\infty} \frac{j_{2j+1}(ka)}{\sqrt{k}(p_1s_1 - p_2s_2)} \Big\{ s_2(k - ip_1s_1)e^{ip_2z(\text{sgn } z)}$$
$$- s_1(k - ip_2s_2)e^{ip_1z(\text{sgn } z)} \Big\} \left\{ e^{ikr - \frac{3i\pi}{4}} + e^{-ikr + \frac{3i\pi}{4}} \right\} dk. \quad (32)$$

So, oscillations in the integrand are introduced by the factors of the form

$$e^{\pm ip_jz \pm ikr}.$$

This type of oscillatory integrals can easily be computed by the method of stationary phase (Lighthill, 1978; Jeffreys and Jeffreys, 1950).

For this purpose, a spherical coordinate system R, ϕ, θ is now introduced as shown in Fig. 1. Then

$$e^{ip_jz \pm ikr} = e^{iR(p_j \cos \phi \pm k \sin \phi)}$$
$$= e^{iR\psi_j(k)} \text{ or } e^{iR\chi_j(k)} \quad (33)$$

where

$$\psi_j(k) = p_j \cos \phi + k \sin \phi$$
$$\chi_j(k) = p_j \cos \phi - k \sin \phi. \quad (34)$$

The stationary points can be obtained from $\psi_j'(k) = 0$ and $\chi_j'(k) = 0$, which gives

$$\frac{2b_1k - (-1)^j(2B_1k^3 + B_2k)/B(k)}{(b_2 - b_1k^2 + (-1)^jB(k))^{1/2}} = \pm\sqrt{8c_4c_5} \tan \phi \quad (35)$$

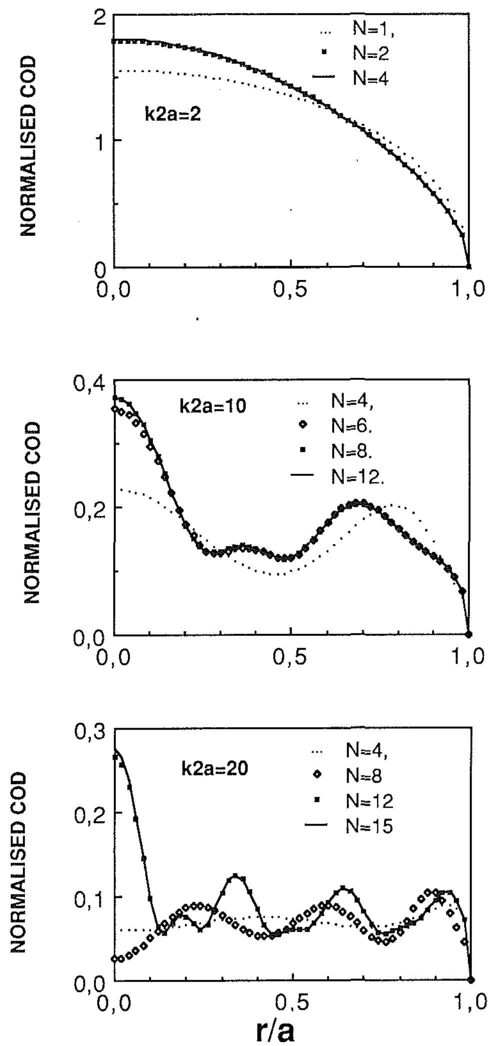


Fig. 4 COD, in the isotropic solid, computed by considering different truncations in equation (27) for $k_2a = 2$ (top figure), 10 (middle figure), and 20 (bottom figure).

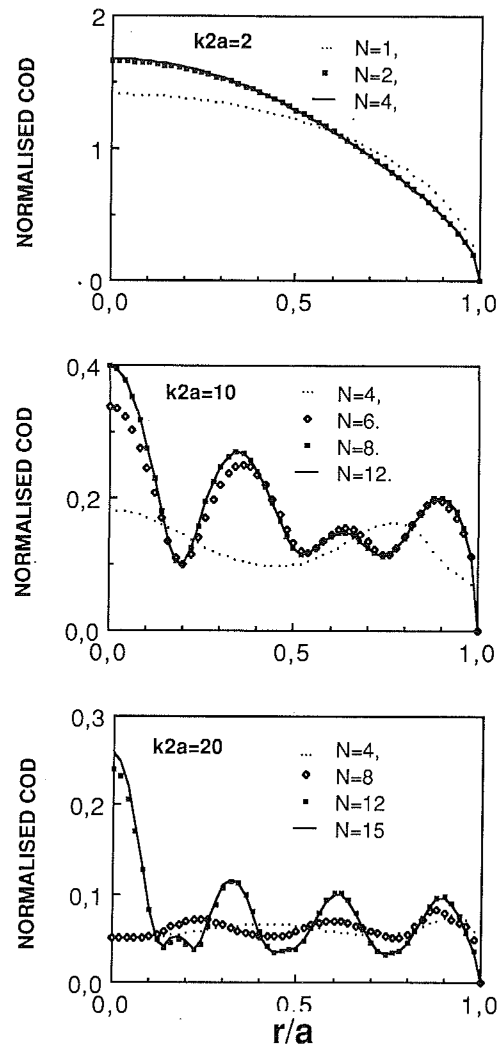


Fig. 5 Same as Fig. 4, but the material is the graphite-epoxy composite whose material properties are given below equation (35)

where

$$B(k) = (B_1k^4 + B_2k^2 + B_3)^{1/2}.$$

The roots of equation (35) can be obtained numerically. For an isotropic material, these roots become simply equal to $\pm k_j \sin\phi$, but for an anisotropic material, no such simple relation exists. For a graphite-epoxy composite material no positive root of k is obtained from $\chi_j' = 0$. $\psi_1'(k) = 0$ gives one positive root and $\psi_2'(k) = 0$ gives multiple positive roots for some values of ϕ and a single positive root for other values of ϕ. Existence of these multiple roots is in consistence with the theory of elastic wave propagation in anisotropic materials that states that in a transversely isotropic solid along certain directions, the shear wave can propagate with more than one velocity (Van der Hijden, 1987). Denote these roots by k_{01} and k_{02}^m, respectively. For multiple values of k_{02}, the superscript m takes the values 1, 2, 3. The variation of k_{01} and k_{02}^m with ϕ for the graphite-epoxy composite ($c_1 = 13.92$ GPa, $c_2 = 6.92$ GPa, $c_3 = 6.44$ GPa, $c_4 = 160.73$ GPa, $c_5 = 7.07$ GPa, $\rho = 1578$ kg/m^3, Mal, Yin, and Bar-Cohen, 1989) and for an isotropic solid (Young's modulus $E = 69.15$ GPa, Poisson's ratio $\nu = 0.25$, $\rho = 2770$ kg/m^3) are shown in Fig. 2. For the isotropic solid the k_{01} and k_{02} curves match exactly with $k_1 \sin\phi$ and $k_2 \sin\phi$ curves, respectively.

During the far-field computation one can neglect all terms containing e^{-ikr} in equation (32) since $\chi_j'(k) = 0$ does not give any positive root. Collecting everything the stationary phase method then gives the far-field components

$$u_z^s \simeq (\operatorname{sgn} z)\,\frac{ia}{2R\sqrt{\sin\phi}}\sum_{j=0}^{\infty}\alpha_j\left\{\sum_{m=1}^{M}F_j(k_{02}^m)e^{iR\psi_2(k_{02}^m)} - G_j(k_{01})e^{iR\psi_1(k_{01})}\right\}$$

$$u_r^s \simeq \frac{ia}{2R\sqrt{\sin\phi}}\sum_{j=0}^{\infty}\alpha_j\left\{is_1G_j(k_{01})e^{iR\psi_1(k_{01})} - \sum_{m=1}^{M}is_2F_j(k_{02}^m)e^{iR\psi_2(k_{02}^m)}\right\} \quad (36)$$

where M is the number of roots of the equation $\psi_2'(k) = 0$ and

$$F_j(k) = \frac{j_{2j+1}(ka)\,(k - ip_1s_1)}{|\psi_2''(k)|^{1/2}\sqrt{k}\,(p_1s_1 - p_2s_2)}\,e^{-i\frac{\pi}{4}+i\frac{\pi}{4}\operatorname{sgn}[\psi_2''(k)]}$$

$$G_j(k) = \frac{j_{2j+1}(ka)\,(k - ip_2s_2)}{|\psi_1''(k)|^{1/2}\sqrt{k}\,(p_1s_1 - p_2s_2)}\,e^{-i\frac{\pi}{4}+i\frac{\pi}{4}\operatorname{sgn}[\psi_1''(k)]} \quad (37)$$

and

$$\psi_j''(k) = -\frac{\cos\phi}{\sqrt{8c_4c_5}}\,[\{2b_1 - (-1)^jB(k)^{-3}(2B_1^2k^6 + 3B_1B_2k^4$$

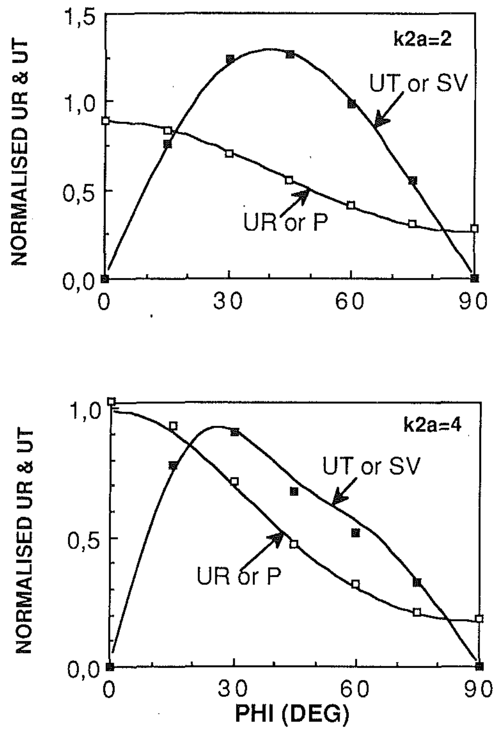


Fig. 6 Normalized radial and tangential displacements (continuous lines) in the far field as a function of the angle ϕ. The material is the isotropic solid. Results of Krenk and Schmidt (1982) are shown by white squares for scattered P waves and by black squares for scattered SV waves.

$$+6B_1B_3k^2+B_2B_3)\}\{b_2-b_1k^2+(-1)^jB(k)\}+\frac{1}{2}\{2b_1k -(-1)^jB(k)^{-1}(2B_1k^3+B_2k)\}^2]\{b_2-b_1k^2 +(-1)^jB(k)^{-1}\}^{-3/2} \quad (38)$$

where b_1, b_2, $B(k)$, B_1, B_2, and B_3 have been defined previously.

Energy conservation generally gives a good check on a numerical procedure and it is often easiest to apply in the far field. It has been briefly investigated for the present case, but due to the rather implicit form of the far field, it would be somewhat cumbersome computationally and has therefore been abandoned.

5 Numerical Results

Numerical results are presented for two types of materials, an isotropic solid and an anisotropic solid. Their material properties are given in the previous section (see below equation (35)).

Figure 3 shows the COD of a circular crack in an isotropic solid when a vertically propagating longitudinal wave strikes the crack. Different curves are shown for different values of k_2a. In this and the following figures, the COD has been normalized with respect to W_0, where W_0 is the COD at the crack center when the crack is subjected to a static stress of magnitude equal to the amplitude of the incident stress field. Equations (21) and (31) gives

$$W_0=-2ik_1a. \quad (39)$$

In Fig. 3, crosses are results obtained by Mal (1970) and white squares are those obtained by Budreck and Achenbach (1988). The present results are closer to Budreck and Achenbach's results than to Mal's results.

Equation (27) gives a set of linear equations containing an infinite number of terms. But for all practical purposes one needs to consider only a finite number of terms. This needed

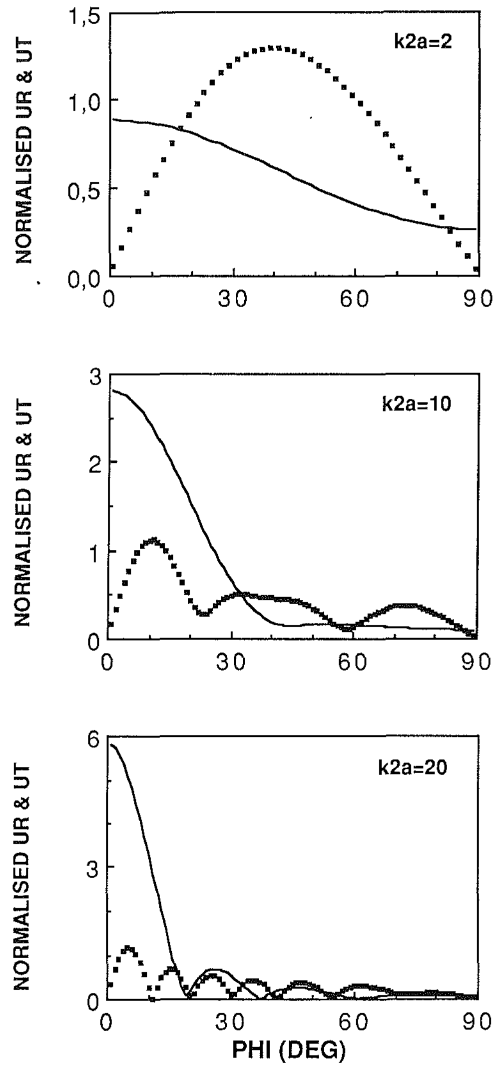


Fig. 7 Normalized radial (continuous line) and tangential (black squares) components of the scattered field displacement in the isotropic solid

truncation increases with increasing k_2a. Figures 4 and 5 show the COD evaluated in the isotropic and the anisotropic solids, respectively, for different values of k_2a when different truncations are considered in equation (27). It can be seen from these two figures that the truncations required for proper convergence is a function of k_2a; these truncations are 2, 8, and 12 for k_2a equal to 2, 10, and 20, respectively. For these calculations, the integrals in equation (29) could be accurately computed with the 60-point Gaussian quadrature scheme.

Next, for the isotropic solid the radial and tangential components (u_R^s and u_T^s) of displacement are computed using equation (36) and the relations

$$u_R^s=u_z^s\cos\phi+u_r^s\sin\phi$$
$$u_T^s=-u_z^s\sin\phi+u_r^s\cos\phi. \quad (40)$$

Thus, the contributions of scattered P and SV waves are separated by the u_R^s and u_T^s components. They are computed for k_2a = 2 and 4 and plotted in Fig. 6. Squares in these plots have been obtained by Krenk and Schmidt (1982). Scattered displacements in this figure and in the subsequent figures have been normalized with the factor U_0a/R, where U_0 is the displacement amplitude of the incident wave, a is the crack radius, and R is the radial distance of the point of interest from the crack center.

Figure 7 shows the scattered displacements in the far field

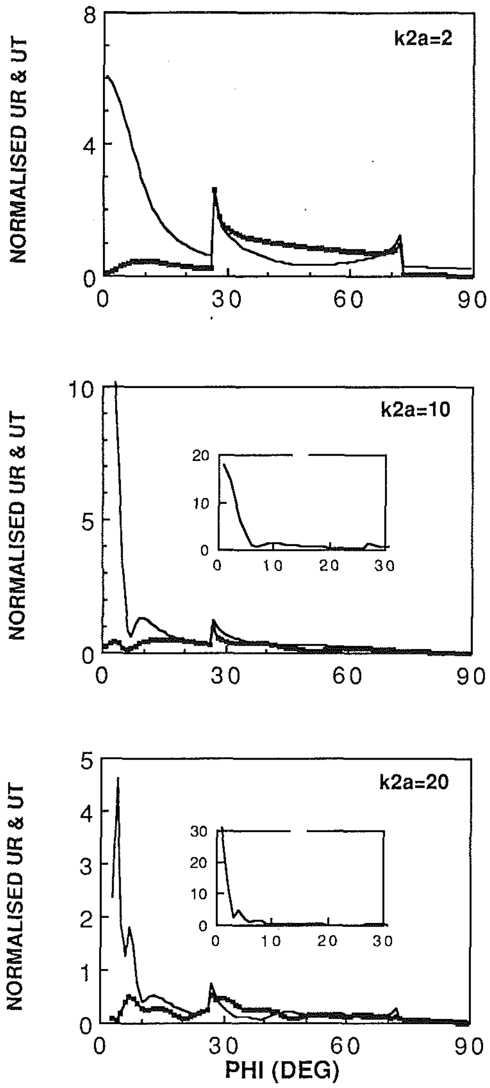


Fig. 8 Same as Fig. 7, but the material is the graphite-epoxy composite solid. In the middle and bottom plots the smaller inserted curves show radial displacements plotted from ϕ = 1 deg, whereas the bigger curves are plotted from ϕ = 3 deg.

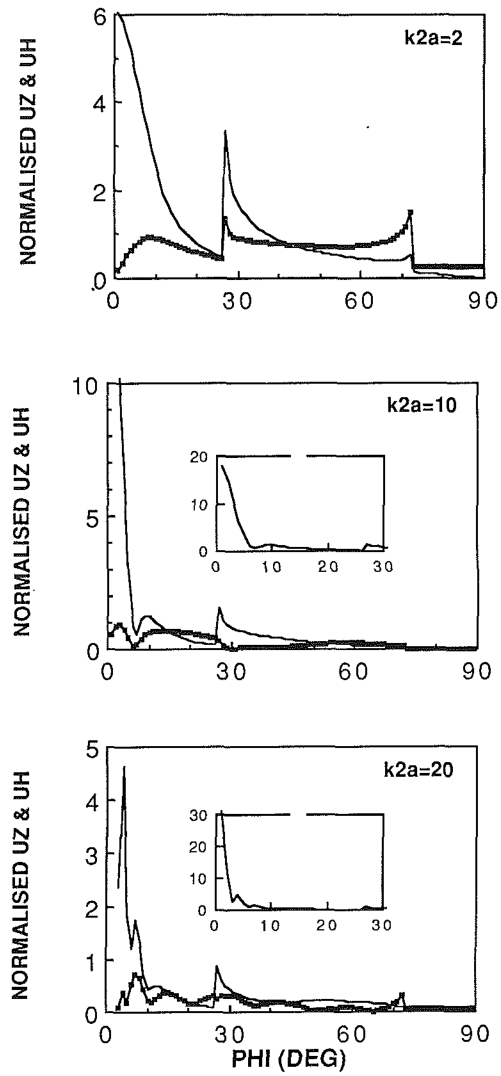


Fig. 9 Same as Fig. 8, but instead of radial and tangential components, the vertical (along the fiber) and horizontal (perpendicular to the fiber) components of the scattered field displacements are plotted

for the isotropic solid for k_2a = 2, 10, and 20. Both radial (scattered P wave) and tangential (scattered SV wave) components of displacement are shown in solid and dotted lines, respectively. As the frequency increases, the P wave curve becomes narrower and the SV wave curve gradually diminishes. Thus, as expected, the results become closer to the ray approximations as the frequency increases.

Figure 8 is similar to Fig. 7, but for this figure the material is the anisotropic solid. In Fig. 8 the radial components decay with the angle much more rapidly than those in Fig. 7. This phenomenon is due to the presence of strong fibers in the z-direction. Because of these fibers, the disturbance quickly propagate in the z-direction or, in other words, the scattered wave remains closer to the symmetry axis of the crack. As the frequency increases, these peaks become sharper. In the middle and bottom plots the smaller inserted curves show radial displacements plotted from ϕ = 1 deg, whereas the bigger curves are plotted from ϕ = 3 deg to magnify the tail portion of the curves. Figure 9 is similar to Fig. 8, but here vertical (along the fiber) and horizontal (perpendicular to the fiber) components of displacements have been plotted.

6 Concluding Remarks

In the present paper the scattering by a circular crack in an anisotropic solid has been considered. The integral equation method employed is rather direct and it has the virtue that the unknown is a physically interesting quantity, namely the COD. The numerical examples show that it is possible to consider relatively high frequencies and the results then become closer to the expectations from simple ray theory. Both COD and far-field scattered displacements in isotropic solids compare favorably with those available in the literature. It is interesting to note that in an isotropic solid the scattered field is spread over a wider region as compared to an anisotropic solid.

The present work is limited to the axisymmetric case. Generalization of this technique, to the case of inclined waves and multilayered anisotropic solids, will be more cumbersome. The case with inclined incident waves are presently being investigated, and the main extra complication is that also the antisymmetric part of the problem must also be solved and this involves two coupled integral equations (cf., Krenk and Schmidt, 1982).

Acknowledgment

This research was partially supported by the National Swedish Board for Technical Development (STU) and the National Science Foundation (NSF) under contract number DMC-8807661 and this is gratefully acknowledged.

References

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

Boström, A., and Peterson, L., 1989a, "Wave Scattering by a Circular Crack in the Interface Between Two Elastic Media," Report 1989:7, Division of Mechanics, Chalmers University of Technology, Göteburg, Sweden.

Boström, A., and Peterson, L., 1989b, "Scattering of Acoustic Waves by a Circular Disc in the Interface Between Two Fluids," *J. Acoust. Soc. Am.*, submitted for publication.

Budreck, D. E., and Achenbach, J. D., 1988, "Scattering from Three-Dimensional Planar Cracks by the Boundary Integral Equation Method," ASME Journal of Applied Mechanics, Vol. 55, pp. 405–412.

Jeffreys, H., and Jeffreys, B. S., 1950, *Methods of Mathematical Physics*, 2nd ed., Cambridge University Press, Cambridge, U.K., pp. 505–507.

Keogh, P. S., 1986, "High-Frequency Scattering of a Normally Incident Plane Compressional Wave by a Penny-Shaped Crack," *Q. J. Mech. Appl. Math.*, Vol. 39, pp. 535–566.

Krenk, S., 1982, "Some Integral Relations of Hankel Transform Type and Applications to Elasticity Theory," *Integral Equations and Operator Theory*, Vol. 5, pp. 548–561.

Krenk, S., and Schmidt, H., 1982, "Elastic Wave Scattering by a Circular Crack," *Phil. Trans. Roy. Soc. London*, Vol. A308, pp. 167–198.

Kristensson, G., and Waterman, P. C., 1982, "The T Matrix for Acoustic and Electromagnetic Scattering by Circular Discs," *J. Acoust. Soc. Am.*, Vol. 72, pp. 1612–1625.

Kuo, A.-Y., 1984, "Transient Stress Intensity Factors of an Interfacial Crack Between Two Dissimilar Anisotropic Half-Spaces," ASME Journal of Applied Mechanics, Vol. 51, pp. 71–76.

Lighthill, J., 1978, *Waves in Fluids*, Cambridge University Press, Cambridge, U.K.

Mal, A. K., 1968a, "Diffraction of Elastic Waves by a Penny Shaped Crack," *Q. J. Appl. Math.*, Vol. 26, pp. 231–237.

Mal, A. K., 1968b, "Dynamic Stress Intensity Factor for an Axisymmetric Loading of the Penny Shaped Crack," *Int. J. Eng. Sci.*, Vol. 6, pp. 623–629.

Mal, A. K., 1968c, "Dynamic Stress Intensity Factor for a Non-Axisymmetric Loading of the Penny Shaped Crack," *Int. J. Eng. Sci.*, Vol. 6, pp. 725–733.

Mal, A. K., 1970, "Interaction of Elastic Waves with a Penny-Shaped Crack," *Int. J. Eng. Sci.*, Vol. 8, pp. 381–388.

Mal, A. K., Yin, C.-C., and Bar-Cohen, Y., 1989, "A Theoretical and Experimental Study of Guided Waves in Unidirectional Fiber-Reinforced Composite Plates," *Wave Motion*, submitted for publication.

Martin, P. A., 1981, "Diffraction of Elastic Waves by a Penny Shaped Crack," *Proc. R. Soc. London*, Vol. A378, pp. 263–285.

Martin, P. A., and Wickham, G. R., 1983, "Diffraction of Elastic Waves by a Penny-Shaped Crack: Analytical and Numerical Results," *Proc. R. Soc. London*, Vol. A390, pp. 91–129.

Nishimura, N., and Kobayashi, S., 1988, "An Improved Boundary Integral Equation Method for Crack Problems," *IUTAM Symposium on Boundary Elements*, Springer-Verlag, New York.

Niwa, Y., and Hirose, S., 1987, "Application of the BEM to Elastodynamics in a Three Dimensional Half Space," *Recent Applications in Computational Mechanics*, ASCE, New York.

Robertson, I. A., 1967, "Diffraction of a Plane Longitudinal Wave by a Penny-Shaped Crack," *Proc. Camb. Phil. Soc.*, Vol. 63, pp. 229–238.

Sih, G. C., and Loeber, J. F., 1968, "Torsional Vibration of an Elastic Solid Containing a Penny-Shaped Crack," *J. Acoust. Soc. Am.*, Vol. 44, pp. 1237–1245.

Sih, G. C., and Loeber, J. F., 1969, "Normal Compression and Radial Shear Waves Scattering at a Penny-Shaped Crack in an Elastic Solid," *J. Acoust. Soc. Am.*, Vol. 46, pp. 711–721.

Srivastava, K. N., Palaiya, R. M., and Gupta, O. P., 1979, "Interaction of Longitudinal Wave with a Penny-Shaped Crack at the Interface of Two Bonded Dissimilar Solids—II," *Int. J. Fracture*, Vol. 15, pp. 591–599.

Vinson, J. R., and Sierakowski, R. L., 1986, *The Behavior of Structures Composed of Composite Materials*, Martinus Nijhoff Publishers, The Netherlands.

Transient Analysis for Antiplane Crack Subjected to Dynamic Loadings

Chien-Ching Ma
Professor.

Ying-Chung Hou
Graduate Student.

Department of Mechanical Engineering,
National Taiwan University,
Taipei, Taiwan 10764

The problem considered here is the antiplane response of an elastic solid containing a half-plane crack subjected to suddenly applied concentrated point forces acting at a finite distance from the crack tip. A fundamental solution for the dynamic dislocation is obtained to construct the dynamic fracture problem containing a characteristic length. Attention is focused on the time-dependent full-field solutions of stresses and stress intensity factor. It is found that at the instant that the first shear wave reaches the crack tip, the stress intensity factor jumps from zero to the appropriate static value. The stresses will take on the appropriate static *value instantaneously upon arrival of the shear wave diffracted from the crack tip, and this static value is thereafter maintained. The dynamic stress intensity factor of a kinked crack from this stationary semi-infinite crack after the arrival of shear wave is obtained in an explicit form as a function of the kinked crack velocity, the kink angle, and time. A perturbation method, using the kink angle as the perturbation parameter, is used. If the maximum energy release rate is accepted as the crack propagation criterion, then the crack will propagate straight ahead of the original crack when applying point load at the crack face.*

1 Introduction

Most of the analysis done regarding cracked bodies are quasi-static. Because of loading conditions and material characteristics, there are numerous problems for which the assumption that the deformation is quasi-static is invalid and the inertia of the material must be taken into account. The inherent time dependence of the dynamic fracture problems makes them more complex than equivalent quasi-static models. Both the case of a stationary crack in a body subjected to dynamic loading and the case of a rapidly propagating crack in a stressed body are considered as dynamic fracture problems.

When dynamic loading is applied to a body with an internal crack, the resulting stress waves may initiate crack growth. Few solutions for a cracked elastic solid subjected to dynamic loading are available. The most notable of these are the analysis of diffraction of a plane pulse for a semi-infinite crack by de Hoop (1958) and the equivalent problem for a finite length crack by Thau and Lu (1971). The study of propagation crack in a brittle solid began with the pioneering analysis of Yoffe (1951), and considerable progress has been made in the area of dynamic brittle fracture. Transient problems for constant crack propagation velocity along the fracture plane have been studied by Baker (1962), Broberg (1960), and Achenbach and Nuismer (1971). In a series of papers, Freund (1972a, 1972b, 1973, 1974a) developed important analytical methods for evaluation of the transient stress field of a propagating crack in a two-dimensional geometric configuration. In Freund's papers, a fundamental solution is obtained and is used to develop the solution for negation the stress distribution on the prospective fracture plane by superposition. A generalization of this idea also led to solutions of crack kinking problems under dynamic loading analyzed by Ma and Burgers (1986, 1987, 1988) and Ma (1988).

The difficulty in determining the transient stress field in a cracked body subjected to dynamic loading is well known. The complete solution of a spatially uniform traction distribution acting on the crack faces can be obtained by integral transformation methods. If the problem is modified by replacing a nonuniform distribution having a characteristic length, then the same solution procedure using integral transformation methods does not apply. Freund (1974b) developed a technique which makes it possible to solve this modified problem. Freund solved the problem of an elastic solid containing a half-plane crack subjected to concentrated impact loading on the faces of the crack. An exact expression for the dynamic stress intensity factor was derived by superposition over a one-parameter family of continuously distributed moving dislocations. The complete elastic solution can also be determined by this scheme, but only the stress intensity factor was studied in detail by Freund (1974b). Freund found that if the applied point loading on the crack faces is a step function of time dependence, the dynamic stress intensity factor is zero until the longitudinal wave, which was generated at the loading point, arrives at the crack tip. At the instant the Rayleigh wave arrives,

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 12, 1990; final revision, June 15, 1990.

the stress intensity factor takes on its appropriate static value, and this value in maintained thereafter. Dynamic stress wave interaction with cracks was analyzed by Brock (1982, 1984, 1986) and Brock et al. (1985). The analysis of the interaction of dynamic dislocations with stationary semi-infinite cracks was studied by Brock (1983a,b). Brock also focused his attention mainly on the investigation of the dynamic stress intensity factor.

In some classes of dynamic problems of impact loading, the ability to find a static field may hinge on waiting for the wave front to pass and the transient effect to die away. For point dynamic loading with the step function suddenly applied on the surface of a half plane, the stress field becomes a static value as the time tends to infinity. When a propagating antiplane crack subjected to dynamic loading suddenly stops, Eshelby (1969) and Ma and Burgers (1988) found that the stationary crack solution is radiated out behind the shear wave centered at the stopped crack tip. In this study, the problem to be considered is the antiplane response of an elastic solid containing a half-plane crack subjected to impact loading on the crack faces with finite distance to the crack tip as shown in Fig. 1. The techniques used in this study were first described by Freund (1974b) who investigated the same problem in the plane case but focused only on the dynamic stress intensity factor. In this study, we analyze not only the dynamic stress intensity factor but also the transient full-field solution. The main results are that the stress intensity factor is zero before the shear wave arrives at the crack tip, and then it jumps from zero to its appropriate static value at the instant of the wave arrival. The full-field solution of stresses will take on the appropriate static value instantaneously upon arrival of the secondary shear wave (SS wave) diffracted from the shear wave (S wave) which is generated by the suddenly applied load. A kinked crack which suddenly propagates out of the original semi-infinite crack with constant velocity is also considered. The direction of propagation, as well as the velocity of crack propagation will depend on the local stress field around the crack tip. To understand the observed bifurcation events, the dynamic stress intensity factor for cracks which suddenly kink is obtained in closed form by a perturbation method. The energy flux into the propagating kinked crack tip is derived and these results are discussed in terms of an assumed fracture criterion.

2 Statements of the Problem

Consider a stress-free elastic homogeneous isotropic infinite medium that contains a semi-infinite crack, a Cartesian coordinate system is defined in the body in such a way that the antiplane deformation is in the y-direction. The planar crack lies in the plane $z=0$, $x<0$. At time $t=0$, a concentrated force of magnitude σ_0 (per unit length in the y-direction) acts at $x=-l$ on each face of the crack as shown in Fig. 1. The relevant stress components are denoted by σ_{yz} and σ_{xy}, and the nonzero out-of-plane displacement is denoted by w. In a stationary coordination systems of x and z, two-dimensional antiplane wave motions are governed by

$$\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial z^2}=b^2\frac{\partial^2 w}{\partial t^2}, \quad (1)$$

where b is the slowness of the transverse wave given by

$$b=\frac{1}{v_s}=\sqrt{\frac{\rho}{\mu}}.$$

Here, μ and ρ are the shear modulus and the mass density of the material, respectively. The nonvanishing shear stresses are

$$\sigma_{xy}=\mu\frac{\partial w}{\partial x},\quad \sigma_{yz}=\mu\frac{\partial w}{\partial z}.$$

The crack faces are traction-free, except for the point of application of the concentrated forces. Because of symmetry with respect to the plane $z=0$, the boundary conditions can be written as

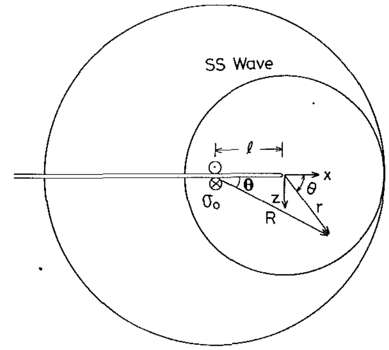


Fig. 1 Wavefronts for a stationary crack subject to dynamic point loading at the crack faces

$$\sigma_{yz}(x,\,0,\,t)=\sigma_0\delta(x+l)H(t),\ x<0 \quad (2)$$

$$w(x,\,0,\,t)=0,\ x>0, \quad (3)$$

where H is the Heaviside step function and δ is the Dirac delta function. The formulation is completed by specifying zero initial condition. Because of the presence of the characteristic length l in the formulation, the standard Wiener-Hopt technique cannot be used. Therefore, some other approach must be followed. If the boundary condition (2) is extended to the entire boundary, then the problem is reduced to the antiplane analog of Lamb's problem in the plane case. Hence, the problem described in (2) and (3) is Lamb's problem with a concentrated loading at $x=-l$, but with surface displacement negated for $x>0$. By solving for the fundamental solution of a distribution dislocation, this problem can be solved by superposition. This methodology was first discussed and used to solve the correspondence problem in plane strain by Freund (1974b).

3 Transient Solutions for Impact Loading on the Crack Faces

Now let us consider the same unbounded body containing a semi-infinite crack. At time $\tau=0$, a screw dislocation of strength 2Δ begins to move from the crack tip at constant speed v in the positive x-direction. This problem is also symmetric with respect to the plane $z=0$, the boundary conditions can be written as

$$\sigma^F_{yz}(x,\,0,\,\tau)=0,\ x<0 \quad (4)$$

$$w^F(x,\,0,\,\tau)=\Delta H(v\tau-x),\ x>0. \quad (5)$$

The solution of this fundamental problem can be obtained by using integral transformation and the standard Wiener-Hopf technique. The exact full-field solutions can be expressed as follows:

$$\sigma^F_{yz}(x,\,z,\,\tau)=-\frac{\Delta\mu(b+h)^{1/2}}{\pi}\mathrm{Im}\left[\frac{(b+\lambda)^{1/2}}{\lambda+h}\frac{\partial\lambda}{\partial\tau}\right]H(\tau-br), \quad (6)$$

$$\sigma^F_{xy}(x,\,z,\,\tau)=\frac{\Delta\mu(b+h)^{1/2}}{\pi}\mathrm{Im}\left[\frac{\lambda}{(b-\lambda)^{1/2}(\lambda+h)}\frac{\partial\lambda}{\partial\tau}\right]H(\tau-br), \quad (7)$$

$$w^F(x,\,z,\,\tau)=\frac{\Delta(b+h)^{1/2}}{\pi}\int_{br}^{\tau}\mathrm{Im}\left[\frac{1}{(b-\lambda)^{1/2}(\lambda+h)}\frac{\partial\lambda}{\partial\tau}\right]d\tau\,H(\tau-br), \quad (8)$$

and

$$\lim_{x\to 0^+}[x^{1/2}\sigma_{yz}^F(x,\,0,\,\tau)]=-\frac{\mu\Delta(b+h)^{1/2}}{\pi\tau^{1/2}},\tag{9}$$

where

$$h=1/v,$$

$$\lambda=-\frac{\tau}{r}\cos\theta+i\sin\theta\left(\frac{\tau^2}{r^2}-b^2\right)^{1/2},$$

$$r=\sqrt{x^2+z^2}.$$

Now, consider the transient wave in a half-space generated by a impact point loading σ_0 at $x=0$, $z=0$, which can be viewed as the Lamb's problem in antiplane analogy. The solutions are

$$\sigma_{yz}^L(x,\,z,\,t)=\frac{\sigma_0 t\sin\theta}{\pi br^2[(t/br)^2-1]^{1/2}}H(t-br),\tag{10}$$

$$\sigma_{xy}^L(x,\,z,\,t)=\frac{\sigma_0 t\cos\theta}{\pi br^2[(t/br)^2-1]^{1/2}}H(t-br),\tag{11}$$

$$w^L(x,\,z,\,t)=-\frac{\sigma_0}{\mu\pi}\ln\left[\frac{t}{br}+\sqrt{\left(\frac{t}{br}\right)^2-1}\right]H(t-br).\tag{12}$$

With the fundamental solution and Lamb's solution at hand, it is now possible to construct the field solutions for the problem of impact loading on the crack faces at $x=-l$. As described in the previous section, this solution can be superimposed by two solutions, one is the Lamb's problem with a concentrated force at $x=-l$, $z=0$, the other problem is that which cancels out the surface displacement for $z=0$, $x>0$ of Lamb's problem. The surface displacement for Lamb's problem by applying concentrated loading σ_0 at $x=-l$, $z=0$ and at time $t=0$ is obtained from (12)

$$w^L(x,\,0,\,t)=-\frac{\sigma_0}{\mu\pi}\mathrm{Im}\left[\frac{t}{b(x+l)}+\sqrt{\left(\frac{t}{b(x+l)}\right)^2-1}\right]H(t-b(x+l)).\tag{13}$$

It is observed from (13) that w^L depends only on the ratio $t/(x+l)$, which means any given displacement level radiates out at a constant speed $(x+l)/t$ along the x-axis for $t>0$. The speed varies between zero and the shear wave speed. For a particular speed v_τ arriving at $x=0$, at any time t will be $v_\tau=l/t$. Then, the full-field transient solution of stress $\sigma_{yz}(x,\,z,\,t)$ can be constructed by superposition over a one-parameter family of dislocation velocity. The result is

$$\sigma_{yz}(x,\,z,\,t)=\sigma_{yz}^L(x,\,z,\,t)-\int_{v_s}^{v_\tau}\sigma_{yz}^F(x,\,z,\,\tau-\tau_0,\,v)\frac{dw^L(v)}{dv}dv.\tag{14}$$

The coordinate systems and the wavefronts are shown in Fig. 1. The wavefront for Lamb's problem consists of only the shear wavefront, denoted by S, propagating away from the loading point. When this shear wave reaches the crack tip, an additional shear wave indicated by SS diffracted from the crack tip is generated. The first term in (14) represents the contribution from the S wave while the second term is from the SS wave. It is convenient to change the integration variable from v to $h=1/v$. After the substitution of the explicit expressions for σ_{yz}^L, σ_{yz}^F and dw^L/dv into (14), the expression becomes

$$\sigma_{yz}(r,\,\theta,\,t)=\frac{\sigma_0 t\sin\Theta}{\pi bR^2[(t/bR)^2-1]^{1/2}}H(t-bR)-\frac{\sigma_0}{\pi^2}\int_b^{(t-br)/l}\mathrm{Im}\left[\frac{(b+\eta_1)^{1/2}}{\eta_1+h}\frac{\partial\eta_1}{\partial\tau}\right]\frac{dh}{(h-b)^{1/2}}H(t-b(r+l)),\tag{15}$$

where

$$\eta_1=-\frac{t-lh}{r}\cos\theta+i\sin\theta\left[\left(\frac{t-lh}{r}\right)^2-b^2\right]^{1/2},\tag{16}$$

$$\frac{\partial\eta_1}{\partial\tau}=\frac{1}{r}\left[-\cos\theta+i\sin\theta\frac{t-lh}{((t-lh)^2-b^2r^2)^{1/2}}\right],$$

$$R=[(l+r\cos\theta)^2+r^2\sin^2\theta]^{1/2}.\tag{17}$$

By a similar procedure, we also get the result for σ_{xy}

$$\sigma_{xy}(r,\,\theta,\,t)=\frac{\sigma_0 t\cos\Theta}{\pi bR^2[(t/bR)^2-1]^{1/2}}H(t-bR)+\frac{\sigma_0}{\pi^2}\int_b^{(t-br)/l}\mathrm{Im}\left[\frac{\eta_1}{(\eta_1+\mathrm{h})(\mathrm{b}-\eta_1)^{1/2}}\frac{\partial\eta_1}{\partial\tau}\right]\frac{dh}{(h-b)^{1/2}}H(t-b(r+l)).\tag{18}$$

The stress intensity factor is defined by the following limit

$$K=\lim_{x\to 0^+}(2\pi x)^{1/2}\sigma_{yz}(x,\,0,\,t).\tag{19}$$

It is clear that the stress in Lamb's problem is not singular at $x=0$, so that the stress intensity factor is determined by the second term in (15). From the result of (9) and in the same manner as the construction of the full-field solution, it is found that the stress intensity factor is given by

$$K(t)=-\sigma_0\sqrt{\frac{2}{\pi l}}\frac{1}{\pi}\int_b^{h^*}\frac{dh}{(h^*-h)^{1/2}(h-b)^{1/2}}H(t-bl)=-\sigma_0\sqrt{\frac{2}{\pi l}}H(t-bl),\tag{20}$$

where $h^*=t/l$. The interesting result of (20) is that the stress intensity factor jumps from zero to the static value after the shear wave generated from the loading point arrives at the crack tip. In the plane-strain case, Freund (1974b) found that the stress intensity factor takes on its static value instantaneously upon arrival of the Rayleigh wave generated by the suddenly applied load.

Now, we focus our attention on the stress along the crack-tip line. After making the indicated change of variable, the result is

$$\sigma_{yz}(x,\,0,\,t)=\frac{\sigma_0}{\pi x^{1/2}}\int_b^{1/v_\tau}\frac{(t-lh-xb)^{1/2}}{(hx+hl-t)\sqrt{h-b}}dh\,H(t-b(l+x)).\tag{21}$$

After some suitable change of variables and working out the details, it is found that

$$\sigma_{yz}(x,\,0,\,t)=-\frac{\sigma_0}{\pi}\sqrt{\frac{l}{x}}\frac{1}{x+l}H(t-b(x+l)).\tag{22}$$

The remarkable result shown in (22) is that the stress at any point along the crack line takes on its static value instantaneously after the shear wave has passed this point.

The static full-field solutions of stresses for applied point loading σ_0 at the crack faces of the same problem are

$$\sigma_{yz}^s(r,\,\theta)=-\frac{\sigma_0}{\pi}\sqrt{\frac{l}{r}}\frac{r\cos(3\theta/2)+l\cos(\theta/2)}{r^2+2rl\cos\theta+l^2},\tag{23}$$

$$\sigma_{xy}^s(r,\,\theta)=\frac{\sigma_0}{\pi}\sqrt{\frac{l}{r}}\frac{r\sin(3\theta/2)+l\sin(\theta/2)}{r^2+2rl\cos\theta+l^2}.\tag{24}$$

The numerical calculations of the full-field solutions of stress σ_{yz} in (15) are shown in Fig. 2. The most interesting feature as shown in this figure is that the full-field solutions of stresses jump from the dynamic transient solution to the appropriate static value expressed in (23) instantaneously upon arrival of

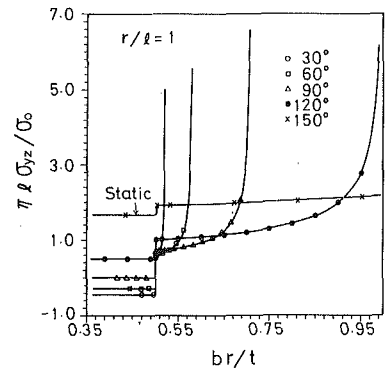


Fig. 2 The transient shear stresses σ_{yz} for applying dynamic point loading on the crack faces

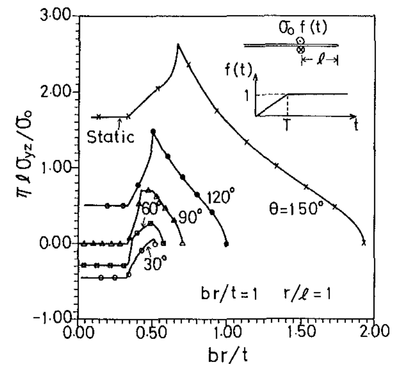


Fig. 3 Nondimensionalized transient shear stress σ_{yz} for applied point loading with a finite rise time

the secondary shear wave (SS) diffracted from the shear wave (S) which is generated by the suddenly applied load. Along the crack-tip line, the SS and S wave coincide. From the general features of the numerical results of full-field solutions indicated above, it is then very easy to draw the conclusion just made regarding the stress along the crack-tip line. Hence, the transient full-field solution of stress σ_{yz} can be expressed as follows:

$$\begin{cases} \sigma_{yz}(r,\theta,t)=0 & \text{for } t<bR, \\ \sigma_{yz}(r,\theta,t)=\sigma_{yz}^{L} & \text{for } t>bR \text{ and } b(r+l)>t, \\ \sigma_{yz}(r,\theta,t)=\sigma_{yz}^{s} & \text{for } b(r+l)<t, \end{cases} \tag{25}$$

where

$$R=[(x+l)^2+z^2]^{1/2}.$$

In order to indicate the correctness of the solutions for point loading shown previously, these solutions for point loading are regarded as the Green function and are used to construct the solutions of uniform loading σ_0 applied to the crack faces. The method used to obtain solutions for stresses is very straightforward. The results are

$$\sigma_{yz}(r,\theta,t)=-\frac{\sigma_0}{\pi}\left[2\cos(\theta/2)\sqrt{\frac{t}{br}-1}-\tan^{-1}\sqrt{\frac{t/br-1}{1-\sin\theta}}-\tan^{-1}\sqrt{\frac{t/br-1}{1+\sin\theta}}\right], \tag{26}$$

$$\sigma_{xy}(r,\theta,t)=\frac{2\sigma_0}{\pi}\sin(\theta/2)\sqrt{\frac{t}{br}-1}. \tag{27}$$

The result of the numerical integration of stress σ_{yz} in (15) over the crack faces is compared with the solution shown in (26). The two results are in excellent agreement as indicated in Table 1.

Table 1 Comparison of the numerical results for σ_{yz} from integration of the solution of point loading and the solutions of applying uniform loading on the crack faces

br/t	Green function	Uniform loading
	$\theta=0^0$	
0.9	$-7.385E-3$	$-7.385E-3$
0.8	$-2.316E-2$	$-2.316E-2$
0.7	$-4.778E-2$	$-4.778E-2$
0.6	$-8.393E-2$	$-8.393E-2$
0.5	$-1.366E-1$	$-1.366E-1$
0.4	$-2.157E-1$	$-2.157E-1$
0.3	$-3.416E-1$	$-3.416E-1$
0.2	$-5.678E-1$	$-5.678E-1$
0.1	$-1.115E-0$	$-1.115E-0$
	$\theta=30^0$	
0.9	$1.982E-2$	$1.982E-2$
0.8	$1.180E-2$	$1.181E-2$
0.7	$-8.609E-3$	$-8.593E-3$
0.6	$-4.217E-2$	$-4.214E-2$
0.5	$-9.302E-2$	$-9.294E-2$
0.4	$-1.700E-1$	$-1.698E-1$
0.3	$-2.927E-1$	$-2.925E-1$
0.2	$-5.132E-1$	$-5.132E-1$
0.1	$-1.046E-0$	$-1.042E-0$
	$\theta=60^0$	
0.9	$1.275E-1$	$1.275E-1$
0.8	$1.348E-1$	$1.348E-1$
0.7	$1.190E-1$	$1.190E-1$
0.6	$8.713E-2$	$8.714E-2$
0.5	$3.807E-2$	$3.806E-2$
0.4	$-3.510E-2$	$-3.510E-2$
0.3	$-1.494E-1$	$-1.494E-1$
0.2	$-3.512E-1$	$-3.512E-1$
0.1	$-8.293E-1$	$-8.294E-1$

Up to this point, the time dependency of the point loading profile is a simple step function in time. The stress for spatially distributed traction on crack faces or for more general time dependence can be obtained by superposition. Suppose that the rate of increase in magnitude of the point loading from zero is taken to be linear, after some finite rise time, say T, the magnitude of the loading is held constant. In this case, the stress σ_{yz} can be obtained from superposition of (15) over time. The numerical results are shown in Fig. 3. The transient solutions will become static at time $t=b(l+r)+T$ as expected. Now, let us consider uniform loading of step function applied on parts of the crack faces from $x=-l_1$ to $x=-(l_1+l_2)$ at $t=0$. The transient solution of σ_{yz} can also be obtained from superposition of (15), the numerical results are shown in Fig. 4. The solutions become a static value after time $t=b(l_1+l_2+r)$ as expected.

4 Dynamic Crack Kinking

In this section, we will analyze the dynamic crack growth

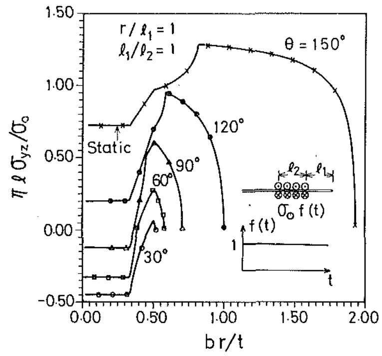


Fig. 4 Nondimensionalized transient shear stress σ_{yz} for suddenly applied uniform loading over part of the crack faces

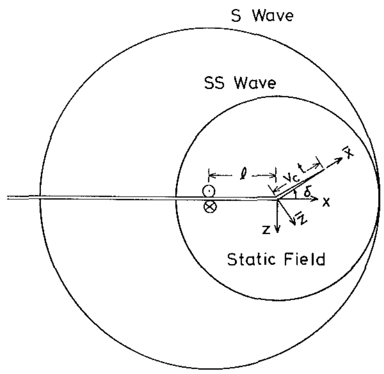


Fig. 5 Geometry of wavefronts for a kinking crack subject to point loading on the crack faces

out of the original semi-infinite crack at an angle to the original crack after applying dynamic point loadings on the original crack faces. From the results indicated in the previous section, we know that the dynamic stress intensity factor jumps from zero to the static value after the S wave arrives at the crack tip. Hence, if the critical stress intensity factor criteria for crack growth is adopted, the crack will be expected to grow immediately after the S wave reaches the crack tip.

Now, let us consider a stationary, semi-infinite crack in an initially stress-free isotropic elastic full space. A sharp crack, which will be referred to as the original crack, is subjected to a point dynamic loading at the crack faces with a distance l to the crack tip. A short time later, at $t=0$, the S wave generated from the point loading arrives at the crack tip. A crack referred to as the new crack, propagates out from the tip of the semi-infinite crack. The velocity of propagation v_c is constant and less than the shear wave speed v_s. The line of propagation is straight, making an angle δ with the original crack, thus producing a kinked crack. The pattern of wavefronts and the position of the crack tip for $t>0$ are shown in Fig. 5. The field solution for a kinked crack geometry can be considered as the superposition of the field for the S and SS waves of the stationary crack and the field from the new crack faces subjected to crack-face tractions which are opposite in sign to the stresses computed from the stationary crack. This computation involves coupled integral equations which must be solved numerically.

The method used to solve the problem in this study relies on an asymptotic approach. The perturbation procedure indicated by Kuo and Achenbach (1985), by using the kinking angle δ as the perturbation parameter, is adopted to construct the solution. The first-order approximation of the dynamic stress intensity factor for a kinked crack can be expressed by the stress intensity factor for a straight crack, propagating in its own plane, subjected to the negative of the traction computed from the stationary crack problem along the line of the kinked crack.

$$\sigma_{yz}^k = 0 \quad \text{for} \quad \bar{x} < 0,$$
$$\sigma_{yz}^k = -\sigma_{\theta y}(\theta=\delta) \quad \text{for} \quad 0 < \bar{x} < v_c t. \tag{28}$$

A fundamental solution needed to construct the problem indicated in (28) is similar to that proposed by Freund (1972a) in the plane-strain case. Consider a crack extending straight at a constant speed v_c in the $\bar{x}$-direction. For $t<0$, there are no body forces or tractions acting on the body. At time $t=0$, the position of the crack tip is $\bar{x}=0$ and concentrated forces of unit magnitude appear at the crack tip. For $t>0$, the crack tip continues to move in the positive $\bar{x}$-direction, but the concentrated forces continue to act at $\bar{x}=0$. The shear stress on the plane $\bar{z}=0$ can be obtained from results analyzed by Ma and Burgers (1986) as follows:

$$\sigma_{yz}^f(\xi,\,0,\,t) = -\frac{\sqrt{d}[\lambda^2 - b^2 - b^2\lambda^2/d^2 - 2b^2\lambda/d]^{1/2}}{\xi\pi(d+\lambda)[b+\lambda(1+b/d)]^{1/2}}\ t > b_2\xi, \tag{29}$$

where

$$\lambda = t/\xi,\ \xi = \bar{x} - v_c t,$$
$$b_2 = b/(1-b/d),\ d = 1/v_c.$$

As indicated in the previous section, stresses for material points behind the SS wave are essentially a static value. With this special feature in mind, the kinked crack problem can be greatly simplified. Hence, if the kinked crack velocity is less than the shear wave speed, the dynamic stress intensity factor for the kinked crack tip subjected to dynamic point loading at the crack faces is the same as that for the static point loading case. For this reason the dynamic stress $\sigma_{\theta y}$ in (28) can be replaced by the equivalent static value $\sigma_{\theta y}^s$ which can be obtained from (23) and (24) as follows:

$$\sigma_{\theta y}^s(r,\,\theta) = -\frac{\sigma_0}{\pi}\sqrt{\frac{l}{r}}\frac{(r+l)\cos(\theta/2)}{r^2+2rl\cos\theta+l^2}.$$

The dynamic problem as indicated in (28) can be solved in a similar manner to that considered above, with the exception that a traction which gives rise to $\sigma_{\theta y}^s(\bar{x})$ on $0<\bar{x}<v_c t$, instead of concentrated forces appears through the moving crack tip. The solution for the case of a distributed traction $\sigma_{\theta y}^s(\bar{x})$ appearing through the crack tip on $0<\bar{x}<v_c t$ is given by the following superposition integral

$$\sigma_{yz}^k(\bar{x},\,0,\,t) = \int_0^{v_c t}\sigma_{yz}^f(\bar{x}-x_0,\,0,\,t-x_0/v_c)\sigma_{\theta y}^s(x_0)dx_0. \tag{30}$$

The first-order approximation of the dynamic stress intensity factor K^d is obtained by considering the limiting behavior $\xi\to 0^+$ of (30) at the moving crack tip

$$K^d(t,\,v_c,\,\delta) = \lim_{\xi\to 0^+}\sqrt{2\pi\xi}\,\sigma_{yz}^k(\bar{x},\,0,\,t). \tag{31}$$

Taking the limit $\xi\to 0^+$ in (29) first and working out the detail of (31) yields

$$K^d(t,\,v_c,\,\delta) = -\frac{\sigma_0}{\pi}\sqrt{\frac{2l}{\pi}}(1-b/d)^{1/2}\int_0^{v_c t}$$
$$\frac{(v_c t + l - \eta)\cos(\delta/2)}{\sqrt{\eta}\sqrt{v_c t - \eta}[(v_c t-\eta)^2 + 2(v_c t - \eta)l\cos\delta + l^2]}d\eta. \tag{32}$$

Equation (32) can be worked out and expressed in explicit form as follows:

$$K^d(t,\,v,\,\delta) = -\sigma_0\frac{\sqrt{1-V}}{\sqrt{\pi l}}$$
$$\left\{\frac{\sqrt{(VT)^2+2VT\cos\delta+1}+VT+\cos\delta}{(VT)^2+2VT\cos\delta+1}\right\}, \tag{33}$$

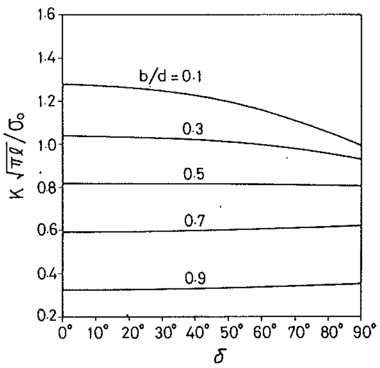


Fig. 6 Dynamic stress intensity factor for various values of the crack kinking angle and crack speed

where $V=b/d$ and $T=t/bl$ are nondimensional quantities. Figure 6 shows the dimensionless dynamic stress intensity factor for various values of the crack kinking angle δ and the normalized crack speed b/d. The error by using this approximation method is less than 10 percent for the any kinking angle less than 90 deg; if the kinking angle is less than 45 deg, the error is less than 2 percent, as indicated by Ma and Burgers (1986) for kinking crack under stress wave loading. We believe that the result in (33) will also be comparably accurate. For the kinking angle $\delta=0$, the crack propagates straight out of the original crack, the solution in (33) is an exact result without any approximation.

$$K^d(t,\ v,\ 0)=-\sigma_0\sqrt{\frac{2}{\pi(vt+l)}}(1-b/d)^{1/2}=K^s\kappa(d). \quad (34)$$

The expression for K^d in (34) has the interesting form of the product of a function of the crack velocity $\kappa(d)$ and the corresponding static stress intensity factor K^s for applying concentrated loading σ_0 at the crack face with a distance $l+vt$ from the crack tip. The value $\kappa(d)=(1-b/d)^{1/2}$ is an universal function which depends only on crack speed and material properties.

At the instant that the kinking has just occurred, we have

$$\lim_{t\to 0^+} K^d(t,\ v,\ \delta)=-\sqrt{\frac{2}{\pi l}}(1-b/d)^{1/2}\cos(\delta/2)$$
$$=K^s\kappa(d)\cos(\delta/2). \quad (35)$$

That is, the stress intensity factor just after the initiation of the kinked crack has the form of the universal function of the crack-tip speed $\kappa(d)$ times the stress intensity factor appropriate for static value K^s times the spatial angular dependence of the stationary crack field.

The energy flux into the propagating crack tip can be written in terms of the corresponding dynamic stress intensity factor by

$$E=\frac{K^2}{2\mu d(1-b^2/d^2)^{1/2}}=\frac{\sigma_0^2}{2\mu lb}E^*, \quad (36)$$

where

$$E^*=\frac{V(1-V)^{1/2}}{\pi(1+V)^{1/2}}\frac{\sqrt{(VT)^2+2VT\cos\delta+1}+VT+\cos\delta}{(VT)^2+2VT\cos\delta+1}.$$

If the maximum energy release rate criterion is accepted as the kinking condition, then the combination of the kinking angle and the crack speed can be determined at which the energy flux into the propagating crack tip achieves a maximum value. The conditions for this to occur are

$$\frac{\partial E^*}{\partial V}=0,\quad \frac{\partial^2 E^*}{\partial V^2}<0,$$

and

$$\frac{\partial E^*}{\partial \delta}=0,\quad \frac{\partial^2 E^*}{\partial \delta^2}<0.$$

If one wants to study the criterion for a crack kinking event, it is clear that the most significant time involved will be when the crack kinking has just occurred, i.e., $T\to 0$. From the maximum energy release rate criterion, it is found that the crack will tend to propagate straight ahead of the original crack with a constant crack speed $v_c=0.618v_s$ which makes $E^*_{\max}=0.191$.

5 Conclusions

The difficulty in determining the transient stress field in a cracked body subjected to dynamic loading is well known. The problem considered in this study is the antiplane response of an elastic solid containing a half-plane crack subjected to impact loading on the crack faces. Attention is focused on the transient stress fields for an applied load with step-function time dependence. The remarkable results are that the full stress takes on its static value a very short time after the *SS* wave diffracted from the crack tip has passed, while the stress intensity factor takes on the appropriate static value after the shear wave generated from the point loading reaches the crack tip. Generalizations are discussed for spatially distributed and time-varying impact loads.

Because of the interesting feature that the static field propagates out behind the *SS* wave front, the evaluation of the dynamic stress intensity factor of the kinking crack propagating with constant crack speed is the same for the dynamic point loading and for the static point loading on the crack faces. A perturbation method is used to obtain the first-order analytic closed-form solution of the dynamic stress intensity factor for the kinking crack. The elastodynamic stress intensity factors of the kinking crack tip are used to compute the corresponding fluxes of energy flux into the propagating crack tip. With these theoretical results for the stress intensity factor at hand, an attempt can be made to determine the kink angle and the new kinked crack speed using different fracture criteria and to compare them with the experimental results available. An energy based fracture criterion is used to look at the initiation of the crack-tip motion. The energy criterion suggests that the crack will choose to propagate in the direction and at the velocity for which the energy flux into the crack tip has a maximum value. Based on the maximum energy release rate criterion, it is found that the crack will tend to propagate straight ahead with crack speed of 0.618 v_s.

Acknowledgment

This investigation was supported by the National Science Council, Republic of China through Grant NSC76-0401-E002-17 at National Taiwan University and is gratefully acknowledged.

References

Achenbach, J. D., and Nuismer, R., 1971, "Fracture Generated by a Dilatational Wave," *Int. J. Frac. Mech.*, Vol. 7, pp. 77–88.

Baker, B. R., 1962, "Dynamic Stresses Created by a Moving Crack," ASME Journal of Applied Mechanics, Vol. 29, pp. 449–458.

Broberg, K. B., 1960, "The Propagation of a Brittle Crack," *Arkiv for Fysik*, Vol. 18, pp. 159–192.

Brock, L. M., 1982, "Shear and Normal Impact Loadings on One Face of a Narrow Slit," *Int. J. Solids Struct.*, Vol. 18, pp. 467–477.

Brock, L. M., 1983a, "The Dynamic Stress Intensity Factor for a Crack Due to Arbitrary Rectilinear Screw Dislocation Motion," *J. of Elasticity*, Vol. 13, pp. 429–439.

Brock, L. M., 1983b, "The Dynamic Stress Intensity Factor Due to Arbitrary Screw Dislocation Motion," ASME Journal of Applied Mechanics, Vol. 50, pp. 383–389.

Brock, L. M., 1984, "Stresses in a Surface Obstacle Undercut Due to Rapid Indentation," *J. of Elasticity*, Vol. 14, pp. 415–424.

Brock, L. M., Jolles, M., and Schroedl, M., 1985, "Dynamic Impact Over a Subsurface Crack: Applications to the Dynamic Tear Test," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 287–290.

Brock, L. M., 1986, "Transient Dynamic Green's Functions for a Cracked Plane," *Q. Appl. Math.*, Vol. 44, pp. 265–275.

de Hoop, A. T., 1958, "Representation Theorems for the Displacement in an Elastic Solid and Their Application to Elastodynamic Diffraction Theory," Ph.D. Dissertation, Technische Hogeschool, Delft.

Eshelby, J. D., 1969, "The Field of a Crack Extending Non-uniformly under General Anti-Plane Loading," *J. Mech. Phys. Solids*, Vol. 17, pp. 177–199.

Freund, L. B., 1972a, "Crack Propagation in an Elastic Solid Subjected to General Loading-I. Constant Rate of Extension," *J. Mech. Phys. Solids*, Vol. 20, pp. 129–140.

Freund, L. B., 1972b, "Crack Propagation in an Elastic Solid Subjected to General Loading-II. Non-uniform Rate of Extension," *J. Mech. Phys. Solids*, Vol. 20, pp. 141–152.

Freund, L. B., 1973, "Crack Propagation in an Elastic Solid Subjected to General Loading-III. Stress Wave Loading," *J. Mech. Phys. Solids*, Vol. 21, pp. 47–61.

Freund, L. B., 1974a, "Crack Propagation in an Elastic Solid Subjected to General Loading-IV. Obliquely Incident Stress Pulse," *J. Mech. Phys. Solids*, Vol. 22, pp. 137–146.

Freund, L. B., 1974b, "The Stress Intensity Factor Due to Normal Impact Loading of the Faces of a Crack," *Int. J. Engng Sci.*, Vol. 12, pp. 179–190.

Kuo, M. K., and Achenbach, J. D., 1985, "Perturbation Method to Analyze the Elastodynamic Field Near a Kinked Crack," *Int. J. Solids Struct.*, Vol. 21, pp. 273–278.

Ma, C. C., and Burgers, P., 1986, "Mode III Crack Kinking with Delay Time: An Analytical Approximation," *Int. J. Solids Struct.*, Vol. 22, pp. 883–899.

Ma, C. C., and Burgers, P., 1987, "Dynamic Mode I and Mode II Crack Kinking Including Delay Time Effects," *Int. J. Solids Struct.*, Vol. 23, pp. 897–918.

Ma, C. C., and Burgers, P., 1988, "Initiation, Propagation, and Kinking of an Antiplane Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 111–119.

Ma, C. C., 1988, "Initiation, Propagation, and Kinking of an In-plane Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 587–595.

Tau, S. A., and Lu, T. H., 1971, "Transient Stress Intensity Factors for a Finite Crack in an Elastic Solid Caused by a Dilatational Wave," *Int. J. Solids Struct.*, Vol. 7, pp. 731–750.

Yoffe, E. H., 1951, "The Moving Griffith Crack," *Phil. Mag.*, Vol. 42, pp. 739–750.

Torsional Stress Waves in a Circular Cylinder With a Modulated Surface

Jin O. Kim
Assoc. Mem. ASME

Haim H. Bau[1]
Assoc. Mem. ASME

Department of Mechanical Engineering
and Applied Mechanics,
University of Pennsylvania,
Philadelphia, PA 19104-6315

The paper describes a theoretical and experimental study pertaining to torsional stress wave motion in an axisymmetric waveguide whose cross-sectional area varies periodically as a function of the axial coordinate. Dispersion relations for the phase speed are obtained for both nonresonant and resonant conditions, using perturbation techniques for small amplitude, sinusoidal modulation. Resonant conditions exist when the modulation wave number is proportional to the sum or difference of wave numbers corresponding to various modes of the torsional stress wave. The experiments consist of measuring the stress wave speed in waveguides with threaded surfaces. The experimental observations verify the general trends predicted by the theory.

1 Introduction

We study theoretically and experimentally transmission of torsional stress waves in a waveguide with a circular cross-section whose radius varies periodically as a function of the axial coordinate. The analysis consists of a perturbation expansion in the amplitude of the modulation (ϵ). The experiments consist of wave speed measurements in unmodulated and modulated waveguides. The experiments confirm the general trends predicted by the theory.

The study was motivated by our efforts to develop a real-time, on line, viscosimeter (Kim et al., 1989; Kim, 1989). In our prior investigation, we studied the effect of an adjacent viscous fluid on the characteristics of torsional stress waves transmitted in submerged waveguides. The speed and attenuation of the stress wave were correlated with the fluid's viscosity and density. We found that one method of increasing the sensor's sensitivity was to increase the waveguide's surface area in contact with the fluid. This can be accomplished by corrugating or modulating the waveguide's surface. In order to achieve a better understanding of the operation of the modulated waveguide, we embarked first on studying the effect of modulation on torsional stress wave transmission in waveguides in vacuum. The results of this study are reported herein. Our results are also applicable for the design of delay lines.

2 Mathematical Model

Consider a torsional stress wave propagating in a cylinder with a circular cross-section (Fig. 1) whose radius ($r = 1 + \epsilon h(z)$) varies periodically (with wave number k_w) as a function of the axial coordinate z, where $h(z) = h(z + 2\pi/k_w)$. All quantities presented here are nondimensional. The average radius of the cross-section is the length scale. The speed of a torsional wave in a straight (smooth) cylinder, $C_0 = (G/\rho)^{1/2}$, is the velocity scale. G and ρ are, respectively, the waveguide's shear modulus and density.

The equations describing the dependence of the circumferential displacement (u) on the radial (r) and axial (z) coordinates, time t and induced-frequency ω are (Love, 1927):

$$Lu = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2} + \frac{\partial^2}{\partial z^2} + \omega^2\right)u = 0, \quad 0 \le r \le 1 + \epsilon h(z) \qquad (1)$$

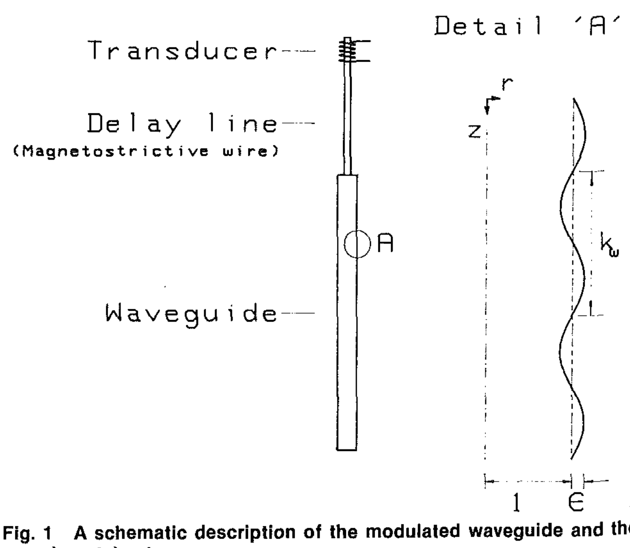


Fig. 1 A schematic description of the modulated waveguide and the experimental setup

[1] To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Professor Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 11, 1989; final revision, Apr. 18, 1990.

$$B(r)u=\left(\frac{\partial}{\partial r}-\frac{1}{r}\right)u=\epsilon\frac{dh}{dz}\frac{\partial u}{\partial z}\quad \text{at } r=1+\epsilon h(z) \tag{2}$$

$$u \text{ is finite at } r=0. \tag{3}$$

The significance of the linear operators L and $B(r)$ is clear from the context. Equation (2) represents the stress-free outer surface condition $\tau_{r\theta}=\tau_{\theta z}\cdot dr/dz$ (Timoshenko and Goodier, 1970). We assume that the waveguide is sufficiently long so as to render end conditions unimportant. The solution presented will be valid away from the waveguide's ends.

3 Perturbation Solution for Nonresonant Condition

Since an exact solution to this problem is unlikely, we resort to approximate, perturbative analysis. To this end, we expand the displacement u into a power series in terms of ϵ:

$$u(r,z,\epsilon)=u_0(r,z)+\epsilon u_1(r,z)+\epsilon^2 u_2(r,z)+O(\epsilon^3). \tag{4}$$

Such an expansion allows us to specify the boundary conditions at $\epsilon=0$ (i.e., Van Dyke, 1975). Next, we substitute (4) into equations (1)–(3), equate coefficients of like powers in ϵ, and obtain an infinite set of differential equations which we solve recursively.

3.1 $O(\epsilon^0)$ Solution. To the leading order $O(\epsilon^0)$, we obtain the classical torsion problem, $Lu_0=0$, with the boundary conditions $B(1)u_0=0$ and u_0 finite at $r=0$. This problem admits the well-known solution (Kolsky, 1953):

$$u_0=U_0F_{0,n}(r)\exp[ik_{0,n}z], \tag{5}$$

where $i=\sqrt{-1}$; U_0 is the wave amplitude which depends on initial conditions and $k_{0,n}$ is the leading order approximation for the torsional wave number of the nth mode, k_n, which we also expand into a power series in terms of ϵ,

$$k_n=k_{0,n}+\epsilon k_{1,n}+\epsilon^2 k_{2,n}+O(\epsilon^3). \tag{6}$$

For the fundamental (nondispersive) mode,

$$F_{0,0}(r)=2r \text{ and } k_{0,0}=\omega, \tag{7}$$

while for the dispersive modes $(n=1,2,3,\cdots)$,

$$F_{0,n}(r)=\sqrt{2}J_1(p_nr)/J_2(p_n) \text{ and } k_{0,n}=(\omega^2-p_n^2)^{1/2}, \tag{8}$$

where J_k is a Bessel function of order k and p_n $(\neq 0)$ is a zero of

$$J_1(p_n)-p_nJ_0(p_n)/2=0, \tag{9}$$

i.e., $p_1=5.136$, $p_2=8.417$, $p_3=11.62$, etc. In the above, the functions $F_{0,n}(r)$ were chosen so as to maintain orthonormality under the inner product

$$<v(r),w(r)>=\int_0^1 rvw\,dr.$$

That is, $\|F_{0,n}(r)\|^2=<F_{0,n}(r),F_{0,n}(r)>=1$ and $<F_{0,n}(r),F_{0,m}(r)>=0$ for $n\neq m$, where $\|\cdot\|$ is the norm.

To obtain higher order corrections, it is convenient to expand $h(z)$ into its Fourier series

$$h(z)=\sum_{j=1}^{\infty}h_j\sin(jk_wz),$$

where h_j are constant coefficients. For conciseness in the derivation below, we take $h_1=1$ and $h_j=0$ for $j>1$. A similar derivation can be carried out for a more general $h(z)$.

3.2 $O(\epsilon^1)$ Solution. To obtain the $O(\epsilon^1)$ solution, we solve the boundary value problem:

$$Lu_1=2k_{0,n}k_{1,n}U_0F_{0,n}(r)\exp[ik_{0,n}z], \tag{10}$$

$$\begin{aligned}B(1)u_1&=U_0\exp[ik_{0,n}z]\{ik_{0,n}k_wF_{0,n}(1)\cos(k_wz)+F''_{0,n}(1)\sin(k_wz)\}\\&=\frac{i}{2}U_0[k_{0,n}k_wF_{0,n}(1)+F''_{0,n}(1)]\exp[i(k_{0,n}+k_w)z]\\&\quad+\frac{i}{2}U_0[k_{0,n}k_wF_{0,n}(1)-F''_{0,n}(1)]\exp[i(k_{0,n}-k_w)z],\end{aligned} \tag{11}$$

$$\text{and } u_1 \text{ is finite at } r=0, \tag{12}$$

where the primes (′) in (11) indicate differentiation with respect to r. In order to solve equations (10)–(12), we decompose u_1 into a sum of three terms,

$$u_1=G_a(r)\exp[ik_{0,n}z]+G_b(r)\exp[i(k_{0,n}+k_w)z]+G_c(r)\exp[i(k_{0,n}-k_w)z], \tag{13}$$

to obtain three boundary value problems (BVPs) for G_x ($x=a$, b, or c) with r being the only independent variable. The homogeneous part of the BVP for G_a

$$L_nG_a=\left(\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr}-\frac{1}{r^2}+\omega^2-k_{0,n}^2\right)G_a=2k_{0,n}k_{1,n}U_0F_{0,n}(r),$$

$$B(1)G_a=0, \text{ and } G_a(0) \text{ finite} \tag{14}$$

is identical to the BVP for u_0. The corresponding solvability condition requires that

$$<L_nG_a(r),F_{0,n}(r)>=B(1)G_a\cdot F_{0,n}(1) \tag{15}$$

or

$$2k_{0,n}k_{1,n}U_0<F_{0,n}(r),F_{0,n}(r)>=F_{0,n}(1)\cdot B(1)G_a=0.$$

Since $\|F_{0,n}(r)\|\neq 0$, we have $k_1=0$. In order to obtain the first correction to the wave number, we will need to proceed to $O(\epsilon^2)$.

Before calculating the $O(\epsilon^2)$ approximation, we shall calculate u_1 explicitly. After some straightforward algebra, we obtain

$$G_a=U_1F_{0,n}(r),\ G_b=i\zeta_{b,n}U_0J_1(\alpha_{b,n}r), \text{ and } G_c=i\zeta_{c,n}U_0J_1(\alpha_{c,n}r), \tag{16}$$

where U_1 is an unknown constant,

$$\alpha_{b,n}^2=\omega^2-(k_{0,n}+k_w)^2,\ \alpha_{c,n}^2=\omega^2-(k_{0,n}+k_w)^2,$$

$$\zeta_{b,n}=\frac{1}{2}[k_{0,n}k_wF_{0,n}(1)+F''_{0,n}(1)]/[\alpha_{b,n}J_0(\alpha_{b,n})-2J_1(\alpha_{b,n})],$$

$$\text{and } \zeta_{c,n}=\frac{1}{2}[k_{0,n}k_wF_{0,n}(1)-F''_{0,n}(1)]/[\alpha_{c,n}J_0(\alpha_{0,n})-2J_1(\alpha_{c,n})].$$

The arguments of the Bessel functions may be imaginary. That is, we may encounter combinations of frequency and wave number such that either $\alpha_{b,n}^2$ and/or $\alpha_{c,n}^2$ are negative. This should not cause any concern as one can replace the Bessel functions (J) with modified ones (I).

Note that the solution u_1 does not exist when the denominator of either $\zeta_{b,n}$, or $\zeta_{c,n}$ vanishes. This "singularity" occurs at specific wave numbers $k_{0,n}=k_r$ such as $\alpha_{b,n}$ or $\alpha_{c,n}$ equals 0 or p_m (where p_m is defined in equation (9)). This is not a physical singularity, but merely an indication that the perturbation expansion (4) is not uniformly valid for $k_{0,n}\approx k_r$. We shall refer to this circumstance as a resonant condition. In order to obtain a uniformly valid solution in the vicinity of resonance, we need to modify the perturbation expansion. We shall discuss this case in detail in Section 4. In the remaining part of this section, we focus on wave numbers away from resonance, i.e., $k_{0,n}\neq k_r$.

3.3 $O(\epsilon^2)$ Solution. The BVP for u_2 has the form:

$$Lu_2=2k_{0,n}k_{2,n}U_0F_{0,n}(r)\exp[ik_{0,n}z], \tag{17}$$

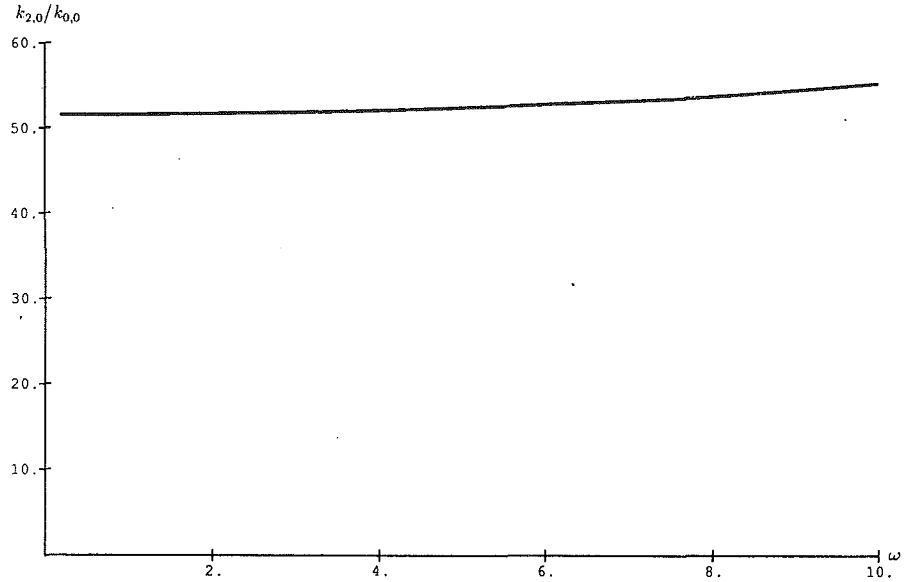


Fig. 2 The fundamental mode phase speed correction $k_{2,n}/k_{0,n}$ depicted as a function of ω for $k_w = 50$

$$B(1)u_2 = \xi_{a,n}U_0\exp[ik_{0,n}z] + i\xi_{b,n}U_1\exp[i(k_{0,n}+k_w)z] + i\xi_{c,n}U_1\exp[i(k_{0,n}-k_w)z] + \xi_{d,n}U_0\exp[i(k_{0,n}+2k_w)z] + \xi_{e,n}U_0\exp[i(k_{0,n}-2k_w)z], \quad (18)$$

where $\alpha_{d,n} = \omega^2 - (k_{0,n}+2k_w)^2$, $\alpha_{e,n} = \omega^2 - (k_{0,n}-2k_w)^2$,

$$\xi_{a,n} = \frac{1}{2}\zeta_{b,n}[-2\alpha_{b,n}J_0(\alpha_{b,n}) - (\alpha_{b,n}^2 - 3)J_1(\alpha_{b,n}) - (k_{0,n}+k_w)k_wJ_1(\alpha_{b,n})] - \frac{1}{2}\zeta_{c,n}[-2\alpha_{c,n}J_0(\alpha_{c,n}) - (\alpha_{c,n}^2 - 3)J_1(\alpha_{c,n}) + (k_{0,n}-k_w)k_wJ_1(\alpha_{c,n})] - \frac{1}{4}[F_{0,n}'''(1) - F_{0,n}''(1)], \quad \xi_{b,n} = \frac{1}{2}F_{0,n}''(1) + \frac{1}{2}k_{0,n}k_wF_{0,n}(1),$$

$$\xi_{c,n} = -\frac{1}{2}F_{0,n}''(1) + \frac{1}{2}k_{0,n}k_wF_{0,n}(1),$$

$$\xi_{d,n} = -\frac{1}{2}\zeta_{b,n}[-2\alpha_{b,n}J_0(\alpha_{b,n}) - (\alpha_{b,n}^2 - 3)J_1(\alpha_{b,n}) + (k_{0,n}+k_w)k_wJ_1(\alpha_{b,n})] + \frac{1}{8}\left[F_{0,n}'''(1) - F_{0,n}''(1)\right] + \frac{1}{4}k_{0,n}k_wF_{0,n}(1),$$

$$\xi_{e,n} = \frac{1}{2}\zeta_{c,n}[-2\alpha_{c,n}J_0(\alpha_{c,n}) - (\alpha_{c,n}^2 - 3)J_1(\alpha_{c,n}) - (k_{0,n}-k_w)k_wJ_1(\alpha_{c,n})] + \frac{1}{8}[F_{0,n}'''(1) - F_{0,n}''(1)] - \frac{1}{4}k_{0,n}k_wF_{0,n}(1),$$

and u_2 is finite at $r=0$. (19)

The form of equations (17)–(18) suggests a solution of the form:

$$u_2 = H_a(r)\exp[ik_{0,n}z] + H_b(r)\exp[i(k_{0,n}+k_w)z] + H_c(r)\exp[i(k_{0,n}-k_w)z] + H_d(r)\exp[i(k_{0,n}+2k_w)z] + H_e(r)\exp[i(k_{0,n}-2k_w)z], \quad (20)$$

where we generate five BVPs for the functions $H_x(x=a, b, c, d$, and $e)$. The correction $k_{2,n}$ to the wave number is obtained from the solvability condition for the BVP for H_a:

$$< L_nH_a(r), F_{0,n}(r) > = B(1)H_a \cdot F_{0,n}(1).$$

That is,

$$2k_{0,n}k_{2,n} < F_{0,n}(r),\ F_{0,n}(r) > = \xi_{a,n}F_{0,n}(1),$$

or $k_{2,n} = \xi_{a,n}F_{0,n}(1)/(2k_{0,n})$.

The functions H_x ($x=a$, b, c, d, and e) were calculated in a straightforward way. Due to space considerations, we do not include the calculations here. The interested reader is referred to Kim (1989). We note in passing, that the $O(\epsilon^2)$ equations also give rise to resonant conditions. The functions $H_x(r)$ in equation (20) contain, in their denominators, expressions of the form $\alpha_xJ_0(\alpha_x) - 2J_1(\alpha_x)$, where $x=b$, c, d, e. A singularity occurs when any of the aforementioned expressions vanishes. The disappearance of the last two expressions ($x=d$ or e) leads to additional, weaker resonant conditions which occur when $k_{0,n} \pm k_{0,m} = 2k_w$. In fact, each order of approximation will give rise to additional, weaker resonant conditions.

3.4 Dispersion Relation. Next, based on the information we have gained thus far, we shall construct an approximation for the phase speed $c_n(=\omega/k_n)$:

$$c_n \approx \frac{\omega}{k_{0,n}}\left(1 - \epsilon^2\frac{k_{2,n}}{k_{0,n}}\right), \quad (21)$$

where c_n is the phase speed of the nth mode. Equation (21) suggests that there are no nondispersive modes in the modulated waveguide and that the modulation leads to a reduction in the phase speed compared to the case of the smooth cylinder. In Fig. 2, we depict the ratio $k_{2,n}/k_{0,n}$ as a function of the frequency ω for the fundamental mode. We also calculated the group velocity but due to the length of the corresponding expressions, we do not reproduce them here. We found the group velocity to be slightly higher than the phase velocity (see Table 1 for a few examples).

One would expect the phase speed to depend on the axial coordinate z. Such a dependence is not reflected in equation (21). In order to obtain such a dependence, we recast our perturbation series (4) into the form:

$$u = U_0F_{0,n}(r)\exp[A(r,z)]\exp[iK(r,z)], \quad (22)$$

where $A(r,z)$, which we do not reproduce here (see Kim (1989)), is a periodic function of z and depends on initial conditions, and

Table 1 Speed corrections for the torsional stress waves along the modulated waveguides. The theoretically predicted results are compared with the experimental observations.

cross-section	thread type	measured speed c (m/s)	c_0 (m/s)	$\frac{c_0-c}{c_0}$ (%)	predicted speed $\frac{k_{2,0}}{k_{0,0}}\times\epsilon^2$ (%) phase	group
solid	NF-UNF 3-56	2696	3015	10.6	13.3	13.2
solid	NC-UNC 3-48	2624	3009	12.8	11.6	11.6
hollow	NF-UNF 3-56	2762	3019	8.5	-	-

$$K(r,z)=(k_{0,n}+\epsilon^2 k_{2,n})z+\epsilon\left[\zeta_b\frac{J_1(\alpha_b r)}{F_{0,n}(r)}+\zeta_c\frac{J_1(\alpha_c r)}{F_{0,n}(r)}\right]\cos k_w z$$
$$+\epsilon^2\left[\left(\zeta_d\frac{J_1(\alpha_d r)}{F_{0,n}(r)}+\frac{\zeta_b^2}{2}\frac{[J_1(\alpha_b r)]^2}{[F_{0,n}(r)]^2}\right)\right.$$
$$\left.-\left(\zeta_e\frac{J_1(\alpha_e r)}{F_{0,n}(r)}+\frac{\zeta_c^2}{2}\frac{[J_1(\alpha_c r)]^2}{[F_{0,n}(r)]^2}\right)\right]\sin 2k_w z$$
$$+O(\epsilon^3).$$

Note that $K(r,z)$ is independent of any constants of integration. We define the effective wave number as $\tilde{k}(r,z)=dK/dz$. Then, the local phase speed $\tilde{c}(r,z)=\omega/\tilde{k}(r,z)$, where $\tilde{c}$ is a periodic function of the axial distance z. The mean of $\tilde{c}$ over the axial distance z is identical to the expression given in (21).

4 Resonant Modes

When the denominators of u_1, $\alpha_{b,n}[J_1(\alpha_{b,n})-2/\alpha_{b,n}J_0(\alpha_{b,n})]\to O(\epsilon)$ or $\alpha_{c,n}[J_1(\alpha_{c,n})-2/\alpha_{c,n}J_0(\alpha_{c,n})]\to O(\epsilon)$, the series ceases to be uniformly valid. We referred to these circumstances as resonant conditions. Similar resonant conditions (albeit weaker resonance) arise at higher-order terms of the perturbation expansion. This unboundedness is not a physical singularity but rather a flaw in the perturbation scheme. Thus, the singularity can be removed by modifying the mathematical procedure.

Resonance of order $O(\epsilon)$ occurs when $k_w=k_{0,n}\mp k_{0,m}$, where $k_{0,l}=(\omega^2-p_l^2)^{1/2}$ ($l=0,1,2,\cdots$ and $p_0=0$). That is, it occurs when the wavelength of the modulation (k_w) is equal to the sum or the difference of two wave numbers. In order to make the perturbation expansion uniformly valid, following Nayfeh (1981), we introduce the detuning parameter σ, so that $k_w=k_{0,n}\mp k_{0,m}+\epsilon\sigma$, and $\sigma\ll\epsilon^{-1}$. Consequently, we have two length scales of $O(1)$and $O(\epsilon)$, respectively. This suggests the use of the method of multiple scales. We replace equation (4) with the expansion,

$$u(r,z,z_1;\epsilon)=u_0(r,z,z_1)+\epsilon u_1(r,z,z_1)+\epsilon^2 u_2(r,z,z_1)+O(\epsilon^3),\quad(23)$$

where $z_1=\epsilon z$. z and z_1 are treated as two independent variables. Next, we introduce (23) into equations (1)–(3), and collect coefficients of like powers in ϵ. For brevity's sake, we shall present the solution technique for $k_w=k_{0,n}-k_{0,m}+\epsilon\sigma$ ($n<m$). The case $k_w=k_{0,n}-k_{0,m}+\epsilon\sigma$ is described in Kim (1989). The zeroth-order problem is identical to the one presented in Section 3. But, now we take u_0 to be the sum of two interacting modes:

$$u_0(r,z,z_1)=A_n(z_1)F_{0,n}(r)\exp[ik_{0,n}z]+A_m(z_1)F_{0,m}(r)\exp[ik_{0,m}z].\quad(24)$$

Since z_1 is a parameter in the $O(\epsilon^0)$ problem, the two integration constants are functions of z_1. The reason for the superposition (24) becomes apparent when we inspect the form of the RHS of equation (25) below.

To $O(\epsilon)$ we have

$$Lu_1=-2ik_{0,n}A_n'(z_1)F_{0,n}(r)\exp[ik_{0,n}z]-2ik_{0,m}A_m'(z_1)F_{0,m}(r)\exp[ik_{0,m}z]\quad(25)$$

$$B(1)u_1=i\zeta_{n_+}A_n(z_1)\exp[i\sigma z_1]\exp[i(2k_{0,n}-k_{0,m})z]$$
$$+i\zeta_{n\_}A_n(z_1)\exp[-i\sigma z_1]\exp[ik_{0,m}z]$$
$$+i\zeta_{m_+}A_m(z_1)\exp[i\sigma z_1]\exp[ik_{0,n}z]$$
$$+i\zeta_{m\_}A_m(z_1)\exp[-i\sigma z_1]$$
$$\times\exp[i(-k_{0,n}+2k_{0,m})z],\quad(26)$$

where $\zeta_{l\pm}=1/2k_w k_{0,l}F_{0,l}(1)\pm 1/2\ F''_{0,l}(1)$ ($l=m,n$). Since the present analysis leads to the $O(\epsilon)$ correction in the wave number expansion (equation (6)), we set, without loss of generality, $k_{1,n}=0$ in equations (25)–(26). (Retention of k_1 leads to identical results.) Equation (26) suggests that we seek a solution in the form:

$$u_1=V_a(r,z_1)\exp[ik_{0,n}z]+V_b(r,z_1)\exp[ik_{0,m}z].\quad(27)$$

Of particular interest are the BVPs associated with $V_a(r,z_1)$ and $V_b(r,z_1)$, as the corresponding BVPs require that certain *solvability conditions* are satisfied. Upon invoking the solvability conditions (which are similar to equation (15)), we obtain two coupled ordinary differential equations for the functions $A_n(z_1)$ and $A_m(z_1)$ of the form:

$$A_n'=iC_1A_m\exp[i\sigma z_1]\text{ and }A_m'=iC_2A_n\exp[-i\sigma z_1],\quad(28)$$

where $C_1=i\zeta_{m+}F_{0,n}(1)/(2k_{0,n})$ and $C_2=i\zeta_{n-}F_{0,m}(1)/(2k_{0,n})$. Primes in (28) denote differentiation with respect to z_1. We proceed to solve equations (28) by introducing a solution of the form $A_l=B_l\exp[i\gamma_l z_1]$($l=m,n$) where B_l are constants. Upon substitution in (28), we obtain:

$$\gamma_nB_n=C_1B_m\exp[i(\sigma+\gamma_m-\gamma_n)z_1],\quad(29)$$

$$\text{and }\gamma_mB_m=C_2B_n\exp[i(-\sigma+\gamma_n-\gamma_m)z_1].\quad(30)$$

For consistency with the assumptions of B_l constants, we need $\sigma+\gamma_m-\gamma_n=0$. Thus, we obtain two, homogeneous, algebraic equations for the coefficients B_l. The solvability condition requires that the determinant of the coefficients vanish. Upon solving the quadratic equation, we obtain

$$\gamma_{n\pm}=1/2(\sigma\pm\sqrt{\sigma^2+4C_1C_2}),$$
$$\text{and }\gamma_{m\pm}=1/2(-\sigma\pm\sqrt{\sigma^2+4(C_1C_2}).\quad(31)$$

In Fig. 3, we depict, for example, $\gamma_{0\pm}$ and $\gamma_{1\pm}$ as functions of k_w for wave frequency $\omega=10$ and $\epsilon=0.2$.

We can now write the first approximation for the deformation field as

$$u(rz)=B_{m-}\left[F_{0,n}(r)+\frac{\gamma_{n-}}{C_1}F_{0,m}(r)\exp(-i\sigma z)\right]$$
$$\times\exp\left[i(k_{0,n}+\epsilon\gamma_{n-})z\right]$$
$$+B_{m+}\left[F_{0,m}(r)+\frac{\gamma_{m+}}{C_2}F_{0,n}(r)\exp(i\sigma z)\right]$$
$$\times\exp[i(k_{0,m}+\epsilon\gamma_{m+})z]+O(\epsilon^1),\quad(32)$$

which represents two interacting, dispersive modes. At resonance ($\sigma=0$), both modes coincide and $\gamma_m=\gamma_n=\pm2\sqrt{C_1C_2}$. As we move away from the resonant conditions, γ_{m+} and γ_{n-} go to zero (Fig. 3) so that the $O(\epsilon)$ correction to the wave speed vanishes. This trend is expected since, as we move away from resonance, we expect the resonant and nonresonant solutions to match asymptotically.

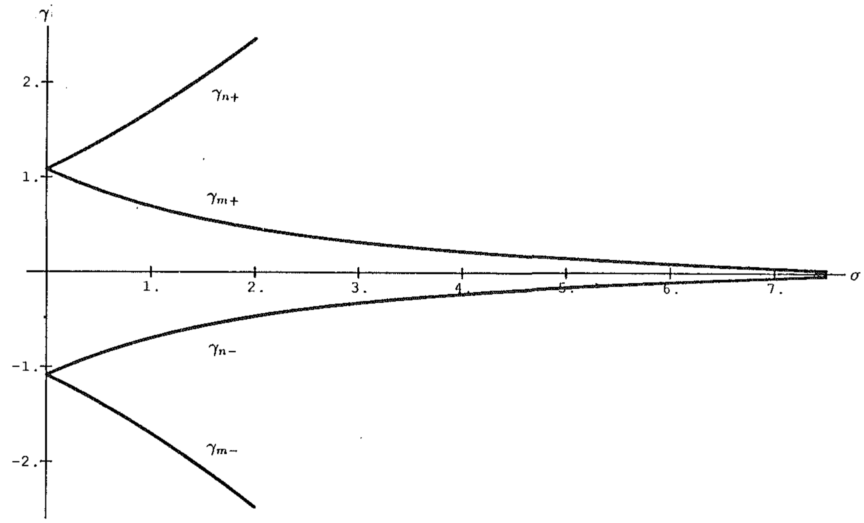


Fig. 3 $\gamma_{n\pm}$ **and** $\gamma_{m\pm}$ **depicted as functions of** σ **for** $\omega = 10$, $\epsilon = 0.2$, $n = 0$, **and** $m = 1$

5 Experiments

In this section, we describe a few experiments we conducted to check the trends predicted by the theory. Due to budget constraints, we were not able to employ waveguides with sinusoidal modulation of the kind described in the theoretical model. Instead, we used threaded waveguides.

In our experiments, we introduced torsional pulses (Fig. 1) in the cylinder (or waveguide) using a torsional-wave sensor, which was designed and manufactured by Lynnworth (1980 and 1989). The pulses were conveniently induced utilizing the magnetostrictive phenomenon. Briefly, one end of a delay line, made of a magnetostrictive material, was soldered or glued to the waveguide while a coil was placed around its other end (Fig. 1). The delay line was electrically polarized so as to develop a circumferential, permanent magnetic field inside the magnetostrictive wire. The introduction of a current pulse in the coil caused a time varying axial magnetic field to develop. The interaction between the two aforementioned magnetic fields led to a twisting force on the magnetostrictive wire and the generation of a torsional pulse. This is known as the Wiedemann effect (Tzannes, 1966). The resulting torsional stress wave traveled in the magnetostrictive wire. Part of the wave was reflected at the magnetostrictive wire-waveguide interface. The other part traveled through the waveguide and was reflected from its other end. The reflected waves caused electromotive forces in the coil which then acted as a receiver (the inverse Wiedemann effect). The signal was viewed on an oscillioscope's screen (see Kim and Bau (1989). By measuring the time which elasped between the two signals, we calculated the speed of the torsional stress wave in a waveguide of known length. In our experiments, the time span was measured peak to peak with a precision of 5 ns. The dominant frequency was estimated as 90 kHz. The delay line was made of Remendur (Fe-Co-V-Mn) of length about 1000 mm. The reflectivity of the delay line-waveguide interface was controlled by controlling the mechanical impedance mismatch at the interface. This is typically done by soldering a small ring around the waveguide. Our experiments suggest insensitivity to eccentric placement of the delay line with respect to the waveguide.

In our experiments, we employed waveguides made of solid and hollow aluminum rods (nominal density $\rho = 2.70 \times 10^3$ kg/m^3, and shear modulus $G = 2.59 \times 10^{10}$ N/m^2). The outer radius of the solid and hollow waveguides $r_o = 1.22$ mm and their length $L = 306$ mm. The inner radius of the hollow waveguide $r_i = 0.78$ mm. The speed of the torsional stress wave in the smooth solid and hollow waveguides submerged in air at 25°C was, respectively, 3014 m/s and 3019 m/s, which agrees within 3 percent with the corresponding speed calculated from nominal material properties $(G/\rho)^{1/2}$.

Subsequently, we threaded the waveguides with standard thread dies; one is NF-UNF 3-56, where $k_w = 15.8$, and another is NC-UNC 3-48, where $k_w = 13.6$. The magnitude of ϵ is commonly 0.087 ± 0.002. The frequency of the wave is 90 kHz, thus the dimensionless frequency $\omega = 0.269$. Since $k_w >> \omega$, resonant effects are not likely to be important. The data (with precision of 0.2 percent) of our experimental observations and the theoretical predictions are documented in Table 1. We observed that the wave speed in the threaded waveguides was 8–12 percent slower than in the corresponding smooth waveguides.

6 Discussion and Conclusion

In this paper, we described an approximate theory which allows us to obtain the dispersion relation for torsional stress waves transmitted in waveguides with modulated surfaces. The theory predicts that the phase speed in the modulated cylinder is smaller than in a smooth one. The analysis predicts the nonresonant and resonant modes will be slowed down, respectively, by a factor of $O(\epsilon^2)$ and $O(\epsilon)$, where ϵ is the normalized amplitude of the modulation.

In parallel, we also conducted a few experiments to measure the effect of modulation on wave speed. The experiments were conducted on a somewhat different geometry than the one employed in the analysis. Thus, we can discuss only qualitative verification of the theoretical results. The experiments verified that the phase speed in the modulated waveguide was lower than the one in the smooth cylinder. The reduction observed in the experiments was of $O(\epsilon^2)$, which corresponds to the theoretical predictions for the nonresonant modes.

Acknowledgment

This work has been supported by the National Science Foundation (Grant CBT83-51658), Panametrics, Inc., and the Electric Power Research Institute (RP-2515-03). We gratefully acknowledge useful discussions with L. C. Lynnworth of Panametrics, Inc.

References

Kim, J. O., 1989, "The Interaction Between Stress Waves Transmitted in Solid Waveguides and Adjacent Media," Ph.D. Dissertation, University of Pennsylvania.

Kim, J. O., and Bau, H. H., 1989, "On Line, Real-Time Densimeter—Theory and Optimization," *J. Acoustical Society of America,* Vol. 85, No. 1, pp. 432-439.

Kim, J. O., Wang, Y., and Bau, H. H., 1991, "The Effect of an Adjacent Viscous Fluid on the Transmission of Torsional Stress Waves in a Submerged Waveguide," *J. Acoustical Society of America,* Vol. 89, pp. 1414-1422.

Kolsky, H., 1953, *Stress Waves in Solids,* Clarendon Press, Oxford, pp. 65-68.

Love, A. E. H., 1927, *A Treatise of the Mathematical Theory of Elasticity,* 4th ed., Cambridge University Press, Oxford, pp. 287-289.

Lynnworth, L. C., 1980, "Slow Torsional Wave Densimeter," United States Patent 4,193,291.

Lynnworth, L. C., 1989, *Ultrasonic Measurement for Process Control,* Academic Press, New York.

Nayfeh, A. H., 1981, *Introduction to Perturbation Techniques,* John Wiley and Sons, New York, pp. 418-426.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity,* 3rd ed., McGraw-Hill, New York, pp. 341-343.

Tzannes, N. S., 1966, "Joule and Wiedemann Effects—The Simultaneous Generation of Longitudinal and Torsional Stress Pulses in Magnetostrictive Materials," *IEEE Trans. on Sonics and Ultrasonics,* Vol. SU-13, No. 2, pp. 33-41.

Van Dyke, M., 1975, *Perturbation Methods in Fluid Mechanics,* The Parabolic Press, Stanford, Calif., pp. 36-37.

H. T. Banks[1]
Center for Control Sciences,
Division of Applied Mathematics,
Brown University,
Providence, RI 02912

D. J. Inman
Mechanical Systems Laboratory,
Department of Mechanical and
Aerospace Engineering,
State University at New York at Buffalo,
Buffalo, NY 14260
Fellow ASME

On Damping Mechanisms in Beams

A partial differential equation model of a cantilevered beam with a tip mass at its free end is used to study damping in a composite. Four separate damping mechanisms consisting of air damping, strain rate damping, spatial hysteresis, and time hysteresis are considered experimentally. Dynamic tests were performed to produce time histories. The time history data is then used along with an approximate model to form a sequence of least squares problems. The solution of the least squares problem yields the estimated damping coefficients. The resulting experimentally determined analytical model is compared with the time histories via numerical simulation of the dynamic response. The procedure suggested here is compared with a standard modal damping ratio model commonly used in experimental modal analysis.

I Introduction

In this paper a variety of damping mechanisms for a quasi-isotropic pultruded composite beam are examined. The approach taken here is a physical one. The beam is modeled by a partial differential equation describing the transverse vibration of a beam with tip mass. The damping mechanisms considered are all physically based as opposed to the usual modal model. In total, four possible damping mechanisms are considered, one external and three internal. They are: viscous damping (air damping); strain rate damping; spatial hysteresis; and time hysteresis. In addition, various combinations of these mechanisms are considered.

These physical damping models are incorporated into the Euler-Bernoulli beam equation, with care taken to formulate boundary conditions that are compatible with the various damping models. The resulting partial differential equation (integro partial differential equation in the case of time hysteresis damping) is approximated using cubic splines. The time histories of the measured experimental responses are then used to form a least squares fit-to-data parameter estimation problem. The mathematical arguments underlying this procedure are complete and imply convergence of a sequence of parameter estimates obtained from finite dimensional models to a set of best fit coefficients of the partial differential equation model. The least squares estimates of the various different damping parameters are then used in the partial (integro partial) differential equation to numerically simulate the response of the system. This numerically generated time response of the estimated system is then compared with the actual experimental time histories. These comparisons allow several conclusions to be drawn regarding the physical damping mechanisms present in the composite beam.

In particular, it is shown that the spatial hystereis model combined with a viscous air damping model results in the best quantitative agreement with the experimental time histories. The results also support the physically intuitive notion that air damping should play a more significant role in lower modes, while internal damping plays a more significant role for higher modes. It is also shown explicitly that the proposed damping models listed above cannot be modeled with any degree of success or consistency by using standard modal damping ratios, as the traditional modal analysis approach completely masks the physics of damping mechanisms.

II Basic Beam Model

The particular beam considered here is a pultruded quasi-isotropic composite beam constructed for use in the proposed space station (Wilson and Miserentino, 1986). As such, the configuration of interest is a cantilevered beam with a mass attached to the free end. The beam is constructed of a biaxial (0 deg/90 deg) fiberglass roving held in place with knitted polyester yarn with an equal volume of fibre in both orientations. An isophtalic polyester resin system was used as the matrix. This material provides an alternative to aluminum which is lower in cost, has higher specific strength, but is dynamically similar. As is illustrated here, this material also has interesting damping properties, dissimilar to those of aluminum.

The equation of motion for the flexural vibration of a beam is easily calculated from consideration of the equilibrium of forces acting on a differential segment of beam (see, for instance, Clough and Penzien, 1975). In this formulation, damping can easily be included by adding the appropriate force or moment to the equations of equilibrium. A partial differential equation model of the beam with general damping is of the form

[1] Current address: Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, CA 90089.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 24, 1989; final revision, Mar. 14, 1990.

$$u_{tt}(x,t)+L_1u_t(x,t)+L_2u(x,t)+\frac{\partial^2}{\partial x^2}\left[\frac{EI(x)}{\rho}u_{xx}(x,t)\right]=f(x,t) \quad (1)$$

for $x\epsilon(0,l), t>0$, subject to the appropriate boundary conditions and the initial conditions (taken to be $u=u_t=0$ at $t=0$). Here, ρ is the linear mass density (mass per unit length) of the beam, $EI(x)$ is the spatial varying flexural stiffness of the beam, the subscript indicates partial differentiation with respect to the indicated variable and $u(x,t)$ is the beam displacement in the transverse direction at location x, time t. The function $u(x,t)$ is assumed to be smooth enough so that all the appropriate derivatives exist. The term $L_1u_t(x,t) + L_2u(x,t)$ is the focus of attention in this paper. The nature of the operator L_1 is usually determined by external damping mechanisms (although internal damping may contribute to this term—see the example involving spatial hysteresis modeled in (13)–(15) below) while the nature of term $L_2u(x,t)$ is most often determined by internal damping mechanisms.

The boundary conditions of interest here are those for a beam clamped at the end $x=0$ and with a free end at $x=l$. In addition a mass of mass, m_T, and rotational inertia, J, is attached at $x=l$. The fixed end requires that the displacement and the slope of the displacement both be zero. This yields:

$$u(0,t)=0 \quad (2)$$

$$u_x(0,t)=0. \quad (3)$$

The free end requires that the sum of the moments at $x=l$ and the sum of the forces acting at $x=l$ must each be zero. For the case of a tip mass at the free end, these boundary conditions become

$$EI(l)u_{xx}(l,t)=-Ju_{ttx}(l,t) \quad (4)$$

$$[EI(x)u_{xx}(x,t)]_x=m_Tu_{tt}(x,t),\ x=l \quad (5)$$

as long as only external damping is present.

Equations (1)–(5) describe the transverse vibration of a beam satisfying the Bernoulli-Euler assumption that the bending wave length is several times larger than the cross-sectional dimensions of the beam, and that only lower frequency excitations are applied to the beam. It is assumed that rotary inertia of the beam, shear displacement of the beam and axial displacements are negligible.

If the tip mass is not present, the boundary conditions of equations (4) and (5) change accordingly. In addition, the nature of the internal damping operator will effect the boundary conditions. For the case of $L_1=L_2=0$, the vibration analysis problem is very simple as is the inverse problem addressed here. The nature of the damping mechanisms drastically changes the nature of the solution to the vibration problem and hence controls the response of the beam. The following section discusses several possible choices for modeling the operators L_1 and L_2 in equation (1) and hence the internal damping mechanism in the beam.

III Damping Models

As mentioned in the Introduction, four models for the damping mechanisms are examined. Two of these are time-independent proportional models lending themselves to modal expansions, the other two are nonproportional hysteretic models. Various combinations of these models are also considered.

Viscous Air Damping. The most straightforward method of modeling the damping of a beam (or other object) vibrating in air is to use a viscous model with damping force assumed proportional to velocity. In this case the operator L_1 becomes

$$L_1=\gamma I_0 \quad (6)$$

where I_0 is the identity operator and γ is the viscous damping constant of proportionality. The physical basis of this approach is a simple model of air resistance. As the beam vibrates it must displace air causing the force $\gamma u_t(x,t)$ to be applied to the beam. Mathematically, this form of damping is used because it is proportional and easily treated using the same methods of analysis used for undamped systems (see Inman, 1989, for instance). This form of damping is often called viscous external damping.

Kelvin-Voigt Damping. Kelvin-Voigt damping, or *strain-rate damping* as it is sometimes called, is damping of the form

$$L_2=c_dI\frac{\partial^5}{\partial x^4\partial t} \quad (7)$$

where I is the moment of inertia and c_d is the strain-rate damping coefficient. This model also satisfies a proportional damping criteria and hence is mathematically convenient. (We note that this form of damping could also be written in terms of L_1 as $L_1=c_dI\,(\partial^4)/(\partial x^4)$. It is compatible with theoretical modal analysis and is also, along with viscous damping, widely used in finite element modeling. This form of damping is most often referred to as internal damping and represents energy dissipated by friction internal to the beam.

Unlike viscous external damping, inclusion of this form of damping affects the free-end boundary conditions because it is strain dependent. The strain-rate dependence results in a damping moment M_D of the form

$$M_D=c_dI(x)\frac{\partial^3u}{\partial x^2\partial t}, \quad (8)$$

which is included in the derivation of the equation of motion (Clough and Penzien, 1975) and, hence, must also be included in any boundary conditions (such as a free-end condition) depending on the moment.

The full equation of motion and boundary conditions for the case including linear viscous external damping and Kelvin-Voigt internal damping can thus be written

$$\rho u_{tt}(x,t)+\frac{\partial^2}{\partial x^2}[(EIu_{xx}(x,t)+c_dIu_{xxt}(x,t)]+\gamma u_t(x,t)=f(x,t),\ x\epsilon(0,l),\ t>0$$

$$u(0,t)=u_x(0,t)=0,\ t>0$$

$$EIu_{xx}(l,t)+c_dIu_{xxt}(l,t)=-Ju_{xtt}(l,t),\ t>0$$

$$\frac{\partial}{\partial x}[EIu_{xx}(x,t)+c_dIu_{xxt}(x,t)]=m_Tu_{tt}(x,t),\ x=l,\ t>0. \quad (9)$$

Here, note that the tip mass as well as the internal damping moment are represented in the boundary conditions. The total damping mechanism used in (9) is the analog to proportional damping (i.e., a linear combination of mass and stiffness operators).

Time Hysteresis. Hysteretic damping terms are most commonly associated with sinusoidal loadings. The generic idea of including a mechanism in the beam vibration constitutive equation indicating that stress is proportional to strain plus the past history of the strain can be accomplished by introducing an integral term of the form

$$\int_{-r}^{0}g(s)u_{xx}(x,t+s)ds \quad (10)$$

with the history kernel $g(s)$ defined by

$$g(s)=\frac{\alpha}{\sqrt{-s}}\exp(\beta s) \quad (11)$$

where α and β are constants. Since the introduction of the hereditary integral occurs in the stress strain relationship, the boundary conditions must also be modified. In this case the boundary value problem of interest becomes

$$\rho u_{tt}(x,t)+\frac{\partial^2}{\partial x^2}\left[EIu_{xx}(x,t)-\int_{-r}^{0}g(s)u_{xx}(x,t+s)ds\right]$$
$$=f(x,t),\quad x\in(0,l),\ t>0$$
$$u(0,t)=u_x(0,t)=0\ t>0$$
$$EIu_{xx}(l,t)-\int_{-r}^{0}g(s)u_{xx}(l,t+s)ds=Ju_{xtt}(l,t),\ t>0$$
$$\frac{\partial}{\partial x}\left[EIu_{xx}(x,t)-\int_{-r}^{0}g(s)u_{xx}(x,t+s)ds\right]$$
$$=m_T u_{tt}(x,t),\ x=l,t>0. \quad (12)$$

It is emphasized again that the inclusion of an internal damping moment in the equation of motion also affects the boundary condition.

Spatial Hysteresis. Another type of damping (proposed by Russell, 1991) is based on interpreting the energy lost in the transverse vibration of a beam as resulting from differential rates of rotation of neighboring beam sections causing internal friction. This can be modeled by the damping expression

$$\frac{\partial}{\partial x}\left[\int_{0}^{l}h(x,\xi)\{u_{xt}(x,t)-u_{xt}(\xi,t)\}d\xi\right] \quad (13)$$

where the kernel $h(x,\xi)$ may be defined, for example, by

$$h(x,\xi)=\frac{a}{b\sqrt{2\pi}}\exp[-(x-\xi)^2/2b^2]. \quad (14)$$

Under these circumstances the model for the beam vibration becomes (including viscous external damping)

$$\rho u_{tt}(x,t)+\frac{\partial^2}{\partial x^2}[(EIu_{xx}(x,t)]+\gamma u_t(x,t)$$
$$-\frac{\partial}{\partial x}\left[\int_{0}^{l}h(x,\xi)\{u_{xt}(x,t)-u_{xt}(\xi,t)\}d\xi\right]$$
$$=f(x,t),\ x\in(0,l),\ t>0$$
$$u(0,t)=u_x(0,t)=0,\ t>0$$
$$EIu_{xx}(l,t)=-Ju_{xtt}(l,t)\ t>0$$
$$\frac{\partial}{\partial x}\{(EIu_{xx}(x,t)]-\int_{0}^{l}h(x,\xi)\{u_{xt}(x,t)-u_{xt}(\xi,t)\}dx$$
$$=m_T u_{tt}(x,t),\ t>0,\ x=l \quad (15)$$

where, again, the internal damping mechanism is reflected in the boundary conditions.

In total, the models described by systems (9), (12), and (15) represent four possible mechanisms of damping taken in various combinations. The approach taken here is to attempt to fit each of the combinations of damping models listed above to experimentally measured data. By examining each model's numerical solution in comparison with measured data, a best model is chosen from these as being most representative of the cantilevered quasi-isotropic beam. As is discussed in Section V, these models all admit reasonable mathematical treatment.

IV Problem Statement

The various damping coefficients introduced in the preceding discussion cannot be measured by static experiments. Thus, the damping constants γ, c_d, α, β, a, and b must all be estimated based on measurements taken from dynamic experiments. The procedure suggested here is to estimate various groups of damping parameters such as indicated in the three models of (9), (12), and (15). Once these coefficients are estimated they can be used in the model to produce a numerical simulation of the response of the structure under consideration subjected to identical experimental inputs. The analytical time response (with the estimated coefficients) is then compared to the experimentally measured time response. The model that best agrees with (predicts) the experimental response is then considered to be a valid and desirable physical model.

In particular, several vectors of parameters $\mathbf{q}$ are defined, one for each model of interest. For the three cases discussed here they are:

$$\mathbf{q}_1=[EI,c_dI,\gamma] \quad (16)$$

which delineates the first damping model as defined by system (9). Here c_d is the internal strain rate damping coefficient and γ is the linear air damping coefficient. The second model considered, as defined by system (12), is characterized by the parameter vector

$$\mathbf{q}_2=[EI,\alpha,\beta] \quad (17)$$

where α and β characterize the time hysteretic damping term. The last model considered contains a combination of linear air damping, defined by the coefficient γ, and spatial hysteresis defined by the constants a and b. The parameter vector for the third system defined by (15) is

$$\mathbf{q}_3=[EI,\gamma,a,b]. \quad (18)$$

Other combinations of the four damping mechanisms were considered but were dismissed as discussed in the later section on results. Even though the techniques (Spline Inverse Procedures) described below can readily be used to treat spatially varying coefficients EI and c_dI, the experiments for the effort presented in this paper were performed on uniform beams. Thus, consideration in this paper will henceforth be restricted to constant EI and c_dI.

Note that in each case the parameter vector contains the flexural stiffness constant EI. For most common materials EI is tested, tabulated, and well known. However, in this case the material is a prototype composite with unknown material properties. Thus, EI is also estimated. Because of the relative size of the strain rate damping coefficient c_d, the term c_dI is estimated.

V Mathematical Foundation of the Estimation Problem

Two approaches to solving the problem of determining the coefficients in the vectors $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$ are formulated here. The first approach involves application of experimental modal analysis methods (see Ewins, 1988 or Inman, 1989) to a theoretical modal analysis of equation (1). This procedure, suggested by Clough and Penzien (1975), can only be applied to the problem of estimating $\mathbf{q}_1$ because modal analysis is not applicable to the hysteresis terms in $\mathbf{q}_2$ and $\mathbf{q}_3$. This modal approach is presented for comparison and because it represents a standard approach for measuring damping. However, modal approaches cannot be used to solve the inverse problem for general constitutive elements. The inverse procedures suggested as the second approach here do not have this limitation and can be applied to systems with spatially dependent physical parameters (EI, etc.) as well as exotic damping mechanisms.

The second approach taken here is a nonmodal procedure applicable to all three estimation problems, and forms our proposed method. This method is based on a careful consideration of the distributed parameter nature of the test article. It consists of forming a sequence of finite dimensional approximations (Galerkin-type approximation with cubic B-splines) to equation (1) with an associated least squares fit-to-data (see Banks and Kunisch, 1989, for a general discussion of these ideas.) For each of the damping models presented in this paper, a corresponding sequence of approximate estimates

$\mathbf{q}_i^N$ can be shown to converge to a best-fit parameter $\mathbf{q}_i$ for the original distributed parameter system (1) (or specifically for (9), (12), or (15)). Detailed convergence arguments are given in Banks et al. (1983, 1986, 1988, 1989), Banks and Ito (1988) for a variety of continuum models. The computational algorithms proposed here are based on these considerations of the distributed parameter nature of the estimation problem and is called the Spline-based Inverse Procedure (SIP).

Modal Analysis (EMA). The typical experimental approach to measuring the damping in a structure is to use experimental modal analysis (EMA) to determine modal damping ratios and natural frequencies. These quantities can then be used to determine the physical parameters contained in the vector $\mathbf{q}_1$. In the case discussed here, the tip mass is removed (for ease in exposition) so that $m_T = J = 0$, and the unit interval is used (i.e., $l = 1$). The damping operators L_1 and L_2 become

$$L_1 = \frac{\gamma}{\rho},\ L_2 = \frac{c_d I}{\rho}\frac{\partial^5}{\partial x^4 \partial t} \tag{19}$$

which commute with the stiffness operator so that a modal representation is possible (Caughey and O'Kelley, 1965). According to this theory (see, for instance, Inman, 1989), the solution of (9) can be written as a series of products of two functions, $u_m(x,t) = a_m(t)\ \phi_m(x)$, which satisfy (9) for each m and whose sequence of partial sums converges to the unique solution of (9). Here the normalized functions $\phi_m(x)$ are the eigenfunctions (mode shapes) of the stiffness operator

$$\frac{\partial^2}{\partial x^2}\left(\frac{EI}{\rho}\frac{\partial^2}{\partial x^2}\right) \tag{20}$$

subject to the boundary conditions of system (9). These eigenfunctions satisfy (for EI constant)

$$\frac{\partial^4}{\partial x^4}\phi_m(x) = \beta_m^4 \phi_m(x) \tag{21}$$

where $\beta_m^4 = \omega_m^4 (\rho/EI)^2$ and the ϕ_m satisfy the orthogonality condition

$$\int_0^1 \phi_n(x)\phi_m(x) = \delta_{nm}. \tag{22}$$

Here, δ_{nm} is the Kronecker delta and ω_m are the undamped natural frequencies of the system. Substituting $u_m(x,t)$ into equation (1), multiplying by $\phi_m(x)$, and integrating with respect to x over the interval (0,1), one finds that each $a_m(t)$ must satisfy

$$\ddot{a}_m(t) + \left(\frac{\gamma}{\rho} + \frac{c_d I}{\rho}\beta_m^4\right)\dot{a}_m(t) + \frac{EI}{\rho}\beta_m^4 a_m(t) = \int_0^1 f(x,t)\phi_m(x)dx, t>0 \tag{23}$$

for all $m = 1,2,3,\ldots$. Equation (23) has a direct relationship to the frequency domain measured modal data available from EMA. The experimental modal analysis procedure assumes that the structure consist of some finite number of single degree-of-freedom oscillators of the form:

$$\ddot{a}_m(t) + 2\hat{\zeta}_m\hat{\omega}_m\dot{a}_m(t) + \hat{\omega}_m^2 a_m(t) = f_m \tag{24}$$

where $\hat{\zeta}_m$ and $\hat{\omega}_m$ are measured damping ratios and natural frequencies, respectively. Comparing coefficients of a_m and $\dot{a}_m$ in equations (23) and (24), one obtains

$$\beta_m^4 \frac{EI}{\rho} = \hat{\omega}_m^2 \tag{25}$$

and

$$\frac{\gamma}{\rho} + \frac{c_d I}{\rho}\beta_m^4 = 2\hat{\zeta}_m\hat{\omega}_m \tag{26}$$

for each m.

As outlined in Cudney and Inman (1989), the elastic modulus E may be estimated from equation (25) by a linear least squares fit by which one obtains

$$E = \frac{1}{K}\sum_{m=1}^{K}\frac{4\pi^2\rho}{I\beta_m^4}\hat{f}_m^2. \tag{27}$$

where $\hat{f}_m = \hat{\omega}_m/2\pi$ hertz with $(\hat{\zeta}_m, \hat{\omega}_m)$, $m = 1,2,\ldots K$, a given measured set. Equation (26) can be written down once for each pair in this measured set $(\hat{\zeta}_m, \hat{\omega}_m)$, $m = 1,2,\ldots K$, to produce a least squares determination of the damping parameters of the form

$$\begin{bmatrix} \gamma \\ c_d \end{bmatrix} = B^\dagger \mathbf{z}. \tag{28}$$

Here, $B\dagger$ denotes the generalized inverse (least square) of the $2 \times K$ matrix

$$B = \frac{1}{\rho}\begin{bmatrix} 1 & \beta^4{}_1 I \\ 1 & \beta^4{}_2 I \\ \vdots & \vdots \\ 1 & \beta_K^4 I \end{bmatrix} \tag{29}$$

and $\mathbf{z}$ is the $K \times 1$ vector $\mathbf{z} = [2\hat{\zeta}_1\hat{\omega}_1,\ 2\hat{\zeta}_2\hat{\omega}_2,\ \ldots,\ 2\hat{\zeta}_K\hat{\omega}_K]^T$ of measured modal information. The entries in B are calculated from the analytical solution of the eigenvalue problem for the stiffness operator above with appropriate boundary conditions.

Equation (28) (and/or a weighted version of it) can be used to estimate the distributed damping parameters for the problem involving $\mathbf{q}_1$. While intuitively obvious and straightforward, this modal-based method requires that EI must be constant and that β_m must be known in closed form. In addition, this approach is not applicable to the noncommuting damping models involving $\mathbf{q}_2$ and $\mathbf{q}_3$ or to problems with spatially varying coefficients. This provides the motivation for the nonmodal approach developed next.

Spline Inverse Procedure (SIP). An alternative to estimating $\mathbf{q}_i$ from measured modal data (frequency domain) is to formulate a parameter estimation problem based on measured time histories of the test structure's response. Let $\hat{u}_{tt}(l,t_i)$ denote the acceleration measurements at the tip of the beam $(x = l)$ at various times t_i. The inverse problem of interest is then to find the vector of parameters $\mathbf{q}$ such that

$$J(\mathbf{q}) = \sum_{i=1}^{m} |u_{tt}(l,t_i,\mathbf{q}) - \hat{u}_{tt}(l,t_i)|^2 \tag{30}$$

is minimized where $u(x,t,\mathbf{q})$ denotes the solution of equation (1) with the appropriate boundary and initial conditions corresponding to parameter values $\mathbf{q}$. Here, m is the number of tip acceleration measurements.

This estimation problem cannot, of course, be solved analytically. However, an iterative optimization scheme coupled with an approximation method for the infinite dimensional system of equation (1) may be used. The procedure suggested here is outlined as follows. First, equation (1) is approximated via Galerkin procedures using cubic spline elements (N is used to denote the approximation index) to yield an approximate finite dimensional version which is solved for $u^N(t)$. The approximate accelerations u_{tt}^N are then used in the cost function of equation (30) to define the finite dimensional estimation problem of minimizing

$$J^N(\mathbf{q}) = \sum_{i=1}^{m} |u_{tt}^N(l,t_i,\mathbf{q}) - \hat{u}_{tt}(l,t_i)|^2. \tag{31}$$

The solution of this set of estimation problems yields a sequence of estimates, denoted $\{\mathbf{q}^N\}$, of best-fit parameters. Under appropriate assumptions, this sequence is then shown

Table 1 Theoretical eigenvalues and experimentally measured modal data for a 1-m long clamped free beam

Moment of inertia = $1.64 \times 10^{-9} m^4$, linear mass density = 1.02 kg/m					
Mode No.	Eigenvalue β_i	Theoretical Freq. (Hz.)	Experimental Freq. (Hz.)	Damping Ratio (%)	Std. Dev.
1	1.875	-	-	-	-
2	4.694	23.096	22.8	.218	.015
3	7.855	64.675	65.3	.227	.016
4	10.996	126.740	127	.154	.003
5	14.137	209.488	212	.228	.023
6	17.279	312.955	314	.120	.008
7	20.420	437.075	435	.131	.021
8	23.562	581.475	580	.155	.010
9	26.704	747.475	733	.202	.015

Table 2 Estimates of damping based on using a successive number of modes (all values in kg/msec)

Modes	Viscous Only γ	Strain Only $c_d \times 10^6$	Viscous and Strain γ	Viscous and Strain $c_d \times 10^6$
1-2	.3619	.8024	.0724	.7092
1-3	.8755	.3127	.2014	.2699
1-4	1.2829	.1179	.6053	.0873
1-5	2.2693	.0978	.6157	.0856
1-6	2.6962	.0451	1.3901	.0323
1-7	3.3565	.0304	1.6867	.0221
1-8	4.3778	.0251	1.8039	.0199
1-9	6.0027	.0236	1.7561	.0205

to converge to $\mathbf{q}^*$, a vector of parameters for the fully distributed parameter model of equation (1) which minimizes $J(\mathbf{q})$ of equation (30). The theoretical formulation of this approach is presented next. The required proofs are omitted but can be found in Banks and Ito (1988), Banks et al. (1983, 1986, 1989), and Banks and Kunisch (1989).

The SIP estimation algorithm is formulated in weak or variational form by multiplying equation (1) by $\psi(x)$ and integrating over the interval $(0,l)$. This yields

$$<u_{tt}, \psi> + <L_1 u_t, \psi> + <L_2 u, \psi> + <D^2\left(\frac{EI}{\rho}D^2 u\right), \psi> = <f, \psi> \quad (32)$$

where $D = \partial/\partial x$ and the inner product $<\cdot,\cdot>$ for the Hilbert space $H = L_2\ (0,l)$ is defined by

$$<\phi_1, \phi_2> = \int_0^l \phi_1(x)\phi_2(x)dx. \quad (33)$$

Equation (32) must hold in a generalized sense (this involves integration by parts in certain terms) for all ψ in V, a Hilbert space continuously and densely imbedded in the Hilbert space H, containing the solution of (1) subject to the appropriate boundary conditions. (In this case $V = \{\psi \in H^2(0,l): \psi(0) = \psi'(0) = 0\}$ where H^2 is the Sobolev space of functions possessing first and second derivatives in $L_2(0,l)$). The terms in equation (32) are now identified with elements $\ddot{u}(t;\mathbf{q})$, $\dot{u}(t;\mathbf{q})$ and $u(t;\mathbf{q})$ which evolve in time, and satisfy (in a generalized or weak sense—see Banks and Ito (1988) and Banks and Kunisch (1989) the evolution equation

$$\ddot{u}(t;\mathbf{q}) + \mathcal{B}(\mathbf{q})\dot{u}(t;\mathbf{q}) + \mathcal{A}(\mathbf{q})\ u(t;\mathbf{q}) = f(t),\ t>0 \quad (34)$$

subject to the appropriate initial conditions. Here, the explicit dependence on the parameter vector $\mathbf{q}$ is emphasized, while the solution of equation (34) is a function of time for each x (or, alternatively, is thought of as a function of x for each t, the function being an element of V for each t). The operators $\mathcal{A}$ and $\mathcal{B}$ can be appropriately defined using the corresponding terms in (32).

The Galerkin approach employing cubic spline subspaces to solve (34) (or, equivalently (32)) approximately is explained next. Given a value of N and a vector $\mathbf{q}$, an approximate solution to (34) in $X^N = \text{span}\ \{B_1^N, \ldots, B_{NH}^N\}$ is sought of the form

$$u^N(t;\mathbf{q}) = \sum_{j=1}^{N+1} w_j^N(t)B_j^N = \sum_{j=1}^{N+1} w_j^N(t;\mathbf{q})B_j^N \quad (35)$$

where $\{B_j^N\}$ is the set of cubic spline basis functions appropriately modified to be in the domain of definition of the operators in equation (34). More precisely, let $\Delta^N = \{x_i\}_{i=0}^N$ with $x_i = il/N$ for $i = 0, 1, \ldots, N$, and let $\tilde{B}_j^N, j = -1, \ldots, N+1$ denote the standard $C^2(0,l)$ basis elements for the cubic B-spline subspaces of dimension $N+3$ corresponding to the grid Δ^N (see Prenter, 1975). Here $C^2(0,1)$ is the set of all continuous functions with continuous first and second derivatives on the interval $(0,l)$. Then B_j^N is given by

$$B_j^N = \tilde{B}_j^N, \text{ for } j = 2 \ldots NH$$
$$B_1^N = \tilde{B}_0^N - 2\tilde{B}_1^N - 2\tilde{B}_{-1}^N.$$

Note then that $X^N = S_L^3(\Delta^N) = \{\phi \in S^3(\Delta^N): \phi(0) = \phi^1(0) = 0\}$ where $S^3(\Delta^N) = \{\phi \in C^2(0,l);\ \phi$ is a cubic polynomial on each interval $[x_i, x_{i+1}]\}$.

The approximate solutions to (34) are determined from requiring that for all functions $\mathbf{z} \in X^N$

$$<\ddot{u}^N(t), \mathbf{z}> + <\mathcal{B}(\mathbf{q})\dot{u}^N(t), \mathbf{z}> + <\mathcal{A}(\mathbf{q})u^N(t), \mathbf{z}> = <f(t), \mathbf{z}> \quad (36)$$

with appropriate projected initial condition. Choosing $\mathbf{z} = B_i^N$ and using equation (35) one may write this in an equivalent matrix form as

$$M^N\ddot{\mathbf{w}}^N(t) + D^N\dot{\mathbf{w}}^N(t) + K^N\mathbf{w}^N(t) = \mathbf{F}^N(t) \quad (37)$$

where $\mathbf{w}^N$ is the $(N+1)\times 1$ vector $[w_0^N, w_1^N, \ldots, w_N^N]^T$ and where the $(N+1)\times(N+1)$ "mass," "damping," and "stiffness" matrices M^N, D^N, and K^N are defined by

$$M_{ij}^N = <B_i^N, B_j^N>$$
$$D_{ij}^N = <B_i^N, \mathcal{B}(\mathbf{q})B_j^N>$$
$$K_{ij}^N = <B_i^N, \mathcal{A}(\mathbf{q})B_j^N>.$$

Here, the subscript ij denotes the ijth element of the matrix and the vector $\mathbf{F}^N$ is defined as the $N\times 1$ vector of elements $F_j^N = <f(t), B_j^N>$. Each approximate identification problem now reduces to calculating $\mathbf{q}^N$ that minimizes (31) subject to the vector differential equation (37). These calculations result in the sequence $\{\mathbf{q}^N\}$ which, as mentioned above, converges to a $\mathbf{q}$ minimizing (30) subject to systems (9), (12), or (15), as appropriate.

It is important to note here that this approximation differs from a standard finite element method in two fundamental ways. First, the damping mechanism produced by the matrix D^N converges to a physical damping model. Typical finite element models treat damping in an ad hoc fashion. Secondly, the entire model converges to a strength of materials/continuum mechanics model having more physical significance than the rather arbitrary node model produced by standard finite element approximations.

VI Results

The results of estimating the various damping parameters from experimental data using EMA are discussed first as it is limited to the problem defined by $\mathbf{q}_1$. In the EMA approach,

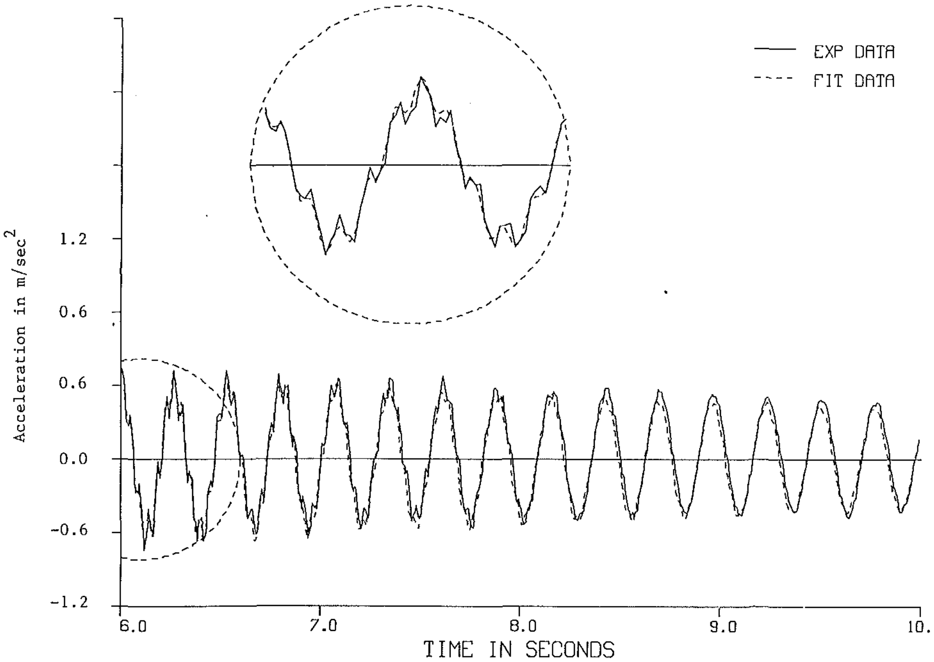


Fig. 1 A comparison of the experimentally measured time response and model using q_1. Estimated values are $E = 2.75 \times 10^{10}$ N/m^2, $c_d = 1.357 \times 10^6$ kg/m sec, $\gamma = .2025$ kg/m sec.

the correctness of a given estimate is judged by the ability of the solution to produce frequency-independent parameters γ and c_d. Next, the results of the SIP approach are used to examine the damping mechanism. In this case the success of a given estimate is judged on the ability of the estimate to numerically simulate the experimental time history of the structure's response, which is taken as the fundamental goal of a parameter identification procedure.

Experimental Modal Analysis. The geometric values of the beam tested and the analytical values of β_n are given in Table 1. The results of performing 15 modal tests, as outlined in Cudney and Inman (1989), are also listed in Table 1 for the system without a tip mass. Note the large damping ratios exhibited by this material when compared with calculated values of aluminum or steel. These data were first used in equation (27) to determine the values of the modulus (E) calculated to be

$$E = 2.68 \times 10^{10}\ \mathrm{N/m^2} \tag{38}$$

with a variance of 0.6 N/m^2.

The excellent fit provided by the Bernoulli-Euler beam stiffness to the measured frequencies indicated that this is a suitable stiffness model for this particular composite with 0 deg/90 deg orientations. Some researchers have suggested that a Timoshenko model might be more appropriate for composites. However, the inclusion of rotary inertia and shear effects did not provide a more convincing fit to the frequency data obtained here.

It has been shown by Cudney and Inman (1988) that attempting to use just air damping, γ, or just strain rate damping, c_d, alone fails to match the measured modal data. In each case the attempt to fit a single damping parameter is measured by the ability of the estimated values of γ and c_d, to reproduce the measured damping ratios $\hat{\zeta}_m$. This situation is discussed later in the context of the SIP estimates. The significance of this result is that a single modal damping ratio cannot logically be used to model the damping mechanism of the composite beam.

Next, the generalized inverse of the data matrix B defined by equation (29) and the theoretical values of the eigenvalues given in Table 1 are used to calculate the desired damping coefficients γ and c_d from equation (28) and the vector $\mathbf{z}$. The vector $\mathbf{z}$ contains the experimentally determined modal data of Table 1. The results of γ and c_d are

$$\gamma = 1.7561\ \mathrm{kg/m\ sec},\ c_d = 2.05 \times 10^5\ \mathrm{kg/m\ sec} \tag{39}$$

for 9 modes of data.

The effect of natural frequency on the measured modal damping ratio $\hat{\zeta}_m$ is seen by substituting the analytical expression for β_m into equation (26). This yields that

$$\hat{\zeta}_m = \frac{\gamma}{2\hat{\omega}_m} + \frac{c_d}{2E^2 I}\hat{\omega}_m. \tag{40}$$

This indicates clearly that the effect of air damping (8) decreases with increasing mode number($\hat{\omega}_m \rightarrow \infty$). Thus, for higher modes the strain rate damping makes a more significant contribution to the measured damping ratio. This agrees with the physically intuitive notion that the large-amplitude low-frequency modes are pushing more air than the higher-frequency lower-amplitude modes. In fact, for a free-free configuration, it is claimed by Vinson (1989) that the effect of air damping can be subtracted based on Blevins' equation (Blevins, 1977), which considers flow effects.

To check the validity of this approach, the frequency dependence of the coefficients γ and c_d was examined by recalculating them using a different number of modes. Successive least squares was performed using first 2 modes, then 3 modes, etc., up to the total of 9 modes. The result is illustrated in Table 2. Table 2 indicates that the estimates of γ and c_d depend somewhat on the frequency range of interest. This is inconsistent with the physical model put forth in Section III for estimating $\mathbf{q}_1$. In the modal testing community an acceptable error is measured damping ratios is typically 20–30 percent (Ewins, 1988). In this sense the modal data obtained (see Table 1) is valid because the damping ratios of modes 2 through 9 are within 23 percent of their mean value. However, the resulting values of γ and c_d obtained with these damping ratios are *not* consistent and their average values do not provide a reasonable reproduction of the measured time response when used in the original equation of motion (9). In fact, as shown in the last two columns of Table 2, the lack of convergence of the solution of equation (28) as the number of modes in-

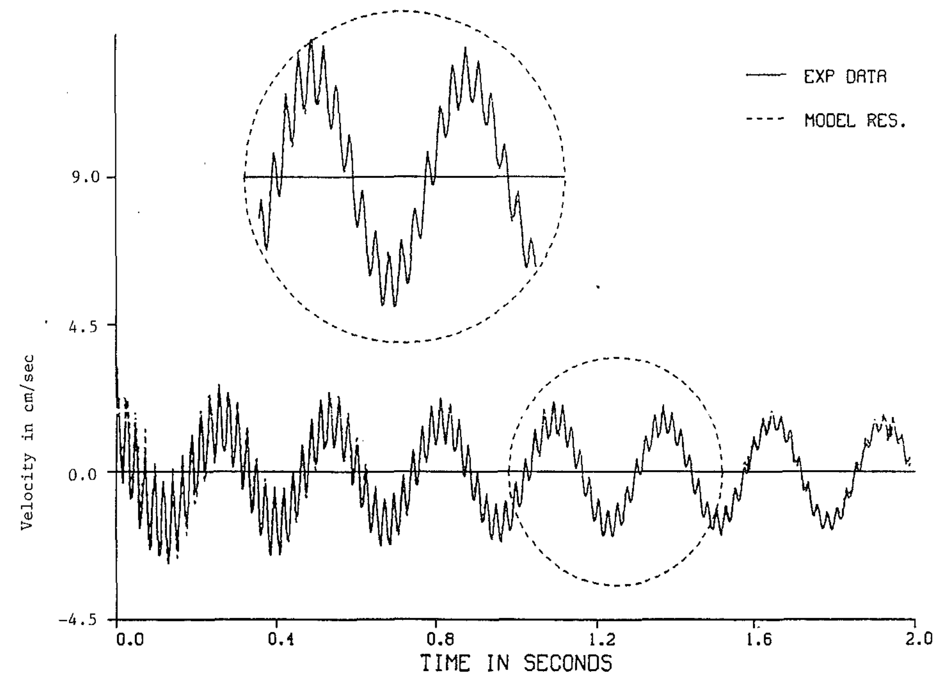


Fig. 2 A comparison of the experimentally measured time response and model. Estimated values are $E = 2.62 \times 10^{10}$ N/m^2, $\alpha = 20.02$ Nm$^2\sqrt{\text{sec}}$, $\beta = 6.54.108$ sec^{-1}.

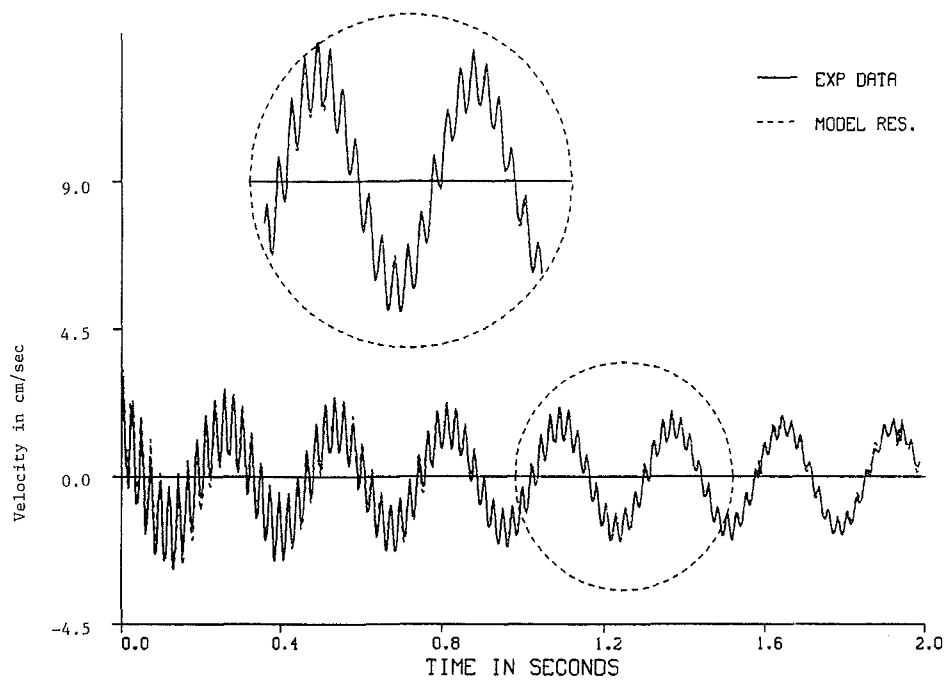


Fig. 3 A comparison of the experimentally measured time response using the spatial hysteresis model q_3. Estimated values are $E = 2.539 \times 10^{10}$ Nm^2, $a = 4.624$ N sec, $b = .0196$m, $\gamma = .5006$ kg/m sec.

creases (even when weighted) illustrates that the concept of modal damping has relatively little correlation with the physical damping mechanism of the structure under test and therefore traditional EMA cannot be used. Furthermore, this approach is not applicable to the problem of estimating $\mathbf{q}_1$ and $\mathbf{q}_3$. Hence, one must conclude that the modal approach is not satisfactory in attempting to model, in composite beams, any of the damping mechanisms proposed in this paper.

Spline Inverse Procedure. The use of the SIP provides a nonmodal approach appropriate for each of the problems of estimating $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$. The problem of estimating $\mathbf{q}_1$ is solved first for comparison with the modal approach. The stiffness parameter (elastic modulus) E was estimated to be 2.75×10^{10} N/m^2, in good agreement with the modal estimation results above. Estimates of air damping alone or strain rate alone as a damping model proved to be inadequate in reproducing time histories matching those of the experimental data, indicating a poor model. This is again in agreement with the results obtained by using the modal approach.

The time domain approach provided by SIP allows a convenient comparison between the measured time response and the analytical time response generated by the model of equation (9) with the experimentally determined parameter vector $\mathbf{q}_1$. The difference between the numerical solution for the time history of the acceleration $u_{tt}(x_i,t)$ for the analytical model with the estimated parameter $\mathbf{q}_1$ and the experimentally measured accelerations define the residual which is generally small (Banks et al., 1987). The analytical time response is plotted along with the measured time response versus time in Fig. 1. While the agreement in fair, the residual is larger for longer time intervals, warranting further modeling.

Next, the temporal hysteresis model involving $\mathbf{q}_2$ was considered as a possible candidate for modeling the damping in the composite. In this case, we used a slightly more complicated beam with tip mass. The estimation procedure produces a good value for E (i.e., consistent with our previous methods for estimating $\mathbf{q}_1$) but drives the air damping coefficient to zero. The residual, however, is better than that for the model with $\mathbf{q}_1$. Figure 2 illustrates a plot of the measured velocity versus time as well as the velocity predicted by the model with the estimate $\mathbf{q}_2$. The difference between the measured and predicted value over the time interval of interest is almost negligible. Because this model drives the air damping coefficient to zero (violating physical intuition), a third model ($\mathbf{q}_3$) was considered.

The last model considered is based on a concept of spatial hysteresis as defined by the estimation problem for $\mathbf{q}_3$. Again (using the same beam with tip mass), the resulting estimate of the elastic modulus E is consistent with those estimated previously. The values estimated for the spatial hysteresis parameters ($a = 4.624$ N sec, $b = .0196$m) and an air damping coefficient ($\gamma = .5006$ kg/m sec) produce an excellent match between predicted and measured response as indicated in Fig. 3 (Banks et al. 1988). However, the external damping coefficient γ differs from that estimated using the parameter vector $\mathbf{q}_1$ ($\gamma = .2025$ kg/m sec) emphasizing the fact that air damping should not be estimated independently whenever internal damping mechanisms are present.

VII Conclusion and Discussion

Three different models of damping have been presented to account for the experimentally observed dissipation in a pultruded composite beam. A spline-based inverse procedure (SIP), which relies on the distributed parameter nature of the damping mass and stiffness parameters, was proposed and used to estimate the form of each damping mechanism. External air damping, strain rate damping, spatial hysteresis, and time hysteresis models were considered. The spline-based method was also compared to a standard experimental modal analysis (EMA) approach. The EMA approach is not applicable to the various hysteresis models, nor is it applicable to systems with spatially varying parameters in general. Both the SIP and EMA approaches yield consistent values for the elastic modulus (E) for all three estimation models. This is consistent with the fact that frequencies are much more robust to estimates than are damping quantities.

Both hysteresis models produce better results than the strain rate damping model. However, the spatial hysteresis model permits a nonzero air damping term while time hysteresis does not. Since air damping is obviously present, the time hysteresis result is less satisfying. A comparison of the hysteresis models is given in Banks et al. (1988). As indicated in that presentation, further analysis and modeling is required before a conclusive decision can be made about a best model. It is clear from the results presented here that hysteretic damping is able to reproduce experimental time responses with more accuracy than the standard Kelvin-Voigt model. It is also clear that the standard method of measuring damping, EMA, does not provide an accurate method for investigating the energy dissipation in the composite beam tested here.

In summary, a new method of determining damping mechanisms in a distributed parameter model has been proposed and applied to a beam. The SIP method proposed here also yields estimates of the stiffness parameters. This method has been compared to the standard method of determining damping in structures using modal methods. When compared on the same experimental test data, the SIP approach produces more consistent estimates of the Kelvin-Voigt damping parameter than those obtained by using modal methods. In addition, the proposed procedure is applicable to hysteretic damping models and to systems with spatially varying parameters that cannot be treated by modal methods.

VIII Acknowledgments

This research was supported in part under the following grants and contracts: NASA: NAG-1-517 (HTB) and NAG-1-785 (DJI); NSF: MSM-8351807 (DJI) and MCS-8504316 (HTB); AFOSR: F49620-86-C-0111 (HTB and DJI) and several DOD equipment grants: AFOSR: 85-0119 (DJI) and AFOSR 87-0099 (DJI). The authors also wish to acknowledge H. H. Cudney, R. H. Fabiano, and Y. Wang for numerous discussions and for their efforts on both the experimental and computational aspects of this project.

References

Banks, H. T., Crowley, J. M., and Kunisch, K., 1983, "Cubic Spline Approximation Techniques for Parameter Estimation in Distributed Systems," *IEEE Transactions on Automatic Control*, Vol. AC-28, No. 7, pp. 773–786.

Banks, H. T., Crowley, J. M., and Rosen, I. G., 1986, "Methods for the Identification of Material Parameters in Distributed Models for Flexible Structures," *Mat. Aplicada e Comput.*, Vol. 5, pp. 139–168.

Banks, H. T., Fabiano, R. H., Wang, Y., Inman, D. J., and Cudney, H. H., 1988, "Spatial Versus Time Hysteresis in Damping Mechanisms," *Proceedings of the 27th IEEE Conference on Decision and Control*, pp. 1674–1677.

Banks, H. T., Fabiano, R. H., and Wang, Y., 1989, "Inverse Problem Techniques for Beams with Tip Body and Time Hysteresis Damping," *Mat. Aplicada e Comput.*, Vol. 8, pp. 101–118.

Banks, H. T, and Inman, D. J., 1989, "On Damping Mechanisms in Beams," ICASE Report No. 89-64, NASA Contractor Report 181904, NASA Langley Res. Ctr., Hampton, Va.

Banks, H. T., and Ito, K., 1988, "A Unified Framework for Approximation and Inverse Problems for Distributed Parameter Systems," *Control, Theory and Advanced Technology*, Vol. 4, pp. 73–90.

Banks, H. T., and Kunisch, K., 1989, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston, Mass.

Banks, H. T., Wang, Y., Inman, D. J., and Cudney, H. H., 1987, "Parameter Identification Techniques for the Estimation of Damping in Flexible Structures Experiments," *Proceedings of the 26th IEEE Conference on Decision and Control*, Vol. 2, pp. 1392–1395.

Blevins, R. D., 1977, *Flow-Induced Vibration*, Van Norstrand Reinhold, New York.

Caughey, T. K., and O'Kelly, M. E. J., 1965, "Classical Normal Modes in Damped Linear Systems," ASME Journal of Applied Mechanics, Vol. 37, pp. 583–588.

Clough, R. W., and Penzien, J., 1975, *Dynamics of Structures*, John Wiley and Sons, New York.

Cudney, H. H., and Inman, D. J., 1989, "Experimental Vertification on Damping Mechanisms in a Composite Beam," *Proceedings of the 7th International Modal Analysis Conference*, pp. 904–910.

Ewins, D. J., 1988, *Modal Testing Theory and Practice*, Research Studies Press, Letchworth, U. K.

Inman, D. J., 1989, *Vibration with Control, Measurement and Stability*, Prentice-Hall, Englewood Cliffs, N.J.

Prenter, P. M., 1975, *Splines and Variational Methods*, John Wiley and Sons, New York.

Russell, D. L., 1991, "On Mathematical Models for the Elastic Beam with Frequency Proportional Damping," *Control and Estimation in Distributed Parameter Systems*, SIAM, to appear.

Sprinak, G. T., and Vinson, J. R., 1988, "The Effect of Temperature on the Material Damping of Graphite; Epoxy Composites in a Simulated Space Environment," *Recent Advances in the Macro- Micro-Mechanics of Composite Material and Structures*, AMD-Vol. 4, ASME, New York, pp. 189–192.

Wilson, M. L., and Miserentino, R., 1986, "Pultrusion Processes Development for Long Space Boom Models," *Proceedings of the 41st Annual Conference of the Reinforced Plastics/Composites Industry*, Paper #6-D.

Vibrations of Moderately Elliptic Clamped Plates: A Perturbation Scheme for Eigenvalues[1]

R. Parnes
Department of Solid Mechanics, Materials and Structures,
Faculty of Engineering,
Tel Aviv University,
69978 Ramat Aviv, Israel

Frequencies of vibration of an elliptic plate, clamped along the edge, are determined by means of a perturbation scheme based on a boundary perturbation method (B.P.M.). Eigenvalues are obtained corresponding to higher modes of vibration containing elliptic nodes, in addition to the fundamental mode. Comparison with previously derived values in the fundamental mode reveals that the present scheme leads to accurate results for moderately elliptic plates.

Introduction

Frequencies of vibration of elastic plates, even according to classical simplified theories, have been determined analytically only for a limited class of plate geometries. An extensive survey and detailed review of this subject is given in a monograph by Leissa (1969). Vibrations of circular plates were first studied over a century ago by Lord Rayleigh (1945). Frequencies of vibration for circular plates subject to classical boundary conditions (i.e., fully constrained or free displacements at the outer edge) were obtained by Airey (1911), Carrington (1925), and others; corresponding cases for plates with elastic edge restraints were considered by Kantham (1958) and Parnes (1970). Thus, while circular plates have been extensively studied for a wide variety of boundary conditions, in the case of elliptic plates, very few results appear in the literature.

McLachlan (1947) investigated a number of vibrational problems in an elliptic domain using elliptic coordinates. However, while he obtained general solutions, no numerical results were presented. The major difficulty encountered with the use of this coordinate system evidently is due to the nature of the solution which, expressed in terms of Mathieu functions, gives rise to frequency equations represented in the form of infinite determinants. Using this coordinate system, Shibaoka (1956) determined an exact value (to four significant figures) for the fundamental frequency of clamped plates by evaluating zeros of truncated determinants of successively higher order. Eigenvalues which converged numerically, were thus obtained by means of this truncation technique. However, the method, while rather complex, did not lead to eigenvalues for higher modes of vibration. McNitt (1962) treated the same problem using the Galerkin method and a two-term deflection function expressed in terms of Cartesian coordinates. While this treatment resulted in reasonably accurate approximate eigenvalues for the first mode, the second mode results clearly are questionable, since the effect of ellipticity appears to be independent of the mode.

In this paper, we investigate the vibrations of a moderately elliptic clamped plate by means of a perturbation scheme. Boundary perturbation methods first appeared a number of years ago; an early exposition is given by Morse and Feshbach (1953). Recently, a higher-order boundary perturbation method (B.P.M.) was developed by Parnes and Beltzer (1986) to treat, among others, problems existing in an elliptic domain when the boundary conditions are of the Dirichlet or Neumann type. Considering the ellipse to be a perturbation of a circumscribing circle, it was shown that problems in the elliptic domain can be treated by solving a sequence of problems in the corresponding circular domain for which equivalent boundary conditions are prescribed. Explicit higher-order expressions for these boundary conditions were presented. Based on these results, we apply the B.P.M. to extract the eigenvalues for moderately elliptic plates. The technique yields simple analytic expressions in terms of the ellipticity for the eigenvalues of elliptic plates. Comparisons with results given by Shibaoka (1956) and McNitt (1962) for the fundamental mode show that the present scheme leads to highly accurate results for moderate ellipticities. Moreover, a significant feature of this method is the capability to extract eigenvalues in higher modes containing nodal ellipses in addition to the fundamental mode.

2 General Formulation and Solution

2.1 Formulation. We consider an elastic elliptic plate of thickness h with semi-major and minor radii a and b, respec-

[1]Major portions of this work were performed while the author was Visiting Professor at Laboratoire de Mécanique Théorique, Université Pierre et Marie Curie, Paris VI, and at Laboratoire de Mécanique des Solides, Ecole Polytechnique, Palaiseau, France.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Nov. 24, 1987; final revision, Nov. 22, 1988.

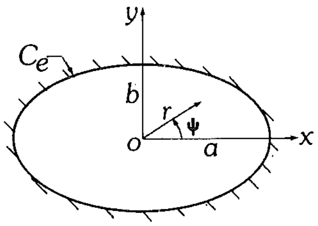


Fig.1 Geometry of problem

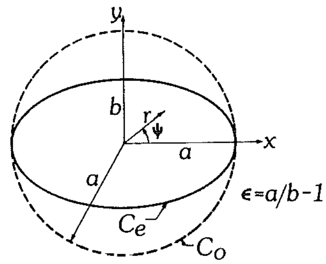


Fig. 2 Perturbed geometry

tively, clamped along the boundary C_e. Using a polar coordinate system (r, ψ) with center at 0 (Fig. 1), and denoting the transverse displacement by $W(r, \psi, t)$, the equation governing free vibrations of the plate, $W(r, \psi, t) = w(r, \psi)e^{i\omega t}$, is

$$(\nabla^4 - k^4)w(r,\psi) = 0 \tag{1}$$

where

$$k^4 = m\omega^2/D \tag{2a}$$

with

$$D = E\,h^3/12(1-\nu^2). \tag{2b}$$

In the above, ∇^4 is the bi-harmonic operator, E and ν are the modulus of elasticity and Poisson ratio of the plate material, respectively, m is the mass of the plate per unit area, and ω is the frequency.

Equation (1) is to be solved subject to the boundary conditions on C_e,

$$w\Big]_{C_e} = 0, \; \frac{\partial w}{\partial n}\Big]_{C_e} = 0, \tag{3a,b}$$

where $\mathbf{n}$ is the unit normal to C_e.

Mathematically, equations (1)–(3) represent a boundary value problem in an elliptic domain containing Dirichlet or Neumann-type boundary conditions. The system clearly does not lend itself in the given elliptic domain to a direct analytic and tractable solution yielding the required eigenvalues. We proceed, therefore, with a perturbation scheme, adopting the boundary perturbation method (B.P.M.) presented by Parnes and Beltzer (1986) for this class of boundary conditions. In accordance with the B.P.M., for moderately elliptic plates, we consider the boundary C_e as a perturbation of a circumscribing circle C_0 of radius a (Fig. 2), and define the ellipticity parameter

$$\epsilon = \frac{a}{b} - 1. \tag{4}$$

We assume $w(r, \psi)$ to be analytic throughout the x-y plane and let $w(r, \psi)$ be expandable in the ellipticity parameter ϵ:

$$w(r,\psi) = \sum_{j=0}^{n} w^{(j)}\epsilon^j. \tag{5}$$

The functions $w^{((j))}(r, \psi), j = 0, 1, \ldots, n$, must then satisfy equivalent boundary conditions on the boundary $C_0(r=a)$; explicit expressions for these boundary conditions, derived in the above-referenced paper for a second-order scheme, are as follows:

$$w^{(0)}]_{C_0} = w]_{C_e}, \; w^{(0)}_{,r}\Big]_{C_0} = \frac{\partial w}{\partial n}\Big]_{C_e} \qquad (6a, b)^{2}$$

$$w^{(1)}]C_0 = -{}_0\Psi]_1^{(0)}, \; w^{(1)}_{,r}]_{C_0} = -{}_n\Psi_1^{(0)} \qquad (6c, d)$$

$$w^{(2)}]C_0 = -{}_0\Psi_1^{(1)} - {}_0\Psi_2^{(0)} \tag{6e}$$

$$w^{(2)}_{,r}]_{C_0} = -{}_n\Psi_1^{(1)} - {}_n\Psi_2^{(0)} \tag{6f}$$

where, for $i = 0, 1$,

$${}_0\Psi_1^{(i)} = -a\sin^2\psi\, w^{(i)}_{,r} \tag{7a}$$

$${}_0\Psi_2^{(i)} = \frac{a}{2}\sin^2\psi\left[(2\sin^2\psi - \cos^2\psi)\, w^{(i)}_{,r} + a\sin^2\psi\, w^{(i)}_{,rr}\right] \tag{7b}$$

$${}_n\psi_1^{(i)} = -a\sin^2\psi\, w^{(i)}_{,rr} + \frac{\sin 2\psi}{a} w^{(i)}_{,\psi} \tag{7c}$$

$${}_n\psi_2^{(i)} = \frac{a^2}{2}\sin^4\psi\, w^{(i)}_{,rrr} - \frac{a}{2}(\cos^2\psi - 2\sin^2\psi)\sin^2\psi\, w^{(i)}_{,rr}.$$
$$-\frac{1}{2}(\sin 2\psi)^2\, w^{(i)}_{,r} - \sin^2\psi \sin 2\psi\, w^{(i)}_{,r\psi} + \frac{1}{4a}\sin 4\psi\, w^{(i)}_{,\psi}. \tag{7d}$$

Invoking a procedure suggested by Millman and Keller (1969) in their study of nonlinear boundary value problems, we expand k^4 similarly in terms of the ellipticity,

$$k^4 = \sum_{j=0}^{n} k_j^4 \epsilon^j \tag{8}$$

so as to be able to satisfy appropriate solvability conditions which appear below. Substituting equations (5) and (8) in (1), leads to a set of equations for $w^{(j)}(r, \psi)$

$$(\nabla^4 - k_0^4)w^{(0)} = 0 \qquad j = 0 \tag{9a}$$

$$(\nabla^4 - k_0^4)w^{(j)} = \sum_{i=0}^{j-1} k_{j-i}^4 w^{(i)} \; j \geq 1. \tag{9b}$$

We observe that for any $j \geq 1$, a sequential solution of equation (9b) is possible since the inhomogeneous terms $i = 1, \ldots (j-1)$ are then known. In the following development we proceed with a second-order scheme.

2.2 The axisymmetric case, $j = 0$. The governing equation, (9a), with boundary conditions given by (6a, b) where $w]_{C_e} = \frac{\partial w}{\partial n}\Big|_{C_e} = 0$, is recognized as that for the free vibration of a circular plate of radius a.

This problem admits the solution (Leissa, 1969)

$$w^{(0)}(r) = A_0[J_0\,(k_0\,r) + \alpha_0 I_0(k_0\,r)] \tag{10a}$$

where

$$\alpha_0 = -J_0(\lambda_0)/I_0(\lambda_0) \tag{10b}$$

and where the eigenvalues,

$$\lambda_{0,s} = k_{0,s}\,a$$

represent the s roots ($s = 0, 1, 2, \ldots$) of the frequency equation

$$I_0(\lambda_0)\;J_1(\lambda_0) + I_1(\lambda_0)J_0(\lambda_0) = 0. \tag{11}$$

In the foregoing equations J_n and I_n denote, respectively, the Bessel and modified Bessel functions of order n.

For use in the treatment that follows, it is noted that

$$w^{(o)}_{,r}(r) = A_0 k_0[-J_1(k_0 r) + \alpha_0 I_1(k_0\,r)] \tag{12a}$$

$$w^{(o)}_{,rr}(r) = A_0 k_0^2[-J_0(k_0 r) + \alpha_0 I_0(k_0\,r)] - \frac{1}{r} w^{(o)}_{,r}(r). \tag{12b}$$

[2] In these, and all subsequent equations, a comma followed by a subscripted r or ψ denotes derivatives with respect to the variable.

Hence, by equation (10*b*),

$$w^{(o)}_{,rr}(a) = -\frac{2\lambda_0^2 A_0}{a^2} J_0(\lambda_0). \quad (12c)$$

Making use of the standard relations for the Bessel functions, we note that

$$w^{(o)}_{,rrr}(a) = -\frac{1}{a} w^{(o)}_{,rr}(a) + 2A_0 k_0^3 J_1(\lambda_0). \quad (13a)$$

Hence, we may write

$$w^{(o)}_{,rrr}(a) = -\frac{C}{a} w^{(o)}_{,rr}(a) \quad (13b)$$

where

$$C = -[1 + \lambda_0 J_1(\lambda_0)/J_0(\lambda_0)]. \quad (13c)$$

2.3 Perturbed solutions; $j \geq 1$. We observe that the governing equations, equation (9*b*), contain inhomogeneous terms which may be interpreted physically as "forcing" terms. Since we are concerned here with free vibrations, we shall seek, by means of appropriate transformations, to find (while still satisfying the required boundary conditions) related auxiliary functions which are governed by corresponding homogeneous equations. This physical reasoning will be seen to be consistent with the required solvability conditions on the explicit equations derived as follows.

2.3.1. Case j = *1:* The explicit governing equation on $w^{(1)}$ (r, ψ), from (9*b*), is

$$(\nabla^4 - k_0^4)\, w^{(1)}(r,\psi) = k_1^4 w^{(o)}(r) \quad (14)$$

subject to the boundary conditions, (6*c*,*d*). Upon noting that $w^{(o)}_{,r}(a) = w^{(o)}_{\psi}(a) = 0$, these yield the simpler expressions

$$w^{(1)}(a) = \frac{a}{2}(1 - \cos 2\psi)\, w^{(o)}_{,r}(a) = 0 \quad (15a)$$

$$w^{(1)}_{,r}(a) = \frac{a}{2}(1 - \cos 2\psi)\, w^{(o)}_{,rr}(a). \quad (15b)$$

We now define the function $\tilde{w}^{(1)}$ by means of the transformation

$$w^{(1)}(r,\psi) = \tilde{w}^{(1)}(r,\psi) + \frac{r}{2} w^{(o)}_{,r}(r). \quad (16)$$

It is observed that equations (15) are satisfied if

$$\tilde{w}^{(1)}(a,\psi) = 0 \quad (17a)$$

$$\tilde{w}^{(1)}_{,r}(a,\psi) = -\frac{a}{2} w^{(o)}_{,rr}(a) \cos 2\psi. \quad (17b)$$

Substituting equation (16) in (14),

$$(\nabla^4 - k_0^4)\left[\tilde{w}^{(1)} + \frac{r}{2} w^{(o)}_{,r}\right] = k_1^4 w^{(o)}. \quad (18)$$

Now, using equation (12*a*), it is seen that

$$(\nabla^4 - k_0^4)[r\, w^{(o)}_{,r}] = 4\, k_0^4 w^{(o)}. \quad (19)$$

Hence, equation (18) becomes

$$(\nabla^4 - k_0^4)\tilde{w}^{(1)}(r,\psi) = (k_1^4 - 2k_0^4)\, w^{(o)}(r). \quad (20)$$

The solvability conditions for possible solutions of equation (20) subject to the boundary conditions of equation (17) requires that the right-hand side of equation (20) be made to vanish[3]. (This condition is evident, in this particular case upon recognizing that equation (17) contain a ψ-dependency while the right-hand side of equation (20) is ψ-independent.) Thus, we are led to the condition

$$k_1^4 = 2\, k_0^4. \quad (21)$$

The function $\tilde{w}^{(1)}$, satisfying the corresponding homogeneous equation and consistent with the boundary conditions (17), is

$$\tilde{w}^{(1)}(r,\psi) = \chi^{(1)}(r) \cos 2\,\psi \quad (22a)$$

where

$$\chi^{(1)}(r) = A_1[J_2(k_0 r) + \alpha_1 I_2(k_0 r)]. \quad (22b)$$

From equation (17),

$$\chi^{(1)}(a) = 0,\ \chi^{(1)}_{,r}(a) = -\kappa, \quad (23a, b)$$

where

$$\kappa = \frac{a}{2} w^{(o)}_{,rr}(a). \quad (23c)$$

Hence, the constants are readily determined:

$$A_1 = -\frac{\kappa\, I_2(\lambda_0)}{4\, a\, I_1(\lambda_0)\, J_1(\lambda_0)} \quad (24a)$$

$$\alpha_1 = J_2(\lambda_0)/I_2(\lambda_0). \quad (24b)$$

Thus, finally we obtain

$$w^{(1)}(r,\psi) = \chi^{(1)}(r) \cos 2\,\psi + \frac{r}{2} w^{(o)}(r). \quad (25)$$

Using standard expressions for derivatives of the Bessel functions and recurrence relations, and making use of equation (11), we find the following relation (necessary in the subsequent treatment)

$$\chi^{(1)}_{,rr}(a) = B\, w^{(o)}_{,rr}(a) \quad (26a)$$

where

$$B = \frac{\lambda_0^2 J_2(\lambda_0) I_2(\lambda_0)}{4\, J_1(\lambda_0)\, I_1(\lambda_0)}. \quad (26b)$$

2.3.2 Case j = *2*: Combining equations (9b), (16), and (21), the explicit governing equation for $w^{(2)}$ becomes

$$(\nabla^4 - k_0^4) w^{(2)}(r,\psi) = [2\,\tilde{w}^{(1)}(r,\psi) + r\, w^{(o)}(r)]k_0^4 + k_2^4 w^{(o)}(r). \quad (27)$$

Substituting equation (25) in the appropriate boundary conditions, (6*e*, *f*), we obtain explicitly

$$w^{(2)}(a,\psi) = \frac{a^2}{16}(3 - 4\cos 2\psi + \cos 4\,\psi)\, w^{(o)}_{,rr}(a) \quad (28a)$$

$$w^{(2)}_{,r}(a,\psi) = \frac{a}{16}\{[3\, w^{(o)}_{,rr}(a) + a\, w^{(o)}_{,rrr}(a) - 4\chi^{(1)}_{,rr}(a)] \quad (28b)$$
$$+ 8\chi^{(1)}_{,rr}(a) \cos 2\psi - [3 w^{(o)}_{,rr}(a) + 4\chi^{(1)}_{,rr}(a) + a\, w^{(o)}_{,rrr}(a)] \cos 4\,\psi\}.$$

Making use of the relations given by equations (13) and (26), (28*b*) can be written alternatively as

$$w^{(2)}_{,r}(a,\psi) = \frac{a}{16}\{2[1 - 2B - C]\, w^{(o)}_{,rr}(a) + 2\, a\, w^{(o)}_{,rrr}(a)$$
$$+ 2A_0 k_0^3 a\, J_1(\lambda_0) + 8\chi^{(1)}_{,rr}(a) \cos 2\psi \quad (28b')$$
$$- [3 + 4\,B + C] w^{(o)}_{,rr}(a) \cos 4\,\psi\}.$$

Taking note of equation (12*c*), we define the auxiliary function $\tilde{w}\,(r,\,\psi)$ via the transformation

$$w^{(2)}(r,\psi) = \tilde{w}^{(2)}(r,\psi) + \Gamma_0(r) + \Gamma_2(r) \cos 2\psi \quad (29)$$

where

[3]It is noted that this mathematical condition is compatible with the physical argument given above.

$$\Gamma_0(r)=\frac{1}{8}\left[r^2\, w_{,rr}^{(o)}(r)-8\Lambda\, r\, w_{,r}^{(o)}-A_0\,\lambda_0^2\, J_0\,(k_0\, r)\right] \quad (30a)$$

$$\Gamma_2(r)=\frac{1}{4}\left[(1-2\gamma)\, r\, w_{,r}^{(o)}(r)-4\gamma\,\chi^{(1)}(r)+2r\,\chi_{,r}^{(1)}(r)\right] \quad (30b)$$

with

$$\Lambda(\lambda_0)=\frac{1}{8}(1+2B+C) \quad (30c)$$

and where γ is an arbitrary constant.

By direct substitution, we find that equations (28a) and (28b') are satisfied for all values of γ provided that $\tilde{w}^{(2)}$ satisfy the boundary conditions

$$\tilde{w}^{(2)}(a,\psi)=\frac{a^2}{16}\, w_{,rr}^{(o)}(a)\cos 4\,\psi \quad (31a)$$

$$\tilde{w}_{,r}^{(2)}(a,\psi)=-\frac{a}{16}\,[3\, w_{,rr}^{(o)}(a)+4\,\chi_{,rr}^{(1)}(a)+a\, w_{,rrr}^{(o)}(a)]\cos 4\,\psi. \quad (31b)$$

Substituting equation (29) in (27) and making use of (22a), we obtain

$$(\nabla^4-k_0^4)\;[\tilde{w}^{(2)}(r,\psi)+\Gamma_0(r)+\Gamma_2(r)\cos 2\,\psi]$$
$$=[2\chi^{(1)}(r)\cos 2\psi+r\, w_{,r}^{(o)}(r)]\, k_0^4+k_2^4\, w_{(r)}^{(o)}. \quad (32)$$

From equation (12b) and equation (19) we can show that

$$(\nabla^4-k_0^4)\;[r^2\, w_{,rr}^{(o)}(r)]=12\, k_0^4\, w^{(o)}(r)+8\, k_0^4 r\, w_{,r}^{(o)}(r). \quad (33)$$

Hence,

$$(\nabla^4-k_0^4)\Gamma_0(r)=k_0^4\left[\left(\frac{3}{2}-4\Lambda\right)w^{(o)}(r)+r\, w_{,r}^{(o)}(r)\right]. \quad (34)$$

Similarly, upon letting the arbitrary constant γ take the convenient value $\gamma=\frac{1}{2}$ (thus simplifying Γ_2 by the elimination of the $w_{,r}^{(o)}$ term) and using the recurrence relations for Bessel functions, equation (30b) leads to

$$(\nabla^4-k_0^4)\,[\Gamma_2(r)\cos 2\,\psi]=2\, k_0^4\tilde{w}^{(1)}(r,\psi). \quad (35)$$

Equations (34)–(35), when substituted in equation (32), then yield

$$(\nabla^4-k_0^4)\;\tilde{w}^{(2)}(r,\psi)=\left[\left(4\Lambda-\frac{3}{2}\right)k_0^4+k_2^4\right]w^{(o)}(r). \quad (36)$$

As with the case $j=1$ [(equation 20)], the solvability condition for solutions of equation (36) consistent with the boundary conditions on $\tilde{w}^{(2)}$ (r,ψ), equation (31), requires that the inhomogeneous term be made to vanish. Thus, we are led to the condition

$$k_2^4=\left(\frac{3}{2}-4\,\Lambda\right)k_0^4 \quad (37)$$

yielding the solution

$$\tilde{w}^{(2)}(r,\psi)=A_2[J_4(k_0\, r)+\alpha_2\, I_4(k_0\, r)]\cos 4\,\psi \quad (38)$$

where A_2 and α_2 are constants which can be determined by equation (31).[4]

3 Eigenvalues: Numerical Results and Discussion

From equations (21) and (37), the required second-order expansion for the eigenvalues $\lambda^4=(ka)^4$, according to equation (8), is

$$\lambda^4=\lambda_0^4\{1+2\epsilon+\left[\frac{3}{2}-4\Lambda(\lambda_0)\right]\epsilon^2\}. \quad (39)$$

Substituting equations (13c), and (26b) in the definition of Λ, equation (30c), and again making use of the recurrence relations for the Bessel functions, yields

$$\lambda_s^4=\lambda_{0,S}^4\;[1+2\epsilon+\beta(s)\epsilon^2] \quad (40)$$

where $\beta(s)\equiv 3/2-4\,\Lambda(\lambda_{0,s})$ is given explicitly by

$$\beta(s)=\frac{1}{8}\left\{2-\lambda\left[\frac{J_1(\lambda)}{J_0(\lambda)}-\lambda\frac{J_0(\lambda)\, I_0(\lambda)}{2\, J_1(\lambda)\, I_1(\lambda)}\right]\right\}_{\lambda=\lambda_{0,S}}. \quad (41)$$

In the foregoing equation we note that the coefficient of the quadratic term ϵ^2 depends explicitly on the sth root $\lambda_{0,S}$ of the frequency equation of the circular plate, equation (11). Values of $\beta(s)$ corresponding to the first six modes ($s=0, 1, \ldots, 5$) are given in Table 1.

Table 1

s	0	1	2	3	4	5
$\beta(s)$	2.9499	9.4352	20.959	37.429	58.840	85.188

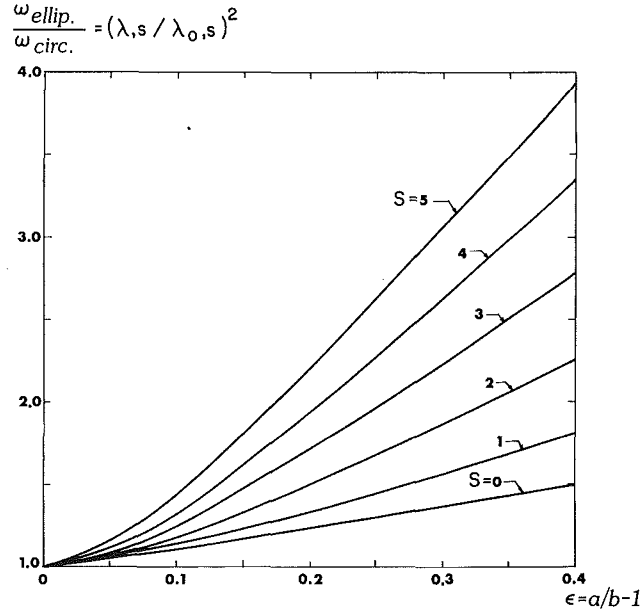


Fig. 3 Frequency ratios versus ellipticity

Realizing that the ellipticity does not result merely in a translation of the spectrum of frequencies of the circular plate, it is evident that the ellipticity should affect each mode differently as shown here. This result is in contradistinction with that given by McNitt (1962) where the ellipticity is presented as affecting the first two modes equally.

The effect of ellipticity on the eigenvalues is presented numerically in Fig. 3, where the family of curves, $(\lambda/\lambda_{0,s})^2=\frac{\omega\text{ ellipt}}{\omega\text{ circle}}$, is plotted as a function of ϵ for $\epsilon\le 0.4$ for the first six modes, $s=0, 1, \ldots, 5$. It is seen that the ellipticity has an increasingly greater effect upon the higher modes. For example, for an ellipse with $\frac{a}{b}=1.4$, the fundamental frequency increases by 50 percent while the frequency corresponding to the fourth mode ($s=3$) increases by a factor of 2.8. A plausible interpretation for the increased effect of ellipticity on the higher

[4]It is noted that the present second-order scheme does not require evaluation of these constants.

Table 2

$\frac{a}{b}$	$\left(\frac{\lambda}{\lambda_0}\right)^2$ BPM	$\left(\frac{\lambda}{\lambda_0}\right)^2$ Shibaoka	Δ (%)	$\frac{a}{b}$	$\left(\frac{\lambda}{\lambda_0}\right)^2$ BPM	$\left(\frac{\lambda}{\lambda_o}\right)^2$ McNitt	Δ (%)
1.25	1.2978	1.2824	1.2	1.1	1.1088	1.1074	0.13
1.5	1.655	1.61*	3.0	1.2	1.232	1.2299	0.17
2.0	2.439	2.692	9.4	1.5	1.655	1.6663	0.68

*Value taken from Fig. 4 of Shibaoka (1956)

modes of vibration may be given as follows: We first note that with increasing ellipticity, the wavelengths in the higher modes, because of the greater number of nodes, become increasingly smaller, particularly along the minor axis. Moreover, we recall, for example, as in the case of circular plates, that the *sensitivity* of eigenvalues corresponding to higher s-modes increases with the wave number ((Leissa, 1969)), i.e., as the wavelengths in the higher modes diminish, the corresponding eigenvalues increase at a faster rate. By analogy with these results, we may conclude that in elliptic plates the resulting smaller wavelengths in the higher modes, due to increased ellipticity, correspond to larger eigenvalues which increase even more rapidly with s.

For the *fundamental* mode, it is possible to compare the results obtained here with the exact frequencies given by Shibaoka (1956) and the approximate frequencies derived by McNitt (1962); these are shown in Table 2, together with the percent differences.

We note that for $\frac{a}{b} \le 1.5$, the results are in agreement to within 3 percent, and moreover even for a relatively large ellipticity, $\frac{a}{b} = 2.0$, the solutions differ by less than 10 percent. We may thus conclude that the perturbation scheme presented here leads to results of acceptable accuracy for moderate ellipticities in the range $\epsilon < 0.6$.

Finally, we observe that the perturbation scheme, while being a relatively simple technique, possesses a distinct advantage over previous solutions, namely the capability of yielding eigenvalues for higher modes of vibration.

Acknowledgment

The author wishes to acknowledge several fruitful discussions on this subject with his colleague, Prof. A. Whitman formerly of Tel-Aviv University.

References

Airey, J. 1911, "The Vibration of Circular Plates and their Relation to Bessel Functions," *Proceedings Physical Society (London)*, Vol. 23, pp. 225—232.

Carrington, H., 1925, "The Frequencies of Vibration of Flat Circular Plates Fixed at the Circumference", *Philosophical Magazine*, Vol. 50, No. 6, pp. 1261-1264.

Kantham, C. L., 1958, "Bending and Vibration of Elastically Restrained Circular Plates," *Journal of the Franklin Institute*, Vol. 265, No. 6, pp. 483-491.

Leissa, A. W., 1969, "*Vibration of Plates*," NASA Report SP-160.

Lord Rayleigh, 1945, *Theory of Sound*, Vol. 1 and Vol. 2, Dover.

McLachlan, N. W., 1947, "Vibrational Problems in Elliptic Coordinates," *Quarterly of Applied Mathematics*, Vol. 5, pp. 289-297.

McNitt, R. P., 1962, "Free Vibration of a Damped Elliptic Plate," *Journal of the Aerospace Sciences*, Vol. 29, No. 9, pp. 1124-1125.

Millman, M. H., and Keller, J. B., 1969, "Perturbation Theory of Nonlinear Boundary-Value Problems," *Journal of Mathematical Physics*, Vol. 10, No. 2, pp. 342-361.

Morse, P. M., and Feshbach, H., 1953, *Methods of Theoretical Physics*, Vol. 2, McGraw-Hill, New York.

Parnes, R., 1970, "Axi-Symmetric Dynamic Response of Circular Plates with Arbitrary Elastic Boundary Restraints," *International Journal of Solids and Structures*, Vol. 6, pp. 21-36.

Parnes, R., and Beltzer, A. I., 1986, "A Higher-Order Boundary Perturbation Method for Asymmetric Dynamic Problems in Solids—Part I: General Formulation," *International Journal of Solids and Structures*, Vol. 22, No. 11, pp. 1177-1187.

Shibaoka, Y., 1956, "On the Transverse Vibration of an Elliptic Plate with Clamped Edge," *Journal of the Physical Society of Japan*, Vol. 11, No. 7, pp. 797-803.

H. P. W. Gottlieb
Division of Science and Technology,
Griffith University,
Nathan, Queensland 4111
Australia

Inhomogeneous Clamped Circular Plates With Standard Vibration Spectra

Theory

In this paper, we extend a method of special conformal maps developed by the author (Gottlieb, 1988) for circular membranes and discover a family of *inhomogeneous* clamped circular plates having the same vibration spectrum as for a *homogeneous* plate.

Let $h(x, y) = g(u, v)$ where $u + iv = f(x + iy)$, where $f(z)$ is a one-to-one analytic mapping of complex variables from the $z = x + iy$ plane to the $u + iv$ plane. Then (c.f., Gottlieb, 1988) $\nabla^2 h = |f'(z)|^2 \tilde{\nabla}^2 g$ where $\nabla^2 = (\partial^2/\partial x^2) + (\partial^2/\partial y^2)$, $\tilde{\nabla}^2 = (\partial^2/\partial u^2) + (\partial^2/\partial v^2)$. A repeated application gives $\tilde{\nabla}^4 g = |f'(z)|^{-2} \nabla^2 [|f'(z)|^{-2} \nabla^2 h(x, y)]$.

Suppose that in Cartesian u, v coordinates, g is the amplitude of small vibrations of a thin *homogeneous* plate. Then (c.f., Meirovitch, 1967)

$$D_0 \tilde{\nabla}^4 g = \omega^2 \sigma_0 g(u,v) \tag{1}$$

where ω is the radian frequency of the modal vibration, and $D_0 = E_0 H_0^3/[12(1-\upsilon_0^2)]$, with standard definitions of the symbols. Substitution gives

$$\nabla^2[(D_0|f'(z)|^{-2})\nabla^2 h] = \omega^2(\sigma_0|f'(z)|^2)h(x,y). \tag{2}$$

The equation governing the small transverse vibrations of amplitude $W(x, y)$ with radian frequency ω for a thin *inhomogeneous* elastic plate has been given by Lang and Nemat-Nasser (1978) in terms of $\sigma(x, y)$, the mass density per unit area, $\upsilon(x, y)$, the Poisson's ratio, and the flexural rigidity of the plate given by $D(x, y) = EH^3/[12(1-\upsilon^2)]$ where $H(x, y)$ is the plate thickness and $E(x, y)$ is Young's modulus. If we impose the condition

$$D(1-\upsilon) = c, \text{ constant (positive)}, \tag{3}$$

then we may show that it becomes a "conditional" plate equation

$$\nabla^2[D\nabla^2 W] = \omega^2 \sigma W. \tag{4}$$

Now, equation (2) is *precisely of this form*, with

$$D(x,y) = D_0|f'(z)|^{-2}, \quad \sigma(x,y) = \sigma_0|f'(z)|^2. \tag{5a, b}$$

(Thus, $\sigma(x, y)\, D(x, y) = \sigma_0 D_0$.) By (3) and (5$a$),

$$\upsilon(x,y) = 1 - c|f'(z)|^2/D_0, \tag{6}$$

$$EH^3 = 12c[2 - c|f'(z)|^2/D_0]. \tag{7}$$

For a physical theory, $0 < \upsilon(x, y) < 1/2$. From (3), this implies $D(\mathrm{MAX})/2 < c < D(\mathrm{MIN})$ where $D(\mathrm{MAX})$ (MIN) is the maximum (minimum) value attained by $D(x, y)$ over the physical region of the plate coordinates. From (5a), this yields the condition on $f(z)$: $|f'(z)|(\mathrm{MAX}) < \sqrt{2}|f'(z)|(\mathrm{MIN})$.

A vibrating plate will be subject to certain boundary conditions. For a *clamped* plate

$$W = 0 = \partial W/\partial n \tag{8}$$

on the boundary, where n is measured along the normal to the bounding curve. Now, if $g(u, v)$ satisfies the conditions (8) on the boundary of a region in the (u, v) plane, then $h(x, y)$ satisfies (8) on the corresponding boundary in the (x,y) plane mapped by f (Saff and Snider, 1976). Thus, our theory allows us to relate certain *inhomogeneous* clamped plates to *homogeneous* clamped plates via the complex mapping function f.

Clamped Circular Plates

For our mapping function, we take the bilinear (Möbius) transformations

$$f(z) = R(z - \alpha R)/(R - \alpha z), \quad 0 \le \alpha < 1. \tag{9}$$

Up to rotations of coordinates, these are the only one-to-one analytic mappings of the disk of radius R onto itself (c.f., Saff and Snider, 1976), with perimeter mapped onto perimeter: $|z| = R \Leftrightarrow |f(z)| = R$.

Then from $u + iv = \rho\exp(i\phi) = f(z)$, corresponding to (9), (c.f., Gottlieb, 1988)

$$\rho\cos\phi = R[(x - \alpha R)(R - \alpha x) - \alpha y^2]/[d(x,y)]^2, \tag{10a}$$

$$\rho\sin\phi = R^2(1-\alpha^2)y/[d(x,y)]^2, \tag{10b}$$

$$[d(x,y)]^2 = (R - \alpha x)^2 + \alpha^2 y^2; \tag{10c}$$

$$|f'(z)| = R^2(1-\alpha^2)/[d(x,y)]^2. \tag{11}$$

The corresponding material properties for these plates are now given by

$$D(x,y) = D_0[d(x,y)]^4/[(1-\alpha^2)^2 R^4], \tag{12}$$

$$\sigma(x,y) = \sigma_0[D_0/D(x,y)], \tag{13}$$

$$\upsilon(x,y) = 1 - [c/D(x,y)], \tag{14}$$

$$E(x,y)[H(x,y)]^3 = 12c\{2 - [c/D(x,y)]\}. \tag{15}$$

In terms of the *inhomogeneous* plate, polar coordinates $x =$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer. The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, Oct. 11, 1989; final revision, Apr. 20, 1990.

$r\cos\theta$, $y = r\sin\theta$, then $[d]^2 = R^2 + \alpha^2 r^2 - 2R\alpha r\cos\theta$, where d has the physical meaning of the distance from the point on the rim ($r = R$, $\theta = 0$) to the internal, *scaled* point (αr, θ). Thus, this resulting *inhomogeneous* plate is, unexpectedly, *not* axisymmetric.

The previous inequality involving $|f'|$ (so that Poisson's ratio is physical) imposes a stricter condition on the parameter α in (9). Using (11), $1+2\alpha+2\alpha^2 < \sqrt{2}(1-\alpha)^2$, i.e.,

$$0 \le \alpha < [(17+13\sqrt{2})^{1/2}+3+2\sqrt{2}]^{-1} = 0.08491. \quad (16)$$

Then the constant c introduced in the special condition (3), and appearing in (14), (15), must satisfy, for any α chosen in the range (16), the condition

$$[1+2\alpha+2\alpha^2]^2/[2(1-\alpha^2)^2] < c/D_0 < (1-\alpha)^2/(1+\alpha)^2. \quad (17)$$

The angular frequencies for (1) with clamped boundary conditions (8) are given by $\omega_n^m = (\gamma_n^m)^2(D_0/\sigma_0)^{1/2}/R^2$ where γ_n^m is the mth root of the characteristic equation $J_n(\gamma)I_{n-1}(\gamma) - J_{n-1}(\gamma)I_n(\gamma) = 0$ where J_n and I_n ($n = 0, 1, 2, \ldots$) are, respectively, the ordinary and modified Bessel functions of the first kind. The natural mode functions $g(u,v) = G_n^m(\rho, \phi)$ are given, for example, in Meirovitch (1967) (Sect. 5·12(b), with change of notation). By expressing ρ, ϕ in terms of r, θ using (10*a*, *b*), these become the mode functions for our inhomogeneous plate, which therefore has different mode shapes.

By equations (1) and (8), the modal frequency spectrum of our *inhomogeneous* circular plate of equations (2), (3), (4) with material properties given by (12)–(15) with any parameters α, c in the ranges (16), (17), is *identical* with the spectrum of the vibrations of the standard clamped homogeneous circular plate given above.

Discussion

We have shown that there is a family of *inhomogeneous* circular plates with exactly the same *complete* vibration spectrum as a standard clamped circular *homogeneous* thin plate. This is in itself an interesting phenomenon. The explicit formulae also provide a benchmark against which various numerical schemes for analysing inhomogeneous plate vibrations might be checked.

It is worth noting that the first (theory) section of this paper remains valid for any one-to-one complex analytic function. Thus one may, in general, map a (u, v) region of a standard homogeneous clamped plate into the resulting (x, y) region to obtain a differently shaped inhomogeneous plate (the condition (3) with $0 < \upsilon < 1/2$ must hold) with the same vibration frequency spectrum.

Acknowledgment

The author extends thanks to the staff in the Department of Mathematics at the University of Queensland for their hospitality during his stay on an Outside Studies Programme when this work was carried out.

References

Gottlieb, H. P. W., 1988, "Density Distribution for Isospectral Circular Membranes," *SIAM Journal on Applied Mathematics*, Vol. 48, pp. 948–951.

Lang, K.-W., and Nemat-Nasser, S., 1978, "Vibration and Stability of Rectangular Strip-Plates," *Journal of Sound and Vibration*, Vol. 61, pp. 9–24.

Meirovitch, L., 1967, *Analytical Methods in Vibrations*, Macmillan, London, Chapter 5.

Saff, E. B., and Snider, A. D., 1976, *Fundamentals of Complex Analysis for Mathematics, Science and Engineering*, Prentice-Hall, Englewood Cliffs, N.J., Chapter 7.

Chi-Sing Man
Department of Mathematics,
University of Kentucky,
Lexington, KY 40506

Mehran Jahed

Stephen J. Lai-Fook[1]
Mem. ASME

Pramode K. Bhagat[2]

Biomedical Engineering Center,
University of Kentucky,
Lexington, KY 40506

Effect of Pleural Membrane on the Propagation of Rayleigh-Type Surface Waves in Inflated Lungs

We model the lung parenchyma as an elastic half-space and the pleura as a taut elastic membrane in smooth or in welded contact with the half-space. In each instance we deduce that the presence of a sufficiently high surface tension T *in the pleural membrane will lead to the existence of a cutoff frequency* f_0 *for the Rayleigh-type surface waves, and we derive an equation which gives* T *in terms of* f_0 *and parameters that characterize the layered medium. We performed experiments on four inflated horse lungs at transpulmonary pressures of 5, 10, and 15* cmH_2O*. A comparison of the experimental results and the theoretical predictions provides an empirical test to the validity of the modeling.*

1 Introduction

In a recent study, Jahed et al. (1989) measured speeds of stress waves propagated along the surface of inflated sheep lungs. They observed that signals were transmitted by two waves, which they called the "fast" and the "slow" wave, respectively. Both waves, however, were much slower than those reported in several other studies (Rice, 1983; Kraman, 1983; Yen et al., 1986) on stress waves in lungs, although Butler et al. (1987) did measure a wave speed compatible with that of the "fast" wave. Since the transmitted signals in the experiments of Jahed et al. had frequencies substantially lower than those in the studies that reported higher wave speeds, they conjectured that "[t]he frequency content of the source may be an important determinant of the type of wave transmitted and is probably responsible for the difference in wave velocities measured."

An inflated lung is covered by the taut pleural membrane. This membrane and the lung parenchyma form a layered medium. It is well known that surface waves in layered elastic media are dispersive; moreover, a mode of propagation may have cutoff frequencies beyond which transmission by that mode is no longer possible. In this paper we shall treat the inflated lung as a layered elastic medium, study the propagation of Rayleigh-type waves along its surface, and explore whether the cutoff of Rayleigh-type waves could shed new light on the question about the speeds of stress waves in lungs. The Rayleigh-type surface waves included in this study are straight-crested, and have real frequencies but possibly complex phase velocities. Since the conditions of contact between the pleura and the lung parenchyma could play an essential role, we shall examine two cases, namely those of smooth contact and welded contact. It turns out that in both instances the presence of a sufficiently high tension T in the pleural membrane will lead to the existence of a cutoff frequency f_0; no Rayleigh-type wave that has a frequency $f>f_0$ and has a real phase velocity c may propagate along the surface of the inflated lung; at cutoff, $c=c_s$, where c_s is the velocity of shear waves in the lung parenchyma. For both models of smooth contact and welded contact, we derive an equation which gives T in terms of f_0 and parameters (i.e., the densities and elastic constants) that characterize the layered medium. Measurement of f_0 will deliver an estimate of T if the values of the other parameters are known. Rayleigh-type surface waves are thus potentially useful as a means for the nondestructive evaluation of the tension T in the pleural membrane.

Cutoff of a wave whose speed roughly approximated c_s was indeed observed in the experiments that we performed on four inflated horse lungs at transpulmonary pressures (Ptp) of 5, 10, and 15 cmH_2O. Assuming that this wave was the Rayleigh-type surface wave under study, we used the measured values of f_0 to predict the membrane tension T at Ptp of 5 and 15 cmH_2O, both for the conditions of smooth contact and for those of welded contact. We stripped the pleura from three horse lungs and used indentation tests to measure the membrane tension at those transpulmonary pressures. A comparison of the results of these measurements and the predicted values of T provides an empirical test to the validity of our modeling.

2 Theory

We consider small-amplitude stress waves superimposed on

[1]To whom correspondence should be addressed: Biomedical Engineering, Wenner-Gren Research Laboratory, University of Kentucky, Lexington, KY 40506-0070.

[2]Presently at the Materials Laboratory, Wright-Patterson Air Force Base, Wright-Patterson, OH 45433.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscripts received by the ASME Applied Mechanics Division, Oct. 4, 1989; final revision, June 5, 1990.

an inflated lung at a given transpulmonary pressure. We model the lung parenchyma as an elastic half-space and the pleura as a taut elastic membrane. We choose a Cartesian coordinate system under which the material points of the pleura and of the lung parenchyma have in their reference configuration coordinates $(x, y, 0)$ and (x, y, z) with $z \geq 0$, respectively. We assume that both the elastic surface and elastic half-space are homogeneous and are isotropic at their reference configuration. This assumption of homogeneity and isotropy for the elastic surface entails that its residual stress is a constant surface tension.

Remark 1. Justification of these idealizations will ultimately come from the corroboration between theoretical predictions and experiments. While inflated lungs respond inelastically to slower deformations (see, for example, Hildebrandt, 1969), it may still be a good approximation to treat superimposed small-amplitude stress waves of sufficiently high frequencies as elastic. In experiments we measured stress waves with wavelengths of 4–13 cm which are much longer than the diameter (~ 100 μm) of alveoli, the unit cells comprising the lung parenchyma. For the scale of deformation pertaining to such waves, past experience suggests that we may indeed treat the lung parenchyma as homogeneous and isotropic (Ardila et al., 1974).

In our model the small displacements $U = (u, v, w)$ superimposed on the reference configuration of the pleura obey the following constitutive equation (cf., Gurtin and Murdoch (1975), equation (8.7)):

$$S = Tl + T\nabla U + \alpha(\mathrm{tr}E)l + 2\beta l E, \tag{1}$$

$$l = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \nabla U = \begin{bmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \\ \partial w/\partial x & \partial w/\partial y \end{bmatrix},$$

$$E = \begin{bmatrix} \partial u/\partial x & (\partial u/\partial y + \partial v/\partial x)/2 \\ (\partial u/\partial y + \partial v/\partial x)/2 & \partial v/\partial y \end{bmatrix}; \tag{2}$$

here S is the Piola-Kirchhoff surface stress; T is the constant surface tension; E is the infinitesimal surface strain; α and β are elastic constants. In referential coordinates the equations of motion that govern (u, v, w) are

$$(\alpha + 2\beta + T)\frac{\partial^2 u}{\partial x^2} + (\alpha + \beta)\frac{\partial^2 v}{\partial x \partial y} + (\beta + T)\frac{\partial^2 u}{\partial y^2} + b_1 = \sigma\frac{\partial^2 u}{\partial t^2},$$

$$(\beta + T)\frac{\partial^2 v}{\partial x^2} + (\alpha + \beta)\frac{\partial^2 u}{\partial x \partial y} + (\alpha + 2\beta + T)\frac{\partial^2 v}{\partial y^2} + b_2 = \sigma\frac{\partial^2 v}{\partial t^2},$$

$$T\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right) + b_3 = \sigma\frac{\partial^2 w}{\partial t^2}. \tag{3}$$

Here, $b = (b_1, b_2, b_3)$ is the force per unit area acted upon the membrane by its environment, and σ is the mass per unit area of the membrane in the reference configuration. It follows from our assumptions about the lung parenchyma that the Piola-Kirchhoff stress $\mathbf{S}$ in the reference configuration is related to the infinitesimal strain $\mathbf{E}$ through the familiar constitutive equation

$$\mathbf{S} = \lambda(\mathrm{tr}\ \mathbf{E})\mathbf{I} + 2\mu\mathbf{E}, \tag{4}$$

where λ and μ are the Lamé constants, and $\mathbf{I}$ is the identity tensor. In referential coordinates the equation of motion that governs the infinitesimal displacements $\mathbf{u} = (u, v, w)$ of the lung parenchyma is

$$\mathrm{Div}\ \mathbf{S} + \mathbf{b} = \rho\ \partial^2\mathbf{u}/\partial t^2, \tag{5}$$

where ρ is the density and $\mathbf{b}$ the body force per unit volume at the reference configuration.

The phenomena that we shall study result from the interaction between the taut elastic membrane and the elastic half-space; the conditions of contact between the two are crucial. In this paper we consider two simple instances, namely, smooth contact and welded contact. We ignore gravitation and assume that the membrane is subjected to no external forces other than that acted upon it by the elastic half-space. The conditions of smooth contact are:

$$\text{at } z = 0, \quad w = w, \quad \mathbf{S}_{33} = b_3,\ \mathbf{S}_{13} = \mathbf{S}_{23} = 0; \tag{6}$$

those of welded contact are:

$$\text{at } z = 0, \quad \mathbf{u} = U, \quad \mathbf{S}\ \mathbf{e}_3 = b. \tag{7}$$

We may regard equation (6) or (7) as boundary conditions to be imposed on $\mathbf{u}$ and $\mathbf{S}$, which should satisfy equations (4) and (5); on the other hand, U and b must observe equation (3).

Henceforth, we assume that the body force $\mathbf{b} = \mathbf{0}$. We study the possibility of free Rayleigh-type waves propagating along the surface of an inflated lung. It has been shown (see, for example, Achenbach and Epstein, 1967) that displacements given by the real part of

$$u = (Ae^{-kqz} - sBe^{-ksz})e^{i(\omega t + kx)},$$

$$v = 0,$$

$$w = (iqAe^{-kqz} - iBe^{-ksz})e^{i(\omega t + kx)}, \tag{8}$$

satisfy the equation of motion (5), provided that the parameters q and s satisfy

$$q^2 = 1 - \rho c^2/(\lambda + 2\mu), \quad s^2 = 1 - \rho c^2/\mu; \tag{9}$$

here, A and B are arbitrary complex constants; ω, which we take to be real, is the angular frequency; k, which we allow to be complex, is the wave number; $c = \omega/k$ is the phase velocity. Without loss of generality, we consider only waves that propagate in the negative x-direction; hence we impose the condition

$$\mathrm{Re}\ k > 0. \tag{10}$$

Equation (8) will not describe a Rayleigh-type surface wave unless

$$\mathrm{Re}\{k[1 - \rho c^2/(\lambda + 2\mu)]^{1/2}\} > 0,$$
$$\mathrm{Re}[k(1 - \rho c^2/\mu)^{1/2}] > 0, \quad \text{and } \mathrm{Im}\ k \leq 0. \tag{11}$$

The first two conditions in (11) guarantee that $\mathbf{u} \to \mathbf{0}$ as $z \to \infty$; the last condition says that the amplitude does not grow as the wave propagates.

Finally, the boundary conditions (6) or (7) must be observed. We investigate the cases of smooth contact and welded contact in turn. In what follows, when the variable $W \neq 0$ ranges over the complex plane, $W^{1/2}$ denotes the doubled-valued square root function. When W is real, $\sqrt{W}$ stands for the positive square root.

(a) Smooth Contact. Equation $(3)_3$ and the conditions that $w = w$ and $\mathbf{S}_{33} = b_3$ at $z = 0$ lead to the equation

$$(-k^2qT - k(1 + s^2)\mu + q\sigma\omega^2)A + (k^2T + 2ks\mu - \sigma\omega^2)B = 0. \tag{12}$$

From equations (4), (8), and the condition $\mathbf{S}_{13}(x, y, 0, t) = 0$, we obtain the equation

$$-2qA + (1 + s^2)B = 0. \tag{13}$$

In order that equations (12) and (13) have a nontrivial solution for A and B, we deduce that the phase velocity c of a Rayleigh-type surface wave must satisfy the dispersion equation

$$(2 - \rho c^2/\mu)^2 - 4[1 - \rho c^2/(\lambda + 2\mu)]^{1/2}(1 - \rho c^2/\mu)^{1/2}$$
$$- (\rho\omega c/\mu^2)[1 - \rho c^2/(\lambda + 2\mu)]^{1/2}(T - \sigma c^2) = 0. \tag{14}$$

Let $Z = \rho c^2/\mu$, and let $f = \omega/2\pi$ be the frequency. Let us recast equation (14) as a relation between f and Z, namely:

$$f = \frac{(2 - Z)^2 - 4(1 - Z)^{1/2}[1 - \mu Z/(\lambda + 2\mu)]^{1/2}}{2\pi[1 - \mu Z/(\lambda + 2\mu)]^{1/2}(T/\mu - \sigma Z/\rho)Z} \cdot \left(\frac{\mu Z}{\rho}\right)^{1/2}. \tag{15}$$

Henceforth, we assume that

$$T/\mu > \sigma/\rho, \text{ and } \lambda + 2\mu > \mu; \qquad (16)$$

the second inequality will follow if both the shear modulus and the bulk modulus of the elastic half-space are positive. Let

$$R(Z) = (2-Z)^2 - 4(1-Z)^{1/2}[1 - \mu Z/(\lambda+2\mu)]^{1/2}. \qquad (17)$$

It is well known (see, for example, Achenbach, 1973, p. 191) that R has a zero at $Z = Z_R \equiv \rho c_R^2/\mu < 1$, where c_R is the Rayleigh wave velocity pertaining to the elastic half-space. Moreover, $R(Z) > 0$ for real values of Z in the range $Z_R < Z < 1$, if $(\bullet)^{1/2}$ is replaced by $\sqrt{(\bullet)}$ in equation (17). Restricting our attention to this range of Z and taking the positive root of all the square roots in equation (15), we obtain a positive real value of f for each value of $c = \sqrt{(\mu Z/\rho)}$. Each such pair of f and c satisfies equation (14) and the conditions (10) and (11). Thence, such pairs of f and c define the dispersion curve of a branch of free, straight-crested, Rayleigh-type waves (cf., equation (8)) that may propagate along the surface of the inflated lung.

As $Z = \rho c^2/\mu \to 1$, we see from equation (15) that f approaches the value

$$f_0 = \{2\pi(T/\mu - \sigma/\rho)\sqrt{[1 - \mu/(\lambda+2\mu)]}\}^{-1}\sqrt{(\mu/\rho)}; \qquad (18)$$

We call f_0 the cutoff frequency, because condition $(11)_2$, namely $\mathrm{Re}[k(1-Z)^{1/2}] > 0$, cannot be satisfied for any real $Z \geq 1$. As $f \to f_0$, the phase velocity $c \to \sqrt{(\mu/\rho)}$, which delivers the shear modulus of the half-space, if the density ρ is known. Furthermore, if the parameters λ, μ, σ, and ρ are given, measurement of f_0 will give an estimate of the surface tension T.

It will be interesting to ascertain whether the aforementioned branch exhausts all the possible Rayleigh-type surface waves of the form (8). We can easily show that there are no more waves with real phase velocity. It suffices to examine the range $0 < Z < Z_R$. Conditions (10) and (11) dictate that we must take positive square roots in equation (15). It follows that $R(Z) < 0$ for $0 < Z < Z_R$. Hence, equation (15) does not give a positive f.

The possibility of waves with complex c ($\mathrm{Im}\ c \neq 0$) remains to be investigated. Pending a more thorough mathematical study, we examined this problem numerically for a specific choice of material parameters typical of an inflated lung.

Example 1. Let $T/\mu = 1/2$ cm, $\mu(\lambda+2\mu) = 1/6$, $\mu = 4{,}000$ dynes/cm^2, $\sigma = 3 \times 10^{-3}$ gm/cm^2, and $\rho = 0.2$ gm/cm^3. These values are typical of those of dog lungs at Ptp of 5 cmH_2O. Figure 1 depicts the dispersion curve obtained by the procedure described above. The cutoff frequency f_0 is 51 Hz. We sought complex solutions c of equation (14) for the following frequencies $f = 1$ Hz, 10 Hz, 20 Hz, 30 Hz, ..., (and by 10 Hz increments up to) 200 Hz. We used the double precision IMSL (International Mathematical and Statistical Library) subroutine DZANLY (Muller's method). For each complex solution c thus found, we checked whether it satisfies also the conditions (10) and (11). In all instances considered we found no complex c ($\mathrm{Im}\ c \neq 0$) that satisfies all the conditions which a Rayleigh-type surface wave should observe.

(b) Welded Contact. The case of welded contact with real c was analyzed previously by Murdoch (1976). (Cf., also Tiersten, 1969, for the special instance where $T = 0$.)

From the conditions of welded contact, namely equation (7), we deduce from equations (3), (4), and (8) the following dispersion equation

$$A_0 k^2 + A_1 k + A_2 = 0; \qquad (19)$$

here

$$A_0 = (T/\mu - \sigma Z/\rho)(\gamma/\mu - \sigma Z/\rho)\{1 - (1-Z)^{1/2} \times [1 - \mu Z/(\lambda+2\mu)]^{1/2}\},$$

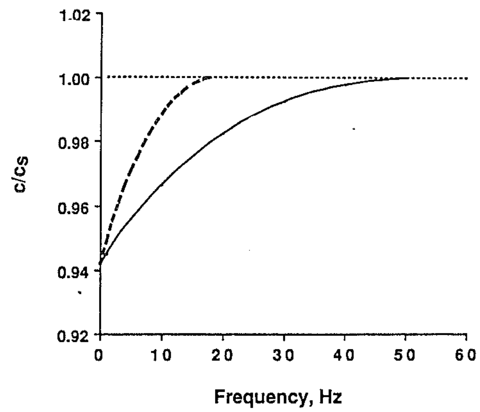


Fig. 1 Dispersion curves pertaining to Example 1 (smooth contact, full line) and Example 2 (welded contact, dashed line)

$$A_1 = Z\{(\gamma/\mu - \sigma Z/\rho)(1-Z)^{1/2} + (T/\mu - \sigma Z/\rho) \times [1 - \mu Z/(\lambda+2\mu)]^{1/2}\},$$
$$A_2 = -R(Z), \qquad (20)$$

where $\gamma \equiv \alpha + 2\beta + T$ and $R(Z)$ is defined in equation (17). For a displacement of the form $U = (u(x, t), 0, w(x, t))$, we see from equation (1) that $\gamma = (S_{11} - T)/E_{11}$.

In addition to the inequalities (16), henceforth we assume also that

$$\gamma/\mu > \sigma/\rho. \qquad (21)$$

For real values of Z in the range $Z_R < Z < 1$, if we take positive square roots in (20), we have $A_0 > 0$, $A_1 > 0$, and $A_2 < 0$. The quadratic equation (19) delivers one positive real root k, which gives a positive real frequency

$$f = \frac{c}{2\pi} \cdot \frac{[-A_1 + \sqrt{(A_1^2 - 4A_0A_2)}]}{2A_0}, \quad (Z_R < Z < 1) \qquad (22)$$

for each value of $c = \sqrt{(\mu Z/\rho)}$. Such pairs of (f, c) satisfy the conditions (10) and (11), so they define the dispersion curve of a branch of Rayleigh-type surface waves.

For real values of Z in the range $0 < Z < Z_R$, all the coefficients A_0, A_1, and A_2 are positive, so equation (19) has no solution k that is both real and positive. Hence, pairs of (f, c) given by equation (22) deliver all Rayleigh-type surface waves of the form (8) that have a real phase velocity.

As $Z = \rho c^2/\mu \to 1$, equation (19) approaches the limiting form

$$(T/\mu - \sigma/\rho)(\gamma/\mu - \sigma/\rho)k^2 + \{(T/\mu - \sigma/\rho)\sqrt{[1 - \mu/(\lambda+2\mu)]}\}k - 1 = 0. \qquad (23)$$

Hence, the cutoff frequency f_0 is given by

$$f_0 = (k_0/2\pi)\sqrt{(\mu/\rho)}, \qquad (24)$$

where k_0 is the positive root of equation (23). Since the k^2 term in equation (23) is positive, we have

$$T/\mu < \sigma/\rho + \{k_0\sqrt{[1 - \mu/(\lambda+2\mu)]}\}^{-1}. \qquad (25)$$

Given a measured value of k_0 or f_0, we see from equation (18) that the right-hand side of (25) is nothing but the value of T/μ predicted by the model of smooth contact. Thus, for a measured value of f_0 and a given set of parameters μ, λ, σ, and ρ, the model of welded contact gives a lower value of T than that predicted by the model of smooth contact.

Following what we did for the case of smooth contact, we investigated numerically the possibility of Rayleigh-type waves with complex c.

Example 2. We took the same numerical values for T/μ, $\mu/(\lambda+2\mu)$, μ, σ, and ρ as in Example 1. As an example, the dispersion curve pertaining to Rayleigh-type surface waves with real c for the case $\gamma = 8000$ dynes/cm is shown in Fig. 1. For $\gamma = 2500$, 5000, 7500, and 10,000 dyne/cm, we sought complex solutions c of equation (19) for the following frequencies: $f = 1$ Hz, 10 Hz, 20 Hz, 30 Hz, (and by 10 Hz increments up to) 200 Hz. Again, we used the IMSL subroutine DZANLY (Muller's method). In all instances considered we found no complex c (Im $c \neq 0$) that satisfies all the conditions, namely equations (10), (11), and (19), which a Rayleigh-type surface wave should observe.

Remark 2. Examples 1, 2, and our analysis about f_0 suggest the following assertion: For both the models of smooth and welded contact, there is a range of frequencies immediately above f_0 for which no Rayleigh-type wave of the form (8) could propagate along the surface of an inflated lung. As we shall see below, our experiments on inflated horse lungs did indicate the existence of a frequency above which transmission of Rayleigh-type surface waves was not observed.

Remark 3. For both cases of smooth and welded contact it is the presence of a sufficiently high surface tension in the pleural membrane, viz., $T > \sigma\mu/\rho$, that accounts for the existence of a finite cutoff frequency f_0. For either case, if we have $0 \leq T \leq \sigma\mu/\rho$ instead of $(16)_1$, an elementary analysis of equation (15) or equation (19) reveals that there is, for each frequency f, a Rayleigh-type surface wave with real phase velocity, $c < \sqrt{(\mu/\rho)}$.

Remark 4. The wave speeds that we measured in our experiments were group velocities. The group velocity c_g is related to the phase velocity c and the frequency f by the formula

$$c/c_g = 1 - (f/c)\ (dc/df). \tag{26}$$

For the case of smooth contact, by equation (15), $df/dZ \to +\infty$ as $Z \to 1^-$; in other words, $dc/df \to 0^+$ as $f \to f_0$. For the case of welded contact, by using implicit differentiation, we deduce from equations (19) and (20) that $dk/dZ \to +\infty$ as $Z \to 1^-$; hence, we again have $dc/df \to 0^+$ as $f \to f_0$. In either case, since $c \to c_s = \sqrt{(\mu/\rho)}$ as $f \to f_0$, equation (26) dictates that c_g also approaches c_s as $f \to f_0$; thus, measurement of c_g just before cutoff will deliver c_s. A glance at Fig. 1 reveals that the dispersion will typically be small. Indeed, for Examples 1 and 2, numerical calculations show that $c_g \geq c$, but the difference is within a few percent of c. Such a difference is well within the margin of error in our experiments.

3 Experimental Method

To see how well our modeling would fare in practice, we performed experiments on horse lungs, which were chosen for their size. Compared with the lungs of smaller animals, horse lungs would allow a greater distance between the source and the receivers. Since the Rayleigh-type waves that we looked for should suffer significantly less geometrical attenuation than the dilatational and the shear waves, larger lungs would work better for our present purpose.

We obtained isolated lungs from horses *post mortem*. The procedure used by Jahed et al. (1989) was followed. We injected isopreterenol (2 mg/10 ml saline) into the bronchus of each lung to reduce bronchoconstriction and to minimize gas trapping. After degassing, the lung was inflated to a transpulmonary pressure (Ptp) of 40 cmH_2O and then deflated to a test Ptp of 15 cmH_2O. The group velocities of stress waves propagated along the surface of the lung were measured at test Ptp values of 15, 10 and 5 cmH_2O by the procedure described below. After the stress wave measurements, the collapsed lung was weighed and displaced in water to determine its residual volume. The air volume withdrawn between the test Ptp values was noted. Lung density was calculated by dividing lung mass by the total lung volume (air plus tissue volume) at each Ptp.

The procedure used to distort the lung surface to produce stress waves was as follows: The lung was oriented with its dorsal surface superior and horizontal. An electromagnetic vibrator (Ling Dynamic Systems, Inc.; Model No. 102A) was used to distort the lung surface. We used two types of input distortions. One was a 3-cycle sinusoidal displacement. The amplitude of distortion was ~ 1 mm. The frequency of distortion was varied between 30 and 120 Hz in 10 Hz increments. The other input distortion was an impulse displacement (frequency content, 0–200 Hz). In one lung at a Ptp of 5 cmH_2O, we used both methods of distortion and compared the velocities of the resulting stress waves as a function of frequency.

The input signal was generated by a computer (IBM PC-AT), amplified and sent to the vibrator via an A/D converter. The input distortion into the lung surface was measured by a linear variable differential transducer (LVDT, Trans-Tek, Model No. 240, 3 db response: 0–300 Hz). The core (1.5 mm radius) of the LVDT was connected rigidly to the shaft of the vibrator and used to distort the lung. The magnetic coil of the LVDT surrounding the core was fixed to the casing of the vibrator.

The stress waves propagated along the surface of the inflated lung were measured by a microphone (Realistic condenser-type; 0.75 cm diameter; ± 6 db response: 20–15000 Hz) located at 35–40 cm from the vibrator. The vibrator and the microphone were situated along the dorsal surface of the lung. The length of the dorsal surface was ~ 50 cm and the height of the lung was ~ 20 cm. The microphone was embedded ~ 1 mm into the lung parenchyma. The signals received by the microphone and recorded by the LVDT were amplified, digitized, and stored on the computer. The sampling rate was 2 KHz and the record length was 0.5–1 second.

We used an indentation test to measure the shear modulus of lung parenchyma in three horse lungs at 5, 10, and 15 cmH_2O Ptp (cf., Lai-Fook et al., 1976). In brief, at a fixed Ptp a rod (2.5-cm diameter) was pressed into the surface of the lung in 1-mm increments. The force of indentation was measured by a load cell. We allowed one minute to elapse between two consecutive displacement increments. The force at one minute after each increment was used to construct the force-displacement curve. The slope of the force-displacement curve was used to estimate the shear modulus of the lung parenchyma.

The pleural membrane tension at 5 and 15 cmH_2O Ptp, respectively, was measured in three horse lungs by the following procedures (cf., Hajji et al., 1979). In brief, the rim of a tube (1.85-cm radius, thickness 0.2 cm) was coated with a dye. Concentric circles were marked on the dorsal surface of the lung at 15 and 5 cmH_2O Ptp by using the tube. The pleural membrane surrounding the marked area was stripped from the lung. The membrane was stretched over the rim of the tube so that the circle marked at 5 cmH_2O Ptp coincided with the rim; it was then fixed in position by a rubber band. A rod (0.35-cm radius) was used to indent the center of the stretched membrane by 0.5-mm increments. The force of indentation was measured by a load cell. The indentation test was repeated after stretching the membrane to the circle marked at 15 cmH_2O Ptp. The slope of the force-displacement curve was used to estimate the pleural membrane tension. The thickness of the pleural membrane was estimated by weighing the membrane enclosed by a marked circle and by dividing its volume by the surface area. The tissue density was assumed to be 1 g/cm^3.

4 Analysis and Results

The input sinusoidal signal of three cycles had a bandwidth of ± 10 percent of its center frequency. Therefore, the microphone signal was filtered at the input center frequency with the same bandwidth. This was repeated for each input center frequency. Figure 2 shows an example of the filtered signals received by the microphone 35 cm from the vibrator at 30, 40, 50, 60, 70, and 80 Hz input frequency. The transit time at each frequency was the difference between the peak of the envelope of the filtered microphone signal and that of the input signal.

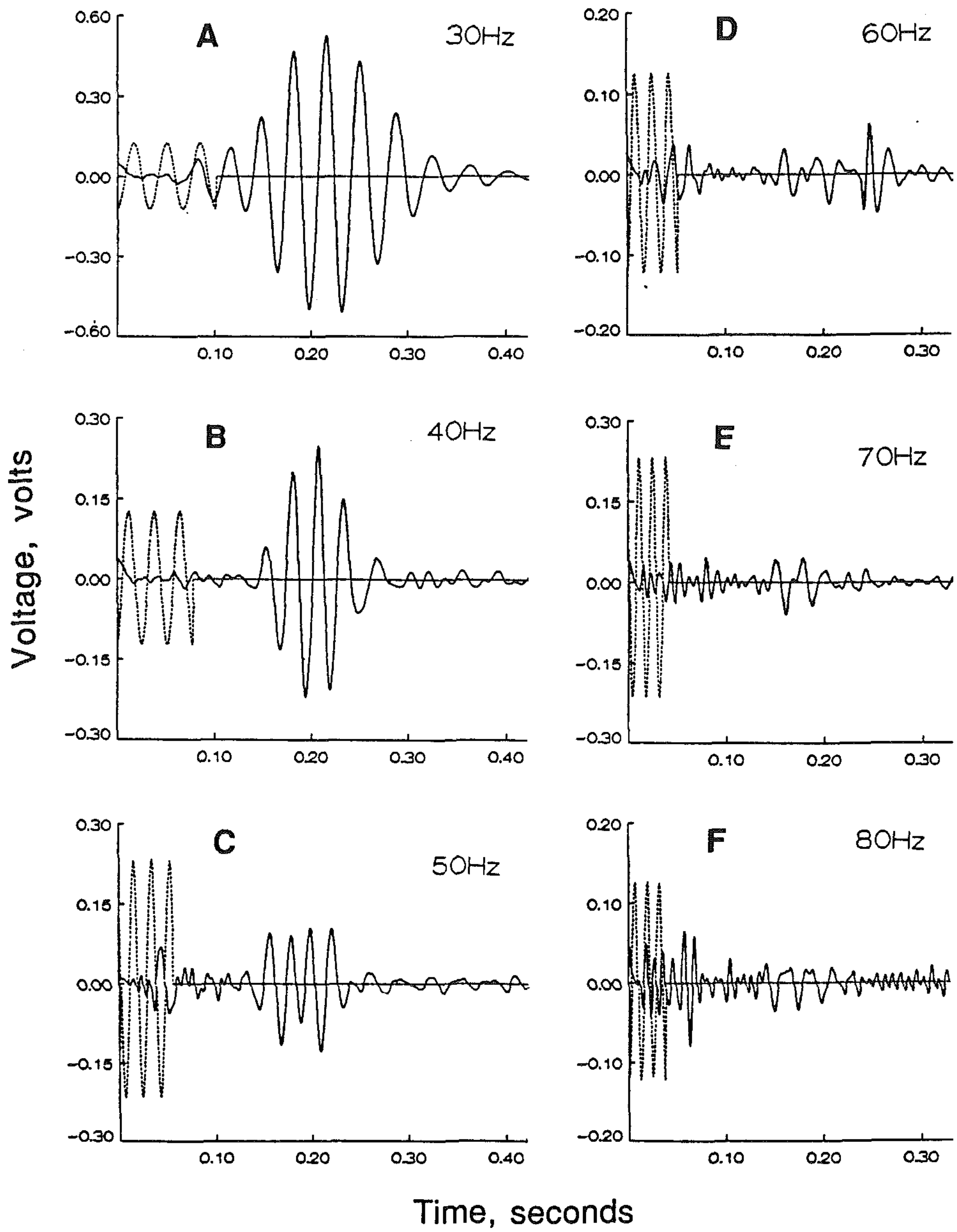


Fig. 2 Microphone signals (full line) measured 35 cm from input sinusoidal distortion for sample H1 at 5 cmH_2O Ptp. Dotted line is the 3-cycle input signal. Frames A through F show responses to input signal frequencies from 30 Hz through 80 Hz in 10 Hz increments. The responses have been filtered to within a bandwidth of ±10 percent of the input center frequency and have undergone the same factor of amplification.

The group velocity of the wave packet was calculated by dividing the distance between the vibrator and the microphone by the transit time.

Above an input frequency of 50 Hz, a definitive interpretation of the transit time of the propagated signal at the center frequency was sometimes not possible (for example, sample H1, at 5 cmH_2O Ptp; see Frame D, Fig. 2). This was due to the presence of signals propagated at frequencies lower than the input frequency or signals of waves of velocities greater than the Rayleigh-type surface waves. The presence of these waves often prevented a clear-cut determination of the cutoff frequency by looking at the filtered signals in the time domain. When this occurred we used a power spectrum analysis of the unfiltered signal to determine whether there was any signal transmitted at the input frequency. Figure 3 shows the normalized power spectrum of the microphone signal given in Frame E of Fig. 2; the input frequency of the signal was 70 Hz. In the transmitted signal there was no significant power in the frequencies above 50 Hz. Thus, we decided that the Rayleigh-type surface wave was not transmitted at 70 Hz. Since we found by the same method that there was transmission at 50 Hz and no transmission at 60 Hz, we took the cutoff frequency as 55 Hz for sample H1 at Ptp of 5 cmH_2O.

To obtain the group velocity of the propagated signal as a function of frequency from the response to an impulse distortion, we filtered the propagated signal at center frequencies between 30 and 150 Hz in 20 Hz increments with a bandwidth of ±10 percent center frequency.

Table 1 summarizes the group velocities versus frequency measured in four horse lungs at Ptp of 5, 10, and 15 cmH_2O. In the experiment H1 at 5 cmH_2O Ptp, similar results were obtained from the method of single-frequency distortions and that of impulse distortion. The cutoff frequency averaged 59 ± 5 (SD) Hz, 70 ± 9 Hz, 83 ± 20 Hz at Ptp of 5, 10, 15 cmH_2O, respectively. A linear regression analysis indicated a significant increase in cutoff frequency with Ptp: $f_0 = 2.38$ Ptp + 46.7 ($r^2 = 0.422$, $P < 0.05$); here, f_0 is in Hz and Ptp is in cmH_2O.

From theoretical considerations, the group velocity of Rayleigh-type surface waves at the cutoff frequency is equal to the shear wave velocity of the lung parenchyma. The shear wave velocity c_s predicted in this way averaged 233 ± 22 (SD) cm/s, 300 ± 59 cm/s, 425 ± 18 cm/s at Ptp of 5, 10, 15 cmH_2O, respectively.

The shear modulus of the lung parenchyma measured by indentation tests in three inflated horse lungs averaged 5.3 ± 1.6 (SD), 8.7 ± 2.7, 12.9 ± 3.8 cmH_2O at Ptp of 5, 10, and 15 cmH_2O, respectively. To calculate the shear modulus from the force-displacement data, we used a Poisson ratio of 0.4 in the elasticity solution of the indentation of an elastic half-space by a rigid rod (cf., Lai-Fook et al., 1976). Pleural membrane tension measured by indentation tests on stripped pleura from three horse lungs averaged 4530 ± 960 (SD) dynes/cm and $11{,}100 \pm 4{,}370$ dynes/cm at Ptp of 5 and 15 cmH_2O, respectively. The thickness of the pleural membrane averaged 81 ± 15 (SD) μm at 5 cmH_2O Ptp and 64 ± 12 μm at 15 cmH_2O Ptp. The density of the horse lungs averaged 0.165 ± 0.013 (SD), 0.129 ± 0.010, and 0.117 ± 0.008 g/ml at Ptp of 5, 10, 15 cmH_2O.

By substituting in the formula $\mu = \rho c_s^2$ these values of the density ρ and the corresponding values of c_s predicted from the velocity of the Rayleigh-type surface wave at cutoff, the shear modulus μ of the lung parenchyma was calculated to be 8.9, 11.6, and 21.1 cmH_2O, respectively. These values for the shear modulus are 30–70 percent greater than the values delivered by the quasi-static indentation tests.

Equations (18) and (24) can be used to give a prediction of the tension T in the pleural membrane for the models of smooth contact and welded contact, respectively. To calculate T, we need the values of f_0, ρ, σ, μ, and λ for the instance of smooth contact; we require in addition the value of γ for the case of welded contact. The measured values of ρ and σ, and the values of f_0 and μ as obtained from the cutoff of the Rayleigh-type surface wave have already been reported above. We assumed a value of four Ptp (cf., Lai-Fook et al., 1976) for λ. For the model of smooth contact, the predicted values of T were 6900 and 21,000 dynes/cm at 5 and 15 cmH_2O Ptp, respectively. These values are 50–90 percent greater than the values obtained from the indentation tests on ablated strips of pleura. The value of γ raises a problem for the model of welded contact. While there are experimental studies on the elastic properties of canine pleura, we do not know of any work on the equine pleural membrane that we can call upon. Here, we are content to tabulate the predicted values of T as γ range from T to $6T$. See Table 2.

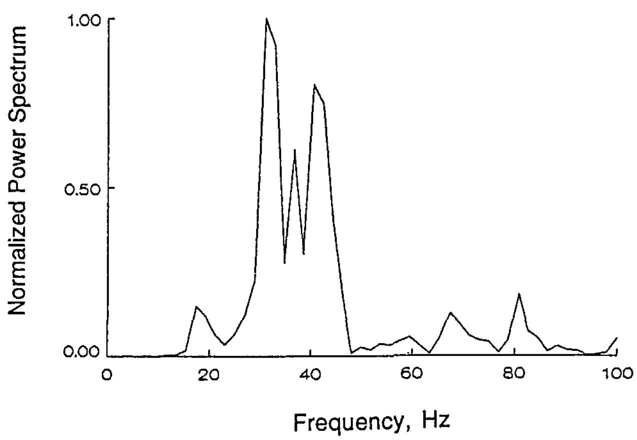


Fig. 3 Normalized power spectrum of unfiltered microphone signal from 70 Hz input distortion for sample H1 at 5 cmH_2O Ptp (cf., Frame E of Fig. 2)

5 Discussion

The first question one might ask is: How did we know that the particular wave which we took note of was the Rayleigh-type surface wave at issue? Our judgement was based on (i) the existence of a cutoff frequency, (ii) the wavelength and speed of the wave just before cutoff, and (iii) the agreement between the predicted and measured values of the membrane tension. We have already presented the empirical evidence that indicates the existence of a cutoff frequency for the wave in question (see Fig. 2 and the related discussion). Items (ii) and (iii), however, call for further comments.

According to our theory (see Remark 4), both the phase velocity c and the group velocity c_g of the Rayleigh-type surface wave should approach $c_s = \sqrt{(\mu/\rho)}$ as $f \to f_0$; thence, by measuring c_g just before cutoff, we obtain an estimate of the shear modulus of the lung parenchyma if ρ is ascertained by an independent measurement. The values of c_s and μ thus delivered were 15–30 percent and 30–70 percent higher than the corresponding values we predicted from the quasi-static indentation tests, respectively. We believe that the wave whose speed we measured was not the dilatational wave, because signals were transmitted by another wave with a higher velocity (see Fig. 2, Frames E and F). In fact, the discrepancy between the predicted values of μ can be understood if we examine the results of the quasi-static tests more closely. Figure 4 shows a typical force versus time curve when the force required to maintain a fixed indentation was recorded. It is clear that significant stress relaxation occurred during the quasi-static indentation tests. Since the pleural membrane shows little stress relaxation in the indentation tests to measure membrane tension, the stress relaxation observed here must represent a parenchymal effect. This conclusion is not at all surprising, since the lung parenchyma is known to exhibit viscoelastic behavior (see, for example, Hildebrandt, 1969). The quasi-static tests, by the procedure they were performed, should deliver a value of μ much lower than that which pertains to the instantaneous

Table 1 Group velocity versus frequency for Rayleigh-type waves propagated along lung surface

Freq	Ptp = 5 cmH_2O				Ptp = 10 cmH_2O				Ptp = 15 cmH_2O			
(Hz)	H1^	H2	H3	H4	H1	H2	H3	H4	H1	H2	H3	H4
30	212+	148	227	239	320	222	314	346	396	355	640	410
40	205	255	199	----	----	228	333	----	----	370	572	----
50	208	255	221	250	332	238	345	361	414	285	422	436
60	*	253	*	----	----	238	307	----	----	342	*	----
70		*		*	*	*	264	365	429	349		447
80							*	----	----	403		----
90								*	*	408		466
100										*		----
110												*

Ptp, transpulmonary pressure; ^Horse number; +velocity in cm/s
----No data from pulse input distortion; signal was filtered at 20 Hz increments.
*Lowest frequency at which no transmission occurred.

Table 2 Pleural membrane tension for γ between T and $6T$

	$\gamma=T$	$\gamma=2T$	$\gamma=3T$	$\gamma=4T$	$\gamma=5T$	$\gamma=6T$
Ptp = 5^	4.1*	3.3	2.9	2.6	2.4	2.2
Ptp = 15	12.3	9.9	8.6	7.5	7.1	6.6

*Tension, 10^3 dynes/cm, ^Ptp, transpulmonary pressure, cmH_2O

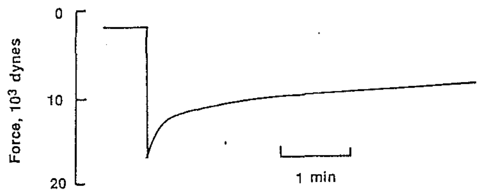


Fig. 4 Time dependency of force required to maintain a 1-mm step indentation of lung surface

elastic response of the lung parenchyma. The wave propagation experiments, on the other hand, might give a value of μ that closely approximates the value pertaining to the instantaneous elastic response. While the foregoing interpretation of the experimental results sounds plausible, further studies are required to determine whether the viscoelasticity of the lung parenchyma could indeed provide a valid explanation for the discrepancy in the predicted values of μ.

In our experiments the wave in question had a wavelength of 4-5 cm just before cutoff, while the horse lungs measured approximately 20 cm in height along the wave paths. We deem that the horse lungs were thick enough for the waves we observed to be Rayleigh-type surface waves.

The model of smooth contact gave a prediction of T that was 50–90 percent greater than the values measured by indentation tests on ablated strips of pleura in the present study. Since we are dealing with biological samples, we can say that the agreement is good. The predicted values of T are comparable to the values estimated by Hajji et al. (1979) from indentation tests on inflated horse lungs under similar assumptions concerning smooth contact. However, the values of T measured by Hajji et al. were larger than those measured in the present indentation experiments probably because the lungs used by Hajji et al. were bigger.

Physically, the pleura seems to be rigidly attached to the lung parenchyma; from this standpoint, smooth contact will not be an acceptable model. Should γ fall in the range $T \leq \gamma \leq 6T$ for the values of Ptp in question, the model of welded contact would provide good predictions of T (see Table 2). If the studies of Stamenovic (1984) and Humphrey et al. (1986, 1987) on the canine pleura could be used as a guide, then γ would typically assume a value of $\sim 5T$ in the range of Ptp under consideration, and the model of welded contact would be credible. Nevertheless, we cannot lay claim to a corroboration of the model of welded contact without verification of the elastic constant γ for horse pleura.

In spite of the uncertainties that remain to be clarified, the empirical evidence gathered above indicates that the wave with the cutoff frequency, which we observed in our experiments, was indeed the Rayleigh-type surface wave in question. Moreover, as far as propagation of Rayleigh-type surface wave is concerned, our experimental results strongly suggest that our modeling of the inflated lung as a layered elastic medium has fared well, and the effect of the taut pleural membrane cannot be ignored.

During quiet breathing the pleural tension may contribute as much as 20 percent of the transpulmonary pressure (Hajji et al., 1979). The present study suggests that stress surface waves could be a promising means for the nondestructive evaluation of the tension in the pleural membrane.

If we reexamine the work of Jahed et al. (1989), it seems likely that their "slow" wave is the Rayleigh-type surface wave studied here. Not only does the speed of the "slow" wave match that of the Rayleigh wave, but the frequency content (below 70 Hz) of the transmitted signals is also consistent with the existence of a cutoff frequency f_0 whose value lies in the expected range. The existence of a low cutoff frequency explains why the "slow" wave of Jahed et al. was not observed by other researchers who used sources that excited waves of higher frequencies.

Acknowledgment

This research was supported by Research Grant HL-40362 from the National Heart, Lung, and Blood Institute. The work of C.-S. Man was supported in part by Grant DMS-8911767 from the National Science Foundation.

Address reprint requests to: Stephen J. Lai-Fook, Ph.D., Biomedical Engineering, Wenner-Gren Research Laboratory, University of Kentucky, Lexington, KY, 40506.

References

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

Achenbach, J. D., and Epstein, H. I., 1967, "Dynamic Interaction of a Layer and a Half-Space," *Journal of the Engineering Mechanics Division, ASCE*, Vol. 93, No. EM5, pp. 27–42.

Ardila, R., Horie, T., and Hildebrandt, J., 1974, "Macroscopic Isotropy of Lung Expansion," *Respiratory Physiology*, Vol. 20, pp. 105–115.

Butler, J. P., Lear, J. L., and Drazen, J. M., 1987, "Longitudinal Elastic Wave Propagation in Pulmonary Parenchyma," *Journal of Applied Physiology*, Vol. 62, pp. 1349–1355.

Gurtin, M. E., and Murdoch, A. I., 1975, "A Continuum Theory of Elastic Material Surfaces," *Archive for Rational Mechanics and Analysis*, Vol. 57, pp. 291–323.

Hajji, M. A., Wilson, T. A., and Lai-Fook, S. J., 1979, "Improved Measurements of Shear Modulus and Pleural Membrane Tension of the Lung," *Journal of Applied Physiology*, Vol. 47, pp. 175–181.

Hildebrandt, J., 1969, "Dynamic Properties of Air-Filled Excised Cat Lung Determined by Liquid Plethysmography," *Journal of Applied Physiology*, Vol. 27, pp. 246–250.

Humphrey, J. D., Vawter, D. L., and Vito, R. P., 1986, "Mechanical Behavior of Excised Canine Visceral Pleura," *Annals of Biomedical Engineering*, Vol. 14, pp. 451–466.

Humphrey, J. D., Vawter, D. L., and Vito, R. P., 1987, "Pseudoelasticity of Excised Visceral Pleura," *ASME Journal of Biomechanical Engineering*, Vol. 109, pp. 115–120.

Jahed, M., Lai-Fook, S. J., Bhagat, P. K., and Kraman, S. S., 1989, "Propagation of Stress Waves in Inflated Sheep Lungs," *Journal of Applied Physiology*, Vol. 66, pp. 2675–2680.

Kraman, S. S., 1983, "Speed of Low-Frequency Sound through Lungs of Normal Men," *Journal of Applied Physiology*, Vol. 55, pp. 1862–1867.

Lai-Fook, S. J., Wilson, T. A., Hyatt, R. E., and Rodarte, J. R., 1976, "Elastic Constants of Inflated Lobes of Dog Lungs," *Journal of Applied Physiology*, Vol. 40, pp. 508–513.

Murdoch, A. I., 1976, "The Propagation of Surface Waves in Bodies with Material Boundaries," *J. Mech. Phys. Solids*, Vol. 24, pp. 137–146.

Rice, D. A., 1983, "Sound Speed in Pulmonary Parenchyma," *Journal of Applied Physiology*, Vol. 54, pp. 304–308.

Stamenovic, D., 1984, "Mechanical Properties of Pleural Membrane," *Journal of Applied Physiology*, Vol. 57, pp. 1189–1194.

Tiersten, H. F., 1969, "Elastic Surface Waves Guided by Thin Films," *Journal of Applied Physics*, Vol. 40, pp. 770–789.

Yen, R. T., Fung, Y. C., Ho, H. H., and Butterman, G., 1986, "Speed of Stress Wave Propagation in Lung," *Journal of Applied Physiology*, Vol. 61, pp. 701–705.

The Effect of Plasticity on Resonant Pipe Vibration

E. M. Beaney
Central Electricity Generating Board,
Berkeley, Gloucestershire, U.K.

An understanding of the response of pipework systems to high levels of seismic excitation is required to enable aseismic design methods to be securely based. Theoretical and experimental modeling of simple systems and components demonstrate that plasticity in the pipe wall controls the vibration response level and that, because of an unexpected level of material strain hardening, in pressurized pipes, simple elastic modal and frequency analysis are satisfactory. Given the correct material properties an energy balance approach correctly predicts the steady.

1 Introduction

There is clear evidence, both from experimental work and from systems which have been subjected to earthquakes, that pipework only fails in exceptional circumstances due to seismic excitation. Current design codes appear to be overconservative and their use requires excessive numbers of undesirable seismic restraints. With the eventual aim of producing improved but simple design procedures, a series of experimental tests with related theoretical studies are underway within the Central Electricity Generating Board (CEGB) to investigate the effects of high amplitude vibration excitation on lengths of pipework. The objectives of this work are to:

(*i*) study the performance of a very simple single span system to facilitate a basic understanding,

(*ii*) mathematically model its performance,

(*iii*) extend the system, both experimentally and theoretically, by introducing components such as bends, nozzles and tees, and

(*iv*) develop the model to cover multiple spans.

So far the response of single spans of pipework with and without internal pressure and with some nozzle-like components have been studied.

There are two distinct problems in aseismic design; first the correct prediction of the level of response and second, but equally important, the mode of failure. Although the program of work encompasses both aspects, this paper concentrates on the prediction of response and the overriding effect that plasticity in the pipe wall has on the level of vibration.

2 Experimental Work

Lengths of straight pipe, pinned at each end have been shaken, at or near their fundamental resonant frequencies, using sinusoidal excitation, in the rig shown in Fig. 1. The hydraulic actuators, which are displacement controlled, have a range of ± 50 mm and at frequencies up to 5 Hz a load capacity of 15 KN. Due to hydraulic limitations above 5 Hz the maximum displacement falls linearly with frequency. The pipe can be pressurized with air up to a maximum pressure of 770 bar and can be shaken in either a vertical or horizontal plane. The pressure, in the pressurized pipes, was chosen to give the normal ASMEIII design hoop stress in the pipe wall.

The instrumentation is used to record pressure, actuator load and displacement, acceleration at the center of the pipe span and hoop, and axial strains at three locations in the pipe wall. Provision is made to record up to 16 channels of data at a maximum rate of 100 scans/second, digitize the data using 12-bit conversion, and store it directly onto a disk. Once digitized and stored onto a disk, the data can be converted into engineering units and presented in a reduced form by means of a local microprocessor.

The pipes tested so far have been made of two types of material: first, of a carbon steel ferritic material which, as shown in Fig. 2, exhibits almost perfectly elastic/plastic tensile properties with no measurable hardening with strains of less than about two percent, and second, 316 austenitic stainless steel which shows marked strain hardening. Pipe geometries tested have been in the range diameter: wall thickness ratio from 7 to 68 and lengths from about 3000 to 6000 mm but the majority of the results are based on pipes with an outside diameter of 25.4 mm and a wall thickness of 2.64 mm. The lengths of the pipes were chosen to give fundamental resonant frequencies, in the range of 5-10 Hz.

3 Acceleration Response

The most striking feature of the response of a length of pipe shaken at resonance is the way in which plasticity in the pipe wall limits the level of response. As can be seen in Fig. 3, when the pipe is shaken with small inputs, the low level response rises linearly with input and its level is controlled by the small inherent system damping of typically one percent of critical measured in these tests.

When the input level reaches about 0.2 g yielding starts to occur in the pipe wall. This plasticity absorbs the input energy,

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 18, 1989; final revision, Sept. 18, 1990.

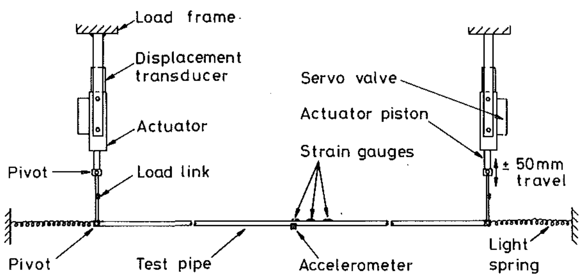


Fig. 1 Experimental test rig

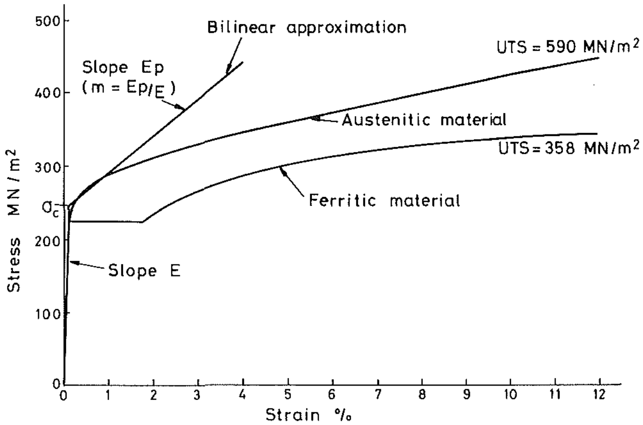


Fig. 2 Material tensile properties

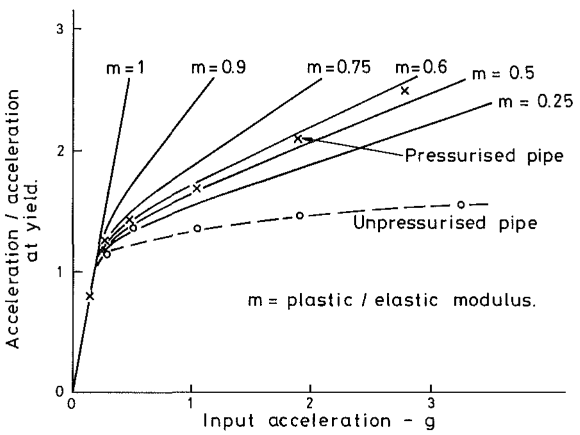


Fig. 3 Acceleration response of ferritic pipe

and hence increases the effective damping of the system so much, that the level of vibration in any of the tests never exceeded twice that at first yield for input acceleration levels in excess of 1 g. Since it is the onset of yielding which controls the response, all the accelerations plotted in Fig. 3 have been normalized by dividing by the acceleration level at first yield. First yield is taken as static yield for the ferritic material and the 0.2 percent proof stress for the austenitic material reduced, to take account of the pressure stress, using the von Mises yield criterion.

This self-limiting effect is common to all the pipes irrespective of the geometry and material properties, although there is a difference between the pressurized and unpressurized pipes. As the input is increased beyond that which causes yield, the unpressurized pipes tend to reach a response plateau. This plateau appears to be associated with large local plastic strains as the limit moment for the pipe is approached. The pressurized pipes do not show this same tendency, and the post-yield response continues to rise slowly with increasing input.

This self-limiting effect of pipework vibration is clearly of great importance since, if it occurs in more general pipework systems, it puts an upper bound on the level of vibration of the system, at realistic inputs, and hence the level of loads applied to components and equipment.

4 Theoretical Response Prediction

Since it appears that plastic work absorbed by the pipe wall material controls the level of response, it seems reasonable that the most attractive way of theoretically modeling response is an energy balance approach equating the input energy at the supports to the energy absorbed in plasticity. This approach is best suited to steady-state excitation but should be most useful in highlighting the important parameters which dictate the level of response.

By making two major assumptions, namely that the mode shape is unaltered by plasticity and that the material properties can be approximated to the bilinear elastic/plastic relationship shown in Fig. 2, Beaney (1989) has shown that the system response k is related to the input acceleration by:

$$\frac{\omega_0^2 A}{g} = \frac{\sigma_c}{\sigma_g}\left[\eta\frac{\pi^3 k}{16} + \frac{4a^2}{a^2+b^2}G(k,m)\right] \tag{1}$$

where

$$G(k,m) = \frac{2(1-m)}{a^2-b^2}\left\{\int_{\sin^{1}\left(\frac{1}{k}\right)}^{\pi/2}\int_{\frac{a}{k\sin\beta}}^{a}\left[\frac{kmy^2}{a^2}\sin^2\beta + \frac{y}{a}(1-2m)\sin\beta - \frac{(1-m)}{k}\right]\sqrt{a^2-y^2}\,dyd\beta - \int_{\sin^{1}\left(\frac{a}{bk}\right)}^{\pi/2}\int_{\frac{a}{k\sin\beta}}^{b}\left[\frac{kmy^2}{a^2}\sin^2\beta + \frac{y}{a}(1-2m)\sin\beta - \frac{(1-m)}{k}\right]\sqrt{b^2-y^2}\,dyd\beta\right\}.$$

A = support displacement amplitude
a = external pipe radius
b = internal pipe radius
E = Young's Modulus
g = acceleration due to gravity
k = response level ratio = displacement: displacement at yield
m = strain-hardening ratio = plastic modulus: elastic modulus
y = distance from neutral axis
β = integration variable
η = proportion of critical damping
σ_c = dynamic elastic stress amplitude
σ_g = deadweight stress
ω_0 = system natural frequency

This expression is for a single span pipe pin jointed at its ends and applies to a particularly simple system with a sinusoidal deflected shape. More recent experimental and theoretical work indicates, however, that the response equation is little different for two equal spans of pipework vibrating in its second mode of vibration. Since the mode shape of this mode of vibration is very different, in that the peak/mean strain is larger and also there are multiple areas where yielding occurs, it is quite likely that the response level depends primarily on yield and strain hardening and, to a much lesser extent, on the complexity of the system.

5 Comparison Between Theory and Experiment

In equation (2) all the parameters except the bilinear material properties σ_c and m are known or measured. The most obvious choice of these variables for the ferritic material would be to take σ_c equal to first yield, since for modest strains no static

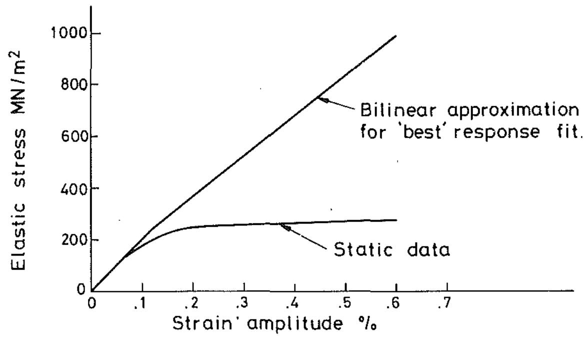


Fig. 4 Comparison of static and dynamic stress strain properties

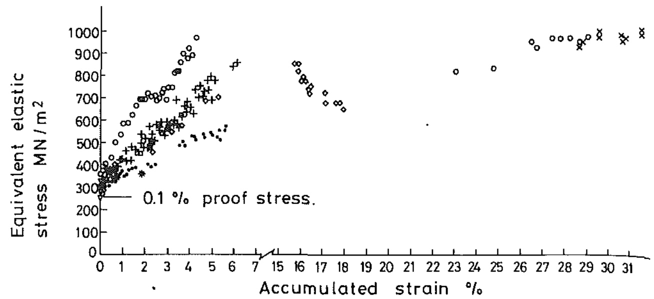


Fig. 6 Equivalent elastic stress versus accumulated strain austenitic pipe

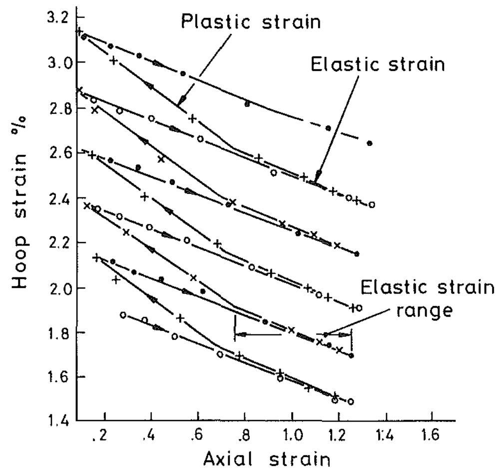


Fig. 5 Hoop and axial strains on pipe

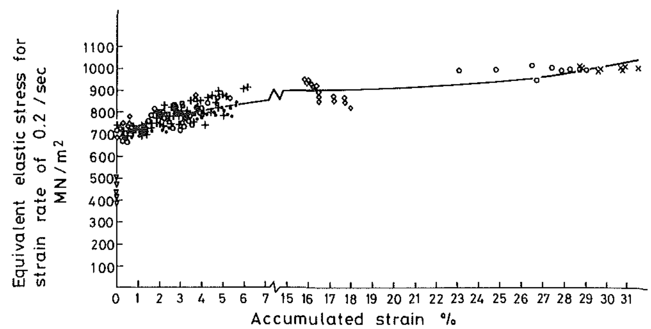


Fig. 7 Equivalent elastic stress versus accumulated strain austenitic pipe

strain-hardening occurs to take $m=0$. However, these values do not give a reasonable response prediction. Similarly, the bilinear model for the strain-hardening austenitic material based on the static tensile properties, Fig. 2, with $m=0.02$, gives a poor response prediction.

By taking a range of values for σ_c and m and fitting the theoretical response to the measured acceleration response, very good agreement is obtained if a reasonable value of σ_c of 10 percent less than static yield is used, together with very high strain-hardening m of 0.6 for the pressurized ferritic pipe. This strain hardening was clearly not expected, but the same value was required for several different ferritic materials, and correctly modeled the results for a wide range of pipe geometries and input levels.

For the austenitic material, the required value of σ_c was reasonable being ten percent less than the 0.2 percent proof stress, but the strain hardening was even higher at 0.75 and compares badly with the static data in Fig. 4. As with the ferritic materials, these same values were required for a wide range of geometries and input levels.

From both pipes it appeared that the dynamic material properties were unexpectedly different from the static uniaxial data.

6 Material Strain-Hardening

6.1 Measurement of Strain Hardening. Beaney (1987) showed that on pressurized pipes hoopwise ratcheting occurs associated with the dynamic axial yielding of the pipe wall and the sustained pressure stress. By extending the work of Edmunds and Beer (1961) he showed that no axial strain accumulation occurs, and that the hoop increments depend upon the amount of axial plastic strain and the level of hoop stress. On all the pressurized pipes this accumulation of hoop strain was very rapid up to a strain of about five percent, and it was thought that this might be the cause of the strain hardening.

To enable a direct measurement of strain hardening, the axial and hoop strains, measured in exactly the same position using stacked strain gauges, were plotted against each other for successive vibration cycles, as shown in Fig. 5. The hoopwise ratchet can clearly be seen from this figure, but more importantly, from the strain-hardening standpoint, so too can the extent of the elastic strain. Fortuitously, there is a marked difference in the ratio of hoop to axial strains in the elastic and plastic parts of the compressive half-cycle. Thus, for each cycle both total axial strain range and elastic axial strain range can be obtained from these plots as well as the total accumulated hoop strain.

Large numbers of elastic strain ranges were obtained in this way for both the ferritic and austenitic materials. Hoop pressure stress will affect this range. In order that pipes of differing pressure could be compared, the elastic strain range was adjusted using the von Mises yield criterion to obtain an equivalent uniaxial elastic stress range. Since the original hypothesis was that accumulated strain caused hardening, the elastic stress range is plotted against hoop strain in Fig. 6. Clearly, there is some connection between strain hardening and accumulated strain but the large spread of data indicates that some other parameter is also affecting the elastic range.

6.2 Dependence of Strain Hardening on Strain Rate. According to a literature survey, typical strain rates of about 0.2/sec should have little effect on strain hardening. However, a regression analysis, relating strain hardening to both accumulated strain and either strain rate or strain range, effected a large reduction in scatter of the data. The most suitable analysis appeared to be using an accumulated strain polynomial of order three and a strain rate or range of order two. The correlation between strain rate and hardening was slightly better than for strain range and hardening but, as all the tests were at approximately the same frequency, the two were linearly related and their relative effects inseparable.

In Fig. 7, the equivalent elastic yield stress, adjusted using

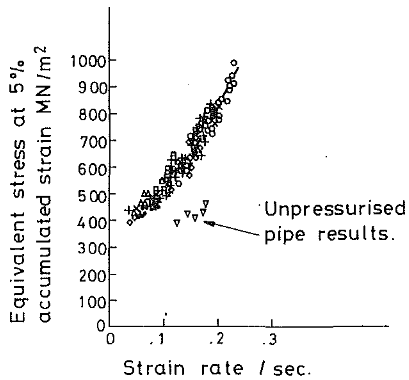


Fig. 8 Effect of strain rate equivalent elastic stress austenitic pipe

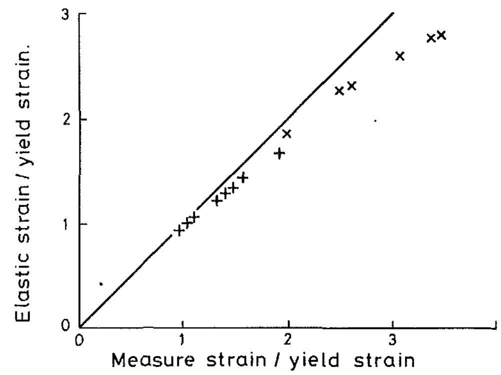


Fig. 10 Elastic strain versus measured strain austenitic material

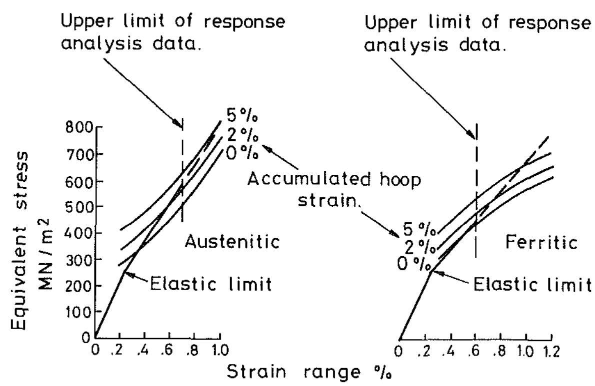


Fig. 9 Comparison of strain hardening from response and regression analyses

the regression constants to a representative strain rate of 0.2/sec, is plotted against accumulated strain. For this austenitic material it appears that accumulated strain is not the predominant parameter in controlling hardening, as it only increases the elastic stress range by about 40 percent. The majority of this hardening appears to take place with accumulated strains of less than 10 percent.

In Fig. 8, equivalent stress, adjusted to an accumulated strain of five percent, is plotted against strain rate. Strain rate, or perhaps strain range, clearly has a predominant role in strain hardening although it is noticeable that, in the absence of accumulated strain on the unpressurized pipe, it has virtually no effect. Herein, perhaps, lies the reason why the literature survey showed no strain-rate effects since all the data available were from uniaxial specimens similar to that in the unpressurized pipes. The only materials data for biaxially loaded materials appear to be that by Hancell and Harvey (1979), but unfortunately, their strain rates are very low.

It is interesting to note the scale of the strain hardening which appears to occur. The 0.2 percent proof stress of the material is 240 MN/m^2 compared with the measured dynamic elastic stresses of up to 1000 MN/m^2, a fourfold increase. This elastic stress is much greater than the UTS of the material which was measured as 590 MN/m^2.

The results for the ferritic material were similar. Hardening of about 40 percent due to accumulated strain and strain rate or strain range effects accounted for the majority of hardening. As with the austenitic material, strain rate gave a better quality fit to the data than strain range, but more controlled materials testing is required to differentiate between the two.

6.3 Comparison of Inferred and Directly Measured Hardening. As seen previously, to obtain a reasonable fit between the measured and calculated responses, values of m of 0.6 and 0.75 were required for ferritic and austenitic materials, respectively. As can be seen in Fig. 9 for both materials, this inferred level of strain hardening agrees quite well with the directly measured regression analysis results. During the response measurements the pipes were ratcheting, and attained maximum accumulated hoop strains of typically five percent for the austenitic pipes and a little more for the ferritic pipes. Also, the axial strain amplitude, and hence strain rate and hardness, varies both along and around the pipe. Thus, the bilinear hardening in the response model is some sort of average result for the materials with various amounts of strain hardening. Bearing this in mind, the inferred and directly measured material properties are very similar.

7 Effect of Strain Hardening on Response

There are two major effects that strain hardening has on response, first, it reduces plasticity with a consequent increase in response and, second, it suppresses the influence of stress raising features by stopping the formation of plastic hinges.

In Fig. 3, the theoretical response level for the same pipe for various hardening factors m are plotted. From this it is clear that the greater m the larger the response level.

Since yielding first occurs at stress concentration components such as bends, it is reasonable to expect these features to dictate system response. If plasticity is limited to these areas, not only will they locally absorb the input energy, but plastic hinges will tend to occur altering the mode of vibration. Strain hardening suppresses both of these effects forcing plasticity along the pipe. For example, if the strain rate at a component with a stress concentration of 1.6 were 0.15/sec, then the strain rate in the adjacent pipe would be $0.15/0.16 = 0.094$/sec. The accumulated strain would also be higher at the concentration. Putting these strain rates and reasonable values of accumulated strain into the regression analysis formula gives the ratio of elastic range at the stress concentration to that in the main pipe at 1.6:1. This is the same as the stress concentration and hence the amount of plastic strain is similar in the component to that in the adjacent pipe.

This result is confirmed by the measured strains at a stress concentration compared with the elastically calculated strains in Fig. 10. The amount of plastic strain is the difference between the measured and calculated elastic strains and, as can be seen in the figure, this is relatively modest even though first yield has been exceeded by a factor of 3.

The implications for the prediction of response are that because plastic hinges are not formed, the mode shape and frequency of the vibration are relatively unchanged from the elastic values, and that because plastic strain is not concentrated at components, it spreads along large sections of the pipe. Thus, it is a "thin smear" of plasticity along large lengths of the pipe that controls the energy absorption rather than concentrated pockets of plasticity at stress concentrations. Hence, apart from changing the elastic stiffness and mass of the system, pipework components do not greatly affect the level of vibration. This is ultimately controlled by the straight pipes, and it is the straight pipe performance which will dictate the loadings on components, equipment, and supports.

This result has only been verified for pressurized pipes with a sustained static hoop stress, normal to the applied dynamic stress, equal to or greater than the normal design level, and it is not applicable to unpressurized pipes. It is possible that there could be a threshold static stress below which hardening does not occur, or alternatively, it could be that hardening progressively increases with increasing normal static stress. Where hardening does not occur, then clearly high-plastic strains will concentrate at components. The effect of these higher local strains is somewhat nullified however, since, in the absence of pressure, fatigue is the mode of ultimate failure, which is a much slower mechanism for failure than ratcheting or fatigue/ratcheting which occurs in pressurized components.

8 Conclusions

Plasticity controls the high-level dynamic response of pipework by absorbing large amounts of energy in plastically deforming the pipe wall material.

Very large material strain hardening occurs in pressurized pipes under the influence of a biaxial stress field consisting of the pressure hoop stress and the dynamic axial bending stress. Since plasticity limits the response and strain hardening controls the amount of plasticity, the hardening also has an overriding influence on response. Its effects are manifest in three ways. First, it stops the formation of plastic hinges at stress concentrations, thus the elastic frequency and mode shape of the vibration are maintained. Second, since it limits plasticity at stress concentrations, it forces the majority of system energy absorption into the straight lengths of pipe, thus these dictate system response. Third, the greater the strain hardening, the higher the level of response. Given the strain-hardening properties of the pipe wall material, the steady-state response of the straight pipe sections can be predicted using an energy balance approach. Since, in the presence of strain hardening, it is these straight sections which control the response, this method of response prediction should be adaptable to predicting the response of complete systems.

Acknowledgments

This work was carried out at the Berkeley Nuclear Laboratories of the Research Division and the paper is published with permission of the Central Electricity Generating Board.

References

Beaney, E. M., 1987, "The Response of Pipes to Seismic Loading," *Trans. of the Ninth International Conf. on Structural Mechanics in Reactor Technology*, Vol. K2, Lausanne, pp. 805–810.

Beaney, E. M., 1989, "Pipework Engineering and Operation," *I. Mech. E. Conf.*, London, Feb., pp. 281–290.

Edmunds, H. G., and Beer, F. J., 1961, "Notes on Incremental Collapse of Pressure Vessels," *Int. Mechanical Engineering Science*, Vol. 3, No. 3, pp. 187–199.

Hancell, P. J., and Harvey, S. J., 1979, "The Use of Kinematic Hardening Models in Multi-Axial Cyclic Plasticity," *Fatique of Engineering Materials and Structures*, Vol. 1, Pergamon Press, pp. 271–279.

Dynamic Pulse Buckling of Imperfection-Sensitive Shells

Herbert E. Lindberg
APTEK, Inc.,
San Jose, CA 95129

The theoretical basis of two related but distinctly different dynamic buckling criteria are summarized with the objective of demonstrating the range of applicability of each, so that together they cover the entire range of dynamic pulse loads from nearly impulsive loads to step loads of infinite duration. The example chosen is a cylindrical shell under elastic axial loads but the approach is applicable more generally. A critical amplification-of-imperfections criterion with a linear shell theory is shown to be applicable for short duration loads, for which a threshold nonlinear divergence criterion gives loads an order of magnitude too conservative. Conversely, the linear theory is inapplicable for long duration loads, for which critical loads are lower than the linear static buckling load because of imperfection sensitivity. In this range the threshold nonlinear divergence criterion is used. For loads of intermediate duration, an extended critical amplification criterion is used with equations that conservatively assume zero static buckling load but give an unchanged formula for critical load amplitude-duration combinations.

Introduction

A critical amplification criterion has been successfully applied to calculate dynamic pulse-buckling loads in a wide variety of structural elements (Lindberg and Florence, 1987), in each case with quite reasonable agreement with experimental loads for thresholds of dynamic buckling. In particular, critical loads for cylindrical shells under axial impact have been predicted for constant elastic axial stresses (Lindberg and Herbert, 1966), sustained constant axial plastic flow (Florence and Goodier, 1968), and oscillating elastic axial stresses (Lindberg, Rubin and Schwer, 1987). Buckling of cylindrical shells at elastic axial levels is very sensitive to initial imperfections. Under static loads, this sensitivity causes large changes in critical buckling stresses for almost imperceptible imperfections. In the present paper it is demonstrated that, under dynamic pulse loads, imperfection sensitivity does not strongly affect critical stress-duration combinations for buckling, but instead affects the range of stresses over which the theory can be applied.

Imperfection sensitivity under static axial loading has an extensive literature of research on the source of this sensitivity and methods of analysis. With the objective of avoiding duplication of this research for dynamic loads of long duration (step loads), Budiansky and Hutchinson (1964) developed a theory to relate critical dynamic loads to static buckling loads of imperfect shells, without specific reference to the imperfections themselves. They found expressions for the ratio of dynamic-to-static buckling loads as a function of the ratio of the static buckling load of the imperfect shell to the classical static buckling load of the perfect shell. They then extended this idea to pulse loads of finite duration (Hutchinson and Budiansky, 1966).

The buckling criterion used by Budiansky and Hutchinson is the transition from oscillatory motion under subcritical loads to divergent motion under buckling loads. For typical imperfect shells (static buckling loads about one fourth the classical loads), they found that critical step loads are about three fourths the slowly applied static load. For rectangular pulse loads of finite duration, their results were more complex, but toward the limit of short durations the critical condition reduces, of course, to a critical loading impulse. This is the same condition found by Lindberg and Herbert (1966), but the impulses from the critical amplification criterion are an order of magnitude larger than from the divergence criterion used by Budiansky and Hutchinson.

Roth and Klosner (1964) also found critical loads for cylindrical shells under axial rectangular pulse loads, based on a criterion of a sudden increase in nonlinear response amplitude, used by Budiansky and Roth (1960) for shallow spherical shells. Roth and Klosner's critical impulse for a cylindrical shell under short duration loads was only slightly larger than that given by Lindberg and Herbert, suggesting that, for short duration loads, their nonlinear response amplitude change criterion was similar to the critical amplification criterion.

In the present paper it is shown that the critical amplification criterion is the more appropriate for pulse loads, while the threshold divergence criterion is appropriate for step loads. An interpolation method is given for loads of intermediate duration. It is further suggested that, with the general source of imperfection sensitivity identified, the critical amplification criterion can be applied by knowing only the ratio of imperfect-to-perfect static buckling loads, just as for the threshold divergence criterion.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be sent to the Technical Editor, Prof. Leon M. Keer, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 12, 1989; final revision, May 18, 1990.

Critical Amplification Theory

Equations of Motion. Both the threshold divergence and critical amplification criteria are applied to cylindrical shells by means of Donnell's equations, which in linear form are:

$$D\nabla^4 w+\overline{N}_x \frac{\partial^2}{\partial x^2}(w+w_i)+\rho h \frac{\partial^2 w}{\partial t^2}-\frac{1}{a}\frac{\partial^2 F}{\partial x^2}=0 \quad (1)$$

$$\nabla^4 F=-\frac{Eh}{a}\frac{\partial^2 w}{\partial x^2} \quad (2)$$

where

$$\nabla^4=\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{a^2\partial\theta^2}\right)^2 \quad (3)$$

In these equations, x is axial coordinate, θ is circumferential coordinate, w is radial displacement, positive inward and measured from an unstressed initial displacement w_i, ρ is material density, E is Young's modulus, h is shell wall thickness, a is shell radius, $D=Eh^3/12(1-\nu^2)$ is shell bending stiffness, ν is Poisson's ratio, F is Airy's stress function for in-plane force resultants produced by the buckling deformation, and $\overline{N}_x$ is the part of the axial force resultant from the applied axial load.

With dimensionless variables defined by

$$\xi=x\left(\frac{\overline{N}_x}{D}\right)^{1/2} \quad \eta=a\theta\left(\frac{\overline{N}_x}{D}\right)^{1/2} \quad \tau=\frac{\overline{N}_x}{(\rho h D)^{1/2}}\cdot t \quad (4)$$

Donnell's equations become

$$\nabla^4 w+\frac{\partial^2}{\partial\xi^2}(w+w_i)+\ddot{w}-\frac{1}{a\overline{N}_x}\frac{\partial^2 F}{\partial\xi^2}=0 \quad (5)$$

$$\nabla^4 F=\frac{Eh\,D}{a\overline{N}_x}\frac{\partial^2 w}{\partial\xi^2} \quad (6)$$

where now

$$\nabla^4=\left(\frac{\partial^2}{\partial\xi^2}+\frac{\partial^2}{\partial\eta^2}\right)^2 \quad (7)$$

and dots indicate differentiation with respect to τ.

With simple-support boundary conditions

$$w=\partial^2 w/\partial x^2=0 \quad \text{at } x=0,\ L \quad (8)$$

dynamic motion following sudden application of $\overline{N}_x$ can be expressed by the Fourier series

$$w(\xi,\eta,\tau)=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}W_{mn}(\tau)\sin\alpha_m\xi\sin\beta_n\eta \quad (9)$$

$$F(\xi,\eta,\tau)=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}F_{mn}(\tau)\sin\alpha_m\xi\sin\beta_n\eta \quad (10)$$

in which

$$\alpha_m=\frac{m\pi}{L}\left(\frac{D}{\overline{N}_x}\right)^{1/2} \quad \beta_n=\frac{n}{a}\left(\frac{D}{\overline{N}_x}\right)^{1/2} \quad (11)$$

and the initial imperfections are also expanded into the Fourier series

$$w_i(\xi,\eta)=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}a_{mn}\sin\alpha_m\xi\sin\beta_n\eta. \quad (12)$$

A similar set of equations results for imperfections of the form $b_{mn}\sin\alpha_m\xi\cos\beta_n\eta$. With these expansions substituted into Donnell's equations (5)–(7), the equations of motion for the modal amplitudes W_{mn} are

$$\ddot{W}_{mn}+k(\alpha_m,\beta_n)W_{mn}=\alpha_m^2 a_{mn} \quad (13)$$

where

$$k(\alpha_m,\beta_n)=(\alpha_m^2+\beta_n^2)^2-\alpha_m^2+\frac{1}{4}\left(\frac{\sigma_c}{\sigma}\right)^2\frac{\alpha_m^4}{(\alpha_m^2+\beta_n^2)^2}. \quad (14)$$

The form for the multiplier in the last term of equation (14) follows from the observation that

$$\frac{EhD}{a^2\overline{N}_x^2}=\frac{1}{4}\left(\frac{\sigma_c}{\sigma}\right)^2 \quad (15)$$

where σ is the axial stress from $\overline{N}_x$ and

$$\sigma_c=\frac{Eh}{a}\frac{1}{\sqrt{3(1-\nu^2)}} \quad (16)$$

is the classical static buckling stress, which can be found by setting $k(\alpha_m,\beta_n)=0$ and minimizing σ with respect to α_m and β_n treated as continuous variables.

Imperfection Growth and Buckling. The solutions to modal equations of motion (13) are

$$\frac{W_{mn}(\tau)}{a_{mn}}=\frac{\alpha_m^2}{k(\alpha_m,\beta_n)}\left[1-\begin{matrix}\cosh p\tau\\ \\ \cos p\tau\end{matrix}\right] \quad (17)$$

in which

$$p=|k(\alpha_m,\beta_n)|^{1/2}. \quad (18)$$

The hyperbolic form is taken for $k(\alpha_m,\beta_n)<0$ and the trigonometric form is taken for $k(\alpha_m,\beta_n)>0$. For $k(\alpha_m,\beta_n)=0$, the function multiplying α_m^2 is replaced by $\tau^2/2$, but this is seldom of concern because only in rare cases do α_m, β_n make k precisely zero with integer values m and n. The quantity given by the right side of equation (17) is called the "amplification function," since it defines the amount by which the imperfection coefficients a_{mn} are amplified by dynamic motion.

If one were to use a threshold divergence criterion for this linear dynamic buckling motion, the dynamic buckling load would be simply the static buckling load, since it separates oscillatory motion from divergent motion. However, for finite duration pulse loads this criterion is far too conservative, as shown by the many examples of pulse buckling of structural elements in Lindberg and Florence (1987), and in particular for the cylindrical shell under axial loading (Lindberg and Herbert, 1966). With finite durations, loads with amplitudes far in excess of static buckling loads can be safely applied as long as the pulse duration is short enough that the magnitude of the motion remains acceptable. Experiments on a wide variety of structural elements have shown that motion is acceptably small if the amplification is less than about 25.

Lindberg and Herbert (1966) evaluated equation (17) over the range α_m, $\beta_n<2$ of significant amplifications, for τ ranging from 0 to 12. The most amplified mode is an axisymmetric mode with axial half-wavelength $l_x\approx\sqrt{2}l_0$, where $l_0=\pi\sqrt{ah}/[12(1-\nu^2)]^{1/4}$ is the axisymmetric classical static buckle half-wavelength. This mode achieves an amplification of 25 at τ ranging only from 6 to 8 for σ/σ_c ranging from 1.1 to ∞. Also, substantial growth occurs in hundreds of modes for thin shells. A statistical analysis showed that with uniformly distributed initial imperfections a_{mn}, the most probable wavelengths in the buckled form have an axial half-wavelength $l_x\approx\sqrt{2}l_0$ and a circumferential-to-axial wavelength ratio of $l_\theta/l_x\approx 3$ at these buckling times. Measurements of permanent buckled forms and high-speed motion pictures showed wavelengths in good agreement with these predictions, and that at $\tau=7$ buckles were just perceptible. It was therefore suggested that $\tau=7$ be taken as a conservative buckling criterion, based on an amplification of about 25, essentially independent of σ/σ_c but with $\sigma/\sigma_c>1$. Later in the present paper it is suggested that the latter condition can be relaxed to $\sigma/\sigma_{ds}>1$, where σ_{ds} is the dynamic buckling load of the imperfect shell under step loading, from the threshold divergence criterion, for which $\sigma_{ds}/\sigma_s<1$ and σ_s is the static buckling stress of the imperfect shell.

From the definition of τ in equation (4), the critical condition $\tau=7$ at $t=T$ gives

$$\overline{N}_D T = 7(\rho h D)^{1/2} \tag{19}$$

or

$$\sigma_d T = \frac{7}{\sqrt{12}}\,\rho c h \approx 2\rho c h \tag{20}$$

where $c=\sqrt{E/\rho(1-\nu^2)}$ is the membrane wave speed in the shell, $\overline{N}_D$ is the critical dynamic resultant force at duration T, and σ_d is the corresponding dynamic stress. Thus, threshold buckling deformations occur at a critical impulse imparted by the axial load. Also, because of the insensitivity to σ/σ_c noted previously, $\sigma_d T$ does not depend on a.

Nonlinear Divergence Theory

Equations of Motion. The initial work of Budiansky and Hutchinson (1964) focused on buckling from step loads, which is essentially a dynamic perturbation of static buckling since the load is maintained indefinitely. Thus, they used Koiter's (1963) theory of elastic stability and post-buckling behavior to capture the transition from small deformations at subcritical loads to large deformations when a critical buckling load is exceeded, just as for static buckling of imperfection sensitive shells. With a nonlinear form of Donnell's equations, they focused on analysis of motion in two modes, with ζ_1 taken as the amplitude of motion in the axisymmetric classical buckling mode and ζ_2 taken as the amplitude of a nonaxisymmetric classical buckling mode with equal axial and circumferential wavelengths (half wavelengths $l_x = l_\theta = 2l_0$; see, for example, Lindberg and Florence (1987) pp. 284–285). Experiments reported by Almroth, Holmes, and Brush (1964) demonstrate that initial buckling indeed occurs in this mode for step loading, produced in these experiments by a small lateral perturbation impulse applied to a statically loaded shell. (A suddenly introduced additional "imperfection" is not the same as a suddenly applied load, but the resulting dominant response modes are the same.)

With ζ_1 and ζ_2 expressed as fractions of the wall thickness, the equations of motion were found to be

$$\frac{1}{4}\ddot{\zeta}_1 + \left(1-\frac{\lambda}{\lambda_C}\right)\zeta_1 - \frac{3b}{32}\,\zeta_2^2 = \frac{\lambda}{\lambda_C}\,\overline{\zeta}_1 \tag{21}$$

$$\ddot{\zeta}_2 + \left(1-\frac{\lambda}{\lambda_C}\right)\zeta_2 - \frac{3b}{2}\,\zeta_1\zeta_2 = \frac{\lambda}{\lambda_C}\,\overline{\zeta}_2 \tag{22}$$

where λ is the applied axial stress, λ_C is the classical static buckling stress given by equation (16), $b=[3(1-\nu^2)]^{1/2}$, and $\overline{\zeta}_1$ and $\overline{\zeta}_2$ are imperfections, also as fractions of the wall thickness. (The notation ζ and b is used here rather than ξ and c as in the Budiansky and Hutchinson papers, because of other use of these symbols in the present paper.) Finally, here, dots now indicate differentiation with respect to $\tau=\omega_2 t$, where the vibration frequencies associated with the two modes in the unloaded shell are

$$\omega_1 = \sqrt{2}c/a \quad \text{and} \quad \omega_2 = c/a\sqrt{2} \tag{23}$$

which can be found from equation (13) with $\sigma=0$ and wave numbers α_m and β_n corresponding to the half wavelengths given above for these modes.

Numerical analysis showed that minimum dynamic buckling loads based on a threshold divergence criterion (discussed more explicitly in the next subsection) occurred with $\overline{\zeta}_1 = 0$, but nevertheless with $\zeta_1 \neq 0$ because of the nonlinear coupling. Solutions to a good approximation for this case were obtained by neglecting the inertia term $\ddot{\zeta}_1/4$ in equation (21), which allows ζ_1 to be expressed in terms of ζ_2^2 from equation (21) and substituted into equation (22) to obtain the single nonlinear equation

$$\ddot{z}_2 + (1-\lambda/\lambda_C)z_2 - \left[\frac{9b^2\overline{\zeta}_2^2}{64(1-\lambda/\lambda_C)}\right]z_2^3 = \lambda/\lambda_C \tag{24}$$

where $z_2 = \zeta_2/\overline{\zeta}_2$.

Hinged-Rod Model. Budiansky and Hutchinson further observed that the form of equation (24) is very similar to that of the two rigid rod, three-hinge column model of von Kármán, Dunn, and Tsien (1940), with a mass and lateral cubic-softening spring attached to the central hinge. The mass is M, the length of each rod is L_r, and the nonlinear spring force from lateral displacement u is

$$F = KL_r(\zeta - B\zeta^3), \quad B>0. \tag{25}$$

With notation analogous to that for the shell, namely $\zeta = u/L_r$, and P and $P_C = KL_r$ denoting axial force and zero-imperfection buckling load, respectively, the equation of motion is

$$\ddot{z} + (1-P/P_C)z - B\overline{\zeta}^2 z^3 = P/P_C \tag{26}$$

where $z=\zeta/\overline{\zeta}$ and dots now indicate differentiation with respect to $t\sqrt{K/M}$.

For the imperfect structure, with $\overline{\zeta}\neq 0$, the equilibrium displacement increases with increasing load to a maximum and then decreases with further increases in load. States beyond the maximum are therefore unstable, and the maximum is the buckling load P_S of the imperfect structure. This load is found by omitting the $\ddot{z}$ term in equation (26) and setting $dP/dz=0$, which yields

$$(1-P_S/P_C)^{3/2} = \frac{3\sqrt{3}}{2}\,B^{1/2}\overline{\zeta}(P_S/P_C). \tag{27}$$

The dynamic buckling load for step loading with $z=\dot{z}=0$ at $t=0$ is found by first using the identity $\ddot{z} = \dot{z}\,d\dot{z}/dz$ so that equation (26) can be integrated once to obtain

$$\dot{z}^2 + (1-P/P_C)z^2 - \frac{1}{2}\,B\overline{\zeta}^2 z^4 = 2(P/P_C)z. \tag{28}$$

At loads below the dynamic buckling load, the steady-state motion is periodic and equation (28) defines its limit cycle in phase space z, $\dot{z}$. The maximum value, $z_{\max}$, of this limit cycle occurs when $\dot{z}=0$, which gives

$$(1-P/P_C)z_{\max}^2 - \frac{1}{2}\,B\overline{\zeta}^2 z_{\max}^4 = 2(P/P_C)z_{\max}. \tag{29}$$

The dynamic buckling load P_D is defined as the load for which the amplitude (and period) of this limit cycle is infinite, so the motion diverges rather than approaching a limit cycle. This occurs under the condition $dP/dz_{\max}=0$ applied to equation (29), with the result

$$(1-P_D/P_C)^{3/2} = \frac{3\sqrt{6}}{2}\,B^{1/2}\overline{\zeta}(P_D/P_C). \tag{30}$$

A key feature of this simple model is that the imperfection parameter $B^{1/2}\overline{\zeta}^2$ can be eliminated between equations (27) and (30), giving a relationship between the static and dynamic buckling loads with no explicit dependence on the imperfections. The result is

$$P_D/P_S = \frac{\sqrt{2}}{2}\left(\frac{1-P_D/P_C}{1-P_S/P_C}\right)^{3/2}. \tag{31}$$

By a similar procedure, the static buckling load from equation (24) for the cylindrical shell is given by

$$(1-\lambda_S/\lambda_C)^2 = \left[\frac{9\sqrt{3}b}{16}\,|\overline{\zeta}_2|\right](\lambda_S/\lambda_C) \tag{32}$$

and the dynamic buckling load is given by

$$\lambda_D/\lambda_S = \frac{\sqrt{2}}{2}\left(\frac{1-\lambda_D/\lambda_C}{1-\lambda_S/\lambda_C}\right)^2. \tag{33}$$

Plots in Budiansky and Hutchinson (1964) of P_D/P_S versus

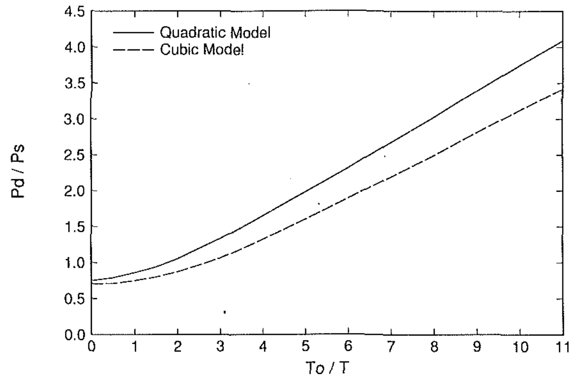


Fig. 1 Conservative dynamic buckling estimates ($\lambda_S/\lambda_C=0$) from the threshold divergence theory, rectangular loading

P_S/P_C from equation (31) and λ_D/λ_S versus λ_S/λ_C from equation (33) are very similar, the only difference being slightly more curvature in the plot from equation (33) than from equation (31) because of the larger exponent. Similar results were also found for the rigid-rod model with a quadratic rather than cubic-softening spring, which gives the same exponent as in equation (33) but a coefficient $3/4=0.75$ in place of $\sqrt{2}/2=0.707$.

Critical Finite Duration Loads. In Hutchinson and Budiansky (1966), the equations of motion of the rigid-rod model for both the quadratic and cubic-softening springs were integrated numerically for rectangular and triangular (sudden jump in load followed by linear decay to zero) finite-duration pulse loads. For each model and pulse shape, they again gave plots of λ_D/λ_S versus λ_S/λ_C but with pulse duration as a parameter, found by use of a divergence threshold criterion just as for step loads. These curves showed a rapid increase in finite duration critical loads with λ_S/λ_C increasing beyond about 0.2, suggesting a high sensitivity of critical dynamic loads to imperfections. In the present paper it is shown that, with the more appropriate critical amplification criterion, critical short pulse loads are insensitive to imperfections for any value of λ_S/λ_C.

For conservative practical application, curves were given of λ_D/λ_S versus T_0/T for the limiting case of imperfect shells with $\lambda_S/\lambda_C\to 0$, based on the observation that for $\lambda_S/\lambda_C<0.2$ critical dynamic load ratios λ_D/λ_S are weakly dependent on λ_S/λ_C. The time T_0 is the vibration period of the dynamic buckling mode in the absence of loading, and T is the rectangular pulse duration. These are repeated here in Fig. 1. The curves approach the static results of Budiansky and Hutchinson (1964) as $T_0/T\to 0$, and approach straight lines toward the impulsive loading limit $T_0/T\to\infty$.

For the cubic model, which more closely approximates the behavior of the cylindrical shell equation (24), a straight line approximation $\lambda_D/\lambda_S=T_0/T$ can be used to good accuracy for $T_0/T>3$. With $T_0=2\pi/\omega_2$, and $\omega_2=c/a\sqrt{2}$ from equation (23) and $\sigma_c=Eh/a\sqrt{3(1-\nu^2)}$ from equation (16), this line gives the following simple formula for critical dynamic axial buckling stress

$$\sigma_d=\pi\left(\frac{2}{3}\right)^{3/2}\frac{\rho ch}{T}\cdot\frac{\sigma_s}{\sigma_c} \tag{34}$$

or

$$\sigma_d T=1.71\ \rho ch(\sigma_s/\sigma_c). \tag{35}$$

For a typical static buckling load ratio $\sigma_s/\sigma_c\approx 0.25$, equation (35) gives $\sigma_d T=0.43\ \rho ch$. If one uses the symmetric mode period from ω_1 in equation (23) as the more conservative estimate suggested by Hutchinson and Budiansky (1966), then the coefficient in equation (35) becomes $1.71/2=0.855$. With this more conservative buckling mode period, $\sigma_d T=0.21\ \rho ch$. These are in exactly the same form as given by equation (20) from the amplification criterion but with a coefficient an order of magnitude smaller. If one were to use for T_0 the period from the most amplified mode, a symmetric mode with half wavelength $l_x=l_0/\sqrt{2}$, the coefficient would be even smaller.

Choices of Buckling Criteria

The excellent agreement between experimental results and critical loads based on a critical amplification criterion suggests that for relatively short finite duration loads, this criterion is more appropriate than the threshold divergence criterion. Furthermore, the formula $\sigma_d=2\ \rho ch$ is also the formula for a bar or flat plate, which corresponds to $a\to\infty$. Thus, there is no reason to suspect that the critical amplification formula is unconservative because of any peculiarity of complex nonlinear shell response; the finite radius of the shell makes the shell "stiffer" than the plate. In the physical buckling process, a sudden-jump dynamic load is applied by impact, and divergent flexural motion takes place only while the stress pulse is maintained. If the shell has a free boundary condition at the opposite end, as in Lindberg and Herbert (1966), then following the compressive pulse the stress jumps to a tensile stress equal to the initial compression and not to zero as assumed in Hutchinson and Budiansky (1966).

Furthermore, as time proceeds, axial waves continue to reverberate between one end of the shell and the other. For the case with one end impacted and fixed to a heavy mass and the other end free, these reverberations result in alternating compressive and tensile pulses near the impacted end of the shell, where the flexural motion is largest. Lindberg, Rubin, and Schwer (1987) showed that further buckling occurs during the first two or three compressive pulses, interspersed by oscillatory motion during the tensile pulses. Also, the corresponding change of the equations of motion between hyperbolic and elliptic forms results in buckle growth fixed in space during the compressive pulses and bending wave propagation away from the impacted end during the tensile pulses. This spread in bending energy during the tensile pulses, together with the finite membrane energy available in the initial compression wave, limits the amount of buckle growth from later compressive pulses such that the single-pulse formula in equation (20) still gives a reasonable estimate for critical impact loading, with T taken as the single round trip transit time of an axial stress wave.

If the shell has an axially fixed boundary at the end opposite the impact, then the compressive stress increases with each axial stress wave reverberation between the impacted and fixed end, resulting eventually in long duration dynamic loading. This is closer to the situation analyzed in Budiansky and Hutchinson (1964), but no attempt was made there to define how one would obtain a sudden increase in load to a fixed load of constant magnitude. The implicit assumption is that the analyst is seeking a conservative estimate for a dynamically applied load and that a step load is a conservative idealization of actual loading through a short series of axial wave reverberations. Thus, in these cases the threshold divergence criterion is appropriate.

Criteria for Intermediate Pulse Durations

In place of Fig. 1, the combined criteria scheme given in Fig. 2 is suggested. The short-dashed line (and the solid extension superimposed on it) is from the critical amplification criterion, given in the form

$$\frac{P_D}{P_S}=\frac{\sqrt{6}}{\pi}\left(\frac{T_0}{T}\right)\left(\frac{P_C}{P_S}\right) \tag{36}$$

which is equation (20) with T_0 taken as the free vibration period $\sqrt{2}\pi a/c$ of the axisymmetric classical buckling mode. (The extension of applicability to $P_D/P_C<1$ is made with the con-

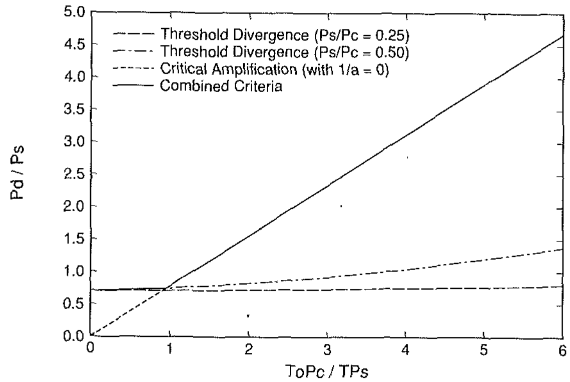


Fig. 2 Dynamic buckling loads for cylindrical shells under rectangular pulse loads of duration T. (Here, T_0 is the unloaded vibration period of the classical axisymmetric buckling mode.)

servative assumption that for pulse buckling one can consider that $1/a=0$, so that an effective $P_C^{\text{eff}} \rightarrow 0$, as discussed more fully in the following paragraphs.) The long-dashed and dash-dot curves are from the threshold divergence criterion for imperfections such that $P_S/P_C=0.25$ and 0.50, respectively. These are essentially the cubic model curve from Fig. 1 with the T_0/T abscissa stretched out to the new abscissa definition in Fig. 2, but with a slight increase in P_D/P_S because of the effect of P_S/P_C as given in Hutchinson and Budiansky (1966).

The combined criteria curve is the solid curve, taken as the upper envelope of the two criteria because each criterion is conservative. This solid curve consists essentially of the critical amplification curve for P_D greater than the step buckling load $P_D/P_S=\sqrt{2}/2$ given by equation (31) (or equation (33)) for $P_S/P_C \rightarrow 0$. For longer pulse durations (smaller T_0/T), this straight line is terminated and the critical load is taken as the conservative value $P_D/P_S=\sqrt{2}/2$ from the step-load theory. For common shells in which $P_S/P_C \approx 0.25$, the corner of the resulting plot occurs at a pulse duration $T \approx 4T_0$. For a 1-m diameter shell, $T \approx 4\sqrt{2}\pi(1\text{ m})/5000\text{ m/s}=0.00355$ s.

For values of T_0P_C/TP_S "near" but greater than the intersection point 0.907, the dynamic-to-static classical buckling load ratio $P_D/P_C=(P_D/P_S)(P_S/P_C)$ is less than 1 but lies on the critical amplification portion of the plot in Fig. 2. For example, with $P_S/P_C=0.25$, $P_D/P_C<1$ for $T_0P_C/TP_S<5.2$. In this range, the roots $k(\alpha_m, \beta_n)$ from equation (14) are positive and response is oscillatory rather than divergent. In this range we appeal to the nonlinear equations of motion (21) and (22) which, by definition of P_S/P_C in equation (32), give unstable motion for all points on the plot in Fig. 2. However, we continue to use $\sigma_d T$ calculated on the basis of a critical amplification rather than threshold divergence, which has been shown to be too conservative. One method to visualize this approach is to conservatively take $\sigma_c/\sigma=0$ in equation (14) which, as mentioned previously, leads to the same critical impulse formula $\sigma_d T=2\rho ch$ as for finite σ_c.

Beyond this appeal to the nonlinear equations of motion, there are two other physical processes that result in divergent motion for $\sigma/\sigma_c<1$. The first is the quasi-nonlinearity of an increase in the local radius of curvature a because of imperfections in modes with wavelengths longer than for the dynamic modes of response. From equation (16), this increase results in a local decrease in σ_c over portions of the shell where imperfections so combine. Calculations of curvature changes with imperfection amplitudes given in Arbocz (1982) give static buckling load reductions of 10 to 20 percent just from this effect. The basis of these calculations and comparisons of buckling load reductions from nonlinear effects (equation (32)) and from curvature changes are given in the Appendix.

The second additional source of divergent motion for $\sigma/\sigma_c<1$ is from hoop stresses produced by the Poisson effect for dynamic loading. Even small hoop stresses result in divergent motion. For the shell with $a/h=500$, $L/2a=1.62$ in Lindberg, Rubin, and Schwer (1987), calculations there showed that the Poisson effect reduced the critical load separating oscillatory from divergent motion to $\sigma_{cr}/\sigma_c=0.204$, for the $m=1$, $n=4$ mode. However, these hoop stresses are quickly relieved because the shell is free to expand. The duration of the initial hoop stress pulse for impact loading is the quarter period $\pi a/2c$ of the breathing mode. The value of τ with $\overline{N}_\theta=\nu\overline{N}_x$ for this period and $\nu=0.3$ is 0.98 σ/σ_c, so initial growth from the Poisson effect is small. Nevertheless, for the multipulse loading in Lindberg, Rubin, and Schwer (1987), the duration of the next compressive swing of the hoop mode is the half period. Strain measurements for an impact load $\sigma/\sigma_c=0.87$ showed both axial and circumferential flexural growth during the second circumferential compressive pulse.

Summary and Conclusions

A linear critical amplification criterion applied to dynamic buckling from pulse loads gives conservative estimates for combinations of stress amplitude and duration that can be safely applied with only modest flexural motion. A nonlinear threshold divergence criterion applied to these same pulse loads gives amplitude-duration combinations an order of magnitude less than from the critical amplification criterion. The latter criterion is therefore overly conservative for these relatively short pulse loads. Conversely, for long duration loads, the linear critical amplification criterion is unconservative because *linear* divergence and hence buckle growth occurs only for $\sigma/\sigma_c>1$. For long duration loads, the nonlinear threshold divergence criterion is appropriate. For loads of intermediate duration, the linear critical amplification criterion is made conservative (but not as conservative as the nonlinear threshold divergence criterion) by letting $\sigma_c=0$ in the equations of motion, which allows the formula from this criterion to be applied to all loads with $\sigma/\sigma_{ds}>1$, where σ_{ds} is the dynamic buckling load from the nonlinear threshold divergence criterion for step loads. The two criteria then give dynamic buckling loads that are not overly conservative over the entire range of pulse durations for which critical stresses are elastic. Furthermore, no specific reference is made to imperfections in either dynamic theory, so imperfection sensitivity need be investigated in detail only for static buckling.

References

Almroth, B. O., Holmes, A. M. C., and Brush, D. O., 1964, "An Experimental Study of the Buckling of Cylinders Under Axial Compression," *Exptl. Mech.*, Vol. 4, pp. 263–270.

Arbocz, J., 1982, "The Imperfection Data Bank, A Means to Obtain Realistic Buckling Loads," *Buckling of Shells, Procedures of a State-of-the-Art Colloquium*, Springer-Verlag, New York.

Budiansky, B., and Hutchinson, J. W., 1964, "Dynamic Buckling of Imperfection-Sensitive Structures," *Proc. 11th Int. Congress of Appl. Mech.*, Springer-Verlag, Berlin, pp. 636–651.

Budiansky, B., and Roth, R. S., 1960, *Axisymmetric Dynamic Buckling of Clamped Shallow Spherical Shells*, NASA Technical Note D-1510.

Florence, A. L., and Goodier, J. N., 1968, "Dynamic Plastic Buckling of Cylindrical Shells in Sustained Axial Compressive Flow," ASME Journal of Applied Mechanics, Vol 35, pp. 80–86.

Hutchinson, J. W., and Budiansky, B., 1966, "Dynamic Buckling Estimates," *AIAA Journal*, Vol. 4, No. 3, pp. 525–530.

von Kármán, T., Dunn, L. G., and Tsien, H. S. 1940, "The Influence of Curvature on the Buckling Characteristics of Structures," *J. Aerospace Sci.*, Vol. 7, pp. 276–289.

Koiter, W. T., 1963, "Elastic Stability and Post-Buckling Behavior," *Non Linear Problems*, R. E. Langer, ed., University of Wisconsin Press, Madison, Wisc.

Lindberg, H. E., and Herbert, R. E., 1966, "Dynamic Buckling of a Thin Cylindrical Shell Under Axial Impact," ASME Journal of Applied Mechanics, Vol. 33, pp. 105–112.

Lindberg, H. E., and Florence, A. L., 1987, *Dynamic Pulse Buckling—Theory and Experiment*, Martinus Nijhoff, Dordrecht, The Netherlands.

Lindberg, H. E., Rubin, M. B., and Schwer, L. E., 1987, "Dynamic Buckling of Cylindrical Shells from Oscillating Waves Following Axial Impact," *Int. J. Solids Structures*, Vol. 23, No. 6, pp. 669–692.

Roth, R. S., and Klosner, J., 1964, "Nonlinear Response of Cylindrical Shells Subjected to Dynamic Axial Loads," *Axial Journal*, Vol. 2, No. 10.

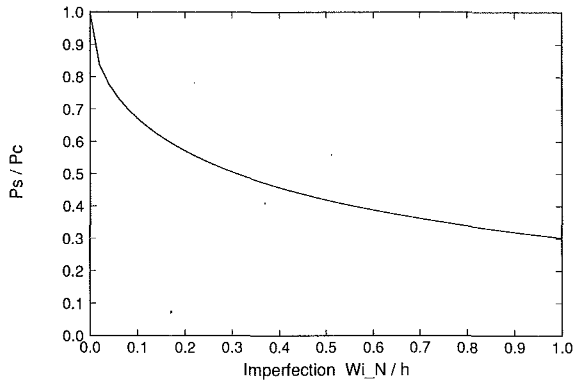


Fig. 3 Static buckling load reduction caused by nonlinear mode-coupling with imperfections in the square classical buckle mode

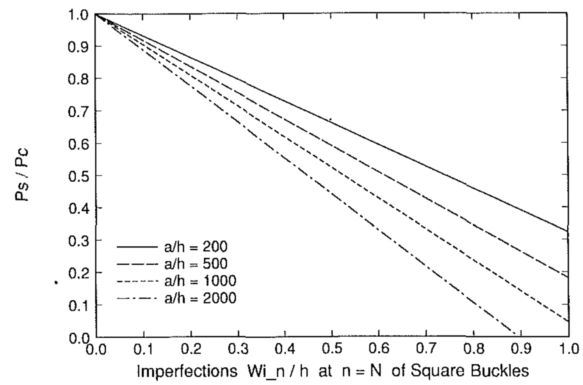


Fig. 4 Static buckling load reduction caused by local curvature reductions from modal imperfections varying as 1/n

APPENDIX

Static buckling loads are reduced by the combined effects of imperfections and nonlinearities, such as given by equation (32), and also by a direct change in the local curvature of the shell caused by imperfections. Since the latter allows straightforward use of the critical amplification criterion for $\sigma < \sigma_c$, it is useful to explore the magnitude of both mechanisms of static buckling load reduction.

Figure 3 gives the static buckling load reduction from imperfections in the square classical buckle mode, calculated from equation (32). The buckling load decreases abruptly for imperfections only a few percent of the wall thickness, and then decreases more slowly for larger imperfections. A reduction to 30 percent of the classical buckling load would require an imperfection equal to the wall thickness for this limited theory.

Reductions from curvature change are not as precipitous as for nonlinear response, but the reductions continue steadily with large imperfections rather than tapering off as for nonlinear response, so if imperfections are a substantial fraction of the wall thickness (crudely made shells, or shells damaged in service) the curvature change could become the dominant effect of imperfections.

The local curvature of the imperfect shell is given by

$$\frac{1}{a_s(x,\theta)} = \frac{1}{a} + \frac{w_i}{a^2} + \frac{\partial^2 w_i}{a^2 \partial\theta^2}. \tag{A1}$$

For this purpose, it is convenient to express the initial imperfections in the form

$$w_i(x,\theta) = a \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \gamma_{mn} \sin\frac{m\pi x}{L} \sin(n\theta + \phi_n) \tag{A2}$$

where $\gamma_{mn} = (a_{mn}^2 + b_{mn}^2)^{1/2}/a$ and ϕ_n is the phase angle of the nth-mode imperfection. With the w_i/a^2 term neglected as small, equation (A1) with equation (A2) gives

$$\frac{1}{a_s} = \frac{1}{a} - \frac{1}{a} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} n^2 \gamma_{mn} \sin\frac{m\pi x}{L} \sin(n\theta + \phi_n). \tag{A3}$$

Consider a specific location some distance from the end of the shell to avoid the complexity of summing over m, and then replace γ_{mn} by $\overline{\gamma}_n$. Also, from equation (16), the local static buckling load is $\sigma_s/\sigma_c = a/a_s$, so

$$\frac{\sigma_s}{\sigma_c} = 1 - \sum_{n=1}^{N/2} n^2 \overline{\gamma}_n \cos(n\theta + \phi_n). \tag{A4}$$

With static buckling in the square classical mode having $n = N$, imperfection modes up to only half this number are included in the final sum expression, to ensure that the local curvature encompasses a buckle. The half wavelength at $n = N$ is $2l_0$, so

$$\frac{\pi a}{N} = 2\pi\sqrt{ah}[12(1-\nu^2)]^{1/4} \tag{A5}$$

or, with $\nu = 0.3$,

$$N = 0.91\sqrt{a/h}. \tag{A6}$$

If one assumes that the imperfections are introduced by random processes, the phase angles ϕ_n can be taken as random, uncorrelated and uniformly distributed. The mean value of the curvature change is therefore the root-mean-square of the coefficients in equation (A4). Somewhere on the shell the curvature change will be as large as about three times this value, so

$$\left.\frac{\sigma_s}{\sigma_c}\right|_{\min} \approx 1 - 3\left[\sum_{n=1}^{N/2} (n^2 \overline{\gamma}_n)^2\right]^{1/2}. \tag{A7}$$

Data from Arbocz (1982) suggests that values for $\overline{\gamma}_n$ can be approximated by $\overline{\gamma}_n = A/n$, where A is about 0.0015. The formula applies only for $n > 8$, below which imperfections are nearly constant, but since $\overline{\gamma}_n$ is multiplied by n^2 in equation (A7) and $N/2 \gg 8$ for the thin shells discussed here, this cutoff is neglected for simplicity. Then, with $\overline{\gamma}_n = A/n$, equation (A7) becomes

$$\left.\frac{\sigma_s}{\sigma_c}\right|_{\min} \approx 1 - 3A\left[\frac{M(M+1)(2M+1)}{6}\right]^{1/2} \tag{A8}$$

in which the expression under the root is the sum of n^2 to $M = N/2$. For $a/h = 1000$, $N = 30$ and $\sigma_s/\sigma_c = 0.84$ for $A = 0.0015$. Thus, for shells of the type described in Arbocz (1982), reductions in buckling loads of about 10 to 20 percent are expected from local curvature changes. From equations (A6) and (A7), this reduction varies as $(a/h)^{1/2}$.

If imperfections are as large as required for substantial buckling load reduction in Fig. 3 from nonlinear effects, reductions from curvature changes are also substantial. This is shown in Fig. 4, which was constructed by again taking a $1/n$ variation of imperfection amplitudes and summing to $N/2$, but with $A = N\overline{\gamma}_N$, where $\overline{\gamma}_N = \zeta_2 h/a$ and ζ_2 is the value of the abscissa w_{iN}/h in Fig. 3. In evaluating equation (A8) for this plot, the 3-sigma factor was omitted, to be reasonably consistent with the effective value of ζ_2 in Fig. 3 being the highest local value for $n = N$, so ζ_2 already contains such a factor. Although this necessarily is a crude comparison because of the vagaries of imperfections, it is reasonable to conclude that for large imperfections curvature change effects must be considered in addition to nonlinear effects.

Structural Computation of an Assembly of Rigid Links, Frictionless Joints, and Elastic Springs

S. Pellegrino

C. R. Calladine

Department of Engineering,
University of Cambridge,
Cambridge CB2 1PZ, U.K.

The aim of the paper is to set up a scheme for efficient computation of the small-displacement response of a plane assembly of rigid links, frictionless joints, and elastic springs to static external forces applied at the joints. The particular assembly of Fig. 1 is used as an example. The conventional "stiffness method"—which becomes singular when, as here, the links are rigid—is abandoned in favor of a method which describes the current state of the assembly in terms of the amplitudes of m (here = 3) independent infinitesimal modes of inextensional deformation of the assembly; and the calculation boils down to the solving of an m×m (here 3×3) set of algebraic equations. The method is particularly straightforward if the inextensional modes (as here) may be obtained by inspection; but a general algorithm is presented for obtaining the inextensional modes of an arbitrary assembly of the same general kind. A major advantage over the conventional stiffness method—which requires, of course, the replacement of rigid links by (stiff) elastic members—is that the number of variables may be reduced substantially. This can be very important for large assemblies.

1 Introduction

Figure 1 shows a plane structural assembly. It consists of rigid links which are freely hinged at their ends to each other and to a rigid foundation, and are restrained from relative rotation at the joints by linear-elastic rotational springs. In the configuration shown, the assembly is stress-free. We wish to compute its response to arbitrary static forces ("loads") applied at the joints, under the general assumption that in the displaced configuration the rotations of the links are small. That is, we wish to compute the joint displacements, spring rotations, and moments of the assembly under an arbitrary loading. We are interested, of course, not only in this particular example but also in general assemblies of the same kind.

At first glance, this problem may appear to be straightforward. After all, practicing engineers use finite element packages on a routine basis to solve structural problems of much greater complexity than this.

A key feature of the assembly shown in Fig. 1 is that the links, or bar elements, are *rigid*. If, instead, these elements were *elastic*, so that they underwent small changes of length but no change of curvature when stressed, then more-or-less standard finite element methods would indeed be satisfactory. Thus, two components of small displacement could be defined for each of the four joints not connected to the ground; the elongations of the five bars and the rotations of the six springs could be expressed by a *compatibility matrix* in terms of the eight nodal displacements variables; and the remainder of the usual "stiffness" method of structural analysis (Livesley, 1975) could be implemented.

In the present example, however, the *inextensibility* of the bars stands in the way of progress along these lines (see Livesley and Charlton, 1955). The kinematic constraints provided by

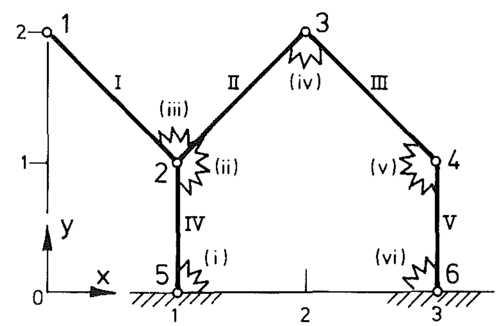


Fig. 1 Layout of example assembly of rigid rods and rotational springs in its original, unstressed configuration. Joints are labelled by large Arabic numerals, bars by large Roman numerals and springs by small Roman numerals. Small Arabic numerals mark divisions on the *x, y* coordinate grid.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 23, 1990; final revision, June 20, 1990.

the rigid links render the eight components of joint displacement nonindependent; and the resulting stiffness matrix becomes singular.

There is, of course, a well-known way around this difficulty, which has been available since the early days of structural computation. The stiffness method can be rescued by endowing the bars with elastic stiffness both high enough to simulate an almost-rigid bar and yet not so high as to render the final stiffness matrix badly ill conditioned.

The aim of the present paper is to overcome this difficulty over rigid links in a more radical, and ultimately more satisfactory, way. If we look at the present problem from the standpoint of classical kinematics, we see that the assembly, in the absence of the springs, has only three degrees-of-freedom as a mechanism. Thus, the state of the assembly can always be described completely in terms of three variables. Accordingly, it should be possible for us to solve this particular problem by dealing with a 3×3 matrix, instead of the 8×8 matrix of the usual stiffness method.

Our aim in this paper is to implement a computational scheme which exploits this idea, and which thus makes a virtue of that same rigidity of the links which proves a stumbling block for the usual method.

Our presentation will be in general terms, but with frequent reference to the particular assembly of Fig. 1 as an illustrative example. In Section 3 it will be assumed that the details of the independent mechanisms are known, and the task will be to perform the subsequent computations. This would be entirely satisfactory for the present example, in which three independent kinematic mechanisms may be found by inspection. In Section 4 we shall address the general problem of finding a set of kinematic mechanisms for a given assembly by a process of automatic computation. Again, we shall illustrate it with respect to the example of Fig. 1; but the real value of the algorithm, of course, lies in its power to deal with more complex assemblies, for which there is no simple intuitive way of establishing either the number of independent mechanisms or the details of them.

Our main motivation in this work is a desire to set up the most efficient scheme of computation for assemblies of this general kind, by making full use of the methods of kinematics. In one sense, the method of Section 3 is an adaptation from the well-known scheme of analysis of the plastic collapse of structural frameworks by the "method of combination of elementary mechanisms" (see, e.g., Baker and Heyman, 1969). There are, of course, many differences between our method and that of plastic collapse analysis; but the two methods have the common feature of making a virtue of the inextensibility of the component members. Our approach is closely related to some recent studies of constrained dynamical systems, as discussed in Section 5.

Part of our motivation in addressing the present problem lies in the possibility that the method may find a useful practical application in the field of molecular dynamics of long-chain molecules (McCammon and Harvey, 1987). In such calculations one conventionally allocates three degrees-of-freedom to each atom of the assembly; but if one takes account of the practical extensional rigidity of the covalent backbone links, it should be possible to reduce substantially the total number of degrees-of-freedom in general. The present paper may therefore be regarded as a first step in the process of producing a compaction of the degrees-of-freedom in a general calculation, by considering a simple analog problem.

2 Notation

Italic lowercase letters denote scalars; boldface lowercase letters denote vectors; and boldface uppercase letters denote matrices. In the list that follows, and indeed elsewhere in the paper, remarks in parentheses refer to the particular assembly shown in Fig. 1.

b = number of (rigid) bars ($b = 5$)

j = number of joints, excluding those embedded into the foundation ($j = 4$)

m = number of degrees-of-freedom of the assembly as an inextensional mechanism ($m = 3$)

s = number of rotational springs ($s = 6$)

$\mathbf{d}$ = nodal displacement vector, of size $2j(= 8)$: the joints are taken in numerical order, and with the x-component preceding the y-component

$\mathbf{e}$ = bar elongation vector, of size $b(= 5)$: the bars are taken in numerical order, and the condition $\mathbf{e} = 0$ corresponds to an inextensional mechanism

$\mathbf{l}$ = nodal force vector, also of size $2j(= 8)$ and with same numbering scheme as for $\mathbf{d}$

$\mathbf{q}$ = spring-moment vector, of size $s(= 6)$: the springs are taken in numerical order

$\mathbf{r}$ = spring-rotation vector, of size $s(= 6)$, and with the same numbering scheme as $\mathbf{q}$

$\boldsymbol{\phi}$ = generalized mode-displacement vector, of size $m(= 3)$

$\boldsymbol{\psi}$ = generalized mode-force vector, also of size $m(= 3)$

$\mathbf{C}$ = "compatibility" matrix which enables the elongations of the b bars to be computed for an arbitrary nodal displacement vector $\mathbf{d}$ by means of the equations $\mathbf{Cd} = \mathbf{e}$. The size of $\mathbf{C}$ is $2j \times b(= 8 \times 5)$. (The set of m-independent mechanisms of the assembly in $\mathbf{M}$ spans the null space of this matrix.)

$\mathbf{K}$ = diagonal matrix of spring stiffnesses, of size $s \times s(= 6 \times 6)$

$\mathbf{M}$ = matrix of $m(= 3)$ columns and $2j(= 8)$ rows, in which each column contains a set of joint-displacement components which constitute an independent inextensional mechanism

$\mathbf{R}$ = matrix of $m(= 3)$ columns and $s(= 6)$ rows, in which each column contains a set of spring rotation angles for the corresponding mechanism in $\mathbf{M}$

$\mathbf{S}$ = generalized $m \times m$ stiffness matrix ($= 3 \times 3$)

3 Structural Computation When the Mechanisms are Known

Figure 2 shows sketches of three inextensional modes of the assembly, which are clearly both independent and exhaustive. These modes (apart from a change of sign) were produced by the algorithm to be described below in Section 4; but in the present example they may be established by intuitive methods. The corresponding matrices $\mathbf{M}$ and $\mathbf{R}$ are given below. They contain details of the three independent mechanisms in terms of the eight components of joint displacement and the six components of rotation of the bars. In general, the magnitudes of the entries for a given mode are arbitrary; here one component of displacement has been assigned magnitude 1. Note that the bars have length 1 or $\sqrt{2}$ in this example: The rotations are calculated by small-displacement algorithms, which are given in detail in Section 4. The sign convention for spring rotation is as follows. Rotations of bars are reckoned positive when counterclockwise; and the rotation of a spring connected

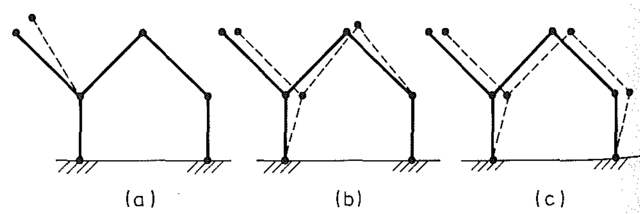


Fig. 2 Three independent modes of inextensional deformation of the assembly, as described by the three columns of matrix M. The amplitudes shown are arbitrary.

to two bars is defined as the rotation of the higher-numbered bar less that of the lower-numbered bar. The foundation counts as "bar 0" for the purposes of this calculation.

$$\mathbf{M}^T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \quad (1)$$

$$\mathbf{R}^T = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ -2 & -3 & 1 & -2 & 1 & 0 \\ -1 & -1 & 0 & 0 & -1 & -1 \end{bmatrix} \quad (2)$$

(It is obvious by inspection that the three columns of **M** are linearly independent.)

The most general kinematically admissible state of small displacement of the assembly is found by taking an arbitrary linear combination of the three mechanisms. Thus, we may write

$$\mathbf{d} = \mathbf{M}\phi \quad (3)$$

where ϕ is a "generalized displacement vector." The corresponding spring rotations are then given by

$$\mathbf{r} = \mathbf{R}\phi. \quad (4)$$

The external forces applied to the joints are defined by the load vector **l**, and the bending moments in the springs are given by the moment vector **q**. Thus **l**, **q** constitute an equilibrium set of external forces and internal moments, while **d**, **r** constitute a compatible set of external displacements and internal rotations. Thus, applying the principle of virtual work, we have

$$\mathbf{r}^T\mathbf{q} = \mathbf{d}^T\mathbf{l}. \quad (5)$$

Substituting for **d** and **r** from (3) and (4) we find, since ϕ is arbitrary,

$$\mathbf{R}^T\mathbf{q} = \mathbf{M}^T\mathbf{l}. \quad (6)$$

Now, the spring moments are given by the elastic law

$$\mathbf{q} = \mathbf{K}\mathbf{r}. \quad (7)$$

Thus, using (4) and (7) we find that (6) may be written

$$\mathbf{R}^T\mathbf{K}\mathbf{R}\phi = \mathbf{M}^T\mathbf{l}. \quad (8)$$

It is convenient to define the generalized (symmetric) stiffness matrix **S** by

$$\mathbf{S} = \mathbf{R}^T\mathbf{K}\mathbf{R}, \quad (9)$$

and also the generalized load vector ψ by

$$\psi = \mathbf{M}^T\mathbf{l}. \quad (10)$$

Our method of solution is as follows: Given **M**, ψ may be computed for any prescribed load vector **l**. Having evaluated the $m \times m$ square matrix **S** by (9), we then solve

$$\mathbf{S}\phi = \psi \quad (11)$$

for ϕ. Then, finally, we calculate **d** by means of (3). If **r** and **q** are also required, we find them by use of (4) and (7).

For example, suppose that we are required to find the components of joint displacement of the given assembly when loading is applied as shown in Fig. 3, and when all six rotational springs have equal stiffness k.

First we evaluate

$$\mathbf{S} = \mathbf{R}^T\mathbf{K}\mathbf{R} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 19 & 4 \\ 0 & 4 & 4 \end{bmatrix} k. \quad (12)$$

From Fig. 3, we have

$$\mathbf{l}^T = [0 \quad -2 \quad 0 \quad 0 \quad 3 \quad -1 \quad -1 \quad 0]. \quad (13)$$

Hence, from (10)

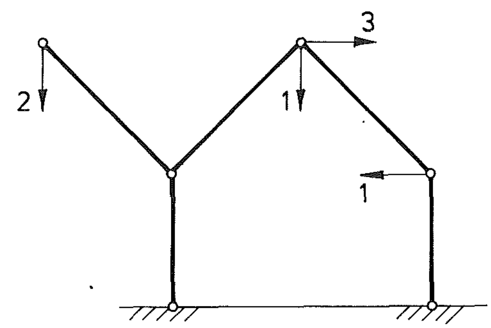


Fig. 3 System of loads applied to the joints in the worked example

$$\psi = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}. \quad (14)$$

Solving (11) we find

$$\phi = \begin{bmatrix} -2.143 \\ 0.143 \\ 0.357 \end{bmatrix} \frac{1}{k} \quad (15)$$

from which, by (3),

$$\mathbf{d}^T = [-1.5 \quad -2.143 \quad 0.643 \quad 0 \quad 0.5 \quad 0.143 \quad 0.357 \quad 0]/k. \quad (16)$$

It is interesting to note that if the loading vector were to be replaced by, say,

$$\mathbf{l}^T = [-2 \quad 0 \quad 0 \quad 0 \quad 5 \quad 1 \quad -1 \quad 0], \quad (17)$$

ψ, ϕ, and **d** would not be changed. This is because the difference between the two loading cases is a "balanced" loading which may readily be shown to require, for equilibrium, tension in some bars, but no bending moments.

4 Determination of the Inextensional Modes of an Arbitrary Assembly

It is obvious that our intuitive ability to "spot" three independent inextensional mechanisms greatly reduced the labor of computation for the example of Fig. 1. In general, for an arbitrary assembly, we must not expect to be able to obtain either the number or the details of the inextensional mechanisms by such methods. Thus we need, in general, an automatic procedure for generating the inextensional mechanisms of a given assembly. The following account of such a procedure is an adaptation of that which is described fully by Pellegrino and Calladine (1986).

First we need to set up the compatibility matrix **C** which expresses the elongations **e** of the $b(=5)$ bars in terms of the $2j(=8)$ components of displacement of the joints:

$$\mathbf{C}\mathbf{d} = \mathbf{e}. \quad (18)$$

This equation is simply a collection of the small-displacement calculations of the elongation of each bar in turn, in terms of the components of displacement at its end. Figure 4 sets the scene for calculating the elongation of a typical bar which connects joint i to joint k. The initial coordinates of the joints and the components of small displacement are defined in the diagram.

Suppose that the only nonzero displacement component were d_{xk}. Then, by the geometry of small displacements the elongation of the bar would be given by

$$e = d_{xk}\cos\alpha \quad (19)$$

where

$$\cos\alpha = (x_k - x_i)/a \quad (20)$$

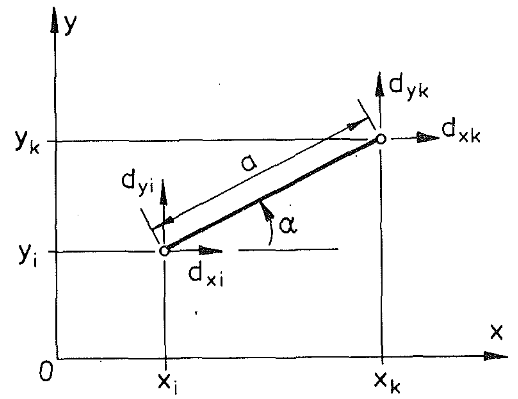


Fig. 4 Typical bar connecting joints i and k, showing original coordinates and components d_{xi}, etc., of small displacement

and the length a is calculated by Pythagoras' theorem. Similarly, d_{yk} alone would give

$$e = d_{yk}\sin\alpha. \tag{21}$$

In this way we find, in general,

$$e = (d_{xk} - d_{xi})\cos\alpha + (d_{yk} - d_{yi})\sin\alpha. \tag{22}$$

Expressions of this sort are written for all bars in order, and the resulting equations are collected as

$$\mathbf{Cd} = \mathbf{e}. \tag{23}$$

For example, we may readily show that the assembly of Fig. 1 has

$$\mathbf{C} = \begin{bmatrix} -0.7 & 0.7 & 0.7 & -0.7 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.7 & -0.7 & 0.7 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.7 & 0.7 & 0.7 & -0.7 \\ 0 & 0 & 0 & 1.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \end{bmatrix} \tag{24}$$

Here, 0.7 stands for $0.7071\ldots = 1/\sqrt{2}$.

Now we are particularly interested in the circumstances in which $\mathbf{d}$ corresponds to $\mathbf{e} = 0$. Such a displacement vector characterizes an *inextensional mode*. The first stage in the computation of the inextensional modes is to perform a Gaussian elimination on matrix $\mathbf{C}$ (see Strang, 1976). The aim is to combine rows in a linear fashion so that the matrix is transformed to the "echelon" form, with ones on the "leading diagonal" and zeros in the triangle below. In the present example it is easy to verify that the echelon form of the matrix is:

$$\begin{array}{cccccccc} & * & & & & * & * & \\ \multicolumn{8}{c}{\left[\begin{array}{cccccccc} 1 & -1 & 0 & 0 & 0 & -2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -2 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]} \\ (1) & (2) & (3) & (4) & (5) & (6) & (7) & (8) \end{array}. \tag{25}$$

The zigzag line marks the upper extent of zeros in the lower left-hand corner. Note that this "descending staircase" is irregular, and that the columns without pivot, i.e., those columns where there is no descending step, have been distinguished by an asterisk. It is shown in Strang (1976) that the corresponding columns of the original matrix (24) are linearly dependent on the other columns, viz. 1, 3, 4, 5, 8, which are themselves linearly independent. This gives a partitioning of the eight variables into five independent and three dependent ones. Precisely which emerge as which depends on the (arbitrary) numbering system used to identify the bars and joints and also on the pivoting strategy; but the null space of the matrix $\mathbf{C}$ (see below) is unaffected.

The three columns marked * have positions 2, 6, and 7. An empty 8×3 matrix $\mathbf{M}$ is prepared, into which the three required vectors will go. The entries are made in stages as follows:

$$\begin{bmatrix} -1 & -2 & -1 \\ * & * & * \\ 0 & -2 & -1 \\ 0 & 0 & \cdot 0 \\ 0 & -1 & -1 \\ * & * & * \\ * & * & * \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} -1 & -2 & -1 \\ -1 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

(*a*) (*b*) (*c*)

Fig. 5

First, rows number 2, 6, and 7—corresponding to *—are "blanked off," as shown by *** in Fig. 5(*a*). The five entries from each of the columns marked * in (25) are then entered, in the same order, into the spaces which are still vacant, as shown.

Second, a 3×3 diagonal matrix—3 being the number of columns marked *—with all nonzero entries—1 is set up, as shown in Fig. 5(*b*).

Finally, the 9 entries in Fig. 5(*b*) are fitted, in the same order, into the blanked-off spaces of Fig. 5(*a*). This gives the required matrix $\mathbf{M}$, as shown in Fig. 5(*c*).

What we have done here is to find three independent solutions of the equations $\mathbf{Cd} = \mathbf{0}$. Each solution has been arrived at by giving zero values to two of the three dependent variables and the value—1 to the third. The remaining independent variables are then obtained by back substitution, which is trivial in this case. This strategy is the simplest way of generating a basis $\mathbf{M}$ for the null space of matrix $\mathbf{C}$. In physical terms it can be seen that any linear combination of the three columns of $\mathbf{M}$ satisfies $\mathbf{Cd} = \mathbf{0}$, and hence is an inextensional mode for the assembly.

Returning to our particular example, we see that all of the entries in $\mathbf{M}$ are negative, and so the signs of $\mathbf{M}$ in (1) have been reversed.

Having obtained the m columns of $\mathbf{M}$, we must now find the corresponding columns of $\mathbf{R}$. To do this, we must calculate the relative rotation of the two bars to which each spring is attached for each particular mechanism. This is straightforward once we have an algorithm for calculating the rotation r of the typical bar shown in Fig. 4, in terms of the four components of displacement.

Suppose that the only nonzero displacement component were d_{xk}; then the counterclockwise rotation of the bar would be given by

$$r = -d_{xk}\sin\alpha/a. \tag{26}$$

Similarly, d_{yk} alone would give

$$r = d_{yk}\cos\alpha/a. \tag{27}$$

In this way we find, in general,

$$r = ((-d_{xk} + d_{xi})\sin\alpha + (d_{yk} - d_{yi})\cos\alpha)/a. \tag{28}$$

5 Discussion

A particular example, such as the one which we have used, has the advantage of providing a sense of definiteness to the calculations. But it has the disadvantage that it may conceal certain points which may occur in general application.

Thus, in the present calculation we found that

$$m=2j-b \quad (3=8-5); \tag{29}$$

and indeed this seems obvious, since the addition of each bar removes one degree-of-freedom from the set of j isolated joints. However, relation (29) is not universally true for two-dimensional assemblies of this sort; in particular, it will be violated if the assembly contains one or more *states of self-stress*. In this case the matrix **C** will not be of full rank and the process of Gaussian elimination will "run out," leaving one or more zero rows at the bottom of the transformed matrix. Points of this sort have been dealt with fully by Pellegrino and Calladine (1986). They are revealed clearly by the process of Gaussian elimination.

In this paper we have dealt exclusively with two-dimensional assemblies. The methods of the paper may be extended to assemblies in three-dimensional Euclidean space. We are currently working on a scheme which will enable us to deal with assemblies which include, *inter alia*, rigid members containing "dog-leg" bends, and linear springs which connect arbitrary points.

We have also considered in this paper only the analysis of assemblies in "small deflection" conditions. We believe that relatively simple iterative schemes will enable us to extend our work to "large deflection" conditions in cases where this is warranted.

The present work has grown out of our earlier paper (Pellegrino and Calladine, 1986). From a computational point of view the main difference is that here we compute a full set of independent mechanisms directly from the "compatibility" matrix **C**, whereas in the earlier paper we performed manipulations on the augmented matrix $[\mathbf{C}^T/\mathbf{I}]$. The present algorithm is somewhat more elaborate, requiring a few more steps and a few more lines of computer program; but since it operates with smaller matrices there is likely to be a net saving of computational effort and storage. The previous paper dealt with situations in which both the null space and the left null space (Strang, 1976) are required; but here only the null space is needed and so the computations can be curtailed.

Ider and Amirouche (1988) have recently devised a method for reducing the number of degrees-of-freedom in the analysis of constrained dynamical systems. Their work was done quite independently of ours (Pellegrino and Calladine, 1986), but their method is in some respects rather similar to ours. For example, they use a method of Gaussian elimination on an augmented matrix $[\mathbf{C}^T/\mathbf{I}]$.

Wehage and Haug (1982) have devised computational methods for studying the dynamics of mechanical assemblies consisting of rigid bodies connected by a variety of undeformable links, elastic springs, and viscous dashpots. In the course of this work they separate the kinematic variables into dependent and independent kinds, by means of a Gauss-Jordan transformation. This separation of variables keeps the calculation well conditioned, but it does not directly make use of the inextensional modes.

Orthogonal decomposition algorithms, including QR and SVD methods, have also been used in the study of constrained dynamical systems (see, eg., Kamman and Huston, 1984; Singh and Likins, 1985; Kim and Vanderploeg, 1986). These methods are more powerful and accurate than the Gaussian elimination scheme which we have described in Section 4. But in examples where rank deficiency is not likely to be a problem—as in linkages of the kind shown in Fig. 1—Gaussian elimination is satisfactory. The method also works in problems with rank deficiency, eg, involving prestressable linkages, as in Pellegrino and Calladine (1986); but if ambiguity over small pivots is likely to create difficulties, then, as Golub and Van Loan (1983) have advised, it would be better to use SVD. However, once the matrix **M** has been obtained—by whatever method—the calculations of Section 3 proceed unchanged.

References

Baker, J. F., and Heyman, J., 1969, *Plastic Design of Frames*, Cambridge University Press.

Golub, G. H., and Van Loan, C. F., 1983, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Md.

Ider, S. K., and Amirouche, F. M. L., 1988, "Coordinate Reduction in the Dynamics of Constrained Multibody Systems: A New Approach," ASME Journal of Applied Mechanics, Vol. 55, pp. 899–904.

Kamman, J. W., and Huston, R. L., 1984, "Dynamics of Constrained Multibody Systems," ASME Journal of Applied Mechanics, Vol. 51, pp. 899–903.

Kim, S. S., and Vanderploeg, M. J., 1986, "QR Decomposition for State Space Representation of Constrained Mechanical Dynamic Systems," *ASME Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 108, pp. 183–188.

Livesley, R. K., 1975, *Matrix Methods of Structural Analysis*, 2nd ed., Pergamon, Oxford, U.K.

Livesley, R. K., and Charlton, T. M., 1955, "The Use of a Digital Computer with Particular Reference to the Analysis of Structures," *Transactions of the North-East Coast Institution of Engineers and Shipbuilders*, Vol. 71, pp. 76–89.

McCammon, J. A., and Harvey, S. C., 1987, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press.

Pellegrino, S., and Calladine, C. R., 1986, "Matrix Analysis of Statically and Kinematically Indeterminate Frameworks," *International Journal of Solids and Structures*, Vol. 22, pp. 409–428.

Singh, R. P., and Likins, P. W., 1985, "Singular Value Decomposition for Constrained Dynamical Systems," ASME Journal of Applied Mechanics, Vol. 52, pp. 943–948.

Strang, G., 1976, *Linear Algebra and its Applications*, Academic Press, New York.

Wehage, R. A., and Haug, E. J., 1982, "Generalized Coordinate Partitioning for Dimensional Reduction in Analysis of Constrained Dynamical Systems," *ASME Journal of Mechanical Design*, Vol. 104, pp. 247–255.

Predicting Rebounds Using Rigid-Body Dynamics

Charles E. Smith
Department of Mechanical Engineering,
Oregon State University,
Corvallis, OR 97331
Mem. ASME

The observation by Thomas Kane a few years ago, that long-used relationships for predicting post-collision motion of a system of rigid bodies can imply a significant increase in kinetic energy during collision, has revived interest in this type of problem. This paper is intended to clarify understanding of the sources of this difficulty, and to suggest an alternative to some of the previously used assumptions for making such predictions. An organization of the pertinent equations of kinetics is presented, which provides a more direct means of examining the aforementioned question and of obtaining rebound predictions.

Introduction

Collisions of bodies within mechanical systems often do not deform any of the bodies significantly, so that equations of rigid body mechanics have long been used for predicting post-collision motion. Beyond the assumption of rigidity of components, further simplification is possible if the configuration may be expected to undergo little change while velocities undergo the changes necessary for separation at the point of collision. Thus, there are many situations in which the common assumption of constant configuration during contact would not be responsible for serious discrepancies in the prediction of the rebound. In the absence of detailed knowledge of the deformations induced by the impulsive reaction force where the bodies contact one another, additional assumptions about the nature of the reaction must be made, since the equations of rigid-body kinetics are three too few to predict the impulse and the velocity changes. These assumptions are based on speculations about such things as sliding with friction and the capacity of the bodies to return energy of deformation. Because the post-collision motion depends so heavily on the unknown impulse, the assumptions that form a "contact law," to supplement equations of rigid body mechanics, have a profound effect on the predicted motion.

That the implied change of kinetic energy should be non-positive provides a boundary for the impulse (or, equivalently, for the changes in velocities). However, a great deal of latitude remains within this boundary, so that satisfaction of the laws of thermodynamics and impulse-momentum relationships is not sufficient to assure an accurate prediction.

In spite of this hazard, an alternative to previously used contact laws is offered here. For collisions of solid, elastic spheres, for which a fairly detailed study of the mechanics of local deformation and sliding has been made, the proposed assumption appears to lead to improvement in the prediction. However, before it is accepted for a wide class of collisions, results need to be compared with those from experiments and analyses that model the transient behavior in sufficient detail.

Impulse, Momentum, and Kinetic Energy

Relationships to be used follow from those developed in Kane and Levinson (1985); for completeness, this development is outlined here. Two rigid bodies B and B' collide as the points P and P', on their respective surfaces, move into coincidence. Denoting by u_r ($r=1, 2, \ldots, n$) a set of generalized speeds for the system, the velocities of the contact points may be written as:

$$\mathbf{v}^P=\sum_r \mathbf{v}_r^P u_r \quad \mathbf{v}^{P'}=\sum_r \mathbf{v}_r^{P'} u_r.$$

It is convenient to deal with the difference between these velocities, defined as

$$\mathbf{v} \triangleq \mathbf{v}^P-\mathbf{v}^{P'}.$$

With the additional definition $\mathbf{v}_r \triangleq \mathbf{v}_r^P-\mathbf{v}_r^{P'}$, the velocity difference may be expressed in terms of the generalized speeds as

$$\mathbf{v}=\sum_r \mathbf{v}_r u_r. \tag{1}$$

With the impulse of the force exerted on B by B', denoted as $\mathbf{g}$, and with changes in configuration and contributions from forces other than the action-reaction at the contact point neglected, the components of generalized impulse can be expressed as

$$I_r=\mathbf{v}_r\cdot\mathbf{g}. \tag{2a}$$

By expressing the kinetic energy in terms of the selected generalized speeds, the inertia coefficients m_{rs} can be evaluated from[1]

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Dec. 28, 1989; final revision, May 15, 1990.

[1]Situations in which K contains terms that are linear in generalized speeds will not be considered here.

$$K = \frac{1}{2} \sum_r \sum_s m_{rs} u_r u_s \tag{3}$$

from which expressions for components of generalized momentum can be written as

$$p_r = \sum_s m_{rs} u_s. \tag{4}$$

Under the aforementioned assumptions, the impulse-momentum laws can be expressed as

$$I_r = \Delta p_r = \sum_s m_{rs} \Delta u_s \tag{5a}$$

wherein Δu_s denotes the change in u_s that occurs during contact. From here on, $\mathbf{v}$ will denote the value of $\mathbf{v}^P - \mathbf{v}^{P'}$ at the time contact begins, and $\mathbf{w}$ will denote the value of $\mathbf{v}^P - \mathbf{v}^{P'}$ at the time contact ends. Thus,

$$\mathbf{w} = \mathbf{v} + \Delta\mathbf{v} \tag{6}$$

in which

$$\Delta\mathbf{v} = \sum_r \mathbf{v}_r \Delta u_r. \tag{7a}$$

Three of the above relationships can be organized to our advantage using matrices defined as follows. Let m be the symmetric $n \times n$ matrix of inertia coefficients m_{rs} and I, and Δu be column matrices with elements I_r and Δu_r, respectively; introduce a set of mutually perpendicular, unit vectors $\mathbf{e}_i$ $(i = 1,2,3)$; let l be the $n \times 3$ matrix having elements $\mathbf{v}_r \cdot \mathbf{e}_i$; let Δv and g be column matrices with elements $\mathbf{e}_i \cdot \Delta\mathbf{v}$ and $\mathbf{e}_i \cdot \mathbf{g}$, respectively. Then, equations (2a), (5a), and (7a) can be written as

$$I = lg \tag{2b}$$

$$I = m\Delta u \tag{5b}$$

$$\Delta v = l^T \Delta u \tag{7b}$$

and combined, with the result

$$\mathbf{g} = \mathbf{M} \cdot \Delta\mathbf{v} \tag{8}$$

in which $\mathbf{M}$ is the symmetric dyadic having components in the $\mathbf{e}_i$ basis that are the elements in the matrix

$$M = (l^T m^{-1} l)^{-1}. \tag{9}$$

The inertia operator $\mathbf{M}$ depends on the configuration of the system at the time of contact, but not on the motion. Also, if the configuration does not change significantly during impact, the small dynamic deformations during contact, and consequently $\mathbf{g}$, may be expected to depend on $\mathbf{v}$, but not on the particular set of generalized speeds that contribute to $\mathbf{v}$. That is, all precontact motions having the approach velocity $\mathbf{v}$ and the same configuration at the initiation of contact will result in the same impulse and corresponding separation velocity $\mathbf{w}$. Once $\Delta\mathbf{v}$ has been determined, changes in the generalized speeds can be evaluated from

$$\Delta u = m^{-1} l M \Delta v, \tag{10}$$

and corresponding changes in velocities and angular velocities of interest can be evaluated using the appropriate partial velocities and partial angular velocities.

The change in kinetic energy induced by the impulse is given by

$$\Delta K = \frac{1}{2}(u + \Delta u)^T m (u + \Delta u) - \frac{1}{2} u^T m u$$

and, with the help of the above equations, can be expressed also as follows:

$$\Delta K = \mathbf{g} \cdot \mathbf{v} + \frac{1}{2} \mathbf{g} \cdot \mathbf{M}^{-1} \cdot \mathbf{g} \tag{11a}$$

$$\Delta K = \mathbf{g} \cdot \frac{\mathbf{v} + \mathbf{w}}{2} \tag{11b}$$

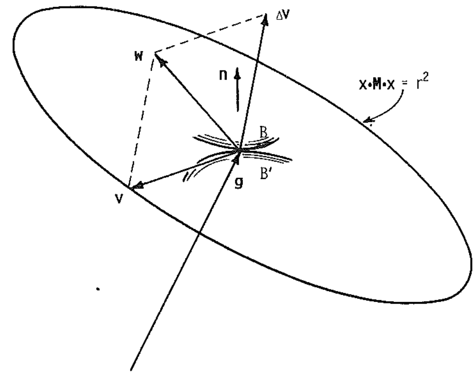


Fig. 1 Impulse, velocities of approach and separation, and boundary for change in kinetic energy, for a typical collision

$$\Delta K = \frac{1}{2}(\mathbf{w} \cdot \mathbf{M} \cdot \mathbf{w} - \mathbf{v} \cdot \mathbf{M} \cdot \mathbf{v}). \tag{11c}$$

Along with the last equation, the Cauchy quadric surface associated with $\mathbf{M}$ provides a convenient means for visualizing the constraint that the predicted change in kinetic energy should be nonpositive. With $\mathbf{x}$ denoting the position vector of a point on the surface, the equation for the quadric is

$$f(\mathbf{x}) = \mathbf{x} \cdot \mathbf{M} \cdot \mathbf{x} = r^2.$$

Because f is positive definite (see equations (3) and (9)), the quadric is an ellipsoid. If the scaling factor r is chosen so that $\mathbf{v}$, with its tail placed at the origin, has its head coincident with a point on the quadric surface, then $\mathbf{w}$ must lie within the ellipsoid if its tail is also placed at the origin. Equation (11c) shows further that the largest possible loss of kinetic energy would occur for $\mathbf{w} = \mathbf{0}$, i.e., if the impact ended with P and P' moving at the same velocity. For a given $\mathbf{v}$ and $\mathbf{w}$, the ellipsoid also indicates the direction of the impulse, since $\mathbf{g}$ must be perpendicular to the surface where the line parallel to $\Delta\mathbf{v}$ and through the origin intersects the surface. These properties are illustrated in the two-dimensional section shown in Fig. 1.

To facilitate formulation of a contact law, one of the basis vectors is chosen to be perpendicular to the surfaces at contact; specifically, let $\mathbf{n}$ be a unit vector perpendicular to the common tangent to the surfaces at P and P' and directed from B' into B, and let $\mathbf{t}_1$ and $\mathbf{t}_2$ be unit vectors parallel to this tangent plane and satisfying $\mathbf{n} = \mathbf{t}_1 \times \mathbf{t}_2$. Normal and tangential components of impulse will then be denoted as

$$\mathbf{g}_n = \mathbf{n} \cdot \mathbf{g} \mathbf{n}$$

$$\mathbf{g}_t = \mathbf{n} \times (\mathbf{g} \times \mathbf{n}) = g_1 \mathbf{t}_1 + g_2 \mathbf{t}_2$$

and those of $\mathbf{v}$ and $\mathbf{w}$ denoted similarly.

Contact Assumptions

Equation (8) provides three relationships among the six unknown components of $\mathbf{g}$ and $\Delta\mathbf{v}$. In the absence of detailed analysis of surface forces in the region of contact and related deformations, assumptions about impulse and relative motion are needed to provide the additional three equations from which $\mathbf{g}$ and $\mathbf{w}$ can be predicted.

A common assumption is that the ratio e of the normal component of $\mathbf{w}$ to the normal component of $\mathbf{v}$ is known, so that

$$\mathbf{w}_n = -e\mathbf{v}_n \tag{12}$$

can be used directly in the calculation. This ratio is related quite simply to the loss of kinetic energy in impacts in which the tangential component of impulse is absent, for example, when the surfaces are perfectly smooth. Equations (8) and (11) can be used to show that, in these cases,

$$\Delta K = -\frac{(1-e^2)\,v_n^2}{2\mathbf{n}\cdot\mathbf{M}^{-1}\cdot\mathbf{n}}. \tag{13}$$

Thus, in the absence of tangential impulse, e^2 cannot exceed 1. Perhaps this relationship, or special cases, have led to the common belief that, in general, $0 \le e \le 1$.[2] Although the lower bound must hold in the absence of penetration, the laws of thermodynamics do not preclude occurrence of a rebound with $e > 1$ when tangential impulse is possible, as Fig. 1 illustrates.

Additional assumptions are usually made in terms of the coefficient of friction μ between the contacting surfaces. An ingenious analysis of sliding and sticking, under the assumption that the tangential force obeys Coulomb's law of friction, is presented in Routh (1905). This analysis deals with two phases of the contact, a "compression" phase in which the normal velocity difference passes from $\mathbf{v}_n$ to $\mathbf{0}$, and a "restitution" phase in which the normal velocity difference passes from $\mathbf{0}$ to $\mathbf{w}_n$. In this analysis Routh defines the coefficient of restitution as the ratio of normal impulse during restitution to normal impulse during compression. In some cases this is equal to e as defined previously, but in others it is not, despite the contrary statement in Brach (1989). An assumption implicit in Routh's analysis is that the tangential velocities are as given by rigid body kinematics, i.e., that sliding or the lack of sliding is unaffected by deformations. The importance of such deformations is indicated in Maw et al. (1976). Keller (1986) has recast Routh's analysis in more modern nomenclature and divorced the procedure from the graphical guidance presented in Routh (1905). Avoiding much of the detail considered by Routh, Whittaker (1904) assumes that $\mathbf{w}_t$ is zero if the magnitude of the tangential impulse is less than μ times the magnitude of the normal impulse, and that when $\mathbf{w}$ has a tangential component, the magnitude of the tangential impulse will equal μ times the magnitude of the normal impulse. Kane and Levinson (1985) use the same criteria, but distinguish between coefficients of static and kinetic friction and are more specific about direction, stating that if and only if

$$|\mathbf{g}_t| \le \mu_o |\mathbf{g}_n|, \tag{14a}$$

then $\mathbf{w}_t = \mathbf{0}$, and if the inequality is violated,

$$\mathbf{g}_t = -\mu |\mathbf{g}_n| \frac{\mathbf{w}_t}{|\mathbf{w}_t|} \tag{14b}$$

in which μ_o is the coefficient of static friction and μ is the coefficient of kinetic friction.

These relationships determine the tangential impulse solely from the normal impulse and the separation velocity. However, if the tangential velocity undergoes a change in direction during contact, a more reasonable speculation would be that the tangential impulse would be proportional to some kind of average of the tangential components of approach and separation velocities. This line of thought leads to the following relationship, which will agree with the last equation above when the directions of $\mathbf{v}_t$ and $\mathbf{w}_t$ are the same, but gives a more realistic estimate of the rebound when they differ.

$$\mathbf{g}_t = -\mu |\mathbf{g}_n| \frac{|\mathbf{v}_t|\mathbf{v}_t + |\mathbf{w}_t|\mathbf{w}_t}{|\mathbf{v}_t|^2 + |\mathbf{w}_t|^2} \tag{15}$$

This, together with a reasonable estimate of the normal velocity ratio e, defined by equation (12), and the impulse-momentum relationship (8), permit estimation of the impulse and corresponding separation velocity.

A bound on the loss of kinetic energy implied by (12) and (15) can be obtained by combining these equations with equation (11b). This leads to the relationship

$$\Delta K \le -\frac{1}{2}|\mathbf{g}_n|\left[(1-e)|\mathbf{v}_n| + \mu\frac{(|\mathbf{v}_t| + |\mathbf{w}_t|)(|\mathbf{v}_t| - |\mathbf{w}_t|)^2}{|\mathbf{v}_t|^2 + |\mathbf{w}_t|^2}\right]$$

which shows that, with $e \le 1$, these equations will not predict a gain in kinetic energy. Of course, it remains to be determined whether an improvement in the accuracy of the prediction can be expected.

Collision of Two Unconstrained Spheres

A collison between two spheres illustrates the use of the above relationships and, because a detailed analysis of elastic deformation and interface slipping has been carried out (Maw et al., 1976), provides an example to begin an evaluation of the proposed contact assumption (15).

The masses, radii, and central radii of gyration of the spheres B and B' will be designated as m_B and $m_{B'}$, b and b' and k and k', respectively. With the basis vectors defined as above, generalized speeds are introduced as follows, to specify the velocities of the centers C and C' and the angular velocities of the spheres.

$$\mathbf{v}^c = \mathbf{t}_1 u_1 + \mathbf{t}_2 u_2 + \mathbf{n} u_3 \quad \omega^B = \mathbf{t}_1 u_4 + \mathbf{t}_2 u_5 + \mathbf{n} u_6$$

$$\mathbf{v}^{c'} = \mathbf{t}_1 u_7 + \mathbf{t}_2 u_8 + \mathbf{n} u_9 \quad \omega^{B'} = \mathbf{t}_1 u_{10} + \mathbf{t}_2 u_{11} + \mathbf{n} u_{12}$$

Then the velocity difference can be expressed as

$$\mathbf{v} = \mathbf{t}_1 u_1 + \mathbf{t}_2 u_2 + \mathbf{n} u_3 + b\mathbf{t}_2 u_4 - b\mathbf{t}_1 u_5 + \mathbf{0} u_6 - \mathbf{t}_1 u_7 - \mathbf{t}_2 u_8 - \mathbf{n} u_9 + b'\mathbf{t}_2 u_{10} - b'\mathbf{t}_1 u_{11} + \mathbf{0} u_{12}$$

from which

$$l^T = \begin{bmatrix} 1 & 0 & 0 & 0 & -b & 0 & -1 & 0 & 0 & 0 & -b' & 0 \\ 0 & 1 & 0 & b & 0 & 0 & 0 & -1 & 0 & b' & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}.$$

From the expression for kinetic energy,

$$K = \frac{1}{2} m_B[u_1^2 + u_2^2 + u_3^2 + k^2(u_4^2 + u_5^2 + u_6^2)] + \frac{1}{2} m_{B'}[u_7^2 + u_8^2 + u_9^2 + k'^2(u_{10}^2 + u_{11}^2 + u_{12}^2)]$$

the matrix m is readily identified and, with the previous equation, used to construct the matrix of the impulse-velocity change relationship (8):

$$M = \overline{m}\begin{bmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

in which

$$\overline{m} \triangleq \frac{m_B m_{B'}}{m_B + m_{B'}} \text{ and } \kappa \triangleq \frac{(m_B + m_{B'})k^2 k'^2}{m_B k^2(k'^2 + b'^2) + m_{B'} k'^2(k^2 + b^2)}.$$

Two special cases may be of interest. When the spheres are identical, $\overline{m} = 1/2\ m_B$, and as the ratio of the mass of B to that of B' approaches zero (as in the case of a tennis ball striking a court surface), $\overline{m} = m_B$. In both of these cases, $\kappa = k^2/(k^2 + b^2)$.

The absence of coupling within the tangent plane permits reduction of the analysis of the impact to one of two dimensions, unless a friction or deformation characteristic implies an anisotropic contact law. To take advantage of this, $\mathbf{t}_1$ and $\mathbf{t}_2$ can be replaced with a unit vector $\mathbf{t}$, defined to be in the direction of $\mathbf{v}_t$, and $\mathbf{s} = \mathbf{t} \times \mathbf{n}$. With the angle between $\mathbf{t}_2$ and $\mathbf{t}$ denoted as γ, this change of basis may be expressed as

$$\mathbf{t}_1 = \cos\gamma\,\mathbf{s} + \sin\gamma\,\mathbf{t}$$

$$\mathbf{t}_2 = -\sin\gamma\,\mathbf{s} + \cos\gamma\,\mathbf{t}$$

and the approach velocity as

[2] Brach (1989) indicates that the reasons for the limitations on e are kinematic, but does not explain.

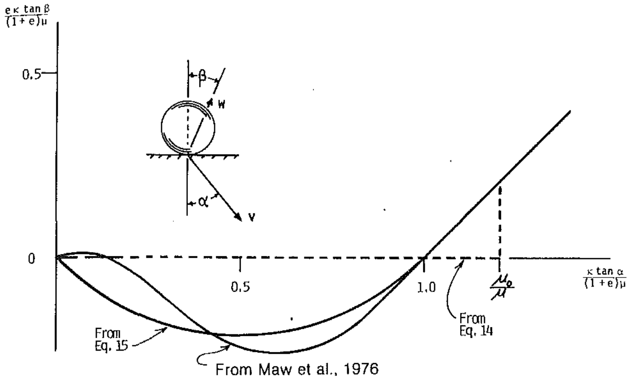


Fig. 2 Rebound predictions for colliding spheres

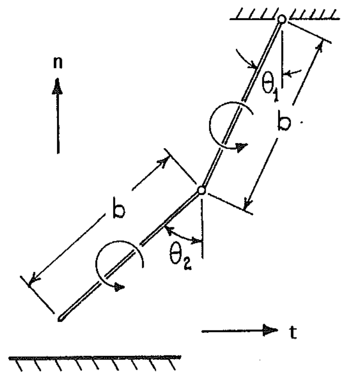


Fig. 3 Kane's example

$$\mathbf{v} = v_t\mathbf{t} + v_n\mathbf{n}, \ v_t \geq 0, \ v_n < 0.$$

The impulse-velocity change law becomes

$$\begin{bmatrix} g_s \\ g_t \\ g_n \end{bmatrix} = \overline{m} \begin{bmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta v_s \\ \Delta v_t \\ \Delta v_n \end{bmatrix} \quad (16)$$

and the first row will be satisfied by $g_s = \Delta v_s = 0$, as will the contact law if it is isotropic within the tangent plane. Equations (12), (15), and the last two of (16) can then be combined to yield

$$a\,\frac{v_t - w_t}{v_t} = \frac{v_t^2 + |w_t| w_t}{v_t^2 + w_t^2}$$

in which

$$a \triangleq \frac{\kappa \tan\alpha}{(1+e)\mu}$$

and α is the angle between $\mathbf{v}$ and $-\mathbf{n}$. This can be solved to give the tangential component of separation velocity, as

$$\frac{w_t}{v_t} = \begin{cases} \dfrac{1}{2a}\,[1 - \sqrt{1 + 4a(1-a)}] & 0 < a < 1 \\ 1 - \dfrac{1}{a} & a \geq 1. \end{cases} \quad (17)$$

If the criteria (14) are used instead of equation (15), the result is

$$\frac{w_t}{v_t} = \begin{cases} 0 & 0 \leq a \leq \dfrac{\mu_o}{\mu} \\ 1 - \dfrac{1}{a} & a > \dfrac{\mu_o}{\mu}. \end{cases}$$

This, together with equation (12), gives the separation velocity in terms of e and the approach velocity. Other quantities of interest may be evaluated from equation (10), which gives the changes in motion of B as

$$\begin{bmatrix} \Delta u_1 \\ \Delta u_2 \\ \Delta u_3 \\ \Delta u_4 \\ \Delta u_5 \\ \Delta u_6 \end{bmatrix} = \frac{-v_n}{1 + \dfrac{m_B}{m_{B'}}} \begin{bmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & 1 \\ 0 & \dfrac{b\kappa}{k^2} & 0 \\ -\dfrac{b\kappa}{k^2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\left(1 - \dfrac{w_t}{v_t}\right)\tan\alpha \sin\gamma \\ -\left(1 - \dfrac{w_t}{v_t}\right)\tan\alpha \cos\gamma \\ 1 + e \end{bmatrix}.$$

Judgement of the validity of the assumption (15) may be begun by comparison of (17) with the results of a detailed study of deformation and frictional sticking and sliding during contact of two elastic spheres (Maw et al., 1976). For this purpose, the above results are expressed in terms of β, the angle between $\mathbf{w}$ and $\mathbf{n}$:

$$\frac{e\kappa \tan\beta}{(1+e)\mu} = \begin{cases} \dfrac{1}{2}\,[1 - \sqrt{1 + 4a(1-a)}] & 0 \leq a \leq 1 \\ a - 1 & a \geq 1. \end{cases}$$

Figure 2 permits comparison with the predictions resulting from the assumptions (14) and with the detailed study (Maw et al., 1976). All three predictions agree for $a > \mu_o/\mu$, where slipping at the end of contact is in the same direction as that at the beginning of contact. For values of between about 1/4 and μ_o/μ, the prediction from (15) agrees better with that of (Maw et al., 1976) than does the prediction from (14).

The result (17) can be combined with equation (11) to determine the implied loss of kinetic energy:

$$\Delta K = -\frac{1}{2}\,\overline{m}\,v_n^2 \left[1 - e^2 + \kappa \tan^2\alpha \left(1 - \frac{w_t^2}{v_t^2}\right)\right]$$

which, for $a > 1$, becomes[3]

$$\Delta K = -\frac{1}{2}\,\overline{m}\,v_n^2\,(1+e)\left\{1 - e + \mu\left[2 \tan\alpha - \frac{(1+e)\mu}{\kappa}\right]\right\}.$$

Kane's Example

Figure 3 depicts two slender rods, connected together and to a fixed support with hinges, the system free to swing in the plane. The lower end strikes the rigid surface, inducing a sudden change in motion of the system. This example was used by Professor Kane to point out the difficulty mentioned in the abstract above (see Kane and Levinson, 1985, p. 348).

The rods are identical, each with mass m_R and length b. Using $\dot\theta_1$ and $\dot\theta_2$ as generalized speeds, the kinetic energy is written as

$$K = \frac{1}{2}\,m_R b^2 \left[\frac{4}{3}\,u_1^2 + \cos(\theta_2 - \theta_1) u_1 u_2 + \frac{1}{3}\,u_2^2\right]$$

and the velocity of the lower end as

$$\mathbf{v} = b(-\cos\theta_1 \mathbf{t} + \sin\theta_1 \mathbf{n})u_1 + b(-\cos\theta_2 \mathbf{t} + \sin\theta_2 \mathbf{n})u_2.$$

From these, the matrices m and l are identified and the matrix of the impulse-velocity change dyadic constructed:

$$M = \frac{m_R}{6s^2\theta} \begin{bmatrix} (2s_1^2 + 8s_2^2 - 6c\theta s_1 s_2) & (2c_1 s_1 + 8c_2 s_2 - 3c\theta s\psi) \\ (2c_1 s_1 + 8c_2 s_2 - 3c\theta s\psi) & (2c_1^2 + 8c_2^2 - 6c\theta c_1 c_2) \end{bmatrix}$$

in which

[3] If κ were replaced with 1, this result would agree with equation (19) of Brach (1984). For spheres, the largest possible value of κ is 2/5, and for analogous planar collisions of disks, the largest possible value of κ is 1/2. The error in Brach (1984) appears to stem from a peculiar concept of a particle.

$$(c_i, s_i) \triangleq (\cos\theta_i, \sin\theta_i) \quad i=1,2;$$

$$\theta \triangleq \theta_2 - \theta_1; \quad \psi \triangleq \theta_2 + \theta_1.$$

Insertion of equation (12) into equation (8) leads to

$$\begin{bmatrix} g_t \\ g_n \end{bmatrix} = \begin{bmatrix} M_{tt} & M_{tn} \\ M_{nt} & M_{nn} \end{bmatrix} \begin{bmatrix} w_t - v_t \\ -(1+e)\, v_n \end{bmatrix}.$$

The criteria (14) then imply that if

$$\mu_o \geq \frac{|M_{tn}(1+e) - M_{tt}\tan\,\alpha|}{M_{nn}(1+e) - M_{nt}\tan\,\alpha},$$

the tangential velocity at separation is zero, and otherwise satisfies

$$x + (Ax+B)\,\frac{x}{|x|} + C = 0$$

in which α is the angle between $\mathbf{v}$ and $-\mathbf{n}$, and

$$x = \frac{w_t}{-v_n} \qquad A = \frac{\mu M_{nt}}{M_{tt}}$$

$$B = \frac{\mu}{M_{tt}}\,[M_{nn}(1+e) - M_{nt}\tan\,\alpha] \quad C = \frac{M_{tn}(1+e)}{M_{tt}} - \tan\,\alpha.$$

Alternatively, the proposed contact law (15) implies that w_t must satisfy

$$x + (Ax+B)\,\frac{\tan^2\alpha + |x|x}{\tan^2\alpha + x^2} + C = 0.$$

The following configuration and motion at initiation of contact will be used to illustrate the implications of these two contact assumptions.

$$\theta_1 = 18 \text{ deg} \quad \theta_2 = 30 \text{ deg} \quad \dot{\theta}_1 = -\omega \quad \dot{\theta}_2 = -2\omega$$

With these values,

$$M = \begin{bmatrix} 4.951\,312 & 7.214\,469 \\ 7.214\,469 & 11.471\,051 \end{bmatrix} m_R$$

and

$$\mathbf{v} = (2.683\,107\,\mathbf{t} - 1.309\,017\,\mathbf{n})b\omega.$$

If, in addition, $\mu_o < 0.4489$, $\mu = 0.4$, and $e = 0.7$, the criteria (14) predict

$\mathbf{w} = (-0.146\,\mathbf{t} + 0.916\,\mathbf{n})b\omega$

$\mathbf{g} = (2.046\,\mathbf{t} + 5.116\,\mathbf{n})m_R b\omega$

$\Delta K = 1.591\, m_R b^2\omega^2$ (an increase of 68.8 percent of the initial kinetic energy).

With the same coefficient of kinetic friction and normal velocity ratio, the proposed contact law (15) predicts

$\mathbf{w} = (-0.660\,\mathbf{t} + 0.916\,\mathbf{n})b\omega$

$\mathbf{g} = (-0.499\,\mathbf{t} + 1.407\,\mathbf{n})m_R b\omega$

$\Delta K = -0.781\, m_R b^2\omega^2$ (a decrease of 33.8 percent of the initial kinetic energy).

These results, along with the Cauchy quadric associated with $\mathbf{M}$, are shown in Fig. 4. The difficulty with the criteria (14) is readily apparent from the figure: The tangential velocity reverses direction during impact, driven by the normal impulse through the strong coupling expressed in $\mathbf{M}$. The criteria (14) set the direction of tangential impulse based only on the tangential *separation* velocity, taking no account of the friction force that would be expected during the time the tangential velocity is brought from v_t to zero. The difference between the two predicted values of separation velocity is modest, while the differences between predicted values of impulse and change in kinetic energy are both striking.

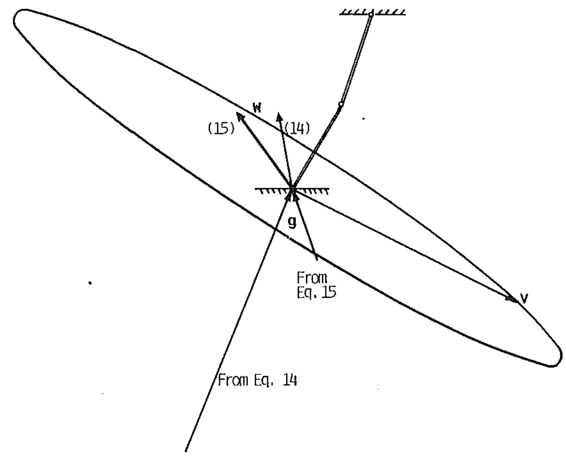


Fig. 4 Impulse and rebound predictions for Kane's example

Summary

Equations of rigid-body mechanics provide a means by which rebounds may be predicted without recourse to highly complex, detailed analysis of deformations during contact. To retain the simplicity of the method, assumptions about the impulse and velocity change at the point of contact must be made to supplement the equations of kinetics. Equation (8) and the Cauchy ellipsoid associated with $\mathbf{M}$ provide a convenient vehicle with which implications of such contact assumptions can be examined. Equations (8), (9), and (10) also suggest a better-organized computational procedure than normally results from an unguided consideration of equations (2) through (7).

The proposed assumption (15) permits retaining the simplicity of analysis based on rigid body kinetics, while leading to results that appear to be more reasonable than those from the use of the assumption (14). However, any such assumption will require validation through more extensive comparisons with results of experiments and of more detailed analysis, such as presented in Fig. 2. There appears to be a further need to examine the common assumption that e is a constant (independent of, for example, μ, $\mathbf{v}$, and $\mathbf{M}$), to make available better guidance for estimating its value.

Acknowledgment

This research was supported by the Northwest College and University Association for Science (University of Washington) under Contract DE-AM06-76-RL02225 with the U.S. Department of Energy.

References

Brach, R. M., 1984, "Friction, Restitution, and Energy Loss in Planar Collisions," ASME Journal of Applied Mechanics, Vol. 51, pp. 164–170.

Brach, R. M., 1989, "Rigid Body Collisions," ASME Journal of Applied Mechanics, Vol. 56, pp. 133–138.

Kane, T. R., and Levinson, D. A., 1985, *Dynamics: Theory and Applications*, McGraw-Hill, New York.

Keller, J. B., 1986, "Impact with Friction," ASME Journal of Applied Mechanics, Vol. 53, pp. 1–4.

Maw, N., Barber, J. R., and Fawcett, J. N., 1976, "The Oblique Impact of Elastic Spheres," *Wear*, Vol. 38, No. 1, pp. 101–114.

Routh, E. J., 1905, *Dynamics of a System of Rigid Bodies*, Vol. 1, Macmillan and Co., London, U.K.

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

Successive Synthesis of Substructure Modes

Luis E. Suarez[1]
Assistant Professor.

Mahendra P. Singh
Professor.

Department of Engineering Science and Mechanics,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061

A mode synthesis approach is presented to calculate the eigenproperties of a structure from the eigenproperties of its substructures. The approach consists of synthesizing the substructures sequentially, one degree-of-freedom at a time. At each coupling stage, the eigenvalue is obtained as the solution of a characteristic equation, defined in closed form in terms of the eigenproperties obtained in the preceding coupling stage. The roots of the characteristic equation can be obtained by a simple Newton-Raphson root finding scheme. For each calculated eigenvalue, the eigenvector is defined by a simple closed-form expression. The eigenproperties obtained in the final coupling stage provide the desired eigenproperties of the coupled system. Thus, the approach avoids a conventional solution of the second eigenvalue problem. The approach can be implemented with the complete set or a truncated number of substructure modes; if the complete set of modes is used, the calculated eigenproperties would be exact. The approach can be used with any finite element discretization of structures. It requires only the free interface eigenproperties of the substructures. Successful application of the approach to a moderate size problem (255 degrees-of-freedom) on a microcomputer is also demonstrated.

Introduction

For the dynamic analysis of large structures, the component mode synthesis approaches are now commonly used. In these approaches the modes or the eigenproperties of the components or substructures are synthesized to obtain the modes or eigenproperties of the complete system. This involves the implementation of the following steps in a sequence (Craig, 1981): (1) division of the structure into two or more substructures, (2) calculation of the component modes of the individual substructures, (3) synthesis of a first few modes of the substructures to obtain and solve a new reduced size (transformed) eigenvalue problem, and (4) back transformation of the calculated eigenproperties to obtain the original system eigenproperties. The first papers on this subject were by Gladwell (1964) and Hurty (1965) and since then several improvements have been proposed and the literature on this topic is quite rich now (Hurty et al., 1971; Engels and Harcrow, 1981).

The currently used mode synthesis approaches are essentially matrix order reduction techniques. Invariably, they require the solution of a second eigenvalue problem of smaller size, which provides a first few eigenproperties of the combined system. The writers (Suarez and Singh, 1987) have developed an alternative mode synthesis approach whereby exact and complete sets of eigenproperties of the combined system can be obtained without conventionally solving a second eigenvalue problem. The approach consists of sequentially coupling two subsystems, one degree-of-freedom at a time. But this approach has some limitations; in particular, it can be used only with spring, truss, or beam elements which couple the subsystems. That is, its application to generalized finite element models consisting of plate, shell, or solid elements is not possible. In this paper, we now generalize this approach by carrying out the sequential coupling procedure in the modal space.

Eigenvalue Analysis

The approach is formulated for the substructure eigenvalues obtained with free boundary conditions at the interfaces. Consider a system composed of two substructures that we will refer to as the p and s structures, respectively, shown in Figs. 1 and 2. The p structure is modeled as an np degree-of-freedom system described by the stiffness matrix $[K_p]$ and mass matrix $[M_p]$. The s system is modeled as an ns degree-of-freedom system with the stiffness matrix $[K_s]$ and mass matrix $[M_s]$. The equations of motion for the free vibration of the uncoupled system are

$$\begin{bmatrix} M_p & 0 \\ 0 & M_s \end{bmatrix}\ddot{\mathbf{x}}+\begin{bmatrix} K_p & 0 \\ 0 & K_s \end{bmatrix}\mathbf{x}=0. \quad (1)$$

The combined uncoupled system possesses $n = np + ns$ degrees-of-freedom. These two subsystems are coupled at m degrees-of-freedom. The compatibility of displacements at the interface of the two systems can be represented by a set of holonomic constraints in the form:

[1]Presently at the Department of General Engineering, University of Puerto Rico, Mayaguez, PR 00709.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Aug. 17, 1987; final revision, Apr. 18, 1990.

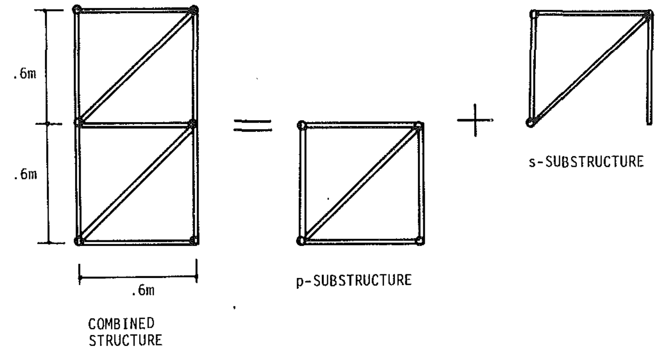


Fig. 1 Free-free truss considered for numerical example

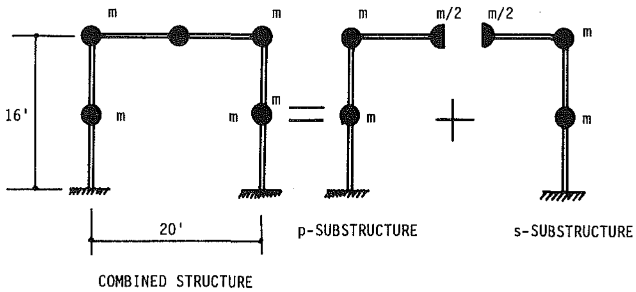


Fig. 2 Frame considered for numerical example

$$[B]\mathbf{x}=0 \tag{2}$$

where $[B]$ is an $(m \times n)$ dimensional matrix. The equations of motion for the two subsystems subjected to the set of holonomic constraints can be obtained through Lagrange's equations (Meirovitch, 1970) with Lagrange multipliers η_k

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right)-\frac{\partial L}{\partial x_i}=\sum_{k=1}^{m}\eta_k B_{ki};\ i=1,\ \ldots,\ n \tag{3}$$

where B_{ki} are the elements of matrix $[B]$. The Lagrangian L of the uncoupled system is defined as

$$L=\frac{1}{2}\,\dot{\mathbf{x}}^T\begin{bmatrix} M_p & 0 \\ 0 & M_s\end{bmatrix}\dot{\mathbf{x}}-\frac{1}{2}\,\mathbf{x}^T\begin{bmatrix} K_p & 0 \\ 0 & K_s\end{bmatrix}\mathbf{x} \tag{4}$$

from which we obtain the coupled equations of motion as

$$\begin{bmatrix} M_p & 0 \\ 0 & M_s\end{bmatrix}\ddot{\mathbf{x}}+\begin{bmatrix} K_p & 0 \\ 0 & K_s\end{bmatrix}\mathbf{x}-\sum_{i=1}^{m}\eta_i\mathbf{b}_i=0 \tag{5}$$

where we have defined matrix $[B]$ as

$$[B]^T=[\mathbf{b}_1\ \ \mathbf{b}_2\ .\ .\ .\ \mathbf{b}_m]. \tag{6}$$

Equation (5) is the equation of motion of the combined system with constraints. The last term of this equation represents the effect of the constraints as free vibration constraint forces. There are m terms in the summation, each corresponding to the m constraints represented by equation (2). To solve the eigenvalue problem corresponding to this equation, we will couple the two substructures successively by considering increasing number of terms in the summation. That is, at the first coupling stage we consider only one term of the summation, at the second stage two terms, and so on till all the terms are included. The development of the eigenvalue problem at these successive coupling stages is described next.

First Coupling Stage

Assuming that the subsystems are connected at only one degree-of-freedom, the holonomic constraints of equation (2) can be written as

$$\mathbf{b}_1^T\mathbf{x}=0 \tag{7}$$

where $\mathbf{b}_1$ is an n-dimensional vector with only two nonzero entries. The equations of motion for this *ad hoc* uncoupled system are

$$\begin{bmatrix} M_p & 0 \\ 0 & M_s\end{bmatrix}\ddot{\mathbf{x}}+\begin{bmatrix} K_p & 0 \\ 0 & K_s\end{bmatrix}\mathbf{x}-\eta_1\mathbf{b}_1=0. \tag{8}$$

We introduce in this equation the following transformation of coordinates

$$\mathbf{x}=\begin{bmatrix} \Phi_p & 0 \\ 0 & \Phi_s\end{bmatrix}\mathbf{q}=[U]\mathbf{q} \tag{9}$$

where $[\Phi_p]$ and $[\Phi_s]$ are the matrices of eigenvectors of the p and s systems, respectively, obtained with free interface coordinates. The eigenvector matrix of the p system is assumed to be normalized such that

$$[\Phi_p]^T[M_p][\Phi_p]=[I]. \tag{10}$$

A similar normalization scheme is adopted for the s system. Substituting equation (9) in (8) and premultiplying by $[U]^T$ we obtain the following transformed equation

$$[I]\ddot{\mathbf{q}}+[\Lambda_o]\mathbf{q}-\eta_1\mathbf{c}=0 \tag{11}$$

where

$$[\Lambda_o]=\begin{bmatrix} \Lambda_p & 0 \\ 0 & \Lambda_s\end{bmatrix} \tag{12}$$

in which $[\Lambda_p]$ and $[\Lambda_s]$ are, respectively, the diagonal matrices containing the eigenvalues of the p and s systems. Vector $\mathbf{c}$ is defined as

$$\mathbf{c}=[U]^T\mathbf{b}_1. \tag{13}$$

If, say, the uth degree-of-freedom of the p structure is connected to the vth degree-of-freedom of the s structure, the elements of vector $\mathbf{c}$ are simply defined as

$$c_i=U_{u,i}-U_{np+v,i};\quad i=1,\ .\ .\ .\ .\ ,\ n. \tag{14}$$

The compatibility condition in equation (7) is also transformed into

$$\mathbf{c}^T\mathbf{q}=0 \tag{15}$$

which implies that one of the generalized coordinates $\mathbf{q}$ can be expressed in terms of the remaining coordinates. If q_v is the dependent coordinate, we can write

$$q_v=-\frac{1}{c_v}\,\hat{\mathbf{c}}^T\hat{\mathbf{q}} \tag{16}$$

where the vectors $\hat{\mathbf{c}}$ and $\hat{\mathbf{q}}$ are obtained by eliminating the vth elements from the vectors $\mathbf{c}$ and $\mathbf{q}$; that is

$$\mathbf{q}=\begin{Bmatrix}\hat{\mathbf{q}}\\ q_v\end{Bmatrix};\quad \mathbf{c}=\begin{Bmatrix}\hat{\mathbf{c}}\\ c_v\end{Bmatrix}. \tag{17}$$

In view of equation (17) we can also rewrite equation (11) as follows

$$[I]\begin{Bmatrix}\ddot{\hat{\mathbf{q}}}\\ \ddot{q}_v\end{Bmatrix}+\begin{bmatrix}\hat{\Lambda}_o & \mathbf{0}\\ \mathbf{0}^T & \omega_{ov}^2\end{bmatrix}\begin{Bmatrix}\hat{\mathbf{q}}\\ q_v\end{Bmatrix}-\eta_1\begin{Bmatrix}\hat{\mathbf{c}}\\ c_v\end{Bmatrix}=0 \tag{18}$$

where $[\hat{\Lambda}_0]$ contains the eigenvalues ω_{0i} of the uncoupled system with the exception of the vth eigenvalue. From the last row of the above set of equations we obtain

$$\eta_1=\frac{1}{c_v}\,(\ddot{q}_v+\omega_{ov}^2 q_v) \tag{19}$$

which, with the help of equation (16), can be also written

$$\eta_1=-\frac{1}{c_v^2}\,\hat{\mathbf{c}}^T(\ddot{\hat{\mathbf{q}}}+\omega_{ov}^2\hat{\mathbf{q}}). \tag{20}$$

Substituting equation (20) in equation (18) we obtain

$$\left[[I]+\frac{1}{c_v^2}\,\hat{\mathbf{c}}\hat{\mathbf{c}}^T\right]\ddot{\hat{\mathbf{q}}}+\left[[\hat{\Lambda}_o]+\frac{\omega_{ov}^2}{c_v^2}\,\hat{\mathbf{c}}\hat{\mathbf{c}}^T\right]\hat{\mathbf{q}}=0. \tag{21}$$

The eigenvalue problem associated with the above system of equations is

$$\left[[\hat{\Lambda}_o]+\frac{\omega_{ov}^2}{c_v^2}\hat{\mathbf{c}}\hat{\mathbf{c}}^T\right]\hat{\boldsymbol{\phi}}_j^{(1)}=\lambda_j^{(1)}\left[[I]+\frac{1}{c_v^2}\hat{\mathbf{c}}\hat{\mathbf{c}}^T\right]\hat{\boldsymbol{\phi}}_j^{(1)} \tag{22}$$

where the superscript (1) in equation (22) now associates the eigenproperties with the first coupling stage when the two substructures are coupled by only one degree-of-freedom.

Instead of attempting to solve this eigenproblem by conventional methods, we will now develop a closed-form expression for the characteristic equation of equation (22) and then solve it to obtain the eigenvalues. For each calculated eigenvalue, the eigenvector can be obtained from a closed-form expression. First, however, we will assume here that neither the p nor the s structure possess any rigid body mode. The case of the free-free uncoupled subsystems with rigid body modes will be examined in a later section.

Eigenvalues and Eigenvectors. By rearranging equation (22) in the following form

$$[[\hat{\Lambda}_o]-\lambda_j^{(1)}[I]]\hat{\boldsymbol{\phi}}_j^{(1)}=-\frac{\omega_{ov}^2-\lambda_j^{(1)}}{c_v^2}\hat{\mathbf{c}}\hat{\mathbf{c}}^T\hat{\boldsymbol{\phi}}_j^{(1)}, \tag{23}$$

we can solve for $\hat{\boldsymbol{\phi}}_j^{(1)}$ directly, as the matrix on the left is diagonal

$$\hat{\boldsymbol{\phi}}_j^{(1)}=-\frac{\omega_{ov}^2-\lambda_j^{(1)}}{c_v^2}\left[\frac{1}{\hat{\delta}_i}\right]\hat{\mathbf{c}}\hat{\mathbf{c}}^T\hat{\boldsymbol{\phi}}_j^{(1)} \tag{24}$$

where

$$\hat{\delta}_i=\omega_{0i}^2-\lambda_j^{(1)};\quad i=1,\ldots,n-1. \tag{25}$$

Multiplying by $\hat{\mathbf{c}}^T$ and realizing that the dot product $(\hat{\mathbf{c}}^T\boldsymbol{\phi}_j^{(1)})$ is a scalar and thus it can be cancelled out, we obtain:

$$1=-\frac{\omega_{ov}^2-\lambda_j^{(1)}}{c_v^2}\hat{\mathbf{c}}^T\left[\frac{1}{\hat{\delta}_i}\right]\hat{\mathbf{c}}. \tag{26}$$

Equation (26) can be expanded to obtain:

$$\sum_{i=1}^{n-1}\frac{\hat{c}_i^2}{\hat{\delta}_i}+\frac{c_y^2}{\omega_{ov}^2-\lambda_j^{(1)}}=0. \tag{27}$$

Realizing that the elements of $\hat{\mathbf{c}}$ are the same as the elements of $\mathbf{c}$ except for the element c_v, the above expression can be written more simply as:

$$f(\lambda_j)=\sum_{i=1}^{n}\frac{c_i^2}{\delta_i}=0 \tag{28}$$

where δ_i is defined by equation (25), but for all i between 1 and n. Equation (28) is the characteristic equation of order $n-1$. It can be easily solved to obtain its roots by a simple Newton-Raphson scheme, as described later.

Once the $(n-1)$ eigenvalues of the system are obtained from the solution of equation (28), the elements of the eigenvectors can be obtained directly from equation (24). We choose to normalize the eigenvectors $\hat{\boldsymbol{\phi}}_j$ with respect to the matrix on the right-hand side of equation (22). With some simplification, we obtain a simple expression for all n elements of the eigenvector $\boldsymbol{\phi}_j^{(1)}$ as:

$$\phi_{ij}^{(1)}=-\frac{c_i}{\delta_i}\tau_j;\quad i=1,\ldots,n;\quad j=1,\ldots,n-1 \tag{29a}$$

where

$$\tau_j=\left[\sum_{i=1}^{n}(c_i/\delta_i)^2\right]^{-1/2}. \tag{29b}$$

The eigenvectors $\psi_j^{(1)}$ of the original system, equation (8), are then obtained with:

$$\psi_j^{(1)}=[U]\boldsymbol{\phi}_j^{(1)};\quad j=1,\ldots,n=1. \tag{30}$$

If the subsystems are connected at the uth degree-of-freedom of the p system and the vth degree-of-freedom of the s system, it can be verified that the elements of the eigenvectors $\psi_j^{(1)}$ do satisfy the following compatibility condition:

$$\psi_{u,j}^{(1)}=\psi_{v+np,j}^{(1)}. \tag{31}$$

It can be shown from the following transformation that

$$\mathbf{x}=[\Psi^{(1)}]\mathbf{z}. \tag{32}$$

Utilizing the $(n\times n-1)$ modal matrix, $[\Psi^{(1)}]$, containing $\psi_j^{(1)}$ in its columns, can indeed decouple equation (8) for the purposes of dynamical analysis as

$$[I]\ddot{\mathbf{z}}+[\lambda_j^{(1)}]\mathbf{z}=0. \tag{33}$$

Subsequent Coupling Stages

We now consider a generic case where the two systems are connected by, say l degrees-of-freedom. In such a case, the equations of motion can be written as follows:

$$\begin{bmatrix}M_p & 0\\ 0 & M_s\end{bmatrix}\ddot{\mathbf{x}}+\begin{bmatrix}K_p & 0\\ 0 & K_s\end{bmatrix}\mathbf{x}-\sum_{i=1}^{l-1}\eta_i\mathbf{b}_i-\eta_l\mathbf{b}_l=0. \tag{34}$$

For example, $l=2$ for the second coupling stage. We introduce again the transformation of equation (9):

$$\mathbf{x}=[U]\mathbf{q}, \tag{35}$$

but now we define the transformation matrix $[U]$ in terms of the eigenvectors obtained in the preceding coupling stage as follows:

$$[U]=[\Psi^{(l-1)}]. \tag{36}$$

Matrix $[\Psi^{(l-1)}]$ is the eigenvector matrix of the system at the previous coupling stage, without its vth column; that is, $[U]$ is an $(n\times n-l+1)$ dimensional matrix. The transformation of equation (35) with premultiplication by $[U]^T$ leads to the following transformed system of equations:

$$[U]^T\begin{bmatrix}M_p & 0\\ 0 & M_s\end{bmatrix}[U]\ddot{\mathbf{q}}+[U]^T\begin{bmatrix}K_p & 0\\ 0 & K_s\end{bmatrix}[U]\mathbf{q}-[U]^T\left\{\sum_{i=1}^{l-1}\eta_i\mathbf{b}_i-\eta_l\mathbf{b}_l\right\}=0. \tag{37}$$

Since $[U]$ is the eigenvector matrix for the previous coupling stage, we can show that the first three terms of equation (37) reduce to the same decoupled form as equation (33), thereby simplifying it as follows:

$$[I]\ddot{\mathbf{q}}+[\Lambda_{l-1}]\mathbf{q}-\eta_l\mathbf{c}=0 \tag{38}$$

where $[\Lambda_{l-1}]$ is a diagonal matrix of dimension $(n-l+1)\times(n-l+1)$ containing the eigenvalues obtained at the preceding coupling stage. Vector $\mathbf{c}$ is now defined as

$$\mathbf{c}=[U]^T\mathbf{b}_l. \tag{39}$$

Equation (38) is of the same form as equation (11). Thus, proceeding as before, we obtain again the characteristic equation (28), except that the summation term is now for only $(n-l+1)$ terms. That is

$$f(\lambda_j)=\sum_{i=1}^{n-l+1}\frac{c_i^2}{\delta_i}=0 \tag{40}$$

where c_i are the elements of vector $\mathbf{c}$ defined in equation (39) and

$$\delta_i=\lambda_i^{l-1}-\lambda_j^l. \tag{41}$$

The roots of this $(n-l)$th order polynomial will provide the eigenvalues for the lth coupling stage. The eigenvectors of the transformed system are still calculated with equation (29) except that the ranges of the indices are different as:

$$\phi_{ij}^{(l)}=-\frac{c_i}{\delta_i}\tau_j;\quad i=1,\ldots,n-l+1;\quad j=1,\ldots,n-l. \tag{42}$$

The eigenvectors of the original system of equation (34) are retrieved from the transformation of equation (35). We again redefine matrix $[U]$ and vector $\mathbf{c}$ according to equations (36) and (39), respectively, for the next coupling stage. The process continues till the subsystems are coupled by all the m degrees-of-freedom.

Rigid Body Modes

As it was mentioned before, the expressions obtained for the eigenvectors in the previous section are not valid for the rigid body modes of the combined system. Let's suppose that at a certain previous stage of coupling, the combined system has nr rigid body modes. Then at the current stage, the eigenvalue problem in equation (22) for the case $\lambda_j = 0$ becomes:

$$[\omega_{io}^2]\hat{\phi}_j = -\frac{\omega_{ov}^2}{c_v^2}\,\hat{\mathbf{c}}\,\theta_j; \quad j=1,\ldots,nr-1 \tag{43}$$

where

$$\theta_j = \hat{\mathbf{c}}^T\hat{\phi}_j. \tag{44}$$

Because in this case $\omega_{o1} = \omega_{o2} = \ldots = \omega_{onr} = 0$, we cannot solve for $\hat{\phi}_j$ as we did previously. However, if we examine the first nr rows of equation (43) we observe that

$$\theta_j = \hat{\mathbf{c}}^T\hat{\phi}_j = 0; \quad j=1,\ldots,nr-1 \tag{45}$$

which when substituted in the remaining $(n - nr)$ rows of equation (43), defines some elements of ϕ_j as

$$\hat{\phi}_{ij} = 0; \quad i = nr+1,\ldots,n-l+1; \quad j=1,\ldots,nr-1. \tag{46}$$

The first nr elements of $\hat{\phi}_j$ are still undefined. However, we know that they must satisfy the condition in equation (45). In addition, we normalize the eigenvectors with respect to the matrix on the right-hand side of equation (22) which, with due consideration of equation (45) and (46), requires that

$$\sum_{i=1}^{nr} (\hat{\phi}_{ij})^2 = 1. \tag{47}$$

It is straightforward to show that any rigid body mode that satisfies equation (45) is automatically orthogonal to the remaining eigenvectors corresponding to the nonzero frequencies.

In order to define the first nr elements of $\hat{\phi}_j$, $j = 1, \ldots, nr - 1$, satisfying equations (45) and (47), we choose to define all but two elements as

$$\hat{\phi}_{ij} = 0; \quad i=1,\ldots,nr-1; \quad i \neq j; \quad j=1,\ldots,nr-1. \tag{48}$$

The condition in equation (45) then becomes

$$\sum_{i=1}^{nr} \hat{c}_i\hat{\phi}_{ij} = \hat{\phi}_{jj}\hat{c}_j + \hat{\phi}_{nr,j}\hat{c}_{nr} = 0 \tag{49}$$

which, with the normalization as in equation (47), provide the remaining two elements as

$$\hat{\phi}_{j,j} = \frac{\hat{c}_{nr}}{\sqrt{\hat{c}_j^2 + \hat{c}_{nr}^2}}; \quad \hat{\phi}_{nr,j} = -\frac{\hat{c}_j}{\sqrt{\hat{c}_j^2 + \hat{c}_{nr}^2}}. \tag{50}$$

All the elements of the eigenvectors $\hat{\phi}_j$ are now defined. However, these eigenvectors, although orthogonal to the remaining eigenvectors with nonzero frequencies, are not orthogonal to each other. In order to render these eigenvectors orthogonal, the Gram-Schmidt orthogonalization process (Meirovitch, 1980) can be easily implemented. Once the nr eigenvectors are calculated with equations (46), (48), and (50), the orthogonal eigenvectors, identified here as $\hat{\phi}'$, are obtained as follows:

$$\hat{\phi}_j' = \hat{\phi}_j - \sum_{k=1}^{j-1} (\hat{\phi}_j^T\hat{\phi}_k')\hat{\phi}_k'; \quad j=2,\ldots,nr-1. \tag{51}$$

Note that the first eigenvector does not need to be orthogonalized. Once the eigenvectors $\hat{\phi}_j'$ are computed, they need to be renormalized according to equation (47).

Subsystems With Equal Frequencies

When subsystems with the same properties, geometry and boundary conditions, or some particular subsystems with rigid body modes are being coupled for the first time, the eigenvalue problem from which the characteristic equation was derived, takes a special form that requires a different analysis than presented before. To examine this case, we will assume that two frequencies ω_{01} and ω_{02}, in the diagonal matrix $[\hat{\Lambda}_0]$ in equation (22), are equal. In such a case, the first two terms of the characteristic equation (28) will be the same. Thus, the characteristic polynomial will be of order $(n - l - 1)$, thereby giving only $(n - l - 1)$ roots or eigenvalues. The remaining eigenvalue cannot be obtained from this equation. To obtain this eigenvalue we subtract the vector $\omega_{01}^2[I + 1/c_v^2\hat{\mathbf{c}}\hat{\mathbf{c}}^T]\phi_j$ from both sides of equation (22) written for the case of $\omega_{01} = \omega_{02}$ and rearrange the terms to obtain

$$\left[\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \omega_{03}^2-\omega_{01}^2 & \\ & & & \omega_{0n}^2-\omega_{01}^2 \end{bmatrix} + \frac{\omega_{ov}^2-\omega_{01}^2}{c_v^2}\,\hat{\mathbf{c}}\hat{\mathbf{c}}^T\right]\hat{\phi}_j = (\lambda_j - \omega_{01}^2)\left[[I] + \frac{1}{c_v^2}\,\hat{\mathbf{c}}\hat{\mathbf{c}}^T\right]\hat{\phi}_j. \tag{52}$$

The first two rows of the matrix on the left-hand side of the above equation are linearly dependent. Thus, the rank of this matrix will be less by one, meaning thereby that one of the eigenvalues of equation (52) must be zero. That is

$$\lambda_j - \omega_{01}^2 = 0 \quad \text{or} \quad \lambda_j = \omega_{01}^2. \tag{53}$$

Therefore, one of the two equal eigenvalues remains unchanged after coupling.

Equation (37) can still be used to calculate the eigenvector except for the unchanged eigenvalue since $\delta_1 = 0$. From equation (52), with $\lambda_1 = \omega_{01}^2$, we obtain

$$\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \omega_{03}^2-\omega_{01}^2 & \\ & & & \omega_{0n}^2-\omega_{01}^2 \end{bmatrix}\hat{\phi}_1 = -\frac{\omega_{ov}^2-\omega_{01}^2}{c_v^2}\,\hat{\mathbf{c}}\hat{\mathbf{c}}^T\hat{\phi}_1. \tag{54}$$

Examining the first two rows we obtain

$$\hat{\mathbf{c}}^T\hat{\phi}_1 = 0 \tag{55}$$

and the remaining rows, in view of equation (55), give

$$\hat{\phi}_{i,1} = 0; \quad i=3,\ldots,n-1 \tag{56}$$

The two remaining elements of ϕ_j have to satisfy the constraint conditions (55)

$$\hat{\mathbf{c}}^T\hat{\phi}_1 = \hat{c}_1\hat{\phi}_{11} + \hat{c}_2\hat{\phi}_{21} = 0. \tag{57}$$

With the condition of orthonormality of this eigenvector with respect to the matrix on the right-hand side of equation (52), we also have

$$\hat{\phi}_1^T\hat{\phi}_1 = \hat{\phi}_{1,1}^2 + \hat{\phi}_{2,1}^2 = 1. \tag{58}$$

Equations (57) and (58) are solved to define the remaining elements of the eigenvectors as

$$\hat{\phi}_{1,1} = -\frac{\hat{c}_2}{\sqrt{\hat{c}_1^2 + \hat{c}_2^2}}; \quad \hat{\phi}_{2,1} = \frac{\hat{c}_1}{\sqrt{\hat{c}_1^2 + \hat{c}_2^2}}. \tag{59}$$

Solution of the Characteristic Equation

The roots of the nonlinear characteristic equation (28) can be easily obtained with the standard Newton-Raphson tech-

Table 1 Eigenvalues of the combined system in Fig. 1 at consecutive levels of coupling

Eigenvalues of the uncoupled system, $x\ 10^7$	Number of connected DOF: 1, Initial values	1, Final values	2, Initial values	2, Final values	3, Initial values	3, Final values	4, Initial values	4, Final values
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.014849	0.026802
0.0	0.0	0.0	0.0	0.0	0.014871	0.029698	0.034124	0.030170
0.0	0.0	0.0	0.015758	0.029742	0.034729	0.038551	0.073084	0.074807
0.0	0.024986	0.031557	0.046671	0.039715	0.073856	0.107618	0.129608	0.128331
0.049973	0.065613	0.061785	0.112890	0.107997	0.141713	0.151598	0.183912	0.155728
0.081254	0.137782	0.163995	0.189368	0.175430	0.196810	0.216225	0.222952	0.223230
0.194310	0.215144	0.214741	0.225444	0.218191	0.227882	0.229678	0.242720	0.232359
0.235978	0.243296	0.236146	0.243380	0.237574	0.248086	0.255761	0.265452	0.268941
0.250614	0.250614	0.250614	0.255203	0.258597	0.267436	0.275143	0.346317	0.293614
0.250614	0.255855	0.259793	0.276112	0.276274	0.347196	0.417490		
0.261095	0.277245	0.292431	0.361803	0.418120				
0.293395	0.363276	0.431174						
0.433156								

niquè (Householder, 1970). Because the derivative $f'(\lambda_j)$ of the function $f(\lambda_j)$ in equation (28) or (40) is always positive, the function $f(\lambda_j)$ is monotonically increasing in the intervals $\omega_{oj}^2 < \lambda_j < \omega_{oj+1}^2$. This fact provides us with bounds to check the convergence to the proper roots, as well as with initial values for the iterative process. Assuming that the eigenvalues ω_{oj}^2 at the previous stage of coupling are arranged in ascending order of their magnitudes $\omega_{01}^2 \le \omega_{02}^2 \le \ldots \le \omega_{0l}^2$, the initial trial values for the Newton-Raphson algorithm can be obtained as follows:

$$\lambda_j = \frac{\omega_{oj}^2 + \omega_{oj+1}^2}{2}. \tag{60}$$

If a root is very close to a pole of the function $f(\lambda_j)$, the root calculated in an iterative step can be larger than ω_{0j+1}^2 or smaller than ω_{0j}^2. In this situation, the process is restarted with the new initial value being taken as the average of the previous initial value and the bounding pole.

If some of the coefficients c_i are zero, the corresponding eigenvalues will remain unchanged after the coupling process. This can be seen by inspecting the eigenvalue problem in equation (22). For the coefficient c_k equal to zero, all the elements of the corresponding eigenvector ϕ_k are zero except $\phi_{kk} = 1$ or, in other words, the kth eigenvector ψ_k of the original system and its corresponding eigenvalue will remain unchanged during the coupling process. In such a case, therefore, only the eigenvalues which will change should be considered in the solution of the characteristic equation. The bounds for the roots of the characteristic equation are now defined only by the eigenvalues which will change; that is, by those eigenvalues which correspond to the c_i which are different from zero.

Numerical Results

The proposed procedure was implemented to obtain the eigenproperties of the structures shown in Figs. 1 and 2. The area of cross-section, mass density, and modulus of elasticity of each member of the truss were taken to be $A = 16.0$ cm^2, $\rho = 0.00783$ Kg/cm^3, $E = 210$ GPa, respectively. In the eigenvalue analysis of both substructures, the consistent mass matrix was used.

Table 1 shows the eigenvalues obtained in the synthesis process at each successive coupling stage. The eigenvalues of the uncoupled system are shown in column 1. Seven of these eigenvalues are zero which correspond to the three rigid body modes of substructure 1 and four of substructure 2. The next two columns show the initial estimates and the final eigenvalues in the first coupling stage calculated according to the proposed approach. The initial estimates of the eigenvalues were obtained from equation (60). For the rigid body modes the eigenvectors were calculated according to the method described in the section on "rigid body modes." It is noted that at each coupling stage the rigid body modes of the combined structures are reduced by one. The results shown in the last column of the fourth coupling stage now give the eigenvalues of the completely assembled structure. As one would expect, at this stage the structure is left with only three rigid body modes. The final eigenvalues are exactly the same as those obtained by a direct eigenvalue analysis of the combined structure.

For the structure shown in Fig. 2, the cross-sectional and material properties were taken to be uniform for all members as: $A = 17.1$ in.^{2}, moment of inertia $I = 228.0$ in.^{4}, $\rho = 0.105 \times 10^{-3}$ kips $-$ sec^2/in.^{4}, and $E = 30000$ ksi. In addition to the distributed mass considered in the consistent mass matrix, the concentrated masses of 0.5 kips $-$ sec^2/in. were also placed

Table 2 Frequencies of the combined system in Fig. 2 at consecutive levels of coupling

	Frequencies of Uncoupled Systems	Coupling Stage (1) Initial Values	(1) Final Values	(2) Initial Values	(2) Final Values	(3) Initial Values	(3) Final Values
1	1.4323	--	1.4323	3.3441	2.5609	--	2.5608
2	1.4323	3.3441	4.0917	--	4.0917	10.8760	8.1553
3	4.5071	--	4.5071	--	14.8268	--	15.5141
4	4.5071	11.6474	14.8268	11.6474	15.5141	38.6619	15.6880
5	15.8433	--	15.8433	38.8639	51.5294	--	51.5294
6	15.8433	38.8639	52.6275	--	52.6275	60.3742	56.3182
7	52.6289	--	52.6289	60.4518	65.3453	--	65.3453
8	52.6289	60.4518	67.2342	--	67.2342	74.8531	79.8246
9	67.3723	--	67.3725	--	81.7651	106.3695	96.9603
10	67.3723	98.7809	81.7651	98.7809	121.0358	--	121.0358
11	122.3776	--	122.3776	--	126.2673	--	150.0695
12	122.3776	136.9592	126.2673	136.9592	150.0695	141.3746	152.1695
13	150.1312	--	150.1312	152.5937	154.9007	--	154.9007
14	150.1312	152.5937	155.0165	--	155.0165	196.6019	176.7026
15	155.0171	--	155.0171	196.6659	228.4974	--	228.4974
16	155.0171	196.6658	230.8127	--	230.8127	--	--
17	230.9212	--	230.9212	--	--	--	--
18	230.9212	--	--	--	--	--	--

at each node as shown in the figure. The numerical results for this structure are presented in Table 2.

Column 1 shows the frequencies of the two substructures, arranged in increasing order of magnitude. Since the two substructures are identical and thus have the same frequencies, there are pairs of equal frequencies in this column. When the two substructures are joined, one of the two equal frequencies remains unchanged, as indicated in the paper. The initial estimates of these frequencies are, therefore, omitted from column 2 in the table. The final eigenvalues calculated for the first coupling stage are shown in column 3. In the subsequent coupling stages, several other eigenvalues are also seen to remain unchanged because the corresponding elements of the vector **c** were zero. These unchanged eigenvalues can be identified by the blank space in the columns of the initial estimates. The initial estimates of the eigenvalues which will change in the coupling process are again calculated by equation (60) as the average of the two consecutive eigenvalues obtained in the previous coupling stage. In calculating this average, however, the eigenvalues which remain unchanged are disregarded, as mentioned in the text. It is therefore noticed that in the case of this particular structure, there were always some eigenvalues which remained unchanged in all three coupling stages. The eigenvalues shown in column 7, obtained in the final coupling stage, are the true eigenvalues of the combined structure. Again, these eigenvalues were exactly the same as those obtained by a direct eigenvalue analysis of the combined structures. Once the eigenvalues were known, the corresponding eigenvectors were calculated by the closed-form expressions given in the paper. These were also exactly the same as those obtained by a direct eigenvalue analysis.

As another example, a moderate size problem of a structural frame with five bays and five stories was also analyzed by this approach. The frame consisted of a total of 110 frame elements with 85 unrestrained nodes. Each node had three degrees-of-freedom, with a total of 255 degrees-of-freedom in the structure. The structure was divided into two identical substructures, each with 45 modes and 135 degrees-of-freedom. The synthesis problem, therefore, has 135 equal eigenvalues and several instances of unchanged roots requiring special treatment, as noted in the paper earlier. The problem was successfully analyzed by the proposed synthesis scheme, utilizing all modes and eigenvalues of the two substructures, on an Apollo DN 3000 work station. The eigenvalues and eigenvectors of the combined structure obtained by the proposed synthesis approach and those obtained by a direct eigenvalue analysis of the complete structure were in excellent agreement, with some minor differences in the higher digits in only the higher eigenvalues. It can also be shown that the roots of the characteristic equation developed in the paper are not very sensitive to the errors in its coefficients. Since, the coefficients are calculated very simply, it is easy to control the error in the calculated eigenproperties in a successive step of the proposed approach. It is therefore expected that the accuracy of the calculated eigenvalues in the proposed approach will be better than the accuracy of the eigenvalues calculated by a straightforward algebraic matrix eigenvalue analysis.

Concluding Remarks

A generalized approach is presented for synthezing the eigenproperties of two structures to obtain the eigenproperties of the combined structure. The eigenproperties of the two substructures obtained with free boundary conditions at the interface are utilized. The two structures are coupled successively one degree-of-freedom at a time. At each coupling stage, the eigenvalues of the coupled systems are obtained as the solution of a characteristic equation. This characteristic equation is defined in closed-form. The roots of this equation can be easily obtained by the Newton-Raphson approach as the bounds on each root are precisely known as defined by the Raleigh's inclusion principle. Once the eigenvalues are known, the eigenvectors can be calculated by utilizing closed-form expressions, also provided in the paper. The eigenproperties calculated at a coupling stage are subsequently utilized in the next coupling stage. The eigenproperties calculated in the final coupling stage, when the substructures are connected at all the interfacing degrees-of-freedom, then provide the desired combined system eigenproperties.

The approach can be used with any finite element discretization of the substructures. The methods to treat the special cases involving the rigid body modes and equal frequencies are also presented. Numerical examples demonstrating the application of the proposed method is also given.

Herein we present a new mode synthesis approach without any claim about the computational superiority of the proposed method. Such claims can only be verified by solving large problems. This approach can provide an exact solution of the problem if a complete set of substructures' modes are used. However, a set of truncated modes can also be used in the approach if only a first few important eigenproperties are to be determined. For large eigenvalue problems, the proposed computational scheme will require a small memory storage requirement than a straightforward eigensolution scheme.

Acknowledgments

This work is financially supported by the National Science Foundation through Grant No. CEE-8208897, CEE-8412830 and CES-8719078 with Dr. S. C. Liu as the program Director. This support is gratefully acknowledged.

References

Craig, R. R., 1981, *Structural Dynamics*, John Wiley and Sons, New York.

Engels, R. C., and Harcrow, H. W., 1981, "A Survey of Payload Integration Methods," *AIAA/ASME/AHS Structures, Structural Dynamics and Materials Conference*, AIAA Paper No. 810568, pp. 303-311.

Gladwell, G. M. L., 1964, "Branch Mode Analysis of Vibrating Systems," *Journal of Sound and Vibration*, Vol. 1, pp. 41-59.

Householder, A. S., 1970, *The Numerical Treatment of a Single Non-linear Equation*, McGraw-Hill, New York.

Hurty, W. C., 1965, "Dynamic Analysis of Structural Systems Using Component Modes," *AIAA Journal*, Vol. 3, pp. 678-685.

Hurty, W., Collins, J. D., and Hart, G. C., 1971, "Dynamic Analysis of Large Structures by Modal Synthesis Techniques," *Computers and Structures*, Vol. 1, pp. 535-563.

Meirovitch, L., 1970, *Methods of Analytical Dynamics*, McGraw-Hill, New York.

Meirovitch, L., 1980, *Computational Methods in Structural Dynamics*, Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands.

Suarez, L. E., and Singh, M. P., 1987, "An Exact Component Mode Synthesis Approach," *Earthquake Engineering and Structural Dynamics*, to appear.

APPENDIX

As suggested by an anonymous reviewer, the characteristic equation (28) can also be developed quite elegantly as follows. For a single coupling case, equation (11) can also be written as

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} \ddot{\mathbf{q}} \\ \ddot{\eta} \end{Bmatrix} + \begin{bmatrix} \Lambda_0 & -\mathbf{c} \\ -\mathbf{c}^T & 0 \end{bmatrix} \begin{Bmatrix} \mathbf{q} \\ \eta \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}. \tag{61}$$

Its corresponding eigenvalue problem is

$$\left[\begin{bmatrix} \Lambda_0 & -\mathbf{c} \\ -\mathbf{c}^T & 0 \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \right] \phi_j = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}. \tag{62}$$

Reducing the matrix on the left-hand side of the equation (62) to the upper triangular form and setting its determinant formed as a product of the diagonal terms to zero will provide the characteristic equation (28). This approach avoids the use of the quantities with a hat.

Multi-Flexible Body Dynamics Capturing Motion-Induced Stiffness

A. K. Banerjee
Senior Staff Engineer.

M. E. Lemak
Senior Research Engineer.

Lockheed Missiles and Space Company
Sunnyvale, CA 94089

This paper presents a multi-flexible-body dynamics formulation incorporating a recently developed theory for capturing motion-induced stiffness for an arbitrary structure undergoing large rotation and translation accompanied by small vibrations. In essence, the method consists of correcting dynamical equations for an arbitrary flexible body, unavoidably linearized prematurely in modal coordinates, with generalized active forces due to geometric stiffness corresponding to a system of 12 inertia forces and 9 inertia couples distributed over the body. Computation of geometric stiffness in this way does not require any iterative update. Equations of motion are derived by means of Kane's method. A treatment is given for handling prescribed motions and calculating interaction forces. Results of simulations of motions of three flexible spacecraft, involving stiffening during spinup motion, dynamic buckling, and a slewing maneuver, demonstrate the validity and generality of the theory.

1 Introduction

The subject of simulation of motion of a system of rigid bodies connected in a topological tree has reached a stage of maturity, thanks to the work of numerous investigators over the last 25 years, so that books (Wittenburg, 1977; Roberson and Schwertassek, 1988) and efficient computer programs (Rosenthal and Sherman, 1986; Levinson and Kane, 1990) are now available. When the bodies in the system must be considered as deformable, the situation is quite different. Public domain software (Bodley et al., 1978; Singh et al., 1985) on flexible multibody dynamics has been available for some time, and it has been widely believed, in the absence of any disclaimers in these codes, that they simulate correctly the dynamics of flexible bodies over all ranges of motion. Recent simulations with rotating beams (Kane et al., 1987) and plates (Banerjee and Kane, 1989) have shown that, on the contrary, these codes can produce incorrect results such as predicting dynamic softening of a rotating structure when dynamic stiffening is to be expected, for *all rotational speeds*, with the frequency error growing from 0 to 100 percent as the speed grows from zero to the fundamental bending frequency. Eke and Laskin (1987), while confirming the predictions of Kane et al. (1987) with the simulation of one publicly available software, makes the recommendation that for rotation rates that are low compared to the fundamental bending vibration frequency, the analyst is better off deleting the erroneous dynamic softening terms from the formulations. Similar recommendations are given by Padilla and von Flotow (1989). In the light of this situation, the need for developing multibody formulations that correctly reflect motion-induced stiffness of arbitrary structures can hardly be overemphasized.

The mechanics of spin-stiffening of beams has long been known, see, for example, Bisplinghoff et al. (1955), and Meirovitch (1967). For more general rotating structures, Likins (1974) was one of the earliest to suggest the use of geometric stiffness augmenting the structural stiffness to correctly represent the dynamic response. The researches of Levinson and Kane (1976), Turcic and Midha (1984) and Turcic et al. (1984), Modi and Ibrahim (1988), Ider and Amirouche (1989) illustrate the use of geometric stiffness. Simo and Vu-Quoc (1986) and Housner et al. (1986) capture dynamic stiffening by using nonlinear stiffness and mass formulations, respectively, in finite element theory. All of these works use the elastic displacements at discrete points of a continuum as generalized coordinates; with the use of displacement-dependent geometric stiffness, this approach becomes unwieldy for large space structures. Conversely, the modal approach to representing motion-induced stiffness is highly effective from the point of view of model reduction. Wu and Haug (1990) use component modes and produce coupling between axial deformation and bending, necessary for geometric stiffness, by breaking a structure into multiple substructures and using constraint conditions and Lagrange multipliers in the manner of Bodley et al. (1978); however, this introduces additional modal coordinates for the substructures as well as associated multipliers, increasing the dimension of the problem and leaving the simulation susceptible to constraint violations. Zeiler and Buttrill (1988) consider generalized modal stiffness corresponding to geometric stiff-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 20, 1989; final revision, Feb. 1, 1990.

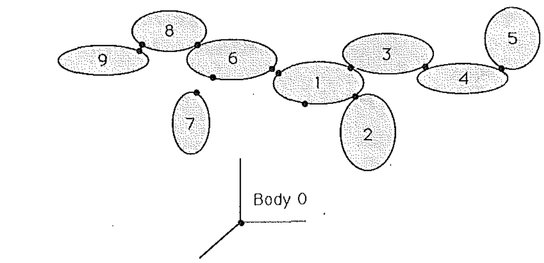


Fig. 1(*a*) A system of hinge-connected rigid and flexible bodies in a topological tree

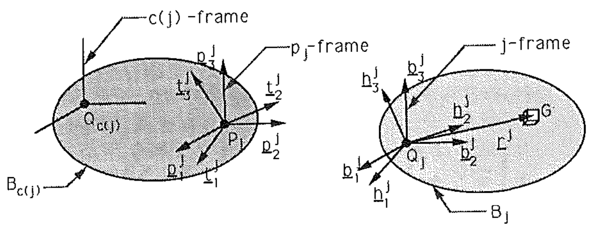


Fig. 1(*b*) Two hinge-connected adjacent bodies in the topological tree

ness associated with an approximate representation of the rotational inertia force. In a recent paper, Banerjee and Dickens (1990) considered geometric stiffness due to a system of 12 inertia forces and 9 inertia couples, representing the most general motion of the reference frame, to capture motion-induced stiffness for an arbitrary structure. This manner of geometric stiffness computation does not require iterative update of the geometric stiffness matrix as with nonlinear finite element codes (which makes these codes more accurate but also more expensive).

This paper extends the theory of Banerjee and Dickens (1990) to a multibody formulation, and shows in the process how to update existing multibody software to correctly represent dynamic stiffening. In Section 2, a system of hinge-connected flexible bodies in a topological tree is considered and generalized speeds are introduced. In Section 3, we review the kinematics and generalized inertia force expressions for an arbitrary flexible body. In Section 4, the recurrence relations for the jth hinge acceleration and angular acceleration for the reference frame of the jth body are written. The generalized active forces due to motion-induced stiffness and structural elasticity are given in Section 5. The final form of the flexible multibody dynamics equations are given in Section 6 with an explicit form of the dynamical "mass matrix"; use of the theory for a treatment of prescribed motion and associated interaction forces and moments are explained. Application of the theory to the simulation of three motions of a flexible spacecraft are given at the end.

2 System Description

A system of rigid and flexible bodies connected in a topological tree is shown in Fig. 1(*a*). We number the bodies arbitrarily, with an inertial frame denoted as body 0; following Huston and Passerello (1980) a topological array is defined such that body j has inboard adjacent body $c(j)$ in the path going from body j to body 0. Thus, for the system of Fig. 1(*a*), we have the array

j	1	2	3	4	5	6	7	8	9
$c(j)$	0	1	1	3	4	1	6	6	8

In Fig. 1(*b*) we show two adjacent bodies B_j and $B_{c(j)}$ connected at a hinge allowing rotation and translation, with Q_j a hinge on B_j and P_j the corresponding hinge point on $B_{c(j)}$. Reference frames j and p_j are attached to Q_j and P_j, respectively. Following Kane we introduce generalized speeds as motion variables, as many in number as there are degrees-of-freedom in the system. If there are R_j rotations along hinge axis unit vectors $\bar{h}^i_j$ (i = 1, ., R_j), and T_j translations along unit vectors $\bar{t}^i_j$, (i = 1, ., T_j), and if M_j modal coordinates are used to describe the small elastic deformation of body B_j in its reference frame j, then the following generalized speeds are selected.

$$u^j_i=\bar{\omega}^{j/p_j}\cdot\bar{h}^j_i \quad (j=1,..,N;\ i=1,..,R_j)$$

$$u^j_{R_j+i}=\bar{v}^{Q_j/P_j}\cdot\bar{t}^j_i \quad (j=1,\ldots,N;\ i=1,..,T_j)$$

$$u^j_{R_j}+T_j+i=\dot{\eta}^j_i \quad (j=1,\ldots,N;\ i=1,\ldots,M_j) \tag{1}$$

where $\bar{\omega}^{j/p_j}$ denotes the angular velocity of frame j in the frame p_j fixed at P_j and $\bar{v}^{Q_j/P_j}$ is the velocity of Q_j with respect to the point P_j (see Fig. 1(*b*)). Note that this defines n generalized speeds corresponding to the total number of degrees-of-freedom of the system, where n is given by

$$n=\sum_{j=1}^{N}(R_j+T_j+M_j). \tag{2}$$

3 Kinematics and Generalized Inertia Force for a Single Flexible Body

Let the elastic displacement $\bar{d}_j$ of the body B_j in Fig. 1(*b*) at a generic point G, located relative to Q_j by $\bar{r}_j$ in the undeformed configuration, be given in terms of space-dependent modal functions $\bar{\phi}^j_i$ and generalized coordinates η^j_i by

$$\bar{d}^j=\sum_{i=1}^{M_j}\bar{\phi}^j_i\eta^j_i \tag{3}$$

Then, the velocity of G in the inertial frame is written as

$$\bar{v}^G=\bar{v}^{Q_j}+\bar{\omega}^j x(\bar{r}^j+\bar{d}^j)+\sum_{i=1}^{M_j}\bar{\phi}^j_i\dot{\eta}^j_i \tag{4}$$

where $\bar{v}^G$ and $\bar{v}^{Q_j}$ denote the velocity in the *inertial* frame of points G and Q_j, respectively, and $\bar{\omega}^j$ is the *inertial* angular velocity of frame j. If G is treated as a rigid body rather than a material point, then the *inertial* angular velocity of G is written, for small elastic rotations of the body in frame j given in terms of the space-dependent functions $\bar{\psi}^j_i$, as

$$\bar{\omega}^G=\bar{\omega}^j+\sum_{i=1}^{M_j}\bar{\psi}^j_i\dot{\eta}^j_i. \tag{5}$$

The partial velocity of G (where for a nodal rigid body G stands for its mass center) and the partial angular velocity of nodal rigid body G, in the inertial frame, are obtained from equations (4) and (5) as (see Kane and Levinson, 1985)

$$\bar{v}^G_i=\bar{v}^{Q_j}_i+\bar{\omega}^j_i x(\bar{r}^j+\bar{d}^j)+\bar{\phi}^j_k\delta_{ik} \quad (i=1,\ldots..,n)$$

$$\bar{\omega}^G_i=\bar{\omega}^j_i+\bar{\psi}^j_k\delta_{ik} \quad (i=1,\ldots..,n) \tag{6}$$

where δ_{ik} is the Kronecker delta symbol, having the value unity only when the ith system generalized speed is the kth modal coordinate rate for the jth body, and is otherwise zero. Note that the velocity expressions in equations (4) and (5) do not represent the kinematical constraints of elastic deformation and are linear in the modal coordinates; partial velocities and partial angular velocities obtained from such prematurely linearized equations give rise to incorrect equations. However, for an arbitrary flexible body, there is no general way of writing the kinematical constraints; the resulting error will be compensated for in our consideration of the generalized active forces.

The acceleration of G of body B_j is obtained by differentiation in the inertial frame of velocity of G given by equation (4)

$$\bar{a}^G = \bar{a}^{Q_j} + \bar{\alpha}^j x(\bar{r}^j + \bar{d}^j) + \bar{\omega}^j x[\bar{\omega}^j x(\bar{r}^j + \bar{d}^j)] + 2\bar{\omega}^j x \sum_{i=1}^{M_j} \bar{\phi}_i^j \dot{\eta}_i^j + \sum_{i=1}^{M_j} \bar{\phi}_i^j \ddot{\eta}_i^j \quad (7)$$

The angular acceleration of G of body B_j follows from equation (5) as

$$\bar{\alpha}^G = \bar{\alpha}^j + \bar{\omega}^j x \sum_{i=1}^{M_j} \bar{\psi}_i^j \dot{\eta}_i^j + \sum_{i=1}^{M_j} \bar{\psi}_i^j \ddot{\eta}_i^j \quad (8)$$

The generalized inertia force contribution from body B_j is given by

$$F_i^{j*} = -\int_{B^j} \bar{v}_i^G \cdot \bar{a}^G dm - \int_{B^j} \bar{\omega}_i^G \cdot [dI^G \cdot \bar{\alpha}^G + \bar{\omega}^G x(\widetilde{dI}^G \cdot \bar{\omega}^G)] \quad (i=1,\ldots,n). \quad (9)$$

With equations (6)–(8) available, the operations indicated in equation (9) become straightforward, but massive and extremely laborious. What one must do is perform all the integrations over the spatial domain in equation (9) in such a way that the integrands do not contain any time-varying quantities. This is done by invoking certain vector-dyadic identities, and this yields the modal integrals defined in the Appendix. In terms of these integrals, equation (9) can be written as follows, where all terms nonlinear in the modal coordinates and their time derivatives have been dropped.

$$\begin{aligned}
F_i^{j*} = \; & -\bar{v}_i^{Q_j} \cdot \Big\{ m^j \bar{a}^{Q_j} - \bar{s}^j x \bar{\alpha}^j + \sum_{k=1}^{M_j} \bar{b}_k^j \ddot{\eta}_k^j + \bar{\omega}^j x \Big[\bar{\omega}^j x \bar{s}^j + 2 \sum_{k=1}^{M_j} \bar{b}_k^j \dot{\eta}_k^j \Big] \Big\} \\
& - \bar{\omega}_i^j \cdot \Big\{ \bar{s}^j x \bar{a}^{Q_j} + \bar{I}^j \cdot \bar{\alpha}^j + \sum_{k=1}^{M_j} \bar{c}_k^j \ddot{\eta}_k^j + \bar{\omega}^j x \bar{I}^j \cdot \bar{\omega}^j + 2 \sum_{k=1}^{M_j} \tilde{N}_k^{j*} \dot{\eta}_k^j \cdot \bar{\omega}^j \Big\} \\
& - \delta_{ik} \Big[\bar{b}_k^j \cdot \bar{a}^{Q_j} + \bar{g}_k^j \cdot \bar{\alpha}^j + \sum_{l=1}^{M_j} e_{lk}^j \ddot{\eta}_l^j - \bar{\omega}^j \cdot \tilde{D}_k^j \cdot \bar{\omega}^j + 2\bar{\omega}^j \cdot \sum_{l=1}^{M_j} \bar{d}_{lk}^j \dot{\eta}_l^j \Big] \\
& - \Big\langle \bar{\omega}_i^j \cdot \Big\{ \Big(\widetilde{W1}^j + \sum_{k=1}^{M_j} \widetilde{W2}_k^j \eta_k^j \Big) \cdot \bar{\alpha}^j + \bar{\omega}^j \cdot \sum_{k=1}^{M_j} \widetilde{W3}_k^j \dot{\eta}_k^j + \sum_{k=1}^{M_j} \widetilde{W4}_k^j \ddot{\eta}_k^j \\
& \quad + \bar{\omega}^j x \Big[\Big(\widetilde{W1}^j + \sum_{k=1}^{M_j} \widetilde{W2}_k^j \eta_k^j \Big) \cdot \bar{\omega}^j + \sum_{k=1}^{M_j} \overline{W4}_k^j \dot{\eta}_k^j \Big] \\
& \quad - \bar{\omega}^j \cdot \sum_{k=1}^{M_j} \widetilde{W5}_k^j \dot{\eta}_k^j \Big\} \\
& + \delta_{ik} \Big\{ \Big(\overline{W6}_k^j + \sum_{l=1}^{M_j} \overline{W7}_{lk}^j \eta_l^j \Big) \cdot \bar{\alpha}^j + \bar{\omega}^j \cdot \sum_{l=1}^{M_j} \overline{W8}_{lk}^j \dot{\eta}_l^j + \sum_{l=1}^{M_j} W9_{lk}^j \ddot{\eta}_l^j \\
& \quad - \bar{\omega}^j \cdot \Big[\widetilde{W3}_k^j \cdot \bar{\omega}^j + \sum_{l=1}^{M_j} \widetilde{W10}_{lk}^j \eta_l^j \cdot \bar{\omega}^j \\
& \quad - \sum_{l=1}^{M_j} (\overline{W11}_{lk}^j - \overline{W12}_{lk}^j) \dot{\eta}_l^j \Big] \Big\} \Big\rangle \quad (j=1,\ldots,N;\; i=1,\ldots,n) \quad (10)
\end{aligned}$$

If rotatory inertia is not significant, terms within the angular brackets ⟨ ⟩ are deleted.

4 Kinematical Recurrence Relations

In this section we review the recursive relations for $\bar{\omega}^j$, $\bar{v}^{Q_j}$, $\bar{\alpha}^j$, $\bar{a}^{Q_j}$ needed in equation (10) for body B_j in terms of the same kinematical variables for body $B_{c(j)}$. Refer to Fig. 1 (b). Introducing an orthogonal triad of basis vectors $\bar{p}_i^j$ fixed to the frame p_j, and assuming R_j hinge rotations about (possibly) nonorthogonal axes one can write the angular velocity of frame j as

$$\bar{\omega}^j = \bar{\omega}^{c(j)} + \sum_{i=1}^{M_{c(j)}} \bar{\psi}_i^{c(j)} (P_j) \dot{\eta}_i^{c(j)} + \sum_{i=1}^{3} \sum_{k=1}^{R_j} \bar{p}_i^j G_{ik}^j \dot{\theta}_k^j$$

$$\text{where} \quad G_{ik}^j = \bar{p}_i^j \cdot \bar{h}_k^j \quad (11)$$

where $\bar{\psi}^{c(j)}(P_j)$ is the ith modal small rotation at P_j; $\dot{\theta}_k^j$ is the time derivative of the kth relative rotation of the hinge at Q_j. The velocity of Q_j can be written assuming T_j translational motion at P_j along (possibly) nonorthogonal axes as follows.

$$\begin{aligned}
\bar{v}^{Q_j} = \; & \bar{v}^{Q_{c(j)}} + \bar{\omega}^{c(j)} x \Big[\bar{r}^{Q_{c(j)} P_j} + \sum_{k=1}^{M_{c(j)}} \bar{\phi}_k^{c(j)} (P_j) \eta_k^{c(j)} \Big] \\
& + \sum_{k=1}^{M_{c(j)}} \bar{\phi}_k^{c(j)} (P_j) \dot{\eta}_k^{c(j)} \\
& + \Big[\bar{\omega}^{c(j)} + \sum_{k=1}^{M_{c(j)}} \bar{\psi}_k^{c(j)} (P_j) \dot{\eta}_k^{c(j)} \Big] x \sum_{i=1}^{3} \sum_{k=1}^{T_j} \bar{p}_i^j L_{ik}^j \delta_k^j \\
& + \sum_{i=1}^{3} \sum_{k=1}^{T_j} \bar{p}_i^j (\dot{L}_{ik}^j \delta_k^j + L_{ik}^j \dot{\delta}_k^j)
\end{aligned}$$

$$\text{where} \quad L_{ik}^j = \bar{p}_i^j \cdot \bar{t}_k^j. \quad (12)$$

Here, δ_k^j is the kth translational degree-of-freedom at the hinge point Q_j. Equations (11) and (12) provide the mechanism for developing expressions for velocities of a point Q_j and the angular velocity for the jth reference frame for all bodies B_j in the tree structure in Fig. 1 (a), starting at body 1. The partial velocities needed in equation (9) are obtained from

$$\bar{v}^{Q_j} = \sum_{i=1}^{n} \bar{v}_i^{Q_j} u_i$$

$$\bar{\omega}^j = \sum_{i=1}^{n} \bar{\omega}_i^j u_i. \quad (13)$$

The remaining variables needed in equation (10), namely, $\bar{a}^{Q_j}$, $\bar{\alpha}^j$, represent, respectively, the acceleration of point Q_j and the angular acceleration of the jth frame; they are obtained by differentiation of $\bar{v}^Q$ and $\bar{\omega}^j$ in the inertial frame N, and written on the basis of equation (13) as

$$\bar{a}^{Q_j} = \frac{{}^N d}{dt} \bar{v}^{Q_j} = \sum_{i=1}^{n} \bar{v}_i^{Q_j} \dot{u}_i + \bar{h}^j$$

$$\bar{\alpha}^j = \frac{{}^N d}{dt} \bar{\omega}^j = \sum_{i=1}^{n} \bar{\omega}_i^j \dot{u}_i + \bar{f}^j. \quad (14)$$

Note that $\bar{h}^j$ and $\bar{f}^j$ in equation (14) represent terms free of u_i ($i = 1, \ldots, n$). Use of equations (11)–(14) in equation (10) completes the formulation for the ith generalized inertia force contribution from the jth flexible body. The system generalized inertia force is obtained by summing up these contributions over all bodies in the system,

$$F_i^* = \sum_{j=1}^{N} F_i^{j*} \qquad (i=1,\ldots,n). \quad (15)$$

5 Generalized Active Force due to Nominal and Motion-Induced Stiffness

In this section we develop the expressions for the generalized active forces due to nominal structural stiffness and motion induced (geometric) stiffness. Because both these quantities are typically developed via the finite element method that deals with matrices, we will use the matrix notation here.

Generalized active forces due to nominal, structural elasticity are customarily written on the basis of the component modes of the jth body in the matrix form,

$$F_n^j = -\Phi^{jT} K_n^j \Phi^j \eta^j = -\Lambda_n^j \eta^j \tag{16}$$

where Φ^j is the modal matrix (with superscript T denoting transpose), K_n^j is the stiffness matrix in the reference state, and η^j is the $(M_j \times 1)$ matrix of modal coordinates for the jth body. The matrix Λ_n^j premultiplying η^j in equation (16) is diagonal if the columns of Φ^j are mass-normalized vibration modes, and nondiagonal if constraint modes are used in addition to the normal modes (Craig, 1981). It should be emphasized here that the modes reflect deformations about the undeformed state of the body; if on the other hand, the equilibrium state of the body corresponds to a prestressed state, then modes have to be computed with this state as a reference.

Motion-induced stiffness, as conceived by Banerjee and Dickens (1990), is a special case of geometric or initial stress stiffness (Cook, 1985) caused by inertia loading. It has its origin in the strain energy term,

$$P_{NL} = \int_B \epsilon_{NL}^T \sigma_0 dv \tag{17}$$

where σ_0 is the reference state stress (matrix) due to inertia loading and ϵ_{NL} is a matrix representing the nonlinear terms in the Lagrangian strain tensor,

$$\sigma_0 = (\sigma_{110}, \sigma_{120}, \sigma_{130}, \sigma_{220}, \sigma_{230}, \sigma_{330})^T$$

$$\epsilon_{NL} = \begin{Bmatrix} 1/2(w_{1,1}^2 + w_{2,1}^2 + w_{3,1}^2) \\ 1/2(w_{1,2}^2 + w_{2,2}^2 + w_{3,2}^2) \\ 1/2(w_{1,3}^2 + w_{2,3}^2 + w_{3,3}^2) \\ w_{1,3}w_{1,2} + w_{2,3}w_{2,2} + w_{3,3}w_{3,2} \\ w_{1,3}w_{1,1} + w_{2,3}w_{2,1} + w_{3,3}w_{3,1} \\ w_{1,2}w_{1,1} + w_{2,2}w_{2,1} + w_{3,2}w_{3,1} \end{Bmatrix}. \tag{18}$$

Here, w_1, w_2, w_3 are the components in the orthogonal 1, 2, and 3-directions of the elastic displacement w, and a subscript preceded by a comma denotes differentiation with respect to that coordinate. In the standard procedure of the finite element method (Zienkiewicz, 1977) one assumes interpolation functions $N(x, y, z)$ between the nodal displacements d for the displacement at a point within an element, and derives elemental stiffness matrices based on the potential energy function. In this case equation (17) leads to the geometric or initial stress stiffness matrix k_g^e in

$$\{w\} = [N(x_1, x_2, x_3)]\{d\}$$

$$k_g^e = \int_e \begin{bmatrix} N_{,1}^T \; N_{,2}^T \; N_{,3}^T \end{bmatrix} \begin{bmatrix} \sigma_{110}I_3 & \sigma_{120}I_3 & \sigma_{130}I_3 \\ \sigma_{120}I_3 & \sigma_{220}I_3 & \sigma_{230}I_3 \\ \sigma_{130}I_3 & \sigma_{230}I_3 & \sigma_{330}I_3 \end{bmatrix} \begin{Bmatrix} N_{,1} \\ N_{,2} \\ N_{,3} \end{Bmatrix} dv \tag{19}$$

where I_3 is a 3×3 identity matrix. In a finite element program such as NASTRAN or EISI-EAL (Whetstone, 1983), the reference stress stiffness matrix is computed by first evaluating the stresses due to a reference state distributed loading, and then generating elemental stiffness matrices as per equation (19), and finally assembling the latter into a global geometric stiffness matrix for the structure. The distributed loading in this case is the system of inertia forces and inertia torques associated with the motion itself. To produce a theory that is finally linear in the modal coordinates (in consistency with the small elastic displacement assumption) we consider inertia force and torque expressions that involve only zeroth-order terms in the modal coordinates. In other words, we neglect the terms related to elastic displacement in equations (6)–(8) to develop expressions for the inertia force and torque on a nodal rigid body G in body j.

$$\bar{f}^{jG*} = -dm[\bar{a}^{Q_j} + \bar{\alpha}^j x \bar{r}^j + \bar{\omega}^j x(\bar{\omega}^j x \bar{r}^j)]$$

$$\bar{t}^{jG*} = -[\widetilde{dI}^G \cdot \bar{\alpha}^j + \bar{\omega}^j x \widetilde{dI}^G \cdot \bar{\omega}^j] \tag{20}$$

Equation (20) is written in the matrix form after introducing

$$\bar{b}^{jT} = [\bar{b}_1^j \; \bar{b}_2^j \; \bar{b}_3^j] \quad (j = 1, \ldots, N)$$

$$C_{1j}(i,k) = \bar{b}_i^1 \cdot \bar{b}_k^j. \tag{21}$$

We now define a $(n \times 1)$ column matrix U formed by stacking the generalized speeds given by equation (1) for all the bodies in the system. Then equation (13) is rewritten so as to introduce the partial velocity matrix for Q_j and the partial angular velocity matrix for frame-j,

$$\bar{\omega}^j = \bar{b}^{1T} \omega_u^j U$$

$$\bar{v}^{Q_j} = \bar{b}^{1T} v_u^{Q_j} U. \tag{22}$$

Equation (14) is now written as

$$\bar{a}^{Q_j} = \bar{b}^{jT} C_{lj}^T (v_u^{Q_j} \dot{U} + h^j)$$

$$\bar{\alpha}^j = \bar{b}^{jT} C_{lj}^T (\omega_u^j \dot{U} + f^j). \tag{23}$$

At this stage we introduce scalars z_i^j, $i = 1, \ldots, 15$; $j = 1, \ldots, N$:

$$\bar{\omega}^j = z_1^j \bar{b}_1^j + z_2^j \bar{b}_2^j + z_3^j \bar{b}_3^j$$

$$\bar{r}^j = x_1^j \bar{b}_1^j + x_2^j \bar{b}_2^j + x_3^j \bar{b}_3^j$$

$$\tilde{\omega}^j = \begin{bmatrix} 0 & -z_3^j & z_2^j \\ z_3^j & 0 & -z_1^j \\ -z_2^j & z_1^j & 0 \end{bmatrix}$$

$$\begin{bmatrix} z_4^j & z_5^j & z_6^j \\ z_5^j & z_7^j & z_8^j \\ z_6^j & z_8^j & z_9^j \end{bmatrix} = C_{1j}^T \tilde{\omega}^j \tilde{\omega}^j C_{1j}$$

$$z_{10}^j = z_1^j z_2^j$$

$$z_{11}^j = z_2^j z_3^j$$

$$z_{12}^j = z_3^j z_1^j$$

$$z_{13}^j = (z_1^j)^2 - (z_2^j)^2$$

$$z_{14}^j = (z_2^j)^2 - (z_3^j)^2$$

$$z_{15}^j = (z_3^j)^2 - (z_1^j)^2. \tag{24}$$

Now, the matrix form of the inertia force in equation (20) can be written as representing the following set of 12 distributed loadings:

$$\begin{Bmatrix} f_1^{j*} \\ f_2^{j*} \\ f_3^{j*} \end{Bmatrix} = -dm[I_3 \quad x_1 I_3 \quad x_2 I_3 \quad x_3 I_3] \begin{Bmatrix} A_1 \\ \vdots \\ A_{12} \end{Bmatrix} \tag{25}$$

where A_i ($i = 1, .., 12$) are expressed in terms of the previous definitions and e_i, the ith row of the 3×3 unity matrix I_3, as

$$\begin{Bmatrix} A_1^j \\ A_2^j \\ A_3^j \end{Bmatrix} = C_{1j}^T (v_u^{Q_j} \dot{U} + h^j)$$

$$A_4^j = z_4^j$$

$$A_5^j = z_5^j + e_3 C_{1j}^T (\omega_u^j \dot{U} + f^j)$$

$$A_6^j = z_6^j - e_2 C_{1j}^T (\omega_u^j \dot{U} + f^j)$$

$$A_7^j = z_5^j - e_3 C_{1j}^T(\omega_u^j \dot{U} + f^j)$$

$$A_8^j = z_7^j$$

$$A_9^j = z_8^j + e_1 C_{1j}^T(\omega_u^j \dot{U} + f^j)$$

$$A_{10}^j = z_6^j + e_2 C_{1j}^T(\omega_u^j \dot{U} + f^j)$$

$$A_{11}^j = z_8^j - e_1 C_{1j}^T(\omega_u^j \dot{U} + f^j)$$

$$A_{12}^j = z_9^j. \tag{26}$$

The inertia torque on a nodal rigid body, given in equation (20), can be similarly rewritten in terms of the inertia components of the nodal body in frame-j, as constituting the following set of nine torque loadings on the body.

$$\begin{Bmatrix} t_1^* \\ t_2^* \\ t_3^* \end{Bmatrix} = -\begin{bmatrix} I_{11} & I_{12} & I_{13} & I_{13} & I_{33}-I_{22} & -I_{12} & 0 & I_{23} & 0 \\ I_{12} & I_{22} & I_{23} & -I_{23} & I_{12} & I_{11}-I_{33} & 0 & 0 & I_{13} \\ I_{13} & I_{23} & I_{33} & I_{22}-I_{11} & -I_{13} & I_{23} & I_{12} & 0 & 0 \end{bmatrix} \begin{Bmatrix} A_{13} \\ \vdots \\ A_{21} \end{Bmatrix} \tag{27}$$

where A_i, (i = 13, ..., 21) are defined as

$$\begin{Bmatrix} A_{13}^j \\ A_{14}^j \\ A_{15}^j \end{Bmatrix} = C_{1j}^T(\omega_u^j \dot{U} + h^j)$$

$$A_{15+i}^j = z_{9+i}^j \quad (i = 1, \ldots, 6). \tag{28}$$

With the force and torque loadings given by equations (25) and (27), the finite element code has to generate geometric stiffness matrices $K_g^{(i)}$ for unit values of each of A_i, (i = 1, ..., 21); these are premultiplied by Φ^T and postmultiplied by Φ to yield the generalized geometric stiffness $S^{(i)}$, (i = 1, ..., 21); the time-varying nature of A_i is accounted for in the multibody computer program. This produces the generalized active force due to motion induced stiffness in the modal coordinates of the jth body in the form of the matrix equations,

$$S^{j(i)} = \Phi^{jT} K_g^{j(i)} \Phi^j$$

$$F_g^j = -\sum_{i=1}^{21} A_i^j S^{j(i)} \eta^j. \tag{29}$$

Finally, the scalar form of the generalized force due to nominal and motion-induced stiffness is written using equations (16) and (29) as

$$F_i^j = -\delta_{jk} \sum_{l=1}^{M_j} \left[\Lambda_{n(il)}^j + \sum_{m=1}^{21} S_{m(il)}^j A_m^j \right] \eta_l^j \quad (i = 1, \ldots, n). \tag{30}$$

Recall that the index of the second summation in equation (30) goes only up to 12 if rotatory inertia is ignored. System generalized active forces are obtained by summing individual body contributions over all bodies, a step that accomplishes the elimination of all nonworking constraint forces. Restricting ourselves only to forces associated with structural elasticity and motion-induced stiffness, we have

$$F_i = \sum_{j=1}^{N} F_i^j \quad (i = 1, \ldots, n). \tag{31}$$

Kane's dynamical equations (Kane and Levinson, 1985) for the system are

$$F_i^* + F_i = 0 \quad (i = 1, \ldots, n). \tag{32}$$

Considering the forms of equations (14) embedded in equations (10), and of equations (26) and (28) in equations (29), it is seen that equations (32) can be written in the matrix form

$$D\dot{U} + E = 0. \tag{33}$$

The coefficient matrix D is unsymmetric because of acceleration dependent stiffness and is a function of the generalized coordinates, while the column matrix E is a function of both the generalized coordinates and the generalized speeds.

6 Treatment of Prescribed Motion and Internal Forces

Here we consider two types of problems that can be treated alike. In one, relative motion between two contiguous bodies may be prescribed and the interaction force required to realize this motion is to be found; in another, the nonworking internal force at a joint of the structure may be of interest. If some of the generalized coordinates are prescribed, which means prescribing the corresponding generalized speeds and their derivatives, then equation (33) can be written in the partitioned marix form as follows, with a new term added to represent the generalized active forces due to the (working) interaction forces needed to realize the prescribed motion.

$$\begin{bmatrix} D_{ff} & D_{fp} \\ D_{pf} & D_{pp} \end{bmatrix} \begin{Bmatrix} \dot{U}_f \\ \dot{U}_p \end{Bmatrix} = -\begin{Bmatrix} E_{fp} \\ E_{pf} \end{Bmatrix} - \begin{Bmatrix} 0 \\ F \end{Bmatrix} \tag{34}$$

Here, subscripts f and p are associated with free and prescribed variables. The method of solution proceeds in the steps

$$D_{ff}\dot{U}_f = -D_{ff}\dot{U}_p - E_{fp}$$

$$F = -[D_{pf} D_{pp}] \begin{Bmatrix} \dot{U}_f \\ \dot{U}_p \end{Bmatrix} - E_{pf}. \tag{35}$$

For our choice of generalized speeds, which are time derivatives of relative angles and distances (see equation (1)), the elements of the column matrix F are directly the torques and forces that do work over the prescribed motion. If the elastic displacement is prescribed as in an antenna shape control problem, $F = \Phi^T f$, where f is a column matrix describing the actuator force distribution over the finite element nodal degrees

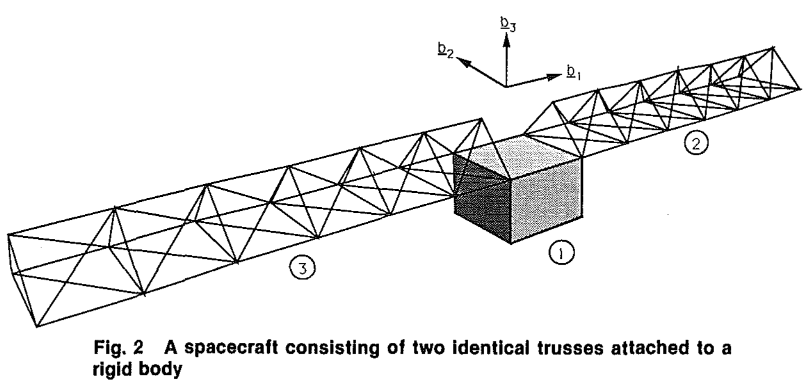


Fig. 2 A spacecraft consisting of two identical trusses attached to a rigid body

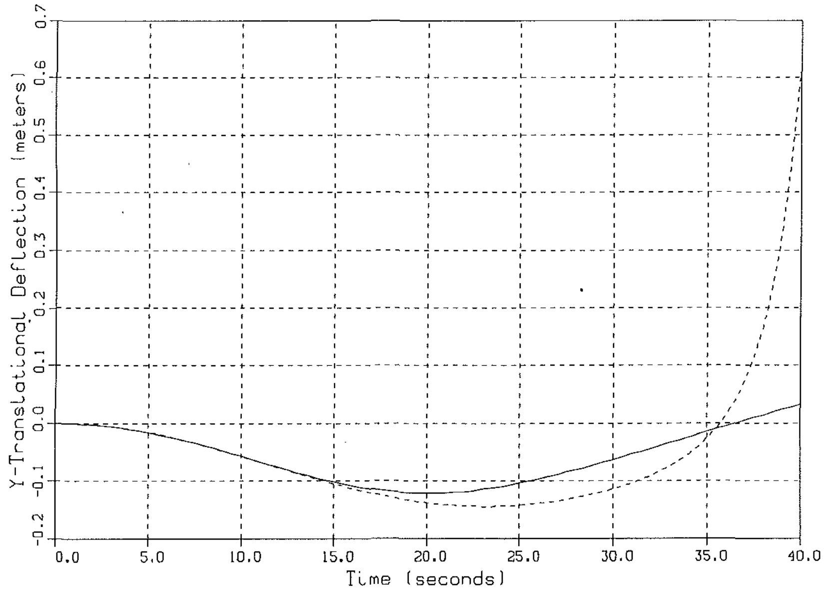


Fig. 3 Transverse tip deflection of a corner of the truss in spinup motion (solid line: present theory; dashed line: theory with no geometric stiffness)

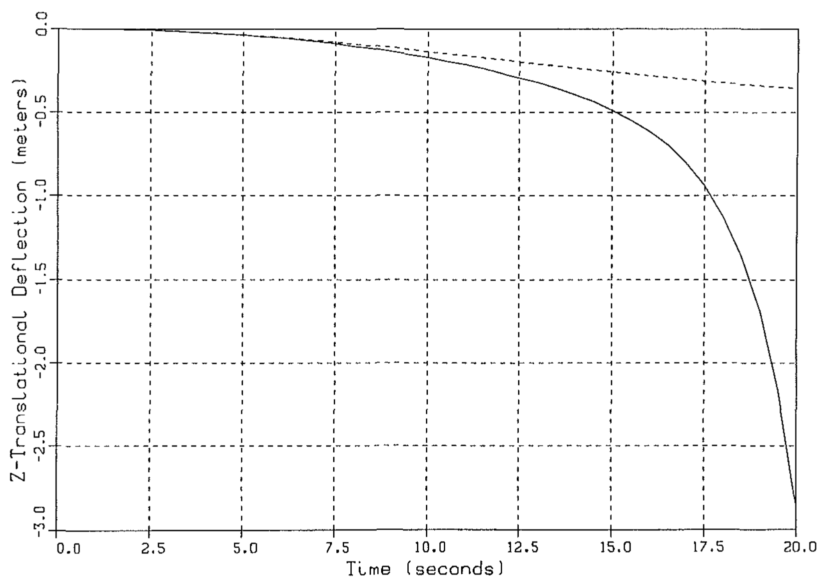


Fig. 4 Dynamic buckling of truss under translational acceleration (solid line: present theory; dashed line: theory with no geometric stiffness)

of freedom, in such a case a least square solution to f can be found by the use of a pseudo-inverse of Φ^T.

If the internal force at a hinge point in a structure is of interest, the hinge is assigned degrees-of-freedom and associated interaction forces and couples such that these forces and couples do work when motion involving those degrees-of-freedom takes place. This prescribed motion is subsequently set equal to zero. The interaction forces are found from equation (35).

7 Simulation Results

To establish the correctness and generality of the foregoing theory, we have performed motion simulations for three spacecraft problems, each bringing out a special feature of the theory. The first problem is that of spinup motion of a spacecraft (see Fig. 2) consisting of a rigid body to which two identical trusses are attached. Each truss is 12-m long and consists of tubular members interconnected as shown at 21 joints; the cross-section of the truss is a right-isosceles triangle, the members at the right angle being 2-m long. Young's modulus for all members is 10^9 N/m^2, the cross-sectional area of each member is 8.46×10^{-4} m^2 and each member has a mass density 10^4 kg/m^3. The rigid body has a mass of 10^4 kg, and a spin-axis moment of inertia of 1200 kg-m^2. The angular velocity of the rigid body in spinup is prescribed as follows, where ω is the angular speed in simple spin along the vector $\bar{b}_3$, Ω is the steady-state spin speed, and T is the rise time,

$$\omega = \frac{\Omega}{T}\left(t - \frac{T}{2\pi}\sin\frac{2\pi t}{T}\right) \quad ,t<T$$
$$= \Omega \quad ,t>T. \tag{36}$$

The vibration modes of the truss were obtained by means of the EAL (Whetstone, 1983) finite element program. The first ten modes for each truss are included with the lowest in-plane frequency (which is the second-mode frequency of the truss) of 0.592 Hz or 3.72 rad/sec. The spinup speed is purposefully chosen at 4 rad/sec to be higher than this lowest frequency to accentuate the capture of motion-induced stiff-

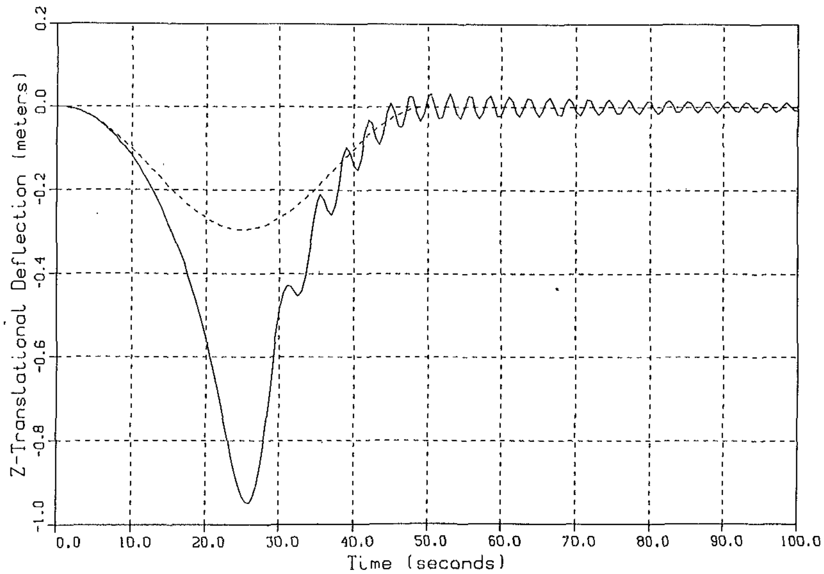


Fig. 5 Dynamic weakening of the truss under compression due to subcritical translational acceleration (solid line: present theory; dashed line: theory with no geometric stiffness)

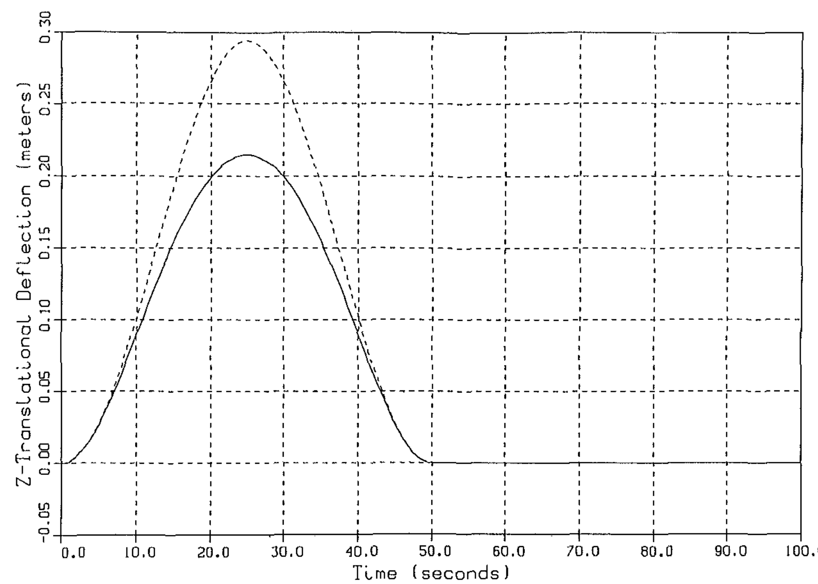


Fig. 6 Dynamic stiffening of the truss under tension due to translational acceleration (solid line: present theory; dashed line: theory with no geometric stiffness)

ness in this formulation; formulations not having any provision of geometric stiffness are known (see Kane et al., 1987) to produce unbounded response at spinup speeds exceeding the fundamental bending vibration frequency. The spinup speed in this case is attained over a rise time of 50 secs. The solid line in Fig. 3 shows the $\bar{b}_2$-direction tip deflection of the right-corner point of one truss at its free end. Note that these results are the same for both trusses in this case. The response may be recognized as the typical signature of dynamic stiffening of a structure, see for example, (Kane et al., 1987; Banerjee and Kane, 1989; Banerjee and Dickens, 1990). The dashed line in Fig. 3 shows the incorrect, unbounded response that is obtained when the geometric stiffness terms are omitted from equation (30).

Consideration of motion-induced stiffness in the present theory allows one to represent dynamic weakening and incipient buckling under appropriate conditions. To demonstrate this, we considered the problem of rectilinear motion of the system in Fig. 2, with the speed of the mass center of the rigid body given by equation (36) with ω now standing for speed in translation along the vector $\bar{b}_1$, and Ω the maximum speed attained. The maximum acceleration is then given by $2\Omega/T$.

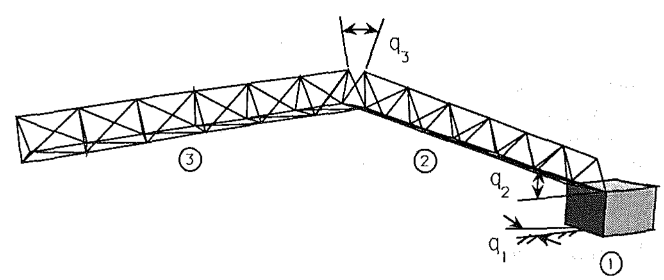


Fig. 7 A space crane with three rigid body degrees-of-freedom in a slewing maneuver

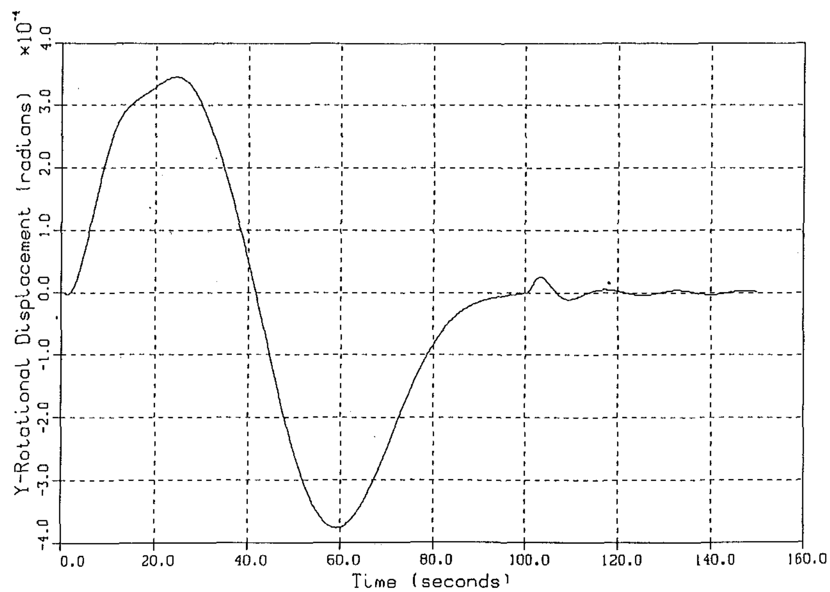


Fig. 8 Elastic rotation due to bending at the tip of the end truss during slewing of the space crane

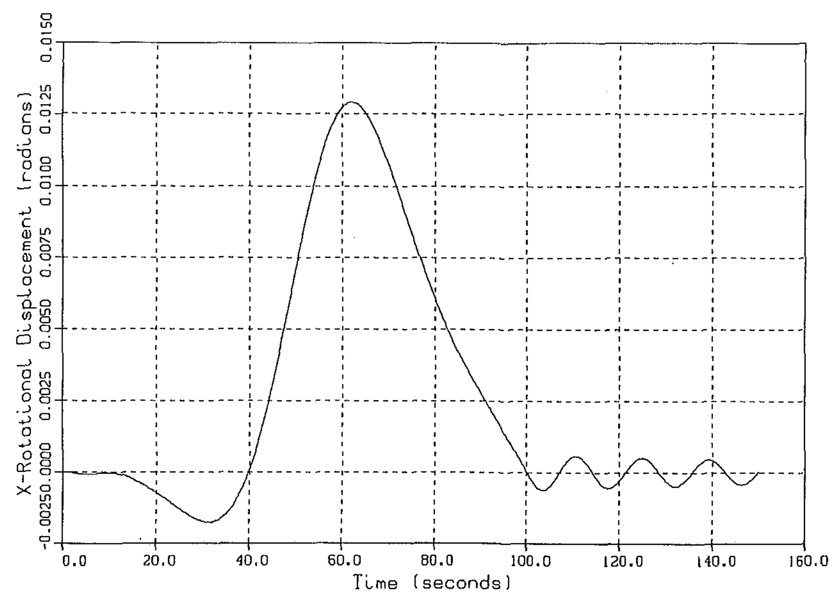


Fig. 9 Torsional rotation at the end of the inner truss during slewing of the space crane

The truss at the right end of the rigid body is now in dynamic compression, while the truss at the left end is in tension. The critical buckling acceleration for a uniform, constant loading applied at all nodes of the truss is computed by the EAL (Whetstone, 1983) finite element code as 53.6 m/s^2. Figure 4 shows the $\bar{b}_3$-direction tip deflection of the truss under compression. Here, we clearly see the evolution of buckling when the maximum acceleration during speedup corresponds to the critical acceleration. The solid line correctly demonstrates the expected large deflections of the buckling truss. The dashed line shows the erroneous solution with no geometric stiffness terms in equation (30).

Figures 5 and 6 demonstrate the expected softening (at the compression end) and stiffening (at the tension end) for a case where the system is accelerated to 75 percent of the critical buckling acceleration. Once again, the $\bar{b}_3$-directional deflections are plotted as a solid line for the correct solution and as a dashed line for the solution ignoring geometric stiffness. Figure 5 shows a larger deformation, indicating a softer struc-

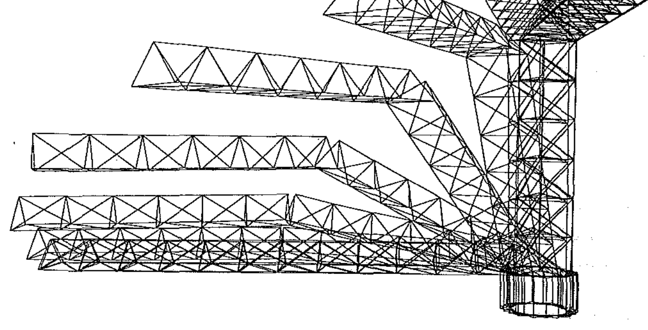


Fig. 10 Animation of the slewing maneuver for the space crane

ture, as compared to the incorrect result. On the other hand, the truss under dynamic tension shows that the correct solution corresponds to a stiffer structure than the incorrect solution predicts. In passing, note that the dashed line curves in Figs. 3–6, given by a formulation in which the geometric stiffness terms were deleted from our present formulation, effectively predict the response that is expected to be given by public domain software that do not have any provision for dynamic stiffness for an arbitrary structure.

Kane et al. (1989) point out the importance of accounting for motion-induced stiffness for a slewing maneuver of a flexible body and documents a case where lack of such stiffness results in a 180-deg phase error in the response. Accordingly, we investigate a slewing maneuver with the present formulation as it incorporates motion-induced stiffness for an arbitrary structure. Figure 7 shows a space crane consisting of a rigid body with two hinge-connected articulated trusses.

We assume that the rigid body undergoes a simple rotation through an angle q_1 in inertial space about the axis shown, while the trusses go through the relative angles q_2 and q_3. In a repositioning maneuver, the commanded values of these angles are q_{1c}, q_{2c}, and q_{3c}, respectively, and the following joint torques are applied at the hinges to realize the desired commands:

$$T_i = k_i(q_{ic} - q_i) - c_i\dot{q}_i \quad i = 1, 3. \tag{37}$$

Initially, the angles are all equal to zero, corresponding to both trusses being aligned in the plane of rotation of the base body. The commanded angles use the integral of the expression in equation (36) to reposition each body such that all three relative angles go from zero to 90 deg in 100 sec. The control gains in equation (37) are chosen for the "worst case" or largest effective inertia at the respective revolute joints, for a bandwidth of 0.1 Hz and damping factor of 0.707. This controller bandwidth is below the first natural frequency of the synthesized system in the initial configuration.

The data for an individual truss used in the previous two problems are also used here. However, a change is made to the inner truss by assigning a large stiffness at the free end. This is done to model a possible gimbal fixture that attaches to all three nodes of the truss tip, so that, the three nodes are "fixed" in a plane. The modal data used for the inner beam include the first ten fixed-fixed natural modes and six constraint modes (Craig, 1981) representing unit displacements for this end plane. With three rigid-body motion freedoms and ten modes for the outboard truss, this represents a system having 29 degrees-of-freedom.

Figure 8 shows the rotation due to bending at the free end of the second truss at the end of the maneuver. Figure 9 displays the angle of twist at the end of the inner truss at the elbow. Plots such as these can obviously be used to assess control-structure interactions corresponding to various control law designs for meeting a specified control objective. Finally, an animation of the motion of the structure during the repositioning maneuver is shown in Fig. 10, where the structure, initially in the plane of the paper, comes out of the plane during the maneuver.

Conclusion

The illustrative examples given here establish the validity and generality of the theory presented in this paper for capturing motion-induced stiffness (or softness) in a hinge-connected system of *arbitrary* flexible bodies undergoing large rotation and translation. The formulation consists of correcting flexible body dynamical equations, unavoidably linearized prematurely in modal coordinates, through the addition of generalized active forces associated with geometric stiffness due to a system of *12 inertia forces and 9 inertia couples* that corresponds to the most general motion of the reference frame of the body. Computation of geometric stiffness in this way does not require displacement-dependent updates, and is produced by online multiplication, by the values of the actual inertia forces and couples, of precomputed geometric stiffness matrices for unit values of these inertia forces and couples. It has been shown that the present formulation captures dynamic stiffening, which is at the heart of producing correct simulation of multi-flexible body dynamics, while making use of vibration modes to reduce the number of coordinates. Existing formulations for dynamics of a system of arbitrary flexible bodies reproduce dynamic stiffening either by introducing additional coordinates and enforcing constraints (a procedure susceptible to constraint violation) or by introducing geometric stiffness in a nonlinear finite element approach requiring iterative updates. Both these methods use large numbers of dependent variables, and hence require excessive computations. The contribution of the present paper in producing correct simulations of large motions of multiflexible-body systems with reduced-order modal models of the structures is significant in this context.

Acknowledgment

We thank Prof. Thomas R. Kane of Stanford University for a critical review of the manuscript.

References

Banerjee, A. K., and Kane, T. R., 1989, "Dynamics of a Plate in Large Overall Motion," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 887–892.

Banerjee, A. K., and Dickens, J. M., 1990, "Dynamics of an Arbitrary Structure in Large Rotation and Translation," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 2, pp. 221–227.

Bisplinghoff, R. L., Ashley, H., and Halfman, R. L., 1955, *Aeroelasticity*, Addision-Wesley, pp. 95–97.

Bodley, C. S., Devers, A. D., Park, A. C., and Frisch, H. P., 1978, "A Digital Computer Program for the Dynamic Interaction Simulation of Controls and Structures (DISCOS)," Vols. I and II, NASA TP-1219.

Cook, R. D., 1985, *Concepts and Applications of Finite Element Analysis*, McGraw-Hill, New York, pp. 331–341.

Craig, R. R., Jr., 1981, *Structural Dynamics*, John Wiley and Sons, New York.

Eke, F. O., and Laskin, R. A., 1987, "On the Inadequacies of Current Multi-Flexible Body Simulation Codes," AIAA Paper 87-2248, AIAA Guidance, Navigation and Control Conf., Monterey, Calif.

Housner, J., McGowan, P. E., Abrahamson, L. A., and Powell, M. G., 1986, *LATDYN User's Manual*, NASA Technical Memorandum 87635.

Huston, R. L., and Passerello, C. E., 1980, "Multibody Structural Dynamics Including Translation Between the Bodies," *Computers and Structures*, Vol. 12, pp. 713–720.

Ider, S. K., and Amirouche, F. M. L., 1989, "Influence of Geometric Nonlinearities in the Dynamics of Flexible Tree-Line Structures," *Journal of Guidance, Control, and Dynamics*, Vol. 12, No. 6, Nov.-Dec., pp. 830–837.

Kane, T. R., and Levinson, D. A., 1985, *Dynamics: Theory and Applications*, McGraw-Hill, New York.

Kane, T. R., Ryan, R. R., and Banerjee, A. K., 1987, "Dynamics of a Cantilever Beam Attached to a Moving Base," *Journal of Guidance, Control, and Dynamics*, Vol. 10, No. 2, pp. 139–151.

Kane, T. R., Ryan, R. R., Banerjee, A. K., 1989, reply by authors to K. W. London, *Journal of Guidance, Control, and Dynamics*, Vol. 12, No. 2, pp. 286–287.

Levinson, D. A., and Kane, T. R., 1976, "Spin Stability of a Satellite Equipped with Four Booms," *Journal of Spacecraft and Rockets*, Vol. 13, April, pp. 208–213.

Levinson, D. A., and Kane, T. R., 1990, "AUTOLEV—A New Approach to Multibody Dynamics," *Multibody Systems Handbook*, W. Schiehlen, ed., Springer-Verlag, Berlin, pp. 81–102.

Likins, P. W., 1974, "Geometric Stiffness Characteristics of a Rotating Elastic Appendage," *International Journal of Solids and Structures*, Vol. 10, No. 2, pp. 161–167.

Meirovitch, L., 1967, *Analytical Methods in Vibrations*, Macmillan, New York, pp. 443–445.

Modi, V. J., and Ibrahim, A. M., 1988, "On the Dynamics of Flexible Orbiting Structures," *Large Space Structures: Dynamics and Control*, S. N. Atluri and A. K. Amos, eds., Springer-Verlag, New York, pp. 93–114.

Padilla, C. E., and von Flotow, A. H., 1989, "Nonlinear Strain-Displacement Relations and Flexible Multibody Dynamics," *Proceedings of the AIAA/ASME/ASCE/AHS/ASC 30th Structures, Structural Dynamics, and Materials Conf.*, Mobile, Ala., Apr. 3–5, pp. 405–413.

Roberson, R. E., and Schwertassek, R., 1988, *Dynamics of Multibody Systems*, Springer-Verlag, New York.

Rosenthal, D. E., and Sherman, M. A., 1986, "High Performance Multibody Simulations via Symbolic Equation Manipulation and Kane's Method," *Journal of the Astronautical Sciences*, Vol. 34, No. 3, July-Sept., pp. 223-239.

Simo, J. C., and Vu-Quoc, L., 1986, "On the Dynamics of Flexible Bodies Under Large Overall Motions—The Plane Case: Parts I and II," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, Dec., pp. 849-863.

Singh, R. P., van der Voort, R. J., and Likins, P. W., 1985, "Dynamics of Flexible Bodies in Tree Topology—A Computer Oriented Approach," *Journal of Guidance, Control, and Dynamics*, Vol. 8, Sept.-Oct., pp. 584-590.

Turcic, D. A., and Midha, A., 1984, "Generalized Equations of Motion for the Dynamic Analysis of Elastic Mechanism Systems," ASME *Journal of Dynamic Systems, Measurement, and Control*, Vol. 106, Dec., pp. 243-248.

Turcic, D. A., Midha, A., and Bosnik, J. R., 1984, "Dynamic Analysis of Elastic Mechanism Systems. Parts I and II," ASME *Journal of Dynamic Systems, Measurement, and Control*, Vol. 106, December, pp. 249-260.

Vu-Quoc, L., and Simo, J. C., 1987, "Dynamics of Earth-Orbiting Satellites with Multibody Components," *Journal of Guidance, Control, and Dynamics*, Vol. 10, Nov.-Dec., pp. 549-558.

Whetstone, W. D., 1983, *EISI-EAL Engineering Analysis Language Reference Manual*, Engineering Information Systems, Inc., San Jose, CA.

Wittenburg, J., 1977, *Dynamics of Systems of Rigid Bodies*, B. G. Teubner, Stuttgart.

Wu, S.-C., and Haug, E. J., "Geometric Nonlinear Substructuring for Dynamics of Flexible Mechanical Systems," to appear in the *International Journal for Numerical Methods in Engineering*.

Zeiler, T., and Buttrill, C., 1988, "Dynamic Analysis of an Unrestrained Rotating Structure Through Nonlinear Simulation," AIAA 29th Structures, Structural Dynamics and Materials Conference, Williamsburg, Va., Paper No. 88-2232, Apr. 18-20.

Zienkiewicz, O. C., 1977, *The Finite Element Method*, McGraw-Hill, New York, pp. 517-519.

APPENDIX

Modal Integrals Used in Equation (10)

When certain vector-dyadic identities are used in carrying out the operations indicated in equation (9), we get the following modal integrals reported in equation (10). Here, all integrations are over the volume occupied by body B^j; the ranges of the indices are: $(j = 1, \ldots, N)$, $(l, k = 1, \ldots, M_j)$; superscript * on a dyadic indicates its transpose.

1. $\bar{b}^j_k = \int \bar{\phi}^j_k \, dm$

2. $\bar{s}^j = \int \bar{r}^j dm + \sum_{k=1}^{M_j} \bar{b}^j_k \eta^j_k$

3. $\bar{c}^j_k = \int \bar{r}^j \, x \, \bar{\phi}^j_k \, dm$

4. $\bar{d}^j_{lk} = \int \bar{\phi}^j_l \, x \, \bar{\phi}^j_k \, dm$

5. $\bar{g}^j_k = \bar{c}^j_k + \sum_{l=1}^{M_j} \bar{d}^j_{lk} \eta^j_l$

6. $e^j_{lk} = \int \bar{\phi}^j_l \cdot \bar{\phi}^j_k \, dm$

7. $\tilde{N}^j_l = \int [(\bar{r}^j \cdot \bar{\phi}^j_l) \tilde{U} - \bar{r}^j \bar{\phi}^j_l] dm$

8. $\tilde{I}^j = \int [(\bar{r}^j \cdot \bar{r}^j) \tilde{U} - \bar{r} \bar{r}^j] dm + \sum_{i=1}^{M_j} (\tilde{N}^j_l + \tilde{N}^{j*}_l) \eta^j_l$

9. $\tilde{D}^j_l = \tilde{N}^j_l + \sum_{k=1}^{M_j} \eta^j_k \int \left[\left(\bar{\phi}^j_k \cdot \bar{\phi}^j_l \right) \tilde{U} - \bar{\phi}^j_k \bar{\phi}^j_l \right] dm.$

The following additional integrals need to be considered only if rotatory inertia at the nodes is taken into account. Here, the modal expansion of the nodal moment of inertia dyadic involves a coordinate transformation representing small elastic rotations. We also introduce a skew-symmetric matrix $\tilde{\psi}^j_k$ formed with the measure numbers of the small elastic rotation vector $\bar{\psi}^j_k$, just as $\tilde{\omega}^j$ is formed out of the measure numbers of $\bar{\omega}^j$ in equation (24).

10. $\widetilde{W1}^j = \int \widetilde{dI}^j_0$

11. $\widetilde{W2}^j_k = \int \widetilde{dI}^j_k$, where $\widetilde{dI}^j = \widetilde{dI}^j_0 + \sum_{k=1}^{M_j} \widetilde{dI}^j_k \eta^j_k$

12. $\widetilde{W3}^j_k = \int \tilde{\psi}^j_k \widetilde{dI}^{j*}_0$

13. $\overline{W4}^j_k = \int \widetilde{dI}^j_0 \cdot \bar{\psi}^j_k$

14. $W5^j_k = \int dI^j_0 \, x \, \bar{\psi}^j_k$

15. $\overline{W6}^j_k = \int \bar{\psi}^j_k \cdot \widetilde{dI}^j_0$

16. $\overline{W7}^j_{lk} = \int \bar{\psi}^j_l \cdot \widetilde{dI}^j_k$

17. $\overline{W8}^j_{lk} = \int \tilde{\psi}^j_l \widetilde{dI}^j_0 \cdot \bar{\psi}^j_k$

18. $W9^j_{lk} = \int \bar{\psi}^j_l \cdot \widetilde{dI}^j_0 \cdot \bar{\psi}^j_k$

19. $\widetilde{W10}^j_{lk} = \int \tilde{\psi}^j_l \widetilde{dI}^{j*}_k$

20. $\overline{W11}^j_{lk} = \int (\widetilde{dI}^j_0 \cdot \bar{\psi}^j_l) \, x \, \bar{\psi}^j_k$

21. $\overline{W12}^j_{lk} = \int (\widetilde{dI}^j_0 \, x \, \bar{\psi}^j_l) \cdot \bar{\psi}^j_k.$

Integrals numbered 7, 8, 9 are obtained by applying the identity

$$\bar{a} \, x \, (\bar{b} \, x \, \bar{c}) = [(\bar{c} \cdot \bar{a}) \tilde{U} - \bar{c} \, \bar{a}] \cdot \bar{b}$$

where $\tilde{U}$ is a 3×3 unity dyadic. Integrals 12, 17, 19 are obtained by using the following vector-dyadic relations which can be verified by expansion

$$\tilde{A} \cdot (\bar{b} \, x \, \bar{c}) = \bar{b} \cdot \tilde{C} \tilde{A}^*$$

$$\bar{c} \cdot \bar{d} \, x \, \tilde{A} \cdot \bar{b} = -\bar{d} \cdot \tilde{C} \tilde{A}^* \cdot \bar{b}$$

where $\tilde{C}$ is the skew-symmetric matrix formed out of the components of $\bar{c}$, and $\tilde{A}$ is a dyadic.

An-Chen Lee
Assoc. Professor.

Yuan Kang
Research Assistant.

Shin-Li Liu
Research Assistant.

Department of Mechanical Engineering,
National Chiao Tung University,
Hsinchu 30049, Taiwan

A Modified Transfer Matrix Method for Linear Rotor-Bearing Systems

The steady-state responses of linear flexible rotor-bearing systems are analyzed by the modified transfer matrix method. The transfer matrix has the advantage of solving the problems in frequency domain with fixed matrix size. This makes the method more economical in analyzing a large degree-of-freedom rotor system than many time-marching integrating methods. In this paper, the modifications of transfer matrix method include that the transfer matrix of shaft is derived from the "continuous system" concept instead of conventional "lumped system" concept, and the paper tries to extend the transfer matrix method to fit synchronous elliptical orbit and nonsynchronous multi-lobed whirling orbit. To demonstrate the applications of the method, three examples are presented; two synchronous and one nonsynchronous.

Introduction

Rotor dynamics plays an important role in many fields of engineering, such as the gas and steam turbines, turbogenerators, reciprocating and centrifugal compressors. On account of the ever-increasing demands for high power, high speed, and light weight which are the main reasons of failure in performances and fatigue in structures of the rotor-bearing systems, the designers need to have some new techniques for the prediction of critical speeds, unbalance responses, and threshold speeds of instability for synchronous and nonsynchronous whirling.

For the linearized model, analysis techniques to estimate the dynamic characteristics of rotor systems have been fairly well established. Currently, there are three main methods to analyze rotor bearing systems, i.e., finite element method, mode synthesis method, and transfer matrix method. The transfer matrix method solves dynamic problems in the frequency domain that makes itself suitable to study the steady-state behavior of the rotor with the advantages of small computer memory requirement and satisfactory accuracy. Because the equations of motion derived by such a method are not in explicitly written form, it is also useful in optimizing eigensolutions for adjusting design factors.

The transfer matrix method was first proposed by Prohl (1945). Later, the effects of damping and stiffness of the fluid film bearing were included by Koenig (1961) and Guenther and Lovejoy (1961). Lund (1974) and Bansal and Kirk (1975) applied the transfer matrix method in modal analysis for calculating damped natural frequencies and examining the stability of flexible rotors supported by fluid film bearings. Lund (1980) presented a scheme for estimating the sensitivity of the critical speeds of a rotor to changes in the design factors. In the aforementioned papers, the rotor-bearing system is modeled as a connection of shaft "station" which is normally considered to consist of a mass plus the shaft section immediately to its right. Then, one assigns degrees-of-freedom (coordinates) at the junctions between the stations (i.e., at each concentrated mass). The linear differential equations are written for each station and are arranged in matrix form. A single transfer matrix, which fully represents the entire system, can be obtained by multiplying together all the transfer matrices for the system.

Lund and Orcutt (1967) modified the transfer matrix method by allowing a continuous representation of the shaft and investigating the unbalance vibrations of a rotor analytically and experimentally. In their work, it is assumed that the effects of rotary inertia and gyroscopic moment are neglected in the shaft section. Furthermore, since only eight state variables are considered, the general ellipse of whirling orbits with arbitrary tilt could not be obtained. Gu (1986) proposed an improved transfer matrix-direct integration (ITMDI) method for rotor dynamics. He attempted to combine transfer matrix and direct integration methods while incorporating the advantages of both. The assumptions he made that the whirling orbits are circular and that the bearings have no damping effect are not realistic. In practice, the fact that the bearings may possess considerable damping and anisotropic characteristics, or that the cross-sections of shaft and disk may have asymmetric inertias of moment, would cause the induced whirling orbits to be elliptical in general (Lund, 1974, pp. 525–533).

In this work, attempts have been made to represent the transfer matrix of shaft by a continuous system instead of the

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 7, 1989; final revision, Mar. 7, 1990.

lumped system representation and to describe the whirling orbit of a rotor system in a general way. The rotary inertia, gyroscopic, and transverse shear effects are included. Small deflections are assumed throughout the derivation of the analysis. Material damping in the shaft and external damping effects, such as air resistance, are not included in the formulation. Three examples are presented to show the applicability of the modified transfer matrix for steady-state analysis in both synchronous and nonsynchronous whirlings.

Transfer Matrix Analysis

1 Modified Transfer Matrix for an Uniform Shaft. From the Timoshenko's beam theory, the equations of motion of a uniform and homogeneous shaft are given in Eshleman and Eubanks (1969) and written down as follows: In the XZ-plane it is

$$\frac{\partial^4 X}{\partial Z^4}-\left(\frac{\rho}{KG}+\frac{\rho}{E}\right)\frac{\partial^4 X}{\partial Z^2\partial t^2}+\frac{\rho^2}{KGE}\frac{\partial^4 X}{\partial t^4}+\frac{\rho A}{EI}\frac{\partial^2 X}{\partial t^2}-\frac{2\rho\omega}{E}\left(\frac{\partial^3 Y}{\partial Z^2\partial t}-\frac{\rho}{KG}\frac{\partial^3 Y}{\partial t^3}\right)=0. \quad (1a)$$

The motion equation of shaft in the YZ-plane is

$$\frac{\partial^4 Y}{\partial Z^4}-\left(\frac{\rho}{KG}+\frac{\rho}{E}\right)\frac{\partial^4 X}{\partial Z^2\partial t^2}+\frac{\rho^2}{KGE}\frac{\partial^4 Y}{\partial t^4}+\frac{\rho A}{EI}\frac{\partial^2 Y}{\partial t^2}+\frac{2\rho\omega}{E}\left(\frac{\partial^3 X}{\partial Z^2\partial t}-\frac{\rho}{KG}\frac{\partial^3 X}{\partial t^3}\right)=0. \quad (1b)$$

Because the synchronous whirling orbit is elliptical in general, the solutions of two linear differential equations (LDE) (1a) and (1b) can be represented in the following form

$$X(Z,t)=X_c(Z)\cos\Omega t+X_s(Z)\sin\Omega t$$
$$Y(Z,t)=Y_c(Z)\cos\Omega t+Y_s(Z)\sin\Omega t. \quad (2)$$

The representations shown above are steady-state solutions of the LDE. The whirling orbits due to these displacement functions are ellipses. It is noted that X_c, X_s, Y_c, and Y_s are the mode functions. The representations of the other state variables, i.e., slopes, moments, and shear forces are the same as those of X and Y. Their coefficients are α_c, β_c, M_{xc}, M_{yc}, Q_{xc}, Q_{yc}, and α_s, β_s, M_{xs}, M_{ys}, Q_{xs}, Q_{ys} associated to respective cosine terms and sine terms in the XY or YZ-plane.

Substituting equations (2) into (1), then separating the terms associated with $\cos\Omega t$ and $\sin\Omega t$, respectively, would result in four homogeneous equations. Furthermore, they can be combined into two complex equations by introducing the complex variables, $\bar{X}=X_c+jX_s$, and $\bar{Y}=Y_c+jY_s$.

In the XZ-plane,

$$\frac{d^4\bar{X}}{dZ^4}+\left(\frac{\rho\Omega^2}{E}+\frac{\rho\Omega^2}{KG}\right)\frac{d^2\bar{X}}{dZ^2}+\left(\frac{\rho^2\Omega^4}{KGE}-\frac{\rho A\Omega^2}{EI}\right)\bar{X}-\frac{j2\rho\omega\Omega}{E}\frac{d^2\bar{Y}}{dZ^2}-\frac{j2\omega\rho^2\Omega^3}{KGE}\bar{Y}=0. \quad (3a)$$

In the YZ-plane,

$$\frac{d^4\bar{Y}}{dZ^4}+\left(\frac{\rho\Omega^2}{E}+\frac{\rho\Omega^2}{KG}\right)\frac{d^2\bar{Y}}{dZ^2}+\left(\frac{\rho^2\Omega^4}{KGE}-\frac{\rho A\Omega^2}{EI}\right)\bar{Y}+\frac{j2\rho\omega\Omega}{E}\frac{d^2\bar{X}}{dZ^2}+\frac{j2\omega\rho^2\Omega^3}{KGE}\bar{X}=0. \quad (3b)$$

The solutions of equations (3) are in the forms of

$$\bar{X}(Z)=X_c+jX_s=U_c e^{\lambda Z}+jU_s e^{\lambda Z}$$
$$\bar{Y}(Z)=Y_c+jY_s=V_c e^{\lambda Z}+jV_s e^{\lambda Z} \quad (4)$$

where U_c, U_s, V_c, and V_s are the arbitrary real constants, and λ is characteristic value w.r.t., a specific natural mode. Then, by substituting equation (4) into (3) and separating the real part and imaginary part, the four algebraic equations are obtained. For U_c, U_s, V_c and V_s being nontrivial, the characteristic equations become

$$\begin{vmatrix} \lambda^4+f\lambda^2+g & 0 & 0 & -(h\lambda^2+k) \\ 0 & \lambda^4+f\lambda^2+g & h\lambda^2+k & 0 \\ 0 & h\lambda^2+k & \lambda^4+f\lambda^2+g & 0 \\ -(h\lambda^2+k) & 0 & 0 & \lambda^4+f\lambda^2+g \end{vmatrix}=0 \quad (5)$$

where

$$f=\frac{\rho\Omega^2}{E}+\frac{\rho\Omega^2}{KG},\ g=\frac{\rho^2\Omega^4}{KGE}-\frac{\rho A\Omega^2}{EI},\ h=\frac{2\rho\omega\Omega}{E},\ k=\frac{2\omega\rho^2\Omega^3}{KGE}$$

in which Ω is replaced by ω when the whirling is synchronous; otherwise, Ω is substituted by $n\omega$ when the whirling is nonsynchronous. It can be shown that the roots of the characteristic function of equation (5) are

$$\lambda=\pm\lambda_a,\ \pm j\lambda_b,\ \pm\lambda_c,\ \pm j\lambda_d$$

for a constant value of ω. Normally, the characteristic roots of the Timoshenko's beam are not much different from those of Euler's beam, and they are not equal to each other (Gu, 1986). If we substitute eigenvalues into the characteristic equations the relations of $V_c=U_s$, and $V_s=-U_c$ for λ_a or λ_b, and $V_c=-U_s$, and $V_s=U_c$ for λ_c or λ_d are obtained. Therefore, the mode functions can be written as follows:

Nomenclature

- X, Y = deflection in XOZ and YOZ-plane
- E = Young's modulus
- G, K = shear modulus and shear factor
- α, β = slope in XOZ and YOZ-plane
- M = bending moment
- Q = shearing force
- A = cross-section area of shaft
- ρ = density
- I, J = transverse and polar area moment of inertia of the shaft
- h = thickness of disk
- I_d, J_p = transverse and polar mass moment of inertia of the disk
- ω, Ω = rotating and whirling speed, respectively
- e = eccentricity of disk
- $K_{xx}, K_{yy}, K_{xy}, K_{yx}$ = spring constants of bearing
- $C_{xx}, C_{yy}, C_{xy}, C_{yx}$ = damping coefficients of bearing

Subscripts

- x, y = in XZ, YZ-plane
- b, t = caused by bending, shearing
- r, l = right, left
- c, s = associated to cosine and sine terms
- $0, n$ = stage number

$$
\begin{aligned}
X_c(Z) &= A_1\cosh\lambda_a Z + A_2\sinh\lambda_a Z + A_3\cos\lambda_b Z + A_4\sin\lambda_b Z \\
&\quad + A_5\cosh\lambda_c Z + A_6\sinh\lambda_c Z + A_7\cos\lambda_d Z + A_8\sin\lambda_d Z \\
X_s(Z) &= B_1\cosh\lambda_a Z + B_2\sinh\lambda_a Z + B_3\cos\lambda_b Z + B_4\sin\lambda_b Z \\
&\quad + B_5\cosh\lambda_c Z + B_6\sinh\lambda_c Z + B_7\cos\lambda_d Z + B_8\sin\lambda_d Z \\
Y_c(Z) &= B_1\cosh\lambda_a Z + B_2\sinh\lambda_a Z + B_3\cos\lambda_b Z + B_4\sin\lambda_b Z \\
&\quad - B_5\cosh\lambda_c Z - B_6\sinh\lambda_c Z - B_7\cos\lambda_d Z - B_8\sin\lambda_d Z \\
Y_s(Z) &= -A_1\cosh\lambda_a Z - A_2\sinh\lambda_a Z - A_3\cos\lambda_b Z - A_4\sin\lambda_b Z \\
&\quad + A_5\cosh\lambda_c Z + A_6\sinh\lambda_c Z + A_7\cos\lambda_d Z + A_8\sin\lambda_d Z.
\end{aligned}
\tag{6}
$$

The real constants, A_i and B_i, which can be expressed in terms of the boundary conditions at the left end of the shaft (i.e., $Z=0$), are shown as follows:

$$
\begin{bmatrix} X_c(0) \\ X_c'(0) \\ X_c''(0) \\ X_c'''(0) \\ \hdashline Y_s(0) \\ Y_s'(0) \\ Y_s''(0) \\ Y_s'''(0) \end{bmatrix}
=
\left[\begin{array}{cccc:cccc}
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & \lambda_a & 0 & \lambda_b & 0 & \lambda_c & 0 & \lambda_d \\
\lambda_a^2 & 0 & -\lambda_b^2 & 0 & \lambda_c^2 & 0 & -\lambda_d^2 & 0 \\
0 & \lambda_a^3 & 0 & -\lambda_b^3 & 0 & \lambda_c^3 & 0 & -\lambda_d^3 \\
\hdashline
-1 & 0 & -1 & 0 & 1 & 0 & 1 & 0 \\
0 & -\lambda_a & 0 & -\lambda_b & 0 & \lambda_c & 0 & \lambda_d \\
-\lambda_a^2 & 0 & \lambda_b^2 & 0 & \lambda_c^2 & 0 & -\lambda_d^2 & 0 \\
0 & -\lambda_a^3 & 0 & \lambda_b^3 & 0 & \lambda_c^3 & 0 & -\lambda_d^3
\end{array}\right]
\begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ \hdashline A_5 \\ A_6 \\ A_7 \\ A_8 \end{bmatrix}.
$$

For simplification, the following expression is introduced

$$
\begin{bmatrix} \mathbf{X}_c(0) \\ \hdashline \mathbf{Y}_s(0) \end{bmatrix}
=
\left[\begin{array}{c:c} M_1 & M_2 \\ \hdashline -M_1 & M_2 \end{array}\right]
\{A\} = [M_a]\{A\}.
\tag{7a}
$$

Similarly, the relationships between $X_s(0), \cdots, Y_s(0), \cdots$, and B_i are written as

$$
\begin{bmatrix} \mathbf{X}_s(0) \\ \mathbf{Y}_c(0) \end{bmatrix}
=
\begin{bmatrix} M_1 & M_2 \\ M_1 & -M_2 \end{bmatrix}
\{B\} = [M_b]\{B\}.
\tag{7b}
$$

From equations (7), the coefficients of mode functions, which are in terms of deflections and their derivatives at $Z=0$, are obtained by premultiplication of $[M_a]^{-1}$ or $[M_b]^{-1}$ in both sides of equation (7). Let $\cosh\lambda_a Z = C_1$, $\sinh\lambda_a L = C_2$, $\cos\lambda_b L = C_3$, $\sin\lambda_b L = C_4$, $\cosh\lambda_c Z = C_5$, $\sinh\lambda_c L = C_6$, $\cos\lambda_d L = C_7$, $\sin\lambda_d L = C_8$. The deflections and their derivatives at $Z = L$ can be obtained from equation (6),

$$
\begin{bmatrix} X_c(L) \\ X_c'(L) \\ X_c''(L) \\ X_c'''(L) \\ \hdashline Y_s(L) \\ Y_s'(L) \\ Y_s''(L) \\ Y_s'''(L) \end{bmatrix}
=
\left[\begin{array}{cccc:cccc}
C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 & C_8 \\
\lambda_a C_2 & \lambda_a C_1 & -\lambda_b C_4 & \lambda_b C_3 & \lambda_c C_6 & \lambda_c C_5 & -\lambda_d C_8 & \lambda_d C_7 \\
\lambda_a^2 C_1 & \lambda_a^2 C_2 & -\lambda_b^2 C_3 & -\lambda_b^2 C_4 & \lambda_c^2 C_5 & \lambda_c^2 C_6 & -\lambda_d^2 C_7 & -\lambda_d^2 C_8 \\
\lambda_a^3 C_2 & \lambda_a^3 C_1 & \lambda_b^3 C_4 & -\lambda_b^3 C_3 & \lambda_c^3 C_6 & \lambda_c^3 C_5 & \lambda_d^3 C_8 & -\lambda_d^3 C_7 \\
\hdashline
-C_1 & -C_2 & -C_3 & -C_4 & C_5 & C_6 & C_7 & C_8 \\
-\lambda_a C_2 & -\lambda_a C_1 & \lambda_b C_4 & -\lambda_b C_3 & \lambda_c C_6 & \lambda_c C_5 & -\lambda_d C_8 & \lambda_d C_7 \\
-\lambda_a^2 C_1 & -\lambda_a^2 C_2 & \lambda_b^2 C_3 & \lambda_b^2 C_4 & \lambda_c^2 C_5 & \lambda_c^2 C_6 & -\lambda_d^2 C_7 & -\lambda_d^2 C_8 \\
-\lambda_a^3 C_2 & -\lambda_a^3 C_1 & -\lambda_b^3 C_4 & \lambda_b^3 C_3 & \lambda_c^3 C_6 & \lambda_c^3 C_5 & \lambda_d^3 C_8 & -\lambda_d^3 C_7
\end{array}\right]
\begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ \hdashline A_5 \\ A_6 \\ A_7 \\ A_8 \end{bmatrix}
$$

or, in the simplified form as

$$
\begin{bmatrix} \mathbf{X}_c(L) \\ \mathbf{Y}_s(L) \end{bmatrix}
=
\begin{bmatrix} H_1 & H_2 \\ -H_1 & H_2 \end{bmatrix}
\{A\} = [H_a]\{A\}.
\tag{8a}
$$

Similarly,

$$
\begin{bmatrix} \mathbf{X}_c(L) \\ \mathbf{Y}_c(L) \end{bmatrix}
=
\begin{bmatrix} H_1 & H_2 \\ H_1 & -H_2 \end{bmatrix}
\{B\} = [H_b]\{B\}.
\tag{8b}
$$

Substituting (7a) into (8a), and (7b) into (8b), we obtain

$$
\begin{bmatrix} \mathbf{X}_c(L) \\ \mathbf{Y}_s(L) \end{bmatrix}
= [H_a][M_a]^{-1}
\begin{bmatrix} \mathbf{X}_c(0) \\ \mathbf{Y}_s(0) \end{bmatrix}
=
\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}
\begin{bmatrix} \mathbf{X}_c(0) \\ \mathbf{Y}_s(0) \end{bmatrix}
\tag{9a}
$$

and

$$
\begin{bmatrix} \mathbf{X}_s(L) \\ \mathbf{Y}_c(L) \end{bmatrix}
= [H_b][M_b]^{-1}
\begin{bmatrix} \mathbf{X}_s(0) \\ \mathbf{Y}_c(0) \end{bmatrix}
=
\begin{bmatrix} N_{11}' & N_{12}' \\ N_{21}' & N_{22}' \end{bmatrix}
\begin{bmatrix} \mathbf{X}_s(0) \\ \mathbf{Y}_c(0) \end{bmatrix}
\tag{9b}
$$

Combining two equations (9a) and (9b), results in

$$
\{W(Z=L)\} =
\begin{bmatrix} \mathbf{X}_c(L) \\ \mathbf{X}_s(L) \\ \mathbf{Y}_c(L) \\ \mathbf{Y}_s(L) \end{bmatrix}
=
\begin{bmatrix} N_{11} & 0 & 0 & N_{12} \\ 0 & N_{11}' & N_{12}' & 0 \\ 0 & N_{21}' & N_{22}' & 0 \\ N_{21} & 0 & 0 & N_{22} \end{bmatrix}
\begin{bmatrix} \mathbf{X}_c(0) \\ \mathbf{X}_s(0) \\ \mathbf{Y}_c(0) \\ \mathbf{Y}_s(0) \end{bmatrix}
= [N]\{W(Z=0)\}
\tag{10}
$$

where 0 is 4×4 null matrix.

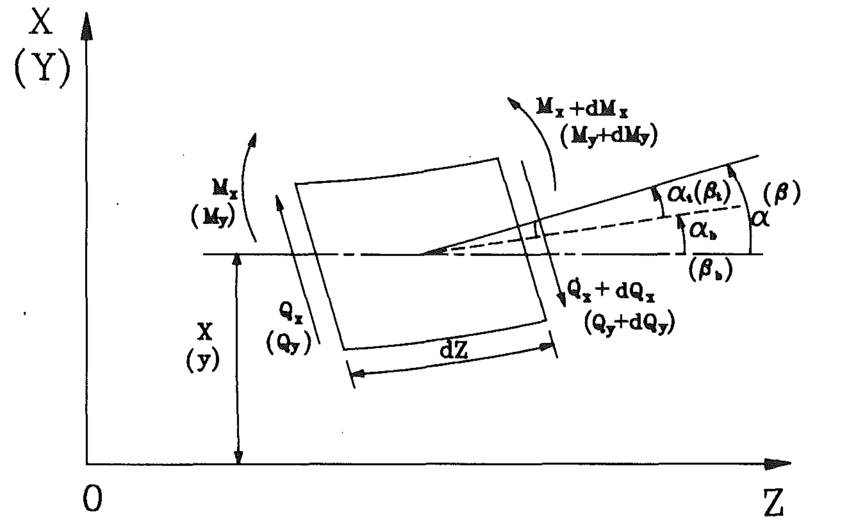


Fig. 1 Element of a shaft in the *XZ*-plane or *YZ*-plane

From the theory of Timoshenko's beam (Dym and Shames, 1973), we have the compatibility relations of the shaft projected in the XZ-plane,

$$X = X_b + X_t$$

$$\alpha = \frac{\partial X}{\partial Z} = \frac{\partial X_b}{\partial Z} + \frac{\partial X_t}{\partial Z} = \alpha_c \cos \Omega t + \alpha_s \sin \Omega t = \alpha_b + \alpha_t$$

$$M_x = EI\frac{\partial^2 X_b}{\partial Z^2} = M_{xc} \cos \Omega t + M_{xs} \sin \Omega t$$

$$Q_x = -KAG\left(\frac{\partial X}{\partial Z} - \frac{\partial X_b}{\partial Z}\right) = Q_{xc} \cos \Omega t + Q_{xs} \sin \Omega t. \qquad (11a)$$

The similar relations in the YZ-plane are

$$Y = Y_b + Y_t$$

$$\beta = \frac{\partial Y}{\partial Z} = \frac{\partial Y_b}{\partial Z} + \frac{\partial Y_t}{\partial Z} = \beta_c \cos \Omega t + \beta_s \sin \Omega t = \beta_b + \beta_t$$

$$M_y = EI\frac{\partial^2 Y_b}{\partial Z^2} = M_{yc} \cos \Omega t + M_{ys} \sin \Omega t$$

$$Q_y = -KAG\left(\frac{\partial Y}{\partial Z} - \frac{\partial Y_b}{\partial Z}\right) = Q_{yc} \cos \Omega t + Q_{ys} \sin \Omega t. \qquad (11b)$$

Considering a uniform shaft segment, the projections of the element in the XZ and YZ-planes are shown in Fig. 1. The force and moment equilibrium conditions are

$$\frac{\partial Q_x}{\partial Z} = -\rho A \ddot{X}$$

$$Q_x = \frac{\partial M_x}{\partial Z} - \rho I\left(\frac{\partial^2 \alpha_b}{\partial t^2} + 2\omega\,\frac{\partial \beta_b}{\partial t}\right)$$

$$\frac{\partial Q_y}{\partial Z} = -\rho A \ddot{Y}$$

$$Q_y = \frac{\partial M_y}{\partial Z} - \rho I\left(\frac{\partial^2 \beta_b}{\partial t^2} + 2\omega\frac{\partial \alpha_b}{\partial t}\right). \qquad (12)$$

Since the element is considered as a circular thin disk in this study, we have $J = 2I$. From the relations of equations (9)–(11), the following relationships presented in the complex form between the differential variables of deflection and the state variables can be obtained.

In the XZ-plane,

$$\bar{X} = X_c + jX_s$$

$$\bar{X}' = \bar{\alpha} + \frac{\bar{Q}_x}{kGA} \qquad (13a)$$

$$\bar{X}'' = \frac{\bar{M}_x}{EI} + \frac{\rho\Omega^2}{KG}\bar{X}$$

$$\bar{X}''' = \left(\frac{\rho\Omega^2}{KG} - \frac{\rho\Omega^2}{E}\right)\bar{\alpha} + \frac{2\rho\omega\Omega}{E} j\bar{\beta} + \left[\frac{1}{EI} - \frac{\rho}{kGA}\left(\frac{1}{E} - \frac{1}{kG}\right)\Omega^2\right]\bar{Q}_x + \frac{2\rho\omega\Omega}{kGAE}\bar{Q}_y.$$

In the YZ-plane,

$$\bar{Y} = Y_c + jY_s$$

$$\bar{Y}' = \bar{\beta} + \frac{\bar{Q}_y}{kGA}$$

$$\bar{Y}'' = \frac{\bar{M}_y}{EI} + \frac{\rho\Omega^2}{KG}\bar{Y}$$

$$\bar{Y}''' = \left(\frac{\rho\Omega^2}{KG} - \frac{\rho\Omega^2}{E}\right)\bar{\beta} + \frac{2\rho\omega\Omega}{E} j\bar{\alpha} + \left[\frac{1}{EI} - \frac{\rho}{kGA}\left(\frac{1}{E} - \frac{1}{kG}\right)\Omega^2\right]\bar{Q}_y + \frac{2\rho\omega\Omega}{kGAE}\bar{Q}_x$$

in which $\bar{\alpha} = \alpha_c + j\alpha_s$, $\bar{\beta} = \beta_c + j\beta_s$, $\bar{M}_x = M_{xc} + jM_{xs}$, $\bar{M}_y = M_{yc} + jM_{ys}$, $\bar{Q}_x = Q_{xc} + jQ_{xs}$, and $\bar{Q}_y = Q_{yc} + jQ_{ys}$. The above equations could be represented in a matrix form

$$\{W\} = [A]\{S\} \qquad (14)$$

where

$\{W\} = (X_c, X_c', X_c'', X_c''', X_s, X_s', X_s'', X_s''', Y_c, Y_c', Y_c'', Y_c''', Y_s, Y_s', Y_s'', Y_s''')^t$, and $\{S\} = (X_c, X_s, Y_c, Y_s, \alpha_c, \alpha_s, \beta_c, \beta_s, M_{xc}, M_{xs}, M_{yc}, M_{ys}, Q_{xc}, Q_{xs}, Q_{yc}, Q_{ys})^t$, in which "$t$" denotes the transpose of an array.

By considering the boundary conditions at $Z = 0$ and $Z = L$, we have

$$\{W(Z=L)\} = [A]\ \{S(Z=L)\} = [A]\{S_1\}$$

$$\{W(Z=0)\} = [A]\ \{S(Z=0)\} = [A]\{S_0\}. \qquad (15)$$

Substituting the above equation into (10) results in

$$\{S_1\} = [A]^{-1}\ [N]\ [A]\ \{S_0\} = [T]\{S_0\}. \qquad (16)$$

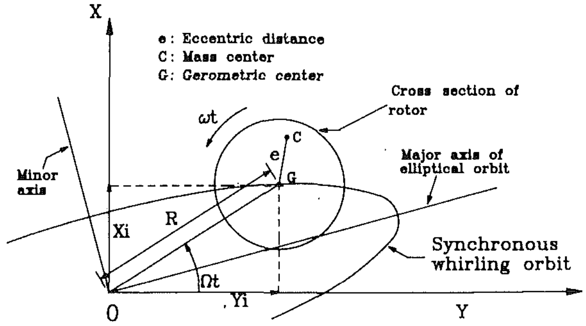


Fig. 2 Response of a synchronous whirling orbit

Hence, the transfer matrix $[T]$ of a shaft is modified by the size of 16×16 to fit general whirling in the elliptical orbits. In consequence, a uniform shaft segment can be taken as long as possible for decreasing the numbers of matrix multiplications and increasing the accuracy of computation.

For the purpose of response calculations we augment the transfer matrix to size of 17×17 and add an identity 1 at the last diagonal element as

$$\begin{Bmatrix} \mathbf{S}_1 \\ 1 \end{Bmatrix} = \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} \mathbf{S}_0 \\ 1 \end{Bmatrix}. \tag{17}$$

2 The Transfer Matrix of the Disk. The disk is assumed to be a rigid body with the gyroscopic effect. The transfer matrix of the disk is defined as a point matrix which considers the effect of mass at the station.

Unbalance mass on the disk induces the exciting force and will cause the synchronous whirling. The whirling orbit of the rotor system at the disk is shown in Fig. 2 with the unbalance given in a plane. From the relations of the forces and deflections of the disk in a plane (see Fig. 3), the equilibrium and compatibility conditions may be shown as follows:

$$\begin{aligned}
X_r &= X_1 + \alpha_1 h \\
\alpha_r &= \alpha_1 \\
M_{xr} &= M_{x1} + I_d\,\ddot{\alpha}_1 + J_p\omega\dot{\beta}_1 - m_d\left(\ddot{X}_1 + \frac{1}{2}\,\ddot{\alpha}_1\,h\right)\frac{h}{2} \\
Q_{xr} &= Q_{x1} - m_d\left(\ddot{X}_1 + \frac{1}{2}\,\ddot{\alpha}_1\,h\right) - m_d\,\omega^2 e_x\,(\cos\Omega t - \sin\Omega t) \\
Y_r &= Y_1 + \beta_1 h \\
\beta_r &= \beta_1 \\
M_{yr} &= M_{y1} + I_d\ddot{\beta}_1 + J_p\omega\dot{\alpha}_1 - m_d\left(\ddot{Y}_1 + \frac{1}{2}\,\ddot{\beta}_1 h\right)\frac{h}{2} \\
Q_{yr} &= Q_{y1} - m_d\left(\ddot{Y}_1 + \frac{1}{2}\,\ddot{\beta}_1 h\right) - m_d\omega^2 e_y\,(\cos\Omega t + \sin\Omega t)
\end{aligned} \tag{18}$$

where m_d is the mass of disk, and $J_p = 2I_d$. The unbalance forces $m_d\omega^2 e_x$, $m_d\omega^2 e_y$ are functions of rotating speed. The relationship of the state variables between the right side and left side of an unbalance disk is directly derived from equation (18). It is shown that

$$\begin{bmatrix} \mathbf{S}_r \\ 1 \end{bmatrix} = [D]_{17\times17} \begin{bmatrix} \mathbf{S}_l \\ 1 \end{bmatrix} \tag{19}$$

where "r" and "l" represent the right and left of disk, respectively. The elements of the matrix $[D]$ are

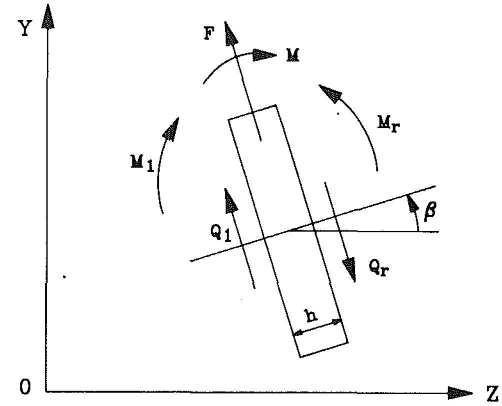


Fig. 3 Forces acting on a disk

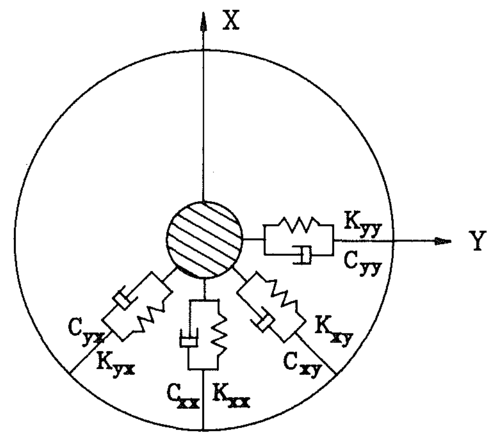


Fig. 4 Model of a bearing

$$\begin{aligned}
&D_{i,\,i} = 1,\ i = 1, 2, 3, \ldots, 17 \\
&D_{1,\,5} = D_{2,\,6} = D_{3,\,7} = D_{4,\,8} = h \\
&D_{9,\,1} = D_{10,\,2} = D_{11,\,3} = D_{12,\,4} = D_{13,\,5} = D_{14,\,6} = D_{15,\,7} \\
&\qquad = D_{16,\,8} = \frac{M_d h\Omega^2}{2} \\
&D_{9,\,5} = D_{10,\,6} = D_{11,\,7} = D_{12,\,8} = \frac{M_d h^2\Omega^2}{4} - J_d\Omega^2 \\
&D_{13,\,1} = D_{14,\,2} = D_{15,\,3} = D_{16,\,4} = M_d\Omega^2 \\
&D_{9,\,8} = -D_{10,\,7} = D_{11,\,6} = -D_{12,\,5} = J_p\omega\Omega \\
&-D_{13,\,17} = D_{14,\,17} = m_d\,\omega^2\,e_x \\
&D_{15,\,17} = D_{16,\,17} = -m_d\,\omega^2\,e_y
\end{aligned} \tag{20}$$

and the others are equal to zero.

3 The Transfer Matrix of a Support Bearing. In the rotor system, the fluid film bearings play an important role in the dynamic behavior of the system. Since the squeezed thin film between the journal and the ring acts to provide the effects of spring and damping, the dynamic properties of the fluid film bearings will dominate the critical speeds, the unbalance responses, and threshold speeds of instability, which is significantly, and totally different from those of the rigid supports. However, these fluid films are very complicated in operation. When the fluid film bearing is analyzed, the physical model of fluid bearing may be simplified as a linear element and represented by eight coefficients, i.e., two direct stiffness K_{xx},

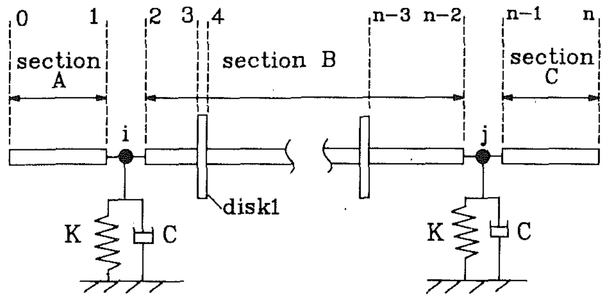


Fig. 5 A general rotor-bearing system

Table 1 The details of the three-disk rotor system

The coefficients of shaft	
Young's modulus E	2.07×10^{7} N/cm^2
shear modulus G	8.1×10^{6} N/cm^2
shear factor K	0.68
density ρ	7.75×10^{-3} Kg/cm^3
The coefficients of disks	
disk mass M_d	13.47 Kg
polar moment of inertia	1020 Kg·cm^2
transverse moment of inertia	512 Kg·cm^2
unbalance at the disk 1	0.01347 Kg·cm
The coefficients of bearings	
direct stiffness	$K_{xx}=K_{yy}=1.0\times10^{7}$
cross stiffness	$K_{xy}=K_{yx}=5.\times10^{6}$
direct damping	$C_{xx}=C_{yy}=2000.$
cross damping	$C_{xy}=C_{yx}=0.$

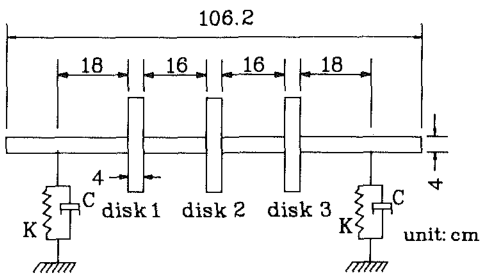


Fig. 6 Configuration of the three-disk rotor

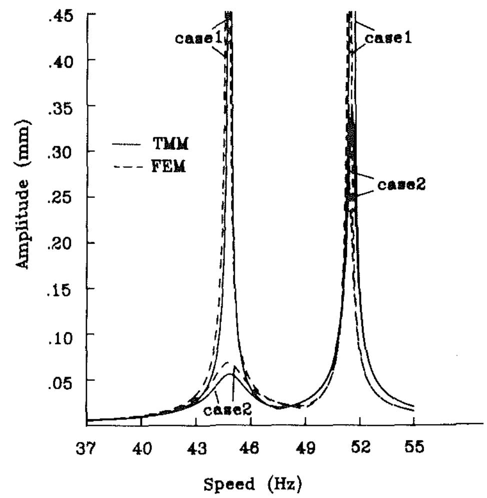


Fig. 7 Unbalance responses of disk 1 by FEM and modified transfer matrix method (TMM)

K_{yy}, two cross stiffness K_{xy}, K_{yx}, two direct damping C_{xx}, C_{yy}, and two cross damping C_{xy}, C_{yx} (see Fig. 4). The dynamic properties of such bearing a model are derived by treating the fluid film as a laminar flow and are governed by a Reynold's equation (Rao, 1983).

The equilibrium relations of the bearing force at the XZ and YZ-planes can be written as

$$Q_{xr}=Q_{x1}-K_{xx}X_1-K_{xy}Y_1-C_{xx}\dot{X}_1-C_{xy}\dot{Y}_1+m_b\,\Omega^2\,X_1$$

$$Q_{y1}=Q_{y1}-K_{yy}Y_1-K_{yx}X_1-C_{yy}\dot{Y}_1-C_{yx}\dot{X}_1+m_b\,\Omega^2\,Y_1 \qquad (21)$$

where m_b is the distributed mass of the shaft supported by the specific bearing. From the foregoing discussion, the relation of the state variables between the right and left of the bearing is

$$\begin{bmatrix}\mathbf{S}_r\\1\end{bmatrix}=[B]_{17\times17}\begin{bmatrix}\mathbf{S}_l\\1\end{bmatrix} \qquad (22)$$

where the elements of the matrix $[B]$ are

$$\begin{aligned}
&B_{i,\,i}=1,\ i=1,\,2,\ldots,\,17\\
&B_{13,\,3}=B_{14,\,4,}=-K_{xy}\\
&B_{14,\,1}=-B_{13,\,2}=C_{xx}\\
&B_{13,\,4}=-B_{14,\,3}=-C_{xy}\\
&B_{16,\,1}=-B_{15,\,2}=C_{yx}\\
&B_{16,\,2}=-B_{15,\,1}=K_{yx}\\
&B_{16,\,3}=-B_{15,\,4}=C_{yy}\\
&B_{13,\,1}=B_{14,\,2}=m_b\,\Omega^2-K_{xx}\\
&B_{15,\,3}=B_{16,\,4}=m_b\,\Omega^2-K_{yy}
\end{aligned} \qquad (23)$$

and the others are equal to zero.

Numerical Examples

The study is first based on the assumption that the whirling frequency is equal to the rotating frequency, i.e., synchronous whirling. In practice, it is predominantly excited by the mass unbalances of a rotor system.

As in the previous sections, since the transfer matrix for each component (i.e., bearing, shaft, and disk) is constructed, it can be used for calculating the unbalance response, critical speeds, modes, whirling orbits, and threshold speeds of instability of a rotor-bearing system. Consider a rotor-bearing system shown in Fig. 5, which has multidisks and bearings, and has an unbalance eccentricity e_i at each disk, with both ends of the shaft free. The overall transfer matrix of the rotor can be represented by

$$\begin{aligned}\{S_n\}&=[U]\ \{S_0\}\\ &=[T][B][T]\ldots[T][D][T]\ldots[B][T]\{S_0\}.\end{aligned} \qquad (24)$$

A simplified rotor-bearing system shown in Fig. 6 is used to demonstrate the applicability of the analytical results. The rotor supported by two bearings has three disks of equal mass with an unbalance mass at disk 1 only. The bearing is approximated by different dynamic characteristics in order to compare their effects on the rotor dynamics. The bearing forces are considered as concentrated forces. The details of the rotor are listed in Table 1. There are two cases to be investigated:

Case 1 bearing without damping and

Case 2 bearing with both stiffness and damping.

Because the shear forces and bending moments are zero at both ends, equation (24) becomes

$$\begin{bmatrix}\mathbf{S}'_n\\\mathbf{0}\\1\end{bmatrix}=\begin{bmatrix}U_{11}&U_{12}&\mathbf{u}_1\\U_{21}&U_{22}&\mathbf{u}_2\\0&0&1\end{bmatrix}\begin{bmatrix}\mathbf{S}'_0\\\mathbf{0}\\1\end{bmatrix} \qquad (25)$$

where $\mathbf{S}'=\{S'\}=(X_c,\,X_s,\,Y_c,\,Y_s,\,\alpha_c,\,\alpha_s,\,\beta_c,\,\beta_s)'$, $\mathbf{0}=(0,\,0,$

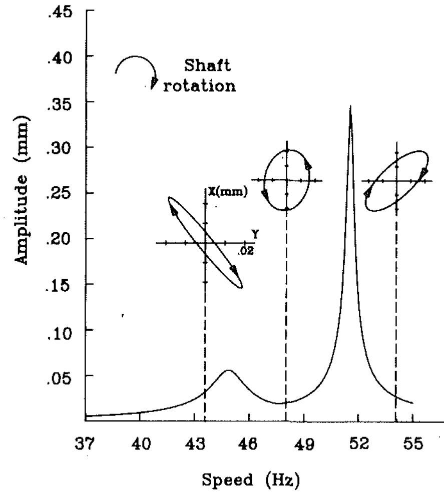
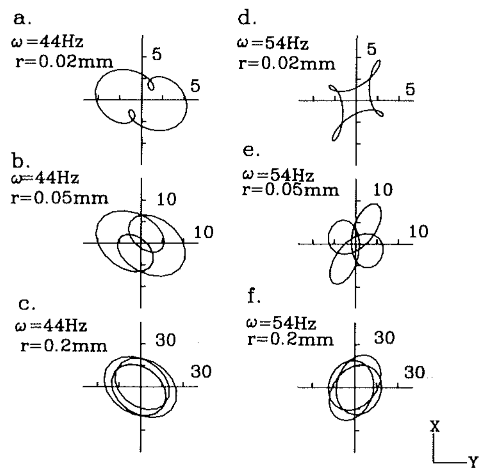


Fig. 8 Synchronous whirling orbits of unbalance responses of disk 1 in Case 2



(r is radius of journal motion with Ω= 3ω)

Fig. 9 Nonsynchronous whirling orbits of disk 1 induced by journal motion in Case 2

0, 0, 0, 0, 0, 0)^t, and subscripts 0 and 1 are labeled for stages. By deleting all elements which are related to moments and shear forces, equation (25) is expanded into

$$\{S_n'\} = [U_{11}]\ \{S_0'\} + \mathbf{u}_1 \qquad (26a)$$

and

$$[U_{21}]\ \{S_0'\} + \mathbf{u}_2 = \mathbf{0}. \qquad (26b)$$

The above system of linear equations (26*b*) can be solved to determine the state variables X_c, X_s, . . . etc. at the stage 0, denoted by $\{S_0\}$. Then, the state vectors at every point of the rotor can be obtained by multiplying transfer matrices between the specific stage and the zeroth stage. Shown in the following equation, to get the unbalance response,

$$\{S_p\} = [U_p]\ \{S_0\}, \qquad (27)$$

where $\{S_p\}$ is state vector of stage p and $[U_p]$ is the transfer matrix from stage 0 to stage p. For instance, the state vector of stage n is obtained from equation (26*a*).

In this paper, the amplitudes of responses are represented by the absolute values of the major axes of the elliptic orbits. The calculated unbalance responses at disk 1 are plotted as solid curves in Fig. 7, for the first natural mode region. The dashed curves are calculated from the finite element method by six elements. It is shown that two curve groups are close to each other. The cross stiffnesses make the critical speed split into two values for its asymmetry, as evaluated in Case 2 at 44.83 Hz and 51.52 Hz. The critical speed can be located from the maximum resonance peak of the frequency response curve. For the unsymmetrical bearing due to cross stiffness and damping, the whirling orbit becomes elliptical and there is a backward whirling phenomena between the two split critical speeds. The whirling orbits in forward and backward directions of Case 2 are shown in Fig. 8.

Lastly, we consider the motion of the journal as a nonsynchronous exciting sources which may be induced by the roughness and unround profile, axial forces or torques, nonlinear bearing performances, or external vibrations. In this example the synchronous response, due to unbalance mass and nonsynchronous response due to journal motion with frequency being three times of the rotating speed, is considered. It is assumed that the journal motion at stage n is represented by $X = r\cos 3\omega t$, and $r \sin 3\omega t$. In this case, $\Omega = 3\omega$ is substituted into all prescribed equations which process Ω. The displacements of the free ends are set to be equal to the displacements of the journals. And the bending moments of the free ends are still zero. Then substituting these boundary conditions into equation (24), we have

$$\{S_n\} = (r, 0, 0, r, \alpha_c, \alpha_s, \beta_c, \beta_s, 0, 0, 0, 0, Q_{xc}, Q_{xs}, Q_{yc}, Q_{ys})^t$$
$$\{S_0\} = (X_c, X_s, Y_c, Y_s, \alpha_c, \alpha_s, \beta_c, \beta_s, 0, 0, 0, 0, 0, 0, 0, 0)^t$$

for nonsynchronous components and

$$\{S_n\} = (0, 0, 0, 0, \alpha_c, \alpha_s, \beta_c, \beta_s, 0, 0, 0, 0, Q_{xc}, Q_{xs}, Q_{yc}, Q_{ys})^t$$
$$\{S_0\} = (X_c, X_s, Y_c, Y_s, \alpha_c, \alpha_s, \beta_c, \beta_s, 0, 0, 0, 0, 0, 0, 0, 0)^t$$

for synchronous components. By further manipulation, we can obtain the excited response of any stage induced by journal motion.

The configuration of Case 2 is considered here. From the superposition of the unbalance response and excited response the multi-lobed whirling orbits for various journal motions in different radii are obtained and shown in Fig. 9. It is obvious that the results can not be obtained from the assumption of a circular orbit in transfer matrix method. The nonsynchronous whirling is a natural phenomenon of the nonlinear rotor system which indicates that this method can be extended to the nonlinear analysis.

Conclusion

A modified transfer matrix method involving the general whirling motions is presented to analyze the dynamic properties of a flexible linear rotor-bearing system. The transfer matrix of a shaft is derived from continuous-system representation in order to decrease the number of segments of the shaft and also to obtain higher accuracy than that of a lumped-system representation.

The computing results show that the developed method agrees satisfactorily with actual performances of the rotor systems. For synchronous whirling, the unbalance response is solved following the track of an elliptical orbit. For nonsynchronous whirling, the motion of the bearing will cause the secondary orbit. Also, the method can be used to investigate the nonsynchronous whirling orbits of subharmonic and superharmonic resonances of the nonlinear rotor system.

Acknowledgment

This study was supported by the National Science Council, Republic of China, under contract number NSC80-0401-E-009-04.

References

Bansal, P. N., and Kirk, R. G., 1975, "Stability and Damped Critical Speeds of Rotor-Bearing Systems," *ASME Journal of Engineering for Industry*, Vol. 97, pp. 1325–1332.

Dym, C. L., and Shames, I. H., 1973, *Solid Mechanics, A Variational Approach*, McGraw-Hill, New York, pp. 187–194.

Eshleman, R. L., and Eubanks, R. A., 1969, "On the Critical Speeds of a Continuous Rotor," ASME *Journal of Engineering for Industry*, Vol. 91, pp. 1180–1188.

Gu, Jialiu, 1986, "An Improved Transfer Matrix-Direct Integration Method for Rotor Dynamics," ASME *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 108, pp. 183–188.

Guenther, T. G., and Lovejoy, D. C., 1961, "Analysis for Calculating Lateral Vibration Characteristics of Rotating Systems with Any Number of Flexible Supports: Part 2—Application of the Method of Analysis," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, pp. 591–600.

Koenig, E. C., 1961, "Analysis for Calculating Lateral Vibration Characteristics of Rotating Systems with Any Number of Flexible Supports: Part 1—The Method of Analysis," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, pp. 585–590.

Lund, J. W., and Orcutt, F. K.,.1967, "Calculations and Experiments on the Unbalance Response of a Flexible Rotor," ASME *Journal of Engineering for Industry*, Vol. 89, No. 4, pp. 785–596.

Lund, J. W., 1974, "Stability and Damped Critical Speeds of a Flexible Rotor in Fluid-Film Bearings," ASME *Journal of Engineering for Industry*, Vol. 96, pp. 509–516.

Lund, J. W., 1974, "Modal Response of a Flexible Rotor in Fluid-Film Bearings," ASME *Journal of Engineering for Industry*, Vol. 96, pp. 509–516.

Lund, J. W., 1974, "Modal Response of a Flexible Rotor in Fluid Film Bearings," ASME *Jounal of Engineering for Industry*, Vol. 96, pp. 525–533.

Lund, J. W., 1980, "Sensitivity of the Critical Speeds of a Rotor to Changes in Design," ASME *Journal of Mechanical Design*, Vol. 102, pp. 115–121.

Prohl, M. A., 1945, "A General Method for Calculating Critical Speeds of Flexible Rotors," ASME JOURNAL OF APPLIED MECHANICS, Vol. 67, pp. A-142–A-148.

Rao, J. S., 1983, *Rotor Dynamics*, Wiley Eastern Ltd., New Delhi.

A Geometrical Approach Assessing Instability Trends for Galloping

P. Yu

N. Popplewell

A. H. Shah

Faculty of Engineering,
The University of Manitoba,
Winnipeg, Manitoba R3T 2N2, Canada

Although the galloping of an iced electrical conductor has been considered by many researchers, no special attention has been given to the galloping's sensitivity to alternations in the system's parameters. A geometrical method is presented in this paper to describe these instability trends and to provide compromises for controlling an instability. The conventional but uncontrollable parameter of the wind speed is chosen as the basis for obtaining the critical conditions under which bifurcations occur for a representative two degrees-of-freedom model. Variations in these critical conditions are found in a two-dimensional parameter space in order to determine the trends for the initiation of galloping as well as to evaluate the stability of the ensuring periodic vibrations.

1 Introduction

The galloping of iced electrical conductors has been considered, since early in this century, by many researchers. Important results have been obtained, for example, by Den Hartog (1932), Simpson (1965), Chadha (1973), and Blevins and Iwan (1974). The main focus of these papers, however, was to find the requirements for the initiation of galloping as well as to determine the conditions for the instability of the vibrations. However, galloping is highly complex as a result particularly of the nonlinear aerodynamic forces and because many physical parameters are involved. Although several analytical expressions have been given for the criteria of instability (e.g., Blevins and Iwan, 1974), the direct application of these expressions to the design of appropriate control devices is still not possible. This serious deficiency suggests that it is necessary to investigate the instability trends with respect to the system's parameters. Moreover, exact solutions and instability conditions may not be meaningful in view of the approximations in the mathematical modeling and analysis and the errors in laboratory experiments and field trials. Instability trends, therefore, should provide more suitable and effective guidelines for a robust control strategy which makes a particular design more tolerant of uncertain parameters.

A geometrical approach will be introduced to consider the instability trends for the galloping of a two degrees-of-freedom model (Blevins and Iwan, 1974). The approach is based upon the instability conditions for the steady-state solutions associated with equilibrium, periodic, and quasi-periodic motions. Of particular interest here are the conditions for the initiation of galloping as well as the critical boundaries where a plunge or torsional motion, or even a mixed-mode motion, loses stability. A two-dimensional parameter space will be chosen to be chosen to demonstrate the instability trends. Thus, instead of expressing the instability conditions in the whole (at least eight-dimensional) parameter space, a single parameter—the wind speed—will be chosen as the critical variable. Then the critical values of this variable, which correspond to the initiation of galloping and to the onset of instability of the periodic motions, will be determined. The influence of the remaining variables will be found in the two-dimensional parameter space by considering their individual effects on the critical values of the wind speed.

A brief derivation of the equations is given in the Appendix and the results obtained are presented in the following section. The geometrical approach and two practically important instability trends are discussed in Section 3. Finally, conclusions are drawn in Section 4.

2 Iniation of Galloping and Critical Boundaries for a Two Degree-of-Freedom System

The two degree-of-freedom model shown in Fig. 1 has mass m and moment of inertia I. It represents a cross-section of an iced conductor where y is the vertical (plunge) displacement and θ is the angle of rotation or twist. The k_y and k_θ are the vertical and torsional stiffness, respectively, and the c_y and c_θ are corresponding viscous dampers (which are not shown in

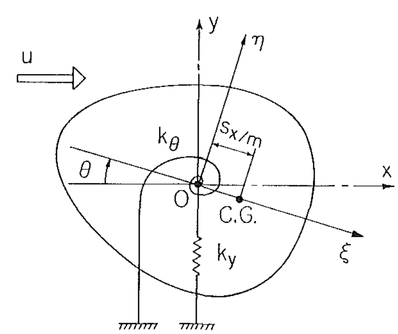


Fig.1 Elastically supported two degrees-of-freedom model

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Sept. 28, 1989; final revision, May 9, 1990.

Table 1 Steady-state solutions

	Solution	Expression
(I)	I.E.S.	$A_y = 0,\ A_\theta = 0$
(II)	H.B.S. (P)	$A_y^2 = -3B_1B_3,\ A_\theta^2 = 0$
(III)	H.B.S. (T)	$A_y^2 = 0,\ A_\theta^2 = -3B_2B_4$
(IV)	2-D Torus	$A_y^2 = B_3B_6,\ A_\theta^2 = B_4B_5$

Fig. 1). Furthermore, s_x generally indicates the eccentricity so that s_x/m is the lateral position of the iced conductor's center of gravity (C.G.) measured from its center of rotation O. However, only the cocentric case corresponding to $s_x = 0$ will be considered here. The rate equations, describing the motion of the cocentric model, expressed in terms of the amplitudes (A_y, A_θ) and phases (ϕ_y, ϕ_θ) of y and θ, respectively, are obtained by setting $e_x = 0$ ($s_x = 0$) in equations (A12) and (A13) as follows:

$$\dot{A}_y=\eta_y A_y\left\{\left(\frac{1}{2}U_y a_1-\frac{\xi_y}{\eta_y}\right)+\frac{3a_3}{8U_y}[A_y^2+2(U_y^2+r_1^2\omega^2)A_\theta^2]\right\},$$

$$\dot{A}_\theta=\eta_\theta A_\theta\left\{\left(\frac{1}{2}U_y r_1 b_1-\frac{\omega\xi_\theta}{\eta_\theta}\right)+\frac{3r_1b_3}{8U_y}[2A_y^2+(U_y^2+r_1^2\omega^2)A_\theta^2]\right\},\qquad(1)$$

and

$$\dot{\phi}_y=1,$$

$$\dot{\phi}_\theta=\omega+\frac{U_y\eta_\theta}{2\omega}\left\{U_yb_1+\frac{3b_3}{4U_y}[2A_y^2+(U_y^2+r_1^2\omega^2)A_\theta^2]\right\},\qquad(2)$$

where

$$\xi_y=\frac{c_y}{2m\omega_y},\ \xi_\theta=\frac{c_\theta}{2I\omega_\theta},\ \eta_y=\frac{\rho d^2}{2m},\ \eta_\theta=\frac{\rho d^4}{2I},\ \eta=\frac{\eta_\theta}{\eta_y},$$

and (3)

$$U_y=\frac{U}{\omega_y d},\ r_1=\frac{R_1}{d},\ \omega_y^2=\frac{k_y}{m},\ \omega_\theta^2=\frac{k_\theta}{I},\ \omega=\frac{\omega_\theta}{\omega_y}.$$

In equations (1) through (3), ρ is the density of air, U is the free-stream wind speed, d is a characteristic length which is usually taken as the maximum width of the cross-section normal to the free stream, and R_1 is the characteristic radius of the section. The aerodynamic lift force and moment are approximated by the best fit cubic polynomials having coefficients a_1, a_3, b_1, and b_3 (Blevins and Iwan, 1974).

The steady-state solutions can be obtained readily from (1) by setting $\dot{A}_y = \dot{A}_\theta = 0$. They are listed in Table 1, where the abbreviations I.E.S., H.B.S.(P), H.B.S.(T), and 2-D Torus represent the initial equilibrium solution, the Hopf bifurcation solution corresponding to a periodic plunge vibration, the Hopf bifurcation solution for a periodic torsional vibration, and a motion of a two-dimensional torus, respectively. In Table 1, the coefficients B_i ($i = 1, 2, \ldots, 6$) are given by

$$B_1=\frac{1}{2}U_ya_1-\frac{\xi_y}{\eta_y},\qquad B_4=\frac{8U_y}{9r_1b_3(U_y^2+r_1^2\omega^2)},$$

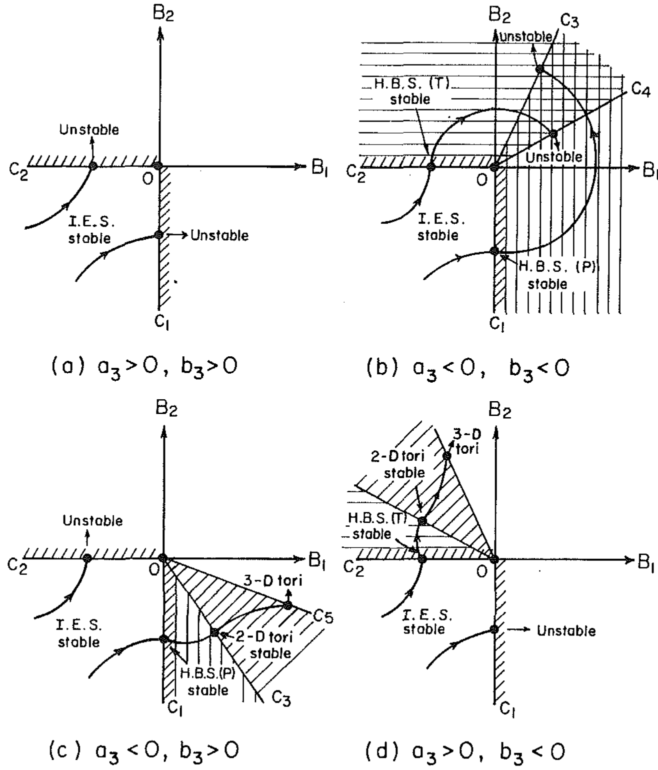


Fig. 2 Bifurcation and stability diagrams

Table 2 The stability conditions and critical boundaries for $e_x = 0$

Solution	(I)		(II)	(III)	(IV)
Existence Condition			$\frac{B_1}{a_3} < 0$	$\frac{B_2}{b_3} < 0$	$\frac{B_6}{a_3} > 0,\ \frac{B_5}{b_3} > 0$
Frequency			1	$\omega + \frac{\xi_\theta U_y}{r_1}$	$\left(1,\ \omega + \frac{\xi_\theta U_y}{r_1}\right)$
Stability Condition	$B_1 < 0,\ B_2 < 0$		$B_1 > 0,\ B_5 < 0$	$B_2 > 0,\ B_6 < 0$	$a_3b_3 < 0,\ B_5 + B_6 < 0$
Critical Boundary	$C_1: B_1 = 0$	$C_2: B_2 = 0$	$C_3: B_5 = 0$	$C_4: B_6 = 0$	$C_1: B_1 + B_6 = 0$
Bifurcation Solution	(II)	(III)	(IV)		3-D Torus or Chaos

Note: The bifurcation solutions (II), (III), (IV) and 3-D torus or chaos given in the last row, denote the solutions bifurcating from (I), (II), (III) and (IV) along the critical boundaries C_1, C_2, C_3 or C_4, and C_5, respectively.

$$B_2=\frac{1}{2}U_yr_1b_1-\frac{\omega\xi_\theta}{\eta_\theta},\quad B_5=B_5-2S_3B_1,\qquad(4)$$

$$B_3=\frac{8U_y}{9a_3},\qquad B_6=B_1-\frac{2}{S_3}B_2,$$

where

$$S_3=r_1\left(\frac{b_3}{a_3}\right).$$

Stability conditions for the steady-state solutions and the critical boundaries, where bifurcations occur, can be obtained from equation (1). The results are summarized in Table 2 and are illustrated graphically in Fig. 2.

Table 2 and Fig. 2 present an overview of the dynamical behavior of the system. Critical boundaries where bifurcations occur are given explicitly. Bifurcations and secondary bifurcations into periodic and nonresonant quasi-periodic motions are also indicated. The aerodynamic coefficients a_3 and b_3 can be seen to play a very important role in determining the stability of these bifurcations. Indeed, depending upon a_3 and b_3, four distinct cases exist. They are categorized in Fig. 2 as case (*a*) $a_3 > 0$, $b_3 > 0$, for which no stable periodic solution exists. Case (*b*) $a_3 < 0$, $b_3 < 0$ where both the plunge and torsional motions can be stable but they cannot exist simultaneously. Case (*c*) $a_3 < 0$, $b_3 > 0$ for which the plunge motion is stable but the torsional motion in unstable. (Furthermore, subsequent bifurcations from the stable plunge motion may lead motion is unstable. (Furthermore, subsequent bifurcations from the stable plunge motion may lead to a family of two-dimensional or even three-dimensional tori.) Case (*d*) $a_3 > 0$, $b_3 < 0$ where the torsional motion is stable but the plunge motion is unstable. Then the same family of the two-dimensional or three-dimensional tori may bifurcate from the stable torsional motion.

3 Instability Trends

3.1 A Geometrical Approach. Although the stability conditions of the steady-state solutions have been given explicitly in the previous section, instability trends with respect to different parameter are still unknown. A geometrical approach will be presented next to give a clear intuitive view of how these instability trends change with variations in the system's parameters. The main thrust will be to first introduce a reference line (or curve) representing a given parameter in the two-dimensional, B_1-B_2 parameter space employed, for example, in Fig. 2. Then critical values will be found, with respect to this reference, which correspond to the initiation of galloping and to the (dynamic) instability of the periodic as well as the quasi-periodic motions.

It has been reported that a change in a steady wind speed is an important factor in causing instability (Edwards and Madeyski, 1956). This observation suggests that the dimensionless wind speed, U_y, is a reasonable choice for the reference line. Indeed, it may be deduced, from the critical boundaries given in Table 2, that all the critical values can be expressed in terms of U_y because every critical boundary is described by an homogeneous equation. The equation of the reference line can be obtained from equation (4) by eliminating U_y in the following form:

$$U_y\colon B_2 = S_1 B_1 + \left(\frac{\xi_y}{\eta_y}\right) (S_1 - Q) \tag{5}$$

where

$$S_1 = r_1 \left(\frac{b_1}{a_1}\right) = r_1 b \text{ and}$$

$$Q = \left(\frac{\omega}{\eta}\right) \left(\frac{\xi_\theta}{\xi_y}\right) = \frac{\left(\frac{\omega\xi_\theta}{\eta_\theta}\right)}{\left(\frac{\xi_y}{\eta_y}\right)} = \frac{1}{d^2}\left(\frac{c_\theta}{c_y}\right). \tag{6}$$

Here, the ratio $(1/d^2)$ (c_θ/c_y) is defined as a new parameter Q. In practice the natural frequency ratio ω $(=\omega_\theta/\omega_y)$ is often manipulated to alleviate galloping by using dampers (Havard, 1988; Sasaki et al., 1986). Now, Q is related directly to the natural frequency ratio because c_θ/c_y = (I/m) ω. Therefore, it is believed that Q plays a significant role. Moreover, changing the ratio ω (e.g., by using dampers) may also simultaneously vary the inertia and mass of the system. It is more practically useful, therefore, to consider Q instead of the simples ratio, ω, as a control parameter. Also, it may be noted from equations (5) and (6) that the slope of the line U_y involves a factor $b(=b_1/a_1)$. The b is an uncontrollable parameter because it is very sensitive to the uncertain geometric shape produced by the weather-dependent ice accumulation. However, b is very important in determining the stability of the conductor's initial configuration because the stability conditions of the initial equilibrium state depend upon both a_1 and b_1. The effects of the two parameters Q and b, therefore, will be considered separately later.

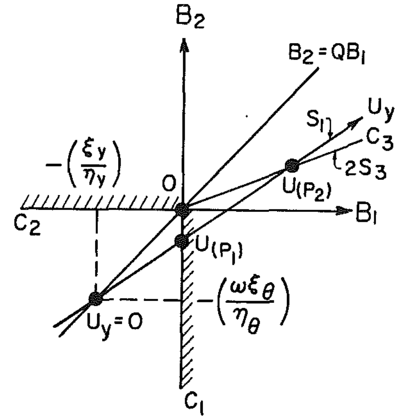
It has been observed that a_1 and b_1 are both positive, whereas a_3 and b_3 are both negative in most practical situations (Novak, 1971). Thus, special attention will be given in the following analysis to the case: $a_1 > 0$, $b_1 > 0$, $a_3 < 0$, and $b_3 < 0$. These stipulations correspond to case (*b*) in Fig. 2. However, it is not difficult to extend the approach to the other cases presented in Fig.2.

The following two distinct cases can be found by comparing the slope of U_y, i.e., S_1, to the control parameter Q:

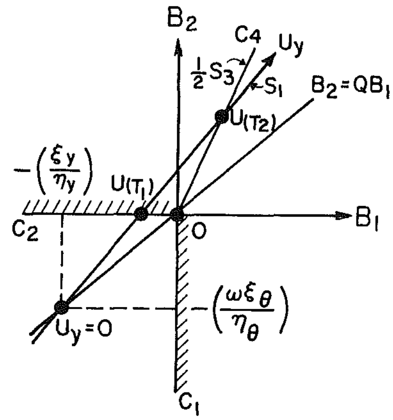
Case (*P*): $Q > S_1$, corresponds to the initiation of plunge vibration, and

Case (*T*): $Q < S_1$, corresponds to the initiation of torsional vibration.

These two distinct cases are illustrated in Fig. 3. The hatched lines in this figure indicate the stability boundaries of the initial equilibrium solution. The $U_{(P_1)}$ and $U_{(T_1)}$ are the initiation values for plunge and torsional vibrations, respectively. The $U_{(P_2)}$ and $U_{(T_2)}$, on the other hand, respectively represent the practically important critical values where the stability of the



(a) Case (P), Q > S1



(b) Case (T), Q < S1

Fig. 3 The critical boundaries

Table 3 Plunge and torsional cases

Case	(P) $Q > S_1$		(T) $Q < S_1$	
Initiation of Galloping	Plunge Vibrations		Torsional Vibrations	
Stability Condition for Periodic Vibrations	$S_1 \le 2S_3$	Stable $(U_{(P_2)} \to \infty)$	$S_1 \ge \frac{1}{2} S_3$	Stable $(U_{(T_2)} \to \infty)$
	$S_1 > 2S_3$	$U < U_{(P_2)}$ Stable	$S_1 < \frac{1}{2} S_3$	$U < U_{(T_2)}$ Stable
		$U \ge U_{(P_2)}$ Unstable		$U \ge U_{(T_2)}$ Unstable

plunge motion and the torsional motion are lost. All the initiation and critical values can be obtained from the critical boundaries listed in Table 2 and equation (4). They are given as follows:

$$U_{(P_1)} = \frac{2\xi_y}{\eta_y a_1},$$

$$U_{(T_1)} = \frac{2\omega\xi_\theta}{\eta_\theta r_1 b_1},$$

$$U_{(P_2)} = \left(\frac{2\xi_y}{\eta_y a_1}\right) \cdot \frac{(Q-2S_3)}{(S_1-2S_3)}, \tag{7}$$

$$U_{(T_2)} = \left(\frac{2\xi_y}{\eta_y a_1}\right) \cdot \frac{(S_2-2Q)}{(S_3-2S_1)} = \left(\frac{2\omega\xi_\theta}{\eta_\theta r_1 b_1}\right) \cdot \frac{\left(\frac{1}{Q}-\frac{2}{S_3}\right)}{\left(\frac{1}{S_1}-\frac{2}{S_3}\right)}.$$

It can be seen from Fig.3 (*a*) that if the slope of C_3, the critical boundary for the plunge motion defined in Table 2 is larger than that of U_y (i.e., $2S_3 > S_1$), there is no intersection of C_3 and U_y in the first quadrant. This implies $U_{(P_2)} \to \infty$ so that the actual wind speed, U, can never exceed $U_{(P_2)}$, and the plunge motion is always beneficially stable.

A similar conclusion can be drawn from Fig. 3(*b*) for the torsional vibration case, H.B.S.(T). The two sets of results are summarized, for convenience, in Table 3.

Table 3 suggests that both the plunge and the torsional motion are stable when $1/2\, S_3 \le S_1 \le 2S_3$, irrespective of the wind speed (i.e., $U_{(P_2)} \to \infty$ and $U_{(T_2)} \to \infty$). This region, therefore, is of no practical interest. Consequently, it is assumed from the following analysis that

$$S_1 > 2S_3 \text{ for H.B.S.(P), and}$$
$$S_1 < \frac{1}{2} S_3 \text{ for H.B.S.(T).} \tag{8}$$

3.2 General Instability Trends. Consider the effects of variations in the system's parameters on the critical values $U_{(P_1)}$, $U_{(T_1)}$, and $U_{(T_2)}$. As the parameters of their combinations change, lines like U_y, C_3, C_4, etc., will move in the $B_1 - B_2$ plane and new values will be obtained for $U_{(P_1)}$, $U_{(T_1)}$, $U_{(P_2)}$ and $U_{(T_2)}$. For illustration, examine the simple example of $c_y \nearrow$, which results in $(\xi_y/\eta_y)\ (= c_y/\omega_y \rho d^2) \nearrow$, where the arrow $\nearrow$ ($\searrow$) indicates an increase (a decrease) in the given parameter. This example is illustrated in Fig. 4 where a superscript prime denotes a new value of a parameter whereas the corresponding unprimed quantity indicates the old values. An "x" in this figure represents the point to which the old $U_{(P_1)}$ or old $U_{(P_2)}$ will move when c_y increases. The movements are indicated by the dashed lines which are parallel to the line passing through the points marked $U_y = 0$ and $U_y' = 0$. These latter points correspond to the old and new values of c_y, respectively.

Table 4 Trends of the key region and points for case (P)

Parameter ($\nearrow$)	S.R. of I.E.S.	$U_{(P_1)}$	$U_{(P_2)}$
c_y	$\nearrow$	$\nearrow$	$\searrow$
c_θ	$\nearrow$	Unchanged	$\nearrow$
a_1	Unchanged	$\searrow$	$\nearrow$
b_1	Unchanged	Unchanged	$\searrow$
$\frac{b_3}{a_3}$	Unchanged	Unchanged	$\nearrow$

Table 5 Trends of the key region and points for case (T)

Parameter ($\nearrow$)	S.R. of I.E.S.	$U_{(T_1)}$	$U_{(T_2)}$
c_y	$\nearrow$	Unchanged	$\nearrow$
c_θ	$\nearrow$	$\nearrow$	$\searrow$
a_1	Unchanged	Unchanged	$\searrow$
b_1	Unchanged	$\searrow$	$\nearrow$
$\frac{b_3}{a_3}$	Unchanged	Unchanged	$\searrow$

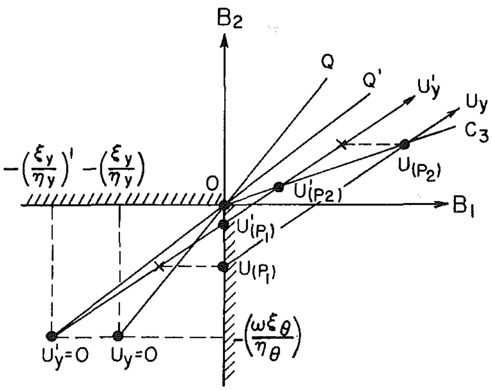


Fig. 4 Instability trends for increasing c_y and $Q > S_1$

The following results can be observed from Fig. 4. When c_y increases, the point marked $U_y = 0$ will move to the left to a point marked $U_y' = 0$. Hence, line U_y', which is parallel to line U_y, will be displaced to the left, too. Thus, the new value $U'_{(P_1)}$ obtained from the moving up of $U_{(P_1)}$ parallel to axis B_2, is greater than the value of the corresponding point marked "x," which moves from $U_{(P_1)}$ parallel to axis B_1. Therefore an increase in c_y causes $U_{(P_1)}$ to grow. Similarly, it can be concluded that an increase in c_y makes $U_{(P_2)}$ diminish. Moreover, it is seen that the stable region for the initial equilibrium solution is a rectangle in the third quadrant which is bounded by the B_1, B_2 axes, and the dashed lines passing through $U_y = 0$ or $U_y' = 0$. So, a greater c_y will enlarge the stable region of the I.E.S. In summary,

$$c_y \nearrow \Rightarrow U_{(P_1)} \nearrow,\ U_{(P_2)} \searrow \text{ and S.R. of I.E.S.} \nearrow, \tag{9}$$

where S.R. denotes the stable region. The same conclusions may also be deduced straightforwardly from equation (7) for this simple example.

Although conclusion (9) was obtained on the basis of a particular point $U_y = 0$ and a small increase in c_y, it can be

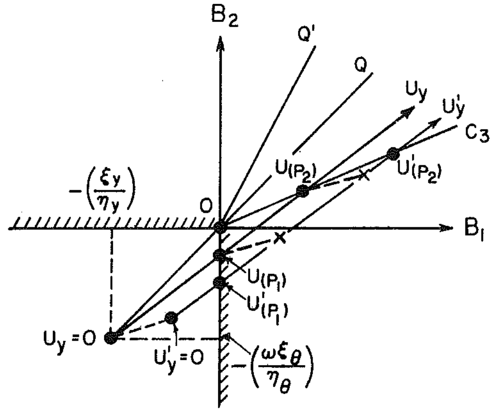


Fig. 5 Instability trends for increasing Q in case (P)

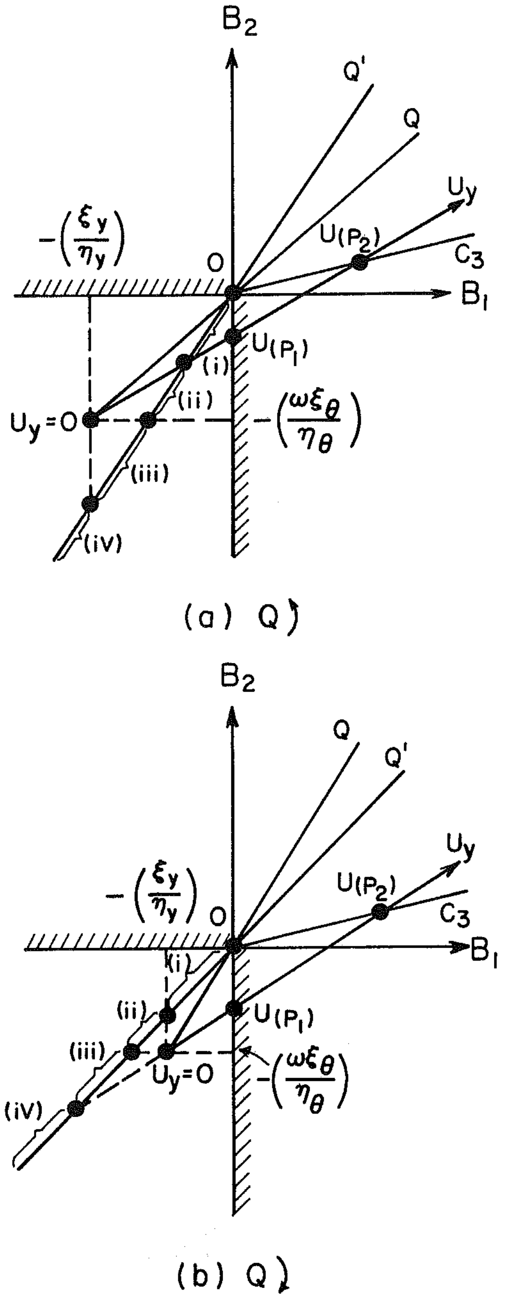


Fig. 6 Different characteristic regions which depend upon the origin of U_y' in case (P)

Table 6 Trends of the I.E.S. stability region (S.R.) and the key points w.r.t. Q in case (P)

Case	Region	Definition	S.R. of I.E.S.	$U_{(P_1)}$	$U_{(P_2)}$
Q ↗	(i)	$c_y' < c_y^*$	↘	↘	↘
	(ii)	$c_y^* < c_y' < \frac{c_\theta}{Q'}$	↘	↘	↗
	(iii)	$\frac{c_\theta}{Q'} < c_y' < c_y$	Indefinite	↘	↗
	(iv)	$c_y < c_y'$	↗	↗	↗
Q ↘	(i)	$c_y' < c_y$	↘	↘	↘
	(ii)	$c_y < c_y' < \frac{c_\theta}{Q'}$	Indefinite	↗	↘
	(iii)	$\frac{c_\theta}{Q'} < c_y' < c_y^*$	↗	↗	↘
	(iv)	$c_y^* < c_y'$	↗	↗	↗

observed from Fig. 4 that this conclusion is true for any arbitrary location of the point $U_y = 0$. Thus, conclusion (9) is true globally so that is robust.

Following the previous procedure, instability trends were derived for separate changes in several simple parameters. These trends are summarized in Tables 4 and 5. Except for the stability region (S.R.) of the I.E.S., the results can be obtained, alternatively, from equation (7). (Later, however, two important cases will be studied in which it may not be feasible to use equation (7).) Tables 4 and 5 suggest that the stable I.E.S. region and the the initiation points $U_{(P_1)}$ and $U_{(T_1)}$ generally grow beneficially with individual increases in c_y or c_θ. On the other hand, such increases, when simultaneous, have disadvantageously counteracting influences on the dynamic stability points $U_{(P_2)}$ and $U_{(T_2)}$. Consequently, the obvious control strategy of simple enlarging both c_y and c_θ does make the initiation of galloping harder, but the wind speed for the onset of a dynamic instability may or may not be affected. Increasing the aerodynamic coefficients a_1, b_1 or the ratio b_3/a_3 produces similar opposing trends for $U_{(P_2)}$ and $U_{(T_2)}$, but the wind speed at the initiation of galloping is either unchanged or reduced disadvantageously. Therefore, the problem of controlling galloping may not be straightforward.

3.3 Two Important Practical Cases. Specific instability trends with respect to (w.r.t.) first Q (rather than c_y or c_θ) and then b will be considered next. It is more appropriate, from a practical viewpoint, to consider the effects of c_y and c_θ simultaneously because, as seen in Section 3.1, Q is the key parameter which distinguishes the torsional from the plunge vibration. Further, Q plays a significant role in controlling the stability trends of galloping.

3.3.1 Instability Trends w.r.t. Parameter Q. First, consider case (P), in which $Q > S_1$, and, further, suppose that Q is increasing. Draw the line U_y and lines having the old slope, Q, and the new slope, Q', in the $B_1 - B_2$ plane as indicated in Fig. 5. To draw the line corresponding to a new $U_y' = 0$, the origin $U_y' = 0$ has to be chosen in the third quadrant on the line having slope $U_y' = 0$, the origin $U_y' = 0$ has to be chosen in the third quadrant on the line having slope Q'. This origin reflects the absolute values of c_y and c_θ and its location will

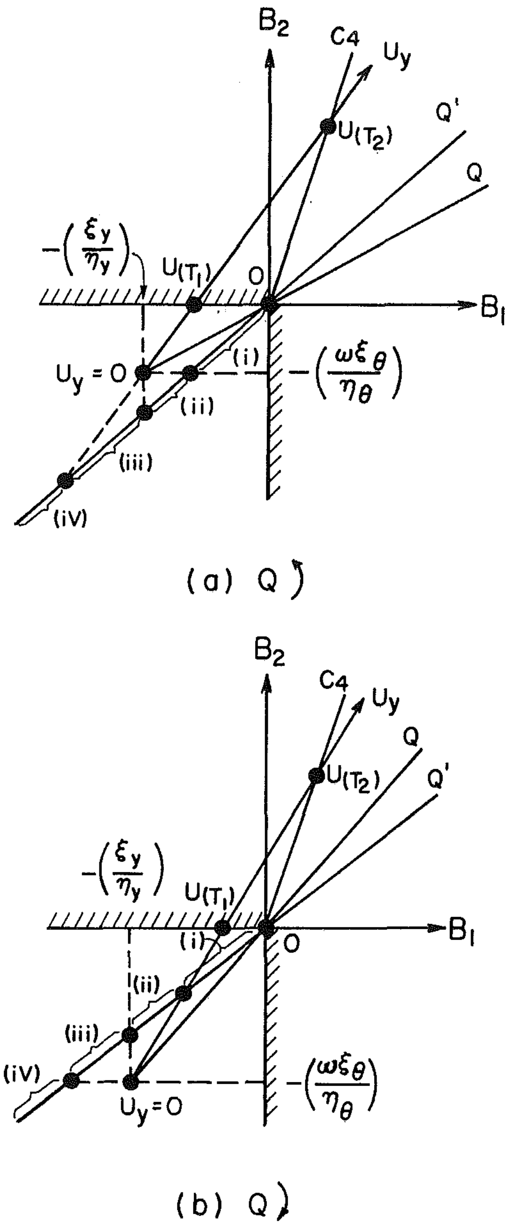


Fig. 7 Different characteristic regions which depend upon the origin of U_y' in case (T)

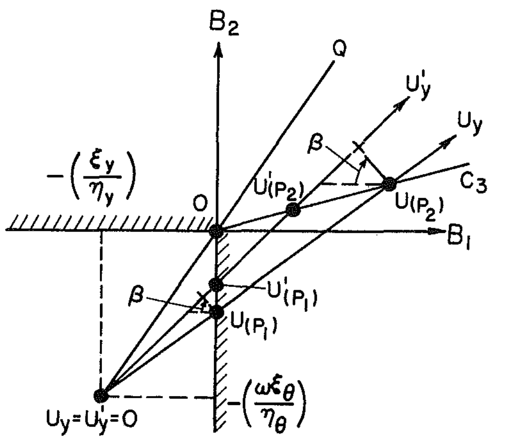


Fig. 8 Instability trends for increasing b in case (P)

Table 7 Trends of the I.E.S. stability region (S.R) and the key points w.r.t. Q in case (T)

Case	Region	Definition	S.R. of I.E.S.	$U_{(T_1)}$	$U_{(T_2)}$
Q ↗	(i)	$c_\theta' < c_\theta$	↘	↘	↘
	(ii)	$c_\theta < c_\theta' < Q'c_y$	Indefinite	↗	↘
	(iii)	$Q'c_y < c_\theta' < Q'c_y^*$	↗	↗	↘
	(iv)	$Q'c_y^* < c_\theta'$	↗	↗	↗
Q ↘	(i)	$c_\theta' < Q'c_y^*$	↘	↘	↘
	(ii)	$Q'c_y^* < c_\theta' < Q'c_y$	↘	↘	↗
	(iii)	$Q'c_y < c_\theta' < c_\theta$	Indefinite	↘	↗
	(iv)	$c_\theta < c_\theta'$	↗	↗	↗

Table 8 Trends of the critical boundaries w.r.t. b

Case	Critical Boundary	Condition	b (↗)
$Q > S_1$	$U_{(P_1)}$	$\beta < \frac{\pi}{2}$	↗
		$\beta > \frac{\pi}{2}$	↘
	$U_{(P_2)}$	$\beta < \beta_3$	↗
		$\beta > \beta_3$	↘
$Q < S_1$	$U_{(T_1)}$	$\beta < 0$	↗
		$\beta > 0$	↘
	$U_{(T_2)}$	$\beta < \beta_4$	↗
		$\beta > \beta_4$	↘

Note: $\beta_3 = -\tan^{-1}(2S_3)$ and $\beta_4 = \pi - \cot^{-1}(2S_3)$ where S_3 is defined in equation (6).

affect the results. Once chosen, the line U_y' can be drawn parallel to the line U_y and the new values of $U'_{(P_1)}$ and $U'_{(P_1)}$ can be obtained. Next, draw dashed lines through the points $U_{(P_1)}$ and $U_{(P_2)}$ which are parallel to the line from the point $U_y = 0$ to the point $U_y' = 0$. These dashed lines intersect the line U_y' at the point marked by "x" in Fig. 5. Thus, an increase or a decrease in $U_{(P_1)}$ and $U_{(P_2)}$ can be determined, as a result of Q increasing, by comparing $U'_{(P_1)}$ and $U'_{(P_2)}$ with the corresponding points marked by "x." For the case depicted in Fig. 5, $U_{(P)1}$ ↘ and $U_{(P)2}$ ↗ when Q ↗. However, it can be shown that the choice of the origin $U_y' = 0$ affects this conclusion. Four different regions can be distinguished depending upon the location of $U'_y = 0$. They are indicated by (i), (ii), (iii), and (iv) in Fig. 6(a). A similar procedure can be followed for the case $Q\searrow$ and the results are shown in Fig. 6 (b). The associated instability trends for $Q\nearrow$ and $Q\searrow$ are listed in Table 6, where

$$c_y^* = \frac{(S_1 - Q)}{(S_1 - Q')} c_y. \tag{10}$$

It should be noted that, although the notation for the different regions given in Fig. 6 and Table 6 are identical, the definitions of these regions are very different. Also, it is interesting to note that, if compared in the reverse order of the rows in Table 6, the trends for $Q\searrow$ are opposite to those for $Q\nearrow$ because of antisymmetry.

The procedure described above can be applied similarly to case (T) where $Q < S_1$. The results are shown in Fig. 7 and they are summarized in Table 7. Due to the symmetry between the equation describing the torsional motion and that giving the plunge motion (see equation (1)), the trends of the critical arrowed values listed in Table 7 for $Q\nearrow$ ($Q\searrow$, respectively) are identical to those given in Table 6 for $Q\searrow$ ($Q\nearrow$, respectively).

The need for four separate regions in Tables 6 and 7 suggests that the effects of changing Q are not described simply. However, it appears from these tables that a larger or a smaller Q generally produces conflicting trends in the initiation and the dynamic stability of either plunge or torsional galloping. Furthermore, these trends are usually opposite for the plunge and corresponding torsional situation. Therefore, a control strategy for galloping may have to be a careful compromise which depends upon individual circumstances.

3.3.2 Instability Trends w.r.t. Parameter b. Parameter Q does not change for this case but the slope of the line U_y does vary. First, consider case (P), $Q > S_1$, which corresponds to the initiation of a plunge vibration. Suppose $b = (= b_1/a_1)$ is increasing as exemplified in Fig. 8.

In order to compare $U'_{(P_1)}$ with $U_{(P_1)}$ and $U_{(P_2)}$ with $U_{(P_2)}$, the positions of the points $U_{(P_1)}$ and $U_{(P_2)}$ have to be found on line U'_y. First, define a new parameter (angle) β given by

$$\tan \beta = \frac{\Delta b_1}{\Delta a_1} = \frac{b'_1 - b_1}{a'_1 - a_1}, \tag{11}$$

where the range of β is found to be

$$-\tan^{-1}(S_1) = \beta_1 \le \beta \le \beta_2 = \pi - \beta_1. \tag{12}$$

Then it is straightforward, in terms of β, to draw similar diagrams to Fig. 8 for the cases presented in Table 8. For example, consider the changing trends of $U_{(P_1)}$. It is easy to observe from Fig. 8 that $\beta = \pi/2$ is a critical value for which the new point $U'_{(P_1)}$ will be located at the place marked by "x." So $U_{(P_1)}$ does not change when b varies such that $\beta = \pi/2$. Here, it should be noted that, for convenience, the direction of measuring angle β in Fig. 8 is always clockwise irrespective of whether parameter b increases or decreases.

The results shown in Table 8 suggest that the possibilities of the initiation as well as the dynamic instability of galloping are less if β is mall, or even negative, when b grows. Of course, the reverse is true when b decreases.

4 Conclusions

A geometrical approach has been introduced to find the instability trends of galloping of an iced, transmission lines when its parameters change. Based on the stability boundary derived from a two degrees-of-freedom model, critical wind speeds are obtained, with respect to (at least eight) system parameters, but in a simple two-dimensional space. Besides considering the general instability trends, important practical cases are studied in detail. Tabulations are presented of the increases or decreases in the stability trends, with respect to various parameter ranges, for the initiation of galloping and periodic vibrations. Not only does the geometrical approach provide robust solutions, but it can also be generalized straightforwardly to accommodate eccentricities, and more than the quite representative, but still only two degrees-of-freedom model employed here. Changing the parameters of an iced conductor has been shown to often lead to contradictory instability trends. Therefore it is not surprising that the most practical way to best control galloping is still debated.

Acknowledgments

The authors gratefully acknowledge the financial support of the Canadian Electrical Association (CEA Project No. 321 T 672) through the auspices of Manitoba Hydro. Partial support was also received from the Natural Sciences and Engineering Research Council of Canada.

References

Blevins, R.D., 1974, "Flow Induced Vibration," *Ph.D. Thesis*, California Institute of Technology, Pasadena, Calif.

Blevins, R.D., and Iwan, W. D., 1974, "The Galloping Response of a Two-Degree-of-Freedom System," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, pp. 1113–1118.

Chadha, J., 1973, "A Dynamic Model Investigation of Conductor Galloping," *Ontario Hydro Research Division Report*, No. 73-399-K.

Den Hartog, J. P., 1932, "Transmission Line Vibration Due to Sleet," *Trans. AIEE*, Vol. 51, Part 4, pp. 1074–1086.

Edwards, A. T., and Madeyski, A., 1956, "Progress Report on the Investigation of Galloping of Transmission Line Conductors," *Trans. AIEE*, Vol. 75, pp. 666–686.

Havard, D. G., 1988, "Conductor Galloping and Its Control," (an article for inclusion in) *Power Technology International Sterling Publications Inc.*, London.

Huseyin, K., and Yu, P., 1988, "On Bifurcations into Non-Resonance Quasi-Periodic Motions," *Appl. Math. Modelling*, Vol. 12, pp. 189–201.

Novak, M., 1971, "Galloping and Vortex Induced Oscillations of Structures," *Proc. 3rd Int. Conf. on Wind Effects on Buildings and Structures*, Tokyo, Japan, Sept. 6–11, Paper IV-16, p. 11.

Sasaki, S., Oishi, M., Komoda, M., Kojima, Y., Maeda, Y., Akiyama, T., and Okumura, T., 1986, "Development of Galloping Control Devices and Its Operation Records in Japan," *Int. Conf. Large High Voltage Electric Systems, CIGRE: Proc 31st. Session 22.07/1-6*, Vol. 1, France.

Simpson, A., 1965, "Aerodynamic Instability of Long-Span Transmission Lines," *Proc. IEE*, Vol. 112, No. 2, pp. 315–324.

APPENDIX

A brief outline of the derivation of equations (1) and (2) is given when $s_x/m \neq o$. The equation describing the motion of a two degree-of-freedom model are (Blevins, 1974):

$$\begin{aligned} m\ddot{y} + c_y\dot{y} - s_x\ddot{\theta} + k_y y &= F_y, \\ I\ddot{\theta} + c_\theta\dot{\theta} - s_x\ddot{y} + k_\theta\theta &= F_M, \end{aligned} \tag{A1}$$

where m, I, and s_x are defined with reference to the center of rotation, O, in Fig. 1 as

$$m = \int_A \mu d\xi\, d\eta,\ I = \int_A (\xi^2 + \eta^2)\mu\, d\xi\, d\eta,\ s_x = \int_A \xi\mu\, d\xi\, d\eta \tag{A2}$$

and μ is the mass density over cross-section A. The F_y and F_M represent the vertical aerodynamic force and the aerodynamic moment, respectively. They are functions of the angle of attack α, and can be expressed by

$$\begin{aligned} F_y &= \frac{1}{2}\rho\, U^2 d\, C_y(\alpha), \\ F_M &= \frac{1}{2}\rho\, U^2 d^2\, C_M(\alpha). \end{aligned} \tag{A3}$$

The α can be approximated by

$$\alpha \cong \theta - \frac{R_1}{U}\dot{\theta} - \frac{1}{U}\dot{y}. \tag{A4}$$

Coefficients C_y and C_M are relatively smooth continuous functions of α and they may be expressed as experimentally determined polynomials in α. The cubic polynomial approximation suggested by Blevins and Iwan (1974) for symmetric ice shapes is used here, i.e.,

$$C_y = a_1\ \alpha + a_3\ \alpha^3,$$

$$C_M = b_1\ \alpha + b_3\ \alpha^3.$$

A combination of equations (A1), (A3), and (A5) produces the following set of dimensionless equations:

$$\begin{aligned} \ddot{Y} - e_x\ \ddot{\theta} + Y &= -2\ \xi_y\ \dot{Y} - \eta_y\ U_y^2\ (a_1\ \alpha + a_3\ \alpha^3) \\ \ddot{\theta} - \eta\ e_x\ \ddot{Y} + \omega^2\theta &= -2\ \xi_\theta\ \omega\ \dot{\theta} - \eta_\theta\ U_y^2\ (b_1\ \alpha + b_3\ \alpha^3) \end{aligned} \tag{A6}$$

where $Y = y/d$, $e_x = s_x/md$, and the definitions of the other coefficients are given by equation (3). Here, dots indicate derivatives with respect to time, $\tau = \omega_y t$. The characteristic frequencies for the homogeneous system (A6) may be defined as

$$\omega_{(1,2)c}^2 = \frac{1}{2}\ T\left[1 + \omega^2 \pm \sqrt{(1+\omega^2)^2 - \frac{4}{T}\omega^2}\right], \tag{A7}$$

where $T = 1/(1-\eta e_x^2)$. Next, the introduction of the linear transformation

$$\begin{bmatrix} Y \\ \dot{Y} \\ \theta \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} 1 & 0 & K_2 & 0 \\ 0 & \omega_{1c} & 0 & K_2\omega_{2c} \\ K_1 & 0 & 1 & 0 \\ 0 & K_1\omega_{1c} & 0 & \omega_{2c} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{A8}$$

into (A6) results in the state variable equations

$$\{\dot{x}\} = [L]\{x\} + [N] \tag{A9}$$

whose Jacobian matrix, evaluated on the initial equilibrium solution $\{x\} = 0$, is now in the standard form

$$J = \begin{bmatrix} 0 & \omega_{1c} & 0 & 0 \\ -\omega_{1c} & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_{2c} \\ 0 & 0 & -\omega_{2c} & 0 \end{bmatrix}. \tag{A10}$$

The L and N in equation (A9) represent the linear and nonlinear parts of $\{\dot{x}\}$, respectively, and K_1 and K_2 in equation (8) are given by

$$K_1 = \frac{\omega_{1c}^2}{\omega_{1c}^2 - \omega^2}\ \eta\ e_x\ K_2 = \frac{\omega_{2c}^2}{\omega_{2c}^2 - \omega^2}\ e_x. \tag{A11}$$

Finally, by applying classical methods (e.g., averaging, multiple scale, normal form theory), rate equations governing the local dynamic behavior, expressed in terms of the amplitudes (A_y, A_θ) and phases (ϕ_y, ϕ_θ) of the perodic vibrations having frequencies ω_1, and ω_2 respectively, are obtained as

$$\begin{aligned} \dot{A}_y &= E_1 U_y\ \omega_{1c}^2\ \eta_y A_y\ [A_1 + E_2(A_3A_y^2 + 2A_4A_\theta^2)], \\ \dot{A}_\theta &= E_1 U_y \left(\frac{\omega_{2c}}{\omega}\right)^2 \eta_\theta A_\theta\ [A_2 + E_3\ (2A_3A_y^2 + A_4A_\theta^2)], \\ \dot{\phi}_y &= \omega_{1c}\left\{1 + K_1E_1U_y^2\eta_y\left[\frac{1}{2}\ U_y\ (a_1 + K_1b_1) + (a_3 + K_1b_3)\left(A_3A_y^2 + 2A_4A_\theta^2\right)\right]\right\} \\ \dot{\phi}_\theta &= \omega_{2c}\left\{1 + E_1\left(\frac{U_y}{\omega}\right)^2 \eta_\theta\left[\frac{1}{2}\ U_y\ (K_2a_1 + b_1) + (K_2a_3 + b_3)\left(2A_3A_y^2 + A_4A_\theta^2\right)\right]\right\}. \end{aligned} \tag{A12}$$

Here,

$$\begin{aligned} A_1 &= \frac{1}{2}\ U_y\ (1 + K_1r_1)\ (a_1 + K_1b_1) - \left[\frac{\xi_y}{\eta_y} + K_1^2\ \frac{\omega\xi_\theta}{\eta_\theta}\right] \\ A_2 &= \frac{1}{2}\ U_y\ (K_2 + r_1)\ (K_2a_1 + b_1) - \left[K_2^2\ \frac{\xi_y}{\eta_y} + \frac{\omega\xi_\theta}{\eta_\theta}\right] \\ A_3 &= \frac{3}{8U_y}\ [K_1^2U_y^2 + (1 + K_1r_1)^2\ \omega_{1c}^2]\ (>0), \\ A_4 &= \frac{3}{8U_y}\ [U_y^2 + (K_2 + r_1)^2\ \omega_{2c}^2]\ (>0), \\ E_1 &= \frac{1}{(1 - K_1K_2)\ U_y}, \\ E_2 &= (1 + K_1r_1)\ (a_3 + K_1b_3), \\ E_3 &= (K_2 + r_1)\ (K_2a_3 + b_3). \end{aligned} \tag{A13}$$

Measurements of Velocity, Velocity Fluctuation, Density, and Stresses in Chute Flows of Granular Materials

Hojin Ahn
Assoc. Mem. ASME

Christopher E. Brennen
Mem. ASME

Rolf H. Sabersky
Mem. ASME

California Institute of Technology,
Pasadena, CA 91125

Experiments on continuous, steady flows of granular materials down an inclined channel or chute have been conducted with the objectives of understanding the characteristics of chute flows and of acquiring information on the rheological behavior of granular material flow. Two neighboring fiber-optic displacement probes provide a means to measure (1) the mean velocity by cross-correlating two signals from the probes, (2) the unsteady or random component of the particle velocity in the longitudinal direction by a procedure of identifying particles, and (3) the mean particle spacing at the boundaries by counting the frequency of passage of the particles. In addition, a strain-gauged plate built into the chute base has been employed to make direct measurement of shear stress at the base. With the help of these instruments, the vertical profiles of mean velocity, velocity fluctuation, and linear concentration were obtained at the sidewalls. Measurements of some basic flow properties such as solid fraction, velocity, shear rate, and velocity fluctuation were analyzed to understand the characteristics of the chute flow. Finally, the rheological behavior of granular materials was studied with the experimental data. In particular, the rheological models of Lun et al. (1984) for general flow and fully developed flow were compared with the present data.

1 Introduction

Recent theoretical research has added greatly to our knowledge of the rheological behavior of rapidly flowing granular materials. For example, Ogawa et al. (1980), Savage and Jeffrey (1981), Jenkins and Savage (1983), and Lun et al. (1984) have led to a comprehension of how stresses and solid fraction in a granular flow are related to velocity gradient and to the kinetic energy associated with random motions of particles (the so-called granular temperature). Moreover, for simple shear flow, all the theoretical analyses predict a rheological behavior which is a natural extension of that originally proposed by Bagnold (1954). Namely,

$$\tau_{ij}=\rho_p f_{ij}(\nu)d^2\left(\frac{du}{dy}\right)^2$$

where τ_{ij} is the stress tensor, ρ_p is the density of the solid particle, f_{ij} is a tensor function of solid fraction, ν, d is the diameter of the particle, and du/dy is the local mean shear rate. Lun et al. (1984) estimated the ratio of the characteristic mean shear velocity, $d(du/dy)$, to the root mean square of velocity fluctuations to be a function only of solid fraction and the coefficient of restitution. In their analysis the velocity fluctuations were assumed to be isotropic. Furthermore, the effect of a variable coefficient of restitution which depends on the particle impact velocity, has been studied by Lun and Savage (1986). It has been found that the coefficient of restitution, which increases with decreasing impact velocity, causes the stresses to vary with the shear rate to a power less than two.

These advances have been greatly aided by computer simulations (for example, Campbell and Brennen (1985a,b), Walton and Braun (1986a,b), Campbell and Gong (1986), and Campbell (1989)). Especially Walton and Braun (1986b), Campbell and Gong (1986), and Campbell (1989) produced results similar to those of the theoretical models. However, velocity fluctuations are found to be anisotropic. That is, as solid fraction decreases, granular temperature deviates from an isotropic distribution. The effect of a variable coefficient of restitution has also been examined in the computer simulation by Walton and Braun (1986b). The results manifested a deviation from those of the constant coefficient of restitution in a manner similar to that of Lun and Savage (1986), but the calculated stresses were significantly lower than those of ex-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 17, 1989; final revision, Sept. 12, 1990.

perimental studies and Lun and Savage. Though there has been significant success, computer simulation faces difficulties in creating realistic boundary conditions. In addition, complicated interactions between particles and between solid walls and particles remain to be explored.

On the other hand, progress in experimental methods for granular materials has been very limited, being hindered by the obvious difficulties involved in making point measurements of velocity, solid fraction, or granular temperature in the interior of granular material flow. For example, granular temperature, in spite of its importance, had not been experimentally measured until Ahn et al. (1988) used fiber-optic displacement probes to measure one component of velocity fluctuations. The present state of the experimental information on granular material flows consists of a number of Couette flow studies (e.g., Savage and McKeown (1983), Hanes and Inman (1985), and Craig et al. (1986)) and several studies of flows down inclined chutes (e.g., Bailard (1978), Augenstein and Hogg (1978), Patton et al. (1987), and Ahn et al. (1988)). Understandably, the initial objective of some of the Couette flow experiments (such as those of Savage and McKeown (1983)) was to produce a simple shear flow with uniform velocity gradient, uniform solid fraction, and hopefully, uniform granular temperature. To this end the surfaces of the solid walls were roughened to create a no-slip velocity condition at the wall. Practical engineering circumstances require the knowledge of how to model the conditions for smooth walls at which slip occurs. This presents some difficulties because the boundary conditions on the velocity and granular temperature at the smooth walls are far from clear (see, for example, Campbell (1988)).

Chute flows differ from Couette flows and have a "conduction" of granular temperature as indicated in Campbell and Brennen (1985b). In their work, a boundary layer next to the wall had a lower solid fraction and higher granular temperature than the bulk further from the wall, indicating a conduction from the boundary layer to the bulk. Granular conduction has more extensively been studied by Ahn et al. (1989). The results show that granular temperature can be conducted either from the wall boundary to the free surface, or from the free surface to the wall, depending on the values of the coefficient of restitution and the angle of chute inclination. Furthermore, the granular conduction term and the dissipation term are found to be comparable in magnitude. The results also show a significant role played by the granular conduction in determining the profiles of granular temperature, solid fraction, and velocity.

This paper contains a study of continuous, steady flows of granular materials down an inclined chute. The objective was to understand the characteristics of granular chute flows and to acquire information on the rheological behavior of granular flow.

2 Review of Rheological Models

In this section, the existing rheological models postulated by Lun et al. (1984) will be reviewed for the purpose of analyzing the present experimental data. Comparisons between simple shear flow and fully developed chute flow will also be included.

For two-dimensional flow which is steady and fully developed in the flow direction, the translational fluctuation energy equation is given as follows (see, for example, Jenkins and Savage (1983)):

$$-P_{yx}\frac{du}{dy}-\frac{\partial q_y}{\partial y}-\gamma=0, \tag{1}$$

where P_{yx} is the shear stress, u is the velocity in the flow direction, y is a coordinate in the direction normal to the flow, and q_y is the y-component of fluctuation energy flux. The rate of the dissipation of fluctuation energy per unit volume is denoted by γ. The first term is the work done to the system by stresses, and the second term represents the conduction of the fluctuation energy.

Following Lun et al., the normal and shear stresses and the dissipation term are given as follows:

$$P_{yy}=\rho_p g_1(\nu,e_p)T, \tag{2}$$

$$P_{yx}=-\rho_p g_2(\nu,e_p)d\frac{du}{dy}T^{1/2}, \tag{3}$$

$$\gamma=\frac{\rho_p}{d}g_3(\nu,e_p)T^{3/2}, \tag{4}$$

where ρ_p is the density of the solid particle, and d is the diameter of the particle. The granular temperature, T, is defined by $1/3\,(\langle u'^2\rangle+\langle v'^2\rangle+\langle w'^2\rangle)$ where u', v', and w' are three velocity fluctuation components. And $g_1(\nu,e_p)$, $g_2(\nu,e_p)$, and $g_3(\nu,e_p)$ are functions of solid fraction, ν, and the particle-particle coefficient of restitution, e_p.

For simple shear flow with uniform density and granular temperature, the conduction term in the energy equation (1) vanishes. Therefore, the shear work term and the dissipation term should balance. Using equation (1) with (3) and (4), the ratio of the characteristic velocity gradient to the granular temperature is obtained as follows:

$$S=\frac{d\dfrac{du}{dy}}{T^{1/2}}=\left(\frac{g_3}{g_2}\right)^{1/2}. \tag{5}$$

Note that S is a function only of ν and e_p. Therefore, (2) and (3) can be written as follows:

$$P_{yy}=\rho_p\left(d\frac{du}{dy}\right)^2\frac{g_1g_2}{g_3}, \tag{6}$$

$$P_{yx}=\rho_p\left(d\frac{du}{dy}\right)^2\frac{g_2^{3/2}}{g_3^{1/2}}. \tag{7}$$

The ratio of shear stress to normal stress, or friction coefficient, is also a function only of ν and e_p.

On the other hand, fully developed chute flow does not have uniform temperature and solid fraction over the depth of the flow. Therefore, the conduction term remains in the energy equation (1) and it plays an important role in determining the profiles of granular temperature, solid fraction, and velocity (see Ahn et al. (1989)). Simple momentum principles are sufficient to demonstrate that for fully developed chute flow, the ratio of shear stress to normal stress is a constant given by tan θ where θ is the angle of chute inclination. From (2) and (3), therefore, S is given by

$$S=\frac{d\dfrac{du}{dy}}{T^{1/2}}=\frac{g_1}{g_2}\tan\theta. \tag{8}$$

It should be noted that S is a function not only for ν and e_p but also of tan θ. And since ν varies over the depth of the chute flow, S also varies, tending to zero at the free surface. Under these circumstances, (2) and (3) can be written as

$$P_{yy}\tan^2\theta=\rho_p\left(d\frac{du}{dy}\right)^2\frac{g_2^2}{g_1}, \tag{9}$$

$$P_{yx}\tan\theta=\rho_p\left(d\frac{du}{dy}\right)^2\frac{g_2^2}{g_1}. \tag{10}$$

These characteristics of chute flows will be important in considering the results presented in this paper.

3 Experimental Measurements

The present experiments were conducted in a long rectangular aluminum channel or chute—7.62 cm wide and 1.2 m

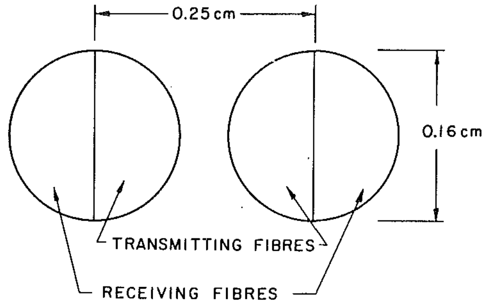


Fig. 1 Geometry of the faces of the two displacement probes used for velocity measurements with the 1.26-mm diameter glass beads

long. The chute was installed in a continuous flow, granular material facility, as previously described in Patton et al. (1987). The material enters the chute from an upper hopper and is collected in a collecting hopper from which, in turn, a mechanical conveyor delivers the material to the upper hopper. The channel is positioned at different angles, θ, to the horizontal. Measurements were taken only after a steady-state flow had been established. The flow into the channel is regulated by a vertical gate, and the opening between the gate and the channel base is referred to as the entrance gate height, h_o.

In these experiments two sizes of glass beads were used as granular materials; one is of mean diameter $d=1.26$ mm with 2.9 percent standard deviation, and the other has $d=3.04$ mm with 7.2 percent standard deviation. The maximum shearable solid fraction, ν^* was estimated from measurements of the values for the two cases in which the materials are densely packed and loosely packed in a container. The 1.26-mm beads had a value of $\nu^*=0.61$. For the 3.04 mm beads, ν^* is 0.59. The density of both granular materials is $\rho_p=2500$ kg/m^3.

Two important instruments were used in the experiments; one is a gauge to measure shear stress, and the other is a set of two fiber-optic probes to measure mean velocity and velocity fluctuation. In order to measure shear stress of flowing material at the chute base, a rectangular hole, 11.4 cm long and 3.8 cm wide, was cut into the chute base and replaced by a plate supported by strain-gauged flexures sensitive to the shearing force applied to the plate. Calibration of this balance was achieved by placing weights on the plate with the channel set at various inclinations. The clearance between the plate and the rest of the chute base was adjusted to be about 0.2 mm, much smaller than the particle sizes. Nevertheless, dirt would occasionally get trapped in the gap and this necessitated cleaning of the gap prior to each measurement.

A system of fiber-optic probes, similar to that originally devised by Savage (1979), was developed to measure particle velocities and their fluctuations at the chute base, the free surface, and the sidewalls. The system consisted of two MTI fiber-optic displacement probes set with their faces flush in a lucite plug which was, in turn, either set flush in the chute base or sidewalls or held close to the free surface of the flowing granular material. The probe faces were 1.6 mm in diameter and of the type in which one semicircle of the face consisted of transmitting fibers and the other of receiving fibers. The specific geometry is shown in Fig. 1. The distance between two displacement probes was selected to be about two particle diameters. This distance was carefully calibrated by placing the probes close to a revolving drum to which particles had been glued, and comparing the drum peripheral velocity with the velocity measured from the probe output.

The output from these velocity measuring devices was processed in the following way. First, the signals from each of the two displacement probes were simultaneously digitized and stored using a data acquisition system. Sampling rate was varied, depending on the mean velocity of particles. For most flows, the record time was about 0.5 second, recording at 3×10^4 samples/sec. Typically, each record detected the passage of 300 ~ 600 particles. The two records were digitally cross-correlated over the entire record in order to obtain the mean particle velocity, u. This information was then used to identify the peaks on the two records corresponding to the passage of a particular particle. When no such correspondence could be established or where the peak was below a certain threshold, the data was discarded for the purposes of this second part of the analysis. However, where positive identification was made, the velocity of that individual particle was obtained from the time interval between the peaks it generated on the two records. In this way, a set of instantaneous particle velocities were obtained, and ensemble-averaging was used to obtain both the mean velocity, u, and the root mean square of velocity fluctuation, u'. Though the latter represents only one component of velocity fluctuations, it should be some measure of granular temperature. Finally, the number of particle passages per unit time detected by the probe was divided by the mean velocity to obtain the characteristic particle spacing, C_{1D}, and in turn the linear concentration, ν_{1D}, was calculated as $\nu_{1D}=d/C_{1D}$ where d is the mean diameter of the particles. An estimate of the local solid fraction near the wall, ν_w, was calculated using $\nu_w=\pi\nu_{1D}^3/6$.

In addition, point probes were used to record the depth, h, of flow at several longitudinal locations in the channel. Mass flow rate, $\dot{m}$, was obtained by timed collection of material discharging from the chute. Mean velocities at the chute base and at the free surface obtained by the fiber-optic probes were averaged to give the average mean velocity, u_m, over the depth of the flow. A mean solid fraction, ν_m, could then be obtained as $\nu_m=\dot{m}/\rho_p hbu_m$ where b is the channel width. Furthermore, mean shear rate was calculated as $\Delta u/h$ where Δu is the difference between the two velocities at the base and at the free surface, and h is the depth of the flow. Normal stress was calculated by $\tau_N=\rho_p\nu_m gh\cos\theta$ where g is the gravitational acceleration, and shear stress, τ_S, was measured directly by the shear gauge. All the above measurements except for shear stress were made at two stations located at 72 cm and 98 cm downstream from the entrance gate. The shear gauge was located in the middle of these two stations. It should be noted that ν_w, u, u', τ_N, and τ_S are local properties while ν_m and $\Delta u/h$ represent quantities averaged over the depth of flow. The data from all the measurements were quite repeatable, and the data presented here are typically averages over two or five measurements.

Preliminary tests suggested that the flow could be influenced by the surface conditions of the chute base. Indeed, the data were quite sensitive to the degree of the cleanliness of the aluminum chute base. Therefore, it was possible to create different surface conditions with the aluminum chute by controlling the cleanliness. In addition, a very thin film of liquid rubber (Latex) was applied to the chute base to give a totally different surface condition. This film was about 0.2 mm thick. Before conducting experiments, the chute was run long enough to achieve a steady-state surface condition. With these precautions, data will be classified in this presentation by whether the chute base was "smooth," "moderately smooth," or "rubberized." The state of being moderately smooth was quite stable, but the smooth surface condition was less stable, requiring careful control of the cleanliness. To systematically characterize these different surface conditions, Coulombic friction coefficients were measured using the shear gauge and a block to which glass beads were glued. The kinematic Coulombic friction coefficient of the smooth surface was 0.15; the moderately smooth and rubberized surface had coefficients of 0.22 and 0.38, respectively. Furthermore, smooth and moderately smooth surfaces yielded coefficients of restitution different from that of the rubber-coated surface; the former was 0.7 while the latter 0.5. Both were measured by observing an

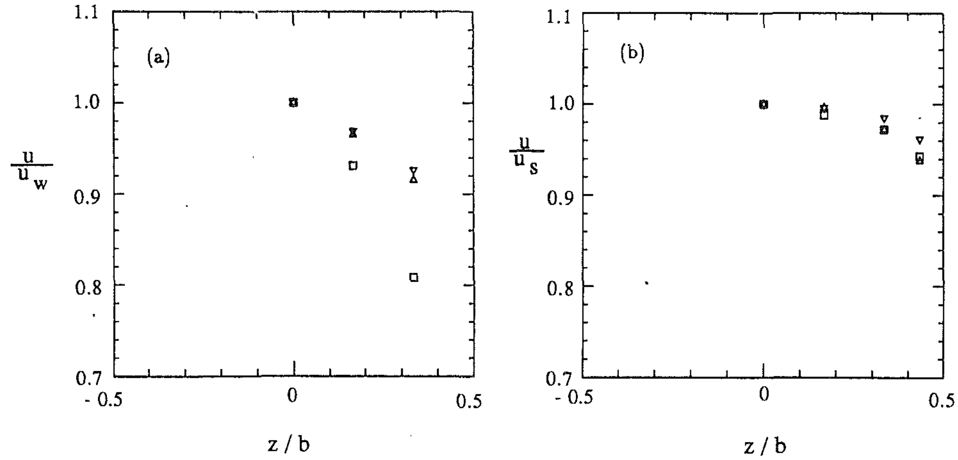


Fig. 2 The transverse velocity profiles (*a*) at the chute base and (*b*) at the free surface; mean velocity normalized by mean velocity at the center against the lateral location, *z*, normalized by the chute width, $b = 76.2$ mm. □, $\theta = 17.8$ deg, $\nu_m = 0.54$, $u_w = 0.898$ m/sec, and $u_s = 1.118$ m/sec; Δ, $\theta = 22.7$ deg, $\nu_m = 0.50$, $u_w = 1.386$ m/sec, and $u_s = 1.639$ m/sec; ∇, $\theta = 32.2$ deg, $\nu_m = 0.49$, $u_w = 2.055$ m/sec, and $u_s = 2.263$ m/sec. u_w, velocity at the center on the chute base; u_s, velocity at the center on the free surface.

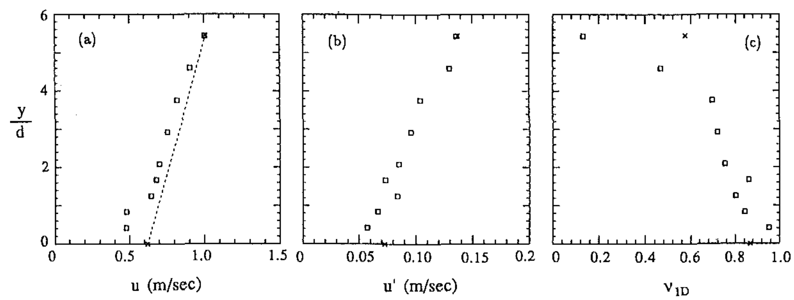


Fig. 3 The vertical profile at the sidewall, □. Vertical location, *y*, normalized by the particle diameter *d* against (*a*) mean velocity, (*b*) velocity fluctuation, and (*c*) linear concentration. ×, data at the center of the chute. Dotted line, the assumed velocity profile at the center of the chute. Data were taken at $\theta = 17.8$ deg on the rubberized surface; $\nu_m = 0.30$, $h_o = 25.4$ mm, and $d = 3.04$ mm.

individual particle colliding with the surface. These coefficients of restitution, e_w, between the wall surface and a particle should be distinguished from that between two particles, e_p, which was not measured here.

4 Preliminary Observations on Profiles

Originally the chute was designed to be wide enough to yield almost two-dimensional flow. To examine the effect of the sidewalls (and the extent to which this objective was achieved), fiber-optic probe measurements were made at several lateral locations with various chute inclinations. The 1.26-mm glass beads were used in measurements of the transverse velocity profiles, and the surfaces of the aluminum chute base as well as the sidewalls were smooth. Velocities normalized by the velocity at the centerline are plotted in Fig. 2. Comparison of the profiles on the free surface and on the chute base indicates that the flow at the free surface is more uniform and less affected by sidewall than the flow at the base. This is a "corner effect" in which particles in the corner are slowed both by the chute base and the side wall. One could visually observe that particles in the corner are arranged in a distinct line which has high solid fraction and low velocity. It should also be noted from Fig. 2 that the higher the velocity (or the higher the chute inclination), the less significant the sidewall effect. Thus non-uniformity, due to the sidewall, was significant only at the base and at low velocities (low inclinations). We were particularly concerned about the sidewall effect on the shear gauge whose width was one half of that of the channel. The foregoing results indicated that this sidewall effect would be very small.

Vertical profiles were obtained by making measurements through lucite windows in the sidewalls. Savage (1979) made similar efforts to obtain velocity profiles at the sidewalls using fiber-optic probes. Bailard (1978) obtained the vertical profiles of velocity and solid fraction by measuring cumulative mass flux profiles. Campbell and Brennen (1985b) in the computer simulation with circular discs obtained the profiles of velocity, granular temperature, and solid fraction. In the present work, fiber-optic probes were used to measure velocity, its fluctuation, and linear concentration. It should be noted that, usually, fully developed flow could not be achieved because of the finite length of the chute.

One typical example of the vertical profiles is included in Fig. 3, the measurements being taken with 3.04-mm glass beads with a chute inclination of 17.8 deg and a rubberized chute base. As illustrated in Fig. 3(*a*), the velocity profile is fairly linear except within a distance of about one particle diameter from the base. The uniform velocity within the distance of one-particle diameter indicates that there is a distinct layer at the corner, preventing particles from entering the layer from above, assuring the existence of the "corner effects." Note that the ratio of velocity at the base to that at the free surface

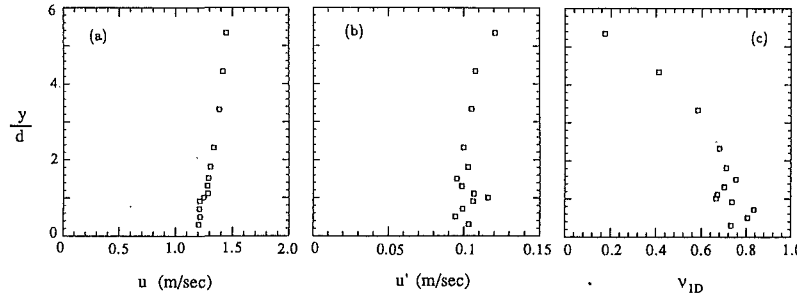


Fig. 4 The vertical profile at the sidewall, □. Vertical location, *y*, normalized by the particle diameter *d* against (*a*) mean velocity, (*b*) velocity fluctuation, and (*c*) linear concentration. Data were taken at $\theta = 22.7$ deg on the smooth surface; $h_o = 15.9$ mm and $d = 1.26$ mm.

is about one half, which is comparable with the results of the computer simulation by Campbell and Brennen (1985b). This result should be distinguished from those of Savage (1979) and Bailard (1978) where almost zero velocity was obtained at the chute surfaces roughened by rough rubber sheets or attached particles.

Velocities at the center of the chute, both at the base and the free surface, are shown in Fig. 3(*a*) for comparison with the velocities at the sidewall. At the free surface, the velocities at the center and at the sidewall are almost equal. But at the base there is some discrepancy due to the corner effect. This characteristic of the data suggests that the velocity profile in the center of the chute is similar to that at the sidewall except within one particle diameter distance from the base. An assumed velocity profile at the center is shown by the dotted line in Fig. 3(*a*). We also conclude from these observations that the shear rate, du/dy, can be approximated by the difference between the base and free-surface velocities, Δu, divided by the depth, h, of the flow. This approximation has been used throughout the analysis which follows.

The profile of velocity fluctuations at the sidewall is plotted in Fig. 3(*b*). It can be seen that the profile is fairly linear, and that fluctuations are larger at the free surface than at the chute base. Comparison between the sidewall and center values is also included in Fig. 3(*b*). At the free surface, no significant difference is encountered between the velocity fluctuations at the sidewall and at the center. But at the base, a small discrepancy is observed which is again believed to be due to the corner effect. This fairly linear profile for velocity fluctuation was observed in most flows. Furthermore, the fluctuations were always higher at the free surface than at the base of the chute. These overall features are in contrast to the results obtained by Campbell and Brennen (1985b). In their computer simulation, granular temperature near the solid wall was substantially higher than near the free surface, and the profile was far from linear. We believe this difference is probably due to that fact that 0.6 was used for e_p in the computer simulation, while e_p for glass beads is more like 0.95 (Lun and Savage (1986); refer to Ahn et al. (1989) for more detail).

When Fig. 3(*b*) is closely examined, it raises some complicated problems in measurements of granular flows. For instance, a slight peak in the velocity fluctuation was consistently observed at a distance of one-particle diameter from the chute base. This location coincides with the interface between the first and second layers of particles which are quite distinct because of the corner effect. Within each distinct layer, the fiber-optic probes measure only longitudinal fluctuations for the particles within that layer. At the interface, however, particles from both layers contribute, and hence the difference in the mean velocities in the two layers enters into the result. Therefore, the fluctuations at the interface were observed to be slightly higher than elsewhere.

The profile of linear concentration, ν_{1D}, is presented in Fig. 3(*c*). Again, in the region near the base, the locations of the first and second layers and their interface can be determined by the details of the profile. The peak at $y/d \simeq 0.5$ indicates the location of the center of the first layer; the interfacial region has a lower concentration; the peak at $y/d \simeq 1.7$ corresponds to the center location of the second layer. This detailed structure seems to disappear above the second layer. Near the free surface, the linear concentration decreases gradually, and as a result the free surface is not clearly defined as it might otherwise be.

The monotonic decrease of solid fraction with distance from the wall as shown in Fig. 3(*c*) was a somewhat unexpected result. From previous experiments (Bailard (1978)) and from computer simulations (Campbell and Brennen (1985b)), it has been observed that solid fraction increases with distance from the base and it vanishes at the free surface after it achieves its maximum in the bulk. The discrepancy between the profile of the present experiments and the results of Bailard may be due to the different surface conditions used in the experiments. The experiments of Bailard used the surface on which particles were glued to create a no-slip condition at the boundary. On the other hand, the present experiments used relatively smooth surfaces. The discrepancy between the present data and the results of Campbell and Brennen may arise from the fact that the value of e_p used by Campbell and Brennen is different from that of the glass beads in the present experiments. The results of Ahn et al. (1989) show that the profile of solid fraction can be either of Campbell and Brennen or of the present one, depending on the value of e_p.

Similar sidewall measurements were made with other sizes of glass beads and at other chute inclinations (Ahn (1989)); the general features of these profiles are similar to those of the preceding example though the data with a smooth aluminum base differed somewhat from that with the rubberized base. To illustrate this, measurements with the 1.26-mm glass beads at a chute inclination of 22.7 deg with the smooth aluminum base are presented in Fig. 4. Compared to the data on the rubberized surface, the profiles of velocity and velocity fluctuation are more uniform. The velocity at the wall is more than 80 percent of that at the free surface. Velocity fluctuation is fairly uniform, although there is a slight increase with distance from the chute base. The detailed structure of the layers due to the corner effect is clearly observed in all the profiles.

5 Presentation of Experimental Data

5.1 Experimental Data on Basic Flow Properties. In this section, we examine how basic flow properties (such as velocities, velocity fluctuation, and shear rate) vary with solid fraction. Two kinds of solid fraction are used in this presentation; mean solid fraction, ν_m, and wall solid fraction, ν_w. The mean solid fraction is an average value over the depth of flow, and the wall solid fraction describes a density in the vicinity of the chute base. Because it is calculated from a measurement of linear concentration, the wall solid fraction may not represent accurately the local solid fraction near the wall, but it is at least a qualitative, comparative measure.

The ratio of velocity at the wall, u_w, to velocity at the free surface, u_s, is plotted against mean solid fraction in Fig. 5. Different symbols are used for different surface conditions. For the smooth surface, the ratio u_w/u_s is fairly constant and greater than 0.9, implying that the velocity profile over the depth is close to uniform. On the other hand, for the rubberized surface, the ratio increases with decreasing ν_m. In other words, the lower the solid fraction, the more uniform the velocity profile. Note the rather sudden change of u_w/u_s at $\nu_m \simeq 0.1$ which will be discussed later. As expected, the data for the moderately smooth surface lie between those for the smooth surface and the rubberized surface.

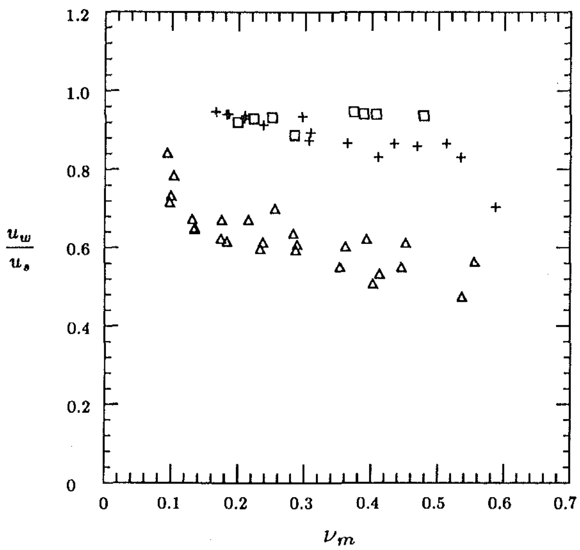


Fig. 5 The ratio of velocity at the chute base wall to velocity at the free surface, u_w/u_s, against mean solid fraction, ν_m. □, the smooth surface, +, the moderately smooth surface; Δ, the rubberized surface.

The mean shear rate, $\Delta u/h$, is plotted against mean solid fraction in Fig. 6. For the smooth surface (see Fig. 6(*a*)), the shear rate monotonically increases with decreasing ν_m. (Recall, however, u_w/u_s remains constant as shown in Fig. 5). On the other hand, the moderately smooth and rubberized surfaces (see Figs. 6(*b*) and (*c*)) yield shear rates which first increase and then decrease as the solid fraction decreases. The values of ν_m at which the shear rate is a maximum are about 0.3 for the moderately smooth surface, and about 0.2 for the rubberized surface regardless of the particle size. Note that the steep change of the shear rate at $\nu_m \simeq 0.1$ for the rubberized surface corresponds to that of u_w/u_s in Fig. 5.

The variation of the velocity fluctuation at the wall, u'_w, with wall solid fraction is examined in Fig. 7(*a*). Regardless of surface conditions, u'_w increases with decreasing ν_u. The use of wall solid fraction was essential for the examination of the local quantity u'_w. To illustrate this, the local quantity u'_w was plotted against the mean quantity ν_m as shown in Fig. 7(*b*). The use of the mean quantity with the local quantity leads to a wide scattering of the data. If examined more closely, the data reflected a strong dependency on the entrance gate opening h_o. As observed in Fig. 7(*b*), the data have a distinct line for each h_o. This is because the mean solid fraction is closely related to h_o. Note that fully developed flow was not achieved in the present experiments (this will be discussed later). Therefore, the test section was directly affected by entrance conditions governed by h_o. When ν_w is used, the dependency on h_o largely disappears as shown in Fig. 7(*c*).

It is also interesting to present the velocity fluctuation in a nondimensionalized form. In Fig. 8, the velocity fluctuation normalized by the mean velocity is plotted against wall solid fraction. Note all the quantities are local values measured at

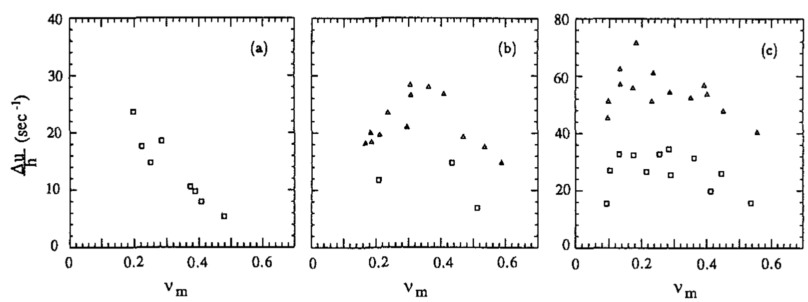


Fig. 6 The shear rate, $\Delta u/h$, against mean solid fraction, ν_m: (*a*) the smooth surface, (*b*) the moderately smooth surface, and (*c*) the rubberized surface. □, $d=3.04$ mm; Δ, $d=1.26$ mm.

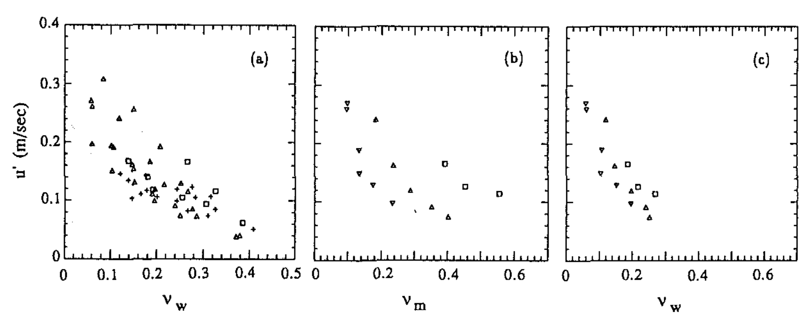


Fig. 7 The longitudinal velocity fluctuation at the wall, u'_w, against solid fraction with the rubberized surface. (*a*) Data for $d=1.26$ mm and $d=3.04$ mm. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface. (*b*) Data for $d=1.26$ mm. □, $h_o=38.1 \sim 50.8$ mm; Δ, $h_o=25.4$ mm; ∇, $h_o=12.7 \sim 15.9$ mm. (*c*) Data as in (*b*).

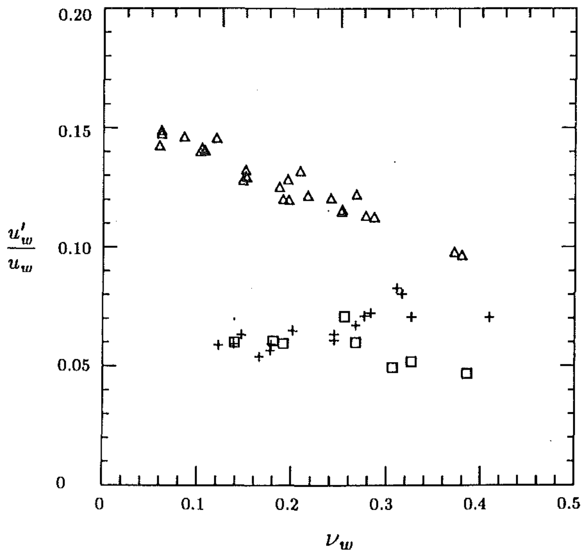


Fig. 8 The longitudinal velocity fluctuation at the chute base wall normalized by mean velocity, u'_w/u_w, against wall solid fraction, ν_w. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface.

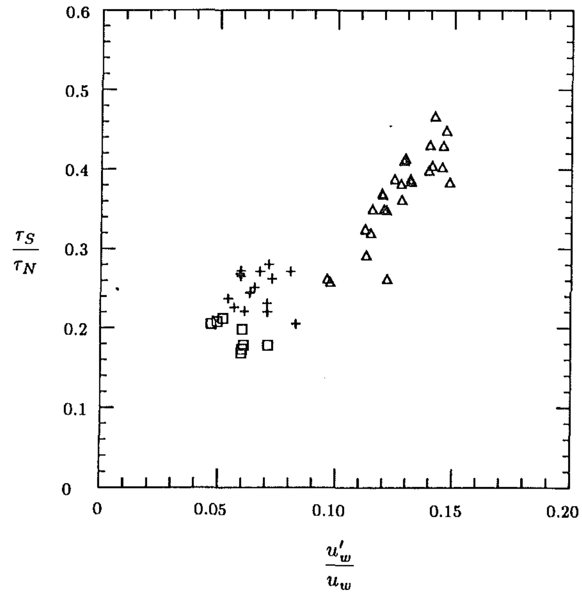


Fig. 10 Friction coefficient at the wall, $f = \tau_S/\tau_N$, against longitudinal velocity fluctuation at the wall normalized by mean velocity, u'_w/u_w. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface.

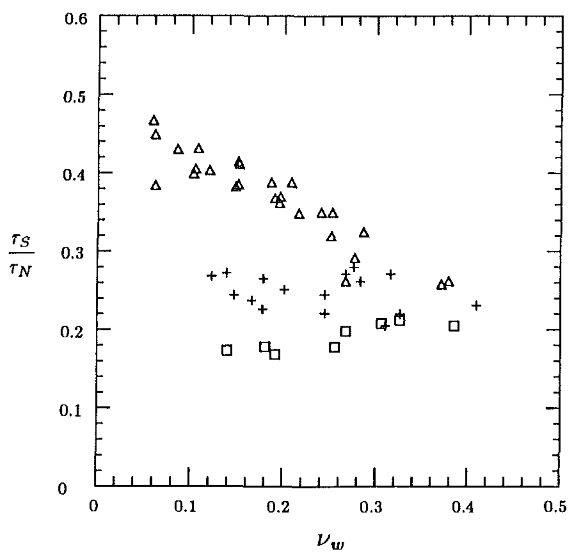


Fig. 9 Friction coefficient at the wall, $f = \tau_S/\tau_N$, against wall solid fraction, ν_w. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface.

the wall. The ratio of u'_w to u_w for the rubberized surface is larger than that for the smooth surface. The ratio u'_w/u_w for the smooth surface shows little variation with ν_w. For the moderately smooth surface, u'_w/u_w changes only mildly with ν_w. However, the rubberized surface clearly shows the increase of u'_w/u_w with decreasing ν_w.

5.2 Experimental Data on Friction Coefficient. As previously mentioned, shear stress was directly measured by the shear gauge, and normal stress was calculated as $\rho_p \nu_m g h \cos \theta$ where ρ_p is the density of particles, g is the gravitational acceleration, h is the depth of flow, and θ is the angle of the chute inclination. Note both stresses were measured at the chute base wall. Recall from Section 3 that kinematic Coulombic friction coefficients, μ_c, were measured for each surface condition; 0.15 for the smooth surface, 0.22 for the moderately smooth surface, and 0.38 for the rubberized surface.

The ratio of shear stress to normal stress, or friction coefficient, f, is plotted against wall solid fraction in Fig. 9. For the smooth and moderately smooth surfaces, friction coefficients appear to be fairly constant. Furthermore, the values of friction coefficients are comparable to the kinematic Coulombic friction coefficients for each surface (though f is slightly higher than μ_c). On the other hand, for the rubberized surface, the friction coefficient is a decreasing function of solid fraction. And the Coulombic friction coefficient for the rubberized surface does not seem to directly affect the friction coefficient for the flowing material. Therefore, it may be concluded that the different surface conditions result in quite different types of boundary condition at the wall.

In Fig. 10, the friction coefficient is plotted against velocity fluctuation normalized by mean velocity, or u'_w/u_w. For the smooth and moderately smooth surfaces, all the data are clustered at one region. For the rubberized surface, f seems to correlate quite well with u'_w/u_w; f increases with increasing u'_w/u_w. This phenomenon is independent of particle size.

5.3 Experimental Results on Rheological Behavior. The data on the normal and shear stresses will be examined by comparison with the rheological model of Lun et al. (1984). In particular we examine the stresses by normalizing by $\rho_p(u'_w)^2$ and $\rho_p(d\Delta u/h)u'_w$ (see equations (2) and (3), for any kind of flow). Other possible normalizing factors which merit investigation are $\rho_p(d\Delta u/h)^2$ (see equations (6) and (7), for simple shear flow), $\rho_p(d\Delta u/h)^2/\tan^2\theta$, and $\rho_p(d\Delta u/h)^2/\tan\theta$ (see equations (9) and (10), for fully developed flow). In this investigation, it is important to recall that Lun et al. assume that the granular temperature is isotropic, and that the effects of particle rotation and surface friction are not included in their model. One could, therefore, expect some discrepancies in comparison with the experimental data.

The theory of Lun et al. suggests that the appropriate normalizing factor of the normal stress should be $\rho_p(u'_w)^2$, and the experimental data thus normalized is plotted against the wall solid fraction in Fig. 11(a). This method of normalization appears to correlate the data quite well and seems to collapse the data for the different surface conditions. When these same values are plotted against the mean solid fraction as in Fig. 11(b), the data are more scattered. This may be explained by realizing that the normalized stress is a local quantity which should be related to the local wall solid fraction rather than the mean solid fraction. In both figures, results for the rheo-

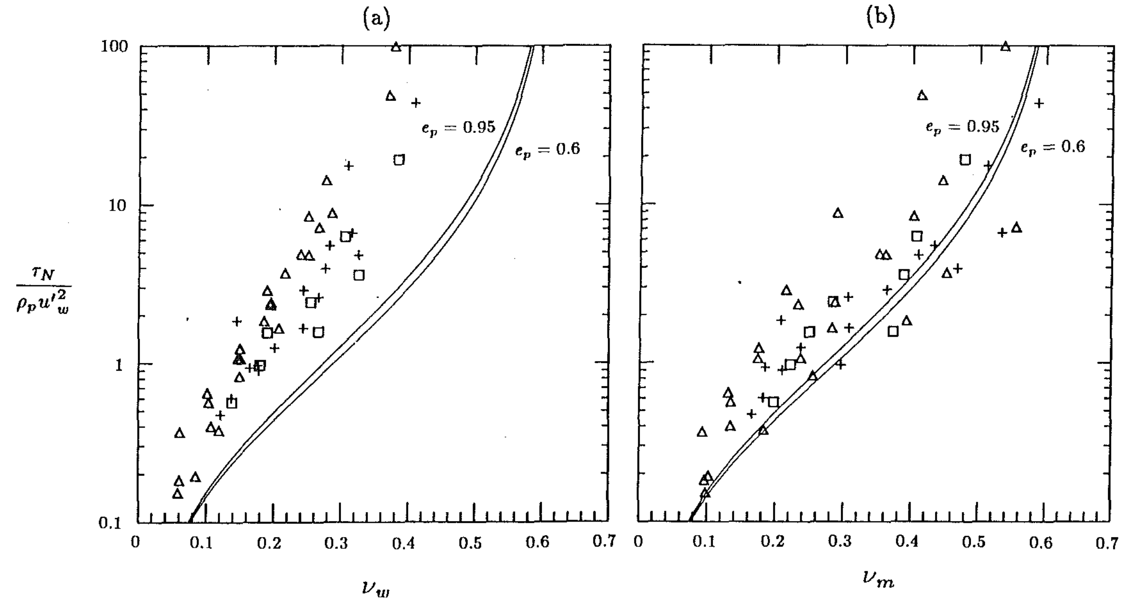


Fig. 11 The normalized normal stress, $\tau_N/\rho_p u_w'^2$, against (*a*) wall solid fraction, ν_w, and (*b*) mean solid fraction, ν_m. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface. The solid lines, the results of Lun et al. (1984).

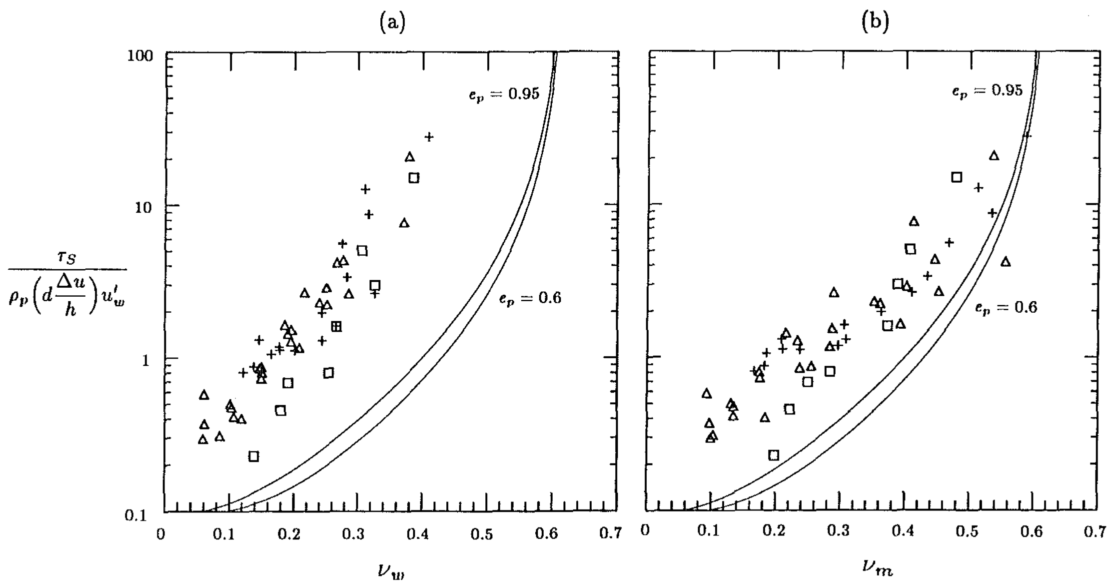


Fig. 12 The normalized shear stress, $\tau_S/\rho_p d(\Delta u/h)u_w'$, against (*a*) wall solid fraction, ν_w, and (*b*) mean solid fraction, ν_m. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface. The solid lines, the results of Lun et al. (1984).

logical model postulated by Lun et al. (1984) (see equation (2)) are also plotted using $T=\langle u'^2\rangle$.

On the other hand, the theory of Lun et al. suggests that the shear stress should be normalized by $\rho_p(d\Delta u/h)u_w'$, and the resulting experimental data is presented in Fig. 12. Again the data is well correlated regardless of surface conditions when plotted against the wall solid fraction, and the data is less satisfactorily correlated with the mean solid fraction. The results of Lun et al. (1984) for any general flow (see equation (3)) are shown in the same figure for comparison. The quantitative discrepancy between the theoretical and experimental results is substantial.

It should also be observed that alternative normalizations with $\rho_p(d\Delta u/h)^2$, $\rho_p(d\Delta u/h)^2/\tan^2\theta$, and $\rho_p(d\Delta u/h)^2/\tan\theta$ yielded less satisfactory correlation of the data than in Figs. 11(*a*) and 12(*a*) (see Ahn (1989)). This strongly implies that the rheological models of Lun et al. have considerable merit in so far as the functional dependence on the flow parameter is concerned though the quantitative values of some of the coefficients may be significantly in error.

We now examine the parameter, S, introduced by Savage and Jeffrey (1981) where

$$S=\frac{d\,\dfrac{du}{dy}}{T^{1/2}}\,.$$

The model of Lun et al. (1984) predicts that S should be a function only of ν and e_p for simple shear flow and that $S/\tan\theta$ should likewise be a function only of ν and e_p in fully

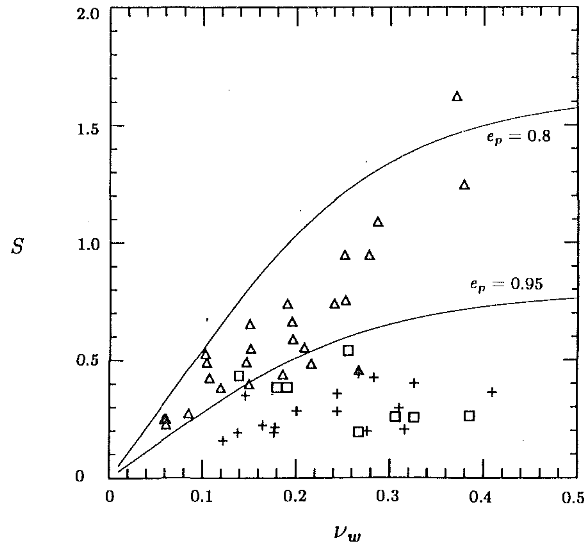


Fig. 13 The parameter, $S = d(\Delta u/h)/u'_w$, against wall solid fraction, ν_w. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface. The solid lines, the results of Lun et al. (1984).

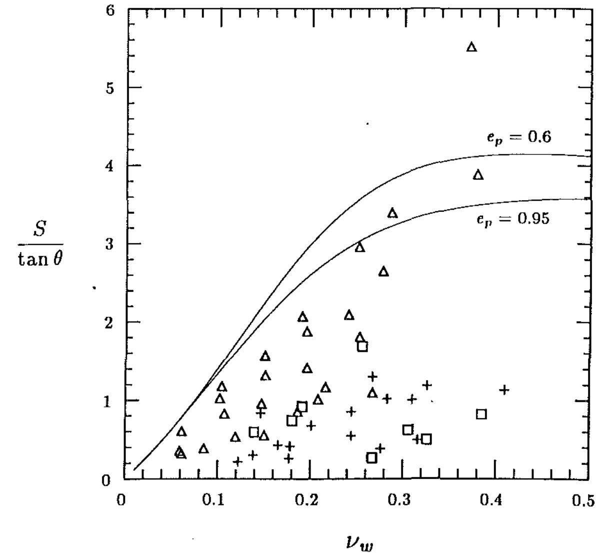


Fig. 14 The parameter, $S/\tan\theta = d(\Delta u/h)u'_w \tan\theta$, against wall solid fraction, ν_w. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface. The solid lines, the results of Lun et al. (1984).

developed flow (see equations (5) and (8)). Here we estimate S by $d(\Delta u/h)/u'_w$. The parameter S is plotted against the wall solid fraction in Fig. 13 while $S/\tan\theta$ is presented in Fig. 14. In both figures the data is widely scattered showing strong dependency on surface conditions. We believe, for reasons stated later, that this is due to the fact that the flow is not a simple shear flow and that only a subset of the data represent fully developed flows.

6 Discussion

6.1 The Characteristics of Chute Flows. It is apparent from Fig. 5 that the surface condition has considerable influence on the characteristics of chute flow. Different results have been achieved by several authors when different surface conditions are used. For example, Bailard (1978) used a surface on which grains were glued, and Savage (1979) applied roughened rubber sheets to the surface. In both cases, the ratio of u_w to u_s was close to zero. Augenstein and Hogg (1978) obtained various u_w/u_s for various surface roughnesses. When a smooth surface with high friction coefficient was used by Campbell and Brennen (1985b) in computer simulations, the ratio of u_w to u_s was about $0.4 \sim 0.5$. The rubberized surface of the present experiments, therefore, is similar to those cases in Campbell and Brennen in which a no-slip condition at the contact surface was assumed. Despite these data, the present state of knowledge does not allow prediction of the slip at the wall. Indeed, the features of the surface or of the flow which determine the slip are not well understood.

The surface conditions also influence velocity fluctuations at the wall as observed in Fig. 8. For the smooth and moderately smooth surfaces, the ratio of u'_w to u_w is low and fairly constant. On the other hand, u'_w/u_w for the rubberized surface is high and increases as solid fraction decreases. These observations may imply the following. The rubberized surface is characterized by large velocity fluctuations particularly at lower solid fractions. The high fluctuations and the low solid fraction allow particles to move more freely from one location to another. One of the effects by these random motions is a decrease of velocity gradient in the direction normal to the flow. That is, when particles move from a layer with low mean velocity to a subsequent layer with high velocity, the mean velocity of the layer with high velocity is reduced. When particles move due to random motion from the upper layer with high mean velocity to the lower layer with low mean velocity, the opposite is true. This phenomenon is consistent with experimental observations. For the rubberized surface, u_w/u_s rather sharply increases at $\nu \simeq 0.1$ as the solid fraction is decreased (see Fig. 5). As also observed in Fig. 6(c), since at low solid fraction velocity fluctuation is high and there is more space for particles to freely move, the shear rate decreases as solid fraction decreases. In Fig. 6(a), however, this decrease of the shear rate is not observed with the smooth surface since no substantial velocity fluctuation exists (see Fig. 8).

6.2 Friction Coefficient and Boundary Conditions. In the present work, an attempt to investigate boundary conditions was made by changing chute surface conditions. The Coulombic friction coefficient, μ_c, was measured for each surface since it was anticipated that μ_c would be a major factor which determines whether or not particles slip when in contact with solid boundary. Here the word "slip" means the tangential slip between the contact surfaces of the particle and the wall. The slip velocity is different from a velocity at the wall, u_w, which is the velocity of the particle center extrapolated to the wall. Clearly, even when slip velocity is zero, a particle touching the wall may roll and thus have a nonzero center velocity.

When a particle collides with a wall such that the shear stress at the contact point exceeds a shear stress limit which the surface can withstand for the given normal stress at the contact point, slip will occur. Then the ratio of the shear stress to the normal stress at the contact point is adjusted to the Coulombic friction coefficient of the surface, i.e. $f = \mu_c$. On the other hand, when the ratio of τ_S to τ_N at the impact does not exceed μ_c, there will be no slip between the contact surfaces of the particle and the wall. In this case f is different from μ_c.

As seen in Fig. 9, friction coefficients for the smooth and moderately smooth surfaces seem to be fairly constant. But for the rubberized surface the friction coefficient decreases with increasing solid fraction. Decreasing friction coefficients with increasing ν were also observed in the shear cell experiments of Savage and Sayed (1984) and in the computer simulations of Campbell (1989). However, the constant friction coefficients of the smooth and moderately smooth surfaces have not been observed previously. To explain these observations, we suggest the following. For the smooth and moderately smooth surfaces, slip occurs at the contact between particles and the surfaces, and the slip condition results in the constant friction coefficient equal to μ_c. On the other hand, the varying friction coefficient for the rubberized surface suggests a no-slip condition at the boundary. The high Coulombic friction coefficient of the rubberized surface would inhibit any slip at the contact between particles and the surface.

When there is no slip, the following relation results from a simple analysis of the oblique impact of a single sphere on the flat surface (see Ahn (1989)):

$$f=\frac{2}{7}\left(1-\frac{\omega_1 r}{u_1}\right)\frac{\tan\alpha_1}{1+e_w}$$

where ω_1 is the rotational rate before impact, r is the radius of the sphere, and u_1 is the velocity tangential to the wall before impact. The impact angle α_1 is defined by $\tan^{-1}(u_1/v_1)$ where v_1 is the velocity normal to the wall before impact, and e_w is the wall-particle coefficient of restitution. In this equation, the friction coefficient or the ratio of τ_S to τ_N at the surface depends on the ratio of rotational velocity to tangential velocity, $\omega_1 r/u_1$, and on the impact angle, $\tan\alpha_1$.

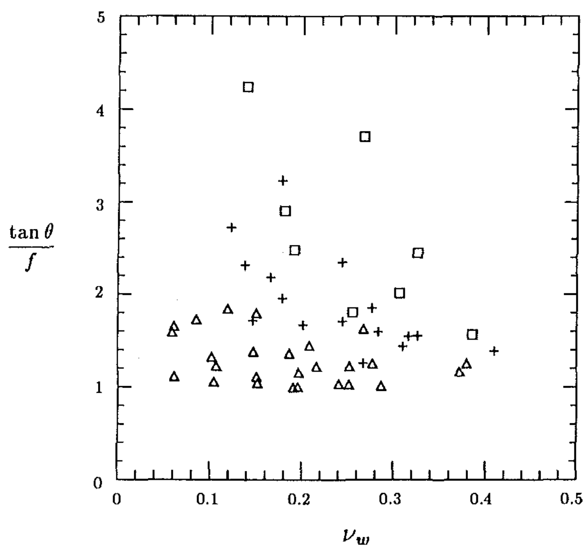


Fig. 15 The ratio of tan θ to f against wall solid fraction, ν_w. □, the smooth surface; +, the moderately smooth surface; Δ, the rubberized surface.

In the present experiments, the values of $\tan\alpha_1$ could not be estimated. Another factor influencing f is the ratio of the rotational velocity ωr to the tangential velocity u before impact. Campbell (1988) has shown that next to the wall ω is considerably larger than the mean value, but that with a small distance from the wall ω is slightly less than the mean value. Therefore, when a particle next to the wall with high ω hits the wall, the friction coefficient will be low, but if a particle at a distance from the wall with low ω comes down and collides with the wall, the friction coefficient will be relatively high.

These phenomena suggests a possible explanation for the decrease in the friction coefficient as the solid fraction increases. At low solid fraction particles move more freely from one layer to another. Thus more particles in the upper layer with small values of $\omega r/u$ move down to the boundary and make collisions with the wall. Because friction is measured in a statistical sense as a sum of frictions due to individual particles colliding with the wall, f is therefore high at low solid fraction. On the other hand, at high solid fraction and low granular temperature, very few particles in the upper layer with low $\omega r/u$ penetrate to the wall. As a result, particles next to the wall with high rotational velocity will dominate collisions at the wall. Thus, f would be smaller at high solid fraction.

This explanation appears to be consistent with the data for the rubberized surface where no slip is expected (see Fig. 9). The general trend of decreasing f with increasing ν holds independent of particle size. In Fig. 10, f is plotted against u_w'/u_w. When higher u_w'/u_w exists, particles with low ω in the upper layer more easily move down to the boundary and collide with the wall. That is, as u_w'/u_w increases, the intrusion of particles with low ω from the upper layer into the boundary becomes more frequent, causing f to increase. As a result, the friction coefficient appears to be a fairly linear function of u_w'/u_w for the rubberized surface.

For the smooth and moderately smooth surfaces, slip occurs and μ_c controls the boundary conditions. Therefore, f is comparable to μ_c (see Fig. 9), and f is unrelated to u_w'/u_w (see Fig. 10). (However, one might argue from Fig. 10 that for the smooth and moderately smooth surfaces f is small because u_w'/u_w is small. Then it appears that regardless of the surface conditions f has a fairly linear relation to u_w'/u_w).

6.3 Stresses and Rheological Behavior. When the experimental data on the normal and shear stresses are normalized in the same way as in the rheological models by Lun et al.

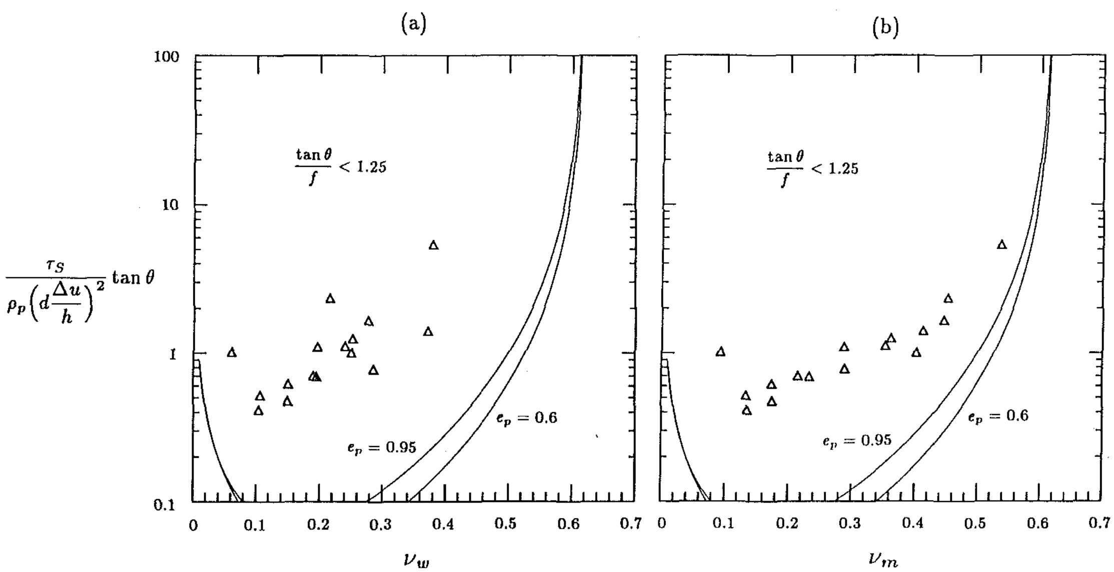


Fig. 16 The normalized shear stress, $\tau_S \tan\theta/\rho_p(d\Delta u/h)^2$, against (a) wall solid fraction, ν_w, and (b) mean solid fraction, ν_m. Data only with $\tan\theta/f < 1.25$. The solid lines, the results of Lun et al. (1984).

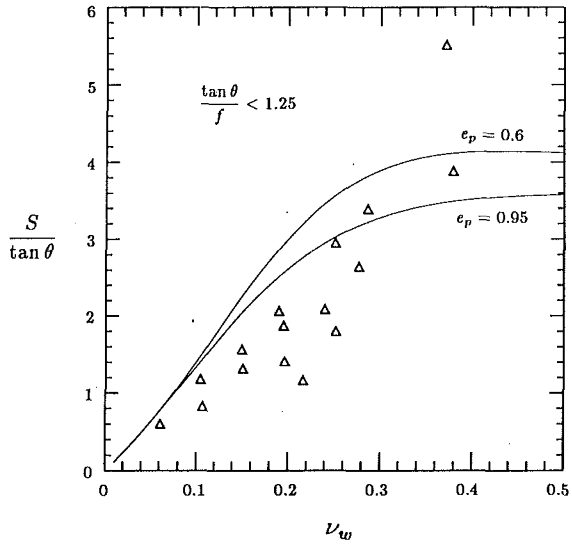


Fig. 17 The parameter, $S/\tan\theta = d(\Delta u/h)/u'_w \tan\theta$, against wall solid fraction, ν_w. Data only with $\tan\theta/f < 1.25$. The solid lines, the results of Lun et al. (1984).

(1984), the measurements turn out to be well correlated as shown in Figs. 11(*a*) and 12(*a*). (The models should hold for any flow whether fully developed flow or not.) The data are also internally consistent, independent of the surface boundary conditions. It is also important to note that other normalizations such as $\tau_N/\rho_p(d\Delta u/h)^2$ and $\tau_S/\rho_p(d\Delta u/h)^2$ (Ahn (1989)) do not lead to such satisfactory collapse of the data and yield curves which appear to depend on the surface boundary condition.

The chute flows in the present experiments were not fully developed. This is confirmed by comparing friction coefficient with the tangent value of the chute inclination angle. The ratio of $\tan\theta$ to f is plotted in Fig. 15. If the flow were fully developed, the ratio would be 1, and the results clearly show that this is not the case. Therefore, when $\tau_N \tan^2\theta/\rho_p(d\Delta u/h)^2$ and $\tau_S \tan\theta/\rho_p(d\Delta u/h)^2$ are plotted against the wall solid fraction, ν_w, considerable scatter is observed since these correlations would only hold for fully developed flows. However, we can select those data points which represent nearly fully developed flow by applying the requirement that $\tan\theta/f < 1.25$ where the 1.25 is somewhat arbitrary. This subset of data is used to present the shear stress normalized by $\rho_p(d\Delta u/h)^2/\tan\theta$ in Fig. 16. It is significant that this subset of data is well correlated in this figure. Though not presented here, the normal stress normalized by $\rho_p(d\Delta u/h)^2/\tan^2\theta$ would also be well correlated when $\tan\theta/f < 1.25$. Furthermore, this subset of data is also used to present $S/\tan\theta$ in Fig. 17 in which the scatter is much less than in Fig. 14. This indicates again that the fully developed flows adhere to the model expected on the basis of the theory of Lun et al. (1984). On the other hand, it is clear that many of the chute flows examined here were not fully developed.

In summary, it may be concluded that the rheological models for general flow (equations (2) and (3)) give good correlation to the present experimental data (see Figs. 11 and 12). The rheological model for fully developed flow (equation (9) or (10)) also agrees with a subset of experimental data which is judged to be fully developed (see Fig. 16).

7 Summary and Conclusion

Experiments on continuous, steady flows of granular materials down an inclined chute have been made with the objectives of understanding the characteristics of chute flows, and of acquiring information on the rheological behavior of granular materials. Two neighboring fiber-optic displacement probes were used to measure mean velocity, one component of velocity fluctuations, and mean particle spacing. The mean particle spacing also gave qualitative information on density near the boundaries. In addition, a strain-gauged plate was employed to directly measure shear stress at the chute base. The surface of the chute base was carefully controlled to yield three distinct surface conditions; smooth aluminum surface; moderately smooth aluminum surface, rubber-coated surface. Each surface condition was characterized by Coulombic friction coefficient and the coefficient of restitution between the chute base and a particle.

The preliminary experiments indicate that the flow at the free surface is less affected by the sidewalls than at the chute base; the transverse velocity profile at the free surface is close to uniform. It is also observed that the higher the velocity (or the higher the chute inclination), the less significant the sidewall effect.

Vertical profiles of velocity, velocity fluctuation, and linear concentration have been measured through lucite windows in the sidewalls. The velocity profile is fairly linear except for the region within the distance of one particle diameter from the chute base. Velocity fluctuation increases with distance from the chute base. This granular conduction from the bulk of the flow to the chute base wall is opposite to what we observe from the results of Campbell and Brennen (1985b). The results of Ahn et al. (1989) indicate that granular temperature can be conducted in either direction, depending on the value of the particle-particle coefficient of restitution and the chute inclination. In the present measurements, linear concentration always decreases monotonically with distance from the chute base. This result is also different from the results found in the other literature. The surface condition of the chute base plays an important role in the above profiles. The profiles of velocity and its fluctuation with the smooth surface (the surface with low Coulombic friction coefficient) are more uniform than those with the rubber-coated surface (the surface with high Coulombic friction coefficient).

The characteristics of the chute flow of granular materials have been studied by measuring various basic flow properties. The experimental data are strongly affected by the surface condition of the chute base. The ratio of velocity fluctuation to mean velocity is fairly constant for the smooth and moderately smooth surfaces, but for the rubberized surface it clearly increases as the solid fraction decreases. And the ratio for the rubberized surface is much larger than those for the smooth and moderately smooth surfaces. Regardless of the surface conditions, the mean shear rate increases at high solid fraction with decreasing solid fraction. But for the rubberized surface the mean shear rate shows a drastic decrease at low solid fraction. The high ratio of velocity fluctuation to mean velocity causes particles to move from one location to another more frequently, and as a result the velocity gradient is reduced. For the smooth surface where the ratio is low, the decrease of mean shear rate is not observed with decreasing solid fraction.

The variation of friction coefficient with solid fraction is similar to that of the ratio of velocity fluctuation to mean velocity. For the smooth and moderately smooth surfaces, the friction coefficient is fairly constant. But for the rubberized surface, it increases with decreasing solid fraction. As a result, the friction coefficient appears to be a linear function of the ratio of velocity fluctuation to mean velocity.

The stress measurements have also been used to study the rheological behavior of granular material. In particular, the rheological models presented by Lun et al. (1984) have been compared. The rheological models for general flow (equations (2) and (3)) give good correlation to the present experimental data. With the smooth and moderately smooth surfaces, it was not possible to create fully developed flow. But some selected experimental data with the rubberized surface, which are close to fully developed flow, are well correlated with the rheological models for fully developed flow (equations (9) or (10)). Since the chute flows of the present experiments are characterized by granular conduction, the rheological models for simple shear flow (equations (6) and (7)) do not provide good correlation for the present experimental data.

References

Ahn, H., 1989, "Experimental and Analytical Investigations of Granular Materials: Shear Flow and Convective Heat Transfer," Ph.D. Thesis, California Institute of Technology, Pasadena, Calif.

Ahn, H., Brennen, C. E., and Sabersky, R. H., 1988, "Experiments on Chute Flows of Granular Materials," *Micromechanics of Granular Materials*, (M. Satake and J. T. Jenkins, eds.), Elsevier, Amsterdam, pp. 339–348.

Ahn, H., Brennen, C. E., and Sabersky, R. H., 1989, "Analysis of the Fully Developed Chute Flow of Granular Materials," submitted to ASME JOURNAL OF APPLIED MECHANICS.

Augenstein, D. A., and Hogg, R., 1978, "An Experimental Study of the Flow of Dry Powders on Inclined Surfaces," *Powder Technology*, Vol. 19, pp. 205–215.

Bagnold, R. A., 1954, "Experiments on a Gravity-Free Dispersion of Large Solid Particles in a Newtonian Fluid Under Shear," *Proc. R. Soc. Lond.*, Vol. A225, pp. 49–63.

Bailard, J., 1978, "An Experimental Study of Granular-Fluid Flow," Ph.D. Thesis, University of California, San Diego, Calif.

Campbell, C. S., 1982, "Shear Flow of Granular Materials," Ph.D. Thesis, California Institute of Technology, Pasadena, Calif.

Campbell, C. S., 1986, "The Effect of Microstructure Development on the Collisional Stress Tensor in a Granular Flow," *Acta Mechanica*, Vol. 63, pp. 61–72.

Campbell, C. S., 1988, "Boundary Interactions for Two-Dimensional Granular Flows: Asymmetric Stresses and Couple Stresses," *Micromechanics of Granular Materials*, M. Satake and J. T. Jenkins, eds., Elsevier, Amsterdam, pp. 163–174.

Campbell, C. S., 1989, "The Stress Tensor for Simple Shear Flows of a Granular Material," *J. Fluid Mech.*, to be published.

Campbell, C. S., and Brennen, C. E., 1985a, "Computer Simulation of Granular Shear Flows," *J. Fluid Mech.*, Vol. 151, pp. 167–188.

Campbell, C. S., and Brennen, C. E., 1985b, "Chute Flows of Granular Material: Some Computer Simulation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 172–178.

Campbell, C. S., and Gong, A., 1986, "The Stress Tensor in a Two-Dimensional Granular Shear Flow," *J. Fluid Mech.*, Vol. 164, pp. 107–125.

Campbell, C. S., and Gong, A., 1987, "Boundary Conditions for Two-Dimensional Granular Flows," *Proceedings of the International Symposium on Multiphase Flows*, Vol. 1, Hangzhou, China, pp. 278–283.

Craig, K., Buckholz, R. H., and Domoto, G., 1986, "An Experimental Study of the Rapid Flow of Dry Cohesionless Metal Powders," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 935–942.

Hanes, D. M., and Inman, D. L., 1985, "Observations of Rapidly Flowing Granular-Fluid Flow," *J. Fluid Mech.*, Vol. 150, pp. 357–380.

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, pp. 187–202.

Lun, C. K. K., and Savage, S. B., 1986, "The Effects of an Impact Velocity Dependent Coefficient of Restitution on Stresses Developed by Sheared Granular Materials," *Acta Mechanica*, Vol. 63, pp. 15–44.

Lun, C. K. K., Savage, S. B., Jeffrey, D. J., and Chepurniy, N., 1984, "Kinetic Theories for Granular Flow: Inelastic Particles in Couette Flow and Slightly Inelastic Particles in a General Flow field," *J. Fluid Mech.*, Vol. 140, pp. 223–256.

Ogawa, S., Umemura, A., and Oshima, N., 1980, "On the Equations of Fully Fluidized Granular Materials," *J. Appl. Math. Phys.*, (ZAMP), Vol. 31, pp. 483–493.

Patton, J. S., 1985, "Experimental Study of Shear Flows and Convective Heat Transfer Characteristics of Granular Materials," Ph.D. Thesis, California Institute of Technology, Pasadena, Calif.

Patton, J. S., Brennen, C. E., and Sabersky, R. H., 1987, "Shear Flows of Rapidly Flowing Granular Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 801–805.

Savage, S. B., 1979, "Gravity Flow of Cohesionless Granular Materials in Chutes and Channels," *J. Fluid Mech.*, Vol. 92, pp. 53–96.

Savage, S. B., 1984, "The Mechanics of Rapid Granular Flows," *Advances in Applied Mechanics*, Vol. 24, J. Hutchinson and T. Y. Wu, eds., Academic Press, pp. 289–366.

Savage, S. B., and Jeffrey, D. J., 1981, "The Stress Tensor in a Granular Flow at High Shear Rates," *J. Fluid Mech.*, Vol. 110, pp. 255–272.

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 Mech.*, Vol. 127, pp. 453–472.

Savage, S. B., and Sayed, M., 1984, "Stresses Developed by Dry Cohesionless Granular Materials in an Annular Shear Cell," *J. Fluid Mech.*, Vol. 142, pp. 391–430.

Sayed, M., and Savage, S. B., 1983, "Rapid Gravity Flow of Cohesionless Granular Materials Down Inclined Chutes," *J. Appl. Math. Phys.*, Vol. 34, pp. 84–100.

Walton, O. R., 1984, "Computer Simulation of Particulate Flow," *Energy and Tech. Rev.*, Lawrence Livermore Lab., May 1984, p. 24.

Walton, O. R., and Braun, R. L., 1986a, "Viscosity, Granular-Temperature, and Stress Calculations for Shearing Assemblies of Inelastic, Frictional Disks," *J. of Rheology*, Vol. 30, No. 5, pp. 949–980.

Walton, O. R., and Braun, R. L., 1986b, "Stress Calculations for Assemblies of Inelastic Spheres in Uniform Shear," *Acta Mechanica*, Vol. 63, pp. 73–86.

A Note on Rotor Instability Caused by Liquid Motions

O. Holm-Christensen

K. Träger

Department of Solid Mechanics,
The Technical University of Denmark,
Lyngby, Denmark

The stability conditions of a hollow rotor partially filled with a Newtonian liquid are investigated. The rotor is considered here to be a rigid body, supported by springs and dampers, and exposed to an external dynamic force in the shape of actions of the encountered liquid. The system has two degrees-of-freedom, defined by deflection in two mutually orthogonal fixed directions perpendicular to the rotor axis. The fluid motions are described by Navier-Stokes equations and comparison is made between the inviscid and viscous case in connection with their predictions of the stability conditions. Experiments are performed with two different rigidity ratios and results are found to be in agreement with theoretical data.

Introduction

Consider a circular cylindrical container, in which an amount of liquid is trapped. The container is rotating about its own axis at a high angular speed, and in the state of equilibrium the liquid is placed along the wall, thereby describing the shape of a ring. Experiments have shown that such a system at certain speeds becomes unstable, as it starts to vibrate with increasing amplitudes.

This problem has been known for many years and the present work was triggered by an application from a local company that produces rotating machinery. The wish was to achieve a better understanding of the mechanisms that govern the instability phenomena, hereby an examination of the effect of anisotropic rotor supports.

The problem has previously received attention, e.g., by Kollmann (1962), Wolf (1968), and Hendricks and Morton (1979), just to mention a few. So far, theoretical results have been established for a rotor containing a viscous liquid, and with the same rigidity and damping in all directions perpendicular to the rotor axis. A concept called the reduced critical speed (r.c.s.) has been introduced to describe a somewhat amazing phenomenon. For the inviscid, undamped rotor system it can be proved that, at r.c.s., the rotor will behave as if it were completely filled with liquid, no matter how much liquid there is actually present in the chamber. The only condition is that there is enough liquid to cover the walls during the whirl.

In this report, the inviscid as well as the viscous flow theory is extended as to describe rotor motion with two degrees-of-freedom (see Fig. 1). This is done by carrying out the linearization of the nonlinear liquid motion equations by use of two perturbation parameters ϵ_1 and ϵ_2. These quantities, which in the algebra are supposed to be complex, describe the deflection of the rotor from the equilibrium position. By using this mathematically convenient method, we have the possibility of prescribing different rigidity and damping coefficients in different directions perpendicular to the rotor axis.

The problem dealing with the inviscid fluid can almost be solved analytically, only the final characteristic equation requires numerical treatment. The procedure is almost the same as the one that Wolf (1968) used in connection with the rotor with one degree-of-freedom. The main difference is that here it has been found convenient to use a set of auxiliary quantities instead of the perturbed pressure and velocity functions.

If the fluid is viscous, the algebra is much more difficult, and the computer is used at a much earlier stage of the solution of the fluid motion equations. Instead of making boundary

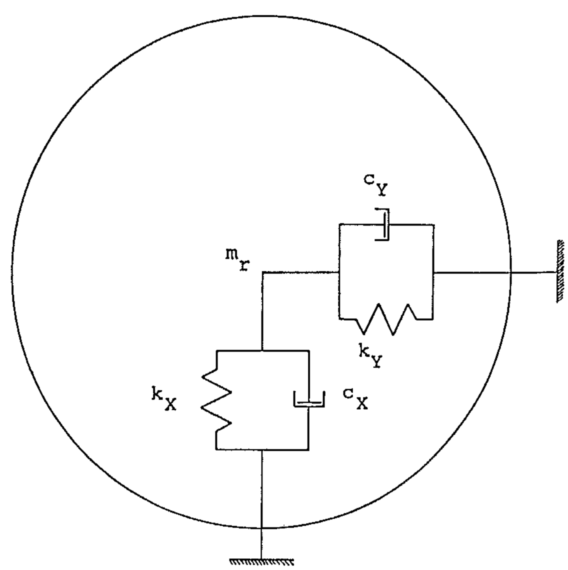


Fig. 1 Sketch of principle, showing the rotor with two degrees-of-freedom

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 1, 1989; final revision, May 8, 1990.

layer approximations as Hendricks and Morton (1979), the results of this report are obtained directly by using the Navier-Stokes equations, including all viscous terms.

In order to make sure that the theory is valid, we have performed some experiments with different rigidity ratios and quite appealing results have been obtained. A brief description of the experimental apparatus is given in the section "Experiments."

General Aspects

Basically, the analytical formulation leads to an eigenvalue problem from which the stability conditions can be deduced. From a stable state, the rotor will be exposed to a deflection, and waves will be produced in the liquid layer. These waves have a net force on the rotor wall depending on the unknown eigenvalue and the rotor speed. This liquid force is incorporated into the equations of motion of the rotor, to give the condition which must be satisfied if whirl is to take place.

In all the analyses the following assumptions are made:

- The rigid liquid chamber has an axisymmetric cylindrical shape and is totally balanced.
- The rotor is driven at constant angular speed Ω.
- The contained liquid is incompressible and surface tension effects are negligible.
- Gravity forces are not present.
- The liquid motion is assumed to be independent of the axial coordinate.
- The whirl motion is described by

$$\{\epsilon(t)\} = \begin{Bmatrix} \epsilon_1 \\ \epsilon_2 \end{Bmatrix} e^{\lambda t} \tag{1}$$

where $\lambda = \alpha + i\omega$, α, and ω are real quantities, and $i^2 = -1$. Using the complex parameters ϵ_1 and ϵ_2 as perturbation parameters, the free surface of the liquid is described by

$$R = b + \eta_1(\phi,t)\cdot\epsilon_1 + \eta_2(\phi,t)\cdot\epsilon_2 \tag{2}$$

where b is the inner radius of the free surface of the liquid for a rigid rotation and η_1 and η_2 are response functions to the disturbance.

Liquid Dynamics: Inviscid Case

By assuming the whirl shape (1), the equations of motion referred to a cylindrical coordinate system fixed to the rotor are:

$$\frac{\partial u}{\partial t} + u\cdot\frac{\partial u}{\partial r} + \frac{v}{r}\cdot\frac{\partial u}{\partial \phi} - \frac{v^2}{r} - r\Omega^2 - 2v\Omega = -\frac{1}{\rho}\frac{\partial p}{\partial r} - \lambda^2 e^{\lambda t}(\epsilon_1\cos(\Omega t+\phi) + \epsilon_2\sin(\Omega t+\phi)) \tag{3a}$$

$$\frac{\partial v}{\partial t} + u\cdot\frac{\partial v}{\partial r} + \frac{v}{r}\cdot\frac{\partial v}{\partial \phi} + \frac{uv}{r} + 2u\Omega = -\frac{1}{\rho r}\frac{\partial p}{\partial \phi} + \lambda^2 e^{\lambda t}(\epsilon_1\sin(\Omega t+\phi) - \epsilon_2\cos(\Omega t+\phi)) \tag{3b}$$

where u and v are the velocities in the r and ϕ-directions, respectively, p is the pressure, and ρ is the liquid density. The last terms on the right-hand side of each equation have been added to account for the acceleration of the origin of the rotor coordinate system produced by the whirling motion. Having an incompressible fluid, the equation of continuity takes the following form:

$$\frac{\partial(ru)}{\partial r} + \frac{\partial v}{\partial \phi} = 0. \tag{4}$$

For the inviscid case, we have the three boundary conditions:

$$u(a)=0;\ p(R)=0;\ u(R)=\frac{\partial R}{\partial t} \tag{5}$$

where "a" denotes the rotor wall and "R" is the earlier defined function for description of the free liquid surface (2). At small deflections $R \simeq b$, leading to the more convenient conditions:

$$u(a)=0;\ p(b)=0;\ u(b)=\frac{\partial R}{\partial t}. \tag{6}$$

The Euler equations (3a)–(3b) are nonlinear and, as a result of this, it is not possible to give a complete solution by any analytical means. To linearize, a perturbation method is employed in which the perturbation parameters are ϵ_1 and ϵ_2, and higher-order terms are neglected

$$\begin{aligned} u &= u_0 + u_1\cdot\epsilon_1 + u_2\cdot\epsilon_2 + \ldots \\ v &= v_0 + v_1\cdot\epsilon_1 + v_2\cdot\epsilon_2 + \ldots \\ p &= p_0 + p_1\cdot\epsilon_1 + p_2\cdot\epsilon_2 + \ldots \end{aligned} \tag{7}$$

Nomenclature

- ϵ_1,ϵ_2 = complex deflections
- λ = complex eigenvalue
- α = parameter of stability
- ω = whirl frequency
- R = position of the disturbed, free surface
- b = position of the undisturbed, free surface
- η_1,η_2 = response functions to the disturbance
- r,ϕ = coordinates in the rotor system of coordinates
- u,v = velocities in the r and ϕ-directions, respectively
- p = pressure
- ρ = liquid density
- Ω = angular velocity of rotor
- a = inner radius of the rotor
- $(\ )_a$ = perturbation functions ($\alpha = 0,1,2$)
- γ,σ = auxiliary quantities
- Γ = quantity to describe the fill ratio
- F_x,F_y = liquid actions in the rotated system of coordinates
- F_X,F_Y = liquid actions in the fixed system of coordinates
- L = length of the rotor
- m_l = mass of liquid needed to completely fill the rotor chamber
- m_r = mass of empty rotor
- c_X,c_Y = external damping coefficients
- k_X,k_Y = rigidities in the main directions
- w = nondimensional eigenvalue
- s = nondimensional rotor speed
- μ = mass ratio
- K = stiffness ratio
- C_X,C_Y = nondimensional external damping coefficients
- ω_x = empty rotor critical speed number 1
- α_{max} = largest real part among the eigenvalues
- ν = kinematic viscosity
- $\sigma_{r\phi}$ = shear stress
- f,g,h,h,l,m = auxiliary quantities
- J_1 = Bessels' function of the first kind, first order
- Y_1 = Bessels' function of the second kind, first order
- A,B = auxiliary quantities to describe liquid reactions
- P = fill ratio
- V = nondimensional kinematic viscosity

The zero-order solutions corresponding to the undisturbed rotor are:

$$u_0=0;\ v_0=0;\ p_0=\frac{1}{2}\rho\Omega^2(r^2-b^2). \tag{8}$$

After perturbation, the momentum equations (3a)–(3b) to be satisfied by the first-order solutions are

$$\frac{\partial u_1}{\partial t}-2v_1\Omega=-\frac{1}{\rho}\bullet\frac{\partial p_1}{\partial r}-\lambda^2e^{\lambda t}\bullet\cos(\Omega t+\phi) \tag{9a}$$

$$\frac{\partial u_2}{\partial t}-2v_2\Omega=-\frac{1}{\rho}\bullet\frac{\partial p_2}{\partial r}-\lambda^2\bullet e^{\lambda t}\bullet\sin(\Omega t+\phi) \tag{9b}$$

$$\frac{\partial v_1}{\partial t}+2u_1\Omega=-\frac{1}{\rho r}\bullet\frac{\partial p_1}{\partial \phi}+\lambda^2\bullet e^{\lambda t}\bullet\sin(\Omega t+\phi) \tag{9c}$$

$$\frac{\partial v_2}{\partial t}+2u_2\Omega=-\frac{1}{\rho r}\bullet\frac{\partial p_2}{\partial \phi}-\lambda^2\bullet e^{\lambda t}\bullet\cos(\Omega t+\phi) \tag{9d}$$

and the corresponding equations of continuity

$$\frac{\partial(ru_1)}{\partial r}+\frac{\partial v_1}{\partial \phi}=0;\ \frac{\partial(ru_2)}{\partial r}+\frac{\partial v_2}{\partial \phi}=0. \tag{10}$$

Boundary conditions:

$$u_1(a)=0;\ u_2(a)=0 \tag{11a}$$

$$p_1(b)=-\rho\Omega^2b\eta_1;\ p_2(b)=-\rho\Omega^2b\eta_2 \tag{11b}$$

$$u_1(b)=\frac{\partial\eta_1}{\partial t};\ u_2(b)=\frac{\partial\eta_2}{\partial t}. \tag{11c}$$

After some rather comprehensive calculations, the pressure at the rotor wall is found to be

$$p(a)=\frac{1}{2}\rho\Omega^2(a^2-b^2)-\frac{1}{2}\rho\lambda^2a(\epsilon_1+i\epsilon_2)\left[\frac{\gamma^2-i2\Omega\gamma+\Omega^2}{\gamma^2\Gamma-i2\Omega\gamma+\Omega^2}\right]\bullet e^{\gamma t-i\phi}$$
$$-\frac{1}{2}\rho\lambda^2a(\epsilon_1-i\epsilon_2)\left[\frac{\sigma^2+i2\Omega\sigma+\Omega^2}{\sigma^2\Gamma+i2\Omega\sigma+\Omega^2}\right]\bullet e^{\sigma t+i\phi} \tag{12}$$

where

$$\gamma=\lambda-i\Omega;\ \sigma=\lambda+i\Omega \tag{13}$$

and

$$\Gamma=\frac{a^2+b^2}{a^2-b^2}. \tag{14}$$

If λ, ϵ_1, ϵ_2, γ, and Γ are replaced by $i\omega$, ϵ, $-i\epsilon$, $i\sigma$, and γ, respectively, the pressure expression (12) will be reduced to the one deduced by Wolf (1968), which is valid if the rotor has only one degree-of-freedom.

When the liquid is considered to be inviscid there are no shear stresses present, and the net force on the rotor wall is found by integrating the pressure on the rotor wall

$$F_x=a\bullet L\int_0^{2\pi}p(a)\bullet\cos\phi d\phi \tag{15a}$$

$$F_y=a\bullet L\int_0^{2\pi}p(a)\bullet\sin\phi d\phi, \tag{15b}$$

where L is the length of the rotor.

Equations (15a)–(15b) give the net force in the x and y-directions, i.e., in the rotating system of coordinates. In order to be able to set up the dynamic equations of equilibrium, we need the net force in the fixed frame of reference X, Y. This is found by rotating F_x and F_y in the following manner:

$$F_X=F_x\cos(\Omega t)-F_y\sin(\Omega t) \tag{16a}$$

$$F_Y=F_x\sin(\Omega t)+F_y\cos(\Omega t) \tag{16b}$$

Employing (12)–(16) finally gives

$$\{\mathbf{F}\}=\begin{Bmatrix}F_X\\F_Y\end{Bmatrix}=-\begin{bmatrix}(A+B); & -i(A-B)\\ i(A-B); & (A+B)\end{bmatrix}\{\epsilon\}=-[F]\{\epsilon\} \tag{17}$$

where

$$A=\frac{1}{2}m_1\lambda^2\left(\frac{\sigma^2+2\Omega i\sigma+\Omega^2}{\sigma^2\Gamma+2\Omega i\sigma+\Omega^2}\right) \tag{18a}$$

$$B=\frac{1}{2}m_1\lambda^2\left(\frac{\gamma^2-2\Omega i\gamma+\Omega^2}{\gamma^2\Gamma-2\Omega i\gamma+\Omega^2}\right) \tag{18b}$$

with m_1 being the mass of liquid needed to completely fill the chamber

$$m_1=\rho\pi a^2L. \tag{19}$$

It is now possible to establish the equations of equilibrium, and referring to Fig. 1, the result takes the well-known form:

$$[M]\{\ddot{\epsilon}\}+[C]\{\dot{\epsilon}\}+[S]\{\epsilon\}=-[F]\{\epsilon\} \tag{20}$$

$$[M]=\begin{bmatrix}m_r; & 0\\0; & m_r\end{bmatrix};\ [C]=\begin{bmatrix}c_X; & 0\\0; & c_Y\end{bmatrix};\ [S]=\begin{bmatrix}k_X; & 0\\0; & k_Y\end{bmatrix}. \tag{21}$$

By inserting (1) into (20), we get a homogeneous system of equations to which nontrivial solutions only exist if the determinant equals zero. After some algebra this determinant conditions leads to the characteristic equation:

$$b_8w^8+b_7w^7+b_6w^6+b_5w^5+b_4w^4+b_3w^3+b_2w^2+b_1w+b_0=0 \tag{22}$$

where

$$\begin{aligned}
b_8&=(\Gamma+\mu)^2\\
b_7&=\Gamma(\Gamma+\mu)(C_X+C_Y)\\
b_6&=\Gamma(K+1)(\mu+\Gamma)+2s^2(2\mu+\Gamma+1)(3\mu+\Gamma+2)\\
&\quad+\Gamma^2C_XC_Y\\
b_5&=\Gamma^2(KC_X+C_Y)+s^2(C_X+C_Y)(2\Gamma^2+5\mu\Gamma+6\Gamma+7\mu+4)\\
b_4&=\Gamma^2K+(1+K)(2\Gamma^2+5\mu\Gamma+6\Gamma+7\mu+4)s^2\\
&\quad+(2\mu+\Gamma+1)^2s^4+2C_XC_Y(\Gamma^2+3\Gamma+2)s^2\\
b_3&=2s^2(\Gamma^2+3\Gamma+2)(KC_X+C_Y)\\
&\quad+(C_X+C_Y)(\Gamma+1)(2\mu+\Gamma+1)s^4\\
b_2&=2Ks^2(\Gamma^2+3\Gamma+2)+(1+K)(\Gamma+1)(2\mu+\Gamma+1)s^4\\
&\quad+C_XC_Y(\Gamma+1)^2s^4\\
b_1&=s^4(\Gamma+1)^2(KC_X+C_Y)\\
b_0&=Ks^4(1+\Gamma)^2.
\end{aligned} \tag{23}$$

The foregoing expressions are shown in a nondimensional form, using the following symbols:

$$w=\frac{\lambda}{\omega_x};\ s=\frac{\Omega}{\omega_x}$$

$$\mu=\frac{m_1}{m_r};\ K=\frac{k_Y}{k_X}$$

$$C_X=\frac{c_X}{m_r\omega_x};\ C_Y=\frac{c_Y}{m_r\omega_x} \tag{24}$$

$$\omega_x^2=\frac{k_X}{m_r}.$$

The stability analysis is now performed by looking at the largest real part $\alpha_{\max}$ occurring among the eight eigenvalues. In agreement with common practice, the system is said to be

stable if $\alpha_{\max}<0$

critical or marginally stable if $\alpha_{\max}=0$

unstable if $\alpha_{\max}>0$.

The solutions to equation (22) can be found in several ways. We have used the numerical procedure of Bairstow. Given a polynomial of the nth degree, this method performs a numerical decomposition resulting in a number of polynomials of the second degree (plus a polynomial of the first degree, if n is odd).

The Bairstow algorithm isn't always applicable as it can be quite sensitive to initial guesses. In the region of parameter space considered in this work, no problems of that kind ever arose. The method was extremely efficient and almost insensitive to the initial guesses.

Liquid Dynamics: Viscous Case

Contrary to earlier investigations of rotor dynamics, including viscous effects in the liquid, we have made no boundary layer approximations. We directly use the full Navier-Stokes equations which here take the form

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial r}+\frac{v}{r}\frac{\partial u}{\partial \phi}-\frac{v^2}{r}-r\Omega^2-2v\Omega= -\frac{1}{\rho}\frac{\partial p}{\partial r}+\nu\left(\frac{\partial^2 u}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \phi^2}+\frac{2}{r}\frac{\partial u}{\partial r}-\frac{1}{r^2}\frac{\partial v}{\partial \phi}\right) -\lambda^2 e^{\lambda t}(\epsilon_1\cos(\Omega t+\phi)+\epsilon_2\sin(\Omega t+\phi)) \quad (25a)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial r}+\frac{v}{r}\frac{\partial v}{\partial \phi}+\frac{uv}{r}+2u\Omega= -\frac{1}{\rho r}\frac{\partial p}{\partial \phi}+\nu\left(\frac{\partial^2 v}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 v}{\partial \phi^2}+\frac{2}{r^2}\frac{\partial u}{\partial \phi}-\frac{v}{r^2}+\frac{1}{r}\frac{\partial v}{\partial r}\right) +\lambda^2 e^{\lambda t}(\epsilon_1\sin(\Omega t+\phi)-\epsilon_2\cos(\Omega t+\phi)) \quad (25b)$$

where ν is the kinematic viscosity, assumed to be constant. Besides the boundary conditions (5), two further restrictions are present when the liquid is viscous. The velocity in the ϕ-direction must be zero at the rotor wall (no slip condition) and the free surface of the liquid must be free from shear stresses, i.e.,

$$\sigma_{r\phi}(b)=\rho\nu\left(\frac{1}{r}\frac{\partial u}{\partial \phi}+\frac{\partial v}{\partial r}-\frac{v}{r}\right)_{r=b}=0 \quad (26)$$

which for $\rho\nu\neq 0$ gives

$$\left(\frac{1}{r}\frac{\partial u}{\partial \phi}+\frac{\partial v}{\partial r}-\frac{v}{r}\right)_{r=b}=0. \quad (27)$$

The complete set of boundary conditions can therefore be written as

$$\begin{gathered} u(a)=0;\ v(a)=0 \\ u(b)=\frac{\partial R}{\partial t}\ ;\ p(b)=0 \\ \frac{1}{b}\frac{\partial u}{\partial \phi}\bigg|_{r=b}+\frac{\partial v}{\partial r}\bigg|_{r=b}-\frac{v(b)}{b}=0. \end{gathered} \quad (28)$$

The expression to describe conservation of mass is the same as in the inviscid case because we still have an incompressible stream

$$\frac{\partial (ru)}{\partial r}+\frac{\partial v}{\partial \phi}=0. \quad (29)$$

Again, we perturb the velocities and the pressure, and by defining a number of auxiliary quantities

$$\begin{gathered} f=u_1+iu_2;\ g=v_1+iv_2;\ l=p_1+ip_2 \\ h=u_1-iu_2;\ k=v_1-iv_2;\ m=p_1-ip_2, \end{gathered} \quad (30)$$

the equations of motion can be written:

$$\frac{\partial f}{\partial t}-2g\Omega= -\frac{1}{\rho}\frac{\partial l}{\partial r}+\nu\left(\frac{\partial^2 f}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 f}{\partial \phi^2}+\frac{2}{r}\frac{\partial f}{\partial r}-\frac{1}{r^2}\frac{\partial g}{\partial \phi}\right)-\lambda^2 e^{\sigma t+i\phi} \quad (31a)$$

$$\frac{\partial g}{\partial t}+2f\Omega= -\frac{1}{\rho r}\frac{\partial l}{\partial \phi}+\nu\left(\frac{\partial^2 g}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 g}{\partial \phi^2}+\frac{1}{r}\frac{\partial g}{\partial r}-\frac{g}{r^2}+\frac{2}{r^2}\frac{\partial f}{\partial \phi}\right)-i\lambda^2 e^{\sigma t+i\phi} \quad (31b)$$

$$\frac{\partial h}{\partial t}-2k\Omega= -\frac{1}{\rho}\frac{\partial m}{\partial r}+\nu\left(\frac{\partial^2 h}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 h}{\partial \phi^2}+\frac{2}{r}\frac{\partial h}{\partial r}-\frac{1}{r^2}\frac{\partial k}{\partial \phi}\right)-\lambda^2 e^{\gamma t-i\phi} \quad (31c)$$

$$\frac{\partial h}{\partial t}+2h\Omega= -\frac{1}{\rho r}\frac{\partial m}{\partial \phi}+\nu\left(\frac{\partial^2 k}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 k}{\partial \phi^2}+\frac{1}{r}\frac{\partial k}{\partial r}-\frac{k}{r^2}+\frac{2}{r^2}\frac{\partial h}{\partial \phi}\right)+i\lambda^2 e^{\gamma t-i\phi}. \quad (31d)$$

Conservation of mass:

$$\frac{\partial (rf)}{\partial r}+\frac{\partial g}{\partial \phi}=0;\ \frac{\partial (rh)}{\partial r}+\frac{\partial k}{\partial \phi}=0. \quad (32)$$

Boundary conditions:

$$\begin{gathered} \tilde f(a)=0\ ;\ \tilde h(a)=0 \\ \tilde l(b)=-\frac{\rho b\Omega^2}{\sigma}\cdot\tilde f(b)\ ;\ \tilde m(b)=-\frac{\rho b\Omega^2}{\gamma}\cdot\tilde h(b) \\ \tilde g(a)=0\ ;\ \tilde k(a)=0 \\ \left(\frac{1}{r}\frac{\partial f}{\partial \phi}+\frac{\partial g}{\partial r}-\frac{g}{r}\right)\bigg|_{r=b}=0\ ;\ \left(\frac{1}{r}\frac{\partial h}{\partial \phi}+\frac{\partial k}{\partial r}-\frac{k}{r}\right)\bigg|_{r=b}=0. \end{gathered} \quad (33)$$

In the expressions it has been used that

$$(f,\ g,\ l)=(\tilde f(r);\ \tilde g(r),\ \tilde l(r))e^{\sigma t+i\phi}\ ;\ (h,\ k,\ m) =(\tilde h(r);\ \tilde k(r);\ \tilde m(r))e^{\gamma t-i\phi}. \quad (34)$$

By applying (32) in connection with (31) and (33), the problem can be converted into two decoupled boundary value problems

$$\nu r^3\cdot\tilde f''''(r)+6\nu r^2\cdot\tilde f'''(r)+(3\nu r-\sigma r^3)\cdot\tilde f''(r) -3(\nu+\sigma r^2)\cdot\tilde f'(r)=0 \quad (35a)$$

$$\nu r^3\cdot\tilde h''''(r)+6\nu r^2\cdot\tilde h'''(r)+(3\nu r-\gamma r^3)\cdot\tilde h''(r) -3(\nu+\gamma r^2)\cdot\tilde h'(r)=0. \quad (35b)$$

Boundary conditions:

$$\begin{gathered} \tilde f(a)=0;\ \tilde f'(a)=0;\ b\tilde f''(b)+\tilde f'(b)=0 \\ \nu b\tilde f'''(b)+4\nu\tilde f''(b)-\sigma b\tilde f'(b)-\frac{\lambda^2}{\sigma}\tilde f(b)-\lambda^2=0 \end{gathered} \quad (36a)$$

$$\begin{gathered} \tilde h(a)=0;\ \tilde h'(a)=0;\ b\tilde h''(b)+\tilde h'(b)=0 \\ \nu b\tilde h'''(b)+4\nu\tilde h''(b)-\gamma b\tilde h'(b)-\frac{\lambda^2}{\gamma}\tilde h(b)-\lambda^2=0. \end{gathered} \quad (36b)$$

Here, "′" means the partial derivative with respect to r. The homogeneous, linear differential equations of fourth order (35) are found to have the following solutions:

$$\tilde f(r)=C_{0f}+\frac{C_{1f}}{r^2}+\frac{C_{2f}}{r}\cdot Y_1\left(i\sqrt{\frac{\sigma}{\nu}}\cdot r\right)+\frac{C_{3f}}{r}\cdot J_1\left(i\sqrt{\frac{\sigma}{\nu}}\cdot r\right) \quad (37a)$$

$$\tilde h(r)=C_{0h}+\frac{C_{1h}}{r^2}+\frac{C_{2h}}{r}\cdot Y_1\left(i\sqrt{\frac{\gamma}{\nu}}\cdot r\right)+\frac{C_{3h}}{r}\cdot J_1\left(i\sqrt{\frac{\gamma}{\nu}}\cdot r\right) \quad (37b)$$

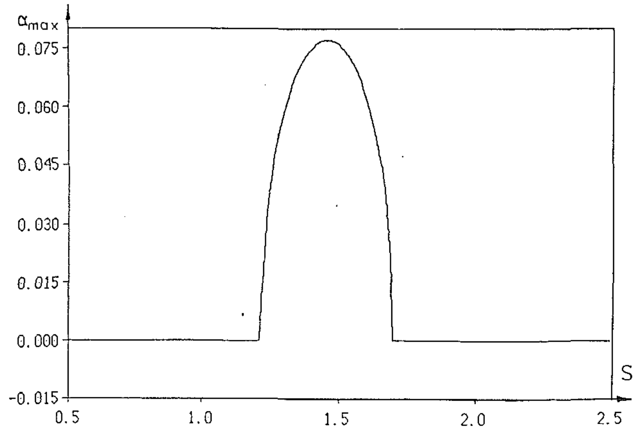


Fig. 2 Parameter of stability versus nondimensional spin speed, one degree-of-freedom, inviscid liquid; $C_X = C_Y = 0$; $K = 1$; $P = 1.5$; $\mu = 0.206$

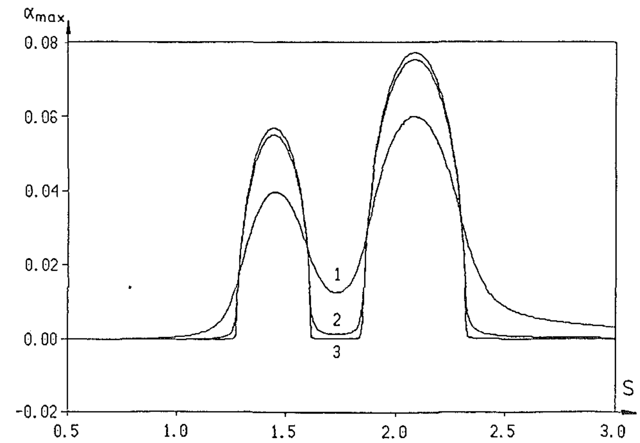


Fig. 4 Parameter of stability versus nondimensional spin speed, two degrees-of-freedom, influence of external damping on rotor containing inviscid liquid; $K = 2$; $P = 1.5$; $\mu = 0.206$; (1) $C_X = C_Y = 0.1$; (2) $C_X = C_Y = 0.01$; (3) $C_X = C_Y = 0.001$

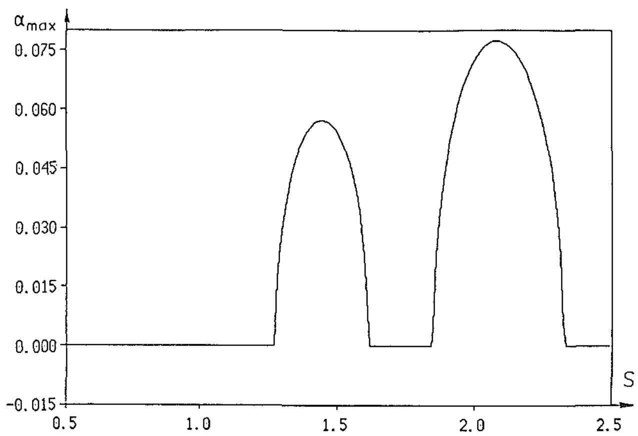


Fig. 3 Parameter of stability versus nondimensional spin speed, two degrees-of-freedom, inviscid liquid; $C_X = C_Y = 0$; $K = 2$; $P = 1.5$; $\mu = 0.206$

in which J_1 is the Bessel function of the first kind, first order and Y_1 is the Bessel function of the second kind, first order. From (37) it appears that evaluation of $\tilde{f}$ and $\tilde{h}$ implies that the Bessel functions have to be calculated for complex arguments. This is rather problematic for arguments with big imaginary parts because the Bessel functions, in these cases, have extreme gradients, and numerical treatment is difficult. Physically, f and h reflect velocity fluctuations, which are supposed to behave properly, and this means that the Bessel terms in some sense must be in mutual balance.

The conclusion of all this is that a numerical treatment is easier to handle if it is brought into action at an earlier stage, that is, at the solution of the boundary value problems (35)-(36). By using a standard finite difference procedure this is a practicable task, especially because the problems are one-dimensional when treated separately.

In the following, we assume that $\tilde{f}$ and $\tilde{h}$ and their derivatives are known quantities. By carrying out the same steps as in the case of the inviscid liquid, and remembering that the net force on the rotor wall now depends on both pressure and shear stress in the following form:

$$F_x = aL \int_0^{2\pi} [p(a)\cos(\phi) + \sigma_{r\phi}(a)\sin(\phi)]d\phi \qquad (38a)$$

$$F_y = aL \int_0^{2\pi} [p(a)\sin(\phi) - \sigma_{r\phi}(a)\cos(\phi)]d\phi, \qquad (38b)$$

the final expression of equilibrium turns out to be

$$[M]\{\ddot{\epsilon}\} + [C]\{\dot{\epsilon}\} + [S]\{\epsilon\} = -[F]\{\epsilon\} \qquad (39)$$

with

$$[M] = \begin{bmatrix} m_r\ ; & 0 \\ 0\ ; & m_r \end{bmatrix};\ [C] = \begin{bmatrix} c_X\ ; & 0 \\ 0\ ; & c_Y \end{bmatrix};\ [S] = \begin{bmatrix} k_X\ ; & 0 \\ 0\ ; & k_Y \end{bmatrix} \qquad (40)$$

$$[F] = -\begin{bmatrix} A+B\ ; & -i(A-B) \\ i(A-B)\ ; & A+B \end{bmatrix}.$$

In order to recreate the same appearance of the system of equations as in the inviscid case, we have defined:

$$A = \frac{1}{2} m_1(\lambda^2 - \nu a \tilde{f}'''(a) - 3\nu \tilde{f}''(a)) \qquad (42a)$$

$$B = \frac{1}{2} m_1(\lambda^2 - \nu a \tilde{h}'''(a) - 3\nu \tilde{h}''(a)). \qquad (42b)$$

The solution procedure to this problem is[1]:

1 A guess is made for the eigenvalue λ.

2 The boundary value problems (35)-(36) are solved, so that $\tilde{f}''(a)$, $\tilde{f}'''(a)$, $\tilde{h}''(a)$, and $\tilde{h}'''(a)$ are known quantities.

3 A and B are calculated.

4 The characteristic equation now becomes of the fourth degree because A and B are given as constants. The four solutions, λ_i, are found numerically.

5 The eigenvalue λ_m with the biggest real part is compared with λ from 1. If the quantities are equal, the problem has converged, and λ_m is the determining eigenvalue from which the stability conditions are deduced. If λ differs from λ_m, then λ_m is used as a new guess at 1 and the procedure is repeated.

The procedure presented above is rather time-consuming and sensitive to the initial guess.

Results

In our system we have seven basic parameters to control the motion of the system:

$s = \Omega/\omega_X$	is the non-dimensional spin speed.
$P = a/b$	is a quantity that describes the amount of liquid present in the rotor. $P = 1$ corresponds to an empty rotor and $P = \infty$ corresponds to a rotor completely filled with liquid.
$\mu = \pi\rho a^2 L/m_r$	is a mass ratio equal to the mass of liquid

[1]The physical meaning of this iteration procedure is to find the values of λ by which the forced vibrations (caused by the liquid motions) correspond to the free vibrations of the rotor.

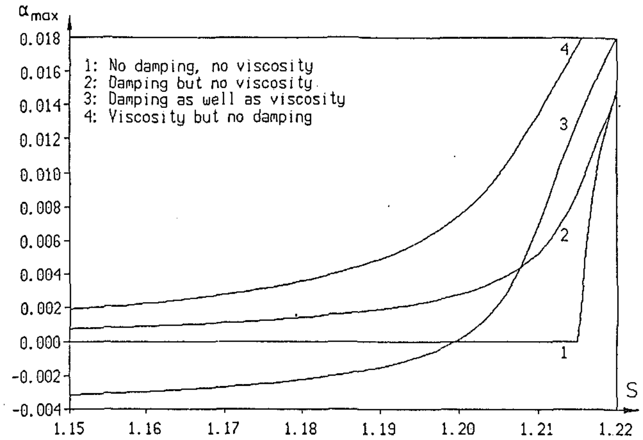


Fig. 5 Parameter of stability versus nondimensional spin speed, interaction of damping and viscosity; $K=1.0$; $P=1.5$; $\mu=0.206$

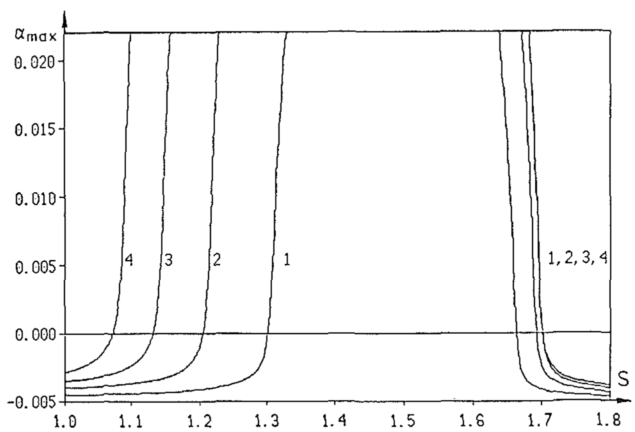


Fig. 6 Parameter of stability versus nondimensional spin speed, influence of mass ratio on stability conditions; $K=1.0$; $P=1.5$; $V=0.0004$, $C_X=C_Y=0.01$; (1) $\mu=0.1$; (2) $\mu=0.2$; (3) $\mu=0.3$; (4) $\mu=0.4$

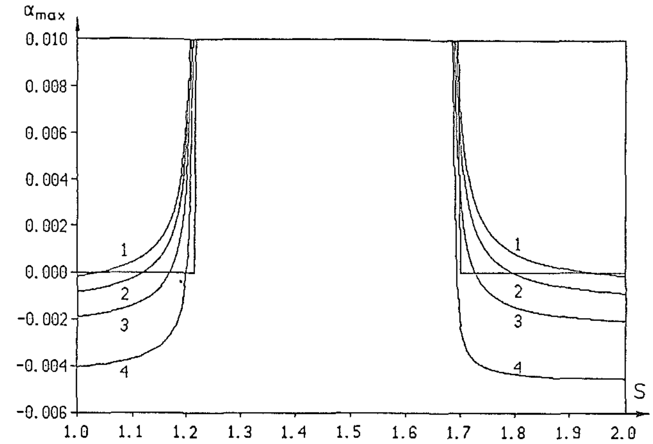


Fig. 7 Parameter of stability versus nondimensional spin speed, influence of damping on stability conditions; $K=1.0$; $P=1.5$; $\mu=0.206$, $V=0.0004$; $C_X=C_Y$; (1) $C_X=C_Y=0.0010$; (2) $C_X=C_Y=0.0025$; $C_X=C_Y=0.0050$; (4) $C_X=C_Y=0.0100$

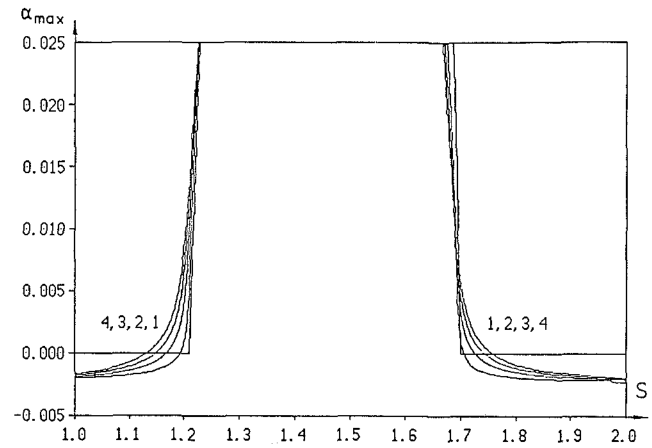


Fig. 8 Parameter of stability versus nondimensional spin speed, influence of viscosity on stability conditions. $K=1.0$; $P=1.5$; $\mu=0.206$; $C_X=C_Y=0.005$; (1) $V=0.0002$; (2) $V=0.0004$; (3) $V=0.0006$; (4) $V=0.0008$

which would completely fill the rotor, divided by the mass of the rotor.

$C_X=\dfrac{c_X}{m_r\omega_x}$, $C_Y=\dfrac{c_Y}{m_r\omega_x}$ are the non-dimensional damping constants, which characterize the damping due to the external dampers in the two main directions.

$V=\nu/a^2\omega_X$ is a non-dimensional viscosity.

$K=\dfrac{k_Y}{k_X}$ is the ratio between the rigidities in the two main directions.

Of course, V does not occur in the inviscid theory. In the graphs, the parameter $\alpha_{\max}$ is depicted as a function of the nondimensional spin speed s in order to get an overall understanding of the stability concept.

Figure 2 shows the behavior of the rotor in the special case, where $k_Y=k_X$, $c_X=c_Y=0$, $V=0$, i.e., a simulation of the rotor with one degree-of-freedom. We observe an unstable frequency band represented by a "bulge" on the $\alpha_{\max}$-curve. From Fig. 3 we discover that if K differs from 1, i.e., if the rigidities in the two main directions are not equal, we get two bulges. Physically, the first interval of instability corresponds to vibration in the direction with low rigidity, and the second interval of instability corresponds to vibration mainly in the direction of high rigidity. In Figs. 2 and 3, the external damping equals zero; but what happens to the predictions of the theory if this is not the case? The answer is given in Fig. 4 on which the $\alpha_{\max}$-curves are drawn at three different degrees of damping, showing the local behavior near the lower end of the bulge.

A quite peculiar result appears. The system turns out to be unstable at any rate of rotation! This is not a totally unknown phenomenon, as Hendricks and Morton (1979) reached the same result for a rotor with only one degree-of-freedom[2]. Their Their explanation is that the external damping introduces a force on the fluid which is out of phase with the acceleration and displacement of the rotor, and that an inviscid fluid has no mechanism for countering this force. If we are to get proper predictions of behavior of the rotor with external damping, it is necessary to take viscosity into consideration. A comparison of theories with and without account of viscous effects is shown in Fig. 5. From this it will be seen that if damping and viscosity occur separately, they will cause the system to be unstable at any rate of rotation, but if they occur together, real stability will be possible and the limits of instability will differ very little from the inviscid, undamped system.

In the light of the viscous theory it is possible to carry out a parameter analysis that examines the effects of the individual parameters in connection with stability conditions. Referring to Figs. 6–8, where only one quantity is changed while the others are kept constant, the following conclusions are made:

[2]Actually, they have used two degrees-of-freedom in their formulation, but by letting the rigidity and the damping be the same in all directions perpendicular to the rotor axis, they reduce their system, so that it is actually treated as having only one degree-of-freedom.

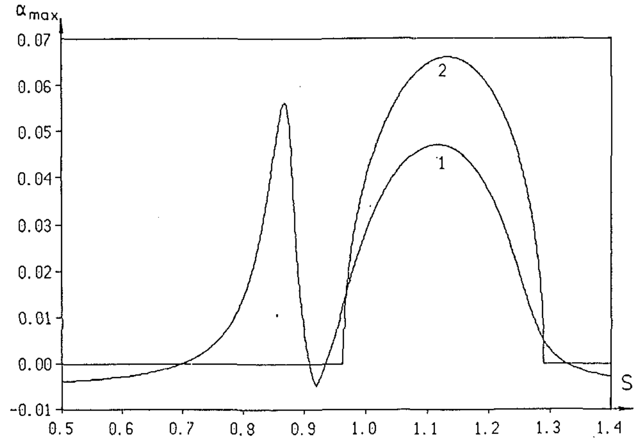


Fig. 9 Parameter of stability versus nondimensional spin speed, illustration of the instability—"top" appearing in the viscous case; $K=1.0$; $P=1.02$; $\mu=0.175$; (1) $V=10^{-5}$, $C_X=C_Y=0.10$; (2) $V=0$, $C_X=C_Y=0$

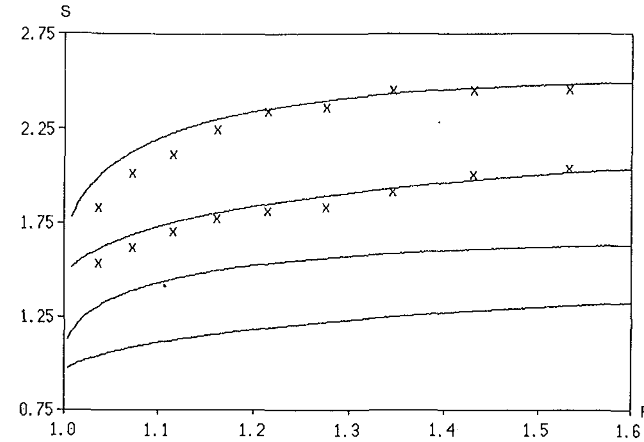


Fig. 11 Nondimensional spin speed versus fill ratio, instability areas determined theoretically (———) and experimentally (x); $K=2.29$; $\mu=0.175$

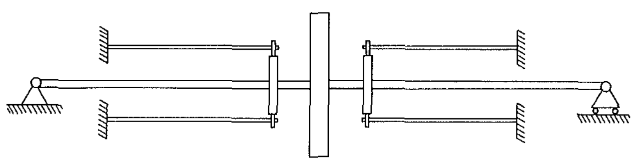


Fig. 10 Sketch of principle showing the spring system. The gravity force is parallel to the rotor axis.

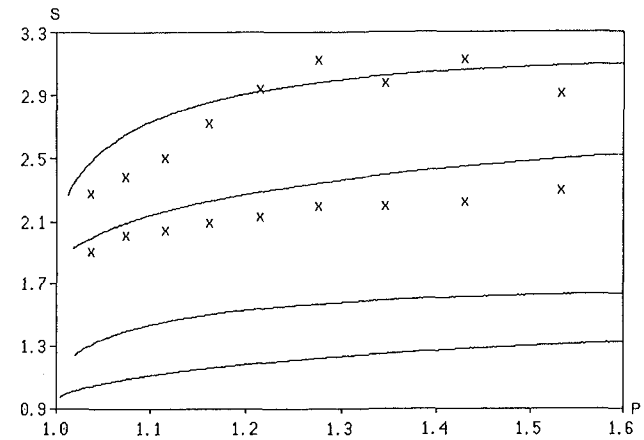


Fig. 12 Nondimensional spin speed versus fill ratio, instability areas determined theoretically (———) and experimentally (x); $K=3.55$; $\mu=0.175$

Fig. 6—an increase in the mass ratio caused the instability area to be wider, and especially the lower limit is seen to move considerably,

Fig. 7—an increase in the external damping tends to make the system more stable by narrowing the area of instability and making α_{max} more negative at stable rates of rotation, and

Fig. 8—an increase of the viscosity tends to have a destabilizing effect, but it appears that the dependence is rather limited.

As a matter of fact, the results of Figs. 7 and 8 are strongly contrary to earlier work by Hendricks and Morton (1979). They claim that the dependence on damping of the stability conditions is enormous. For some fill ratios, the upper limit of the unstable area seems almost to disappear. They have likewise found the viscosity to have a great effect, although not as noticeable as the damping. Apart from that they conclude that the damping has a destabilizing effect and the viscosity the opposite.

It is appropriate to point out some of the differences between the analysis performed by us and that by Hendricks and Morton (1979). As mentioned previously, we have linearized the Navier-Stokes equations, including viscous terms, making no assumptions whatsoever about boundary layers. Hendricks and Morton have separated their equations into three terms. First, they have an inviscid core based upon the classical inviscid theory, then they have a boundary layer, and finally a correction term to the inviscid core. They claim that the error connected with these boundary layer considerations is sufficiently small to determine stability conditions properly. The different results are therefore assumed to be caused by the fact that we have been dealing with different regions of the parameter space.

This is not the only difference. In Fig. 9 we discover a further instability top, and this phenomenon has never been found by a theory dealing with viscosity before. That the top is actually there we discovered when we performed our experiments. It might turn out to be a better explanation to the phenomenon previously named "the reduced critical speed," as the existence of that can only be deduced for the inviscid, undamped rotor model with one degree-of-freedom. Experiments show clearly that such "tops" are present when dealing with rotors with two degrees-of-freedom, and there actually seems to be a top connected to every instability bulge.

The weakness of our procedure is that the numerical treatment is rather sensitive to the parameters used, and for some reason or other, our program has failed to find the instability bulge connected to vibration in the main direction with high rigidity. This is presumed to be caused by the iteration procedure in which we actually have no guarantee that we find the worst eigenvalue, i.e., the eigenvalue with the biggest real part. However, one thing is certain: If the program predicts instability, then there will be instability because at least one eigenvalue with positive real part has been found by convergence.

Experiments

Experiments are performed with water in the container, and our model can be described schematically as shown in Fig. 10. The basis is a flexible shaft with a circular cross-section that provides equal rigidity against deflection in all directions perpendicular to the main axis. To this shaft are attached four small cantilever beams so that rigidity against deflection in the plane of the paper is increased, while rigidity against deflection out of the plane of the paper is unchanged. A motor is placed beneath the model in such a way that it can be removed completely, so that no disturbance appears from the motor when the measurements take place. A guard bearing is placed on each side of the fluid chamber to make sure that the deflections are limited. This is done partly to prevent the apparatus from being damaged, and partly to ensure that the motor is able to accelerate the rotor through unstable speeds. The procedure for carrying out the measurements can now be listed in the following manner:

1 A measured amount of water is poured into the container while it is at rest.

2 The rotor is driven through the unstable rotation speeds, and when rotating supercritically, the motor is removed.

3 The stability limits are determined when the rotor settles down as the speeds of rotation are registered when the transition from stability to instability (or the opposite) is observed.

4 Points 2 and 3 are repeated, and if the measured results are in agreement with each other, point 1 is carried out with a new amount of liquid. Otherwise, points 2 and 3 are repeated until satisfying results are obtained.

Two series of measurements were carried out with two different rigidity ratios, $k=2.29$ in Fig. 11 and $k=3.55$ in Fig. 12. Using water, we got a mass ratio $\mu=0.175$, and the empty rotor critical speed was 49 rad/s for vibration in the weak direction. The small crosses indicate measurements and the drawn curves are the predictions based upon the inviscid, undamped theory with the current rigidity ratio. As it will be seen from the figures we have, in our experiments, only been able to locate the upper area of instability. This is due to the difficulty of separating the "tops" from the "bulges" as mentioned before in connection with Fig. 9.

Our purpose with the experiments has not been to perform precision measurements. We wanted to obtain a qualitative verification of our theory for the rotor with two degrees-of-freedom. It is our conviction that better results can be obtained by use of precision equipment and more accurate control of the rotor speed, so that the "tops" can be separated from the "bulges."

Conclusion

We have found that the inviscid fluid formulation is insufficient when dealing with external damping. By using the viscous liquid theory a comprehensive parameter analysis is carried out, so that dependence on damping, viscosity, and mass ratio is determined in connection with the stability conditions. A new sort of instability region has been found, which is greatly reminiscent of the well-known reduced critical speed, which can only be deduced for the inviscid, undamped rotor system with one degree-of-freedom.

Finally, the theoretical results are qualitatively verified by comparison with experimental measurements.

References

Hendricks, S. L., and Morton, J. B., 1979, "Stability of a Rotor Partially Filled with a Viscous Incompressible Fluid," ASME Journal of Applied Mechanics, Vol. 46, pp. 913–918.

Kollmann, F. G., 1962, "Experimentelle und theoretische Untersuchungen über die kritischen Drehzahlen flüssigkeitsgefüllter Hohlkörper," *Forchung auf dem Gebiete des Ingenieurwessens*, Ausgabe B, Vol. 28, pp. 115–123, 147–153.

Wolf, J. A., Jr., 1968, "Whirl Dynamics of a Rotor Partially Filled with Liquid," ASME Journal of Applied Mechanics, Vol. 35, pp. 676–682.

Free Fall of a Sphere in a Partially Lubricated Cylinder

Roger F. Gans
Professor,
Department of Mechanical Engineering,
University of Rochester,
Rochester, NY 14627
Mem. ASME

The entrainment of lubricant at the entrance of a lubrication zone, such as that of a partially starved slider bearing, is analyzed in a closed system using the method of matched asymptotic expansions. A sphere falling together with a small lens of lubricant in a closely fitting tube is shown to fall under gravity at a speed

$$V = (Mg - F_c)\sqrt{[(R_C - R_S)/R_C]}/(16\pi^2\mu R_C),$$

where M *denotes the total mass of the system, sphere plus lubricant,* g *the acceleration of gravity,* F_c *the differential contact force,* μ *the viscosity of the lubricant, and* R_C *and* R_S *the radii of the tube and the sphere, respectively. Potential biological applications and experimental verification are discussed.*

1 Introduction

This paper introduces a model problem for hydrodynamic lubrication between sliding surfaces with a small amount of lubricant: the generic starved bearing. The model is that of a sphere moving through a closely fitting cylinder. The geometry is closed. The system is axisymmetric rather than planar two-dimensional. The small additional complication is more than compensated for by making the model realizable in the laboratory. This paper is analytic. The outline of an experimental procedure is given at the end. Tests in a simple apparatus are not inconsistent with the analysis presented below.

I consider the fall under gravity of a sphere of radius R_S inside a cylinder of radius R_C in the presence of a small amount of Newtonian lubricant. Both the sphere and its enclosing tube are assumed to be rigid, and steady solutions are sought. Figure 1 shows a static solution, and can serve as a defining sketch. In the sketch r,z denote the usual cylindrical coordinates, the system is assumed to be symmetric about the z-axis, and θ denotes the polar angle. The remaining notation will be defined below as needed.

Static solutions are only possible if the contact angles β_1 and β_2 where the upper and lower liquid lenses intersect the tube are unequal. This is discussed in Section 2. The body of the paper is Section 3, in which the problem given is solved using the method of matched asymptotic expansions. The fundamental small parameter is the square root of twice the ratio between the gap and the cylinder radius

$$\epsilon = \sqrt{\{2[R_C - R_S]/R_C\}}. \tag{1}$$

(The apparently odd factor of the square root of two is chosen to make the spherical boundary in the lubrication region tidy.) I will define a Reynolds number and a capillary number. The former must be proportional to ϵ and the latter to ϵ^3. Account is taken of the force that can arise from contact angle hysteresis. The fundamental result is that the speed of fall V is given by

$$V = \{M_T g - 2\pi\gamma R_C(\cos\beta_1 - \cos\beta_2)\}\epsilon/\{16\sqrt{2}\pi^2\mu R_C\} \tag{2}$$

where M_T denotes the mass of the sphere and its entrained liquid, g the acceleration of gravity, and γ and μ the surface tension and the viscosity of the liquid, respectively. The angles β_1 and β_2 denote the contact angles at the upper and lower boundary between the liquid and the cylinder.

The immediate motivation is to begin to understand the physics of passive swallowing, the process by which a mass of

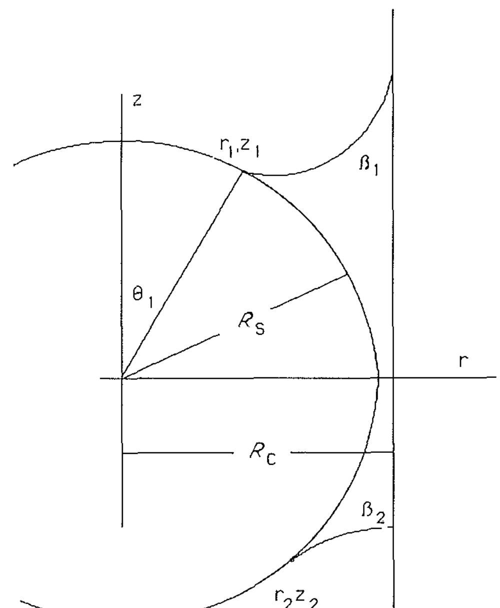


Fig. 1 Definition sketch for the sphere in the cylinder. The state corresponds to that in the fifth line of Table 1.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Nov. 9, 1989; final revision, Mar. 13, 1990.

chewed food, a bolus, passes down the throat. The esophogeal environment is mostly air, with a layer of liquid, mostly saliva, coating the throat, and permeating the bolus. Swallowing is active, though as anyone who has had a bit of food stuck in the throat can testify, not terribly so. It involves a flexible, active tube and a compliant bolus. The liquid is non-Newtonian. This complex system is well beyond simple analytic modeling. The present work is a first attempt to model that system, and it is not entirely satisfactory. The range of validity of the model is not well matched by the physiological parameters.

The model is well suited to understanding entrainment processes for which a free surface and gravity are important. These processes are not well understood. The present model offers the first closed-system model of this phenomenon, a system not requiring simplifying assumptions about two-dimensionality. As such it is an improvement over previous work from the lubrication community (Bonneau and Frene, 1983; Gans and Wang, 1989; Tichy, 1986; Tichy and Bourgin, 1985; Tipei, 1978) and the blade coating community (Campanella and Cerro, 1984; Cerro and Scriven, 1980; Hsu et al., 1985; Sullivan and Middleman, 1986; Sullivan, Middleman and Keunings, 1987.)

The plan of the paper is as follows: In Section 2 I demonstrate the existence of static solutions if there is contact angle hysteresis. In Section 3 I discuss the asymptotic expansion and derive equation (2). Finally, in Section 4, I discuss the oral biology application, and describe a simple order of magnitude experiment.

2 On Static Solutions

In this section the conditions for static equilibrium are found. For equilibrium the net force on the sphere must be zero, and the net force on the sphere-liquid system must be zero. The liquid must also be in hydrostatic equilibrium. Consider first the system. Forces on the system include gravity, possible adhesion between the sphere and the cylinder in the region of close approach, and surface tension forces. Symmetry, with respect to the sphere's equator, shows that adhesion forces cannot balance gravity. Therefore, the contact lines must support the weight of the system. This requires a difference between the advancing and receding contact angles, and a sufficiently large surface tension.

The (dimensionless) contact force on the system is given by

$$F_C = 2\pi(\cos\beta_1 - \cos\beta_2), \qquad (3)$$

where γR_C has been taken as the force scale, γ denotes the surface tension, β the contact angle, and 1 and 2 the top and bottom surfaces. Pressure will be scaled by γ/R_C and lengths by R_C, and the discussion will be conducted in dimensionless units. In these units the radius of the cylinder is unity and that of the sphere is a. Dimensionless cylindrical coordinates r (radial) and z (axial) will be introduced. For the contact force to balance gravity, the top contact angle must be less than the bottom contact angle. Imagine β_2 to be an advancing contact angle and β_1 a receding contact angle, as in the analogous case discussed by Dussan V. and Chow (1983; see also Dussan V., 1985, 1987). The question of advancing and receding contact angles, and their difference, if any, is an area of active research. When a difference is measured between these two angles, the advancing is larger than the receding. The difference can be large.

To complete the global equilibrium, the contact force is equated to that of gravity:

$$M_T g/\gamma R_C = \mathrm{Bo}\{V_L + (4/3)\pi\rho_S a^3\} = \mathrm{Bo}M = 2\pi(\cos\beta_1 - \cos\beta_2), \qquad (4)$$

where M_T denotes the total mass of the system, ρ_S the relative density of the sphere, V_L the liquid volume, M a dimensionless mass (or volume), and Bo a Bond number defined by

$$\mathrm{Bo} = \rho g R_C^2/\gamma.$$

The right-hand side of (4) cannot exceed 4π, therefore Bo cannot grow without bound; some minimum surface tension is require to hold the system in place.

The pressure in the liquid must be continuous, increasing downward at a rate Bo. The pressure at 2 (where $r = r_2$ and $z = z_2$) is $\mathrm{Bo}\Delta z$ greater than that at 1 (where $r = r_1$ and $z = z_1$; see Fig. 1). These pressures are determined by the curvature at the respective points; therefore, the curvatures at the points of intersection are related. Let the upper surface be given by $z = Z_T(r)$ and the lower by $z = Z_B(r)$. Denote the mean curvature by H. Then $2H = \mathbf{div}(\mathbf{T}Z)$, where

$$\mathbf{T}Z = \mathbf{e_r} Z'/\sqrt{(1 + Z'^2)}, \qquad (5)$$

$\mathbf{e_r}$ denotes the radial unit vector and prime the derivative with respect to argument. (This formulation is due to Concus and Finn (1974a,b); see also the monograph by Finn (1986). Many of the mathematical manipulations in this section are straightforward extensions of material found in these sources.)

The equations of the surfaces are

$$\mathbf{div}(\mathbf{T}Z_1) = \mathrm{Bo}Z_1 + P_1,$$
$$\mathbf{div}(\mathbf{T}Z_2) = -\mathrm{Bo}Z_2 + P_2, \qquad (6)$$

where the P are constants. Since Z measures the height of the liquid surface above the reference level, and the pressure in a static liquid is given by

$$p = p_a + \mathrm{Bo}z - \mathbf{div}(\mathbf{T}Z), \qquad (7)$$

where p_a denotes the scaled atmospheric pressure, it can be seen P denotes the difference between atmospheric pressure and the pressure at the reference level. In the event that the liquid lies above the surface, Z is negative, and the discontinuity in pressure caused by the curvature of the surface has the opposite sign. The net effect is that the curvature term in equation (7) has the opposite sign, as written in the second of equations (6). This establishes that $P_1 = -P_2$ for continuity of pressure.

Further conditions on the pressure constants can be established by integrating equations (6) over the projected surface area lying between r_i (the intersection point in each case) and unity. This gives

$$2\pi(\pm\cos\beta_i - r_1\sin\phi_i) = \mathrm{Bo}V_i^* + P_i A_i^*, \qquad (8)$$

where $A_i^* = \pi(1 - r_i^2)$ denotes the projected area of each meniscus and V_i^* the (positive) cylindrical annulus of volume between the meniscus and the midplane of the sphere. The upper (lower) sign is taken in the upper (lower) surface. The angle ϕ_i denotes the angle between the surface and the horizontal at the intersection (see Fig. 1.) Denoting the polar angle θ at each intersection by $\theta_i = \sin^{-1}(r_i/a)$ allows me to write an equation relating the various angles:

$$\phi_{i1} = -\theta_1 + \beta_1';$$
$$\phi_{i2} = \theta_2 - \beta_2'. \qquad (9)$$

The primed contact angles need not be identical to the unprimed contact angles.

For sufficiently large Bo, the left-hand side of (8) is negligible and both P_1 and P_2 will be negative. Their sum cannot vanish, and hydrostatic equilibrium is impossible. Dividing the integrated equations by the projected area and rearranging gives

$$P_1 = P = 2\langle H_1\rangle - \mathrm{Bo}\langle Z_1\rangle,$$
$$P_2 = -P = 2\langle H_2\rangle + \mathrm{Bo}\langle Z_2\rangle, \qquad (10)$$

where the angle brackets denote the mean over the projected surface in each case, and $\langle Z_2\rangle < 0$. Adding these two gives

$$P_1 + P_2 = P - P = 0 = 2[\langle H_1\rangle + \langle H_2\rangle] - \mathrm{Bo}\langle D\rangle, \qquad (11)$$

where $\langle D\rangle$ denotes the mean distance between the two surfaces.

Equations (10) show that $\langle H_1 \rangle$ and $\langle H_2 \rangle$ must be of opposite sign, and (11) that the positive one must be the larger.

It is also necessary that the sphere be in equilibrium. This is assured if the vertical component of the pressure force added to the contact line force balances gravity. The former is the integral over the surface of $\cos\theta$ times the gage pressure, and the latter the line integral of the surface tension. The gage pressure, under conditions of hydrostatic equilibrium, is just

$$p - p_a = -(P + \mathrm{Bo}\ a \cos\theta), \tag{12}$$

where P is the numerical value of $P_1 = -P_2$. Integrating equation (12) then gives

$$\rho_S a^3 \mathrm{Bo} = -[3P(z_2^2 - z_1^2)/4 + \mathrm{Bo}(z_2^3 - z_1^3)/2] + (3/2)(r_1 \sin\phi_1 + r_2 \sin\phi_2). \tag{13}$$

How does this work? In any given case the Bond number and contact angles can be supposed given, as can the mass and size of the sphere. Equation (4) gives a first cut at deciding whether a static solution is possible. If so, one can seek a family of solutions to (6) such that hydrostatic equilibrium is maintained, and then search through that family for candidates satisfying the two equilibrium conditions, equations (4) and (13). Thus, it is necessary to solve equations (6). Fortunately, that is relatively straightforward numerically.

Consider the equation

$$\mathbf{div}(\mathbf{T}Z) = \mathrm{Bo}Z + P \tag{14}$$

subject to boundary conditions on slope at r equal to r_i and unity, and a boundary condition that $Z = z_i$ at $r = r_i$. The equation is second order in r, so the third condition can only be satisfied for specific value(s) of P; this is an eigenvalue problem. Solutions have been obtained by creating an artifical third-order system with the additional equation $P' = 0$. The third-order system is solved implicity after reducing to a set of three first-order equations:

$$Z' = \tan\phi$$
$$(\sin\phi)' + \sin\phi/r = \mathrm{Bo}Z + P$$
$$P' = 0 \tag{15}$$

using routines found in Press et al. (1986). Once separate solutions are obtained for top and bottom they are matched to the same absolute value of P, determining sets of inner radii satisfying local and global hydrostatic equilibrium.

As an example, consider a fully wetted sphere of radius 0.95. "Fully wetted" means that the contact angle at the inner contact lines is zero. Let the upper contact angle be 10 deg and the lower be 90 deg. Table 1 shows a set of possible equilibrium configurations for a Bond number of unity, for which the (scaled) force on the liquid must be equal to the (scaled) liquid volume. The total force on the system is equal to $2\pi\cos(10\ \mathrm{deg}) = 6.1877$. The last column shows the force on the sphere, increasing as the amount of liquid decreases. One can choose sphere densities to match any of these configurations. Figure 1 has been drawn from the fifth line in the table, for which the relative sphere density is 1.3048.

3 Dynamic Solutions

3.1 Introduction. Once the equilibrium described in the previous section is no longer possible, motion sets in. It is the purpose of this section to analyze this motion, and to find steady-state solutions for motion. It will be convenient to analyze the dynamic cases from a reference frame attached to the moving sphere, and to suppose that the cylinder moves upward at some velocity. The contact lines will be in uniform motion and therefore force-free. The expression for contact line force given in Section 2 above holds. The balance of forces on the system will be the difference between the action of gravity on the combined system of sphere and liquid and the contact line forces, and the shear forces developed between the liquid and the wall:

$$\text{gravity force} = \text{contact line support} + \text{viscous drag}.$$

This will determine the velocity of fall. This analysis is completely different in approach from that of Dussan V and coworkers (Dussan V. and Chow, 1983; Dussan V., 1985, 1987.) The essential difference is that the presence of the sphere introduces a lubrication region between the rigid surfaces absent in the earlier work. These results are complementary to her work. The analysis is also different from the usual lubrication analysis in that the free surfaces in the entrance and exit regions are explicitly included. No ad hoc assumptions are needed.

Motion generates lubrication pressures between the sphere and the wall. If the motion is sufficiently slow, this lubrication pressure will be governed by a laminar Reynolds equation. That condition will be assumed. This lubrication pressure is an intermediary between the two trapped liquid volumes, and hydrostatic equilibrium is no longer necessary. Far from the sphere, two situations are possible: either the liquid has settled into a simple layer of uniform thickness (the infinite liquid, perfect wetting, case) or there is no liquid (the finite liquid case). These two cases are quite different. In the former, the effective top and bottom contact angles will be zero, there will be no net contact line force, and there will be a net flux of liquid between the sphere and the cylinder. The latter has contact line forces, in general, and no net flux. To the order to which the problem is solved, details of the motion of the contact lines are not important, nor is the distinction between infinite volume cases. The latter is characterized by zero contact angles. The effects of flux do not appear until higher order in the expansion parameter.

The flow field can be divided into two regions: an inner (lubrication) region, in which Reynolds equation provides an adequate approximation to the flow, and an outer (capillary) region in which the flow has a negligible effect on the pressure distribution. The inner region is characterized by contrasting length scales in the streamwise and normal directions, allowing simple analysis. The outer region does not. In this region one should analyze the full (axisymmetric) three-dimensional flow under the (unknown) free boundary. This is a formidable task, though not necessarily impossible. It is, however, beyond the scope of this paper. In this paper the outer region will be analyzed by neglecting the liquid motions, allowing the use of the analysis outlined in the previous section. This is in the spirit of Levich's (1962) analysis of the withdrawal of a vertical plate from a pool of liquid. The analysis that follows employs a matched asymptotic expansion to make this casual discussion rigorous.

Table 1 Set of possible pairs in hydrostatic equilibrium

rtop	rbottom	P(top)	upper vol	lower vol	total vol	F liquid	F sphere
0.250	0.393	1.169	1.5444	0.7279	2.2723	2.2723	3.9154
0.305	0.481	1.364	1.3911	0.6101	2.0012	2.0012	4.1865
0.351	0.541	1.559	1.2666	0.5326	1.7992	1.7992	4.3885
0.390	0.586	1.754	1.1617	0.4747	1.6363	1.6363	4.5514
0.424	0.621	1.949	1.0700	0.4317	1.5017	1.5017	4.6860
0.454	0.650	2.144	0.9926	0.3921	1.3847	1.3847	4.8031
0.481	0.675	2.339	0.9239	0.3657	1.2896	1.2896	4.8981
0.506	0.696	2.534	0.8654	0.3413	1.2067	1.2067	4.9810
0.528	0.714	2.729	0.8137	0.3188	1.1324	1.1324	5.0553
0.548	0.730	2.924	0.7667	0.2950	1.0617	1.0617	5.1261

Begin by writing the steady, incompressible Navier–Stokes equations in dimensionless form:

$$\mathrm{ReCa}\{\mathbf{u}\cdot\mathbf{u}\} + \nabla p = \mathrm{Ca}\,\nabla^2\mathbf{u} - \mathrm{Bo}\mathbf{e}_z;\ \mathbf{div}\ \mathbf{u} = 0 \quad (16)$$

where $\mathbf{u}$ denotes the vector velocity field, the Reynolds number $\mathrm{Re} = \rho V R_C/\mu$, the capillary number $\mathrm{Ca} = \mu V/\gamma$, and μ denotes the liquid viscosity. The velocity scale V is the falling speed of the sphere, only known *a posteriori*. The Bond number is that defined in Section 2. The three nonzero components of the stress field are

$$\sigma_{rr} = -p + 2\mathrm{Ca}u_r$$
$$\sigma_{zz} = -p + 2\mathrm{Ca}w_z$$
$$\sigma_{rz} = \mathrm{Ca}(u_z + w_r) \quad (17)$$

where u and w denote the radial and axial components of $\mathbf{u}$, respectively. The boundary conditions are those of no slip on solid surfaces and continuity of stress on the free surface.

There are seven free physical parameters: μ, ρ, γ, g, V, R_C, and R_S. The usual arguments from the Π theorem say that there are four possible independent dimensionless groups. The three appearing in equation (16), plus the dimensionless gap ratio introduced in Section 1, will be taken to be fundamental. Bo will be taken to be of order unity, and ϵ will be treated as an asymptotically small parameter about which a matched asymptotic expansion can be constructed. (The reader unfamiliar with this technique is referred to the monograph by Van Dyke (1975).) The Reynolds and capillary numbers will be taken small, and, to facilitate a single asymptotic expansion, write $\mathrm{Re} = R\epsilon$, $\mathrm{Ca} = K\epsilon^3$, where R and K are of order unity. This implies that the velocity-independent parameter $\mathrm{Ca}/\mathrm{Re} = \mu^2/(\rho\gamma R_C)$ is of order ϵ^2.

3.2 The Inner Solution. An appropriate set of spatial variables x, y in the inner region is defined by

$$r = 1 - (1/2)\epsilon^2 y;$$
$$z = \epsilon x. \quad (18)$$

The solid boundaries are given by $y = 0$ and $y = 1 + x^2$. The governing equations in the inner region can be written in terms of these variables as:

$$p_x - K[4w_{yy} + \epsilon^2(-2w_y + w_{zz})] = -\mathrm{Bo}\epsilon - RK\epsilon^3(-2uw_y + \epsilon ww_z);$$
$$-2p_y = K[4u_{yy} + \epsilon^2(-2u_y + u_{zz})] - RK\epsilon^3(-2uu_y + \epsilon wu_z);$$
$$-2u_y + \epsilon^2 u/r + \epsilon w_x = 0; \quad (19)$$

where subscripts have been used to denote partial differentiation. Solutions will be sought in the limit that ϵ is small, so that the momentum equations can be linearized and the usual lubrication approach can be taken.

The explicit boundary conditions for the set (19) are those on the solid boundaries:

$$w = 0 \text{ on } y = 1 + x^2;$$
$$w = 1 \text{ on } y = 0. \quad (20)$$

The boundary conditions in the x-direction take the form of matching conditions connecting this solution to that in the outer regions. These will be discussed after the solution is given.

Let p and $\mathbf{u}$ be expanded in powers of ϵ, e.g.,

$$p = p^{(0)} + \epsilon p^{(1)} + \epsilon^2 p^{(2)} + \ldots. \quad (21)$$

The governing equations for the first three terms are

$$p^{(0)}{}_y = 0; \quad p^{(0)}{}_x = 4Kw^{(0)}{}_{yy}; \quad u^{(0)}{}_y = 0. \quad (22a)$$

$$p^{(1)}{}_y = 0; \quad p^{(1)}{}_x = -\mathrm{Bo} + 4Kw^{(1)}{}_{yy};$$
$$u^{(1)}{}_y = (1/2)w^{(0)}{}_x. \quad (22b)$$

$$p^{(2)}{}_y = -2Ku^{(1)}{}_{yy}; \quad p^{(2)}{}_x = 4Kw^{(2)}{}_{yy};$$
$$u^{(2)}{}_y = (1/2)w^{(1)}{}_x. \quad (22c)$$

Solution of the first two of these is straightforward, and is more than adequate for the purposes of this paper. The usual approach by which Reynolds equations are derived, integrating the axial momentum equation twice with respect to y and applying the boundary conditions, gives formal expressions for w. The y integrals of these expressions from 0 to $1 + x^2$ are proportional to the flux between the sphere and the cylinder, and must be constant. Denoting the constant by q gives

$$2q^{(0)} = 1 + x^2 - p^{(0)}{}_x(1 + x^2)^3/24K; \quad (23a)$$
$$2q^{(1)} = -(p^{(1)}{}_x + \mathrm{Bo})(1 + x^2)^3/24K. \quad (23b)$$

These equations are first integrals of the Reynolds equations for this problem. (Note that this flux represents the inner represention of the actual flux (divided by π). The outer representation of the flux is ϵ^2 smaller than this.)

A second integration of the Reynolds equation gives

$$p^{(0)} = 12KF_1(x) - 6Kq^{(0)}F_2(x) + p^{(0)}{}_0, \quad (24a)$$
$$p^{(1)} = -\mathrm{Bo}x - 6Kq^{(1)}F_2(x) + p^{(1)}{}_0, \quad (24b)$$

where

$$F_1(x) = [x/(1 + x^2) + \tan^{-1}x];$$
$$F_2(x) = 2x/(1 + x^2)^2 + 3[x/(1 + x^2) + \tan^{-1}x]; \quad (25)$$

and the constants of integration $p^{(0)}$,$p^{(1)}$ represent terms in the expansion of the pressure at the midplane of the sphere, $x = 0$. The remaining parts of the solution are

$$w^{(0)} = 1 - y/(1 + x^2) + y(y - 1 - x^2)[3/(1 + x^2)^2 - 6q^{(0)}/(1 + x^2)^3]; \quad (26a)$$
$$w^{(1)} = -6y(y - 1 - x^2)q^{(1)}/(1 + x^2)^3; \quad (26b)$$
$$u^{(1)} = 2xy^2/(1 + x^2)^2 - 6q^{(0)}xy^2/(1 + x^2)^3 - 2xy^3/(1 + x^2)^3 + 6q^{(0)}xy^3/(1 + x^2)^4; \quad (27a)$$
$$u^{(2)} = -6q^{(1)}xy^2/(1 + x^2)^3 + 6q^{(1)}xy^3/(1 + x^2)^4. \quad (27b)$$

Discussion of the matching conditions will be deferred until the outer region has been addressed.

3.3 The Outer Solution and Matching. The equations in outer variables (r,z) are the components of (16)

$$p_r = K\epsilon^3\nabla^2 u - RK\epsilon^4(uu_r + wu_z);$$
$$p_z = -\mathrm{Bo} + K\epsilon^3\nabla^2 w - RK\epsilon^4(uw_r + ww_z);$$
$$(ru)_r + rw_z = 0; \quad (28)$$

and it is immediately clear that the pressure is given by

$$p = c^{(0)} - \mathrm{Bo}z + \epsilon c^{(1)} + \epsilon^2 c^{(2)}, \quad (29)$$

with no viscous effects until the third order in ϵ. The velocity field is undetermined at this level. It is not zero, however, as will be shown by a consideration of the matching conditions. The leading terms of the lowest-order outer representation of the velocity will be required for a formal expression of the viscous drag, and they will be found later in this section.

The matching principle is stated succinctly by Van Dyke (1975) as

the n-term inner expansion {the m-term outer solution} =

the m-term outer expansion {the n-term inner solution}.

The inner expansion of the outer solution is obtained by writing the outer solution in inner variables and examining its behavior as $\epsilon \to 0$. Similarly, the outer expansion of the inner solution is obtained by writing the inner solution in outer variables and taking the same limit. Comparison of the resulting approximations can be made in either system. Matching of the normal stresses requires matching of the pressure (the viscous contribution being negligible by comparison in both inner and outer solutions), which is done by taking the two-term inner expansion of the two-term outer solution:

$$c^{(0)} - \mathrm{Bo}z + \epsilon c^{(1)} \to c^{(0)} - \epsilon(\mathrm{Bo}x - c^{(1)}), \quad (30)$$

and comparing that to the two-term outer expansion of the two-term inner solution:

$$p^{(0)}(x)+\epsilon p^{(1)}(x)\rightarrow 12KF_1(z/\epsilon)-6K\epsilon^2 q^{(0)}F_2(z/\epsilon)$$
$$+p^{(0)}{}_0-\epsilon[\mathrm{Bo}z/\epsilon+6K\epsilon^2 q^{(1)}F_2(z/\epsilon)-p^{(1)}{}_0]. \quad (31)$$

Inspection of F_1 and F_2 reduces the two-term outer expansion to

$$6K\pi-\mathrm{Bo}z+p^{(0)}{}_0+\epsilon p^{(1)}{}_0. \quad (32)$$

Comparison of (30) and (31) shows that the matching conditions on the pressure can be reduced to

$$c^{(0)}=6K\pi+p^{(0)}{}_0;$$
$$c^{(1)}=p^{(1)}{}_0. \quad (33)$$

Repeating the analysis in the lower capillary region leads to two additional conditions, and two new constants

$$c_L{}^{(0)}=-6K\pi+p^{(0)}{}_0;$$
$$c_L{}^{(1)}=p^{(1)}{}_0. \quad (34)$$

The problem is apparently underdetermined: four equations in six unknowns. One pair is simple, but uninformative:

$$c^{(1)}=c_L{}^{(1)}=p^{(1)}{}_0. \quad (35)$$

The other condition relates the sum and difference of the two constants in the capillary regions to the central pressure as seen in the lubrication region. Recall that the sum of these two constants is zero in the absence of a lubrication region. Thus, the lubrication region supports a difference of pressure arising from a nonideal pairing of volumes and Bond numbers. The remaining two degrees-of-freedom, which can be thought of as permission to choose $p^{(0)}{}_0$, $p^{(1)}{}_0$ arbitrarily, are used to satisfy the condition of no net force on the sphere after the falling speed of the sphere is determined.

To determine the falling speed of the sphere and its surrounding capillary regions, it is necessary to find the integrated shear stress between the wall and the liquid. To find the integrated shear stress, it is necessary to find the lowest order (at least) part of w_r on $r=1$ in the capillary regions.

The lowest-order component of the velocity field in outer coordinates (in both the upper and lower capillary regions) satisfies the set of differential equations

$$p^{(3)}{}_r=K\nabla^2 u^{(0)};$$
$$p^{(3)}{}_z=-\mathrm{Bo}K\nabla^2 w^{(0)};$$
$$(ru^{(0)})_r+rw^{(0)}{}_z=0. \quad (36)$$

The last of these can be eliminated by defining the usual streamfunction, ψ, in terms of which

$$u=-\psi_z;\ w=(r\psi)_r/r. \quad (37)$$

The streamfunction satisfies the fourth-order differential equation

$$D^2\{D^2[\psi]\}=0, \quad (38)$$

where

$$D^2[\psi]=\{(1/r)[r\psi]_r\}_r+\psi_{zz}. \quad (39)$$

The matching condition suggests seeking solutions in the form

$$\psi=f_0(r)+f_1(r)/z^2+\ldots+f_j(r)/z^{2j}+\ldots, \quad (40)$$

with the functions $f_j(r)$ in the form of power series in $(1-r)$. Application of the matching principle on the velocity field gives

$$\psi=-(1-r)+a_{02}(1-r)^2+\ldots$$
$$+\{4(1-r)^2+a_{13}(1-r)^3+\ldots\}/z^2$$
$$+\{-4(1-r)^3+a_{24}(1-r)^4+\ldots\}/z^4$$
$$+\{a_{35}(1-r)^5+\ldots\}/z^6+\ldots, \quad (41)$$

which also satisfies the no slip boundary conditions on $r=1$. (Matching conditions on the viscous stress terms are automatically satisfied, as can be easily verified directly.) The unspecified constants in this expression are available to satisfy the differential equation and/or the boundary conditions on the inner edge of the outer regions. In particular, a_{02} is determined by boundary conditions. Fortunately, it does not enter the solution at the lowest order.

The shear stress on the system is found by constructing a one-term composite expansion for w in outer variables. This expression for w is inserted into the third of equations (17), noting that u and all its z derivatives are identically zero on the boundary. The shear stress is then integrated over the cylinder surface between Z_2 and Z_1 at the outer edge (called z_1 and z_2, respectively) at the outer edge and the result equated to the gravity force on the system.

The composite expansion is formed by additive composition (Van Dyke, 1975). The inner and outer expansions are added and their common part subtracted. That result, after eliminating $o(1)$ terms, is

$$w_c{}^{(0)}=1-8(1-r)/(\epsilon^2+z^2)+12\ (1-r)^2/(\epsilon^2+z^2)^2$$
$$-(1-r)/r+2a_{02}(1-r)+a_{02}(1-r)^2/r+\ldots$$
$$+z^{-2}\{4(1-r)^2/r\ \ldots\}$$
$$+z^{-4}\{-4(1-r)^3/r\ \ldots\}+\ldots. \quad (42)$$

The boundary shear stress is

$$\sigma_{rz}=\mathrm{Ca}[(1-2a_{02})+8/(\epsilon^2+z_2)], \quad (43)$$

and the leading term in the integral for the force comes from the second (inverse tangent) factor, making

$$F\sim 16\pi^2\mathrm{Ca}/\epsilon. \quad (44)$$

Equating this to the difference between the gravitational force and the net contact line force, which difference is supposed positive (otherwise a stationary solution as outlined in Section 2 above would be correct), and redimensionalizing gives an expression for the velocity of fall:

$$V=\{Mg-2\pi\gamma R_C(\cos\beta_T-\cos\beta_B)\}\sqrt{[R_C(R_C-R_S)]}/\{16\pi^2\mu R_C{}^2\}. \quad (45)$$

The analysis is completed by requiring the net force on the sphere to vanish. This is obtained by integrating the vertical component of the stress over the wetted surface, and adding to that the net contact line force. Consideration of the uniform composite expansions of w (given above) and u, given by

$$u_c{}^{(0)}=4z(1-r)^2/(\epsilon^2+z^2)^2-16z(1-r)^3/(\epsilon^2+z^2)^3$$
$$+2a_{13}(1-r)^3/z^3+\ldots, \quad (46)$$

shows that the vertical component of the surface stress is dominated by the pressure term. Integration of the pressure term leads to a net (scaled) pressure force (correct to $O(\epsilon)$)

$$F_P=\pi[3p_0(z_1{}^2-z_2{}^2)-2\mathrm{Bo}(z_1{}^3$$
$$-z_2{}^3)]/3+6K\pi^2(z_1{}^2+z_2{}^2). \quad (47)$$

To this is added the contact line force on the sphere, given by $2\pi(r_1\cos\theta_1+r_2\cos\theta_2)$ as in Section 2. That sum is equated to the force of gravity on the sphere, giving the analog of equation (14)

$$(4/3)\pi\rho_S a^3\mathrm{Bo}=2\pi(r_1\cos\theta_1+r_2\cos\theta_2)$$
$$+\pi[3p_0(z_1{}^2-z_2{}^2)-2\mathrm{Bo}(z_1{}^3-z_2{}^3)]/3+6K\pi^2(z_1{}^2+z_2{}^2) \quad (48)$$

which is to be solved for p_0. (In the limit that $K\rightarrow 0$, this is exactly equation (14); p_0 is that required for hydrostatic equilibrium and there is no motion).

3.4 Physicalizing the Expansion. As is often the case in analyses carried out using matched asymptotic expansions, the mathematics obscures somewhat the nature of the solution. Before moving on to look at the fall rates of spheres, I will spend a little time examining the solution.

The velocity field in the lubrication region—the gap—is a combination of (slowly varying) Couette and Poiseuille terms. For the case of finite liquid volumes, the relative size of these two terms is such that the het flux in the gap is zero, as it must be for steady motion. (The volume in the upper and lower menisci is constant, so there can be no net flow in the gap.) For illustrative purposes, the flow is essentially given by $w^{(0)}$, written in inner variables as equation (26*a*). Figure 2 shows a set of velocity profiles corresponding to that equation with zero net flux. The horizontal axis is x, the inner axial variable, and the vertical axis is $1-r$, where r is the outer radial variable. Unit velocity corresponds to one-half unit on the abscissa. The surface of the sphere does not appear spherical because of the different scales for the ordinate and abscissa. The surface is spherical; the distortion in the $r \rightarrow y$ transformation has been removed. The velocity is a function of the single variable $y/(1+x^2)$, lying on (0,1), so all the profiles are "similar," differing only in their scale of variation.

The outer limit of this solution corresponds to its behavior as $|x| \rightarrow \infty$. The velocity obviously persists into the outer region, as noted above. However, the shear decreases, and its divergence decreases even more rapidly. In the inner region the divergence of the stress tensor dominates gravity; in the outer region the relative roles reverse. Thus, the pressure field in the outer region is determined by gravity and surface tension, the inner region functioning as a jump condition linking the upper and lower menisci. There is a nonzero velocity field in each meniscus region, but it is not directly accessible analytically. It is accessible numerically, but that is beyond the scope of this paper. The point to be noted is that it does not affect the pressure field to lowest order, nor does it contribute to the viscous drag on the falling sphere.

This discussion can be made clearer by an example. Let $\epsilon = 0.1$, and choose K such that the pressure jump through the gap is π. Figure 3 shows the pressure jump in the gap plotted against both x and z. This is, in fact, a plot of $F_1(x)$ given by equation (25). Pressure changes are confined to the region in which $|x| = O(1)$.

Figure 4 shows the lowest-order composite pressure expansion—the hydrostatic balance plus the gap pressure shown in Fig. 3—again for $\epsilon = 0.1$. I have taken Bo = 1, and I have chosen the minimum height of the upper meniscus to be 0.75 and that of the lower to be 0.55, the same geometry shown in Fig. 1. The pressure in the liquid is everywhere below atmospheric, as determined by the curvatures of the two menisci. The difference in curvature between the two menisci balances the jump in pressure through the gap. The flow in the gap is the combined Couette and Poiseuille flow shown in Fig. 2. The flow in the two menisci is undetermined at this time, but it is most likely an axisymmetric roll or vortex with liquid rising along the cylinder wall and moving down and across along the spherical surface. There will be an inward flow along under the upper free surface and an outward flow over the lower free surface.

4 Discussion

The analysis presented depends on the two dimensionless groups Re and Ca having specified relations to the asymptotic parameter ϵ. Are these reasonable? From the previous section

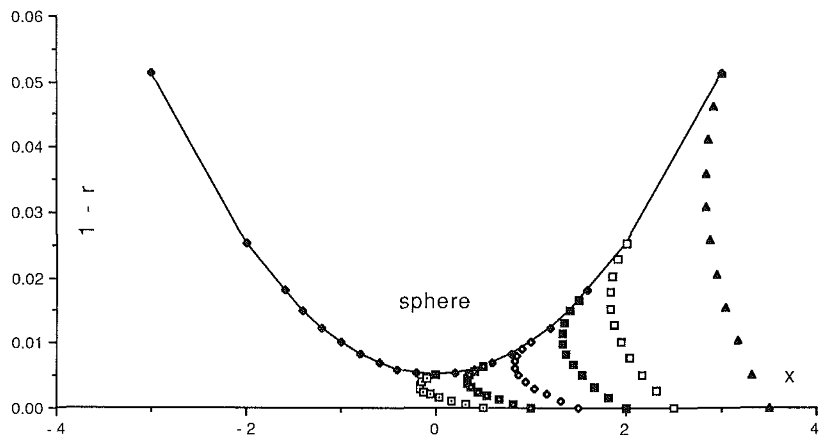


Fig. 2 Dimensionless velocity profiles. The sphere surface is distorted by the scaling, here corresponding to $\epsilon = 0.1$.

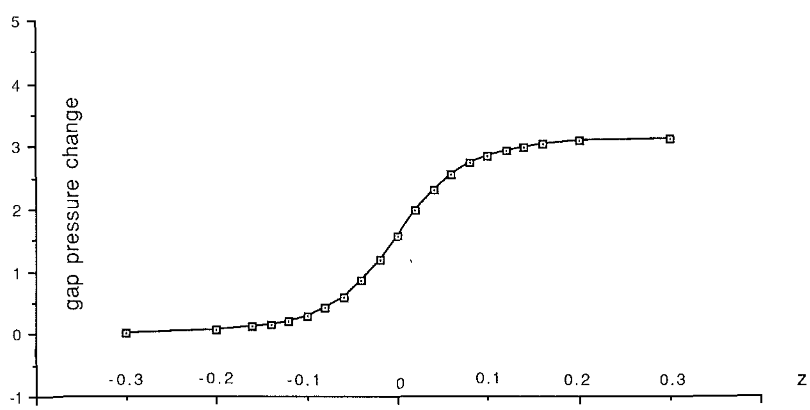


Fig. 3 Inner solution pressure at lowest order. Both inner coordinate x and outer coordinate z are shown for $\epsilon = 0.1$.

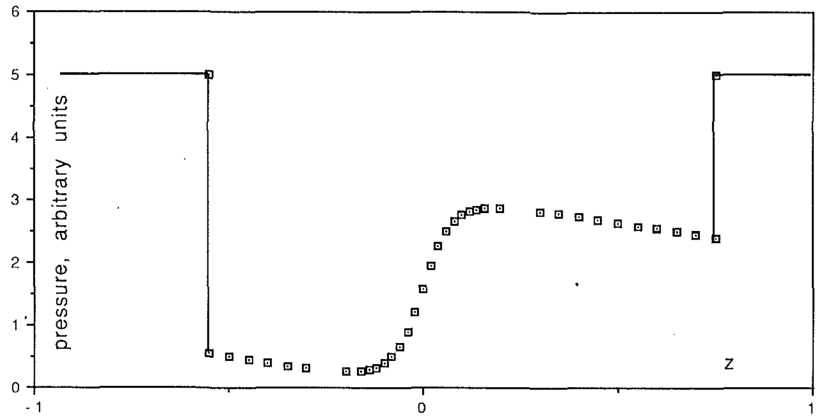


Fig. 4 Pressure from the lowest-order composite solution with $\epsilon = 0.1$

$$V = \{Mg - \Gamma\}\sqrt{[R_C(R_C - R_S)]}/\{16\pi^2 \mu R_C^2\},$$

where $\Gamma = \gamma R_C F_c = 2\pi\gamma R_C(\cos\beta_1 - \cos\beta_2)$ and the Reynolds and capillary numbers are

$$\mathrm{Re} = \epsilon p\{Mg - \Gamma\}/\{16\sqrt{2}\pi^2\mu^2\}; \qquad (49)$$

$$\mathrm{Ca} = \epsilon\{Mg - \Gamma\}/\{16\sqrt{2}\pi^2\gamma R_C\}. \qquad (50)$$

The former must be comparable to or small compared to ϵ, so that the numerical part of (49) can be at most $O(1)$. This puts a lower bound on the viscosity. Similarly, (50) puts a lower bound on the surface tension, a lower bound dependent on ϵ^2, the dimensionless gap. This equation can be rewritten in a more convenient and suggestive form

$$\mathrm{Ca} = \{M\mathrm{Bo} - F_c\}\epsilon/(16\sqrt{2}\pi^2), \qquad (50)'$$

the dimensionless force imbalance times ϵ divided by a large numerical factor: 223.32.... For moderate force imbalance, the condition on the capillary number is not hard to satisfy.

The most stringent bounds arise when there is no contact angle hysteresis, so that the entire equilibrium is supported by the viscous forces. The capillary number is still well behaved for moderate Bond number, but the relation between the viscosity and the size of the system begins to be quite limiting. That can be written

$$\mu^2 \geq \rho^2 g R_C^3[(4/3)\pi\rho_S a^3 + V_L]\mathrm{Bo}/(16\sqrt{2}\pi^2). \qquad (51)$$

For water, for example, $\mu = 1$ mPas, $\rho = 1000$ kg/m^3 and $\gamma = 72$ mPam, and the scale for R_C is of the order to a few hundred μm.

Does too large a Reynolds number compromise the solution fatally? The nonlinear term in the inner variables is of the order of ReCa, so that a reasonably large Re will not upset the solution in that region. If $\mathrm{Re} = O(1)$, the nonlinear term in the outer region, as in equations (36), will be comparable to the viscous terms, but this is not particularly disastrous, as that set of equations is not solved, and the representation of the outer velocity field in terms of a stream function is still valid. Thus, the condition on the Reynolds number can be relaxed with the expectation that the solution for the speed of fall will remain valid. The condition on the capillary number cannot be relaxed.

The biological application which originally motivated this work, swallowing, is worth examining, at least qualitatively. The diameter of the esophagus is between one and three cm in humans. The viscosity of saliva is about 5 cP. I have been unable to find a value for its surface tension; let me use 60 dyne/cm, a value appropriate to very dirty water. The esophagus is certainly wetted and both advancing and receding contact angles are likely to be near zero. I assume a spherical bolus of unit specific gravity, so that the mass of the bolus in grams is equal to its volume in cm^3. Combining these assumptions gives a capillary number

$$\mathrm{Ca} \approx 0.3. \qquad (52)$$

The coefficient of ϵ, here 0.3., is meant to be $O(\epsilon^2)$, so that application to swallowing is a bit strained. If, however, the attempt is made, one finds that the free fall swallowing speed is

$$V_{\mathrm{swallow}} \approx 368\epsilon \mathrm{cm/sec}. \qquad (53)$$

The actual value of ϵ is indeterminate, but it is likely to be quite small. (It is probably determinable by a squeeze film estimate taking into account the dynamic behavior of the real system: a problem well beyond the scope of this work.) Actual swallowing speeds are faster than that indicated by the estimate (53) for any reasonable value of ϵ. The esophagus is of the order of 25 cm long,and swallowing certainly takes less than one second, perhaps as little as 0.1 sec, giving speeds of 100–200 cm/sec. These speeds are, however, in the ballpark, suggesting that the qualitative aspects of this work are applicable to the problem of analyzing swallowing.

It would be of interest to verify this work experimentally, and to measure what happens outside the range of formal validity of the analysis. This turns out to be a difficult problem, not addressable within the scope of this paper. There is only space to report some suggestive early results of experiments dropping steel spheres inside small glass tubes using silicon oil (Dow-Corning 200 Fluid: density of 970 kg/m^3, surface tension 22 mPam, kinematic viscosity 10 mm^2/sec) as the lubricant. The spheres are chrome steel, and come in diameters of 1.5875 mm (1/16 in.), 2.38125 mm (3/32 in.), and further increments of 0.79375 mm (1/32 in.). Sphericity is given by the supplier (Small Parts, Inc.) as 0.635 μm (0.000025 in.). Common laboratory glass tubing is sized by outside diameter and wall thickness. The bores are not particularly precise, nor uniform in the axial direction, nor very round.

Several 3/32 in. balls were dropped down a closely fitting glass tube. The two major contributions to uncertainty for this casual experiment are the inside geometry of the tube and the amount of liquid entrained. The outside diameter of the tube was measured by caliper to be 4.1 mm. The inside diameter was indistinguishable by caliper measurement from the outside diameter of the balls. Two ml of 10-cS silicon oil were inserted in the tube, and the balls were wetted before insertion into the tube. The amount of liquid entrained is hard to judge. It was small, but nonzero. If the liquid were uniformly distributed over the inside of the tube, it would form a layer a little less than 400 μm thick. This is the same order of magnitude as the gap.

After some practice, three balls were dropped without replenishment of the liquid. The experiment was video taped and the balls were observed frame by frame. The position versus time plot was constructed, and the slope of this plot was taken to be the speed. The balls ran freely for about 435 mm before entering the camera field of view. The observations covered about 150 mm. In each case the upper 100 mm showed a higher speed than the lower 50. Observed speeds (upper/lower) were 173/134, 192/152, and 196/146 mm/s, respectively. The differences between the upper and lower segments of each trial are larger than the differences between trials. The Reynolds and capillary numbers based on the observed velocity are 23 and 0.084, respectively, within the range allowed by the analysis. Predicted speeds, based on equation (2), are hard to calculate because contact angles, entrainment volumes and gap are unknown. Under the balancing assumptions of zero contact angle and zero entrainment, equation (2) predicts a speed of 207 mm/sec, a number suggesting that the analysis is germane, but hardly demonstrating its validity.

Acknowledgments

The trial experiments were performed by Rob Athans and Larry Salva. The work was supported in part by USPHS grant DE 08511.

References

Bonneau, D., and Frene, J., 1983, "Film Formation and Flow Characteristics at the Inlet of a Starved Contact—Theoretical Study," *ASME Journal of Lubrication Technology*, Vol. 105, pp. 178–186.

Campanella, O. H., and Cerro, R. L., 1984, "Viscous Flow on the Outside of a Horizontal Rotating Cylinder: The Roll Coating Regime with a Simple Fluid," *Chemical Engineering Science*, Vol. 39, pp. 1443–1449.

Cerro, R. L., and Scriven, L. E., 1980, "Rapid Free Surface Film Flows: An Integral Approach," *Industrial and Engineering Chemistry Fundamentals*, Vol. 19, pp 40–50.

Concus, P., and Finn, R., 1974, "On Capillary Surfaces in a Gravitational Field," *Acta Math*, Vol. 132, pp. 207–223.

Dussan V., E. B., 1985, "On the ability of drops or bubbles to stick to non-horizontal surfaces of solids. Part 2. Small drops or bubbles having contact angles of arbitrary size," *J. Fluid Mech.*, Vol. 151, pp. 1–20.

Dussan V., E. B., 1987, "On the ability of drops or bubbles to stick to non-horizontal surfaces of solids. Part 3. The influences of the motion of the surrounding fluid on dislodging drops," *J. Fluid Mech.*, Vol. 174, pp. 381–397.

Dussan V., E. B., and Chow, R. T-P., 1983, "On the ability of drops or bubbles to stick to non-horizontal surfaces of solids," *J. Fluid Mech.*, Vol. 137, pp. 1–29.

Finn, R., 1986, *Equilibrium Capillary Surfaces*, Springer-Verlag, New York.

Gans, R. F., and Wang, C. H., 1990, "On the Disturbance of a Thin Layer of Liquid by a Moving Obstruction," ASME Journal of Applied Mechanics, Vol. 57, pp. 1066–1072.

Hsu, T. C., Malone, M., Laurence, R. L., and Middleman, S., 1985, "Separating Forces in Blade Coating of Viscous and Viscoelastic Liquids," *Journal of Non-Newtonian Fluid Mechanics*, Vol. 18, pp. 273–294.

Levich, V. G., 1962, *Physicochemical Hydrodynamics*, Prentice-Hall, Englewood Cliffs, N.J., Section 133.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1986, *Numerical Recipes*, Cambridge University Press.

Sullivan, T. M., and Middleman, S., 1986, "Film Thickness in Blade Coating of Viscous and Viscoelastic Liquids," *Journal of Non-Newtonian Fluid Mechanics*, Vol. 21, pp. 13–38.

Sullivan, T. M., Middleman, S., and Keunings, R., 1987, "Use of a Finite-Element Method to Interpret Rheological Effects in Blade Coating," *American Institute of Chemical Engineers Journal*, Vol. 33, pp. 2047–2056.

Tichy, J. A., 1986, "Entry Flow in a Narrow Channel of Varying Gap with One Sliding Surface," *ASME Journal of Tribology*, Vol. 108, pp. 288–293.

Tichy, J. A., and Bourgin, P., 1985, "The Effect of Inertia in Lubrication Flow Including Entrance and Initial Conditions," ASME Journal of Applied Mechanics, Vol. 52, pp. 759–765.

Tipei, N., 1978, "Flow Characteristics and Pressure Build-up at the Inlet of Narrow Passages," *ASME Journal of Lubrication Technology*, Vol. 100, pp. 47–55.

Van Dyke, M., 1975, *Perturbation Methods in Fluid Mechanics*, annotated ed., Parabolic Press, Stanford, Calif.

A. Bogobowicz[1]
Visiting Asst. Professor.

L. Rothenburg
Assoc. Professor.

M. B. Dusseault
Professor.

Department of Civil Engineering,
University of Waterloo,
Waterloo, Ontario N2L 3G1
Canada

Solutions for Non-Newtonian Flow Into Elliptical Openings

A semi-analytical solution for plane velocity fields describing steady-state incompressible flow of nonlinearly viscous fluid into an elliptical opening is presented. The flow is driven by hydrostatic pressure applied at infinity. The solution is obtained by minimizing the rate of energy dissipation on a sufficiently flexible incompressible velocity field in elliptical coordinates. The medium is described by a power creep law and solutions are obtained for a range of exponents and ellipse eccentricites. The obtained solutions compare favorably with results of finite element analysis.

Introduction

Determination of slow closure rates of elliptical openings in a nonlinearly viscous media is a problem of considerable interest to salt rock and potash mining. Although openings in salt rock are rarely excavated in an elliptical shape, development of fractures around deep openings and subsequent detachment of roof and floor slabs and pillar spalling result in openings of nearly elliptical shape (Mraz, 1973). Formation of an elliptical opening around a rectangular room is schematically illustrated in Fig. 1.

In deep salt and potash mines, where ground stresses are high enough to cause fracture of rock around openings, the modification of the initial opening geometry occurs within several months after excavation. Eventually, a couple of years later, the elliptical opening will experience steady-state closure.

In this paper, a semi-analytical solution for steady-state closure rates of infinitely long elliptical openings in an infinite medium is presented. Material behavior is described by steady-state power creep law. The solution is developed for a wide range of ellipse eccentricities. Stress field at infinity is assumed to be hydrostatic.

Closed-form solutions for nonlinear flow problems can be only obtained in effectively one-dimensional cases (flow between two plates, closure of a circular opening). FEM analysis is rather standard in this area, but computations are numerically intensive, and parametric studies that are frequently required in engineering applications are costly in terms of computer time involved.

There are a number of practically important problems that have been addressed in an analytical form by utilizing variational principles for nonlinearly viscous media and using kinematically admissible trial functions with adjustable parameters (Lioboutry, 1987 gives several examples). Gilormini and Montheillet (1986) successfully used trial velocity fields of a corresponding linear problem to study deformations of an incompressible elliptical inclusion in a nonlinearly viscous matrix. The problem of flow into an elliptical opening is addressed here in a similar manner, except it is found that velocity fields of linear and nonlinear problems are substantially different for high exponents of power-law viscosity.

Material Model

It is assumed that the material is described by a power creep law of the form:

$$\dot{\epsilon}'_{ij} = \frac{3}{2}\,\dot{\epsilon}_o\,\frac{\sigma'_{ij}}{\sigma_{ef}}\left(\frac{\sigma_{ef}}{\sigma_o}\right)^M \qquad (1)$$

where $\dot{\epsilon}_{ij}$ is the strain rate, σ'_{ij} is the stress deviator, $\sigma_{ef} = (3/2\sigma'_{ij}\sigma'_{ij})^{1/2}$ is the so-called "effective stress" assumed to govern creep rates in generalized stress conditions, M is the power law exponent, and σ_o, $\dot{\epsilon}_o$ are material constants (only one of which is independent). In the subsequent text the inverted form of the above creep law will be extensively used:

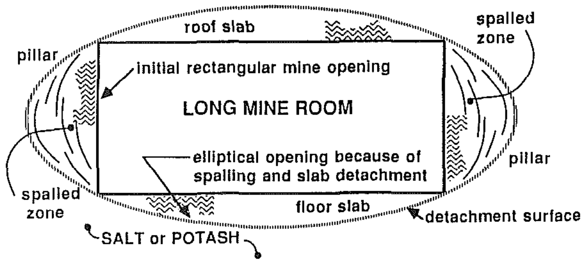


Fig. 1 Formation of an elliptical opening around a rectangular room

[1]Permanent address: Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for publication in the Journal of Applied Mechanics.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, Nov. 9, 1989; final revision, June 20, 1990.

$$\sigma'_{ij} = \frac{2}{3}\,\sigma_o\,\frac{\dot\epsilon'_{ij}}{\dot\epsilon_{ef}}\left(\frac{\dot\epsilon_{ef}}{\dot\epsilon_o}\right)^m\ (\dot\epsilon_{kk}=0) \qquad (2)$$

where $m = 1/M$ and $\dot\epsilon_{ef} = (2/3\,\dot\epsilon'_{ij}\,\dot\epsilon'_{ij})^{1/2}$ is the "effective strain rate." The definitions of effective stress and strain are such that the relationship (1) takes the form $\dot\epsilon_1 = \dot\epsilon_0\left(\frac{\sigma_1-\sigma_3}{\sigma_0}\right)^M$ in conditions of uniaxial compression.

It should be noted that the material law (1) is unlikely to be applicable for the entire range of stresses and certainly not for high stress deviators approaching uniaxial compressive strength. It seems that for salt rocks, the expression (1) is reasonable for $\sigma_{ef} < \sigma_o \approx 10$ MPa with $M = 3$. For higher stress deviators the creep mechanism changes and can be also approximated by a power law, but with a much higher exponent (Dusseault et al., 1987). The present study is limited to a single mechanism power law, although the technique described below is sufficiently general to be applicable for more complex creep models.

Variational Approach to Steady-State Solutions

In general, exact analytical solution for nonlinearly viscous flow can be obtained in effectively one-dimensional cases like flow through a pipe or closure of a circular opening. More complex problems can be addressed by noting that a velocity field in a material obeying the constitutive relationship (1) and resulting in an equilibrium stress field minimizes the following functional (Hill, 1956) on a set of all incompressible velocity fields ($\partial v_i/\partial x_i = 0$):

$$D = \frac{\sigma_o\epsilon_o}{m+1}\int_V\left(\frac{\dot\epsilon_{ef}}{\dot\epsilon_o}\right)^{m+1}dV - \int_B \sigma_{ij}n_jv_i\,dB. \qquad (3)$$

The first term in equation (3) represents the rate of energy dissipation within the volume V of the material, while the second term is the power of boundary tractions $\sigma_{ij}n_j$, where n is the outward normal to the boundary B surrounding V. In order to apply this principle to a particular problem without undue analytical complications, it is necessary to select a sufficiently wide set of physically realistic incompressible velocity fields specific to the problem.

Incompressible Velocity Fields in Elliptical Coordinates

In further analysis, it is convenient to describe the shape of the opening and to seek the solution in elliptical coordinates that transform the exterior of a circle in (ρ, θ) polar coordinates into the exterior of the ellipse on an x, y-plane as follows:

$$x = \left(\rho + \frac{\bar r^2\epsilon}{\rho}\right)\cos\theta;\ y = \left(\rho - \frac{\bar r_e^2}{\rho}\right)\sin\theta \qquad (4)$$

where $\bar r$ is the mean of the ellipse semi-axes $a = \bar r\,(1+e)$, $b = \bar r\,(1-e)$ and e is the ellipse eccentricity. In further analysis, $\bar r$ will be taken as unity and final results for any size opening will be obtained by simple scaling. Figure 2 illustrates families of ρ = const and θ = const lines that form a mesh similar to that used for finite element computations. Elements of length dl_θ, dl_ρ in x, y coordinates along θ, ρ-lines can be calculated by differentiating (3) as follows:

$$dl_\theta = \rho g d\theta;\ dl_\rho = g d\rho$$
$$g = \left(1 - \frac{2e}{\rho^2}\cos 2\theta + \frac{e^2}{\rho^4}\right)^{1/2}. \qquad (5)$$

The simplest incompressible velocity field in elliptical coordinates physically describes D'Arcy flow through porous media. In this case fluid velocities in θ-direction are zero and each flow channel between two θ = const lines carries the same quantity of flow dq, i.e., $v_\rho dl_\theta$ = const = dq. Considering (4), the velocity field of subsequent interest is as follows:

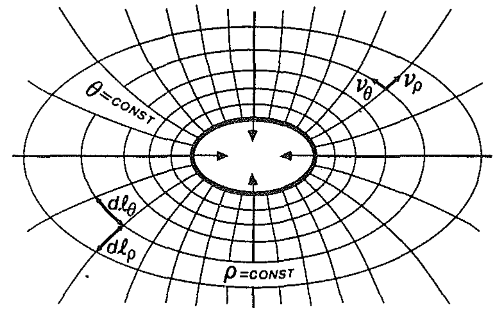


Fig. 2 Elliptical coordinates

$$v_\rho(\rho,\theta) = \frac{\text{const}}{\rho g};\ v_\theta(\rho,\theta) = 0. \qquad (6)$$

A somewhat more complex incompressible velocity field corresponds to a linearly viscous (Newtonian) flow into an elliptical opening. This solution can be obtained by noting a direct analogy between displacements of a linearly elastic incompressible solid (Poisson's ratio 0.5) and velocities of an incompressible viscous fluid in the same boundary value problem. The velocity field obtained using this analogy can be derived by extending complex variables elasticity solution (e.g., Love, 1927) and is as follows:

$$v_\rho(\rho,\theta) = v_o\,\frac{1-\xi\cos 2\theta}{\rho g}\left(\xi = \frac{2e}{1+e^2}\right);\ v_\theta(\rho,\theta) = 0 \qquad (7)$$

where v_o is the mean closure rate of the opening (averaging over the entire opening surface).

In the subsequent analysis it will be convenient to characterize the response of the opening in terms of the average of vertical and horizontal closure rates $\bar v = 1/2(v_{\text{vert}} + v_{\text{horiz}})$ and in terms of the relative difference between closure rates in vertical and horizontal directions, i.e., $\chi = (v_{\text{vert}} - v_{\text{horiz}})/(v_{\text{vert}} + v_{\text{horiz}})$. For the solution (7), $\bar v = v_o/(1+e^2)$, $\chi = e$.

In connection with (7), an important question is the extent to which the velocity field for non-Newtonian flow is different from the velocity field (7) of a corresponding linear problem. While one cannot generally expect that both are identical, except for a circular opening, the incompressibility of flow and identical boundary conditions for linear and nonlinear problems do suggest that the velocity fields in both cases should be not far different.

The velocity field (7) has the property that

$$v_\rho(\rho,0) + v_\rho(\rho,\pi/2) = 2v_\rho(\rho,\pi/4). \qquad (7)$$

Several numerical solutions of nonlinear problems with integer exponents (M = 2, 3, 4, 5, 6) all suggest that this property of the velocity field of the linear problem is preserved in nonlinear cases. Also, detailed analysis of numerical solutions show that velocity fields in nonlinear cases generally contain terms of the order $1/\rho^2$, $1/\rho^3$ and possibly of higher orders of $1/\rho$. Nevertheless, the variation of the velocity field along the coordinate line $\theta = \pi/4$ is exactly according to (7). These observations on numerical solutions suggest that the velocity fields for non-Newtonian flow can be taken in the form:

$$v_\rho(\rho,\theta) = v_o\left(\frac{1}{g\rho} + \frac{f_1(\cos 2\theta)}{g\rho} + \frac{f_2(\cos 2\theta)}{g\rho^2} + \frac{f_3(\cos 2\theta)}{g\rho^3} + \ldots\right). \qquad (8)$$

The fact that f_1, f_2, f_3 are functions of $\cos 2\theta$ follows from the symmetry of the opening geometry. Also, since for $\theta = \pi/4$ the solution degenerates into (6), $f_1(0) = f_2(0) = f_3(0) = 0$. In the subsequent analysis, both functions will be in the form $f_i(\cos 2\theta) = c_i\cos 2\theta$, where c_i are constants to be determined by minimizing (3) numerically. Minimization, with respect to v_o in (8), will be performed analytically.

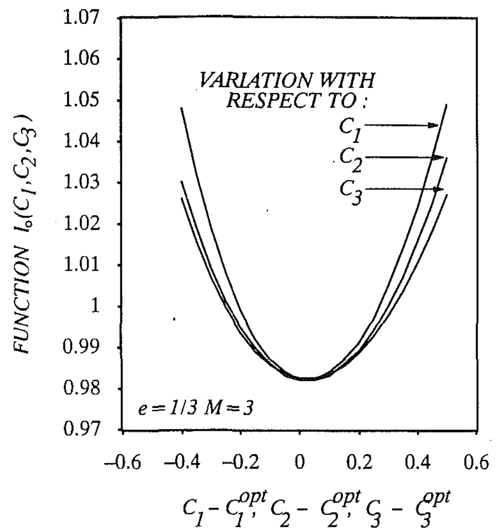


Fig. 3 Function I_o in the vicinity of minimum (e = 1/3, M = 3)

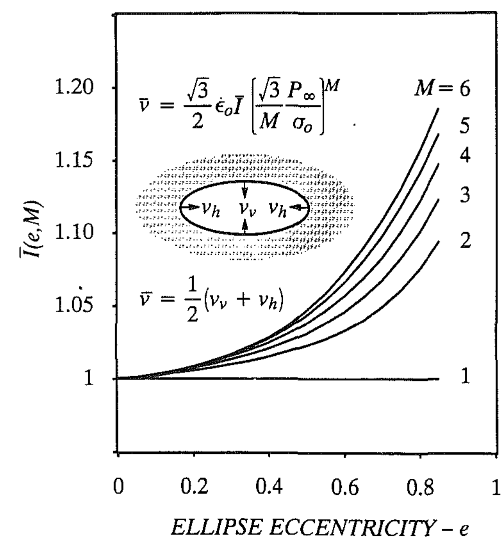


Fig. 4 Calculation of average closure rates of elliptical openings of different eccentricities and power-law exponents

Before proceeding, it is necessary to determine the form of the component v_θ associated with the radial component of the velocity field (8). The incompressibility condition is sufficient to solve this problem. The required expressions for strain rates are as follows:

$$\dot{\epsilon}_{\rho\rho} = \frac{1}{g}\frac{\partial v_\rho}{\partial \rho} + \frac{v_\theta}{\rho g^2}\frac{\partial g}{\partial \theta}$$
$$\dot{\epsilon}_{\theta\theta} = \frac{v_\rho}{\rho g^2}\frac{\partial(\rho g)}{\partial \rho} + \frac{1}{\rho g}\frac{\partial v_\rho}{\partial \rho} \tag{9}$$
$$2\dot{\epsilon}_{\rho\theta} = \rho\frac{\partial}{\partial \rho}\left(\frac{v_\theta}{\rho g}\right) + \frac{\partial}{\partial \theta}\left(\frac{v_\rho}{\rho g}\right).$$

These can be obtained from general expressions for strains in curvilinear coordinates given in (Love, 1927). Calculation of strains for a single term in (8) of the form $v_\rho = f(\theta)/g\rho^n$ and substitution into the incompressibility condition $\dot{\epsilon}_{\rho\rho} + \dot{\epsilon}_{\theta\theta} = 0$ gives the following differential equation for v_θ:

$$\frac{\partial v_\theta}{\partial \theta} + v_\theta = (n-1)v_r.$$

The solution of this equation is as follows:

$$v_\theta(\rho,\theta) = (n-1)\frac{1}{g\rho^n}\int f(\theta)d\theta.$$

The above solution must be a single-valued function of θ. In summary, the following incompressible velocity field will be used to describe flow into an elliptical opening:

$$v_\rho = v_o\left[\frac{1}{g\rho} + \cos 2\theta\left(\frac{c_1}{g\rho} + \frac{c_2}{g\rho^2} + \frac{c_3}{g\rho^3}\right)\right]$$
$$v_\theta = v_o\left[-\frac{1}{2}\sin 2\theta\left(\frac{c_2}{g\rho^2} + \frac{2c_3}{g\rho^3}\right)\right]. \tag{10}$$

Minimization Procedure

Minimization of the functional (3) will be carried out in conditions of a hydrostatic stress field at infinity ($\sigma_{ij} = p_\infty \delta_{ij}$). In this case the boundary tractions term in (3) becomes $p_\infty \int_B v_\rho dB = 2\pi p_\infty v_o$. This can be demonstrated by noting that the element of length in elliptical coordinates is $dB = g\rho d\theta$ and θ-related terms in (10) result in zero contributions, since integration with respect to θ is between 0 and 2π. Further, considering that $\dot{\epsilon}_{ef}$ is linear with respect to v_o, the functional (3) can be written in the form:

$$D = 2\pi\epsilon_o\sigma_o\left[\frac{I}{m+1}\left(\frac{v_o}{\dot{\epsilon}_o}\right)^{m+1} - \frac{p_\infty}{\sigma_o}\left(\frac{v_o}{\dot{\epsilon}_o}\right)\right]$$
$$I = \frac{1}{2\pi}\int_0^\infty\int_0^{2\pi}(\bar{\epsilon}_{ef})^{m+1}\rho g d\rho d\theta, \tag{11}$$

where $\bar{\epsilon}_{ef}$ is the effective strain rate computed with strain rates (9) for the velocity field (10) with v_o taken as unity.

The minimum of (11), with respect to v_o, can be obtained by equating to zero the derivative of D with respect to v_o to obtain:

$$v_o = \epsilon_o\left(I^{-1}\frac{p_\infty}{\sigma_o}\right)^M. \tag{12}$$

At the point of minimum with respect to v_o, (11) becomes $-2\pi\frac{M}{M+1}v_o p_\infty$ with v_o given by the above expression. Further minimization of the functional D with respect to constants c_1, c_2, c_3 is equivalent to minimizing I in the form (11). This can only be done numerically.

It is worth noting that for a circular opening ($e = 0$, $c_1 = c_2 = c_3 = 0$), I in the form (11) can be calculated analytically to obtain $M(2/\sqrt{3})^{1/M}/\sqrt{3}$. If I is normalized by this constant to introduce new I_o that is unity for a circular, the solution (12) can be rewritten as follows:

$$v_o = \frac{\sqrt{3}}{2}\epsilon_o\left(\frac{\sqrt{3}}{MI_o}\frac{p_\infty}{\sigma_o}\right)^M. \tag{13}$$

Expression (14) becomes an exact solution for a closure rate of a circular opening of a unit radius when $e = 0$ and $I_o = 1$. The function $\bar{I}_o(M, e, c_1, c_2, c_3)$ remains close to unity for nonzero ellipse eccentricities and for fixed M, e is very insensitive to variation of the remaining arguments c_1, c_2, c_3, typically changing within one percent within a physically reasonable range of values. Because of this feature, minimization of I_o poses considerable numerical difficulties and has been carried out using the conjugate gradient method with linear search based on a parabolic approximation in the search direction and combined with logarithmic gold section and skew testing. The integral in (11) has been evaluated numerically using Gauss-Legendre quadratures after the transformation $z = 1/\rho$ that give finite integration limits for z. Figure 3 presents a typical variation of I_o in the vicinity of its minimum (for $e = 1/3$ and $M = 3$).

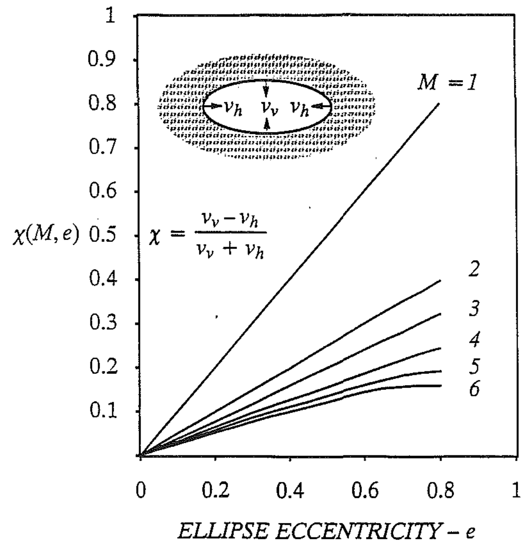


Fig. 5 Relative difference between vertical and horizontal closure rates for different ellipse eccentricities and power-law exponents

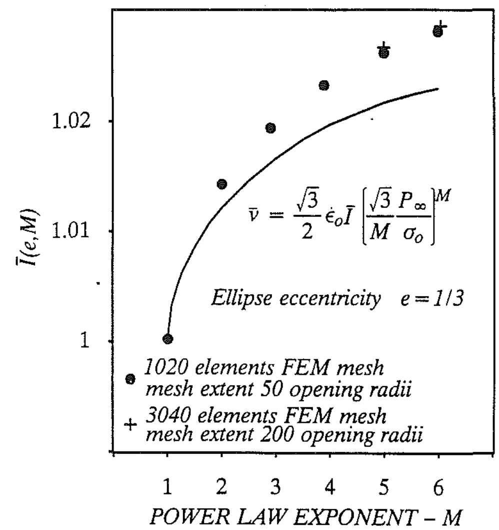


Fig. 7 Comparison of semi-analytical and FEM solutions for the average closure rate ($e = 1/3$)

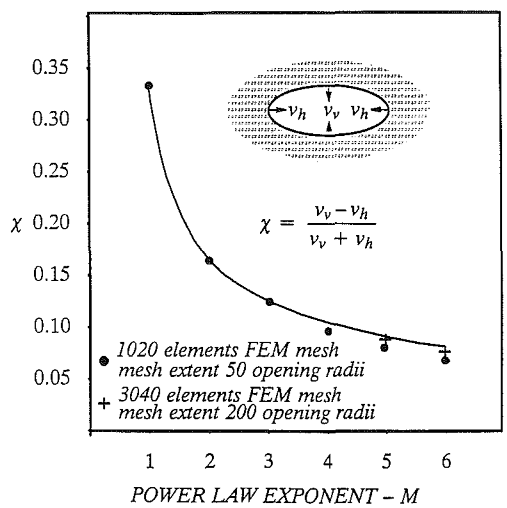


Fig. 6 Comparison of semi-analytical and FEM solutions for the relative difference between vertical and horizontal closure rates ($e = 1/3$)

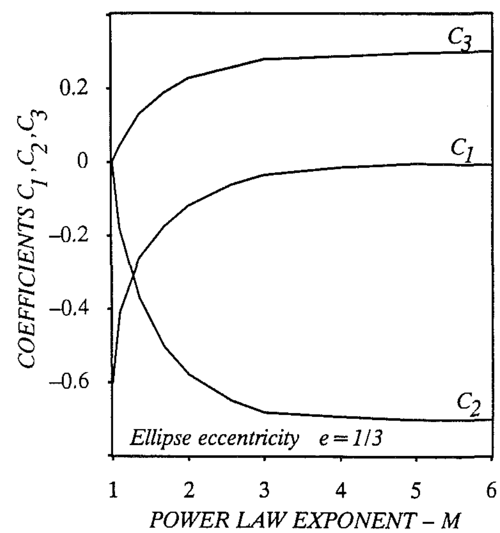


Fig. 8 Variation of the minimization coefficients with the power-law exponent ($e = 1/3$)

Results

The final result is convenient to present in the form similar to (13), but written for the average of vertical and horizontal closure rates $\bar{v}$, since this characteristic can be easily determined in practical situations of closure measurements. The solution for $\bar{v}$ can be presented in a simple form by redefining I_o in (13) as follows:

$$\bar{v} = \bar{r}\,\frac{\sqrt{3}}{2}\,\dot{\epsilon}_o \bar{I}\left(\frac{\sqrt{3}}{M}\frac{p_\infty}{\sigma_o}\right)^M \tag{14}$$

where the form of $\bar{I}\,(e, M)$ is shown in Fig. 4 for a range of ellipse eccentricities and power-law exponents. Note that (14) is presented for an ellipse of an arbitrary mean radius $\bar{r}$.

In practical situations of in-situ closure measurements, horizontal (wall-to-wall) and vertical (roof-to-floor) closure rates frequently become available. The relative difference between vertical and horizontal closure rates $\chi = (v_v - v_h)/(v_v + v_h)$ is independent on creep parameters $\dot{\epsilon}_o$, σ_o and is presented for different ellipse eccentricities and power-law exponents in Fig. 5. The parameter χ was calculated on the basis of coefficients c_1, c_2, c_3 obtained by minimization. In-situ measurements of χ can be used for determination of the power-law exponent while $\bar{v}$, calculated according to (14) using $\bar{I}\,(e, M)$ from Fig. 5, can be used for determination of $\dot{\epsilon}_o$.

Figures 6 and 7 illustrate comparison of steady-state FEM solutions with the solution obtained by the described method. It should be mentioned that for high values of M, very large FEM meshes are required to achieve an accurate solution. This is because of a slow decay of shear stresses away from the opening (of the order of $1/\rho^{2/M}$. To account for the effect of finite boundaries in FEM calculations, the comparison with FEM runs has been made with the solution of the form:

$$\bar{v} = \bar{r}\,\frac{\sqrt{3}}{2}\,\dot{\epsilon}_o \bar{I}\left(\frac{\sqrt{3}}{M}\frac{p_r}{\sigma_o}\,\frac{1}{[1-(\bar{r}/r)^{2/M}]}\right)^M \tag{15}$$

where r denotes the position of the boundary where pressure p_r is applied. The above formula is equivalent to (14) for $r = \infty$. For $e = 0$ ($I = 1$), it is identical to the solution for the closure rate of a thick cylinder with inner radius $\bar{r}$ and outer radius r. It is not difficult to show that the solution technique presented in the paper leads to the solution of the above form for a problem of a thick elliptical cylinder.

To appreciate the effects of finite boundaries, it is worth giving a numerical example: Presence of a boundary 200 mean ellipse radii away from the opening increases the closure rate by a factor 7.2 compared to the infinite media case for the power-law exponent of 6. The discrepancy between infinite media and finite media solutions is still a factor of 2 when the boundary is 500 mean radii away from the opening.

The values of coefficients c_1, c_2, c_3, controlling the shape of the steady-state velocity field around elliptical openings, are illustrated in Fig. 8. It should be noted that for the power-law exponents greater than 3, the dominant terms in the velocity field are controlled by coefficients c_2, c_3, while for exponents close to unity c_1 is dominant.

Closing Remarks

The described technique for obtaining steady-state solutions for closure rates of an elliptical opening in nonlinearly viscous materials, like salt rock, is rather general, and can be used for more complex constitutive models for steady-state creep. The accuracy of the method for the power creep law was assessed by comparing solutions obtained by minimizing functional (3) with results of FEM computations.

An interesting qualitative feature of the obtained solution is that mean closure rate of an elliptical opening is very close to a closure of an equivalent circular opening of a mean radius. Also, it is interesting to note that vertical and horizontal closure rates differ little for a practical range of power-law exponents (M = 3 and higher). The higher the exponent, the less is the difference between horizontal and vertical closure rates (Fig. 5).

Acknowledgments

The research described in this paper was supported by the National Science and Engineering Research Council of Canada.

References

Dusseault, M. B., Mraz, D., and Rothenburg, L., 1987, "The Design of Openings in Salt Rock using a Multiple Mechanism Viscoplastic Law," *Proc. of the 28th U.S. Symposium on Rock Mechanics*, Tucson, Ariz., pp. 633–642.

Gilormini, P., and Montheillet, F., 1986, "Deformation of an Inclusion in a Viscous Matrix and Induced Stress Concentrations," *J. Mech. Phys. Solids*, Vol. 34, No. 1, pp. 97–122.

Hill, R., 1956, "New Horizons in the Mechanics of Solids," *J. Mech. Phys. Solids*, Vol. 5, pp. 97–122.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, Oxford University Press, Oxford, U.K.

Liboutry, L. A., 1987, *Very Slow Flow of Solids*, Martinus Nijhoff Publishers, Dordrecht, The Netherlands.

Mraz, D., 1973, "Behavior of Rooms and Pillars in Deep Potash Mines," *CIM Transactions*, Vol. 76, pp. 138–143.

Jeffrey S. Marshall
Department of Ocean Engineering,
Florida Atlantic University,
Boca Raton, FL 33431
Assoc. Mem. ASME

Downstream Development of Viscous Fluid Wakes Behind Rod-Like Bodies

A new theory of viscous fluid wakes behind rod-like bodies is presented and is used to study the onset and downstream development of vortex street flows. Analytical solutions are obtained for the evolution of wave number, mean centerline velocity, vortex velocity, and vortex "spacing ratio" as a function of downstream distance in a laminar vortex street. A simple criterion for the onset of oscillations in the far wake, which slightly precede vortex street initiation, is also obtained. All of these solutions account for the action of viscous diffusion in spreading the street, and they are found to compare quite well with available experimental results.

1 Introduction

A great many different models of vortex street wakes can be found in the literature, e.g., von Kármán (1911), Hooker (1936), Kida (1982), Saffman and Schatzman (1982a,b). These studies employ highly idealized models which are intended to mimic the nature of the actual flow patterns within the wake. Most existing analytical studies focus on obtaining a criterion for vortex street stability, which is usually expressed in terms of the vortex "spacing ratio," and do not provide predictions for other pertinent features of the street, such as wave number and mean centerline velocity, or for the onset of instability in initially steady wakes. Also, most existing studies assume inviscid flow and do not account for the action of viscous diffusion in spreading the street.

In this paper a new theory of fluid wakes behind rod-like bodies is presented which includes viscous effects, and this theory is then used to obtain simple analytical solutions for the previously mentioned features of vortex street flows which are not provided by existing models. A criterion for stability of viscous vortex streets is also obtained, and although considerably different in form than previous such criterions, it is found to agree well with experimental data and to yield numerical results for spacing ratio immediately behind the body that are quite close to von Kármán's classic solution.

The theory for wake flows in this paper involves a description of the wake in terms of certain characteristic features of the flow, along with development of equations governing the evolution of these features with downstream distance. The general theory is valid throughout all regions of wakes formed behind rod-like bodies of otherwise arbitrary shape, both before and after the onset of vortex shedding from the body; however, specific solutions are obtained only for particular regions and regimes of the wake flow. The wake is assumed to be both laminar and incompressible, and the flow field is assumed to be unbounded in directions normal to the wake.

The theory is patterned after the directed fluid sheet model of Green and Naghdi (1976), which was originally applied to obtain wave equations for inviscid water waves in fluid of small depth. More general theories using directed fluid sheet models have been applied to viscous flow in channels (Green and Naghdi, 1984) and to inviscid waves in waters of finite and infinite depths (Green and Naghdi, 1986, 1987). In the present paper, the wake behind a rod-like body is modeled as a surface which is endowed with some number of additional kinematical variables call "directors" (i.e., a directed surface), and the rod-like body is further modeled as a curve of discontinuity on the directed surface. These additional variables, or their rates, may be identified with characteristic features of the wake flow. It is also noted that the present study somewhat extends the scope of development of directed fluid sheet theories: It is the first such theory (1) to consider a flow which does not possess material surface conditions along the lateral boundaries and (2) to utilize weighting functions that are functions of time and space-dependent variables in addition to the coordinate normal to the directed surface. Theories based on directed fluid sheet models may be constructed such that invariance conditions are identically satisfied and such that jump conditions across curves of discontinuity are immediately obtained from the governing equations. In particular, the jump condition in momentum across the rod is found to provide an ideal vehicle by which to relate features of the flow in the wake to forces acting on the body.

An alternative derivation of the wake equations using approximations from the usual three-dimensional theory is outlined in the Appendix. In this alternative derivation, governing equations for the wake are obtained by certain weighted averages of the full Navier-Stokes equations, together with a plausible representation for velocity variation across the wake. These integrated momentum equations include the effects of

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Technical Editor, Prof. Leon M. Keer, The Technological Institute, Northwestern University, Evanston, IL 60208, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 28, 1989; final revision, May 9, 1990.

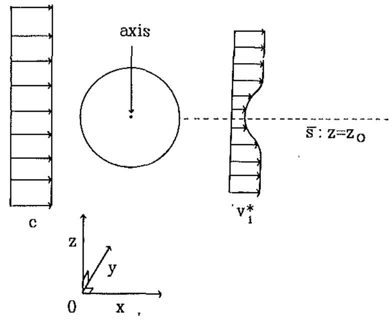


Fig. 1 Schematic diagram of the wake behind a rod-like body

viscous diffusion and pressure changes within the wake. Further, the independent variables that characterize the wake are not allowed to vary completely independently, but are instead constrained such that the incompressibility condition is satisfied identically at every point in the flow field. This alternative derivation is used to motivate certain constraints and constitutive equations in the directed fluid sheet theory. It is noted, however, that this alternative derivation is not unique and that relatively few such approximate formulations from the three-dimensional theory can be represented in terms of a directed fluid sheet model. Some further details regarding restrictions imposed by the directed fluid sheet model on the weighting functions in the alternative derivation of the theory are given by Marshall (1987).

The theoretical development is given mainly in Section 2, including a derivation of the governing equations for the wake flow. Section 3 deals with specification of certain boundary conditions for the independent variables of the theory by considering the effect of forces acting on the rod. Following these preliminaries, the theory is used in various applications in Section 4, including comparison with previous results for steady wakes, prediction of the onset of instability in the far wake, and prediction of the evolution of centerline velocity, wave number, and "spacing ratio" in a fully developed vortex street.

2 Derivation of the Wake Equations

The theory in the present paper deals with fluid flow in incompressible laminar viscous wakes behind rod-like bodies. For simplicity, the velocity $\mathbf{v}^*$ in the three-dimensional theory is allowed to approach a uniform, prescribed value $\mathbf{c}=\mathbf{c}(t)$ at distances infinitely above and infinitely below the plane of the wake. A rod-like body β_r is considered to be any object immersed in the flow which is considerably longer in a direction normal to $\mathbf{c}$ than it is in a direction parallel to $\mathbf{c}$. The axis of the body is identified with the volumetric centerline of β_r, and the body need not necessarily be symmetric about the axis. The middle surface $\bar{s}$ of the wake is chosen to be parallel to $\mathbf{c}$ and to pass through the rod axis, or through the mean location of the rod axis if the rod is in motion.

With reference to a fixed Cartesian coordinate system $x_i=(x,y,z)$ and associated base vectors $\mathbf{e}_i$, let $\mathbf{e}_2$ be parallel to the direction of the axis of the body and let $\mathbf{e}_3$ be normal to the velocity vector $\mathbf{c}$ at infinity. The location of a typical fluid particle in the current configuration is denoted by $\mathbf{x}=x_i\mathbf{e}_i$. A schematic diagram showing a cross-section of the wake in the x-z plane is provided in Fig. 1. The usual summation convention over repeated tensor indices, such that lower case Latin indices take on the values $\{1,2,3\}$ and Greek indices take on the values $\{1,2\}$, is assumed throughout the paper. Also, a superscript star is sometimes used to designate a variable as pertaining to the general three-dimensional flow field rather than to a particular surface in the flow.

The wake is modeled by a surface s which is endowed with three directors $\mathbf{d}_N=\mathbf{d}_N(x_\alpha,t)$, $N=1,2,3$, and their associated velocities $\mathbf{w}_N=\mathbf{w}_N(x_\alpha,t)$. The location and ordinary velocity of material points on s are designated by

$$\mathbf{r}=\mathbf{d}_0=x_\alpha\mathbf{e}_\alpha+\bar{d}\mathbf{e}_3,\ \mathbf{v}=\mathbf{w}_0(x_\alpha,t), \tag{1}$$

where $\bar{d}$ is a prescribed constant. For sufficiently large values of $\bar{d}$, the velocity $\mathbf{v}=\dot{x}_\alpha\mathbf{e}_\alpha$ of material points on the directed surface approaches the uniform value $\mathbf{c}$. The ordinary and director velocities of the surface s may thus be written in component form as

$$\mathbf{v}=c_\alpha\mathbf{e}_\alpha,\ \mathbf{w}_1=w_{1\alpha}\mathbf{e}_\alpha+w_{13}\mathbf{e}_3,$$
$$\mathbf{w}_2=w_{2\alpha}\mathbf{e}_\alpha+w_{23}\mathbf{e}_3,\ \mathbf{w}_3=w_{3\alpha}\mathbf{e}_\alpha+w_{33}\mathbf{e}_3. \tag{2}$$

The first index of quantities associated with the directors (such as $\mathbf{d}_N$, $\mathbf{w}_N$) indicates the director number and should not be confused with tensor indices. At times, the director index $N=0$ attached to a director variable is used to refer to the ordinary kinematic surface variables of s.

In addition to the variables associated with the directors, we also admit a function $\ell=\ell(x_\alpha,t)$ called the "decay coefficient" and defined such that $\ell\geq 0$ for all x_α and t. It is clear from the representation (A1) for the velocity field $\mathbf{v}^*$ of the three-dimensional theory that ℓ is related to the inverse of the "wake width." In order to facilitate understanding of these basic variables a brief summary of these variables, along with their physical interpretations with respect to the three-dimensional flow illustrated in Fig. 1, is given below:

$c_\alpha=c_\alpha(t)$	=	prescribed fluid velocity as $z\to\pm\infty$;
$w_{1\alpha}=w_{1\alpha}(x_\alpha,t)$	=	streamwise velocity at the middle surface minus the velocity at infinity;
$w_{13}=w_{13}(x_\alpha,t)$	=	cross-section velocity at the middle surface;
$w_{21}=w_{21}(x_\alpha,t)$	=	part of the streamwise velocity which is antisymmetric about the middle surface;
$w_{23}=w_{23}(x_\alpha,t)$	=	part of the cross-stream velocity which is symmetric about the middle surface;
$w_{33}=w_{33}(x_\alpha,t)$	=	part of the cross-stream velocity which is antisymmetric about the middle surface and vanishes on this surface;
$\ell=\ell(x_\alpha,t)$	=	coefficient of exponential velocity decay in a direction normal to the middle surface.

For simplicity, the components $w_{3\alpha}$ if $\mathbf{w}_3$ and w_{22} of $\mathbf{w}_2$, which do not appear in the velocity representation (A1), are set identically equal to zero.

Now let $\mathcal{P}$ be an arbitrary part of the surface $\bar{s}$ bounded by a closed curve whose outward unit normal on $\bar{s}$ is $\nu=\nu_\alpha\mathbf{e}_\alpha$. Guided in part by the derivation from the three-dimensional theory in the Appendix and using a fixed surface area on $\bar{s}$, an Eulerian form for the balance of ordinary ($N=0$) and director ($N=1,2,3$) momentum is postulated as

$$\frac{\partial}{\partial t}\int_{\mathcal{P}}\sum_{M=1}^{K}\rho y_{MN}\mathbf{w}_M d\sigma+\int_{\partial\mathcal{P}}\sum_{M=1}^{K}\rho\mathbf{w}_M\mathbf{v}_{MN}\cdot\nu ds$$
$$-\int_{\mathcal{P}}\sum_{M=1}^{K}\rho\mathbf{w}_M a_{MN}d\sigma=\int_{\mathcal{P}}(\rho\ell_N-\mathbf{k}_N)d\sigma+\int_{\partial\mathcal{P}}\mathbf{m}_N ds \tag{3}$$

for $N=0,1,\ldots,K$, where $K=3$ in the present theory. The balance law in equation (3) is similar to but somewhat more general than that proposed by Green and Naghdi (1986). A complete set of balance laws for the directed surface must also include balance laws for mass and moment of momentum. These equations, however, are identically satisfied in the present context using the constitutive equations and incompressibility constraint to be introduced presently (see Marshall, 1987), and so the mass and moment of momentum equations are not listed here.

The scalar fields y_{MN} and a_{MN} and the vector field $\mathbf{v}_{MN}$ in (3) are defined in terms of the velocity fields $\mathbf{w}_N$ and the geometry of the three-dimensional wake flow in (A9) of the Appendix. Also, $\mathbf{m}_0$ is the force vector, $\mathbf{m}_N$ ($N=1,2,3$) are the director force vectors at the curve $\partial\mathcal{P}$, ℓ_0 is the assigned force vector, ℓ_N are the director assigned force vectors, and $\mathbf{k}_N$ ($N=1,2,3$) are the internal director forces (such that $\mathbf{k}_0=\mathbf{o}$).

By usual procedures, from (3) we obtain

$$\mathbf{m}_N=\mathbf{M}_{N\alpha}\nu_\alpha \quad (N=0,1,2,3). \tag{4}$$

With use of the two-dimensional analogue of the divergence theorem, as well as the mass balance law for the directed surface and (4) whenever relevant, the local form of (3) is obtained as

$$\frac{\partial}{\partial t}\left(\sum_{M=1}^{K}\rho y_{MN}\mathbf{w}_M\right)+\frac{\partial}{\partial x_\alpha}\left(\sum_{M=1}^{K}\rho\mathbf{w}_M\mathbf{v}_{MN}\cdot\mathbf{e}_\alpha\right) - \sum_{M=1}^{K}\rho\mathbf{w}_M a_{MN}=\rho\ell_N-\mathbf{k}_N+\mathbf{M}_{N\alpha,\alpha} \tag{5}$$

for $N=0,1,\ldots,K$. In (5) and in the remainder of this paper, the notation $(\ )_{,\alpha}$ or $(\ )_{,t}$ is sometimes used to designate partial differentiation with respect to x_α or t, respectively. It is often convenient to combine the inertia terms on the left-hand side of (5) together into a single term $\rho\mathbf{Y}_N$, defined in (A10), to write the director momentum equation in the more concise form

$$\rho\mathbf{Y}_N=\rho\ell_N-\mathbf{k}_N+\mathbf{M}_{N\alpha,\alpha}. \tag{6}$$

It is argued in the Appendix that the condition of incompressibility in the three-dimensional theory imposes certain constraints on the director velocities, namely

$$\begin{aligned}&\ell_{,\alpha}w_{1\alpha}+\ell^2w_{23}=0, \quad w_{1\alpha,\alpha}+\ell w_{23}=0,\\ &\ell_{,1}w_{21}+\ell^2w_{33}=0, \quad 2\ell w_{13}=\ell w_{33}+w_{21,1}.\end{aligned} \tag{7}$$

The response functions $\mathbf{k}_N$ and $\mathbf{M}_{N\alpha}$ in (6) are written as the sum of indeterminate (or constraint) parts $\overline{\mathbf{k}}_N$ and $\overline{\mathbf{M}}_{N\alpha}$ and determinate parts $\hat{\mathbf{k}}_N$ and $\hat{\mathbf{M}}_{N\alpha}$. With the usual assumption that the constraints (7) are workless, forms for the constraint responses may be obtained as

$$\overline{m}_{0\alpha\beta,\beta}=r_{\alpha\beta,\beta},\ \overline{m}_{03\beta,\beta}=r_{33,3},\ -\overline{k}_{1\alpha}+\overline{m}_{1\alpha\beta,\beta}=p_{1,\alpha}-p_2\ell_{,\alpha},$$
$$-\overline{k}_{13}+\overline{m}_{13\beta,\beta}=p_3,\ -\overline{k}_{23}+\overline{m}_{23\beta,\beta}=-3\ell p_1 \tag{8}$$

where $r_{\alpha\beta}$, r_{33}, p_1, p_2 and p_3 are Lagrange multipliers and $m_{Ni\alpha}$ is defined by

$$m_{Ni\alpha}=\mathbf{M}_{N\alpha}\cdot\mathbf{e}_i. \tag{9}$$

It is shown in the Appendix that these Lagrange multipliers may be identified with various weighted integrals of the pressure defect in the wake, as given by (A12), and furthermore, that these variables must satisfy

$$p_1=\ell p_2/2,\ r_{33,3}=0. \tag{10}$$

Using (8)–(10), the functions p_1, p_2, and p_3 may be eliminated in (6) to obtain the momentum equations for the wake flow as

$$\rho Y_{0\alpha}=\rho\ell_{0\alpha}+\hat{m}_{0\alpha\beta,\beta}+r_{\alpha\beta,\beta}, \tag{11}$$

$$\rho Y_{03}=\rho\ell_{03}+\hat{m}_{03\beta,\beta}, \tag{12}$$

$$\rho\left\{Y_{1\alpha}+\frac{1}{3}\ell^2(Y_{23}/\ell^3)_{,\alpha}\right\}=\rho\left\{\ell_{1\alpha}+\frac{1}{3}\ell^2(\ell_{23}/\ell^3)_{,\alpha}\right\} - \left\{\hat{k}_{1\alpha}+\frac{1}{3}\ell^2(\hat{k}_{23}/\ell^3)_{,\alpha}\right\}+\left\{\hat{m}_{1\alpha\beta,\beta}+\frac{1}{3}\ell^2(\hat{m}_{23\beta,\beta}/\ell^3)_{,\alpha}\right\}. \tag{13}$$

The five velocity components $w_{1\alpha}$, w_{13}, w_{23}, and w_{33} can be determined from the constraints (7) and one of (13) with $\alpha=2$, the remaining velocity component w_{21} can be determined from (12) and the decay coefficient ℓ can be determined from the other of (13) with $\alpha=1$. The two equations (11) are used to determine $r_{\alpha\beta,\beta}$, which is necessary in order to estimate the forces on the rod in Section 3.

For two-dimensional flow in the x-z plane, w_{23} can be eliminated in the constraints $(7)_{1,2}$ and the resulting equation integrated to yield

$$w_{11}=\ell C \tag{14}$$

where the constant C is related to the drag on the rod (see Section 3). Also, in what follows it is convenient to define two additional functions of the variables w_{21} and ℓ as follows:

$$w=w_{21}/\ell,\ \Omega=w_{21}/\ell^2. \tag{15}$$

From the constraints $(7)_{3,4}$ and the definition $(15)_1$, the velocity w_{13} is found to satisfy the relation

$$w_{13}=w_{,x}/2. \tag{16}$$

It should also be noted that in the special case of symmetric wakes for which $w_{21}=w=0$ for all x and t, equation (12) is satisfied identically.

Based on the discussion in the Appendix, constitutive equations for the response coefficients $\mathbf{Y}_N$, $\hat{\mathbf{k}}_N$, and $\hat{\mathbf{m}}_N$ as functions of ℓ and w are obtained and are recorded here for the special case of two-dimensional flow in the x-z plane as follows:

$$Y_{01}=\sqrt{\pi}\left\{C_{,t}+\frac{\sqrt{2}}{8}(\ell w^2)_{,x}+\frac{\sqrt{2}}{2}C^2\ell_{,x}\right\},$$

$$Y_{03}=\sqrt{\pi}\left\{\frac{1}{2}\left(\frac{w}{\ell}\right)_{,xt}+\frac{1}{2}U\left(\frac{w}{\ell}\right)_{,xx}+\frac{\sqrt{2}}{4}C\left[\ell\left(\frac{w}{\ell}\right)_{,x}\right]_{,x}\right\},$$

$$Y_{11}=\sqrt{\pi}\left\{\frac{1}{2}C_{,t}-C\frac{\ell_{,t}}{\ell}-UC\frac{\ell_{,x}}{\ell}+\frac{\sqrt{2}}{8}C^2\ell_{,x} +\frac{3\sqrt{2}}{32}\ell_{,x}w^2+\frac{\sqrt{2}}{16}\ell ww_{,x}\right\},$$

$$\begin{aligned}Y_{23}=\sqrt{\pi}\Bigg\{&-\frac{3}{4}C_{,t}\frac{\ell_{,x}}{\ell^2}-\frac{3}{4}C\frac{\ell_{,xt}}{\ell^2}+\frac{15}{4}C\frac{\ell_{,x}\ell_{,t}}{\ell^3}\\ &-\frac{3}{4}CU\frac{\ell_{,xx}}{\ell^2}+\frac{15}{4}CU\frac{(\ell_{,x})^2}{\ell^3}-\frac{3\sqrt{2}}{32}C^2\frac{\ell_{,xx}}{\ell}\\ &+\frac{3\sqrt{2}}{32}C^2\frac{(\ell_{,x})^2}{\ell^2}+\frac{3\sqrt{2}}{64}ww_{,xx}-\frac{3\sqrt{2}}{64}(w_{,x})^2-\frac{3\sqrt{2}}{32}ww_{,x}\frac{\ell_{,x}}{\ell}\\ &+\frac{15\sqrt{2}}{128}w^2\frac{(\ell_{,x})^2}{\ell^2}-\frac{15\sqrt{2}}{128}w^2\frac{\ell_{,xx}}{\ell}\Bigg\};\end{aligned} \tag{17}$$

$$\begin{aligned}&\hat{k}_{01}=\hat{k}_{03}=0,\\ &\hat{k}_{11}=\sqrt{\pi}\mu\left\{-2C\ell^2-C\frac{\ell_{,xx}}{\ell}-C\frac{(\ell_{,x})^2}{\ell^2}\right\},\\ &\hat{k}_{23}=\sqrt{\pi}\mu\left\{\frac{3}{2}C\ell_{,x}-\frac{9}{4}C\frac{\ell_{,x}\ell_{,xx}}{\ell^3}+\frac{45}{4}C\frac{(\ell_{,x})^3}{\ell^4}\right\};\end{aligned} \tag{18}$$

$$\hat{m}_{011}=0,\ \hat{m}_{031}=\frac{1}{2}\sqrt{\pi}\mu\left(\frac{w}{\ell}\right)_{,xx},\ \hat{m}_{111}=-2\sqrt{\pi}\mu C\frac{\ell_{,x}}{\ell},$$
$$\hat{m}_{231}=\sqrt{\pi}\mu\left\{-\frac{3}{2}C\ell-\frac{3}{4}C\frac{\ell_{,xx}}{\ell^2}+\frac{15}{4}C\frac{(\ell_{,x})^2}{\ell^3}\right\}, \tag{19}$$

where $U(=c_1)$ is used as the more common designation of the "free stream" velocity.

3 Effects of Forces Acting on the Rod

The forces acting on the rod influence various structural aspects of the rod's wake. We recall at this time the model for the wake described in Section 1 in which the rod-like body is regarded as a curve of discontinuity on the directed surface s which models the wake such that the outward unit normal of this curve on the surface s is simply $\mathbf{e}_1$. It is assumed, for simplicity in this section, that the rod is fixed in space and that the flow is two-dimensional in the x-z plane.

From the global form of the ordinary ($N=0$) momentum equation (3), the jump in momentum across the rod is obtained using arguments similar to those of Green and Naghdi (1987) as

$$\mathbf{F} = -\mathbf{S}_0 = -\left[\!\!\left[\sum_{M=1}^{K} \rho \mathbf{w}_M \mathbf{v}_{M0} \cdot \mathbf{e}_1 - \mathbf{m}_0 \right]\!\!\right]. \tag{20}$$

The notation in (20) is defined by $[\![f]\!] = f^+ - f^-$ for any quantity f. The force $\mathbf{F}$ acting on the rod per unit length by the fluid is identified in (20) with the rate of supply $-\mathbf{S}_0$ of ordinary ($N=0$) momentum to the curve of discontinuity from the fluid.

If we suppose the curve of discontinuity which models the rod to be located along the line $x=0$ on the surface s, then using (A5), (A6), (A9), and (8), we find that within the wake (i.e., for $x>0$)

$$\rho \mathbf{v}_{10} \cdot \mathbf{e}_1 = \frac{\rho\sqrt{\pi}}{\ell}\left(U + \frac{\sqrt{2}}{2} C\ell\right),\ \rho \mathbf{v}_{20} \cdot \mathbf{e}_1 = \frac{\sqrt{2\pi}}{8}\rho w,$$

$$\rho \mathbf{v}_{30} \cdot \mathbf{e}_1 = \frac{\rho\sqrt{\pi}}{2\ell}\left(U + \frac{\sqrt{2}}{4} C\ell\right),$$

$$\mathbf{m}_0 = \mathbf{M}_{01} = \overline{m}_{0i1}\mathbf{e}_i + \hat{m}_{0i1}\mathbf{e}_i = r_{11}\mathbf{e}_1 + \frac{\sqrt{\pi}}{2}\mu\left(\frac{w}{\ell}\right)_{,xx}\mathbf{e}_3. \tag{21}$$

Upstream of the body (i.e., for $x<0$), the velocity in three-dimensional theory is assumed to be given simply by $\mathbf{v}^* = c_\alpha \mathbf{e}_\alpha$, so that

$$\mathbf{v}_{10} = \mathbf{v}_{20} = \mathbf{v}_{30} = \mathbf{m}_0 = \mathbf{o}. \tag{22}$$

Using (20)–(22), expressions for drag and lift forces acting on a unit length of the rod are obtained as

$$D = F_1 = -\sqrt{\pi}\rho C\left(U + \frac{\sqrt{2}}{2}C\ell\right) - \frac{\sqrt{2\pi}}{8}\rho\ell w^2 + r_{11},$$

$$L = F_3 = -\frac{\rho\sqrt{\pi}}{2}\left(\frac{w}{\ell}\right)_{,x}\left(U + \frac{\sqrt{2}}{2}C\ell\right) + \frac{\sqrt{\pi}}{2}\mu\left(\frac{w}{\ell}\right)_{,xx}, \tag{23}$$

where all variables in (23) are evaluated at $x=0^+$.

For two-dimensional flow in the x-z plane, an equation for the Lagrange multiplier r_{11} can be obtained from (11) using the constitutive equations $(17)_1$ and $(19)_1$ with $\boldsymbol{\ell}_0 = \mathbf{o}$. Integrating this equation over x and requiring the integrated pressure difference in $(\mathrm{A}12)_1$ to approach zero as $x \to \infty$, we obtain a solution for r_{11} immediately behind the body as

$$r_{11}|_{x=0^+} = \frac{\sqrt{2\pi}}{2}\rho\ell\left(C^2 + \frac{1}{4}w^2\right). \tag{24}$$

Using (23)–(24), the drag and lift forces per unit length on the rod are obtained as

$$D = -\sqrt{\pi}\rho C U,\ L = -\frac{\rho\sqrt{\pi}}{2}\left(\frac{w}{\ell}\right)_{,x}\left(U - \frac{1}{\sqrt{2\pi}}\frac{D\ell}{\rho U}\right) + \frac{\sqrt{\pi}}{2}\left(\frac{w}{\ell}\right)_{,xx}. \tag{25}$$

The results (25) can be used to obtain a value for the constant C and an upstream initial condition for the variable w.

4 Applications

The theory developed in Sections 2–3 is applied in the present section to obtain simple analytical solutions to several fundamental problems pertaining to viscous fluid wakes. As is evident from the momentum equations (11)–(13) and the constitutive equations (17)–(19) for $\mathbf{Y}_N$, $\hat{\mathbf{k}}_N$ and $\hat{\mathbf{m}}_N$, the governing equations for the wake are in general extremely difficult to solve. However, in certain regions of the wake and under certain simplifying assumptions it is possible to obtain approximate solutions of these equations. For simplicity, all of the specific solutions given in this section are developed for wakes behind rigid, fixed cylinders of infinite length. The "trailing edge" of the cylinder is assumed to correspond with the line $x=0$ on the surface $\bar{s}$, and the ordinary and director applied body forces $\boldsymbol{\ell}_N$ are set equal to zero. Additionally, the velocity $\mathbf{c}$ at infinity is assumed to be orthogonal to the cylinder's axis, so that we may write

$$\mathbf{c} = U\mathbf{e}_1, \tag{26}$$

where $U(=c_1)$ is assumed to be a constant. Because all of the problems discussed in this section are concerned with two-dimensional flow in the x-z plane, the middle surface $\bar{s}$ is referred to as the "centerline."

When obtaining specific solutions for fluid flow problems, it is often illuminating to express the results as functions of certain nondimensional parameters. In this section, various such parameters will be used which, for convenience, are defined collectively as follows:

nondimensional downstream distance $= x' = x/d$,
Reynolds number $= R = \rho U d/\mu$,
nondimensional streamwise centerline velocity
$= u = (U + w_{11})/U$,
drag coefficient $= C_d = D/\tfrac{1}{2}\rho U^2 d$,
lift coefficient $= C_L = \mathrm{amp}(L)/\tfrac{1}{2}\rho U^2 d$,
Strouhal number $= S = \omega d/2\pi U$,
nondimensional wave number $= \theta = kd$. (27)

In (27), d is the "projected" diameter of the cylinder and $\omega/2\pi$ and k are the frequency and wave number, respectively, of the vortex street. The lift coefficient is proportional in $(27)_5$ to the amplitude of the oscillating lift force acting on the cylinder per unit length. For vortex street wakes, the value of u is based on the mean value of w_{11}. Using $(25)_1$ and $(27)_4$, the coefficient C is found to be related to the drag coefficient such that

$$C_D = -2\sqrt{\pi}\frac{C}{Ud}. \tag{28}$$

The prime attached to x in $(27)_1$ may at times be dropped if the nondimensional character of x is clear.

4.1 Middle and Far Regions of Steady Symmetric Wakes. The middle and far regions of the wake (occupying a region $x_o < x < \infty$) are characterized by the condition $0 < \epsilon \ll 1$, where ϵ is a parameter of smallness defined by

$$\epsilon = \max_{x_o < x < \infty}\left|\frac{\ell_{,x}}{\ell^2}\right|. \tag{29}$$

To leading order in ϵ, the momentum equation (13) with $\alpha = 1$ and the expressions (17)–(19) can be combined to yield an equation for ℓ for symmetric wakes (for which $w=0$ and (12) is satisfied identically) of the form

$$-U\frac{\ell_{,x}}{\ell} + \frac{\sqrt{2}}{8}C\ell_{,x} = \frac{2\mu}{\rho}\ell^2. \tag{30}$$

Solving for ℓ in (30) and using the results (14) and (28) and the definitions $(27)_{1\text{-}4}$, a solution for the centerline velocity u in the middle and far regions of the wake is obtained as

$$u = 1 - \frac{\sqrt{2}}{64\pi}\frac{C_D^2 R}{2x + A_1}\left[1 + \left\{1 + \frac{256\pi}{C_D^2 R}(2x + A_1)\right\}^{1/2}\right], \tag{31}$$

where the prime on x in $(27)_1$ has been dropped. As $x \to \infty$, the centerline velocity in (31) and the solution for ℓ from (30) become

$$u \to 1 - \frac{C_D}{4\sqrt{\pi}}\left(\frac{R}{x}\right)^{1/2},\ \ell d \to \frac{1}{2}\left(\frac{R}{x}\right)^{1/2}, \tag{32}$$

where $(32)_1$ is identical (in slightly different notation) to the far wake solution recorded by Chang (1961) when the latter is evaluated along the centerline.

For purposes of illustration, we assume that a stagnation

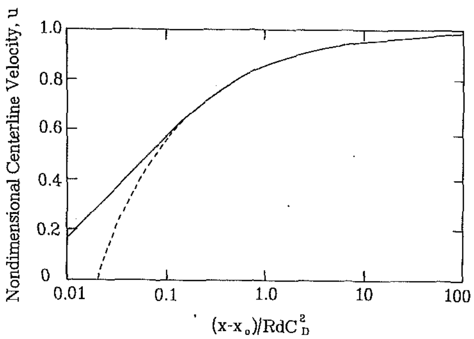


Fig. 2 Centerline velocity prediction in the middle and far regions of steady symmetrical wakes behind cylinders of arbitrary cross-sectional shape, as predicted from equation (31) and equation $(32)_1$ and plotted with a solid and dotted line, respectively

point exists at $x=x_o$ such that $u(x_o)=0$ and use that point as a boundary condition to determine the constant A_1 in (31). The resulting solution for u is shown in Fig. 2 as a function of the nondimensional parameter grouping $(x-x_o)/dRC_D^2$, where for clarity the dimensional distance x is again used in Fig. 2. The far wake solution (32) is seen to lie extremely close to the more general solution (31) for values of x such that $(x-x_o)/dRC_D^2>0.1$.

4.2 Far Region of Steady Antisymmetric Wakes. For antisymmetric wakes, the body is assumed to be subject only to lift and not to drag, so that from (28) the constant C must vanish. In the far region of an antisymmetric wake, the governing equations (12)–(13), along with (17)–(19), reduce to

$$-U\frac{\ell_{,x}}{\ell}=\frac{2\mu}{\rho}\ell^2,\quad \frac{1}{2}\rho U\left(\frac{w}{\ell}\right)_{,xx}=\frac{1}{2}\mu\left(\frac{w}{\ell}\right)_{,xxx}. \tag{33}$$

Equation $(33)_1$ is obtained by first evaluating the appropriate equation for the far region (i.e., letting $\epsilon\to 0$) and then restricting the flow to the antisymmetric case by letting $C\to 0$. Integrating $(33)_1$ gives a solution for ℓ which is the same as the far wake solution $(32)_2$ obtained for steady symmetric wakes in Subsection 4.1. Since w is assumed to remain bounded as $x\to\infty$, a solution for w from $(33)_2$ is given by

$$w=A_2\ell=-\frac{2}{\sqrt{\pi}}\left(\frac{\mu}{\rho^2U^2}\right)L\ell, \tag{34}$$

where the coefficient A_2 is chosen such that the resulting solution for velocity is identical to the asymptotic far wake solution of Chang (1961) for the antisymmetric case. It is not possible in this case to evaluate A_2 from the expression $(25)_2$ for the lift since the solution (34) does not remain valid as x approaches zero.

4.3 Near Region of Steady Symmetric Wakes. For small values of x in a symmetric wake at sufficiently high values of the Reynolds number, the effect of viscosity in the near wake region is assumed to be negligible. (This assumption is also made in many previous models of the flow in the near wake region, as discussed by Wu (1972).) Taking the limit $\mu\to 0$, the momentum equation (13) together with (17)–(19) and with $w=0$ becomes

$$-U\ell_{,x}+\frac{\sqrt{2}}{8}C\ell\ell_{,x}+\frac{1}{4}\ell^3\left\{-\frac{\ell_{,xx}}{\ell^6}\left(U+\frac{\sqrt{2}}{8}C\ell\right)+\frac{5(\ell_{,x})^2}{\ell^6}\left(U+\frac{\sqrt{2}}{40}C\ell\right)\right\}_{,x}=0. \tag{35}$$

If we divide (35) by ℓ^3, integrate over x, multiply the resulting equation by an integrating factor $\{-8\ell_{,x}(U+\sqrt{2}C\ell/8)^7/\ell^6\}$, integrate over x again, and then write the result in terms of u after using (14), $(27)_{1\text{-}4}$ and (28), we obtain an equation for the streamwise centerline velocity of the form

$$\frac{C_D^2}{4\pi}\left\{1-\frac{\sqrt{2}}{8}\left(1-u\right)\right\}^8\left(u_{,x}\right)^2+h(u)=0, \tag{36}$$

where h is defined by

$$h(u)=(1-u)^{10}A_4+(1-u)^{10}\int\frac{\left\{1-\frac{\sqrt{2}}{8}\left(1-u\right)\right\}^7}{(1-u)^7}\times\{4+\sqrt{2}(1-u)+8A_3(1-u)^2\}du. \tag{37}$$

A_3 and A_4 are constants of integration, and the prime on x in $(27)_1$ is again dropped. The solution for u from (36) must satisfy a boundary condition at the trailing edge of the cylinder of the form

$$u|_{x=0}=1-a, \tag{38}$$

where $a=1$ for an impermeable cylinder and $a=0$ for a perfectly permeable cylinder.

For appropriate values of the constants A_3 and A_4, the solution for u from (36) has the form of a solitary wave behind the cylinder. The minimum value of u in the near wake region is obtained by setting $u_{,x}=0$ in (36) and evaluating the lowest positive real-valued root of the function $h=h(u)$ in (37). The constants A_3 and A_4 must be specified as functions of other constants characterizing the wake, such as R, C_D and possibly also the shape of the body; unfortunately, however, available experimental data for the near wake region is not sufficient to determine reliable expressions for these constants. The results do, though, indicate that the theory yields a form for u in the near wake which is consistent with the common observation of a pair of recirculating eddies.

4.4 Onset of Instability in the Far Wake. The onset of instability in the far region of an initially steady wake is studied here by considering the growth or decay of a perturbation in the variable Ω, which is related to the antisymmetric streamwise velocity component w_{21} by $(15)_2$, of the form

$$\Omega=f(x,t)\cos(kx-\omega t+\phi). \tag{39}$$

In (39), f is the amplitude of oscillation, k is the wave number, ω is the frequency, and ϕ is the phase angle. In the present instability analysis, k, ω, and ϕ are all assumed to be prescribed constants. Using the velocity representation (A1), an oscillation of the form (39) for appropriate values of f is found to have the form of two staggered rows of vortices when viewed in the three-dimensional flow field (see also Fig. 6). An equation governing the evolution of Ω is obtained from (12), $(15)_2$, and (17)–(19), after integrating over x, as

$$\rho\left(\Omega_{,t}+U\Omega_{,x}+\frac{\sqrt{2}}{2}C\ell\Omega_{,x}\right)=\mu\Omega_{,xx}. \tag{40}$$

A solution for the oscillation amplitude f is obtained by substituting (39) into (40), using the steady wake solution $(32)_2$ for ℓ (which is valid for sufficiently small f), and solving the resulting system of differential equations to get

$$f(x,t)=A_5\exp\left\{-\frac{\rho(\omega-kU)x}{2k\mu}+\frac{\sqrt{2}}{4}\frac{\rho C}{\mu}\int_0^x\ell(\xi)d\xi\right\}\cdot\exp\left[\left\{\frac{\rho\omega^2}{4\mu k^2}-\frac{\mu k^2}{\rho}-\frac{\rho}{4\mu}\left(U+\frac{\sqrt{2}}{2}C\ell\right)^2+\frac{\sqrt{2}}{4}C\ell_{,x}\right\}t\right], \tag{41}$$

where A_s is a constant of integration.

A "characteristic" coordinate $x_c=x-\omega t/k$ is now introduced such that the wave has zero propagation speed with respect to the characteristic coordinate system. The derivative of f in (41), with respect to t, keeping x_c fixed and evaluated at $x_c=0$, i.e., the rate of change of wave amplitude as we follow a single crest, is obtained from (41) after writing $f=f(x,t)=\bar{f}(x_c,t)$ as follows:

$$\frac{1}{\bar f}\frac{\partial \bar f}{\partial t}\bigg|_{x_C=0} = -\frac{\rho}{4\mu k^2}\left(\omega - Uk - \frac{\sqrt{2}}{2}kC\ell\right)^2 - \frac{\mu k^2}{\rho} + \frac{\sqrt{2}}{4}xC\ell_{,xx} + \frac{\sqrt{2}}{4}\left(1 - \frac{\rho U x}{\mu} - \frac{\sqrt{2}}{2}\frac{\rho x C\ell}{\mu}\right)C\ell_{,x}. \tag{42}$$

The wake is unstable, such that a wave crest grows with time as it travels in the x-direction, whenever the time derivative of $\bar f$ in (42) is positive. Using the nondimensional variables (27) and the solution $(32)_2$ for ℓ, the result (42) yields the following criterion for onset of oscillations in the far wake:

$$1 > \frac{4\pi S}{\theta} + \frac{1}{2}\frac{1}{Rx} + 16\sqrt{2\pi}\frac{\theta\sqrt{x}}{C_D R^{5/2}} + 4(2\pi)^{3/2}\frac{S(2\pi S\theta - 1)^2}{C_D\theta^3 R^{1/2}}\sqrt{x}, \tag{43}$$

where the prime on x in $(27)_1$ is again dropped.

The drag coefficient C_D can be expressed as a function of R and the shape of the cylinder, and the parameters S and θ specify characteristics of the unstable waves. In general, the determination of the region of instability in the far wake entails finding the range of x for which, at given values of R and S, a value of θ can be found such that (43) is satisfied. For instance, as $R \to 0$ the term $1/2Rx \to \infty$ for any finite x, and so the wake is stable for all waves. As $x \to \infty$, the last two terms on the right-hand side of (43) approach positive infinity, so that the wake always becomes stable for sufficiently large x. As $R \to \infty$, (43) becomes $1 > 4\pi S/\theta$, which implies that the wake is always unstable in this limit since θ can always be chosen to satisfy the criterion.

The region of instability depends on the relationship assumed between C_D and R, and hence on the shape of the cylinder. For flow past a circular cylinder with Reynolds number in the range $0.1 < R < 1000$, experimental values of the drag coefficient are well fit by the empirical expression (see data presented by Fleischmann and Sallet, 1981)

$$C_D = 10.0\,R^{-3/4} + 0.95. \tag{44}$$

Using a Newton-Raphson iteration method to determine x from (43), with the use of (44), the greatest values of x for which a value of θ can be found such that the wake is marginally stable for specified values of R and S, are plotted in Fig. 3.

Experimental evidence usually indicates the onset of vortex shedding from stationary circular cylinders at about $R = 41$, although oscillations in the far wake are commonly observed somewhat before this value. Under certain circumstances, such as for a slightly vibrating cylinder, the onset of vortex shedding has been observed for Reynolds numbers as high as $R = 48$ and as low as $R = 20$ (see the review by Fleischmann and Sallet, 1981). The lowest values of S observed for wakes behind fixed circular cylinders lie between $S = 0.10$ and $S = 0.15$, where S increases with increasing R until it levels off at about $S = 0.21$. Also, most vortex streets (even for low values of R) are observed to exhibit oscillations in the velocity field for distances of 70 to 100 diameters downstream of the cylinder.

For an instability region of $x/d < 80$ and a Strouhal number of $S = 0.13$, the results in Fig. 3 predict onset of instability in the far wake at about $R = 30$, and the shedding of vortices from the near wake would be expected to occur at somewhat greater values of R. This result agrees well with the observations of Taneda (1956), who found the onset of small oscillations in the far wake at $R = 30$ while the near wake was still observed to be completely stable. It is recalled that the effects of the near and middle regions are neglected in the calculations leading to Fig. 3 and will tend to stabilize the wake for sufficiently small values of x and R. The results in Fig. 3 also indicate that if the Strouhal number is slightly reduced (for instance, by vibrating the cylinder slightly at appropriate frequencies), the

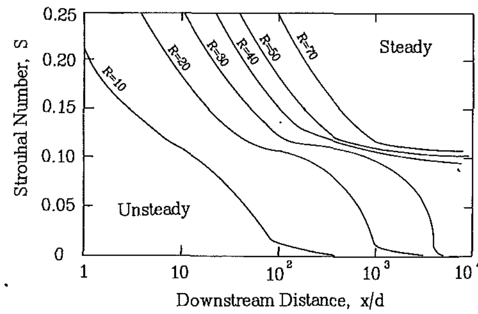


Fig. 3 Curves of marginal stability in the far region of the wake behind a circular cylinder, as predicted from equation (43)

Reynolds number at the onset of wake instability is also reduced. For instance, if $S = 0.10$ and again for an instability region of $x/d < 80$, Fig. 3 indicates that the wake will become unstable at $R = 20$.

4.5 Fully Developed Vortex Streets. A fully developed vortex street is considered here to be a periodic flow pattern for which the amplitude of oscillation of ℓ and Ω at any location x behind the rod is constant in time. In addition to the smallness measure ϵ in (29), two parameters σ and δ are defined by

$$\sigma = \max_{x_0 < x < \infty} |\Omega\ell^2/U|, \quad \delta = \max_{x_0 < x < \infty} |k_{,x}/k^2|, \tag{45}$$

where Ω is again assumed to be of the form (39) with ω and ϕ constant but $k = k(x,t)$. An approximate solution is obtained for the vortex street in which it is assumed that $\epsilon \ll 1$, $\delta \ll 1$ and that both ϵ and δ are smaller than or of the same order as σ. These approximations correspond physically to assuming that the downstream variation of wave number and wake "width" (evaluated over a distance corresponding either to the wavelength or the wake "width," respectively), are small compared to their local values, and that both of these quantities, when nondimensionalized by their local values, are smaller than or of the same order as the amplitude of the antisymmetric streamwise velocity (which is responsible for driving the staggered vortex motion) divided by the free stream velocity. Put another way, it is consistently assumed that changes in the flow pattern due to viscous effects occur over a much longer downstream distance than the wavelength of the vortex street. We do not, however, assume here that $\sigma \ll 1$, which would be equivalent to restricting the analysis to very weak vortex streets.

The leading order equations in ϵ for Ω and ℓ, obtained from (12)–(19), are given by (40) and

$$\rho\left\{\frac{1}{2}C_{,t} - C\frac{\ell_{,t}}{\ell} - UC\frac{\ell_{,x}}{\ell} + \frac{\sqrt{2}}{8}C^2\ell_{,x} + \frac{5\sqrt{2}}{32}\ell^2\Omega^2\ell_{,x} + \frac{\sqrt{2}}{16}\ell^3\Omega\Omega_{,x} + \frac{\sqrt{2}}{64}\ell\Omega\Omega_{,xxx} - \frac{\sqrt{2}}{64}\ell\Omega_{,x}\Omega_{,xx} - \frac{\sqrt{2}}{64}\ell_{,x}(\Omega_{,x})^2 - \frac{3\sqrt{2}}{64}\ell_{,x}\Omega\Omega_{,xx}\right\} = 2\mu C\ell^2. \tag{46}$$

Substituting the form (39) into (40) and using the asymptotic simplifications stated in the previous paragraph, a solution for Ω is obtained as

$$\Omega = A\cos(kx - \omega t + \phi)\exp\left\{-\frac{\rho(\omega - \gamma Uk)x}{2\mu k}\right\} \times \exp\left[\left\{\frac{\rho\gamma U(\omega - \gamma Uk)}{2\mu k} + \frac{\rho(\omega - \gamma Uk)^2}{4\mu k^2} - \frac{\mu k^2}{\rho}\right\}t\right], \tag{47}$$

where A is a constant of integration, $\gamma = \gamma(x,t)$ is a parameter defined by

$$\gamma = 1 - \frac{\sqrt{2}}{2}(1 - u), \tag{48}$$

and u is the nondimensional mean streamwise centerline velocity defined in $(27)_3$. For a fully developed vortex street the amplitude of Ω must be constant at a given value of x, so (47) implies that

$$\frac{\rho\gamma U(\omega-\gamma Uk)}{2\mu k}+\frac{\rho(\omega-\gamma Uk)^2}{4\mu k^2}-\frac{\mu k^2}{\rho}=0. \tag{49}$$

Solving (49) for k, using the nondimensional variables (27) and denoting the resulting nondimensional wave number for a stable vortex street by θ_s, we obtain

$$\theta_s=\frac{\sqrt{2}}{4}\gamma R\left\{-1+\left(1+\frac{64\pi S^2}{R^2\gamma^4}\right)^{1/2}\right\}^{1/2}\simeq\frac{2\pi S}{\gamma}. \tag{50}$$

The result (50) thus provides a criterion for stability of a vortex street which is valid for all downstream distances.

The variable $\ell=\ell(x,t)$ in (46) is written as the sum of a time-averaged mean part $\bar{\ell}=\bar{\ell}(x)$ and an oscillating part $\ell'=\ell'(x,t)$, where we assume that $\ell'\ll\bar{\ell}$. Substituting (47) into (46), averaging the resulting equation over one period of oscillation while neglecting higher-order terms in $\ell'/\bar{\ell}$ and using (14), (28) and the nondimensional variables (27), we obtain an equation for the downstream evolution of u as

$$\frac{RC_D^2}{4\pi}u_{,x}\left[1+\frac{\sqrt{2}}{8}(1-u)+\frac{\sqrt{2}}{16}\frac{\pi A^2}{C_D^2}(1-u)\left\{\theta_s^2+\frac{20\pi}{C_D^2}(1-u)^2\right\}e^{-2\alpha x}\right]=(1-u)^3, \tag{51}$$

where the parameter α in (51) is defined by

$$\alpha=\frac{1}{2}R\left(\frac{2\pi S}{\theta_s}-\gamma\right)\simeq\frac{\pi^2}{4}\frac{S^2}{\gamma^3 R}. \tag{52}$$

Consistent with the previously stated asymptotic assumptions we note that $\alpha\ll\theta_s$ for the range of S and R found in vortex streets. The constant A in (51) and (28) may be evaluated from the result $(25)_2$ for lift and the boundary condition (38) on u. Substituting (47) into $(25)_2$ and again using the previously stated asymptotic simplifications, an expression for A is obtained as

$$A=\left.\frac{C_L}{\sqrt{\pi}\gamma\theta_s}\right|_{x=0^+}\simeq\frac{C_L}{2\pi^{3/2}S}. \tag{53}$$

The differential equation (51), together with (48), (50), (52), and the boundary condition (38) for u, can be used to solve for u in the vortex street given prescribed values for the constants R, S, C_D, and C_L which characterize the wake. A sample calculation for u in the wake behind a circular cylinder is shown in Fig. 4 for $R=150$, $S=0.183$, $C_D=1.18$, and $C_L=0.50$, where the values of S, C_D, and C_L chosen here are based on various experimental observations for $R=150$ as collected by Fleischmann and Sallet (1981). Also plotted in Fig. 4 are the vortex street wave numbers, calculated from (50), and the nondimensional vortex velocity u_s, defined by

$$u_s=\frac{\omega}{Uk_s}=\frac{2\pi S}{\theta_s}. \tag{54}$$

The equation (51) for u indicates that $u\to 1$ as $x\to\infty$, as expected, so that from (50) the vortex street wave number has the limiting value

$$\theta_s\to\frac{\sqrt{2}}{4}R\left\{-1+\left(1+\frac{64\pi S^2}{R^2}\right)^{1/2}\right\}^{1/2}\simeq 2\pi S \tag{55}$$

as $x\to\infty$. The limit (55) provides the minimum value of θ_s for a given value of S, which in turn is a function of R and the cross-sectional shape of the cylinder. Using experimentally determined values of S for wakes behind circular cylinders, the minimum value of θ_s calculated from (55) is plotted in Fig. 5 along with experimental data of Taneda (1959) for various values of R. The data of Taneda (1959) were taken in the range $x/d<50$, although Taneda does not specify the exact location at which his data were taken, so the experimental values are expected to be slightly higher than the minimum value of θ_s predicted by (55). From Fig. 5, we see that for low values of R, the street has nearly attained the asymptotic far field wave number given by (55) at the point of measurement, whereas for large values of R, the street takes longer for the wave number to approach its asymptotic value.

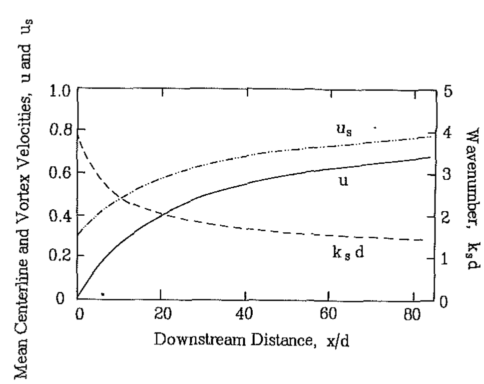


Fig. 4 Downstream development of the mean centerline velocity u, wave number θ_s, and vortex street velocity u_s for a vortex street behind a circular cylinder with $R=150$, $S=0.183$, $C_D=1.18$, $C_L=0.50$, as predicted by equations (50)–(54)

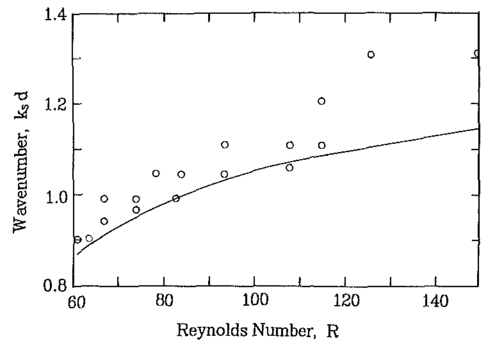


Fig. 5 A lower bound for wave number in the vortex street formed behind a circular cylinder as predicted (shown by a solid line) by equation (55) and as given by the experimental data of Taneda (1959)

From (47) and the definitions (15), the antisymmetric streamwise velocity component w_{21} is obtained as

$$\frac{w_{21}}{U}=\frac{2C_L}{\sqrt{\pi}C_D^2S}e^{-\alpha x}(1-u)^2\cos(\theta_s x-2\pi St+\phi), \tag{56}$$

where t is nondimensionalized by the convective time scale d/U. Substituting (56) into (15)–(16) and again making the same asymptotic simplifications, the cross-stream centerline velocity component w_{13} is obtained as

$$\frac{w_{13}}{U}=-\frac{C_L}{2\pi C_D S}e^{-\alpha x}\theta_s(1-u)\sin(\theta_s x-2\pi St+\phi). \tag{57}$$

It is observed that the condition (50) for vortex street stability is considerably different than the stability criterion of von Kármán (1911), or even later modifications of von Kármán's criterion. In particular, the stability condition (50) is obtained in terms of the wave number of the street rather than the "spacing ratio" of the vortices in the street, as used in the von Kármán criterion. Of course, using the velocity representation (A1) in the alternative derivation of the theory, as well as the results (28) and (56) and the stability criterion (50), the apparent spacing ratio denoted by a, may be estimated as

$$a=\frac{C_D^3e^{\alpha x}S^2}{2(1-u)^2\gamma C_L}. \tag{58}$$

Values of predicted spacing ratio at several downstream lo-

Table 1 **Predicted values of apparent spacing ratio from equation (58) at several downstream locations for the same case as that examined in Fig. 4**

Downstream Distance x/d	Predicted Spacing Ratio from equation (58)
10	0.254
20	0.276
30	0.362
40	0.414
50	0.480

cations are given in Table 1 for the same case as that estimated in Fig. 4. The street was observed by Taneda (1959) to break down at about $x/d=50$, so no spacing ratio estimates for $x/d>50$ are recorded in Table 1. The predicted values of apparent spacing ratio are close to von Kármán's value 0.281 immediately behind the cylinder and increase to 0.48 before the street breaks down. This range of values for spacing ratio, as well as the tendency of the spacing ratio to increase sustantially with downstream distance, is compatible with experimental results of several investigators as summarized by Wille (1960). The gradual predicted increase in wavelength with downstream distance (approaching an asymptotic value at infinity) is also in line with the observation of Wille (1960, p. 280).

Acknowledgment

The author wishes to express his gratitude to Prof. P. M. Naghdi for many helpful discussions during the preparation of this paper. This study was supported by the U.S. Office of Naval Research under Contract N00014-86-K-0057, Work Unit 4322 534 with the University of California, Berkeley.

References

Chang, I-Dee, 1961, "Navier-Stokes Solutions at Large Distances from a Finite Body," *Journal of Mathematics and Mechanics,* Vol. 10, pp. 811-876.

Fleischmann, S. T., and Sallet, D. W., 1981, "Vortex Shedding from Cylinders and the Resulting Unsteady Forces and Flow Phenomenon. Part I," *Shock and Vibration Digest,* Vol. 13, No. 11, pp. 15-24.

Green, A. E., and Naghdi, P. M., 1976, "Directed Fluid Sheets," *Proceedings of the Royal Society of London,* Vol. A347, pp. 447-473.

Green, A. E., and Naghdi, P. M., 1984, "A Direct Theory of Viscous Fluid Flow in Channels," *Archive for Rational Mechanics and Analysis,* Vol. 86, pp. 39-63.

Green, A. E., and Naghdi, P. M., 1986, "A Nonlinear Theory of Water Waves for Finite and Infinite Depths," *Philosophical Transactions of the Royal Society of London,* Vol. A320, pp. 37-70.

Green, A. E., and Naghdi, P. M., 1987, "Further Developments in a Nonlinear Theory of Water Waves for Finite and Infinite Depths," *Philosophical Transactions of the Royal Society of London,* Vol. A324, pp. 47-72.

Hooker, S. G., 1936, "On the Action of Viscosity in Increasing the Spacing Ratio of a Vortex Street," *Proceedings of the Royal Society of London,* Vol. A154, pp. 67-89.

von Kármán, T., 1911, "Über den Mechanismus des Widerstads, den ein bewegter Körper in einer Flüssigkeit erfährt," *Nachrichten der K. Gesellschaft der Wissenschaften zu Göttingen; Mathematisch-Physikalische Klasse,* pp. 509-517.

Kida, S., 1982, "Stabilizing Effects of Finite Core on Kármán Vortex Streets," *Journal of Fluid Mechanics,* Vol. 122, pp. 487-504.

Marshall, J. S., 1987, "Some Theoretical Developments and Predictions for Viscous Fluid Wakes," Ph.D. dissertation, Department of Mechanical Engineering, University of California, Berkeley.

Saffman, P. G., and Schatzman, J. C., 1982a, "Stability of a Vortex Street of Finite Vortices," *Journal of Fluid Mechanics,* Vol. 117, pp. 171-185.

Saffman, P. G., and Schatzman, J. C., 1982b, "An Inviscid Model for the Vortex Street Wake," *Journal of Fluid Mechanics,* Vol. 122, pp. 467-486.

Taneda, S., 1956, "Experimental Investigation of the Wakes Behind Cylinders and Plates at Low Reynolds Numbers," *Journal of the Physical Society of Japan,* Vol. 11, pp. 302-307.

Taneda, S., 1959, "Downstream Development of the Wake Behind Cylinders," *Journal of the Physical Society of Japan,* Vol. 14, pp. 843-848.

Timme, V. A., 1957, "Über die Geschwindigkeitsverteilung in Wirbeln," *Ingenieur-Archiv,* Vol. 25, pp. 205-225.

Wille, R., 1960, "Karman Vortex Streets," *Advances in Applied Mechanics,* Vol. 6, pp. 273-287.

Wu, T. Y., 1972, "Cavity and Wake Flows," *Annual Review of Fluid Mechanics,* Vol. 4, pp. 243-284.

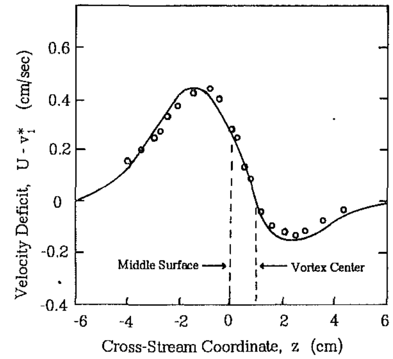


Fig. 6 Downstream velocity variation across a vortex in a vortex street: O, experimental data of Timme (1957, Fig. 17); ———, velocity representation (A1) with $c_2=0$, $C=-0.632$ cm²/sec (set from equation (28)), $\ell=0.38$ 1/cm, $w_{21}=0.68$ cm/sec. Figure 1 of Timme (1957) is used to estimate the apparent location of the vortex center.

APPENDIX

An alternative derivation of the governing equations for the wake from the usual three-dimensional theory is presented in this Appendix. This alternative derivation serves to motivate constitutive equations and certain constraints for the directed fluid sheet model presented in Section 2.

As stated previously in Section 2, the wake is characterized by certain features which serve as independent variables in the theory. In order to derive appropriate equations governing the evolution of these features of the wake flow, it is necessary to assume some representation for the velocity $\mathbf{v}^*$ in the three-dimensional theory as a function of the independent variables of the theory and of the coordinate z normal to the middle surface $\bar{s}$. A representation for $\mathbf{v}^*$ is selected in this paper as

$$v_1^*=c_1+w_{11}e^{-\ell^2z^2}+w_{21}\ell z e^{-\ell^2z^2},\ v_2^*=c_2+w_{21}e^{-\ell^2z^2},$$
$$v_3^*=w_{13}e^{-\ell^2z^2}+w_{23}\ell z e^{-\ell^2z^2}+w_{33}\ell^2z^2e^{-\ell^2z^2}. \quad \text{(A1)}$$

The form (A1) is found to reduce in special cases to the exact far-field solutions for velocity in steady symmetric and antisymmetric wakes. This representation also compares well with data of Timme (1957) for velocity profile through a vortex center in a vortex street, as shown in Fig. 6.

The usual governing equations for incompressible fluid flow in the three-dimensional theory are

$$\rho(\mathbf{v}^*_{,t}+v_i^*\mathbf{v}^*_{,i})=\rho\mathbf{b}+\mathbf{t}_{i,i},\ v^*_{i,i}=0, \quad \text{(A2)}$$

$$\mathbf{t}_i=-p^*\mathbf{e}_i+\sigma_{ij}\mathbf{e}_j,\ \mathbf{t}=n_i\mathbf{t}_i,\ \sigma_{ij}=\mu(v^*_{i,j}+v^*_{j,i}), \quad \text{(A3)}$$

where $\mathbf{t}_i$ are the stress vectors and ρ is the constant fluid density. The stress is composed of a pressure p^*, which is determined by the solution of $(A2)_{1,2}$, and a deviatoric stress response σ_{ij}, which is specified by the constitutive equation $(A3)_3$ for Newtonian fluid flows. The velocity representation (A1) satisfies the condition of incompressibility $(A2)_2$ identically at every point of the three-dimensional flow field only if the independent variables satisfy the restrictions (7) for all x_α and t.

It is convenient to write the representation (A1) for $\mathbf{v}^*$ in terms of the weighting functions $\bar{\lambda}_M$ as

$$\mathbf{v}^*=\sum_{M=0}^{K}\bar{\lambda}_M\mathbf{w}_M\ (K=3), \quad \text{(A4)}$$

where $w_{3\alpha}=w_{03}=w_{22}=0$ and where the weighting functions $\bar{\lambda}_M$ are defined by

$$\bar{\lambda}_0=1,\ \bar{\lambda}_1=e^{-\ell^2z^2},\ \bar{\lambda}_2=\ell z e^{-\ell^2z^2},\ \bar{\lambda}_3=\ell^2z^2e^{-\ell^2z^2}. \quad \text{(A5)}$$

A different set of weighting functions λ_N are also introduced for integration of the balance laws and are defined by

$$\lambda_0=1,\ \lambda_1=\ell^2z^2,\ \lambda_2=\ell^3z^3. \quad \text{(A6)}$$

It is noted that the class of functions from which the weighting functions λ_N and $\overline{\lambda}_M$ may be selected is significantly restricted by the requirement that the constraint response functions determined both by the derivation from (A2)–(A3) and the direct approach in Section 2 be equivalent (see Marshall, 1987).

Assuming that the prescribed velocity $\mathbf{c}=\mathbf{c}(t)$ at infinity satisfies the momentum equation $(A.2)_1$ identically, the expression $(A.2)_1$ evaluated at infinity can be subtracted from the same equation evaluated at an arbitrary point $\mathbf{x}$ in the three-dimensional space to obtain

$$\rho(\mathbf{v}^*_{,t}+v^*_i\mathbf{v}^*_{,i}-\mathbf{c}_{,t}=\rho(\mathbf{b}-\mathbf{b}_\infty)+(\mathbf{t}_i-\mathbf{t}_{\infty i})_{,i}, \qquad \text{(A7)}$$

where the subscript ∞ indicates the value of a variable as $z\rightarrow\pm\infty$. Substitution of (A4) into (A7) and multiplication of the resulting equation by λ_N yields

$$\frac{\partial}{\partial t}\left\{\sum_{M=0}^{K}\rho\lambda_N\overline{\lambda}_M\mathbf{w}_M-\rho\lambda_N\mathbf{c}\right\}+\frac{\partial}{\partial x_i}\left\{\sum_{M=0}^{K}\rho\lambda_N\overline{\lambda}_M v^*_i\mathbf{w}_M-\rho\lambda_N c_\beta\delta_{i\beta}\mathbf{c}\right\}$$
$$-\left\{\left[\sum_{M=0}^{K}\rho\overline{\lambda}_M\mathbf{w}_M-\rho\mathbf{c}\right]\frac{\partial\lambda_N}{\partial t}+\left[\sum_{M=0}^{K}\rho\overline{\lambda}_M v^*_i\mathbf{w}_M-\rho c_\beta\delta_{i\beta}\mathbf{c}\right]\frac{\partial\lambda_N}{\partial x_i}\right\}$$
$$=\rho\lambda_N(\mathbf{b}-\mathbf{b}_\infty)+\frac{\partial}{\partial x_i}\left\{\lambda_N(\mathbf{t}_i-\mathbf{t}_{\infty i})\right\}-\left\{(\mathbf{t}_i-\mathbf{t}_{\infty i})\frac{\partial\lambda_N}{\partial x_i}\right\}. \qquad \text{(A8)}$$

Integration of (A8) across the wake over the entire range of z gives an equation identical in form to (5), where we define the functions y_{MN}, a_{MN}, $\mathbf{v}_{MN}$, ℓ_N, $\mathbf{k}_N$, and $\mathbf{m}_N$ in the three-dimensional derivation of the theory by

$$y_{MN}=\int_{-\infty}^{\infty}\lambda_N\overline{\lambda}_M dz,\ a_{MN}=\int_{-\infty}^{\infty}\overline{\lambda}_M\left(\frac{\partial\lambda_N}{\partial t}+v^*_j\frac{\partial\lambda_N}{\partial x_j}\right)dz,$$
$$\mathbf{v}_{MN}=\mathbf{e}_\alpha\int_{-\infty}^{\infty}\lambda_N\overline{\lambda}_M v^*_\alpha dz,\ \ell_N=\int_{-\infty}^{\infty}\lambda_N(\mathbf{b}-\mathbf{b}_\infty)dz,$$
$$\mathbf{k}_N=\int_{-\infty}^{\infty}\frac{\partial\lambda_N}{\partial x_i}(\mathbf{t}_i-\mathbf{t}_{\infty i})dz,\ \mathbf{m}_N=\left[\int_{-\infty}^{\infty}\lambda_N(t_i-\mathbf{t}_{\infty i})dz\right]\nu_\alpha\delta_{i\alpha}. \qquad \text{(A9)}$$

Also, the function $\mathbf{Y}_N$ in (6) is identified in the alternative derivation of the theory as

$$\mathbf{Y}_N=\int_{-\infty}^{\infty}\lambda_N(\mathbf{v}^*_{,t}+v^*_j\mathbf{v}^*_{,j}-\mathbf{c}_{,t})dz. \qquad \text{(A10)}$$

Using (A3) and $(A9)_{5,6}$, we can identify the constraint and determinate parts of $\mathbf{k}_N$ and $\mathbf{m}_N$ as

$$\overline{\mathbf{k}}_N=-\mathbf{e}_i\int_{-\infty}^{\infty}\frac{\partial\lambda_N}{\partial x_i}(p^*-p^*_\infty)dz,$$
$$\overline{\mathbf{m}}_N=-\mathbf{e}_i\left\{\int_{-\infty}^{\infty}\lambda_N(p^*-p^*_\infty)dz\right\}\nu_\alpha\delta_{i\alpha},$$
$$\hat{\mathbf{k}}_N=\mathbf{e}_j\int_{-\infty}^{\infty}\frac{\partial\lambda_N}{\partial x_i}(\sigma_{ij}-\sigma_{\infty ij})dz,$$
$$\hat{\mathbf{m}}_N=\mathbf{e}_j\left\{\int_{-\infty}^{\infty}\lambda_N(\sigma_{ij}-\sigma_{\infty ij})dz\right\}\nu_\alpha\delta_{i\alpha}. \qquad \text{(A11)}$$

After using the weighting functions (A6), the constraint parts $\overline{\mathbf{k}}_N$ and $\overline{\mathbf{m}}_N$ in $(A11)_{1,2}$ can be written in the form (8), where the Lagrange multipliers are identified by

$$r_{\alpha\beta}=-\int_{-\infty}^{\infty}\delta_{\alpha\beta}(p^*-p^*_\infty)dz,\ r_{33,3}=-\int_{-\infty}^{\infty}(p^*-p^*_\infty)_{,3}dz,$$
$$p_1=-\ell^2\int_{-\infty}^{\infty}z^2(p^*-p^*_\infty)dz,\ p_2=-2\ell\int_{-\infty}^{\infty}z^2(p^*-p^*_\infty)dz,$$
$$p_3=2\ell^2\int_{-\infty}^{\infty}z(p^*-p^*_\infty)dz. \qquad \text{(A12)}$$

From the identifications (A.12), the functions $r_{33,3}$, p_1, and p_2 are found to satisfy the restrictions (10).

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.

Stress Field Around a Rounded Crack Tip

C. R. Chiang[1]

A method based on the eigenfunction (Williams stress functions) expansion is developed to examine the stress distribution around the tip region of a macrocrack which has some specific micro configuration. In particular, a macrocrack running into a micro hole under mode I condition is analyzed in detail. The coefficients associated with each eigenfunction are determined by a collocation procedure and the convergence of the numerical results is shown to be quite satisfactory.

1 Introduction

The present article aims at investigating the stress distribution around a "crack" possessing some specific micro configuration. The (macro) crack is assumed to run into a (micro) hole as depicted in Fig. 1. The radius of the hole ρ is much smaller than any other macro dimensions such as specimen size, crack length, etc. Therefore, the loading condition can be considered to be characterized at infinity by the singular stress field associated with the macro crack without the hole. For simplicity, we further assume the mode I symmetric loading is prevalent so that only K_I is participating.

Considering a blunt crack-tip region as a parabola (an elliptical hole or hyperbolic notch), Creager and Paris (1967) has derived an expression valid for $\theta \approx 0$. Recently, a more general analysis has been presented by Benthem (1987) concerning the stress distribution near a rounded corner. However, their formula are not strictly valid for the crack configuration depicted in Fig. 1. This seems to be overlooked in a well-known handbook (Tada, Paris, and Irwin, 1985). In order to solve this particular boundary value problem, a Trefftze's approach is employed here (Chiang, 1989). The stress distribution is assumed to be a linear combination of the Williams functions (Williams, 1952) which satisfy the field equations and the boundary condition of the crack surfaces. If we assume that the desired Airy stress function has the following product form,

$$\phi_\lambda = r^{\lambda+1} f(\theta, \lambda), \tag{1}$$

where r and θ are the polar coordinates with the origin at the crack tip. The well-known result indicates that the eigenvalues correspond to the roots of

$$\sin 2\lambda\pi = 0. \tag{2}$$

Accordingly, $\lambda = n/2$, where n is any integer. The proper restriction must be placed on the selection of n.

2 Selection of Eigenvalues

We assume that the outer expansion ϕ^0 (which is valid for $r \gg \rho$, disregarding the boundary condition on the micro hole) of the Airy stress function ϕ has been found, i.e.,

$$\theta^0(r, \theta; \rho) = \phi_0(r, \theta) + \rho\phi_1(r, \theta) + \ldots \tag{3}$$

where $\phi_0, \phi_1 \ldots$ may, if necessary, be expanded in terms of ϕ_λ with the proper restriction on λ (here, $\lambda \geq 1/2$ due to the finite value of the displacement at $r = 0$).

Now, by introducing the inner variable $r^* = r/\rho$, we expand the solution as

$$\phi^i(r, \theta; \rho) = \phi_0^* (r^*, \theta) + \rho\phi_1^* (r^*, \theta) + \ldots \tag{4}$$

where $\phi_0^*, \phi_1^*, \ldots$ etc. can be expanded in terms of ϕ_λ with proper restriction on λ (here, $\lambda \leq 1/2$). The coefficients associated with each eigenfunction are determined by enforcing the traction-free boundary condition on the micro hole and the matching condition

$$\lim_{r \to 0} \phi^0 = \lim_{r \to \infty} \phi^i. \tag{5}$$

In particular, this condition provides the proper boundary condition (at infinity) for the inner expansion.

3 Inner Expansion and Collocation Procedure

For each stress function ϕ_λ, there will be the stress components $r^{\lambda-1} S_{ij}(\theta, \lambda)$ corresponding to it. It is noted that the origin of the polar coordinates (r, θ) is at 0 (Fig. 1). Now the stress field (in cartesian coordinates) is assumed to be the following form

$$\sigma_{ij} = \frac{K_I}{\sqrt{(2\pi r)}} \left(\frac{1}{2}\right) T_{ij}^{(1)}(\theta) + A^{(0)} \left(\frac{\rho}{r}\right) T_{ij}^{(0)}(\theta) + \sum_{\alpha=-1,-2,\ldots} A^{(\alpha)} \left(\frac{r}{\rho}\right)^{\frac{\alpha}{2}-1} \left(\frac{\alpha}{2}\right) T_{ij}^{(\alpha)}(\theta) \tag{6}$$

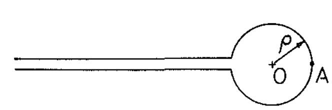


Fig. 1 The micro configuration of the macro crack

[1]Associate Professor, Department of Power Mechanical Engineering, National Tsing Hua University, Hsinchu 30043, Taiwan.

Table 1 Convergence of the numerical results with reference to the maximum stress σ_{yy}

numer or collocation points	2	4	6	8	10	12	15	20
$\frac{\sigma_{max}}{K_I/\sqrt{(2\pi\rho)}}$	1.478	3.011	2.991	2.990	2.991	2.991	2.991	2.991

Table 2 The hoop stress $\sigma_{\theta\theta}$ along the perimeter of the micro hole

θ	0°	9°	18°	27°	36°	45°	54°	63°	72°	81°	
$\frac{\sigma_{\theta\theta}}{K_I/\sqrt{(2\pi\rho)}}$	2.991	2.675	1.871	0.9330	0.2376	0	0.1859	0.5672	0.8711	0.9299	
θ	90°	99°	108°	117°	126°	135°	144°	153°	162°	171°	180°
$\frac{\sigma_{\theta\theta}}{K_I/\sqrt{(2\pi\rho)}}$	0.7456	0.4468	0.1845	0.0398	-0.0003	0	-0.0095	-0.0310	-0.0380	-0.0186	0 (theoretical value)

where functions $T_{ij}^{(\alpha)}(\theta)$ are

$$T_{xx}^{(\alpha)}=\left[\frac{\alpha}{2}+2+(-1)^{\alpha}\right]\cos\left(\frac{\alpha}{2}-1\right)\theta-\left(\frac{\alpha}{2}-1\right)\cos\left(\frac{\alpha}{2}-3\right)\theta$$

$$T_{yy}^{(\alpha)}=\left[2-\frac{\alpha}{2}-(-1)^{\alpha}\right]\cos\left(\frac{\alpha}{2}-1\right)\theta+\left(\frac{\alpha}{2}-1\right)\cos\left(\frac{\alpha}{2}-3\right)\theta$$

$$T_{xy}^{(\alpha)}=\left[-\frac{\alpha}{2}-(-1)^{\alpha}\right]\sin\left(\frac{\alpha}{2}-1\right)\theta+\left(\frac{\alpha}{2}-1\right)\sin\left(\frac{\alpha}{2}-3\right)\theta. \quad (6')$$

Explicit introduction of ρ and $\alpha/2$ is just a matter of convenience and consistency. It is noted that as $r/\rho \to \infty$, only the leading (first) term would remain which is nothing but the known boundary condition at infinity. Since each term in equation (6) satisfies the field equations of linear elasticity and essential boundary condition on the crack surfaces, so does their linear combination. What we have to do now is to adjust the coefficients $A^{(0)}$, $A^{(-1)}$,. . . etc. in terms of K_I so that the remaining boundary condition on the micro hole (i.e., $r=\rho$) is satisfied. To be practical, only N unknown coefficients are sought. The collocation method is employed to accomplish this numerical task. Each collocation point would provide two equations (i.e., $\sigma_{rr}=0$, $\sigma_{r\theta}=0$), except point A (since according to equation (6) $\sigma_{r\theta}$ automatically vanishes at A). Consequently, M collocation points could be used for the determination of 2^*M-1 unknown coefficients.

4 Results and Discussions

In Table 1, the results for σ_{yy} at point A are listed to illustrate the rapid convergence of the present approach. Apparently, only 10 (equally spaced) collocation points are sufficient to yield an accurate result. It is shown that

$$\sigma_{\max}=\frac{2.991K_I}{\sqrt{(2\pi\rho)}} \quad (7)$$

for the present micro configuration. On the other hand, for the parabolical configuration, it was shown by Creager and Paris that

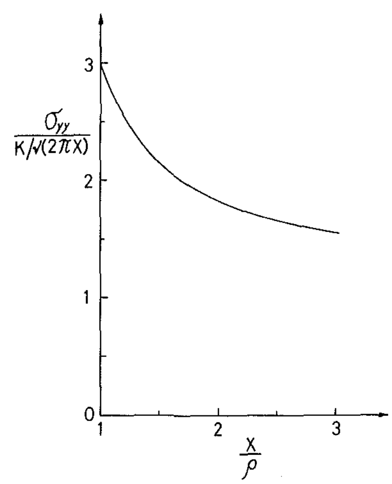


Fig. 2 The variation of σ_{yy} along the x-axis

$$\sigma_{\max}=\frac{2\sqrt{2}K_I}{\sqrt{(2\pi\rho)}}=\frac{2.828K_I}{\sqrt{(2\pi\rho)}}. \quad (8)$$

Therefore, if we use equation (8) in place of equation (7) to discuss the stress distribution for the present crack configuration, there will be about $0.163K_I/\sqrt{(2\pi\rho)}$ error in magnitude in estimating $\sigma_{\max}$.

In the interest of reference, the hoop stress $\sigma_{\theta\theta}$ distribution of the micro hole are recorded in Table 2. In addition, the variation of σ_{yy} ahead of the micro hole along the x-axis is plotted in Fig. 2.

5 Conclusions

A simple and effective method based on the eigenfunction expansion method has been described that enables one to investigate the interaction between the macro crack and micro defects. The present method is illustrated by a specific example of a macro crack running into a micro hole. The convergence of the numerical result is satisfactory. According to the present results, it is concluded (as expected) that the stress field around a macro crack depends on the "whole" crack-tip region, not simply on the radius of the curvature of the blunt tip.

References

Benthem, J. P., 1987, "Stress in the Region of Rounded Corners," *Int. J. Solids Structures*, Vol. 23, pp. 239–252.

Chiang, C. R., 1989, "A Numerical Method for Solving Elastic Fracture Problem," *Computers and Structures*, Vol. 32, pp. 1195–1197.

Creager, M., and Paris, P. C., 1967, "Elastic Field Equations for Blunt Cracks with Reference to the Stress Corrosion Cracking," *Int. J. Fracture*, Vol. 3, pp. 247–252.

Tada, H., Paris, P. C., and Irwin, G. R., 1985, *The Stress Analysis of Cracks: Handbook*, 2nd ed., Paris Production, (and Del. Research Corp.), St. Louis, MO.

Williams, M. L., 1952, "Stress Singularities Resulting From Various Boundary Conditions in Angular Corners of Plates in Extension," ASME Journal of Applied Mechanics, Vol. 19, pp. 526–528.

Formulas for Generating Prescribed Residual Stress Distributions in Center-Wound Rolls

Zine-Eddine Boutaghou[2] and Thomas R. Chase[3]

Altmann's equations for describing the residual stresses in center-wound rolled webs are solved to determine the winding stress necessary to produce prescribed residual stress distributions in the finished roll. A solution for constant circumferential stress is expanded to control the peak winding stress. Two example winding problems are discussed.

Introduction

Webs are sheet materials that exhibit negligible bending stresses when wound into rolls. Paper, magnetic tape, plastic wrap, photographic film base, adhesive tape, and metal foils are common examples. Webs are wound into rolls for processing, transport, and storage. Controlling in-roll stresses is important to prevent damage to the web due to excessive plastic deformations within the roll and to provide rolls that are sufficiently robust to withstand shock during handling.

Altmann (1968) derives exact integral expressions for stresses within a center-wound roll from basic stress-strain relationships. The Altmann model is commonly used as a linear orthotropic model for predicting in-roll stresses developed by a specified winding tension. However, Altmann's model has not been exploited for the "inverse problem" of prescribing a desired stress state in the finished roll and solving for the winding tension required to produce that stress state, as discussed here.

Two earlier authors have presented isotropic formulations for the winding tension necessary to produce prescribed residual in-roll stress distributions. Southwell (1936) solves for the winding tension necessary to obtain prescribed stress distributions in rolls of wire. Catlow and Walls (1962) solve for the winding tension necessary to obtain a constant residual tension in rolls of yarn wound on rigid cores. Catlow and Walls' equations are reproducible as a degenerate case of the general inverse formulas presented here.

Several authors discuss the importance of controlling winding tension to obtain acceptable in-roll stresses. Harland (1967) presents an analytical comparison of isotropic rolls wound with constant tension, constant torque, and constant in-roll tension. Rand and Eriksson (1973) recommend an in-roll stress distribution for newsprint on the basis of their analysis and experiments for determining in-roll stresses. Monk et al. (1975) compare in-roll stresses generated in rolls of cellophane by constant torque, constant tension, and tapered tension winding profiles.

The following section presents the "inverse equations" for generating prescribed residual stresses in the finished roll. The next section examines practical limits associated with a solution for constant circumferential stress (i.e., constant in-roll tension). Finally, the inverse equations are demonstrated with two examples.

Base Equations

Simplified forms of the residual stress formulas developed by Altmann (1968) are presented. These equations are then manipulated to determine the winding stress required to produce a roll with a specified residual radial or circumferential stress distribution.

The assumption of linear elasticity intrinsic to the Altmann model enables simplifying Altmann's original equations by applying Maxwell's reciprocal theorem[4]:

$$E_\theta/E_r = \nu_{\theta r}/\nu_{r\theta} = \beta^2 \tag{1}$$

where E_θ, E_r, $\nu_{\theta r}$, and $\nu_{r\theta}$ correspond to Altmann's E_t, E_r, μ_t, and μ_r, respectively, and β^2 is the modulus ratio. All remaining assumptions of Altmann also apply here.

The in-roll radial stress, σ_r, and circumferential stress, σ_θ, then simplify to[5]:

$$\sigma_r = -\frac{r^{2\beta}+a}{r^{\beta+1}}\int_r^{r_o}\frac{\sigma_w(s)\,s^\beta}{s^{2\beta}+a}\,ds \tag{2}$$

$$\sigma_\theta = \sigma_w(r) - \beta\,\frac{r^{2\beta}-a}{r^{\beta+1}}\int_r^{r_o}\frac{\sigma_w(s)\,s^\beta}{s^{2\beta}+a}\,ds \tag{3}$$

where r is the radius ratio[6] to the point where the stresses are measured and r_o is the outer radius ratio of finished roll. Winding stress, $\sigma_w(r)$, is the tension per unit area of web cross-section as the web initially enters the roll, and s is a variable of integration. Web material and core parameter, a, is defined as:

$$a = \frac{\beta - \nu_{\theta r} - E_\theta/E_C}{\beta + \nu_{\theta r} + E_\theta/E_C} \tag{4}$$

where E_C is the effective radial modulus of the core.

The winding stress required to produce a prescribed radial stress distribution is obtained by taking the derivative of equation (2) with respect to radius ratio, r:

$$\sigma_w(r) = -\frac{\beta\,(r^{2\beta}-a)-(r^{2\beta}+a)}{r^{2\beta}+a}\,\sigma_r + r\,\frac{d\sigma_r}{dr}. \tag{5}$$

The associated circumferential stress is obtained directly from the basic force equilibrium equation for the roll[7]:

[2]General Engineer, IBM Rochester, Rochester, MN 55901.

[3]Assistant Professor, Mechanical Engineering Department, University of Minnesota, Minneapolis, MN 55455. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, Jan. 18, 1990; final revision, June 13, 1990.

[4]See, for example, Tauchert (1974). The necessity of asserting equation (1) to satisfy strain energy constraints is also demonstrated by Willett and Poesch (1988) and noted by Hakiel (1987).

[5]These equations correspond to equations (10) and (11) in Altmann. Parameters r_o, r, σ_w, σ_θ, and $-\sigma_r$ correspond to Altmann's R, r, T_w, T, and P, respectively.

[6]Radius ratio, r, is the actual radius divided by the outer radius of the core.

[7]Altmann's equation (15).

of a macro crack running into a micro hole. The convergence of the numerical result is satisfactory. According to the present results, it is concluded (as expected) that the stress field around a macro crack depends on the "whole" crack-tip region, not simply on the radius of the curvature of the blunt tip.

References

Benthem, J. P., 1987, "Stress in the Region of Rounded Corners," *Int. J. Solids Structures*, Vol. 23, pp. 239–252.

Chiang, C. R., 1989, "A Numerical Method for Solving Elastic Fracture Problem," *Computers and Structures*, Vol. 32, pp. 1195–1197.

Creager, M., and Paris, P. C., 1967, "Elastic Field Equations for Blunt Cracks with Reference to the Stress Corrosion Cracking," *Int. J. Fracture*, Vol. 3, pp. 247–252.

Tada, H., Paris, P. C., and Irwin, G. R., 1985, *The Stress Analysis of Cracks: Handbook*, 2nd ed., Paris Production, (and Del. Research Corp.), St. Louis, MO.

Williams, M. L., 1952, "Stress Singularities Resulting From Various Boundary Conditions in Angular Corners of Plates in Extension," ASME Journal of Applied Mechanics, Vol. 19, pp. 526–528.

Formulas for Generating Prescribed Residual Stress Distributions in Center-Wound Rolls

Zine-Eddine Boutaghou[2] and Thomas R. Chase[3]

Altmann's equations for describing the residual stresses in center-wound rolled webs are solved to determine the winding stress necessary to produce prescribed residual stress distributions in the finished roll. A solution for constant circumferential stress is expanded to control the peak winding stress. Two example winding problems are discussed.

Introduction

Webs are sheet materials that exhibit negligible bending stresses when wound into rolls. Paper, magnetic tape, plastic wrap, photographic film base, adhesive tape, and metal foils are common examples. Webs are wound into rolls for processing, transport, and storage. Controlling in-roll stresses is important to prevent damage to the web due to excessive plastic deformations within the roll and to provide rolls that are sufficiently robust to withstand shock during handling.

Altmann (1968) derives exact integral expressions for stresses within a center-wound roll from basic stress-strain relationships. The Altmann model is commonly used as a linear orthotropic model for predicting in-roll stresses developed by a specified winding tension. However, Altmann's model has not been exploited for the "inverse problem" of prescribing a desired stress state in the finished roll and solving for the winding tension required to produce that stress state, as discussed here.

Two earlier authors have presented isotropic formulations for the winding tension necessary to produce prescribed residual in-roll stress distributions. Southwell (1936) solves for the winding tension necessary to obtain prescribed stress distributions in rolls of wire. Catlow and Walls (1962) solve for the winding tension necessary to obtain a constant residual tension in rolls of yarn wound on rigid cores. Catlow and Walls' equations are reproducible as a degenerate case of the general inverse formulas presented here.

Several authors discuss the importance of controlling winding tension to obtain acceptable in-roll stresses. Harland (1967) presents an analytical comparison of isotropic rolls wound with constant tension, constant torque, and constant in-roll tension. Rand and Eriksson (1973) recommend an in-roll stress distribution for newsprint on the basis of their analysis and experiments for determining in-roll stresses. Monk et al. (1975) compare in-roll stresses generated in rolls of cellophane by constant torque, constant tension, and tapered tension winding profiles.

The following section presents the "inverse equations" for generating prescribed residual stresses in the finished roll. The next section examines practical limits associated with a solution for constant circumferential stress (i.e., constant in-roll tension). Finally, the inverse equations are demonstrated with two examples.

Base Equations

Simplified forms of the residual stress formulas developed by Altmann (1968) are presented. These equations are then manipulated to determine the winding stress required to produce a roll with a specified residual radial or circumferential stress distribution.

The assumption of linear elasticity intrinsic to the Altmann model enables simplifying Altmann's original equations by applying Maxwell's reciprocal theorem[4]:

$$E_\theta / E_r = \nu_{\theta r}/\nu_{r\theta} = \beta^2 \tag{1}$$

where E_θ, E_r, $\nu_{\theta r}$, and $\nu_{r\theta}$ correspond to Altmann's E_t, E_r, μ_t, and μ_r, respectively, and β^2 is the modulus ratio. All remaining assumptions of Altmann also apply here.

The in-roll radial stress, σ_r, and circumferential stress, σ_θ, then simplify to[5]:

$$\sigma_r = -\frac{r^{2\beta}+a}{r^{\beta+1}} \int_r^{r_o} \frac{\sigma_w(s)\, s^\beta}{s^{2\beta}+a}\, ds \tag{2}$$

$$\sigma_\theta = \sigma_w(r) - \beta \frac{r^{2\beta}-a}{r^{\beta+1}} \int_r^{r_o} \frac{\sigma_w(s)\, s^\beta}{s^{2\beta}+a}\, ds \tag{3}$$

where r is the radius ratio[6] to the point where the stresses are measured and r_o is the outer radius ratio of finished roll. Winding stress, $\sigma_w(r)$, is the tension per unit area of web cross-section as the web initially enters the roll, and s is a variable of integration. Web material and core parameter, a, is defined as:

$$a = \frac{\beta - \nu_{\theta r} - E_\theta/E_C}{\beta + \nu_{\theta r} + E_\theta/E_C} \tag{4}$$

where E_C is the effective radial modulus of the core.

The winding stress required to produce a prescribed radial stress distribution is obtained by taking the derivative of equation (2) with respect to radius ratio, r:

$$\sigma_w(r) = -\frac{\beta\,(r^{2\beta}-a) - (r^{2\beta}+a)}{r^{2\beta}+a}\,\sigma_r + r\frac{d\sigma_r}{dr}. \tag{5}$$

The associated circumferential stress is obtained directly from the basic force equilibrium equation for the roll[7]:

[2]General Engineer, IBM Rochester, Rochester, MN 55901.

[3]Assistant Professor, Mechanical Engineering Department, University of Minnesota, Minneapolis, MN 55455. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, Jan. 18, 1990; final revision, June 13, 1990.

[4]See, for example, Tauchert (1974). The necessity of asserting equation (1) to satisfy strain energy constraints is also demonstrated by Willett and Poesch (1988) and noted by Hakiel (1987).

[5]These equations correspond to equations (10) and (11) in Altmann. Parameters r_o, r, σ_w, σ_θ, and $-\sigma_r$ correspond to Altmann's R, r, T_w, T, and P, respectively.

[6]Radius ratio, r, is the actual radius divided by the outer radius of the core.

[7]Altmann's equation (15).

$$\sigma_\theta = \sigma_r + r\frac{d\sigma_r}{dr}. \quad (6)$$

Note that the associated circumferential stress is a function only of the prescribed radial stress; i.e., it is not directly dependent on the material properties of the roll.

The radial stress corresponding to a prescribed circumferential stress can be found by solving equation (6) for σ_r:

$$\sigma_r = -\frac{1}{r}\int_r^{r_o} \sigma_\theta \, ds. \quad (7)$$

The winding stress required to produce the prescribed circumferential stress distribution is found by substituting equation (7) into equation (5):

$$\sigma_W(r) = \sigma_\theta + \frac{\beta}{r}\frac{r^{2\beta} - a}{r^{2\beta} + a}\int_r^{r_o} \sigma_\theta \, ds. \quad (8)$$

As above, the associated radial stress is independent of the material properties of the roll.

Note that the user may prescribe a desired radial or circumferential stress state in the finished roll, but not both.

Limiting the Peak Winding Stress for Constant Circumferential Stress

Winding a finished roll for constant circumferential stress has particular practical value, since circumferential compression of the web is prevented and longitudinal deformation of the web is uniform throughout the roll (Harland, 1967). However, the solution for constant circumferential stress leads to a peak winding stress at or near the core. This peak may be severe enough to damage the web. Controlling this peak is addressed here at the expense of limiting the outer radius ratio of the finished roll. Stiffening the core will be shown to enable increasing the outer radius ratio to a limited extent.

The basic constant circumferential stress solution is obtained by simply prescribing σ_θ to be a constant, $\sigma_{\theta o}$, in equation (8), yielding[8]:

$$\sigma_W(r) = \sigma_{\theta o}\left(1 + \beta\frac{r_o - r}{r}\frac{r^{2\beta} - a}{r^{2\beta} + a}\right). \quad (9)$$

The radius ratio where the winding stress peaks, r_P, is found by setting the derivative of equation (9) to zero. Eliminating r_o between the result and equation (9) yields a quadratic equation in $r_P^{2\beta}$. The analysis given in the Appendix proves that at most one root to this equation can exceed one. In this case:

$$r_P = \left[a\frac{b + \beta(1+\xi)}{\xi + \beta(1-\xi)}\right]^{\frac{1}{2\beta}} \quad (10)$$

where:

$$\xi = 1 - \sigma_{\theta o}/\sigma_{WP} \quad (11)$$

$$b = \sqrt{\xi(4\beta^2 + \xi)} \quad (12)$$

and σ_{WP} is the peak winding stress. Otherwise, inspection of equation (9) indicates that the peak circumferential stress must occur at the core surface ($r_P = 1$). The condition for r_P exceeding one is (from equation (10)):

$$a > \frac{\xi + \beta(1-\xi)}{b + \beta(1+\xi)}. \quad (13)$$

If the peak winding stress occurs at the core surface, the outer radius ratio associated with prescribing a peak winding stress, σ_{WP}, is (from equation (9)):

$$r_o = 1 + \frac{1}{\beta}\frac{1+a}{1-a}\frac{\xi}{1-\xi} \quad (14)$$

Otherwise:

$$r_o = \frac{\xi + \beta(1-\xi)}{\beta(1-\xi)}\frac{b + \xi(2\beta+1)}{b + \xi(2\beta-1)}\left[a\frac{b+\beta(1+\xi)}{\xi+\beta(1-\xi)}\right]^{\frac{1}{2\beta}}. \quad (15)$$

Inspection of equations (14) and (15) reveals that the maximum outer radius ratio can be increased somewhat by stiffening the core. However, examination of equation (4) reveals that material and core parameter, a, is not arbitrarily variable; rather, a must fall within the bounds:

$$-1 < a < \frac{\beta - \nu_{\theta r}}{\beta + \nu_{\theta r}} < 1. \quad (16)$$

The lower limit corresponds to an infinitely soft core ($E_C = 0$) and the upper limit corresponds to a rigid core ($E_C = \infty$). The maximum possible value for the outer radius ratio can be determined by evaluating equations (14) and (15) for a rigid core.

Example 1: Winding a Roll of Cellophane for Constant Circumferential Stress

The inverse equations for circumferential stress are demonstrated by applying them to a roll of cellophane. The basic constant circumferential stress solution produces a high peak winding stress. Therefore, the outer radius ratio is controlled to limit the peak winding stress.

Monk et al. (1975) suggest E_θ, E_r, E_C, and $\nu_{\theta r}$ values of 2.10 GPa, 34.5 MPa, 689 MPa, and 0.10, respectively, for cellophane. They present data for winding such a roll to a radius ratio, r_o, of 2.8. Winding this roll for constant circumferential stress using equation (9) produces a peak winding stress having a magnitude over 11 times the prescribed constant circumferential stress.

The peak winding stress is controlled by limiting it to five times the constant circumferential stress. Inequality (13) then determines that the peak winding stress occurs above the core. Equation (15) fixes the outer radius ratio, r_o, at 1.78.

The winding stress ratio[9] obtained from equation (9) is graphed in Fig. 1(*a*). The winding stress peaks at a radius ratio, r_P, of 1.11.

Figure 1(*b*) illustrates the circumferential stress for several states of winding[10]. The circumferential stress is shown to move smoothly to the final prescribed constant stress. Note, however, that the circumferential stress does not become constant until the roll is fully wound. Also note that the circumferential stress remains at or above the prescribed constant level for all states of winding.

Figure 1(*c*) illustrates the corresponding radial stress for several states of winding. The radial stress steadily decreases to the final state of:

$$\sigma_r = (\sigma_{\theta o}/r)(r - r_o) \quad (17)$$

predicted by equation (7). According to the observations of Connolly and Winarski (1984) and Frye (1967), this distribution appears desirable for avoiding slippage near the core.

A rigid core would enable increasing the outer radius ratio to a maximum of 1.88 for the prescribed winding stress limit. Therefore, the outer radius of the core would have to be increased to store the same amount of web as Monk et al.'s original roll. The actual roll diameter would increase proportionally.

[8] Note that if $r^{2\beta} >> a$ then: $r\sigma_W(r) \approx \sigma_{\theta o}[r(1-\beta)+\beta r_o]$; i.e., the constant circumferential stress profile can be approximated by linearly varying the winding torque as a function of radius.

[9] All stresses in Fig. 1 are made dimensionless by dividing by the specified final circumferential stress, $\sigma_{\theta o}$.

[10] The curves in Figs. 1 (*b*) and (*c*) were obtained by substituting the prescribed winding stress into equations (2) and (3).

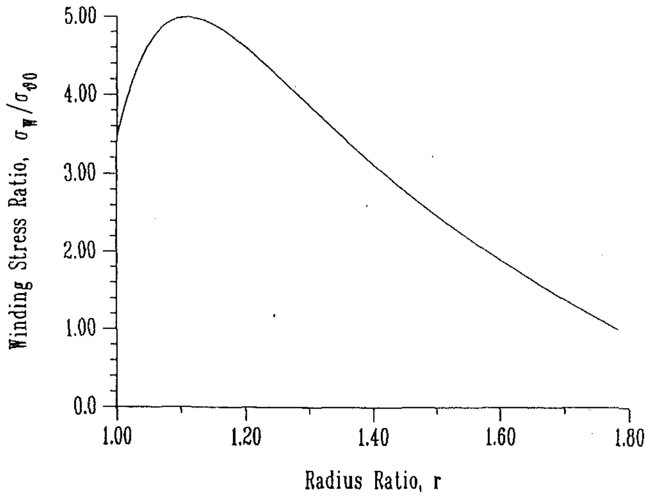


Fig. 1(*a*) Winding stress ratio versus radius ratio

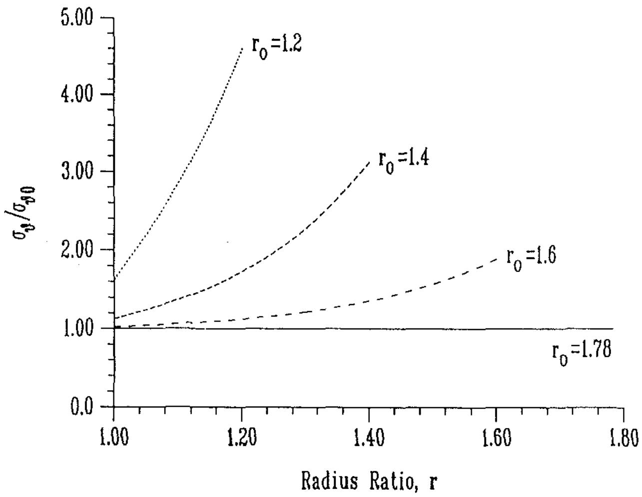


Fig. 1(*b*) Circumferential stress ratio versus radius ratio for several states of winding

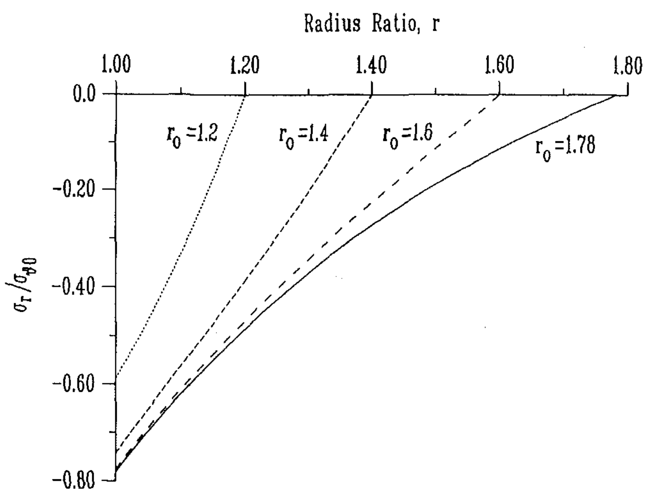


Fig. 1(*c*) Radial stress ratio versus radius ratio for several states of winding

Fig. 1 Winding a roll of cellophane with constant circumferential stress

Example 2: Winding a Roll of Paper to Obtain a Prescribed Radial Stress Distribution

Frye (1967) suggests a radial stress profile for paper rolls[11]. His profile is designed to prevent slipping near the core and the associated winding defects. Equations are developed to represent Frye's suggested profile. These equations are then used to determine the winding stress required to obtain this profile with center winding. The associated circumferential stress distribution is also illustrated.

Frye presents his recommended radial stress profile as a graph for a 750 mm (30 in.) radius roll wound on a 100 mm (4 in.) diameter core. The following functions of radius ratio were developed to approximate his curve:

$$\sigma_r = \{-236.8[2.260e^{-0.2011\,(r-1)} + 1]\}\text{KPa} \qquad 1 \le r \le 14 \quad (18)$$

$$\sigma_r = \{268.9(r-14)^3 - 0.7923(r-14)^2 + 7.880(r-14) - 276.0\}\text{KPa} \quad 14 \le r \le 15 \quad (19)$$

These functions are illustrated in Fig. 2(*a*).

Two functions were joined at an actual radius of 700 mm (r = 14) to duplicate a sharp change in Frye's recommended profile at that point. The exponential function was fit to the radial stress at the core, the radial stress at 700 mm, and the slope at 700 mm measured from Frye's profile. Continuity of the magnitude and slope of the winding stress at 700 mm was guaranteed by using three constants of the cubic polynomial function to match the exponential function through the second derivative. The fourth constant of the cubic polynomial drives the radial stress to zero at the outer surface of the roll.

Values for E_θ, E_r, and $\nu_{\theta r}$ of 4.82 GPa, 31 MPa, and 0.01, respectively, are suggested for paper by Altmann (1968). A core modulus, E_C, of 387 GPa produces a convenient material and core parameter, a, of zero[12].

The winding stress plotted in Fig. 2(*b*) is obtained by substituting equations (18) and (19) in equation (5). Note that the sudden increase in radial stress prescribed near the outer surface requires a corresponding surge in the winding stress. The winding stress peaks to a value of 12 MPa (3.5 pli for 2 mil caliper paper) at the outer edge. This peak may limit the practicality of producing this profile with center winding.

The circumferential stress plotted in Fig. 2(*c*) is obtained by substituting equations (18) and (19) in equation (6). Most of the paper in the roll is subjected to a very small compression, as recommended by Rand and Eriksson (1973). The area under the zero stress line is small, as recommended by Hussain et al. (1968).

Conclusion

The linear in-roll stress equations developed by Altmann (1968) are manipulated to enable determining the winding stress necessary to produce prescribed stress states in the finished roll. The roll modulus in the radial direction is known to exhibit nonlinear behavior (Pfeiffer, 1979 and 1987; Hakiel, 1987; Willett and Poesch, 1988). Furthermore, temperature and aging effects are known to alter the initial wound-in stress state (Tramposch, 1967; Umanskii and Shidlovskii, 1983; Connolly and Winarski, 1984; Lin and Westmann, 1989). Nevertheless, the linear inverse equations presented here can be used to provide a starting point for developing inverse solutions that include these effects in addition to their usefulness for rolls that exhibit essentially linear behavior.

[11]Fig. 8 in Frye (1967).

[12]Specifying an overly stiff core produces negative winding stresses in equation (5).

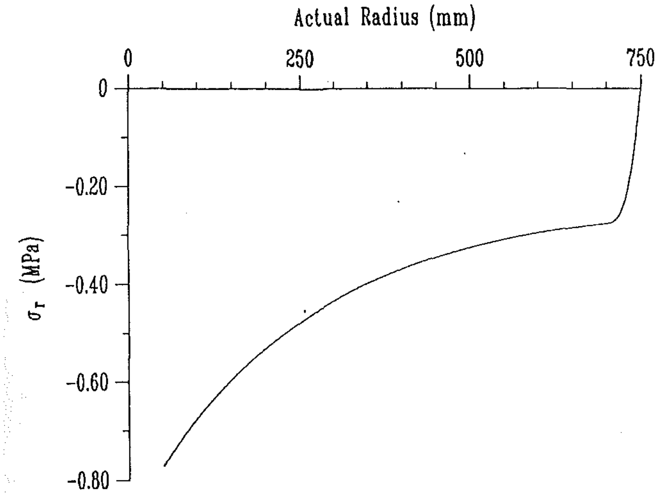


Fig. 2(*a*) Prescribed radial stress versus radius

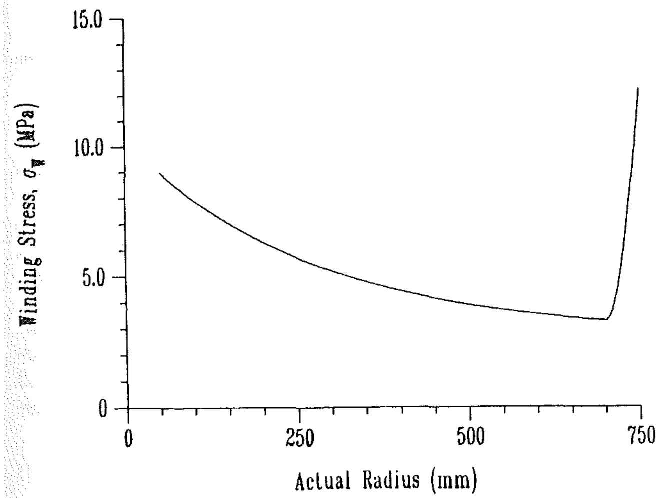


Fig. 2(*b*) Winding stress versus radius

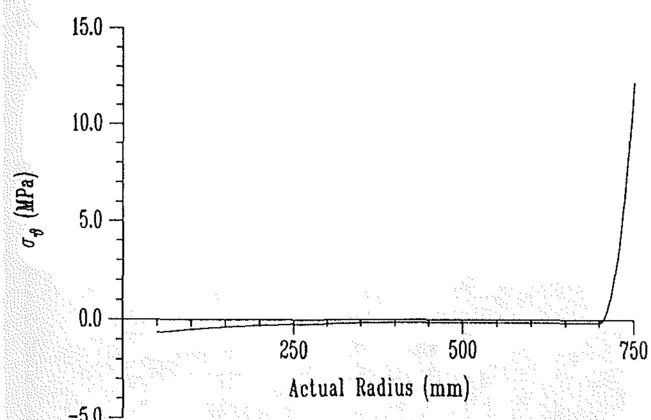


Fig. 2(*c*) Circumferential stress versus radius

Fig. 2 Winding a roll of paper to obtain a prescribed radial stress distribution

Acknowledgments

The suggestions of Mr. Bruce Tait, Dr. James McKinney, and our reviewers are gratefully acknowledged.

References

Altmann, H. C., 1968, "Formulas for Computing the Stresses in Center-Wound Rolls," *TAPPI Journal*, Vol. 51, No. 4, pp. 176–179.

Catlow, M. G., and Walls, G. W., 1962, "A Study of Stress Distribution in Pirns," *Journal of the Textile Institute, Transactions*, Vol. 53, No. 9, pp. T410–429.

Connolly, D., and Winarski, D. J., 1984, "Stress Analysis of Wound Magnetic Tape," *Tribology and Mechanics of Magnetic Storage Systems*, ASLE Special Publication SP-16, pp. 172–182.

Frye, K. G., 1967, "Winding Variables and Their Effect on Roll Hardness and Roll Quality," *TAPPI Journal*, Vol. 50, No. 7, pp. 81A–86A.

Hakiel, Z., 1987, "Nonlinear Model for Wound Roll Stresses," *TAPPI Journal*, Vol. 70, No. 5, pp. 113–117.

Harland, W. G., 1967, "Stress Distribution and Winding Faults in Reels of Plastic Film," *Polymer Engineering and Science*, Vol. 7, No. 1, pp. 58–62.

Hussain, S. M., Farrell, W. R., and Gunning, J. R., 1968, "Most Paper in the Roll is in Unstable Condition," *Canadian Pulp and Paper Industry*, Vol. 21, No. 8, pp. 52–61.

Lin, J. Y., and Westmann, R. A., 1989, "Viscoelastic Winding Mechanics," ASME Journal of Applied Mechanics, Vol. 56, No. 4, pp. 821–827.

Monk, D. W., Lautner, W. K., and McMullen, J. F., 1975, "Internal Stresses Within Rolls of Cellophane," *TAPPI Journal*, Vol. 58, No. 8, pp. 152–155.

Pfeiffer, J. D., 1979, "Prediction of Roll Defects from Roll Structure Formulas," *TAPPI Journal*, Vol. 62, No. 10, pp. 83–85.

Pfeiffer, J. D., 1987, "An Update of Pfeiffer's Roll-Winding Model," *TAPPI Journal*, Vol. 70, No. 10, p. 132.

Rand, T., and Eriksson, L. G., 1973, "Physical Properties of Newsprint Rolls During Winding," *TAPPI Journal*, Vol. 56, No. 6, pp. 153–156.

Southwell, R. V., 1936, *An Introduction to the Theory of Elasticity for Engineers and Physicists*, Oxford University Press, London, pp. 416–423.

Tauchert, T. R., 1974, *Energy Principles in Structural Mechanics*, McGraw-Hill, New York, p. 72.

Tramposch, H., 1967, "Anisotropic Relaxation of Internal Forces in a Wound Reel of Magnetic Tape," ASME Journal of Applied Mechanics, Vol. 34, No. 4, pp. 888–894.

Umanskii, E. S., and Shidlovskii, N. S., 1983 (English 1984), "Experimental Study of Residual Stresses in Reels of Magnetic Tape," *Strength of Materials*, Vol. 15, No. 10, pp. 1402–1409 (in English).

Willett, M. S., and Poesch, W. L., 1988, "Determining the Stress Distributions in Wound Reels of Magnetic Tape Using a Nonlinear Finite-Difference Approach," ASME Journal of Applied Mechanics, Vol. 55, No. 2, pp. 365–371.

APPENDIX

Analysis of the Peak Winding Stress Radius Ratio for Constant Circumferential Stress

This Appendix determines the attributes of the peak winding stress radius ratio, r_P:

$$r_P = \left[a \frac{\beta(1+\xi) \pm \sqrt{\xi(4\beta^2+\xi)}}{\xi+\beta(1-\xi)} \right]^{\frac{1}{2\beta}}. \tag{A1}$$

The fractional expression in equation (A1) must be positive when adding the discriminant. Therefore, if parameter a is negative, r_P will not be real. If parameter a is positive, r_P will exceed one when inequality (13) is satisfied.

Now, consider the fractional expression in equation (A1) corresponding to subtracting the discriminant to be a function of ξ:

$$f(\xi) = \frac{\beta(1+\xi) - \sqrt{\xi(4\beta^2+\xi)}}{\xi+\beta(1-\xi)}. \tag{A2}$$

Since the magnitude of parameter a is less than one, r_P will be less than one or imaginary when subtracting the discriminant if the magnitude of $f(\xi)$ is always less than or equal to one. This is proven by demonstrating that the magnitude of $f(\xi)$ must be less than or equal to one for the extreme values of ξ and proving that $\dfrac{df(\xi)}{d\xi}$ cannot change sign for intermediate values.

Evaluating $f(\xi)$ at the extreme values of ξ yields:

$$f(\xi=0) = 1 \tag{A3}$$

$$-1 < f(\xi=1) < 0. \tag{A4}$$

The derivative of equation (A2) has a double pole and a double zero at:

$$\xi = \beta / (\beta - 1) \tag{A5}$$

in addition to a pole at zero. However, since β exceeds one, equation (A5) must produce a value of ξ exceeding 1. Therefore, the slope must have the same sense for all values of ξ between zero and one; Q.E.D.

Stress-Free Configuration of a Thick-Walled Cylindrical Model of the Artery: An Application of Riemann Geometry to the Biomechanics of Soft Tissues

Keiichi Takamizawa[13]

Introduction

Many researchers have taken account of the finite deformation and the nonlinearity in cardiovascular tissues. They have assumed that living organs are free from stress when the organs are unloaded. Fung (1984) called this assumption "zero initial stress hypothesis" and criticized. Living organs never experience an unloading condition when they are in a living body. The tissues in those organs develop and resorp through the metabolism of cells. The metabolism of cells may depend not only on their chemical circumstances, but also their mechanical conditions (Leung et al., 1976)—stresses and strains in the living organs. Since all processes are performed under loading conditions, the cells cannot take account of a no load condition. An optimal condition of an organ is different from those where there remains no stress in the unloading state. This implies that it is not plausible for an organ, which receives certain mechanical loads under a normal condition, to be free from stress when the organ is unloaded. Indeed, recent investigations (Fung, 1984; Vaishnav and Vossoughi, 1987) have experimentally shown that there are residual stresses in arteries. For example, after longitudinal cutting, the arterial ring soaked in a physiological salt solution springs open and become a shape like a sector as shown in Fig. 1. This implies that there remains compressive stress in the inner side and tensile stress in the outer side in the circumferential direction before the arterial ring was longitudinally cut. Thus, we cannot simply take an unloading state as a stress-free reference configuration in biomechanics.

The residual stress in living organs may be due to an adaptation to physiological loads. It has been shown that the residual stresses in the cardiovascular tissues considerably reduce severe stress concentrations under physiological conditions which occur if there are no residual stresses.

Chuong and Fung (1986) analyzed the stress distribution and the residual stress in the arterial wall by directly taking the sector configuration obtained from the experiment as a stress-free configuration of the artery. On the other hand, Takamizawa and Hayashi (1987) assume the uniform strain distribution as a result of an adaptation to physiological loads in the arterial wall. That is, the residual stress and strain are not the first assumption in their study, but the secondary results from the uniform strain hypothesis. The hypothesis asserts that the circumferential strain uniformly distributes through the wall thickness at a physiological loading configuration, which is the state of the mean blood pressure and the *in vivo* axial stretch.

Since the condition of uniform strain is assumed to be a result of an adaptation to the physiological load, the condition at which the uniform strain occurs may change depending on the physiological circumstance. For instance, the intraluminal pressure at which the uniform strain occurs in normotensive animals may differ from that of hypertensive ones. Even in an individual the condition may change depending on the age, and if the individual is suffering from a disease.

A version of the continuum theories of dislocations is an appropriate means to analyze a solid undergoing residual stress (Takamizawa and Matsuda, 1990). In this paper the deformation of the areterial wall is analyzed by this theory. The important result obtained from the theory is that we can take a Riemannian manifold as a stress-free reference configuration. In this sense the theory is an extension of the ordinary continuum kinematics of solids. The fundamental equations based on this theory are the same as the ones in the orthodox continuum mechanics in the form of a general tensor.

The determination of the Riemannian metric is equivalent to the determination of the residual strain. This implies that we must evaluate the distribution of the residual strain or assume a certain principle to determine the Riemannian metric. The uniform strain hypothesis is one such principle.

Riemannian Stress-Free Configuration and Deformation of the Artery

We shall consider the deformation of a thick-walled cylindrical model for the artery. The wall material is assumed to be homogeneous and incompressible (Carew et al., 1968), and is the uniform strain hypothesis.

It is assumed that there are only normal components of strain in the cylindrical coordinate system [r, θ, z; $g_{rr}=1$, $g_{\theta\theta}=r^2$, $g_{zz}=1$, $g_{kl}=0$ ($k \neq l$)]. Here, r, θ, z denote the radial, circumferential, and longitudinal coordinates, respectively; g_{kl} (k, $l=r$, θ, or z) are the components of metric tensor of the cylindrical coordinate system. The coordinates represent the position of a material point p in a pressure loading and longitudinally stretched state. The deformation is given as follows:

$$r=r(R),\ \theta=\Theta,\ z=\alpha Z, \tag{1}$$

where α is a constant, and R, Θ, Z represent the position of the material point p in a physiological normal state, which is adopted as a reference configuration. The components of the metric tensor of the cylindrical coordinate system at (R, Θ, Z) are denoted by G_{KL}.

The infinitesimal length dS between (R, Θ, Z) and ($R+dR$, $\Theta+d\Theta$, $Z+dZ$) in the stress-free state is given as follows:

$$dS^2 = H_{RR}dR^2 + H_{\Theta\Theta}d\Theta^2 + H_{ZZ}dZ^2 + 2(H_{\Theta Z}d\Theta dZ + H_{ZR}dZdR + H_{R\Theta}dRd\Theta). \tag{2}$$

We shall call the manifold defined by the coordinates R, Θ,

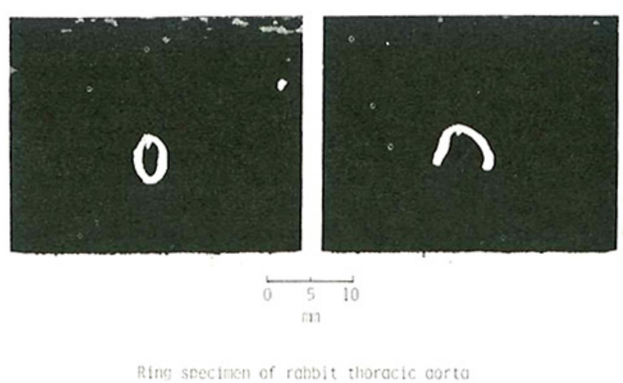


Fig. 1 The left panel shows a ring specimen of the rabbit thoracic aorta soaked in the Krebs solution kept at 37°C. After longitudinal cutting, the ring specimen springs open as shown in the right panel.

[13] Department of Biomedical Engineering, National Cardiovascular Center Research Institute, 6-7-1 Fujishiro-dai, Suita, Osaka 565, Japan.

The derivative of equation (A2) has a double pole and a double zero at:

$$\xi = \beta/(\beta-1) \tag{A5}$$

in addition to a pole at zero. However, since β exceeds one, equation (A5) must produce a value of ξ exceeding 1. Therefore, the slope must have the same sense for all values of ξ between zero and one; Q.E.D.

Stress-Free Configuration of a Thick-Walled Cylindrical Model of the Artery: An Application of Riemann Geometry to the Biomechanics of Soft Tissues

Keiichi Takamizawa[13]

Introduction

Many researchers have taken account of the finite deformation and the nonlinearity in cardiovascular tissues. They have assumed that living organs are free from stress when the organs are unloaded. Fung (1984) called this assumption "zero initial stress hypothesis" and criticized. Living organs never experience an unloading condition when they are in a living body. The tissues in those organs develop and resorp through the metabolism of cells. The metabolism of cells may depend not only on their chemical circumstances, but also their mechanical conditions (Leung et al., 1976)—stresses and strains in the living organs. Since all processes are performed under loading conditions, the cells cannot take account of a no load condition. An optimal condition of an organ is different from those where there remains no stress in the unloading state. This implies that it is not plausible for an organ, which receives certain mechanical loads under a normal condition, to be free from stress when the organ is unloaded. Indeed, recent investigations (Fung, 1984; Vaishnav and Vossoughi, 1987) have experimentally shown that there are residual stresses in arteries. For example, after longitudinal cutting, the arterial ring soaked in a physiological salt solution springs open and become a shape like a sector as shown in Fig. 1. This implies that there remains compressive stress in the inner side and tensile stress in the outer side in the circumferential direction before the arterial ring was longitudinally cut. Thus, we cannot simply take an unloading state as a stress-free reference configuration in biomechanics.

The residual stress in living organs may be due to an adaptation to physiological loads. It has been shown that the residual stresses in the cardiovascular tissues considerably reduce severe stress concentrations under physiological conditions which occur if there are no residual stresses.

Chuong and Fung (1986) analyzed the stress distribution and the residual stress in the arterial wall by directly taking the sector configuration obtained from the experiment as a stress-free configuration of the artery. On the other hand, Takamizawa and Hayashi (1987) assume the uniform strain distribution as a result of an adaptation to physiological loads in the arterial wall. That is, the residual stress and strain are not the first assumption in their study, but the secondary results from the uniform strain hypothesis. The hypothesis asserts that the circumferential strain uniformly distributes through the wall thickness at a physiological loading configuration, which is the state of the mean blood pressure and the *in vivo* axial stretch.

Since the condition of uniform strain is assumed to be a result of an adaptation to the physiological load, the condition at which the uniform strain occurs may change depending on the physiological circumstance. For instance, the intraluminal pressure at which the uniform strain occurs in normotensive animals may differ from that of hypertensive ones. Even in an individual the condition may change depending on the age, and if the individual is suffering from a disease.

A version of the continuum theories of dislocations is an appropriate means to analyze a solid undergoing residual stress (Takamizawa and Matsuda, 1990). In this paper the deformation of the areterial wall is analyzed by this theory. The important result obtained from the theory is that we can take a Riemannian manifold as a stress-free reference configuration. In this sense the theory is an extension of the ordinary continuum kinematics of solids. The fundamental equations based on this theory are the same as the ones in the orthodox continuum mechanics in the form of a general tensor.

The determination of the Riemannian metric is equivalent to the determination of the residual strain. This implies that we must evaluate the distribution of the residual strain or assume a certain principle to determine the Riemannian metric. The uniform strain hypothesis is one such principle.

Riemannian Stress-Free Configuration and Deformation of the Artery

We shall consider the deformation of a thick-walled cylindrical model for the artery. The wall material is assumed to be homogeneous and incompressible (Carew et al., 1968), and is the uniform strain hypothesis.

It is assumed that there are only normal components of strain in the cylindrical coordinate system [r, θ, z; $g_{rr}=1$, $g_{\theta\theta}=r^2$, $g_{zz}=1$, $g_{kl}=0$ $(k\neq l)$]. Here, r, θ, z denote the radial, circumferential, and longitudinal coordinates, respectively; g_{kl} (k, $l=r$, θ, or z) are the components of metric tensor of the cylindrical coordinate system. The coordinates represent the position of a material point p in a pressure loading and longitudinally stretched state. The deformation is given as follows:

$$r=r(R),\ \theta=\Theta,\ z=\alpha Z, \tag{1}$$

where α is a constant, and R, Θ, Z represent the position of the material point p in a physiological normal state, which is adopted as a reference configuration. The components of the metric tensor of the cylindrical coordinate system at (R, Θ, Z) are denoted by G_{KL}.

The infinitesimal length dS between (R, Θ, Z) and ($R+dR$, $\Theta+d\Theta$, $Z+dZ$) in the stress-free state is given as follows:

$$dS^2 = H_{RR}dR^2 + H_{\Theta\Theta}d\Theta^2 + H_{ZZ}dZ^2 + 2(H_{\Theta Z}d\Theta dZ + H_{ZR}dZdR + H_{R\Theta}dRd\Theta). \tag{2}$$

We shall call the manifold defined by the coordinates R, Θ,

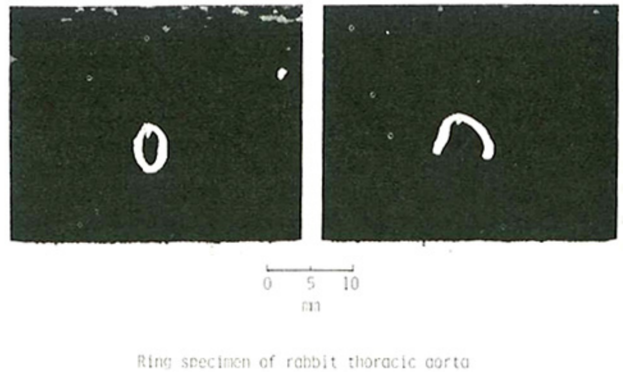


Fig. 1 The left panel shows a ring specimen of the rabbit thoracic aorta soaked in the Krebs solution kept at 37°C. After longitudinal cutting, the ring specimen springs open as shown in the right panel.

[13] Department of Biomedical Engineering, National Cardiovascular Center Research Institute, 6-7-1 Fujishiro-dai, Suita, Osaka 565, Japan.

Z and the metric tensor H_{KL} Riemannian stress-free configuration. Referring to the Riemannian stress-free configuration, the Green's strain tensor of the deformed state is defined by

$$E_{KL}=\frac{1}{2}\left(g_{kl}\frac{\partial x^k}{\partial X^K}\frac{\partial x^l}{\partial X^L}-H_{KL}\right), \tag{3}$$

where the coordinate $X^1=R$, $X^2=\Theta$, $X^3=Z$ represent the material point p in the reference configuration—and those of a deformed configuration $x^1=r$, $x^2=\theta$, $x^3=z$. The radial, circumferential, and longitudinal stretching ratios of the deformed state, referred to the stress-free configuration, are given as follows:

$$\lambda_r=\sqrt{\frac{g_{rr}}{H_{RR}}}\frac{\partial r}{\partial R},$$
$$\lambda_\theta=\sqrt{\frac{g_{\theta\theta}}{\mathrm{H}_{\Theta\Theta}}}\frac{\partial\theta}{\partial\Theta}=\sqrt{\frac{g_{\theta\theta}}{\mathrm{H}_{\Theta\Theta}}}, \tag{4}$$
$$\lambda_z=\sqrt{\frac{g_{zz}}{H_{ZZ}}}\frac{\partial z}{\partial Z}=\alpha\cdot\sqrt{\frac{g_{zz}}{H_{ZZ}}}.$$

If we set $r\equiv R$ and $z\equiv Z$ in the above equations, they give the stretching ratios Λ_K of the physiologically normal state referred to as the Riemannian stress-free configuration:

$$\Lambda_R=\sqrt{\frac{G_{RR}}{H_{RR}}},\quad \Lambda_\Theta=\sqrt{\frac{G_{\Theta\Theta}}{H_{\Theta\Theta}}},\quad \Lambda_Z=\sqrt{\frac{G_{ZZ}}{H_{ZZ}}}. \tag{5}$$

If there is no shear in the residual strain, we obtain the metric tensor from the above equations:

$$H_{RR}=\frac{1}{\Lambda_R^2},\quad H_{\Theta\Theta}=\frac{R^2}{\Lambda_\Theta^2},\quad H_{ZZ}=\frac{1}{\Lambda_Z^2},\quad H_{KL}=0\ (K\neq L). \tag{6}$$

The uniform strain (stretch) hypothesis means that Λ_Θ is constant through the wall thickness. Since the incompressibility $\Lambda_R\Lambda_\Theta\Lambda_Z=1$, and Λ_Z is constant, the radial stretching ratio Λ_R is also constant through the wall thickness.

In contrast with a similar spherical model of the left ventricle (Takamizawa and Matsuda, 1990), the Riemann-Christoffel tensor, which indicates the curvature of the stress-free configuration, vanishes for this thick-walled model of the artery. This is evident from the following transformation of the coordinates:

$$\hat{R}=\Lambda_\Theta\Lambda_Z R,\ \hat{\Theta}=\frac{1}{\Lambda_\Theta^2\Lambda_Z}\Theta,\ \hat{Z}=\frac{1}{\Lambda_Z}Z. \tag{7}$$

With this transformation, the metric tensor of the Riemannian stress-free configuration is transformed into

$$\hat{H}_{RR}=1,\ \hat{H}_{\Theta\Theta}=\hat{R}^2,\ \hat{H}_{ZZ}=1,\ \hat{\mathrm{H}}_{KL}=0(K\neq L). \tag{8}$$

This coordinate system [$\hat{R}$, $\hat{\Theta}$, $\hat{Z}$; $\hat{H}_{RR}=1$, $\hat{H}_{\Theta\Theta}=\hat{R}^2$, $\hat{H}_{ZZ}=1$, $\hat{H}_{KL}=0$ $(K\neq L)$] coincides with the cylindrical coordinate system in Euclidean space, although $\hat{\Theta}$ can take only a value between 0 and $\Phi=2\pi/(\Lambda_O^2\Lambda_Z)$ while Θ lies between 0 and 2π. The Riemannian stress-free configuration is a flat space, but not globally Euclidean. Indeed, it is the product space of one-dimensional Euclidean space and a zone of a cone. The development of this zone of a cone is shown in Fig. 2. It should be noted that the cross-section ($Z=$constant) is regarded as continuous at $\hat{\Theta}=0$ and $\hat{\Theta}=\Phi$. In this cross-section, the parallel translation of a vector along any closed curve, including the Z-axis, makes the angle $\Psi=2\pi-\Phi$ between the starting vector and that of returning to the start point. The section of the Riemannian stress-free configuration is isometric to a sector with the angle Φ, which is the development of the zone of a cone in Fig. 2. Thus, the present analysis is compatible with that of Chuong and Fung (1986) and Takamizawa and Hayashi (1987, 1988).

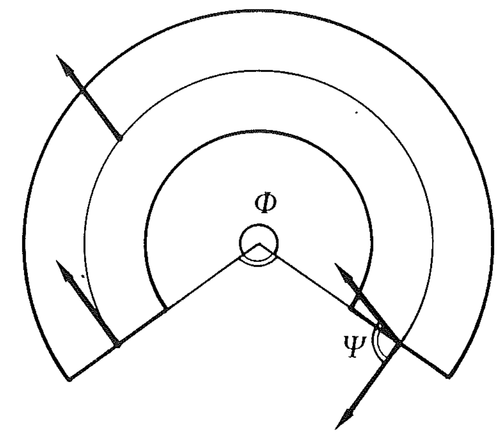


Fig. 2 Cross-section of the stress-free configuration for the cylindrical model of the artery obtained from the uniform strain hypothesis. It is isometric to the sector with a constant thickness, which is a development of a zone of a cone. Then a parallel translation of a vector along a circle, including the top of the cone, makes an angle between the vector at the starting point by the translation.

Strain Energy Density Function

We shall take the following strain energy density function for the arterial wall (Chuong and Fung, 1986)

$$W=C\cdot\exp Q, \tag{9}$$
$$Q=\frac{1}{2}a^{KLMN}E_{KL}E_{MN},$$

where C is a constant which has the dimension of energy density, and a^{*KLMN} the contravariant components of tensor of rank 4 in the Riemannian stress-free configuration. The physical components of this tensor represent the material constants of the arterial wall.

For the incompressible material, Cauchy's stress is derived from the strain energy density function as follows:

$$t^{kl}=-\Pi g^{kl}+\frac{\partial x^k}{\partial X^K}\frac{\partial x^l}{\partial X^L}\frac{\partial W}{\partial E_{KL}}$$
$$=-\Pi g^{kl}+C\cdot\exp Q\cdot\frac{\partial x^k}{\partial X^K}\frac{\partial x^l}{\partial X^L}a^{KLMN}E_{MN}, \tag{10}$$

where Π is the hydrostatic term introduced from the incompressibility condition.

Conclusion

Since it is very plausible that there are residual stresses in living organs, the continuum theory of dislocations based on Riemann geometry may be an appropriate tool in biomechanics. But it should be noted that we need to estimate the distribution of residual strain, or a principle, in order to determine the metric tensor of the stress-free configuration—because the continuum theory of dislocations gives only a kinematical framework of the analysis. The uniform strain hypothesis is one of such principles. Nevertheless, the hypothesis may not be applied to a complex shape model. A principle which is applicable to a general shape has been proposed in a paper by Takamizawa and Matsuda (1990).

Acknowledgment

This study was supported by Grant-in-Aid for Encouragement of Young Scientists from the Ministry of Education, Science, and Culture of Japan.

References

Carew, T. E., Vaishnav, R. N., and Patel, D. J., 1968, "Compressibility of the Arterial Wall," *Circulation Research*, Vol. 23, pp. 61–68.

Chuong, C. J., and Fung, Y. C., 1986, "On Residual Stress in Arteries," *ASME Journal of Biomechanical Engineering*, Vol. 108, pp. 189–192.

Fung, Y. C., 1984, *Biodynamics: Circulation*, Springer-Verlag, New York, pp. 53–68.

Leung, D. Y. M., Glagov, S., and Mathews, M. B., 1975, "Cyclic Stretching Stimulates Synthesis of Matrix Components by Arterial Smooth Muscle Cells in vitro," *Science*, Vol. 191, pp. 475–477.

Takamizawa, K., and Hayashi, K., 1987, "Strain Energy Density Function and Uniform Strain Hypothesis for Arterial Mechanics," *Journal of Biomechanics*, Vol. 20, pp. 7–17.

Takamizawa, K., and Hayashi, K., 1988, "Uniform Strain Hypothesis and Thin-Walled Theory in Arterial Mechanics," *Biorheology*, Vol. 25, pp. 555–565.

Takamizawa, K., and Matsuda, T., 1990, "Kinematics for Bodies Undergoing Residual Stress and Its Applications to the Left Ventricle," ASME Journal of Applied Mechanics, Vol. 57, pp. 321–329.

Vaishnav, R. N., and Vossoughi, J., 1987, "Residual Stress and Strain in Aortic Segments," *Journal of Biomechanics*, Vol. 20, pp. 235–239.

Bending of a Thin Reissner Plate With a Through Crack[14]

P. F. Joseph[15] and F. Erdogan[15,16]

The title problem was first considered by Knowles and Wang (1960) and was shown to be related to the solution given by the classical plate theory. This solution is actually the outer solution of a singular perturbation problem, and therefore is valid only away from the crack-tip region. Within a boundary layer of order h/a, where h is the plate thickness and a is the half-crack length, the two theories differ considerably. In this study the leading order solution is obtained for h/a→0 and it is shown that the limiting stress intensity factor given by the Reissner plate theory is more than 50 percent higher than the asymptotic result (1+ν)/(3+ν) which is obtained from the displacement field as given by the classical plate theory.

1 Introduction

The problem of bending of an elastic plate with a crack as formulated by the Reissner plate theory was first presented by Knowles and Wang (1960) in terms of a singular integral equation. An important contribution from their work was to show that for a vanishingly thin plate, the Reissner plate theory for bending predicts a stress field near the vertex of a crack that is in accordance with the theory of elasticity. They thus showed that the Reissner plate theory can compensate for the well-documented deficiencies of the classical theory (see Williams, 1961). In the case of a thin plate, or when the ratio of the plate thickness h to the half-crack length a approaches zero, Knowles and Wang obtained an approximate closed-form solution to the integral equation by formally setting h/a equal to zero in the Fredholm kernel. This limit transforms the nonsingular Fredholm integral into a singular integral which, in this case, is valid everywhere except within an order of h/a of the crack tips (see Joseph and Erdogan, 1989). The approximate "asymptotic" solution which results, and which is identical to the solution as given by the classical plate theory, is therefore not valid when close to the crack tips. As pointed out by Simmonds and Duva (1981) it is interesting that Knowles and Wang were able to obtain their results, which are valid in the near field, by using a solution valid only in the far field.

The problem is one of singular perturbation with perturbation parameter h/a. There exists a boundary layer of thickness h/a at the crack tips and the classical plate theory solution is the leading order outer solution. The boundary layer difference between the theories exists because the Reissner plate theory allows for the satisfaction of three boundary conditions on the crack surfaces ($N_{12}=0$, $V_1=0$, $M_{12}=0$), while the classical plate theory, which uses the Kirchoff assumption, satisfies only two ($N_{12}=0$, $V_1+\partial M_{12}/\partial x_2=0$). This is true for all h/a, including the limit as $h/a\to 0$. In a region near the crack tips the two solutions must be different; away from these boundaries they asymptotically agree. Consequently, it may be shown that the stress intensity factors obtained from the two theories (as defined in terms of the crack surface displacement in the case of the classical plate theory) are different; that is, the stress intensity factor given by the Reissner plate theory as $h/a\to 0$ is not $(1+\nu)/(3+\nu)$.

The Fredholm kernel of the singular integral equation is sufficiently complicated to allow only a numerical solution for a given Poisson's ratio ν and plate thickness to half-crack length ratio h/a. In terms of numerical convergence, the problem becomes increasingly difficult as $h/a\to 0$. Hartranft and Sih (1968) and Wang (1968) were the first to numerically obtain a solution for finite h/a. In both studies the problem was formulated in terms of dual integral equations and the argument presented by Knowles and Wang (1960) was followed to obtain the "thin plate limit." Some doubt about this limit was raised in a paper by Krenk (1978). Basar and Erdogan (1982), with the help of more refined numerical solutions for h/a as small as 0.01, also suggested that the thin plate limit $(1+\nu)/(3+\nu)$ may not be valid. See also Simmonds and Duva (1981) for a discussion of this boundary layer problem and other references.

In this study the $h/a\to 0$ limit of the integral equation for the derivative of the crack surface rotation as defined by the Reissner's plate theory is obtained and numerically solved. The problem is also set up for the determination of higher-order terms in the perturbation series. Limiting values of the stress intensity factors are presented and the solution is compared to small h/a numerical solutions in order to show the boundary layer nature of the problem.

2 Formulation

Consider a homogeneous elastic plate of thickness h that contains a through crack of length $2a$ in the x_1x_2-plane and is subjected to uniform bending $M_{11}=M_0$ away from the crack region. By applying the Reissner plate theory, the nondimensional integral equation for the derivative of the crack surface rotation β_r is found to be (Knowles and Wang, 1960):

$$\frac{1}{\pi}\int_{-1}^{1}\frac{g_r(t)}{t-y}\,dt+\frac{1/\epsilon}{2\pi(1+\nu)}\int_{-1}^{1}g_r(t)K(z)dt = -1,\ |y|<1,\ y=x_2/a, \tag{1}$$

where

$$K(z)=\frac{-16}{z^3}+\frac{4}{z}+\frac{8}{z}K_2(|z|),$$

$$z=\frac{1}{\epsilon}(t-y),\ \epsilon=h/(\sqrt{10}a), \tag{2}$$

$$g_r(t)=\frac{Eh}{4\sigma a}\frac{d\beta_r}{dt},\ \sigma=\frac{6M_o}{h^2}. \tag{3}$$

$K_2(|z|)$ is the modified Bessel function of the second kind

[14]This study was supported by NSF Grant MSM-8613611 and by NASA-Langley under Grant NAG-1-713.

[15]Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA 18015.

[16]Fellow ASME.

References

Carew, T. E., Vaishnav, R. N., and Patel, D. J., 1968, "Compressibility of the Arterial Wall," *Circulation Research*, Vol. 23, pp. 61–68.

Chuong, C. J., and Fung, Y. C., 1986, "On Residual Stress in Arteries," *ASME Journal of Biomechanical Engineering*, Vol. 108, pp. 189–192.

Fung, Y. C., 1984, *Biodynamics: Circulation*, Springer-Verlag, New York, pp. 53–68.

Leung, D. Y. M., Glagov, S., and Mathews, M. B., 1975, "Cyclic Stretching Stimulates Synthesis of Matrix Components by Arterial Smooth Muscle Cells in vitro," *Science*, Vol. 191, pp. 475–477.

Takamizawa, K., and Hayashi, K., 1987, "Strain Energy Density Function and Uniform Strain Hypothesis for Arterial Mechanics," *Journal of Biomechanics*, Vol. 20, pp. 7–17.

Takamizawa, K., and Hayashi, K., 1988, "Uniform Strain Hypothesis and Thin-Walled Theory in Arterial Mechanics," *Biorheology*, Vol. 25, pp. 555–565.

Takamizawa, K., and Matsuda, T., 1990, "Kinematics for Bodies Undergoing Residual Stress and Its Applications to the Left Ventricle," ASME JOURNAL OF APPLIED MECHANICS, Vol. 57, pp. 321–329.

Vaishnav, R. N., and Vossoughi, J., 1987, "Residual Stress and Strain in Aortic Segments," *Journal of Biomechanics*, Vol. 20, pp. 235–239.

Bending of a Thin Reissner Plate With a Through Crack[14]

P. F. Joseph[15] and F. Erdogan[15,16]

The title problem was first considered by Knowles and Wang (1960) and was shown to be related to the solution given by the classical plate theory. This solution is actually the outer solution of a singular perturbation problem, and therefore is valid only away from the crack-tip region. Within a boundary layer of order h/a, where h is the plate thickness and a is the half-crack length, the two theories differ considerably. In this study the leading order solution is obtained for $h/a \to 0$ and it is shown that the limiting stress intensity factor given by the Reissner plate theory is more than 50 percent higher than the asymptotic result $(1+\nu)/(3+\nu)$ which is obtained from the displacement field as given by the classical plate theory.

1 Introduction

The problem of bending of an elastic plate with a crack as formulated by the Reissner plate theory was first presented by Knowles and Wang (1960) in terms of a singular integral equation. An important contribution from their work was to show that for a vanishingly thin plate, the Reissner plate theory for bending predicts a stress field near the vertex of a crack that is in accordance with the theory of elasticity. They thus showed that the Reissner plate theory can compensate for the well-documented deficiencies of the classical theory (see Williams, 1961). In the case of a thin plate, or when the ratio of the plate thickness h to the half-crack length a approaches zero, Knowles and Wang obtained an approximate closed-form solution to the integral equation by formally setting h/a equal to zero in the Fredholm kernel. This limit transforms the nonsingular Fredholm integral into a singular integral which, in this case, is valid everywhere except within an order of h/a of the crack tips (see Joseph and Erdogan, 1989). The approximate "asymptotic" solution which results, and which is identical to the solution as given by the classical plate theory, is therefore not valid when close to the crack tips. As pointed out by Simmonds and Duva (1981) it is interesting that Knowles and Wang were able to obtain their results, which are valid in the near field, by using a solution valid only in the far field.

The problem is one of singular perturbation with perturbation parameter h/a. There exists a boundary layer of thickness h/a at the crack tips and the classical plate theory solution is the leading order outer solution. The boundary layer difference between the theories exists because the Reissner plate theory allows for the satisfaction of three boundary conditions on the crack surfaces ($N_{12}=0,\ V_1=0,\ M_{12}=0$), while the classical plate theory, which uses the Kirchoff assumption, satisfies only two ($N_{12}=0,\ V_1+\partial M_{12}/\partial x_2=0$). This is true for all h/a, including the limit as $h/a \to 0$. In a region near the crack tips the two solutions must be different; away from these boundaries they asymptotically agree. Consequently, it may be shown that the stress intensity factors obtained from the two theories (as defined in terms of the crack surface displacement in the case of the classical plate theory) are different; that is, the stress intensity factor given by the Reissner plate theory as $h/a \to 0$ is not $(1+\nu)/(3+\nu)$.

The Fredholm kernel of the singular integral equation is sufficiently complicated to allow only a numerical solution for a given Poisson's ratio ν and plate thickness to half-crack length ratio h/a. In terms of numerical convergence, the problem becomes increasingly difficult as $h/a \to 0$. Hartranft and Sih (1968) and Wang (1968) were the first to numerically obtain a solution for finite h/a. In both studies the problem was formulated in terms of dual integral equations and the argument presented by Knowles and Wang (1960) was followed to obtain the "thin plate limit." Some doubt about this limit was raised in a paper by Krenk (1978). Basar and Erdogan (1982), with the help of more refined numerical solutions for h/a as small as 0.01, also suggested that the thin plate limit $(1+\nu)/(3+\nu)$ may not be valid. See also Simmonds and Duva (1981) for a discussion of this boundary layer problem and other references.

In this study the $h/a \to 0$ limit of the integral equation for the derivative of the crack surface rotation as defined by the Reissner's plate theory is obtained and numerically solved. The problem is also set up for the determination of higher-order terms in the perturbation series. Limiting values of the stress intensity factors are presented and the solution is compared to small h/a numerical solutions in order to show the boundary layer nature of the problem.

2 Formulation

Consider a homogeneous elastic plate of thickness h that contains a through crack of length $2a$ in the x_1x_2-plane and is subjected to uniform bending $M_{11}=M_0$ away from the crack region. By applying the Reissner plate theory, the nondimensional integral equation for the derivative of the crack surface rotation β_r is found to be (Knowles and Wang, 1960):

$$\frac{1}{\pi}\int_{-1}^{1}\frac{g_r(t)}{t-y}\,dt + \frac{1/\epsilon}{2\pi(1+\nu)}\int_{-1}^{1} g_r(t)K(z)\,dt = -1,\ |y|<1,\ y=x_2/a, \quad (1)$$

where

$$K(z) = \frac{-16}{z^3} + \frac{4}{z} + \frac{8}{z}K_2(|z|),$$

$$z = \frac{1}{\epsilon}(t-y),\ \epsilon = h/(\sqrt{10}a), \quad (2)$$

$$g_r(t) = \frac{Eh}{4\sigma a}\frac{d\beta_r}{dt},\ \sigma = \frac{6M_o}{h^2}. \quad (3)$$

$K_2(|z|)$ is the modified Bessel function of the second kind

[14]This study was supported by NSF Grant MSM-8613611 and by NASA-Langley under Grant NAG-1-713.

[15]Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA 18015.

[16]Fellow ASME.

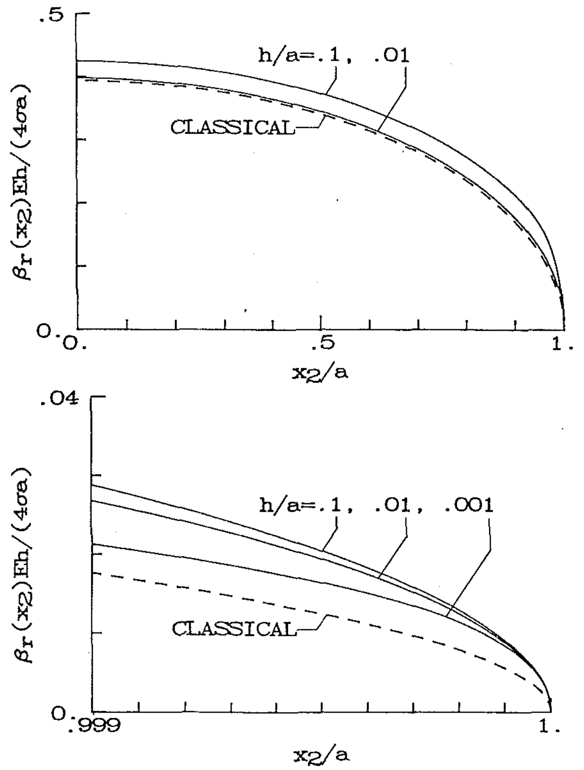


Fig. 1 Normalized crack surface rotation in a plate with a through crack, $-a<x_2<a$, under pure bending, $\overset{\infty}{M}_{xx}=M_0$; $\nu=0.3$, $\sigma=6M_0/h^2$

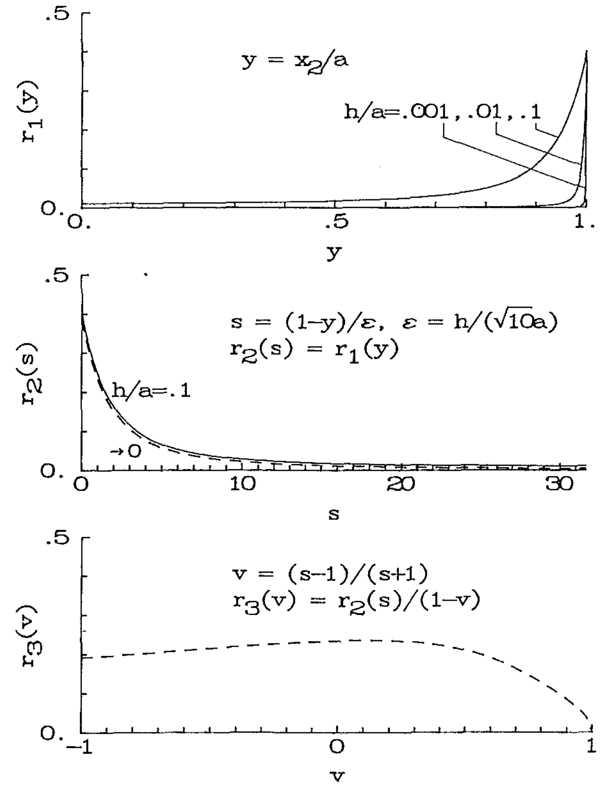


Fig. 2 The right-hand sides $r_1(y)$, $r_2(s)$, and $r_3(v)$ for the integral equations (14), (22), and (25), respectively. Dashed lines correspond to $h/a \to 0$.

which has small z behavior that cancels the apparent singular terms of the Fredholm kernel, $K(z)$. For small z,

$$K(z) = -z\log(|z|/2) + 0(z). \tag{4}$$

The integral equation for the classical plate theory is

$$\frac{3+\nu}{1+\nu}\frac{1}{\pi}\fint_{-1}^{1}\frac{g_c(t)}{t-y}\,dt = -1, \tag{5}$$

where

$$g_c(t) = \frac{Eh}{4\sigma a}\frac{d\beta_c}{dt}. \tag{6}$$

The subscripts r and c refer to the Reissner and the classical plate theories, respectively. Note that equation (5) can be obtained by formally substituting $\epsilon=0$ into equation (1). Equation (5) has a closed-form solution given by

$$g_c(y) = \frac{1+\nu}{3+\nu}\frac{-y}{\sqrt{1-y^2}}, \tag{7}$$

from which it follows that

$$\beta_c(y) = \frac{4\sigma}{E}\frac{1+\nu}{3+\nu}\frac{1}{h}\sqrt{a^2-x_2^2}. \tag{8}$$

We also note that the right-hand sides in equations (1) and (5) represent the bending moment $M_{11}(0, x_2)$ for $|x_2|>a$ as well as $|x_2|<a$. Thus, from equations (5) and (7) the stress intensity factor at the crack tip $x_2=a$ may be obtained as

$$k_1(x_3) = \frac{x_3}{h/2}\sigma\sqrt{a} \tag{9}$$

based on the definition

$$k_1(x_3) = \lim_{x_2\to a}\sqrt{2(x_2-a)}\,\sigma_{11}(0, x_2, x_3),\ x_2>a, \tag{10}$$

and

$$k_1(x_3) = \frac{1+\nu}{3+\nu}\frac{x_3}{h/2}\sigma\sqrt{a} \tag{11}$$

based on the definition

$$k_1(x_3) = -\lim_{x_2\to a}\frac{E}{2}\sqrt{2(a-x_2)}\frac{\partial}{\partial x_2}u_1(0, x_2, x_3),\ x_2<a, \tag{12}$$

where it is assumed that $u_1=x_3\beta_c$. Needless to say, in continuum elasticity solutions, (10) and (12) would give identical results.

The numerically obtained solution of (1) for $\nu=0.3$ and $\sqrt{10}\,\epsilon=h/a=0.1$, 0.01 and 0.001 is compared with the expression given by equation (8) in Fig. 1. It may be observed that as $\epsilon\to 0$ the two curves approach each other everywhere except very near the crack tip as shown in the lower figure where this region is expanded. If we now let

$$g_r(t) = g_c(t) + g(t), \tag{13}$$

and then substitute this expression into equation (1) while making use of (7), we obtain an integral equation for the difference between the two theories as follows:

$$\frac{\epsilon}{\pi}\fint_{-1}^{1}\frac{g(t)}{t-y}\,dt + \frac{1}{2\pi(1+\nu)}\int_{-1}^{1}g(t)K(z)dt = \frac{-2\epsilon}{3+\nu} + \frac{1}{2\pi(3+\nu)}\int_{-1}^{1}\frac{tK(z)}{\sqrt{1-t^2}}\,dt = \frac{-2}{3+\nu}\sqrt{\epsilon}\,r_1(y), \tag{14}$$

where

$$r_1(y) = \sqrt{\epsilon}\left\{1 - \frac{1}{4\pi\epsilon}\int_{-1}^{1}\frac{tK(z)}{\sqrt{1-t^2}}\,dt\right\}. \tag{15}$$

The function $r_1(y)$ is given at the top of Fig. 2 for $\sqrt{10}\,\epsilon=$

$h/a=0.1$, 0.01 and 0.001. As $\epsilon \to 0$ the only nonzero contribution of $r_1(y)$ is within a region of order ϵ of the crack tip. As will be shown below, this contribution cannot be ignored if the correct solution near the crack tip is desired. Note that this figure clearly shows the boundary layer nature of the problem, and why for small values of ϵ, the solution of equation (1) is difficult to obtain by using standard numerical techniques which treat the domain $-1<y<1$, on a nearly equal basis. A method that expands the boundary layer region is obviously required for small ϵ. This difficulty is common to many other crack problems that can be defined by a small geometry parameter such as ϵ.

Noting that $g(t)$ is an odd function of t, the relation

$$\int_{-1}^{1} g(t)L(t-y)dt = \int_{0}^{1} g(t)[L(t-y)-L(-t-y)]dt, \quad (16)$$

which is true for any function L, may be used to change the support of the integrals in equation (14) from -1 to 1, to 0 to 1. Then from equation (14), after making the variable change

$$t = 1-\epsilon r, \quad y = 1-\epsilon s, \quad (17)$$

we obtain the following integral equation valid for all ϵ,

$$\frac{1}{\pi} \int_{0}^{1/\epsilon} \frac{h(r)}{r-s} dr - \frac{1}{2\pi(1+\nu)} \int_{0}^{1/\epsilon} h(r) \times \left[K[|s-r|] - K[|r+s-2/\epsilon|] + \frac{2(1+\nu)\epsilon}{2-(r+s)\epsilon} \right] dr = \frac{2}{3+\nu} r_1(y),\ y=1-\epsilon s,\ 0<s<1/\epsilon, \quad (18)$$

where

$$h(r) = \sqrt{\epsilon}\, g(t). \quad (19)$$

These transformations are made so that $r=0$ corresponds to the crack tip and $r=1/\epsilon$ (later $r\to\infty$) corresponds to the crack center while the right-hand side, $r_1(y)=r_1(s)$, decreases for increasing s (later $r_1(s)\to 0$ for $s\to\infty$).

At this point we consider the solution of equation (18) for small ϵ. Before we solve the problem for the limit as $\epsilon\to 0$, which is the purpose of this study, we consider a more rigorous method that would solve the asymptotic problem for small ϵ. One possible approach for small ϵ would be to use the substitution

$$r = \frac{1+u}{1-u+2\epsilon}, \quad s = \frac{1+v}{1-v+2\epsilon}, \quad (20)$$

which would convert the integral equation (18) to an equation valid for $-1<(u,v)<1$. The asymptotic behavior of the kernels and right-hand side of the resulting equation would then have to be determined. Once these expressions have been found, a proper choice of an asymptotic expansion of the unknown can be made. Substitution of this expansion into the integral equation would give individual equations for the determination of the terms in the series, which are then hopefully solved numerically. Unfortunately, this procedure becomes very complicated after the leading-order term because of the behavior of the kernels at $v=1$ (or $s=1/\epsilon$) which corresponds to the center of the crack. For higher-order terms perhaps a matching technique would be more appropriate. This would separate the boundary layer solution from the solution for the central portion of the crack. The technique of asymptotic matching is not used in this study.

A second method that may give approximate solutions to higher-order terms would be to set the upper limits of the integrals in equation (18) to infinity. The justification for this is that we are only interested in the stress intensity factor which comes from the solution for $s=0$. This approximation to the integral equation would have its greatest influence on the unknown for $s\to\infty$ where we are not really interested in the solution. Recall that the boundary layer to the original problem is at $s=0$ and that the solution decays to zero for large values of s. In this case in order to determine the form of the perturbation series for $h(r)$, only the small ϵ behavior of $r_1(y)$ must first be obtained. The leading term of $r_1(y)$ is of order one. From numerical studies it appears that the second term is of order ϵ, not $\sqrt{\epsilon}$ as may first be assumed by a casual inspection of equation (15). Therefore, the following expression will be chosen to represent $h(r)$ for small ϵ, (and perhaps $r \ll 1/\epsilon$),

$$h(r) = h_0(r) + \epsilon\, h_1(r) + \ldots. \quad (21)$$

By using standard perturbation methods, this expression may be substituted into equation (18) to obtain integral equations for each of the $h_i(r)$, where $h_i(r)$ is dependent on the solutions $h_j(r)$ for $j=0, 1, \ldots, i-1$. For example, the integral equation for $h_0(r)$ is determined by simply retaining terms of order one as $\epsilon\to 0$. Once $h_0(r)$ is known, the integral equation for $h_1(r)$ can be obtained by retaining the order ϵ terms, and so on. The equation for $h_0(r)$ is:

$$\frac{1}{\pi} \int_{0}^{\infty} \frac{h_0(r)}{r-s} dr - \frac{1}{2\pi(1+\nu)} \int_{0}^{\infty} h_0(r)K(s-r)dr = \frac{2}{3+\nu} r_2(s),\ 0<s<\infty, \quad (22)$$

where

$$r_2(s) = \lim_{\epsilon\to 0} r_1(y) = \lim_{\epsilon\to 0} \sqrt{\epsilon} \times \left\{ 1 - \frac{1}{4\pi\sqrt{\epsilon}} \int_{0}^{1/\epsilon} \frac{1-\epsilon r}{\sqrt{1-\epsilon r/2}} \frac{K(s-r)-K(s+r-2/\epsilon)}{\sqrt{2r}} dr \right\} = \frac{-1}{4\pi} \int_{0}^{\infty} \frac{1}{\sqrt{2r}} K(s-r)dr,\ y=1-\epsilon s. \quad (23)$$

The function $r_2(s)$ is given by the dashed line in the center plot of Fig. 2. Also included in the figure is $r_1(y)$ for $h/a=0.1$ ($y=1-\epsilon s$, $\epsilon=h/(\sqrt{10}a)$). The two curves for $h/a=0.01$ and 0.001 lie between these two lines and are not included.

The integral equation for $h_1(r)$ is similar to equation (22) except that its right-hand side is more complicated. In addition to the contribution coming from $r_1(y)$, namely

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \left\{ r_1(y) + \frac{1}{4\pi} \int_{0}^{\infty} \frac{1}{\sqrt{2r}} K(s-r)dr \right] \Bigg\},\ y = 1-\epsilon s, \quad (24)$$

the contribution from $h_0(r)$ due to the second and third terms of the Fredholm kernel of equation (18) must also be taken into account. The solution for $h_1(r)$ will not be given in this study.

3 Solution and Results

The leading order solution $h_0(r)$ will now be determined. A second change in variables is used to rewrite equation (22) in a more convenient form for numerical solution as follows:

$$\frac{1}{\pi} \int_{-1}^{1} \frac{f(u)du}{(1-u)(u-v)} - \frac{1}{\pi(1+\nu)} \int_{-1}^{1} \frac{f(u)}{(1-u)} \bar{K}(u,v)\, du = \frac{2}{3+\nu} r_3(v), \quad (25)$$

where

$$r = \frac{1+u}{1-u}, \quad s = \frac{1+v}{1-v}, \quad (26)$$

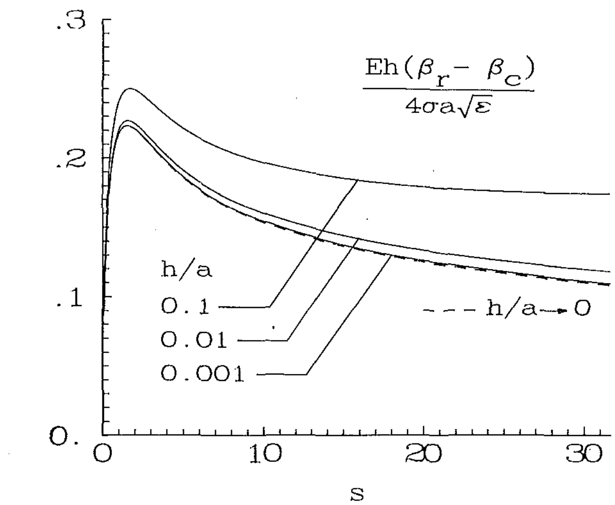


Fig. 3 Difference between the normalized crack surface rotation of the Reissner's plate theory and the classical plate theory for a plate with a through crack, $-a<x_2<a$, under pure bending, $\overset{\infty}{M}_{xx}=M_0$; $\nu=0.3$, $\sigma=6M_0/h^2$, $x_2=a(1-\epsilon s)$, $\epsilon=h/\sqrt{10}a$

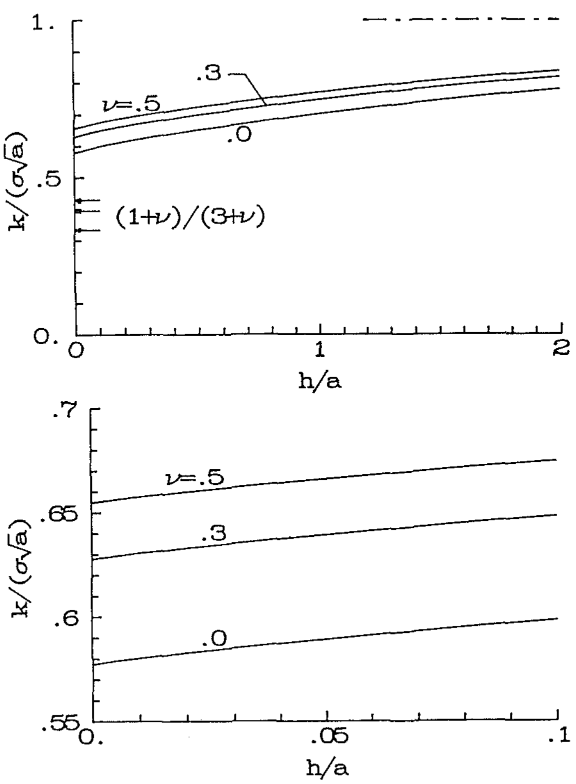


Fig. 4 Normalized stress intensity factors in a plate under bending, $\sigma=6M_0/h$

$$\bar{K}(u,v) = \frac{-2(1-u)^2(1-v)^2}{(v-u)^3} + \frac{2}{v-u} + \frac{4}{v-u}K_2\left|\frac{2(v-u)}{(1-u)(1-v)}\right|, \quad (27)$$

$$r_3(v) = \frac{r_2(s)}{1-v},\ f(u) = h_0(r). \quad (28)$$

The right-hand side of equation (25) is given at the bottom of Fig. 2.

Equation (25) is numerically solved by letting

Table 1 The effect of Poisson's ratio ν and crack-length-to-plate-thickness ratio h/a on the normalized bending stress intensity factor; $\sigma=6M_o/h^2$

h/a	$\frac{k_1(h/2)}{\sigma\sqrt{a}}$, $\nu=0$	$\nu=.3$	$\nu=.5$
$\to\infty$	1.0000	1.0000	1.0000
20.0	0.9851	0.9885	0.9900
10.0	0.9583	0.9676	0.9717
4.0	0.8735	0.8992	0.9111
2.0	0.7804	0.8193	0.8383
1.0	0.7020	0.7475	0.7707
0.5	0.6518	0.6997	0.7247
0.25	0.6211	0.6701	0.6960
0.10	0.5984	0.6481	0.6746
0.01	0.5803	0.6306	0.6575
0.005	0.5790	0.6292	0.6562
0.001	0.5777	0.6280	0.6550
$\to 0$	0.5774	0.6277	0.6547

Table 2 The effect of Poisson's ratio ν on the normalized bending stress intensity factor for the limiting case of $h/a\to 0$ for the Reissner's plate theory; $\sigma=6M_o/h^2$

ν	$\frac{k_1(h/2)}{\sigma\sqrt{a}}$
0.0	0.5774
0.1	0.5957
0.2	0.6124
0.25	0.6202
0.3	0.6277
1.3	0.6325
0.4	0.6417
0.5	0.6547

$$\frac{f(u)}{1-u} = \frac{(1-u)^{\alpha}}{(1+u)^{1/2}}q(u), \quad (29)$$

and writing

$$q(u) = \sum_{n=0}^{N} A_n P_n^{(\alpha,-1/2)}(u), \quad (30)$$

where $P_n^{(\alpha,\beta)}$ are Jacobi polynomials. The value of α that gives the weight function associated with the solution of equation (25) is $+1/2$. The equation can be solved for the unknown coefficients by using the collocation method (see collocation method and Sec. 5.10 in Baker, 1977). The computational effort required for solving equation (25) is more extensive than the numerical solution of equation (1) for $h/a=0.01$, but less difficult than for $h/a=0.001$. The unknown given by equation (29) could also be represented by using $\alpha=-1/2$. For this case the extra condition $q(+1)=0$ must be imposed on the solution. In terms of convergence and numerical efficiency, there is very little difference between using $\alpha=+1/2$ or $\alpha=-1/2$. The advantage of using $\alpha=+1/2$ is that a nonzero value of $q(+1)$ is obtained.

In order to compare the boundary layer solution to the solution for small ϵ, we first use the definitions (3), (6), and (13) along with substitutions from (17), (19), (26), (28), and (29) to obtain the following expression for the difference between the two plate theories in the limit as $\epsilon\to 0$:

$$\frac{Eh}{4\sigma a}(\beta_r-\beta_c) = -\int_y^1 g(t)dt = -2\sqrt{\epsilon}\int_{-1}^{v}\frac{(1-u)^{\alpha-1}}{(1+u)^{1/2}}q(u)du. \quad (31)$$

In Fig. 3 this quantity, divided by $\sqrt{\epsilon}$, is compared to that obtained from equations (1), (3), and (8) for $\sqrt{10}\epsilon=h/a=0.1$, 0.01, and 0.001. The dashed line in this figure corresponds to equation (31).

From the solution of equation (25) we can also obtain the contribution to the stress intensity factor corresponding to $g(t)$ which is defined by equation (13). The total stress intensity factor may then be determined from equation (12) as follows:

$$\frac{k_1}{\sigma\sqrt{a}} = \frac{x_3}{h/2}\left[\frac{1+\nu}{3+\nu} - 2^{1+\alpha}q(-1)\right]. \qquad (32)$$

The stress intensity factors obtained from the Reissner's plate theory for $x_3=h/2$ and $\nu=0.0$, 0.3, 0.5 as functions of h/a are given in Table 1 and Fig. 4. The lower part of Fig. 4 shows that no discontinuity in the stress intensity factor exists for the limiting case of $h/a\to 0$. The boundary layer effects do not seem to be apparent in this quantity. They are, however, apparent in the solution $g_r(t)$ of the integral equation (1) as can be seen from Figs. 1 and 3, (see also equation (3)). In Table 2 the limiting stress intensity factors for various Poisson's ratios are given. It should be emphasized that, unlike the classical plate theory, with the Reissner plate theory both definitions (10) and (12) would give the same results (Joseph and Erdogan, 1989). It should also be pointed out that the through crack results in a plate under bending considered in this study and elsewhere are meaningful only if the plate is under tensile membrane loading of sufficiently high magnitude to keep the crack surfaces fully open. The contact problem for a plate under bending was recently considered by Joseph and Erdogan (1989).

References

Baker, C. T. H., 1977, *The Numerical Treatment of Integral Equations*, Clarendon Press, Oxford.

Civelek, M. B., and Erdogan, F., 1982, "Elastic-Plastic Problem for a Plate with a Part-Through Crack Under Extension and Bending," *International Journal of Fracture*, Vol. 20, pp. 33–46.

Hartranft, R. J., and Sih, G. C., 1968, "Effect of Plate Thickness on the Bending Stress Distribution Around Through Cracks," *Journal of Mathematics and Physics*, Vol. 47, pp. 276–291.

Joseph, P. F., and Erdogan, F., 1989, "Surface Crack Problems in Plates," *International Journal of Fracture*, Vol. 41, pp. 105–131.

Knowles, J. K., and Wang, N. M., 1960, "On the Bending of an Elastic Plate Containing a Crack," *Journal of Mathematics and Physics*, Vol. 39, pp. 223–236.

Krenk, S., 1978, "Influence of Transverse Shear on an Axial Crack in a Cylindrical Shell," *International Journal of Fracture*, Vol. 14, pp. 123–143.

Simmonds, J. G., and Duva, J., 1981, "Thickness Effects Are Minor in the Energy-Release Rate Integral for Bent Plates Containing Elliptic Holes or Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 320–326.

Wang, N. M., 1968, "Effects of Plate Thickness on the Bending of an Elastic Plate Containing a Crack," *Journal of Mathematics and Physics*, Vol. 47, pp. 371–390.

Williams, M. L., 1961, "The Bending Stress Distribution at the Base of a Stationary Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, pp. 78–82.

Matrix Cracking in Fiber-Reinforced Composite Materials

H. A. Luo[17] and Y. Chen[18]

Matrix cracking is a major pattern of the failure of composite materials. A crack can form in the matrix during manufacturing, or be produced during loading. Erdogan, Gupta, and Ratwani (1974) first considered the interaction between an isolated circular inclusion and a line crack embedded in infinite matrix. As commented by Erdogan et al., their model is applicable to the composite materials which contain sparsely distributed inclusions. For composites filled with finite concentration of inclusions, it is commonly understood that the stress and strain fields near the crack depend considerably on the microstructure around it. One notable simplified model is the so-called three-phase model which was introduced by Christensen and Lo (1979). The three-phase model considers that in the immediate neighborhood of the inclusion there is a layer of matrix material, but at certain distance the heterogeneous medium can be substituted by a homogeneous medium with the equivalent properties of the composite. Thus, for the problems of which the interest is in the field near the inclusion, it can reasonably be accepted as a good model. The two-dimensional version of the three-phase model consists of three concentric cylindrical layers with the outer one, labeled by 3, extended to infinity. The external radii a *and* b *of the inner and intermediate phases, labeled by 1 and 2, respectively, are related by* $(a/b)^2=c$, *where* c *is the volume fraction of the fiber in composite.*

In this Note, we shall confine ourselves to the crack which is located on the x-axis in the interval (t_1, t_2). The governing singular integral equations for unknown complex dislocation density $B(x)=B_x(x)+iB_y(x)$ are

$$\int_{t_1}^{t_2}\frac{B(x_0)}{x_0-x}\,dx_0+\int_{t_1}^{t_2}B(x_0)K_1(x,x_0)dx_0 + \int_{t_1}^{t_2}\overline{B}(x_0)K_2(x,x_0)dx_0=\frac{\kappa_2+1}{2\mu_2}\,\pi p(x) \qquad (t_1<x<t_2) \qquad (1)$$

and

$$\int_{t_1}^{t_2}B(x_0)dx_0=0 \qquad (2)$$

where μ_i and ν_i are the shear modulus and the Poisson's ratio of the ith phase, $\kappa_i=3-4\nu_i$ for the homogeneous phase and $\kappa_i=1+2\mu_{12}^i/K_{12}^i$ for the transversely isotropic phase with μ_{12}^i and K_{12}^i being in-plane shear modulus and plane-strain bulk modulus, respectively, $p(x)=p_x(x)+ip_y(x)$ is the traction on the imaginary crack face in the perfect three-phase solid (Fig. 1). In equation (1) the regular kernel

$$K_1(x,x_0)=\frac{1}{2}\,[K_x(x,x_0)+K_y(x,x_0)],$$

$$K_2(x,x_0)=\frac{1}{2}\,[K_x(x,x_0)-K_y(x,x_0)] \qquad (3)$$

with

$$K_x(x,x_0)=\frac{A+B}{2}\,\frac{1}{x-\dfrac{a^2}{x_0}}+A\,\frac{(x_0^2-a^2)(x-x_0)}{x_0^2(xx_0-a^2)}\,\frac{a^2}{\left(x-\dfrac{a^2}{x_0}\right)^2} -\frac{A+B}{2}\,\frac{1}{x}-\frac{B-A}{2x_0}\,\frac{a^2}{x^2}+A\,\frac{a^2}{x^3}-\frac{1}{2a}\sum_{n=1}^{\infty}a_n'n\left(\frac{x}{a}\right)^n$$
$$+\frac{1}{2a}\sum_{n=1}^{\infty}[a_n'(n-1)+a_{-n}']\left(\frac{x}{a}\right)^{n-2}+\frac{A}{2a}\sum_{n=1}^{\infty}a_{-n}'n\left(\frac{a}{x}\right)^n -\frac{1}{2a}\sum_{n=1}^{\infty}[Aa_{-n}'(n+1)-Ba_n']\left(\frac{a}{x}\right)^{n+2}$$

[17]Postdoctoral Research Associate, Department of Civil Engineering, Case Western Reserve University, Cleveland, OH 44106.

[18]Professor, Department of Mechanics and Materials Science, Rutgers University, Piscataway, NJ 08854, Fellow, ASME.

Manuscript received by the ASME Applied Mechanics Division, May 9, 1989; final revision, May 8, 1990.

From the solution of equation (25) we can also obtain the contribution to the stress intensity factor corresponding to $g(t)$ which is defined by equation (13). The total stress intensity factor may then be determined from equation (12) as follows:

$$\frac{k_1}{\sigma\sqrt{a}} = \frac{x_3}{h/2}\left[\frac{1+\nu}{3+\nu} - 2^{1+\alpha}q(-1)\right]. \tag{32}$$

The stress intensity factors obtained from the Reissner's plate theory for $x_3 = h/2$ and $\nu = 0.0$, 0.3, 0.5 as functions of h/a are given in Table 1 and Fig. 4. The lower part of Fig. 4 shows that no discontinuity in the stress intensity factor exists for the limiting case of $h/a \to 0$. The boundary layer effects do not seem to be apparent in this quantity. They are, however, apparent in the solution $g_r(t)$ of the integral equation (1) as can be seen from Figs. 1 and 3, (see also equation (3)). In Table 2 the limiting stress intensity factors for various Poisson's ratios are given. It should be emphasized that, unlike the classical plate theory, with the Reissner plate theory both definitions (10) and (12) would give the same results (Joseph and Erdogan, 1989). It should also be pointed out that the through crack results in a plate under bending considered in this study and elsewhere are meaningful only if the plate is under tensile membrane loading of sufficiently high magnitude to keep the crack surfaces fully open. The contact problem for a plate under bending was recently considered by Joseph and Erdogan (1989).

References

Baker, C. T. H., 1977, *The Numerical Treatment of Integral Equations*, Clarendon Press, Oxford.

Civelek, M. B., and Erdogan, F., 1982, "Elastic-Plastic Problem for a Plate with a Part-Through Crack Under Extension and Bending," *International Journal of Fracture*, Vol. 20, pp. 33–46.

Hartranft, R. J., and Sih, G. C., 1968, "Effect of Plate Thickness on the Bending Stress Distribution Around Through Cracks," *Journal of Mathematics and Physics*, Vol. 47, pp. 276–291.

Joseph, P. F., and Erdogan, F., 1989, "Surface Crack Problems in Plates," *International Journal of Fracture*, Vol. 41, pp. 105–131.

Knowles, J. K., and Wang, N. M., 1960, "On the Bending of an Elastic Plate Containing a Crack," *Journal of Mathematics and Physics*, Vol. 39, pp. 223–236.

Krenk, S., 1978, "Influence of Transverse Shear on an Axial Crack in a Cylindrical Shell," *International Journal of Fracture*, Vol. 14, pp. 123–143.

Simmonds, J. G., and Duva, J., 1981, "Thickness Effects Are Minor in the Energy-Release Rate Integral for Bent Plates Containing Elliptic Holes or Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 320–326.

Wang, N. M., 1968, "Effects of Plate Thickness on the Bending of an Elastic Plate Containing a Crack," *Journal of Mathematics and Physics*, Vol. 47, pp. 371–390.

Williams, M. L., 1961, "The Bending Stress Distribution at the Base of a Stationary Crack," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, pp. 78–82.

Matrix Cracking in Fiber-Reinforced Composite Materials

H. A. Luo[17] and Y. Chen[18]

Matrix cracking is a major pattern of the failure of composite materials. A crack can form in the matrix during manufacturing, or be produced during loading. Erdogan, Gupta, and Ratwani (1974) first considered the interaction between an isolated circular inclusion and a line crack embedded in infinite matrix. As commented by Erdogan et al., their model is applicable to the composite materials which contain sparsely distributed inclusions. For composites filled with finite concentration of inclusions, it is commonly understood that the stress and strain fields near the crack depend considerably on the microstructure around it. One notable simplified model is the so-called three-phase model which was introduced by Christensen and Lo (1979). The three-phase model considers that in the immediate neighborhood of the inclusion there is a layer of matrix material, but at certain distance the heterogeneous medium can be substituted by a homogeneous medium with the equivalent properties of the composite. Thus, for the problems of which the interest is in the field near the inclusion, it can reasonably be accepted as a good model. The two-dimensional version of the three-phase model consists of three concentric cylindrical layers with the outer one, labeled by 3, extended to infinity. The external radii a *and* b *of the inner and intermediate phases, labeled by 1 and 2, respectively, are related by* $(a/b)^2 = c$, *where* c *is the volume fraction of the fiber in composite.*

In this Note, we shall confine ourselves to the crack which is located on the x-axis in the interval (t_1, t_2). The governing singular integral equations for unknown complex dislocation density $B(x) = B_x(x) + iB_y(x)$ are

$$\int_{t_1}^{t_2} \frac{B(x_0)}{x_0 - x}\,dx_0 + \int_{t_1}^{t_2} B(x_0)K_1(x, x_0)dx_0 + \int_{t_1}^{t_2} \overline{B}(x_0)K_2(x, x_0)dx_0 = \frac{\kappa_2 + 1}{2\mu_2}\,\pi p(x) \quad (t_1 < x < t_2) \tag{1}$$

and

$$\int_{t_1}^{t_2} B(x_0)dx_0 = 0 \tag{2}$$

where μ_i and ν_i are the shear modulus and the Poisson's ratio of the ith phase, $\kappa_i = 3 - 4\nu_i$ for the homogeneous phase and $\kappa_i = 1 + 2\mu_{12}^i/K_{12}^i$ for the transversely isotropic phase with μ_{12}^i and K_{12}^i being in-plane shear modulus and plane-strain bulk modulus, respectively, $p(x) = p_x(x) + ip_y(x)$ is the traction on the imaginary crack face in the perfect three-phase solid (Fig. 1). In equation (1) the regular kernel

$$K_1(x, x_0) = \frac{1}{2}\,[K_x(x, x_0) + K_y(x, x_0)],$$

$$K_2(x, x_0) = \frac{1}{2}\,[K_x(x, x_0) - K_y(x, x_0)] \tag{3}$$

with

$$K_x(x, x_0) = \frac{A+B}{2}\,\frac{1}{x - \frac{a^2}{x_0}} + A\,\frac{(x_0^2 - a^2)(x - x_0)}{x_0^2(xx_0 - a^2)}\,\frac{a^2}{\left(x - \frac{a^2}{x_0}\right)^2} - \frac{A+B}{2}\,\frac{1}{x} - \frac{B-A}{2x_0}\,\frac{a^2}{x^2} + A\,\frac{a^2}{x^3} - \frac{1}{2a}\sum_{n=1}^{\infty} a_n' n\left(\frac{x}{a}\right)^n + \frac{1}{2a}\sum_{n=1}^{\infty}[a_n'(n-1) + a_{-n}']\left(\frac{x}{a}\right)^{n-2} + \frac{A}{2a}\sum_{n=1}^{\infty} a_{-n}' n\left(\frac{a}{x}\right)^n - \frac{1}{2a}\sum_{n=1}^{\infty}[Aa_{-n}'(n+1) - Ba_n']\left(\frac{a}{x}\right)^{n+2}$$

[17] Postdoctoral Research Associate, Department of Civil Engineering, Case Western Reserve University, Cleveland, OH 44106.

[18] Professor, Department of Mechanics and Materials Science, Rutgers University, Piscataway, NJ 08854, Fellow, ASME.

Manuscript received by the ASME Applied Mechanics Division, May 9, 1989; final revision, May 8, 1990.

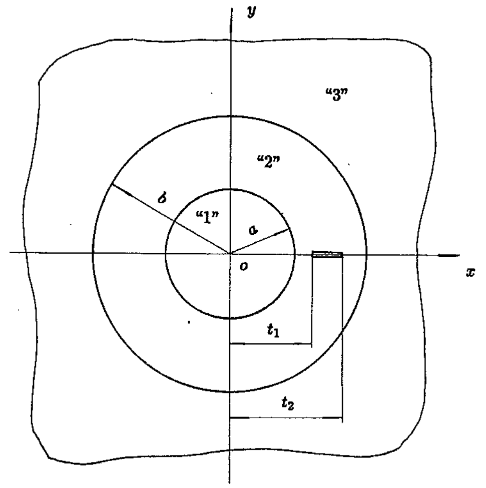


Fig. 1 Three-phase model

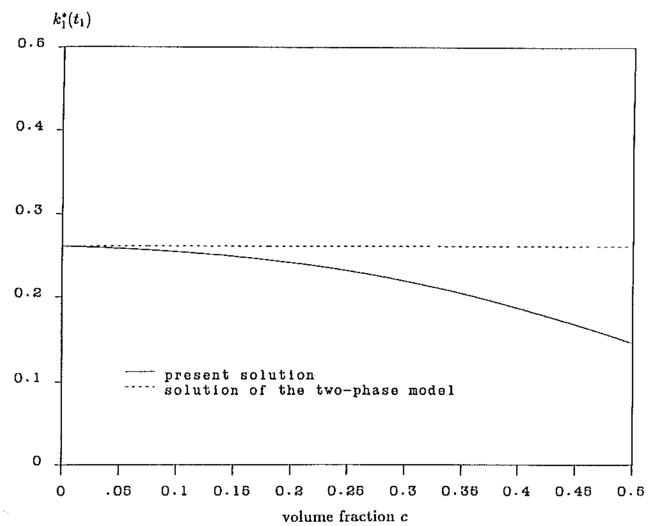


Fig. 2 Effect of volume fraction of fiber on the stress intensity factor

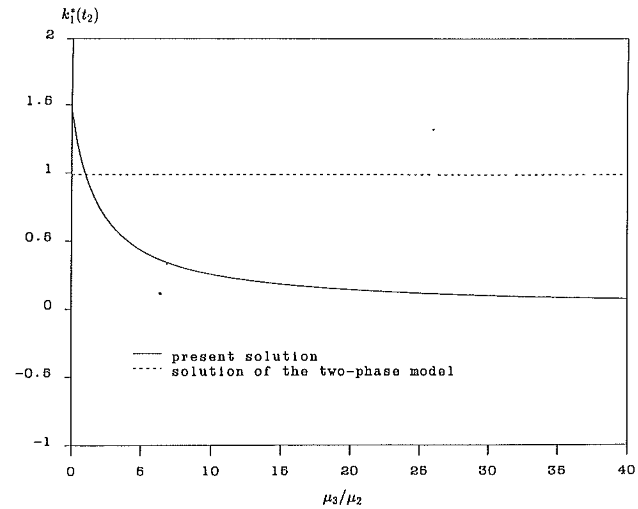


Fig. 3 Effect of μ_3/μ_2 on the stress intensity factor

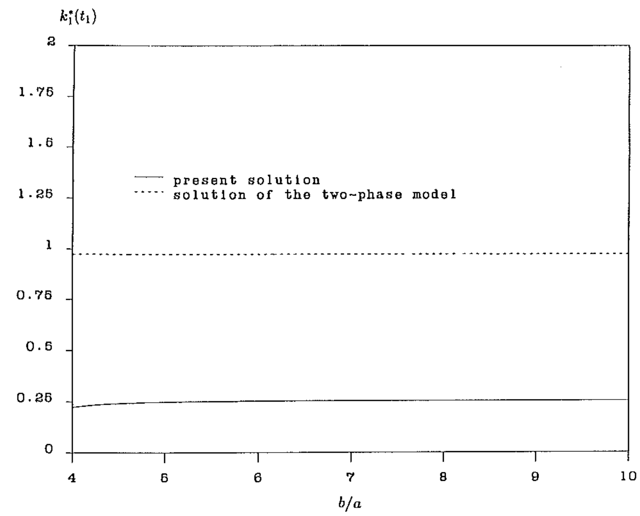


Fig. 4 Effect of b/a on the stress intensity factor

$$
\begin{aligned}
K_y(x, x_0) = {} & \frac{A+B}{2}\,\frac{1}{x-\dfrac{a^2}{x_0}} - \frac{A+B}{2}\,\frac{1}{x} + A\,\frac{x_0^2-a^2}{x_0^3}\left(\frac{x_0^2}{a^2} - \frac{x_0^2-a^2}{xx_0-a^2}\right) \\
& \times \frac{a^2}{\left(x-\dfrac{a^2}{x_0}\right)^2} - A\,\frac{a^2}{x^3} - \frac{1}{2x_0}\left(\frac{a}{x}\right)^2\left[A\left(\frac{2x_0^2}{a^2}-1\right)\right. \\
& \left. + B - 2N\right] - Nb_0\,\frac{a}{x^2} \\
& - \frac{1}{2a}\sum_{n=1}^{\infty} a_n(n+2)\left(\frac{x}{a}\right)^n \\
& + \frac{1}{2a}\sum_{n=1}^{\infty} [a_n(n-1) - a_{-n}]\left(\frac{x}{a}\right)^{n-2} \\
& + \frac{A}{2a}\sum_{n=1}^{\infty} a_{-n}(n-2)\left(\frac{a}{x}\right)^n \\
& - \frac{1}{2a}\sum_{n=1}^{\infty} [Ba_n + A(n+1)a_{-n}]\left(\frac{a}{x}\right)^{n+2} \qquad (4)
\end{aligned}
$$

where the coefficients a_n's and a_n''s are determined by solving a set of linear simultaneous equations given by Luo and Chen (1989), and

$$
A = \frac{1-m}{1+m\kappa_2}, \qquad B = \frac{\kappa_1 - m\kappa_2}{\kappa_1 + m},
$$

$$
N = \frac{\kappa_1 - 1 - m(\kappa_2 - 1)}{\kappa_1 - 1 + 2m}, \quad m = \frac{\mu_1}{\mu_2}. \qquad (5)
$$

Stress intensity factors are readily derived from the stress field due to dislocations. The dimensionless stress intensity factors are defined by $k_i^*(t_j) = k_i(t_j)/(\sigma\sqrt{s})$where $i, j = 1, 2$, σ is the remote uniform stress and $s = (t_2 - t_1)/2$.

The singular integral equations are solved numerically (Erdogan and Gupta, 1972). The fiber-reinforced composite with $\mu_1/\mu_2 = 23$, $\nu_1 = 0.3$, $\nu_2 = 0.35$ and $t_1 = 1.05a$, $t_2 = 1.35a$ is considered. The effective moduli of the composite (the elastic moduli of the outer phase) are evaluated based on the modified Mori-Tanaka method suggested by Luo and Weng (1989). Figure 2 shows the dimensionless stress intensity factors $k_1^*(t_1)$ versus volume fraction of fiber for the above composites under uniaxial tension along the direction perpendicular to the crack face. For comparison, the results calculated based on the two-phase model (Erdogan et al. 1974) are also depicted. It is seen that the volume fraction of fiber has considerable influence on the stress intensity factor of the crack.

It is of interest to study the respective influence of the elastic property and the geometric parameters on the behavior of the matrix crack. Figure 3 shows the variation of $k_1^*(t_1)$ versus μ_3/μ_2 for a solid with $\mu_1=20\mu_2$, $\nu_1=\nu_2=\nu_3=0.3$, $a/b=1/8$, $t_1=3.5a$ and $t_2=4.5a$ under uniaxial tension. Figure 4 shows the variation of $k_1^*(t_1)$ versus b/a for a solid with $\mu_1=20\mu_2$, $\mu_3=10\mu_2$, $\nu_1=\nu_2=\nu_3=0.3$, $t_1=3.5a$ and $t_2=4.5a$ under uniaxial tension. As μ_3/μ_2 increases, the solution of the three-phase model has an asymptotical value which corresponds to the case where the outer-phase is rigid. In Fig. 4 it is observed that the results for the three-phase model approach those of the two-phase model remarkably slowly as the thickness of the intermediate matrix-phase increases. This is because the thickness of the outer phase, which in this case is harder than intermediate phase, is always infinitely large.

References

Christensen, R. M., and Lo, K. H., 1979, "Solutions for Effective Shear Properties in Three Phases Sphere and Cylinder Models," *Journal of Mechanics and Physics of Solids*, Vol. 27, pp. 315–330.

Erdogan, F., and Gupta, G. D., 1972, "On the Numerical Solution of Singular Integral Equations," *Quarterly of Applied Mathematics*, Vol. 29, pp. 525–539.

Erdogan, F., Gupta, G. D., and Ratwani, M., 1974, "Interaction between a Circular Inclusion and an Arbitrarily Oriented Crack," ASME Journal of Applied Mechanics, Vol. 41, pp. 1007–1013.

Luo, H. A., and Chen, Y., 1991, "An Edge Dislocation in a Three-Phase Composite Cylinder Model," ASME Journal of Applied Mechanics, Vol. 51, pp. 75–86.

Luo, H. A., and Weng, G. J., 1989, "On Eshelby's S-Tensor in a Three-Phase Cylindrically Concentric Solid, and the Elastic Moduli of Fiber-Reinforced Composites," *Mechanics of Materials*, Vol. 8, pp. 77–88.

Buckling and Initial Post-Buckling Behavior of Laminated Oval Cylindrical Shells Under Axial Compression

G. Sun[19]

Introduction

The influence of the laminate configuration on the performance of various composite plate, panel, and shell structures has received extensive attention in the literature. Of the papers contributed to the stability of anisotropic laminated shells, the majority pertains to the stability and optimum design problem of laminated composite circular cylinders (e.g., Tennyson et al., 1971; Booton, 1976; Onoda, 1985; Zimmerman, 1986). Recent studies (Tennyson and Hansen, 1983; Sun and Hansen, 1988) have shown that both the buckling load and the degree of imperfection sensitivity of composite circular cylinders are significantly affected by the wall laminate constructions. However, the literature dealing with the stability problem of noncircular cylinders is meager in comparison. Early investigation on the buckling and post-buckling of isotropic oval cylinders was performed by Kempner and Chen (1976). Using the general theory of elastic stability which was first introduced by Koiter (1945) and later written in an equivalent version in terms of the principle of virtual work by Budiansky and Hutchinson (1964), Hutchinson (1968) has studied the initial post-buckling behavior of isotropic oval cylinders, indicating that oval cylinders fabricated from isotropic materials are highly sensitive to small geometric imperfections. The buckling and vibration of laminated composite, noncircular cylindrical shells has been analyzed by Soldatos and Tzivanidis (1982), Soldatos (1984), and Hui and Du (1986), among others.

The present study extends the development of Hutchinson (1968) to include oval cylinders made of anisotropic composite laminates without imposing any restriction on the lamination scheme. The shell analysis is based on Donnell's shallow shell theory and the post-buckling b-coefficient is employed to indicate the sensitivity of laminated oval cylinders to asymmetric geometric imperfections. The intention of this work, however, is to examine the effects of the laminate configuration as well as the eccentricity of the oval cross-section on the buckling and initial post-buckling behavior of laminated composite oval cylinders.

Analysis

Let the midsurface of the cylindrical shell be the reference surface. The coordinates $\bar{x}$ and $\bar{y}$ are measured in the axial and circumferential directions, respectively. As shown in Fig. 1, noncircular cylindrical shells with two types of cross-sections are considered in this study. They are: (1) elliptical cross-section and (2) nonellipsoidal oval cross-section. The points on the middle surface of a cylinder with an elliptical cross-section satisfy

$$\left(\frac{\xi}{A}\right)^1 + \left(\frac{\eta}{B}\right)^2 = 1 \tag{1}$$

where A and B are semi-lengths of the major and minor axes of the ellipse. The circumferential radius of curvature at any point on the ellipse is given by

$$\bar{R}(\bar{y}) = \frac{A^2}{B}\left[1 - \left(1 - \frac{B^2}{A^2}\right)\frac{\xi^2(\bar{y})}{A^2}\right]^{3/2}. \tag{2}$$

The cylinder with a nonellipsoidal oval cross-section, which was considered by Kempner and Chen (1967) and is referred to as an "oval" cylinder in the following context, is characterized by the circumferential radius of curvature

$$\bar{R}(\bar{y}) = \frac{R_o}{1 - e\cos(2\bar{y}/R_o)} \tag{3}$$

where e is the eccentricity parameter. At one limit ($e = 0$) the oval reduces to a circle, while at the other ($e = 1$) corresponds

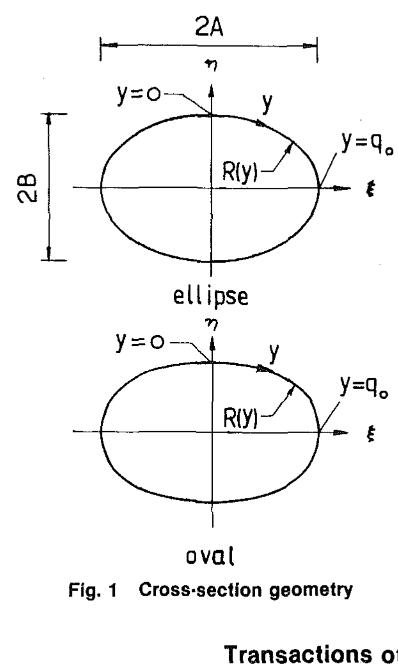


Fig. 1 Cross-section geometry

[19]Department of Engineering Mechanics, Shanghai Jiao Tong University, Shanghai 200030, China.

Manuscript received by the ASME Applied Mechanics Division, Oct. 4, 1989; final revision, June 18, 1990.

that the volume fraction of fiber has considerable influence on the stress intensity factor of the crack.

It is of interest to study the respective influence of the elastic property and the geometric parameters on the behavior of the matrix crack. Figure 3 shows the variation of $k_1^*(t_1)$ versus μ_3/μ_2 for a solid with $\mu_1=20\mu_2$, $\nu_1=\nu_2=\nu_3=0.3$, $a/b=1/8$, $t_1=3.5a$ and $t_2=4.5a$ under uniaxial tension. Figure 4 shows the variation of $k_1^*(t_1)$ versus b/a for a solid with $\mu_1=20\mu_2$, $\mu_3=10\mu_2$, $\nu_1=\nu_2=\nu_3=0.3$, $t_1=3.5a$ and $t_2=4.5a$ under uniaxial tension. As μ_3/μ_2 increases, the solution of the three-phase model has an asymptotical value which corresponds to the case where the outer-phase is rigid. In Fig. 4 it is observed that the results for the three-phase model approach those of the two-phase model remarkably slowly as the thickness of the intermediate matrix-phase increases. This is because the thickness of the outer phase, which in this case is harder than intermediate phase, is always infinitely large.

References

Christensen, R. M., and Lo, K. H., 1979, "Solutions for Effective Shear Properties in Three Phases Sphere and Cylinder Models," *Journal of Mechanics and Physics of Solids*, Vol. 27, pp. 315–330.

Erdogan, F., and Gupta, G. D., 1972, "On the Numerical Solution of Singular Integral Equations," *Quarterly of Applied Mathematics*, Vol. 29, pp. 525–539.

Erdogan, F., Gupta, G. D., and Ratwani, M., 1974, "Interaction between a Circular Inclusion and an Arbitrarily Oriented Crack," ASME Journal of Applied Mechanics, Vol. 41, pp. 1007–1013.

Luo, H. A., and Chen, Y., 1991, "An Edge Dislocation in a Three-Phase Composite Cylinder Model," ASME Journal of Applied Mechanics, Vol. 51, pp. 75–86.

Luo, H. A., and Weng, G. J., 1989, "On Eshelby's S-Tensor in a Three-Phase Cylindrically Concentric Solid, and the Elastic Moduli of Fiber-Reinforced Composites," *Mechanics of Materials*, Vol. 8, pp. 77–88.

Buckling and Initial Post-Buckling Behavior of Laminated Oval Cylindrical Shells Under Axial Compression

G. Sun[19]

Introduction

The influence of the laminate configuration on the performance of various composite plate, panel, and shell structures has received extensive attention in the literature. Of the papers contributed to the stability of anisotropic laminated shells, the majority pertains to the stability and optimum design problem of laminated composite circular cylinders (e.g., Tennyson et al., 1971; Booton, 1976; Onoda, 1985; Zimmerman, 1986). Recent studies (Tennyson and Hansen, 1983; Sun and Hansen, 1988) have shown that both the buckling load and the degree of imperfection sensitivity of composite circular cylinders are significantly affected by the wall laminate constructions. However, the literature dealing with the stability problem of noncircular cylinders is meager in comparison. Early investigation on the buckling and post-buckling of isotropic oval cylinders was performed by Kempner and Chen (1976). Using the general theory of elastic stability which was first introduced by Koiter (1945) and later written in an equivalent version in terms of the principle of virtual work by Budiansky and Hutchinson (1964), Hutchinson (1968) has studied the initial post-buckling behavior of isotropic oval cylinders, indicating that oval cylinders fabricated from isotropic materials are highly sensitive to small geometric imperfections. The buckling and vibration of laminated composite, noncircular cylindrical shells has been analyzed by Soldatos and Tzivanidis (1982), Soldatos (1984), and Hui and Du (1986), among others.

The present study extends the development of Hutchinson (1968) to include oval cylinders made of anisotropic composite laminates without imposing any restriction on the lamination scheme. The shell analysis is based on Donnell's shallow shell theory and the post-buckling b-coefficient is employed to indicate the sensitivity of laminated oval cylinders to asymmetric geometric imperfections. The intention of this work, however, is to examine the effects of the laminate configuration as well as the eccentricity of the oval cross-section on the buckling and initial post-buckling behavior of laminated composite oval cylinders.

Analysis

Let the midsurface of the cylindrical shell be the reference surface. The coordinates $\bar{x}$ and $\bar{y}$ are measured in the axial and circumferential directions, respectively. As shown in Fig. 1, noncircular cylindrical shells with two types of cross-sections are considered in this study. They are: (1) elliptical cross-section and (2) nonellipsoidal oval cross-section. The points on the middle surface of a cylinder with an elliptical cross-section satisfy

$$\left(\frac{\xi}{A}\right)^1 + \left(\frac{\eta}{B}\right)^2 = 1 \tag{1}$$

where A and B are semi-lengths of the major and minor axes of the ellipse. The circumferential radius of curvature at any point on the ellipse is given by

$$\bar{R}(\bar{y}) = \frac{A^2}{B}\left[1 - \left(1 - \frac{B^2}{A^2}\right)\frac{\xi^2(\bar{y})}{A^2}\right]^{3/2}. \tag{2}$$

The cylinder with a nonellipsoidal oval cross-section, which was considered by Kempner and Chen (1967) and is referred to as an "oval" cylinder in the following context, is characterized by the circumferential radius of curvature

$$\bar{R}(\bar{y}) = \frac{R_o}{1 - e\cos(2\bar{y}/R_o)} \tag{3}$$

where e is the eccentricity parameter. At one limit ($e = 0$) the oval reduces to a circle, while at the other ($e = 1$) corresponds

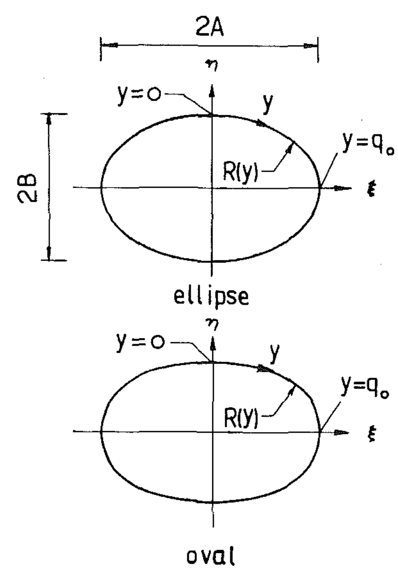


Fig. 1 Cross-section geometry

[19]Department of Engineering Mechanics, Shanghai Jiao Tong University, Shanghai 200030, China.

Manuscript received by the ASME Applied Mechanics Division, Oct. 4, 1989; final revision, June 18, 1990.

to an oval with $B/A \approx 0.485$. A reference radius R_o is defined to be the radius of the circle with exactly the same perimeter length as the ellipse or the oval. Nondimensional parameters are introduced as follows:

$$\begin{aligned}
&(x,y) = (\bar{x},\bar{y})/\sqrt{R_o t}\\
&W = \overline{W}/t\\
&R(y) = \overline{R}(\bar{y})/R_o\\
&(A_{ij}^*, B_{ij}^*, D_{ij}^*) = (E_{11}t\overline{A}_{ij}^*, \overline{B}_{ij}^*/t, \overline{D}_{ij}^*/(E_{11}t^3))\\
&F = \overline{F}/(E_{11}t^3) \qquad (i,j = 1, 2, 6)\\
&N_x = \overline{N}_x R_o/(E_{11}t^2). \qquad (4)
\end{aligned}$$

In the above it is noted that a symbol with an overbar is dimensional, while the same symbol without an overbar is nondimensional. Further, A_{ij}^*, B_{ij}^*, and D_{ij}^* are laminate coefficients as defined in Ashton et al. (1969), t is the wall thickness, W is the normal deflection of the shell middle surface, F is the Airy stress function, E_{11} is the elastic modulus of a lamina in the fiber direction, and N_x is the axial stress resultant.

The buckling analysis is based on Donnell-type compatibility and equilibrium equations for "generally" laminated noncircular cylindrical shells written in terms of the normal deflection W and the stress function F. They are (Tennyson et al., 1971)

$$\left[\begin{array}{c|c} \nabla_{A^*}^4 & -\nabla_{B^*}^4 \\ \hline \nabla_{B^*}^4 & \nabla_{D^*}^4 \end{array}\right]\left[\begin{array}{c} F \\ \hline W \end{array}\right] = \left[\begin{array}{c} (W_{,xy})^2 - W_{,yy}W_{,xx} + \dfrac{1}{R(y)}W_{,xx} \\ \hline F_{,yy}W_{,xx} - 2F_{,xy}W_{,xy} + F_{,xx}W_{,yy} - \dfrac{1}{R(y)}F_{,xx} \end{array}\right] \qquad (5)$$

in which the operators are defined as

$$\nabla_{A^*}^4 = A_{22}^*\frac{\partial^4}{\partial x^4} - 2A_{26}^*\frac{\partial^4}{\partial x^3\partial y} + (2A_{12}^* + A_{66}^*)\frac{\partial^4}{\partial x^2\partial y^2} - 2A_{16}^*\frac{\partial^4}{\partial x\partial y^3} + A_{11}^*\frac{\partial^4}{\partial y^4}$$

$$\nabla_{B^*}^4 = B_{21}^*\frac{\partial^4}{\partial x^4} + (2B_{26}^* - B_{61}^*)\frac{\partial^4}{\partial x^3\partial y} + (B_{11}^* + B_{22}^* - 2B_{66}^*)\frac{\partial^4}{\partial x^2\partial y^2} + (2B_{16}^* - B_{62}^*)\frac{\partial^4}{\partial x\partial y^3} + B_{12}^*\frac{\partial^4}{\partial y^4}$$

$$\nabla_{D^*}^4 = D_{11}^*\frac{\partial^4}{\partial x^4} + 4D_{16}^*\frac{\partial^4}{\partial x^3\partial y} + 2(D_{12}^* + 2D_{66}^*)\frac{\partial^4}{\partial x^2\partial y^2} + 4D_{26}^*\frac{\partial^4}{\partial x\partial y^3} + D_{22}^*\frac{\partial^4}{\partial y^4}.$$

Following the method outlined in Hutchinson (1968), an asymptotic perturbation expansion of the solution valid in the neighborhood of the bifurcation point is assumed to be of the form

$$\left[\begin{array}{c} W \\ \hline F \end{array}\right] = \left[\begin{array}{c} W_o \\ \hline -\frac{1}{2}N_x^o y \end{array}\right] + \epsilon\left[\begin{array}{c} W_1 \\ \hline F_1 \end{array}\right] + \epsilon^2\left[\begin{array}{c} W_2 \\ \hline F_2 \end{array}\right] + \ldots \qquad (6)$$

where N_x^o is the pre-buckling axial stress resultant (positive for compressive stress), W_1, F_1 are the buckling field, W_2, F_2 are the second-order field, and ϵ is the normalized amplitude of the buckling mode W_1 and serves as the perturbation parameter. As a first approximation to the problem the analysis neglects the effects of boundary constraints at both ends of the shell and the prebuckling normal deflection W_o and thus, in effect, treats the case of an infinitely long cylinder. Substituting (6) into the governing equations (5) and setting the coefficients of ϵ to zero yields the following eigenvalue problem for the buckling stress N_{xc}^o and W_1 and F_1:

$$\left[\begin{array}{c|c} \nabla_{A^*}^4 & -\nabla_{B^*}^4 \\ \hline \nabla_{B^*}^4 & \nabla_{D^*}^4 \end{array}\right]\left[\begin{array}{c} F_1 \\ \hline W_1 \end{array}\right] = \left[\begin{array}{c} \dfrac{1}{R(y)}W_{1,xx} \\ \hline N_x^o W_{1,xx} - \dfrac{1}{R(y)}F_{1,xx} \end{array}\right]. \qquad (7)$$

Equating the coefficients of ϵ^2 to zero yields the equations for the second-order field

$$\left[\begin{array}{c|c} \nabla_{A^*}^4 & -\nabla_{B^*}^4 \\ \hline \nabla_{B^*}^4 & \nabla_{D^*}^4 \end{array}\right]\left[\begin{array}{c} F_2 \\ \hline W_2 \end{array}\right] = \left[\begin{array}{c} \dfrac{1}{R(y)}W_{2,xx} + W_{1,xy}^2 - W_{1,xx}W_{1,yy} \\ \hline N_{xc}^o W_{2,xx} - \dfrac{1}{R(y)}F_{2,xx} + F_{1,xx}W_{1,yy} \\ + F_{1,yy}W_{1,xx} - 2F_{1,xy}W_{1,xy} \end{array}\right]. \qquad (8)$$

A solution of equations (7) is possible by means of separation of variables whereby,

$$\begin{aligned} W_1(x,y) &= w(y)e^{iMx} \\ F_1(x,y) &= f(y)e^{iMx} \end{aligned} \qquad (9)$$

where $w(y)$ and $f(y)$ are complex functions of y. Equivalently, the solution can also be assumed in a separable form of real functions

$$\begin{aligned} W_1(x,y) &= w_1(y)\cos(Mx) + w_2(y)\sin(Mx) \\ F_1(x,y) &= f_1(x)\cos(Mx) + f_2(y)\sin(Mx). \end{aligned} \qquad (10)$$

Substituting (10) into equations (7) leads to four coupled fourth-order ordinary differential equations for $w_1(y)$, $w_2(y)$, $f_1(y)$, and $f_2(y)$ (see Sun, 1986). They are discretized using the finite difference technique. The normalized axial wave number M can be treated as a continuous variable since an infinitely long cylinder is considered. This eigenvalue problem can be solved for a number of values of M to find for each value the lowest eigenvalue N_{xc}^o and the associated eigenfunction w_1, w_2, f_1, and f_2. It is convenient to restrict the analysis to a one-quarter segment of the circumference running from $y = 0$ to $y = q_o = \sqrt{R_o/t}\,\pi/2$ (Fig. 1). The circumferential boundary conditions are determined in consideration of continuity and symmetry with respect to the major and minor axes. It can be concluded from the four ordinary differential equations that in case $w_1(y)$ and $f_1(y)$ are symmetric with respect to one of the axes, $w_2(y)$ and $f_2(y)$ are antisymmetric with respect to the same axis and vice versa. Thus, the buckling mode can be identified as one of the following four groups according to the symmetry condition of functions $w_1(y)$ and $f_1(y)$ only: SS (symmetric at both axes), SA (symmetric at the minor axis and antisymmetric at the major axis), AS (antisymmetric at the minor axis and symmetric at the major axis) and AA (antisymmetric at both axes). These abbreviations for circumferential boundary conditions have been used by Hui and Du (1986). The boundary conditions at either the minor axis ($y = 0$) or the major axis ($y = q_o$) for the "S" mode are

$$\begin{aligned} &w_{1,y} = 0, \quad w_{1,yyy} = 0, \quad f_{1,y} = 0, \quad f_{1,yyy} = 0 \\ &w_2 = 0, \quad w_{2,yy} = 0, \quad f_2 = 0, \quad f_{2,yy} = 0. \end{aligned} \qquad (11)$$

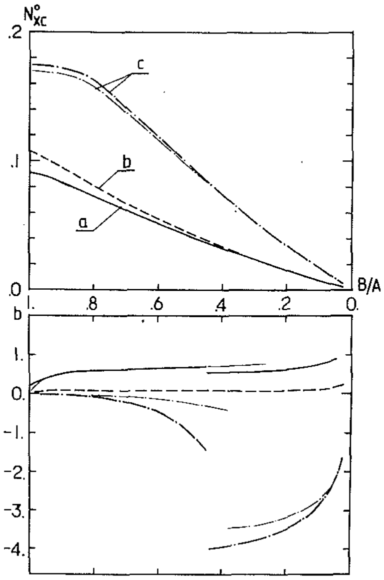


Fig. 2 Buckling load and b-coefficient for elliptical cylinders with (a) (90,0,0,90), (b) (90,45, −45,90), and (c) (45,0, −45,90,90,45,0, −45) lamination

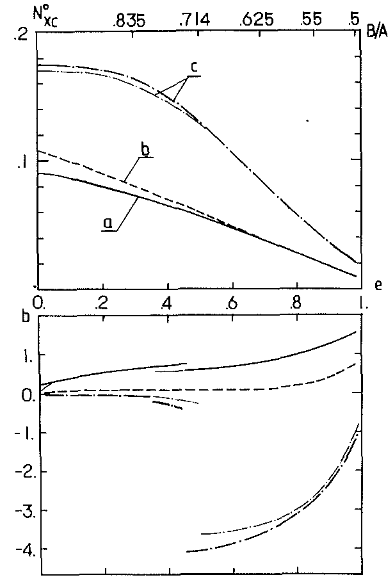


Fig. 3 Buckling load and b-coefficient for oval cylinders with (a) (90,0,0,90), (b) (90,45, −45,90), and (c) (45,0, −45,90,90,45,0, −45) lamination

The boundary conditions for the "A" mode are

$$\begin{aligned} &w_1=0, \quad w_{1,yy}=0, \quad f_1=0 \quad f_{1,yy}=0 \\ &w_{2,y}=0, \quad w_{2,yyy}=0, \quad f_{2,y}=0, \quad f_{2,yyy}=0. \end{aligned} \tag{12}$$

The nonhomogeneous terms on the right-hand sides of the second-order field equations (8) are quadratic in W_1 and F_1. Thus, solutions to these equations are sought in the separable form

$$\begin{aligned} W_2(x,y) &= w_3(y) + w_4(y)\cos(2Mx) + w_5(y)\sin(2Mx) \\ F_2(x,y) &= f_3(y) + f_4(y)\cos(2Mx) + f_5(y)\sin(2Mx) \end{aligned} \tag{13}$$

where M is predetermined in the buckling analysis. Substituting equations (13) into (8) leads to two coupled fourth-order ordinary differential equations for $w_3(y)$ and $f_3(y)$, and four coupled fourth-order ordinary differential equations for $w_4(y)$, $w_5(y)$, $f_4(y)$, and $f_5(y)$. The nonhomogeneous terms to all these ordinary differential equations are quadratic in w_1, w_2, f_1, and f_2, and are symmetric with respect to the major and minor axes. Accordingly, the boundary conditions for w_3, f_3, w_4, and f_4 are symmetric about both axes while the boundary conditions for w_5 and f_5 are antisymmetric about both axes for all four buckling modes, that is (Sun, 1986),

$$\begin{aligned} &w_{3,y}=0, \quad w_{3,yyy}=0, \quad f_{3,y}=0, \quad f_{3,yyy}=0 \\ &w_{4,y}=0, \quad w_{4,yyy}=0, \quad f_{4,y}=0 \quad f_{4,yyy}=0 \\ &w_5=0, \quad w_{5,yy}=0, \quad f_5=0, \quad f_{5,yy}=0. \end{aligned} \tag{14}$$

Solutions to problems (7) and (8) are used to express the equilibrium relation of applied axial load N_x to normalized buckling deflection ϵ in the vicinity of the buckling load N^o_{xc} as

$$\frac{N_x}{N^o_{xc}} = 1 + b\epsilon^2 + \ldots \tag{15}$$

where the postbuckling coefficient "b" is in the form

$$b = \left\{ 2\int_S [F_{1,xx}W_{1,y}W_{2,y} + F_{1,yy}W_{2,x}W_{1,x} - F_{1,xy}(W_{1,x}W_{2,y} + W_{1,y}W_{2,x})]dS + \int_S [F_{2,xx}W^2_{1,y} + F_{2,yy}W^2_{1,x} - 2F_{2,xy}W_{1,x}W_{1,y}]dS \right\} \div \left\{ N^o_{xc}\int_S W^2_{1,x}dS \right\}. \tag{16}$$

A general development of this theory and the significance of the post-buckling coefficient "b" can be found in Hutchinson (1968).

Results and Discussion

One of the important findings reported by Tennyson and Hansen (1983) is that some laminated graphite/epoxy circular cylinders can withstand compressive load more than twice as high as some other circular cylinders of the same size and material, yet different wall laminate configuration. Their post-buckling character also differs substantially from each other. Naturally, it is of interest to examine how laminated cylindrical shells will behave when eccentricity effect is taken into account. The shells considered are assumed to be made of graphite/epoxy laminates with lamina elastic properties $E_{11} = 14.1 \times 10^{10}$N/m^2, $E_{22} = 0.97 \times 10^{10}$N/m^2, $G_{12} = 0.41 \times 10^{10}$N/m^2, and $\nu_{12} = 0.26$, and geometric parameter $R_o/t = 165$. As a check on the analysis and numerical procedure developed, the buckling load and b-coefficient obtained using isotropic material properties coincide well with that presented by Hutchinson (1968).

The effects of the eccentricity parameter (B/A) on the buckling load N^o_{xc} and the value of the post-buckling coefficient b are shown in Fig. 2 for laminated elliptical cylinders with three laminate configurations: (a) (90,0,0,90), (b) (90,45, −45,90), and (c) (45,0, −45,90,90,45,0, −45). (Numbers in parenthesis are orientations of each ply from the inner wall of the shell out; the degree sign notation has been omitted.) In Fig. 3, the results of N^o_{xc} and b versus the eccentricity parameters e are plotted for laminated oval cylinders with these three laminations. It was found that in every case studied, the results for both N^o_{xc} and b of the SS and SA modes, which are plotted in thick lines in Figs. 2 and 3, are indistinguishable, while the results of the AS and AA modes are indistinguishable too. They are plotted using thin lines of the same line style in Figs. 2 and 3 if their divergence from the results of the SS and SA modes is distinguishable. It should be noted that for a given group of circumferential boundary conditions, the buckling mode shape also varies in axial wave length with the increase of eccentricity. It was found that the discontinuities of b-coefficient curves for (45,0, −45,90,90,45,0, −45) cylinders in Figs. 2 and 3 are caused by a jump in axial wave length of the buckling mode. The discontinuities of postbuckling curves for (90,0,0,90) cylinders, however, are caused by a shift of a buckling mode with $w_1 \neq 0$, $w_2 = 0$ to a buckling mode with $w_1 = 0$, $w_2 \neq 0$, or vice versa. All the N_{xc}^o curves appear to be continuous for the full range of eccentricity because the buckling loads are not susceptible to these changes. Obviously, the buckling loads for both elliptical and oval cylinders with (45,0, −45,90,90,45,0, −45) lamination are about twice as high as the other two laminations for the full range of eccentricity. The b-coefficients for cylinders with (90,0,0,90) and (90,45, −45,90) laminations are positive, showing their stable postbuckling behavior. The "b" curves for (45,0, −45,90,90,45,0, −45) cylinders indicate that they are slightly postbuckling unstable for the range of small eccentricity and highly post-buckling unstable for sufficiently large eccentricity. It can also be seen from Figs. 2 and 3 that in the higher range of "e," the buckling load of an oval cylinder is much lower than that of an elliptical cylinder with the same minor-major axis ratio (B/A).

From the above examples we may conclude that the buckling load and the initial post-buckling behavior of laminated composite noncircular cylinders, like laminated circular cylinders, are significantly affected by the wall laminate configuration. An optimization procedure for the selection of wall laminate structure is crucial in the design of laminated composite noncircular cylinders for buckling.

Acknowledgment

This work was started at the Institute for Aerospace Studies, University of Toronto, under the supervision of Profs. J. S. Hansen and R. C. Tennyson to whom the author wishes to express appreciation.

References

Ashton, J. E., Halpin, J. C., and Petit, P. H., 1969, *Primer on Composite Materials: Analysis*, Technomic, Westport, Conn.

Booton, M., 1976, "Buckling of Imperfect Anisotropic Cylinders Under Combined Loadings," UTIAS Report No. 203, University of Toronto, Institute for Aerospace Studies, Downsview, Ontario, Canada.

Budiansky, B., and Hutchinson, J. W., 1964, "Dynamic Buckling of Imperfection Sensitive Structures," *Proc. XI Internat. Congr. Appl. Mech.*, Munich, Springer-Verlag, Berlin, pp. 636–651.

Hui, D., and Du, I. H. Y., 1986, "Effects of Axial Imperfections on Vibrations of Anti-Symmetric Cross-Ply, Oval Cylindrical Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 675–680.

Hutchinson, J. W., 1968, "Buckling and Initial Postbuckling Behavior of Oval Cylindrical Shells Under Axial Compression," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 66–72.

Kempner, J., and Chen, Y. N., 1967, "Buckling and Postbuckling of an Axially Compressed Oval Cylindrical Shell," *Proc. Symp. on the Theory of Shells to Honor Lloyd H. Donnell*, McCutchan Pub. Co., pp. 141–183.

Koiter, W. T., 1945, "On the Stability of Elastic Equilibrium," Thesis, Delft, Amsterdam (English Translation: AFFDL-TR-70-25), 1970.

Onoda, J., 1985, "Optimal Laminate Configurations of Cylindrical Shells for Axial Buckling," *AIAA J.*, Vol. 23, pp. 1093–1098.

Soldatos, K. P., 1984, "A Flugge-type Theory for the Analysis of Anisotropic Laminated Non-Circular Cylindrical Shells," *Int. J. of Solids and Structures*, Vol. 20, No. 2, pp. 107–120.

Soldatos, K. P., and Tzivanidis, G. J., 1982, "Buckling and Vibration of Cross-ply Laminated Non-Circular Cylindrical Shells," *J. of Sound and Vibration*, Vol. 82, No. 32, pp. 425–434.

Sun, G., 1986, "Optimization of Laminated Cylinders for Buckling," Doctoral Thesis, Institute for Aerospace Studies, University of Toronto, Toronto, Ontario, Canada.

Sun, G., and Hansen, J. S., 1988, "Optimal Design of Laminated-Composite Circular-Cylindrical Shells Subjected to Combined Loads," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 136–142.

Tennyson, R. C., Chan, K. H., and Muggeridge, D. B., 1971, "The Effect of Axisymmetric Shape Imperfections on the Buckling of Laminated Anisotropic Circular Cylinders," *CASI Trans.*, Vol. 4, No. 2, pp. 131–139.

Tennyson, R. C., and Hansen, J. S., 1983, "Optimum Design for Buckling of Laminated Cylinders," *Proc. IUTAM Symposium: Collapse: The Buckling of Structures in Theory and Practice*, J. M. T. Thompson, and G. W. Hunt, eds., Cambridge Univ. Press, Cambridge.

Zimmerman, R., 1986, "Optimization of Axially Compressed CFRP Cylinders," *Spacecraft Structures*. CNES, Toulouse, Dec. 3–6, 1985; ESA SP-238, Apr. 1986, pp. 407–412.

Stress Analysis of Smooth Polygonal Holes via a Boundary Perturbation Method

H. Gao[20]

Boundary Perturbation Method

This note presents an analysis of smooth polygon holes in an elastic sheet via a boundary perturbation method that can be formulated by the following steps. Let a sheet containing a circular hole be subjected to a fixed loading system and an additional point force P_i at a position (x, y), with $P_i\delta u_i(x, y)$ equal to the work done by the force at an incremental displacement $\delta u_i(x, y)$. Imagine that the hole boundary, initially at the circle $r = a$, is perturbed to a neighboring position by some variable normal distance $\delta a(\theta')$ (Fig. 1). Treating $\delta a(\theta')$ as infinitesimal, the change in the total energy Γ is

$$\delta\Gamma = P_i\delta u_i - \int_0^{2\pi} w\delta a(\theta')ad\theta' \tag{1}$$

where w is the strain energy density along the unperturbed circular hole boundary. For plane stress conditions, w is related to the hoop stress $\sigma_{\theta\theta}$ by

$$w = \sigma_{\theta\theta}^2/2E. \tag{2}$$

The energy variation due to the shape change of a void or hole was first studied by Rice and Drucker (1967). Let equation (1) be rearranged into the following form known as the Legendre transformation:

$$\delta(P_iu_i - \Gamma) = u_i\delta P_i + \left(\int_0^{2\pi} wg(\theta')ad\theta'\right)\delta A \tag{3}$$

where $\delta a(\theta')$ has been written as $g(\theta')\delta A$. The right side of equation (3) being a perfect differential, the coefficients of δP_i and δA must satisfy the Maxwell reciprocal relation

$$\frac{\partial u_i}{\partial A} = \frac{\partial}{\partial P_i}\left(\int_0^{2\pi} wg(\theta')ad\theta'\right)$$
$$= \frac{1}{E}\int_0^{2\pi} \sigma_{\theta\theta}(\theta')\Sigma_{\theta\theta}^i(\theta';x, y)g(\theta')ad\theta'. \tag{4}$$

Here we have identified $\partial\sigma_{\theta\theta}/\partial P_i$ as the stress Green's function $\Sigma_{\theta\theta}^i(\theta';x, y)$, i.e., the stress at a boundary point θ' due to a unit point force in i direction at x, y. Multiplying both sides of (4) by δA and letting $P_i = 0$ yield

$$\delta u_i(x,y) = \frac{1}{E}\int_0^{2\pi} \sigma_{\theta\theta}^0(\theta')\Sigma_{\theta\theta}^i(\theta';x, y)\delta a(\theta')ad\theta' \tag{5}$$

where $\sigma_{\theta\theta}^0$ denotes the original stress field in absence of the point force P_i.

Equation (5) provides a first-order boundary perturbation formula for calculation of the displacement field of a nearly circular hole. In the perturbation analysis, a hole with a complex shape $r = a(\theta')$ is viewed as being perturbed from a reference circular hole by a small perturbation $\delta a(\theta')$. The displacement field for the actual hole is written as

$$u_i(x, y) = u_i^0(x, y) + \delta u_i(x, y) \tag{6}$$

where $u_i^0(x, y)$ is the reference solution (of a circular hole) and $\delta u_i(x, y)$ is given by (5). This perturbation procedure parallels an earlier development of a crack perturbation theory by Rice (1985).

[20]Professor, Division of Applied Mechanics, Stanford University, Stanford, CA 94305. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, Oct. 27, 1989; final revision, June 22, 1990.

ling mode with $w_1 \neq 0$, $w_2 = 0$ to a buckling mode with $w_1 = 0$, $w_2 \neq 0$, or vice versa. All the N^o_{xc} curves appear to be continuous for the full range of eccentricity because the buckling loads are not susceptible to these changes. Obviously, the buckling loads for both elliptical and oval cylinders with (45,0, −45,90,90,45,0, −45) lamination are about twice as high as the other two laminations for the full range of eccentricity. The b-coefficients for cylinders with (90,0,0,90) and (90,45, −45,90) laminations are positive, showing their stable postbuckling behavior. The "b" curves for (45,0, −45,90,90,45,0, −45) cylinders indicate that they are slightly postbuckling unstable for the range of small eccentricity and highly post-buckling unstable for sufficiently large eccentricity. It can also be seen from Figs. 2 and 3 that in the higher range of "e," the buckling load of an oval cylinder is much lower than that of an elliptical cylinder with the same minor-major axis ratio (B/A).

From the above examples we may conclude that the buckling load and the initial post-buckling behavior of laminated composite noncircular cylinders, like laminated circular cylinders, are significantly affected by the wall laminate configuration. An optimization procedure for the selection of wall laminate structure is crucial in the design of laminated composite noncircular cylinders for buckling.

Acknowledgment

This work was started at the Institute for Aerospace Studies, University of Toronto, under the supervision of Profs. J. S. Hansen and R. C. Tennyson to whom the author wishes to express appreciation.

References

Ashton, J. E., Halpin, J. C., and Petit, P. H., 1969, *Primer on Composite Materials: Analysis*, Technomic, Westport, Conn.

Booton, M., 1976, "Buckling of Imperfect Anisotropic Cylinders Under Combined Loadings," UTIAS Report No. 203, University of Toronto, Institute for Aerospace Studies, Downsview, Ontario, Canada.

Budiansky, B., and Hutchinson, J. W., 1964, "Dynamic Buckling of Imperfection Sensitive Structures," *Proc. XI Internat. Congr. Appl. Mech.*, Munich, Springer-Verlag, Berlin, pp. 636–651.

Hui, D., and Du, I. H. Y., 1986, "Effects of Axial Imperfections on Vibrations of Anti-Symmetric Cross-Ply, Oval Cylindrical Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 675–680.

Hutchinson, J. W., 1968, "Buckling and Initial Postbuckling Behavior of Oval Cylindrical Shells Under Axial Compression," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 66–72.

Kempner, J., and Chen, Y. N., 1967, "Buckling and Postbuckling of an Axially Compressed Oval Cylindrical Shell," *Proc. Symp. on the Theory of Shells to Honor Lloyd H. Donnell*, McCutchan Pub. Co., pp. 141–183.

Koiter, W. T., 1945, "On the Stability of Elastic Equilibrium," Thesis, Delft, Amsterdam (English Translation: AFFDL-TR-70-25), 1970.

Onoda, J., 1985, "Optimal Laminate Configurations of Cylindrical Shells for Axial Buckling," *AIAA J.*, Vol. 23, pp. 1093–1098.

Soldatos, K. P., 1984, "A Flugge-type Theory for the Analysis of Anisotropic Laminated Non-Circular Cylindrical Shells," *Int. J. of Solids and Structures*, Vol. 20, No. 2, pp. 107–120.

Soldatos, K. P., and Tzivanidis, G. J., 1982, "Buckling and Vibration of Cross-ply Laminated Non-Circular Cylindrical Shells," *J. of Sound and Vibration*, Vol. 82, No. 32, pp. 425–434.

Sun, G., 1986, "Optimization of Laminated Cylinders for Buckling," Doctoral Thesis, Institute for Aerospace Studies, University of Toronto, Toronto, Ontario, Canada.

Sun, G., and Hansen, J. S., 1988, "Optimal Design of Laminated-Composite Circular-Cylindrical Shells Subjected to Combined Loads," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 136–142.

Tennyson, R. C., Chan, K. H., and Muggeridge, D. B., 1971, "The Effect of Axisymmetric Shape Imperfections on the Buckling of Laminated Anisotropic Circular Cylinders," *CASI Trans.*, Vol. 4, No. 2, pp. 131–139.

Tennyson, R. C., and Hansen, J. S., 1983, "Optimum Design for Buckling of Laminated Cylinders," *Proc. IUTAM Symposium: Collapse: The Buckling of Structures in Theory and Practice*, J. M. T. Thompson, and G. W. Hunt, eds., Cambridge Univ. Press, Cambridge.

Zimmerman, R., 1986, "Optimization of Axially Compressed CFRP Cylinders," *Spacecraft Structures*. CNES, Toulouse, Dec. 3–6, 1985; ESA SP-238, Apr. 1986, pp. 407–412.

Stress Analysis of Smooth Polygonal Holes via a Boundary Perturbation Method

H. Gao[20]

Boundary Perturbation Method

This note presents an analysis of smooth polygon holes in an elastic sheet via a boundary perturbation method that can be formulated by the following steps. Let a sheet containing a circular hole be subjected to a fixed loading system and an additional point force P_i at a position (x, y), with $P_i\delta u_i(x, y)$ equal to the work done by the force at an incremental displacement $\delta u_i(x, y)$. Imagine that the hole boundary, initially at the circle $r = a$, is perturbed to a neighboring position by some variable normal distance $\delta a(\theta')$ (Fig. 1). Treating $\delta a(\theta')$ as infinitesimal, the change in the total energy Γ is

$$\delta\Gamma = P_i\delta u_i - \int_0^{2\pi} w\delta a(\theta')ad\theta' \tag{1}$$

where w is the strain energy density along the unperturbed circular hole boundary. For plane stress conditions, w is related to the hoop stress $\sigma_{\theta\theta}$ by

$$w = \sigma_{\theta\theta}^2/2E. \tag{2}$$

The energy variation due to the shape change of a void or hole was first studied by Rice and Drucker (1967). Let equation (1) be rearranged into the following form known as the Legendre transformation:

$$\delta(P_iu_i - \Gamma) = u_i\delta P_i + \left(\int_0^{2\pi} wg(\theta')ad\theta'\right)\delta A \tag{3}$$

where $\delta a(\theta')$ has been written as $g(\theta')\delta A$. The right side of equation (3) being a perfect differential, the coefficients of δP_i and δA must satisfy the Maxwell reciprocal relation

$$\frac{\partial u_i}{\partial A} = \frac{\partial}{\partial P_i}\left(\int_0^{2\pi} wg(\theta')ad\theta'\right) = \frac{1}{E}\int_0^{2\pi} \sigma_{\theta\theta}(\theta')\Sigma^i_{\theta\theta}(\theta';x, y)g(\theta')ad\theta'. \tag{4}$$

Here we have identified $\partial\sigma_{\theta\theta}/\partial P_i$ as the stress Green's function $\Sigma^i_{\theta\theta}(\theta';x, y)$, i.e., the stress at a boundary point θ' due to a unit point force in i direction at x, y. Multiplying both sides of (4) by δA and letting $P_i = 0$ yield

$$\delta u_i(x,y) = \frac{1}{E}\int_0^{2\pi} \sigma^0_{\theta\theta}(\theta')\Sigma^i_{\theta\theta}(\theta';x, y)\delta a(\theta')ad\theta' \tag{5}$$

where $\sigma^0_{\theta\theta}$ denotes the original stress field in absence of the point force P_i.

Equation (5) provides a first-order boundary perturbation formula for calculation of the displacement field of a nearly circular hole. In the perturbation analysis, a hole with a complex shape $r = a(\theta')$ is viewed as being perturbed from a reference circular hole by a small perturbation $\delta a(\theta')$. The displacement field for the actual hole is written as

$$u_i(x, y) = u_i^0(x, y) + \delta u_i(x, y) \tag{6}$$

where $u_i^0(x, y)$ is the reference solution (of a circular hole) and $\delta u_i(x, y)$ is given by (5). This perturbation procedure parallels an earlier development of a crack perturbation theory by Rice (1985).

[20]Professor, Division of Applied Mechanics, Stanford University, Stanford, CA 94305. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, Oct. 27, 1989; final revision, June 22, 1990.

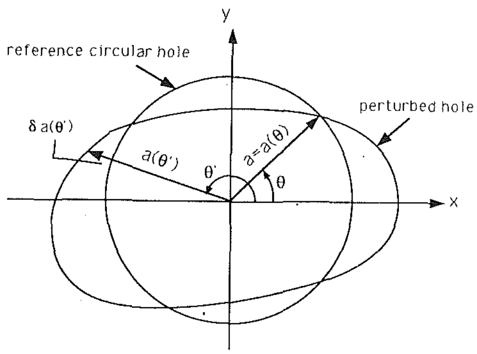


Fig. 1 A nearly circular hole with radius $a(\theta')$ that deviates slightly from constancy; the reference circle with radius $a(\theta)$ at the observation point θ

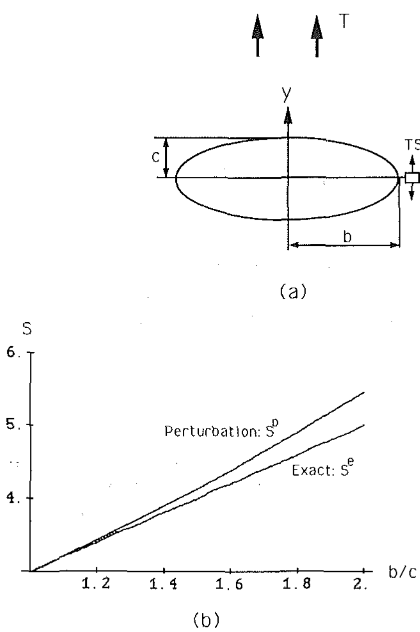


Fig. 2 (*a*) An elliptical hole subjected to remote uniaxial tension T; (b) pertubation versus exact stress concentration factor S^p (0) at the semi-long axis.

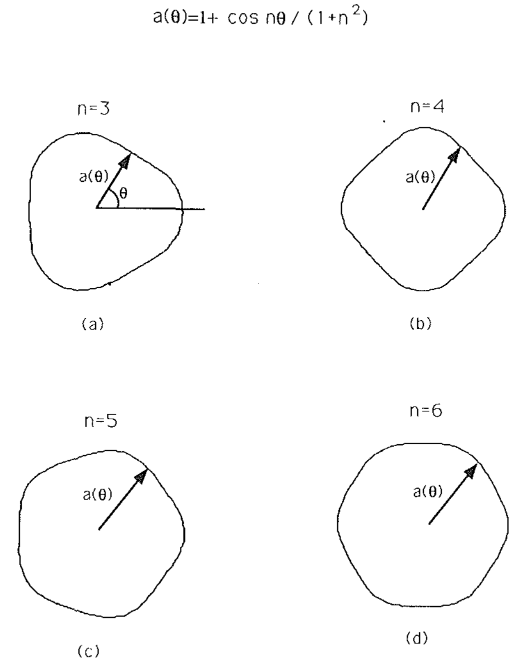


Fig. 3 Polygonal shapes simulated by sinusoidal wavy shapes: (*a*) equilateral triangle, (*b*) square, (*c*) pentagon, (*d*) hexagon

The stress field can be further derived from (5), (6) by Hooke's law. For proper integral convergence in the calculation of stress at a chosen boundary point θ, it is necessary to choose the reference circular hole at $r = a(\theta)$ so that $\delta a(\theta') = a(\theta') - a(\theta)$ vanishes when $\theta' = \theta$. Using the solution for the Green's function $\Sigma^i_{\theta\theta}$ in the literature (e.g., Green and Zerna, 1968) we have obtained the following first-order result

$$\sigma_{tt}(\theta) = \sigma^0_{\theta\theta}(\theta)$$
$$-\frac{1}{2\pi} PV \int_0^{2\pi} \frac{\cos(\theta' - \theta)}{\sin^2[(\theta' - \theta)/2]} \frac{\sigma^0_{\theta\theta}(\theta')[a(\theta') - a(\theta)]}{a(\theta)} d\theta' \quad (7)$$

for the hoop stress distribution along the given hole boundary at $r = a(\theta')$. Here "PV" denotes principal value in the Cauchy sense and $\sigma^0_{\theta\theta}(\theta')$ is the hoop stress distribution for the reference circular hole under the same loading conditions.

Elliptical Hole

To evaluate the accuracy of the perturbation formula (7), consider, for example, an elliptical hole subjected to uniaxial tension T (Fig. 2). The maximum stress occurs at the semi-long axis $\theta = 0$. Define the stress concentration factor of the elliptical hole as $S = \sigma_{tt}(0)/T$. Then the exact result for S is

$$S^e = 1 + 2b/c. \quad (8)$$

The shape function of an ellipse can be written as

$$a(\theta') = \frac{bc}{\sqrt{b^2 \sin^2\theta' + c^2 \cos^2\theta'}}. \quad (9)$$

Substituting (9) and the known reference stress distribution

$$\sigma^0_{\theta\theta}(\theta') = T(1 + 2\cos 2\theta') \quad (10)$$

into the perturbation formula (7) leads to the following first-order result for the stress concentration factor (at $\theta = 0$):

$$S^p = 3 - \frac{1}{2\pi} PV \int_0^{2\pi} \frac{(1 + 2\cos 2\theta')\cos\theta'}{\sin^2(\theta'/2)}$$
$$\times \left(\frac{1}{\sqrt{(b/c)^2 \sin^2\theta' + \cos^2\theta'}} - 1 \right) d\theta'. \quad (11)$$

As plotted in Fig. 2(*b*), the perturbation result calculated from (11) shows agreement with the exact solution (8) for the aspect ratio b/c as large as 1.6 within five percent error and for b/c as large as 2 within ten percent error.

Smooth Polygon Holes

It is interesting to consider holes with the following cosine shape function

$$a(\theta) = a_0 + A\cos n\theta \quad (12)$$

($A/a_0 << 1$). The case $n = 1$ corresponds to a rigid translation of the circle $r = a_0$ and $n = 2$ corresponds to slightly squeezing the circle into an ellipse. The above cosine function can also simulate polygon shapes with smoothed corners as shown in Fig. 3 for $n \geq 3$. The polygon shapes are best approximated by (12) when the curvature is required to vanish at the most concave locations where $\cos n\theta = -1$, corresponding to having

$$A = a_0/(1 + n^2). \quad (13)$$

The fact that $A/a_0 \leq 0.1$ when $n \geq 3$ indicates that the polygon shapes given by (12), (13) represent small perturbations from a circle, and suggests use of the boundary perturbation formula (7) to estimate the stresses. When the holes are subjected to a remote tension $\sigma^\infty_{yy} = T$, the perturbation formula (7) predicts

$$\frac{\sigma_{tt}(\theta)}{T}=1+2\cos 2\theta - \frac{A}{2\pi a_0} PV\int_0^{2\pi} \frac{\cos(\theta'-\theta)(1+2\cos 2\theta')(\cos n\theta'-\cos n\theta)}{\sin^2[(\theta'-\theta)/2]}\,d\theta'. \tag{14}$$

Carrying out the integration yields the final result

$$\frac{\sigma_{tt}(\theta)}{T}=1+2\cos 2\theta-\frac{4A}{a_0}\sin 2\theta\sin\theta \tag{15}$$

for translation mode $n = 1$,

$$\frac{\sigma_{tt}(\theta)}{T}=1+2\cos 2\theta+\frac{2A}{a_0}(2\cos 4\theta+\cos 2\theta-1) \tag{16}$$

for elliptical mode $n = 2$ and

$$\frac{\sigma_{tt}(\theta)}{T}=1+2\cos 2\theta+\frac{2A}{a_0}[(n-1)\cos 2\theta + 2(n-2)\cos 2\theta\cos n\theta-4\sin 2\theta\sin n\theta] \tag{17}$$

for polygon modes $n \geq 3$. Maximum stress concentration for the smooth polygons ($n > 3$, $A/a_0 = 1/(1+n^2)$) occurs at $\theta = 0$ with the stress concentration factor

$$S=3+2(3n-5)/(1+n^2). \tag{18}$$

The above S equals 3.8 for the triangular shape $n = 3$ and approximately 3.82 for the square shape $n = 4$, then it monotonically decreases toward the asymptotical value 3 as the n is increased to approach the circular limit (a convex polygon with infinite number of edges). Therefore, the smooth shapes given by (12), (13) compromise between reasonable stress concentration factor ($3 \leq S \leq 3.82$) and closely matching polygon shapes (Fig. 3). These polygon shapes may find useful applications in a practical design process.

A Fourier analysis for a nearly circular hole can be developed based on our perturbation results. Assume a shape function which is symmetric about the x-axis and given by the series expansion

$$a(\theta)=a_0+\sum_{n=1}^{\infty}A_n\cos n\theta \tag{19}$$

where

$$2\pi a_0=\int_0^{2\pi}a(\theta)d\theta,\quad \pi A_n=\int_0^{2\pi}a(\theta)\cos n\theta d\theta. \tag{20}$$

Then it follows from (17) that the stress concentration factor (at $\theta = 0$) is

$$S=3+\sum_{n=1}^{\infty}\frac{2A_n}{a_0}\,(3n-5). \tag{21}$$

By standard theorems on Fourier series, the last expression will converge when $da(\theta)/d\theta$ is continuous at $\theta = 0$.

Conclusion

The boundary perturbation formula (7) seems to give reasonable predictions for the stress distribution of a nearly circular hole in an elastic sheet. Of particular interest are the cosine shape functions (12), (13) which have been found to give closely matching polygon shapes with only minor stress concentration factor. These shapes may thus find practical applications in a design process.

The methodology developed here provides a perturbation approach to study holes, voids, inclusions, surfaces and interfaces with complex shape profiles, based on known solutions for a simple geometry such as a flat surface or a circular hole. The method has recently been applied to study the stress concentration effect of an undulating surface morphology by Gao (1991).

Acknowledgment

The work reported is supported by the National Science Foundation under Research Initiation Grant No. MSS-9008521.

References

Gao, H., 1991, "Stress Concentration at Slightly Undulating Surfaces," *Journal of the Mechanics and Physics of Solids*, in press.

Green, A. E. and Zerna, W., 1968, *Theoretical Elasticity*, 2nd ed., Oxford University Press, London.

Rice, J. R., 1985, "First Order Variations in Elastic Fields Due to Variation in Location of a Planar Crack Front," ASME Journal of Applied Mechanics, Vol. 52, pp. 571–579.

Rice, J. R., and Drucker, D. C., 1967, "Energy Changes in Stressed Solids due to Void and Crack Growth," *International Journal of Fracture Mechanics*, Vol. 3, pp. 19–27.

Closed-Form Solution for the Composite Sphere Subject to Quadratic Eigenstrains With Radial Symmetry

Mauro Ferrari[21]

Introduction

This paper contains the closed-form solution for the problem of a biphase sphere, in the presence of a polar symmetric eigenstrain field, represented by a quadratic polynomial in each phase. In general, the determination of the residual stress field generated in composite structures during service, as well as during deposition and forming procedures, is essential for failure analysis and for an effective design of these procedures. In particular, the present problem is relevant in the analysis of glass-fiber-reinforced polymers with an inhomogeneous matrix moisture absorption, and of plasma-sprayed ceramic coatings on metallic substrates (Ferrari and Harding, 1990).

A comprehensive discussion of eigenstrain problems is given in Mura's treatise (1982). The case of polynomial eigenstrains in a subdomain of an infinite body is also presented there. The problem of a three-phase spherically concentric solid, subject to polynomial eigenstrains in the included core only, is solved by Luo and Weng (1987) using a procedure of Eshelby's (1957). Multilayered bodies, subject to an homogeneous temperature change, are considered by Mikata and Taya (1986). The analysis of a solid or hollow sphere subject to a radially symmetric eigenstrain, of thermal nature, can be found in Chapter 9 of (Boley and Weiner, 1985). This fundamental text also contains an extensive literature review on the problem of the thermally loaded sphere. An approximate analysis of thermally-induced stresses in multilayered structures is found in (Suhir, 1988).

The General Eigenstrain Problem

Given a body $B+\partial B$, of boundary ∂B, subject to an eigenstrain field ϵ^*, the field equations governing the displacement $\mathbf{u}$, the strain ϵ, and the stress τ are:
the strain-displacement relations

[21]Istituto di Meccanica Teorica ed Applicata, Universita di Udine, via Ungheria 43, 33100 Udine, Italy. Presently at the Department of Civil Engineering, University of California at Berkeley, Berkeley, CA 94720.

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers. Manuscript received by the ASME Applied Mechanics Division, Jan. 29, 1990; final revision, May 20, 1990.

$$\frac{\sigma_{tt}(\theta)}{T} = 1 + 2\cos 2\theta - \frac{A}{2\pi a_0} PV\int_0^{2\pi} \frac{\cos(\theta' - \theta)(1 + 2\cos 2\theta')(\cos n\theta' - \cos n\theta)}{\sin^2[(\theta' - \theta)/2]} d\theta'. \tag{14}$$

Carrying out the integration yields the final result

$$\frac{\sigma_{tt}(\theta)}{T} = 1 + 2\cos 2\theta - \frac{4A}{a_0}\sin 2\theta \sin\theta \tag{15}$$

for translation mode $n = 1$,

$$\frac{\sigma_{tt}(\theta)}{T} = 1 + 2\cos 2\theta + \frac{2A}{a_0}(2\cos 4\theta + \cos 2\theta - 1) \tag{16}$$

for elliptical mode $n = 2$ and

$$\frac{\sigma_{tt}(\theta)}{T} = 1 + 2\cos 2\theta + \frac{2A}{a_0}[(n-1)\cos 2\theta + 2(n-2)\cos 2\theta \cos n\theta - 4\sin 2\theta \sin n\theta] \tag{17}$$

for polygon modes $n \geq 3$. Maximum stress concentration for the smooth polygons ($n > 3$, $A/a_0 = 1/(1+n^2)$) occurs at $\theta = 0$ with the stress concentration factor

$$S = 3 + 2(3n-5)/(1+n^2). \tag{18}$$

The above S equals 3.8 for the triangular shape $n = 3$ and approximately 3.82 for the square shape $n = 4$, then it monotonically decreases toward the asymptotical value 3 as the n is increased to approach the circular limit (a convex polygon with infinite number of edges). Therefore, the smooth shapes given by (12), (13) compromise between reasonable stress concentration factor ($3 \leq S \leq 3.82$) and closely matching polygon shapes (Fig. 3). These polygon shapes may find useful applications in a practical design process.

A Fourier analysis for a nearly circular hole can be developed based on our perturbation results. Assume a shape function which is symmetric about the x-axis and given by the series expansion

$$a(\theta) = a_0 + \sum_{n=1}^{\infty} A_n \cos n\theta \tag{19}$$

where

$$2\pi a_0 = \int_0^{2\pi} a(\theta)d\theta, \quad \pi A_n = \int_0^{2\pi} a(\theta)\cos n\theta d\theta. \tag{20}$$

Then it follows from (17) that the stress concentration factor (at $\theta = 0$) is

$$S = 3 + \sum_{n=1}^{\infty} \frac{2A_n}{a_0}\,(3n-5). \tag{21}$$

By standard theorems on Fourier series, the last expression will converge when $da(\theta)/d\theta$ is continuous at $\theta = 0$.

Conclusion

The boundary perturbation formula (7) seems to give reasonable predictions for the stress distribution of a nearly circular hole in an elastic sheet. Of particular interest are the cosine shape functions (12), (13) which have been found to give closely matching polygon shapes with only minor stress concentration factor. These shapes may thus find practical applications in a design process.

The methodology developed here provides a perturbation approach to study holes, voids, inclusions, surfaces and interfaces with complex shape profiles, based on known solutions for a simple geometry such as a flat surface or a circular hole. The method has recently been applied to study the stress concentration effect of an undulating surface morphology by Gao (1991).

Acknowledgment

The work reported is supported by the National Science Foundation under Research Initiation Grant No. MSS-9008521.

References

Gao, H., 1991, "Stress Concentration at Slightly Undulating Surfaces," *Journal of the Mechanics and Physics of Solids*, in press.

Green, A. E. and Zerna, W., 1968, *Theoretical Elasticity*, 2nd ed., Oxford University Press, London.

Rice, J. R., 1985, "First Order Variations in Elastic Fields Due to Variation in Location of a Planar Crack Front," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 571–579.

Rice, J. R., and Drucker, D. C., 1967, "Energy Changes in Stressed Solids due to Void and Crack Growth," *International Journal of Fracture Mechanics*, Vol. 3, pp. 19–27.

Closed-Form Solution for the Composite Sphere Subject to Quadratic Eigenstrains With Radial Symmetry

Mauro Ferrari[21]

Introduction

This paper contains the closed-form solution for the problem of a biphase sphere, in the presence of a polar symmetric eigenstrain field, represented by a quadratic polynomial in each phase. In general, the determination of the residual stress field generated in composite structures during service, as well as during deposition and forming procedures, is essential for failure analysis and for an effective design of these procedures. In particular, the present problem is relevant in the analysis of glass-fiber-reinforced polymers with an inhomogeneous matrix moisture absorption, and of plasma-sprayed ceramic coatings on metallic substrates (Ferrari and Harding, 1990).

A comprehensive discussion of eigenstrain problems is given in Mura's treatise (1982). The case of polynomial eigenstrains in a subdomain of an infinite body is also presented there. The problem of a three-phase spherically concentric solid, subject to polynomial eigenstrains in the included core only, is solved by Luo and Weng (1987) using a procedure of Eshelby's (1957). Multilayered bodies, subject to an homogeneous temperature change, are considered by Mikata and Taya (1986). The analysis of a solid or hollow sphere subject to a radially symmetric eigenstrain, of thermal nature, can be found in Chapter 9 of (Boley and Weiner, 1985). This fundamental text also contains an extensive literature review on the problem of the thermally loaded sphere. An approximate analysis of thermally-induced stresses in multilayered structures is found in (Suhir, 1988).

The General Eigenstrain Problem

Given a body $B + \partial B$, of boundary ∂B, subject to an eigenstrain field ϵ^*, the field equations governing the displacement $\mathbf{u}$, the strain ϵ, and the stress τ are:
the strain-displacement relations

[21]Istituto di Meccanica Teorica ed Applicata, Universita di Udine, via Ungheria 43, 33100 Udine, Italy. Presently at the Department of Civil Engineering, University of California at Berkeley, Berkeley, CA 94720.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS. Manuscript received by the ASME Applied Mechanics Division, Jan. 29, 1990; final revision, May 20, 1990.

$$\epsilon = \mathrm{sym}(\mathrm{grad}(\mathbf{u})), \tag{1a}$$

the constitutive equations

$$\tau = \mathbf{C}(\epsilon - \epsilon^*), \tag{1b}$$

and the equilibrium conditions

$$\mathrm{div}\ \tau + \mathbf{f} = \mathbf{0}. \tag{1c}$$

Here, $\mathbf{C}$ and $\mathbf{f}$ denote the elastic stiffness tensor field, and the body force vector, respectively, while div(.), grad(.) and sym(.) are the divergence, the gradient, and the symmetric part operators.

Appropriate boundary condition for the field equations (1) are specified as

$$\tau\ \mathbf{n} = \hat{\mathbf{t}} \quad \text{on } \partial B_t \tag{2a}$$

where $\mathbf{n}$ is the outward unit normal, and

$$\mathbf{u} = \hat{\mathbf{u}} \quad \text{on } \partial B_u. \tag{2b}$$

In equations (2), $\hat{\mathbf{t}}$ and $\hat{\mathbf{u}}$ are assigned tractions and displacement vectors, respectively, and ∂B_t and ∂B_u are complementary portions of ∂B.

The field equation system (1) may be reduced to the form

$$\mathrm{div}[\mathbf{C}\ \mathrm{sym}(\mathrm{grad}(\mathbf{u}))] = \mathrm{div}(\mathbf{C}\ \epsilon^*) - \mathbf{f}. \tag{3}$$

The case of zero-body forces and applied boundary tractions is studied below. By virtue of the linearity of the field equations, the solution for the eigenstrained body with applied external forces may be obtained upon superposition of the solution for the eigenstrained body with no external loadings and of the solution for the externally loaded body with no eigenstrains.

The Eigenstrained Composite Sphere

Let the spherical region $0 \le r \le R_1$ be occupied by material 1, and the surrounding region $R_1 \le r \le R_2$ be occupied by material 2. Both materials are isotropic and homogeneous, and are subject to the eigenstrain fields

$$\epsilon_i^* = \alpha_i\ (A_i\ r^2 + B_i\ r + C_i)\mathbf{I} \quad i = 1,2. \tag{4}$$

Here, $\mathbf{I}$ is the identity tensor, A_i, B_i, and C_i are specified eigenstrain parameters, and α_i is a material constant, corresponding to the type of eigenstrain. The assumption of material isotropy and the polar symmetry of the problem reduce the equilibrium requirements (3) for the ith phase to the single equation

$$\frac{d^2\ u_i}{dr^2} + \frac{2}{r}\left[\frac{du_i}{dr} - \frac{u_i}{r}\right] = \alpha_i\ \frac{3\lambda_i + 2\mu_i}{\lambda_i + 2\mu_i}\ (2\ A_i\ r + B_i), \quad i = 1,2, \tag{5}$$

in the ith phase radial displacement u_i. Here, the λ_i and μ_i are Lame's constants. Equation (5) is expressed in a spherical polar coordinate system (r, φ, ϑ). This is employed throughout the present work.

The solution of (5) is, for the ith phase,

$$u_i = \alpha_i\ \frac{3\lambda_i + 2\mu_i}{20(\lambda_i + 2\mu_i)}\ (4\ A_i\ r + 5\ B_i)\ r^2 + k_{i1} r + k_{i2}/r^2, \tag{6}$$

where the k_{ij} are constants of integration. The normal stresses in the radial and transverse directions are, respectively,

$$\tau_i^r = -(3\lambda_i + 2\mu_i)\left\{\alpha_i\left[\frac{\mu_i\ (4A_i\ r + 5B_i)}{5\ (\lambda_i + 2\mu_i)}\ r + C_i\right] - k_{i1}\right\} - \frac{4\mu_i k_{i2}}{r^3} \tag{7}$$

$$\tau_i^t = -(3\lambda_i + 2\mu_i)\left\{\alpha_i\left[\frac{\mu_i\ (16A_i\ r + 15B_i)}{10\ (\lambda_i + 2\mu_i)}\ r + C_i\right] - k_{i1}\right\} + \frac{2\mu_i k_{i2}}{r^3}.$$

The constants of integration are determined upon imposing boundary, interface and boundedness conditions:

$$\begin{gathered} \tau_2^r = 0 \quad \text{at } r = R_2 \\ \left.\begin{aligned} u_1 &= u_2 \\ \tau_1^r &= \tau_2^r \end{aligned}\right\} \quad \text{at } r = R_1 \\ u_1 = 0 \quad \text{at } r = 0. \end{gathered} \tag{8}$$

It is noted that the vanishing of u_1 at the origin is equivalent to its being infinite there, and mathematically given by the vanishing of k_{12}.

Upon imposing equations (8), the integration constants are found:

$$\begin{aligned} k_{11} = \frac{1}{5\ D\ \beta_{13}}\ \{\alpha_1\ \beta_{11}\{L_{11}[\beta_{21}(\mu_1 - \mu_2)R_2^3 + \mu_2\ \beta_{12}\ R_1^3] + \\ + 5\ N_1\ (\beta_{21}\ R_2^3 + 4\ \mu_2\ R_1^3)\} + \\ + \alpha_2\ \beta_{21}\ \beta_{13}\ \mu_2\ (M_{22} - M_{21})\} \end{aligned} \tag{9}$$

$$\begin{aligned} k_{21} = \frac{1}{15\ D\ \beta_{23}}\ \{\alpha_1\ \beta_{11}\ \beta_{23}\ \mu_2\ M_{11} + \\ + \alpha_2\ \beta_{21}[\beta_{22}\ \mu_2\ (L_{22}\ R_2^3 - L_{21}\ R_1^3) + 5\ \beta_{11}\ N_2]\} \end{aligned} \tag{10}$$

$$\begin{aligned} k_{22} = (\beta_{21}\ R_2^3/20\ D\ \beta_{23})\{\alpha_1\ \beta_{11}\ \beta_{23}\ M_{21} + \\ - \alpha_2\ R_1^3[4\ \beta_4\ \mu_2\ L_{22} + + \beta_{21}\ \beta_{22}\ L_{21} + 20\ \beta_{11}\ N_2]\}. \end{aligned} \tag{11}$$

For these equations, the following definitions are introduced:

Material parameters

$$\begin{aligned} &\beta_{i1} \equiv 2\ \mu_i + 3\ \lambda_i = 3\ k_i \\ &\beta_{i2} \equiv 2\ \mu_j + 3\ \lambda_j + 4\ \mu_i = \beta_{j1} + 4\ \mu_i, \qquad i \ne j \\ &\beta_{i3} \equiv 2\ \mu_i + \lambda_i \\ &\beta_4 \equiv \beta_{11} - \beta_{21}. \end{aligned} \tag{12}$$

Loading-geometry parameters (no sum on repeated indices)

$$\begin{aligned} L_{ij} &\equiv [4\ A_i\ R_j + 5\ B_i]R_j \\ M_{ij} &\equiv [3\ L_{ij} + 20\ C_i]R_j^3 \\ N_i &\equiv \beta_{i3}\ C_i \end{aligned} \tag{13}$$

and

$$D \equiv (\beta_{21}\ \beta_{22}\ R_2^3 + 4\ \mu_2\ \beta_4\ R_1^3)/3. \tag{14}$$

In equation (12), k_i is the bulk modulus of the ith phase.

For a solid or hollow sphere, the solution (6)–(7) with (9)–(14) may be obtained following the method of Boley and Wiener (1985, Section 9.14). If the outer phase is not subjected to eigenstrains—i.e., if $A_2 = B_2 = C_2 = 0$ – then the above solution is a special case of the general solution given by Luo and Weng (1987) for the inclusion problem in a three-phase spherical concentric solid: Our solution may be obtained from that memoir upon letting the moduli of their "c" phase vanish, and letting $N = 3$ in their equation (4.1).

In the presence of a spatially uniform eigenstrain throughout the composite sphere ($A_i = B_i = 0$, $C_i \equiv C$, $i = 1,2$), the stresses may be put in the simple form

$$\begin{aligned} \tau_1^r &= \tau_1^t = C\ (\alpha_1 - \alpha_2)\ (1 - R_2^3/R_1^3)\ P \\ \tau_2^r &= C\ (\alpha_1 - \alpha_2)\ (1 - R_2^3/r^3)\ P \\ \tau_2^t &= C\ (\alpha_1 - \alpha_2)\ (1 + R_2^3/2\ r^3)\ P \end{aligned} \tag{15}$$

where P is defined as

$$P = 4\ \beta_{11}\ \mu_2\ R_1^3/(4\ \beta_4\ \mu_2\ R_1^3 + \beta_{21}\ \beta_{22}\ R_2^3). \tag{16}$$

It may be noted that (15) implies that $\mathrm{sign}(\tau_1^r) = \mathrm{sign}(\tau_2^r) = = \mathrm{sign}(\tau_1^t) = -\mathrm{sign}(\tau_2^t)$, that is, the normal stresses in the $\vartheta-$ and $\varphi-$ directions in the cover material are tensile whenever all other stresses are compressive, and vice versa.

For the case of a thin film, upon introducing a measure of smallness $\epsilon \equiv (R_2^3 - R_1^3)/R_1^3$, from (16), one deduces that

$$P(\epsilon) = \frac{4\,\mu_2}{\beta_{21} + 4\,\mu_2}\left[1 - \frac{4\,\beta_{21}\,\beta_{22}\,\mu_2}{\beta_{21} + \mu_2}\,\epsilon\right] \tag{17}$$

to first order in ϵ. Equation (15) then shows that all other normal stresses are negligible, with respect to τ_2^i, in the linear theory.

References

Boley, B. A., and Weiner, J. H., 1985, *Theory of Thermal Stresses*, 2nd ed., Krieger, Malabar.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proc. Roy. Soc. London*, Vol. A241, pp. 376–396.

Ferrari, M., and Harding, J. H., 1990, "Residual Stress Profile in a Plasma-Sprayed Ceramic Coating," *Composite Material Technology 1990*, D. Hui and T. J. Kozik, eds., ASME, New York, pp. 197–202.

Luo, H. A., and Weng, G. J., 1987, "On Eshelby's Inclusion Problem in a Three-Phase Spherically Concentric Solid, and a Modification of Mori-Tanaka's Method," *Mechanics of Materials*, Vol. 6, pp. 347–361.

Mikata, Y., and Taya, M., 1986, "Thermal Stresses in a Coated Short-Fiber Composite," ASME Journal of Applied Mechanics, Vol. 53, pp. 681–685.

Mura, T., 1982, *Micromechanics of Defects in Solids*, Martinus Nijoff, The Hague.

Suhir, E., 1988 "An Approximate Analysis of Stresses in Multilayered Elastic Thin Films," ASME Journal of Applied Mechanics, Vol. 55, pp. 143–148.

Contact of a Smooth Flat Indenter on a Layered Elastic Half-Space: Beam on an Elastic Foundation Model

T. W. Shield[22]

Our previous work (Shield and Bogy, 1989; SB89 in the following) considered the elastic contact of a rigid flat indenter on a layered elastic half-space. It was found that there are three possible contact region configurations depending on the geometrical and material properties. In the case of a single layer the type of solution depends on the layer thickness to indenter half-width ratio, $h^* = h/a$, and the shear modulus of the layer relative to the substrate, $\mu^* = \mu_l/\mu_s$, where μ_l is the shear modulus of the layer and μ_s is the shear modulus of the substrate. Because the calculations involved in the full elastic solution are extensive, it is of interest to consider simplifications of the problem. The problem is greatly simplified if the elastic system is modeled by an elastic beam on a Winkler foundation. The complicated system of partial differential equations for the problem reduces to a single ordinary differential equation in the surface displacement. However, as will be shown, many features of the solution are lost.

This simplified problem has already been considered in another context. The equation for the normal displacement of a beam on an elastic foundation is identical to the equation for shrink-fit problems involving cylindrical shells with a simple redefinition of the parameters. The case of a finite length elastic ring shrunk-fit around a cylindrical shell was first considered by Paul (1962). Becker (1962) in a discussion of Paul's work noted that separation was possible between the ring and the shell for a range of the parameters. Becker and Paul only considered the case of a single region of separation, although Becker noted a two region configuration was possible. Bogy (1987) considered problems involving rigid plugs shrunk-fit inside cylindrical shells that have one or two regions of separation. These two cases are the only configuration found to exist by Bogy. The configuration with no regions of separation is not admissible because simple shell theory does not permit the displacement kink (slope singularity) that would be needed to allow the shell to conform to the plug at all points. The finding that there are no solutions with more than two separation regions agrees with the argument presented in SB89 that there are only three possible contact region configurations.

To allow comparison, the shrink-fit problem considered by Bogy will be reformulated as indenter problem and then his results will be presented along with the results of SB89. Figure 1 shows the geometry of the shrink-fit problem and the corresponding indentation problem. The ODE that governs the behavior of the shell is equation (1) of Bogy (1974),

$$d^4w/dx^4 + 4\beta^4 w = Z/D\ , \tag{1}$$

where w is the normal displacement of the shell and the Z is the loading. The bending stiffness, $D = Eh^3/12(1-\nu^2)$, is unchanged in our application, only the definition of β needs to be modified. In the case of a cylindrical shell, β depends on the shell radius and Poisson's ratio. For the case of an elastic foundation we have

$$\beta = \sqrt[4]{K/4D}, \tag{2}$$

where K is the stiffness of the foundation. To allow comparison to the full elastic solution, we need to write β in terms of h^* and μ^*. This is accomplished by considering the quantity $\alpha = \beta l/2$, which is nondimensional. The length of the plug, l, in Bogy's solution corresponds to the indenter with $2a$. Thus,

$$h^* = 2h/l\ ,$$

where the layer thickness, h, is the same in both cases, as is the modulus of the layer E. The substrate stiffness K cannot be directly related to the actual elastic modulus of the substrate E_s. A plane-strain formulation does not allow calculation of the total indentation (and hence the compliance) of the surface because of the singular nature of the displacements at infinity. A rough approximation is that

$$K = E_s/l\ . \tag{4}$$

Combining the above identifications gives

$$\alpha \equiv \beta l/2 = \sqrt[4]{\frac{6(1-\nu_l)(1+\nu_s)}{\mu^*(h^*)^3}}\ , \tag{5}$$

where ν_l and ν_s are the Poisson's ratios of the layer and substrate, respectively. All of the results of interest from Bogy (1974) are in terms of this single quantity which involves the

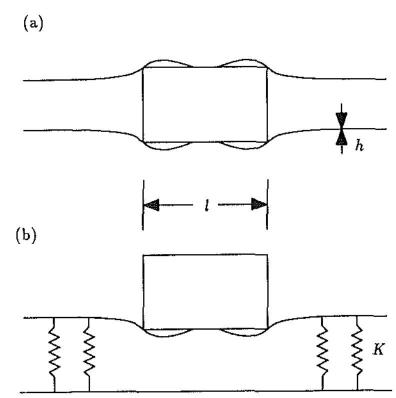


Fig. 1 The geometry of the shrink-fit problem (*a*) and the contact problem (*b*). The thickness of the shell is not shown for clarity.

[22]Brown University, Providence, RI 02912. Currently at the Department of Aerospace Engineering and Mechanics, University of Minnesota, MN 55455. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, Jan. 23, 1990; final revision, Sept. 11, 1990.

stresses in the $\vartheta-$ and $\varphi-$ directions in the cover material are tensile whenever all other stresses are compressive, and vice versa.

For the case of a thin film, upon introducing a measure of smallness $\epsilon \equiv (R_2^3 - R_1^3)/R_1^3$, from (16), one deduces that

$$P(\epsilon) = \frac{4\,\mu_2}{\beta_{21} + 4\,\mu_2}\left[1 - \frac{4\,\beta_{21}\,\beta_{22}\,\mu_2}{\beta_{21} + \mu_2}\,\epsilon\right] \tag{17}$$

to first order in ϵ. Equation (15) then shows that all other normal stresses are negligible, with respect to τ_2^t, in the linear theory.

References

Boley, B. A., and Weiner, J. H., 1985, *Theory of Thermal Stresses*, 2nd ed., Krieger, Malabar.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proc. Roy. Soc. London*, Vol. A241, pp. 376–396.

Ferrari, M., and Harding, J. H., 1990, "Residual Stress Profile in a Plasma-Sprayed Ceramic Coating," *Composite Material Technology 1990*, D. Hui and T. J. Kozik, eds., ASME, New York, pp. 197–202.

Luo, H. A., and Weng, G. J., 1987, "On Eshelby's Inclusion Problem in a Three-Phase Spherically Concentric Solid, and a Modification of Mori-Tanaka's Method," *Mechanics of Materials*, Vol. 6, pp. 347–361.

Mikata, Y., and Taya, M., 1986, "Thermal Stresses in a Coated Short-Fiber Composite," ASME Journal of Applied Mechanics, Vol. 53, pp. 681–685.

Mura, T., 1982, *Micromechanics of Defects in Solids*, Martinus Nijoff, The Hague.

Suhir, E., 1988 "An Approximate Analysis of Stresses in Multilayered Elastic Thin Films," ASME Journal of Applied Mechanics, Vol. 55, pp. 143–148.

Contact of a Smooth Flat Indenter on a Layered Elastic Half-Space: Beam on an Elastic Foundation Model

T. W. Shield[22]

Our previous work (Shield and Bogy, 1989; SB89 in the following) considered the elastic contact of a rigid flat indenter on a layered elastic half-space. It was found that there are three possible contact region configurations depending on the geometrical and material properties. In the case of a single layer the type of solution depends on the layer thickness to indenter half-width ratio, $h^* = h/a$, and the shear modulus of the layer relative to the substrate, $\mu^* = \mu_l/\mu_s$, where μ_l is the shear modulus of the layer and μ_s is the shear modulus of the substrate. Because the calculations involved in the full elastic solution are extensive, it is of interest to consider simplifications of the problem. The problem is greatly simplified if the elastic system is modeled by an elastic beam on a Winkler foundation. The complicated system of partial differential equations for the problem reduces to a single ordinary differential equation in the surface displacement. However, as will be shown, many features of the solution are lost.

This simplified problem has already been considered in another context. The equation for the normal displacement of a beam on an elastic foundation is identical to the equation for shrink-fit problems involving cylindrical shells with a simple redefinition of the parameters. The case of a finite length elastic ring shrunk-fit around a cylindrical shell was first considered by Paul (1962). Becker (1962) in a discussion of Paul's work noted that separation was possible between the ring and the shell for a range of the parameters. Becker and Paul only considered the case of a single region of separation, although Becker noted a two region configuration was possible. Bogy (1987) considered problems involving rigid plugs shrunk-fit inside cylindrical shells that have one or two regions of separation. These two cases are the only configuration found to exist by Bogy. The configuration with no regions of separation is not admissible because simple shell theory does not permit the displacement kink (slope singularity) that would be needed to allow the shell to conform to the plug at all points. The finding that there are no solutions with more than two separation regions agrees with the argument presented in SB89 that there are only three possible contact region configurations.

To allow comparison, the shrink-fit problem considered by Bogy will be reformulated as indenter problem and then his results will be presented along with the results of SB89. Figure 1 shows the geometry of the shrink-fit problem and the corresponding indentation problem. The ODE that governs the behavior of the shell is equation (1) of Bogy (1974),

$$d^4w/dx^4 + 4\beta^4 w = Z/D\ , \tag{1}$$

where w is the normal displacement of the shell and the Z is the loading. The bending stiffness, $D = Eh^3/12(1-\nu^2)$, is unchanged in our application, only the definition of β needs to be modified. In the case of a cylindrical shell, β depends on the shell radius and Poisson's ratio. For the case of an elastic foundation we have

$$\beta = \sqrt[4]{K/4D}, \tag{2}$$

where K is the stiffness of the foundation. To allow comparison to the full elastic solution, we need to write β in terms of h^* and μ^*. This is accomplished by considering the quantity $\alpha = \beta l/2$, which is nondimensional. The length of the plug, l, in Bogy's solution corresponds to the indenter with $2a$. Thus,

$$h^* = 2h/l\ ,$$

where the layer thickness, h, is the same in both cases, as is the modulus of the layer E. The substrate stiffness K cannot be directly related to the actual elastic modulus of the substrate E_s. A plane-strain formulation does not allow calculation of the total indentation (and hence the compliance) of the surface because of the singular nature of the displacements at infinity. A rough approximation is that

$$K = E_s/l\ . \tag{4}$$

Combining the above identifications gives

$$\alpha \equiv \beta l/2 = \sqrt[4]{\frac{6(1-\nu_l)(1+\nu_s)}{\mu^*(h^*)^3}}\ , \tag{5}$$

where ν_l and ν_s are the Poisson's ratios of the layer and substrate, respectively. All of the results of interest from Bogy (1974) are in terms of this single quantity which involves the

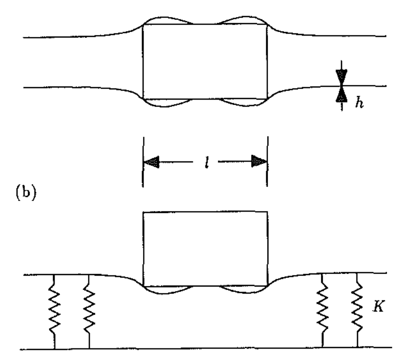


Fig. 1 The geometry of the shrink-fit problem (*a*) and the contact problem (*b*). The thickness of the shell is not shown for clarity.

[22]Brown University, Providence, RI 02912. Currently at the Department of Aerospace Engineering and Mechanics, University of Minnesota, MN 55455. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, Jan. 23, 1990; final revision, Sept. 11, 1990.

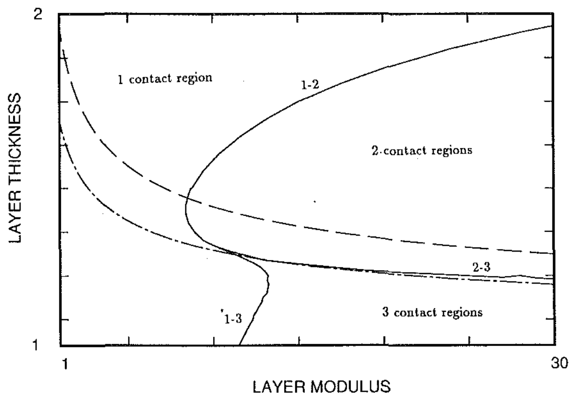


Fig. 2 The (μ^*, h^*) parameter space for the contact problem. The solid curves are the exact solution for $\nu_s = 1/3$, the dashed curve is $\alpha = 0.929$ and the dot-dashed curve is $\alpha = 1.1872$.

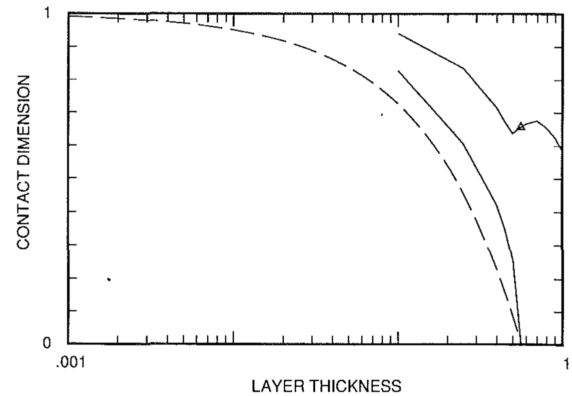


Fig. 3 A comparison of the contact dimension d for the exact solution (lower solid curve) and the approximate solution (dashed curve) for $\mu^* = 15.0$. The contact dimension c, of the exact solution, is the upper solid curve.

two parameters of the full elastic solution. The reduction of the number of parameters from two to one is an indication of the large amount of simplification involved in this approximation.

Bogy found that three possible configurations exist for the shell surface enclosing the rigid plug. Two of them have the same number of separation regions. The short plug solution involves contact of the shell with the plug only at its two corners. This solution is valid for a parameter value of $\alpha \leq 0.929$, equality occurs as the shell surface just touches the center of the plug. For $0.929 \leq \alpha \leq 1.1872$, the shell surface touches the center of the plug (as well as the corners) and a ring load is needed at the center of the plug in the intermediate solution. The long plug solution occurs when $\alpha \geq 1.1872$, and the contact region at the center of the plug has a finite extent over which a uniform pressure acts. At the ends of the central contact region there are ring loads. The boundary points between these three types of solutions are at $\alpha = 0.929$ and $\alpha = 1.1872$. Equation (5) allows these points to be drawn as curves in the (μ^*, h^*) parameter space of Fig. 2. These curves are drawn as the dashed and dot-dashed curves, respectively. The solid curves are the results from SB89 which will be recalled shortly.

For the long plug solution the extent of contact can also be calculated. The size of the separated region for the long plug solution is denoted by Bogy as l'_∞. The quantity $\alpha'_\infty = \beta l'_\infty/2$ has the constant value of 0.5936 for this solution. In SB89 we calculated the distance from the center of the indenter to the point of separation, d. This dimension can be related to l'_∞ as follows:

$$d = 1 - 2l'_\infty/l = 1 - 2(0.5936)/\alpha \ , \tag{6}$$

where α is given by (5).

Figures 2 and 3 also include the results from SB89 for the exact elastic solution. The solid curves in Fig. 2 are the boundaries of the zones of existence of the three types of solutions. We observe that one solid curve generally separates the zone of lower modulus from the rest of the plane. It divides the plane into the domains of low modulus, single contact region solutions and the high modulus, multiple contact region solutions. Because all of the modulus ratios shown in Fig. 2 are greater than one, it follows that multiple contact region solutions occur only when the layer is stiffer than the substrate. The curve that is almost parallel to the modulus axis in the high modulus zone divides the multiple contact region zone into the two and three contact region zones.

The single-multiple contact region solution separation curve in Fig. 2 on which $p(x) = 0$ for some isolated x, can be divided into two parts. If, as μ^* increases for fixed h^*, the zero pressure occurs first for $x = 0$, then crossing the curve in the direction of increasing modulus changes the type of solution from one to two contact regions. This curve is called the 1-2 boundary in Fig. 2. If in this process of increasing μ^* for fixed h^* the zero in pressure occurs first for $x \neq 0$, then further increasing the modulus causes the solution to change from a one to a three contact region solution. This curve is called the 1-3 boundary. The 1-3 boundary occurs for thinner layers than the 1-2 boundary. The point at which the 1-2 and 1-3 boundary curves meet is also the intersection point with the 2-3 boundary curve. The 2-3 boundary is the almost horizontal curve which separates the multiple contact region zone into zones for the two and three contact region solutions.

The simplified solution under discussion here only displays solutions that correspond to the two and three contact region solutions. The curves defined by $\alpha = 0.929$ and $\alpha = 1.1872$ should be compared to the 2-3 boundary curve. Figure 2 shows that the $\alpha = 1.1872$ curve corresponds to the 2-3 boundary curve in the high-modulus zone. Thus, the transition from two to three contact regions in the exact elastic solution occurs at the onset of expansion of the central contact region in the approximate solution. The first contact of the beam with the indenter (at $x = 0$) in the approximate solution occurs before the exact solution predicts contact. Spreading of the contact in the approximate solution cannot occur until the normal ring load at that contact point has increased to a value that will allow a flat section of the beam to be supported with uniform normal pressure. Thus, there exists a range of values for the intermediate solution to the approximate problem. The approximate boundary curves extended into the low modulus zone, but, because the exact solution predicts complete contact in this zone, which the approximate solution cannot model, the approximate solution cannot be used in this zone.

Figure 3 presents the contact dimension d given by both solutions for $\mu^* = 15.0$. The dashed line is the approximate solution. The upper solid curve in this figure is the contact dimension c, while the lower is the contact dimension d. Agreement is reasonable over the range of layer thickness for which the exact solution results are presented. Because the approximate solution cannot have more than a point contact at the indenter corner, the size of the separation region found using the approximate solution will always be larger than the exact result. Thus, the approximate solution gives a bound on the size of the region over which layer effects dominate the behavior of the solution. The approximate solution does not have a quantity that corresponds to the contact region dimension c because the shell only contacts the indenter at its corner. Thus, for the approximate solution "c" is always equal to 1.

It has been shown that an approximation of a beam on a Winker foundation to an elastic layer bonded to an elastic half-space can predict some aspects of the exact solution. The transition from the 2 to 3 contact region solutions is reasonably determined. However, because the approximate solution does not admit a one contact region solution nor spreading of the contact regions at the corners of the indenter, many features of the solution are lost. The range of validity of the approximate solution is not a simple region in the (h^*, μ^*) parameter space. The approximate solution is certainly not valid for low modulus ratios where a one contact region solution is the correct configuration. In the high modulus zone of the parameter space, the approximate solution is only valid in zones where the exact solution has c almost equal to one. Also, in the area of the parameter space with h^* just greater than its value on the 2-3 boundary, the approximate solution predicts contact at the center of the indenter which does not exist in the exact solution.

Acknowledgment

I would like to thank Prof. D. B. Bogy for his helpful comments and Prof. B. Paul for bringing two of the references to my attention.

References

Becker, S. J., 1962, discussion of Paul (1962), *ASME Journal of Engineering for Industry*, Vol. 84, pp. 341–342.

Bogy, D. B., 1974, "The Shrink-Fit of a Cylindrical Shell Onto a Discontinuous Rigid Mandrel," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, pp. 959–962.

Paul, B., 1962, "Shrink Fits of Moderately Long Bands on Thin-Walled Cylinders," ASME *Journal of Engineering for Industry*, Vol. 84, pp. 338–341.

Shield, T. W., and Bogy, D. B., 1989, "Multiple Region Contact Solutions for a Flat Indenter on a Layered Elastic Half Space: Plane-Strain Case," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 251–262.

Influences of Large Amplitudes, Transverse Shear Deformation, and Rotatory Inertia on Free Vibrations of Moderately Thick Polygonal Plates: A New Approach

Rekha Battacharya[23] and B. Banerjee[24]

Introduction

For the nonlinear analysis of moderately thick isotropic plates Bhattacharya and Banerjee suggested a modified strain-energy expression and a new set of differential equations have been obtained in a decoupled form. The accuracy of these equations has been tested for square (Bhattacharya and Banerjee, 1989b) and circular (Bhattcharya and Banerjee, 1989a) plates with immovable as well as movable edge conditions. Results obtained are in excellent agreement with other known results.

The aim of the present paper is to extend this modified approach (Bhattcharya and Banerjee, 1989a,b) to study the influences of large amplitudes, transverse shear deformation, and rotatory inertia on free vibrations of moderately thick regular polygonal plates with movable, as well as immovable, edge conditions.

Formulation of the Differential Equation

Let us consider the free vibrations of thick polygonal plates of thickness, h. Using Reissner's variational theorem and Banerjee's hypothesis, a set of decoupled differential equations in rectangular cartesian coordinates governing the vibrations of thick plates have been derived (Bhattcharya and Banerjee, 1989b).

In a complex coordinate system $Z=x+iy$, $\overline{Z}=x-iy$. These equations in Bhattcharya and Banerjee (1989b) change. Let

$$Z=f(\xi) \tag{1}$$

be the analytic function which maps the given shape in the Z-plane onto a unit circle in the ξ-plane. Substituting the relation (1) into the transformed equations in (Z, $\overline{Z}$) coordinates, the following set of differential equations in (ξ, $\overline{\xi}$) coordinates have been obtained:

$$\begin{aligned}
16[A_1]+\frac{2}{5(1-\nu^2)}\cdot k_1\left(\frac{E}{Gc}\right)\cdot k_2\overline{\alpha}^2h^2\tau^2(t)[(1-\nu)(A_2+\overline{A}_2)\\
+2(1+\nu)A_1]+\frac{96k_3\lambda}{5(1-\nu^2)}k_1\left(\frac{E}{Gc}\right)[4(A_3+A_4)\\
+(A_5+\overline{A}_5)+6(A_6+\overline{A}_6)+10(A_7+\overline{A}_7)+8A_8+2(A_9+\overline{A}_9)]\\
-\frac{24}{5}\frac{\rho}{Gc}\ddot{\tau}(t)[A_{10}]-k_2\overline{\alpha}^2\tau^2(t)[(1-\nu)(A_{11}+\overline{A}_{11})\\
+2(1+\nu)A_{10}]-\frac{48}{h^2}k_3\lambda[4A_{12}+(A_{13}+\overline{A}_{13})\\
-(A_{14}+\overline{A}_{14})]+\frac{12}{h^2cp^2}\ddot{\tau}(t)A_{15}=0
\end{aligned} \tag{2}$$

where $\overline{\alpha}^2$ is obtained from the following relations,

$$\begin{aligned}
\frac{\overline{\alpha}^2h^2}{12}\left\{\left(\frac{dz}{d\xi}\right)\left(\frac{d\overline{z}}{d\overline{\xi}}\right)\right\}^2\tau^2(t)=\frac{1}{2}(1-\nu)\left\{\left(\frac{\partial\omega}{\partial\xi}\right)^2\left(\frac{d\overline{z}}{d\overline{\xi}}\right)^2\right.\\
\left.+\left(\frac{\partial\omega}{\partial\overline{\xi}}\right)^2\left(\frac{dz}{d\xi}\right)^2\right\}+\frac{\partial u}{\partial\xi}\cdot\frac{dz}{d\xi}\left(\frac{d\overline{z}}{d\overline{\xi}}\right)^2+\frac{du}{d\xi}\cdot\frac{d\overline{2}}{d\overline{\xi}}\cdot\left(\frac{dz}{d\xi}\right)^2\\
+\nu i\left\{\frac{\partial v}{\partial\xi}\cdot\frac{dz}{d\xi}\cdot\left(\frac{d\overline{z}}{d\overline{\xi}}\right)^2-\frac{\partial\overline{\xi}}{\partial\xi}\cdot\frac{\partial\overline{z}}{\partial\overline{\xi}}\left(\frac{dz}{d\xi}\right)^2\right\}\\
+(1+\nu)\cdot\frac{\partial\omega}{\partial\xi}\cdot\frac{\partial\omega}{\partial\overline{\xi}}\cdot\frac{dz}{d\xi}\cdot\frac{d\overline{z}}{d\overline{\xi}}.
\end{aligned} \tag{3}$$

Table 1 Mapping Function Coefficients (Loura and Sahady, 1969)

Polygons	L	λ_1
Square	$1.08a$	$-0.11a$
Pentagon	$1.053a$	$-0.07a$
Hexagon	$1.038a$	$-0.05a$
Heptagon	$1.029a$	$-0.036a$
Octagon	$1.022a$	$-0.028a$

Table 2 Linear Time Period

$$T_L^* \text{ (Thick Plate)}=\frac{2\pi}{\sqrt{\alpha}},\ \left(\frac{E}{Gc}\neq 0\right),$$

$$T_L \text{ (Thin Plate)}=\frac{2\pi}{\sqrt{\alpha}},\ \left(\frac{E}{Gc}=0\right)$$

Polygons	$T_L^*\left[\frac{h}{a}=0.2, \frac{E}{Gc}=2.5\right]$	$\frac{T_L^*}{T_L}$
Square	1.5613	1.0261
Pentagon	1.2121	1.0285
Hexagon	1.1185	1.0296
Heptagon	1.0722	1.0303
Octagon	1.0469	1.0308

[23] Lecturer in Physics, P. D. Women's College, Jalpaiguri, West Bengal, India.

[24] A.D.P.I., Jalpaiguri Division, Government of West Bengal, West Bengal, India.

Manuscript received by the ASME Applied Mechanics Division, May 5, 1990; final revision, Sept. 11, 1990.

space can predict some aspects of the exact solution. The transition from the 2 to 3 contact region solutions is reasonably determined. However, because the approximate solution does not admit a one contact region solution nor spreading of the contact regions at the corners of the indenter, many features of the solution are lost. The range of validity of the approximate solution is not a simple region in the (h^*, μ^*) parameter space. The approximate solution is certainly not valid for low modulus ratios where a one contact region solution is the correct configuration. In the high modulus zone of the parameter space, the approximate solution is only valid in zones where the exact solution has c almost equal to one. Also, in the area of the parameter space with h^* just greater than its value on the 2-3 boundary, the approximate solution predicts contact at the center of the indenter which does not exist in the exact solution.

Acknowledgment

I would like to thank Prof. D. B. Bogy for his helpful comments and Prof. B. Paul for bringing two of the references to my attention.

References

Becker, S. J., 1962, discussion of Paul (1962), *ASME Journal of Engineering for Industry*, Vol. 84, pp. 341–342.

Bogy, D. B., 1974, "The Shrink-Fit of a Cylindrical Shell Onto a Discontinuous Rigid Mandrel," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, pp. 959–962.

Paul, B., 1962, "Shrink Fits of Moderately Long Bands on Thin-Walled Cylinders," ASME *Journal of Engineering for Industry*, Vol. 84, pp. 338–341.

Shield, T. W., and Bogy, D. B., 1989, "Multiple Region Contact Solutions for a Flat Indenter on a Layered Elastic Half Space: Plane-Strain Case," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 251–262.

Influences of Large Amplitudes, Transverse Shear Deformation, and Rotatory Inertia on Free Vibrations of Moderately Thick Polygonal Plates: A New Approach

Rekha Battacharya[23] **and B. Banerjee**[24]

Introduction

For the nonlinear analysis of moderately thick isotropic plates Bhattacharya and Banerjee suggested a modified strain-energy expression and a new set of differential equations have been obtained in a decoupled form. The accuracy of these equations has been tested for square (Bhattacharya and Banerjee, 1989b) and circular (Bhattcharya and Banerjee, 1989a) plates with immovable as well as movable edge conditions. Results obtained are in excellent agreement with other known results.

The aim of the present paper is to extend this modified approach (Bhattcharya and Banerjee, 1989a,b) to study the influences of large amplitudes, transverse shear deformation, and rotatory inertia on free vibrations of moderately thick regular polygonal plates with movable, as well as immovable, edge conditions.

Formulation of the Differential Equation

Let us consider the free vibrations of thick polygonal plates of thickness, h. Using Reissner's variational theorem and Banerjee's hypothesis, a set of decoupled differential equations in rectangular cartesian coordinates governing the vibrations of thick plates have been derived (Bhattcharya and Banerjee, 1989b).

In a complex coordinate system $Z=x+iy$, $\overline{Z}=x-iy$. These equations in Bhattcharya and Banerjee (1989b) change. Let

$$Z=f(\xi) \tag{1}$$

be the analytic function which maps the given shape in the Z-plane onto a unit circle in the ξ-plane. Substituting the relation (1) into the transformed equations in (Z, $\overline{Z}$) coordinates, the following set of differential equations in (ξ, $\overline{\xi}$) coordinates have been obtained:

$$\begin{aligned}16[A_1]+\frac{2}{5(1-\nu^2)}\cdot k_1\left(\frac{E}{Gc}\right)\cdot k_2\overline{\alpha}^2h^2\tau^2(t)[(1-\nu)(A_2+\overline{A}_2)\\ +2(1+\nu)A_1]+\frac{96k_3\lambda}{5(1-\nu^2)}k_1\left(\frac{E}{Gc}\right)[4(A_3+A_4)\\ +(A_5+\overline{A}_5)+6(A_6+\overline{A}_6)+10(A_7+\overline{A}_7)+8A_8+2(A_9+\overline{A}_9)]\\ -\frac{24}{5}\frac{\rho}{Gc}\ddot{\tau}(t)[A_{10}]-k_2\overline{\alpha}^2\tau^2(t)[(1-\nu)(A_{11}+\overline{A}_{11})\\ +2(1+\nu)A_{10}]-\frac{48}{h^2}k_3\lambda[4A_{12}+(A_{13}+\overline{A}_{13})\\ -(A_{14}+\overline{A}_{14})]+\frac{12}{h^2cp^2}\ddot{\tau}(t)A_{15}=0\end{aligned} \tag{2}$$

where $\overline{\alpha}^2$ is obtained from the following relations,

$$\begin{aligned}\frac{\overline{\alpha}^2h^2}{12}\left\{\left(\frac{dz}{d\xi}\right)\left(\frac{d\overline{z}}{d\overline{\xi}}\right)\right\}^2\tau^2(t)=\frac{1}{2}(1-\nu)\left\{\left(\frac{\partial\omega}{\partial\xi}\right)^2\left(\frac{d\overline{z}}{d\overline{\xi}}\right)^2\right.\\ \left.+\left(\frac{\partial\omega}{\partial\overline{\xi}}\right)^2\left(\frac{dz}{d\xi}\right)^2\right\}+\frac{\partial u}{\partial\xi}\cdot\frac{dz}{d\xi}\left(\frac{d\overline{z}}{d\overline{\xi}}\right)^2+\frac{du}{d\xi}\cdot\frac{d\overline{z}}{d\overline{\xi}}\cdot\left(\frac{dz}{d\xi}\right)^2\\ +\nu i\left\{\frac{\partial v}{\partial\xi}\cdot\frac{dz}{d\xi}\cdot\left(\frac{d\overline{z}}{d\overline{\xi}}\right)^2-\frac{\partial\overline{\xi}}{\partial\xi}\cdot\frac{\partial\overline{z}}{\partial\overline{\xi}}\left(\frac{dz}{d\xi}\right)^2\right\}\\ +(1+\nu)\cdot\frac{\partial\omega}{\partial\xi}\cdot\frac{\partial\omega}{\partial\overline{\xi}}\cdot\frac{dz}{d\xi}\cdot\frac{d\overline{z}}{d\overline{\xi}}.\end{aligned} \tag{3}$$

Table 1 Mapping Function Coefficients (Loura and Sahady, 1969)

Polygons	L	λ_1
Square	$1.08a$	$-0.11a$
Pentagon	$1.053a$	$-0.07a$
Hexagon	$1.038a$	$-0.05a$
Heptagon	$1.029a$	$-0.036a$
Octagon	$1.022a$	$-0.028a$

Table 2 Linear Time Period

$$T_L^* \text{ (Thick Plate)}=\frac{2\pi}{\sqrt{\alpha}},\ \left(\frac{E}{Gc}\neq 0\right),$$

$$T_L \text{ (Thin Plate)}=\frac{2\pi}{\sqrt{\alpha}},\ \left(\frac{E}{Gc}=0\right)$$

Polygons	$T_L^*\left[\frac{h}{a}=0.2, \frac{E}{Gc}=2.5\right]$	$\frac{T_L^*}{T_L}$
Square	1.5613	1.0261
Pentagon	1.2121	1.0285
Hexagon	1.1185	1.0296
Heptagon	1.0722	1.0303
Octagon	1.0469	1.0308

[23] Lecturer in Physics, P. D. Women's College, Jalpaiguri, West Bengal, India.

[24] A.D.P.I., Jalpaiguri Division, Government of West Bengal, West Bengal, India.

Manuscript received by the ASME Applied Mechanics Division, May 5, 1990; final revision, Sept. 11, 1990.

Table 3 Ratio of nonlinear to linear period for the vibration of simply-supported polygonal plates (square of side $2a$)

$\bar{\beta}=\frac{A_o}{h}$	$\frac{T^*}{T}$ for immovable edges $\cdot$ [$\nu=0.3$, $\lambda=\nu^2$, Banerjee and Datta, 1981), $\frac{h}{2a}=\frac{1}{10}$.									
	Thin Plate									
	$k_1\left(\frac{E}{Gc}\right)=0$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=2.5$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=20$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=30$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=50$	(Bhattacharya and Banerjee, 1989b)
0.6	0.9143	0.9072	0.9346	0.9270	1.0582	1.0469	1.1177	1.1066	1.2191	1.2012
0.8	0.8613	0.8507	0.8784	0.8624	0.9797	0.9636	1.0268	1.0113	1.1046	1.0819
1.00	0.8050	0.7917	0.8191	0.8055	0.9006	0.8809	0.9372	0.9123	0.9959	0.9678

Table 3 Continued

$\bar{\beta}=\frac{A_o}{h}$	$\frac{T^*}{T}$ for movable edges [$\nu=0.3$, $\lambda=\nu^2$, Banerjee and Datta, 1981), $\frac{h}{2a}=\frac{1}{10}$.									
	Thin Plate									
	$k_1\left(\frac{E}{Gc}\right)=0$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=2.5$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=20$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=30$	(Bhattacharya and Banerjee, 1989b)	$k_1\left(\frac{E}{Gc}\right)=50$	(Bhattacharya and Banerjee, 1989b)
0.6	0.9850	0.9779	1.0105	1.0029	1.1717	1.1604	1.2505	1.2394	1.3981	1.3802
0.8	0.9722	0.9616	1.0017	0.9857	1.1498	1.1337	1.2226	1.2071	1.3589	1.3362
1.00	0.9550	0.9416	0.9782	0.9647	1.1213	1.1022	1.1942	1.1693	1.3133	1.2852

Table 4 Ratio of nonlinear to linear period of vibration of different polygons

	$\bar{\beta}=\frac{A_o}{h}$	$\frac{T^*}{T}\left(\nu=0.3,\ \lambda=\nu^2,\text{ (Banerjee and Datta, 1981)},\ \frac{h}{2a}=\frac{1}{10}\right)$									
		Pentagon					Hexagon				
		$k_1\left(\frac{E}{Gc}\right)$					$k_1\left(\frac{E}{Gc}\right)$				
		0	2.5	20	30	50	0	2.5	20	30	50
Immovable edges	0.6	0.9341	0.9546	1.0794	1.1397	1.2391	0.9534	0.9747	1.0994	1.1607	1.2591
	0.8	0.8814	0.8995	1.0007	1.0478	1.1267	0.9016	0.9207	1.0207	1.0691	1.1467
	1.0	0.8255	0.8411	0.9209	0.9577	1.0170	0.8453	0.8621	0.9429	0.9777	1.0381
Movable edges	0.6	1.0050	1.0305	1.1917	1.2708	1.4182	1.0250	1.0515	1.2116	1.2909	1.4383
	0.8	0.9922	1.0227	1.1699	1.2436	1.3784	1.0124	1.0427	1.1900	1.2636	1.3989
	1.0	0.9762	0.9923	1.1416	1.2145	1.3336	0.9962	1.0126	1.1619	1.2346	1.3539

Table 4 Continued

	$\bar{\beta}=\frac{A_o}{h}$	($2a$ is a dimension in length and related to the side of each polygon.)									
		Heptagon					Octagon				
		$k_1\left(\frac{E}{Gc}\right)$					$k_1\left(\frac{E}{Gc}\right)$				
		0	2.5	20	30	50	0	2.5	20	30	50
Immovable edges	0.6	0.9784	0.9994	1.1204	1.1808	1.2791	0.9954	1.0194	1.1414	1.2008	1.3001
	0.8	0.9201	0.9417	1.0417	1.0891	1.1677	0.9429	0.9627	1.0627	1.1101	1.1887
	1.0	0.8632	0.8821	0.9640	0.9987	1.0594	0.8830	0.9031	0.9843	1.0197	1.0804
Movable edges	0.6	1.0451	1.0716	1.2317	1.3110	1.4584	1.0652	1.0917	1.2518	1.3311	1.4785
	0.8	1.0352	1.0627	1.2103	1.2837	1.4190	1.0567	1.0827	1.2303	1.3041	1.4391
	1.0	1.0177	1.0329	1.1822	1.2549	1.3742	1.0391	1.0532	1.2025	1.2752	1.3945

Here, the A is defined in the Appendix and $\bar{A}$ is essentially the same expression having ξ and $\bar{\xi}$ interchanged. k_1 is a tracing constant identifying the effects of transverse shear deformation, while k_2, k_3 are tracing constants identifying nonlinear vibrations. If $k_i=1$ or 0, effects are included or not, respectively. It is to be noted that the effects of rotatory inertia have been neglected because these are considered to be small compared with the effects due to transverse shear deformation as the plate is undergoing flexural vibrations (Wu and Vinson, 1969).

For regular polygons the mapping function is

$$Z=L\xi+\lambda_1\xi^5 \tag{4}$$

where values of L and λ_1 are given in Table 1.

Let us choose the deflection function in the following form:

$$\omega=A_0\tau(t)\left[1-\xi\bar{\xi}\right]\left[1-\frac{1}{3}\xi\bar{\xi}+\frac{1}{2}(\xi^2+\bar{\xi}^2)(1-\xi\bar{\xi})^2\right]. \tag{5}$$

Clearly, ω is θ dependent and satisfies the simply-supported edge conditions, namely, $\omega=0$ and $\frac{\partial^2\omega}{\partial\xi\partial\bar{\xi}}=0$ at $r=1$. Substituting equation (4) and (5) in (2) the error function $\epsilon(\xi, \bar{\xi}, t)$ is obtained. Galarkin's technique requires

$$\int_{\theta=0}^{2\pi} \int_{\nu=0}^{1} \epsilon(\xi, \bar{\xi}, t)\,\omega(\xi, \bar{\xi}, t)\, r\, dr\, d\theta = 0. \tag{6}$$

The orthogonalization of the error function to the spatial function is obvious because of the presence of the sine and cosine functions in the $(\xi, \bar{\xi})$ coordinates.

The constant $\bar{\alpha}$ is determined by putting (5) in (3), using (4), and integrating over the area of the plate.

For the movable edge condition,

$$\bar{\alpha} = 0. \tag{7}$$

Now, for transverse vibrations, the normal displacement $\omega(\xi, \bar{\xi}, t)$ is our primary interest. So, the in-plane displacements u and v have been eliminated through integration by choosing suitable expressions for them compatible with their boundary values namely $u=0$, $v=0$, for immovable edges.

Evaluating the integrals in (6) and considering the values of $\bar{\alpha}$ obtained from (3) after integrating over the area of the plate, one obtains the differential equation with a cubic nonlinearity in the following form

$$\ddot{\tau}(t) + \alpha\tau(t) + \beta\tau^3(t) = 0. \tag{8}$$

The solution of the above equation subject to the boundary conditions

$$\tau(0) = 1$$
$$\bar{\tau}(0) = 0$$

is well known and is obtained in terms of Jacobi's elliptic function. The ratio of the nonlinear and linear time period is

$$\frac{T^*}{T} = \frac{\dfrac{2K}{\pi}}{\left[1 + \dfrac{\beta}{\alpha}\bullet\bar{\beta}^2\right]^{1/2}} \tag{9}$$

where $\bar{\beta} = \dfrac{A_o}{h}$.

Numerical Results

Numerical results are presented here in the tabular forms for both movable as well as immovable edges, for moderately thick, regular polygonal plates.

Discussion

1 The tables show that the new approach given in (Bhattacharya and Banerjee, 1989b) can be conveniently extended to study the dynamic behavior of moderately thick polygonal plates and the results obtained thereby are sufficiently accurate both for movable and immovable edge conditions. If the mapping functions are known, the nonlinear behaviors of thick plates of any shape can be investigated by using the proposed differential equations. This is an advantage of the present study.

2 The results for both movable and immovable edge conditions have been obtained from the same differential equation. This is an additional advantage.

References

Banerjee, B., and Datta, S., 1981, "A New Approach to the Analysis of Large Deflections of Thin Plates," *Int. J. Non-linear Mech.*, Vol. 16, pp. 47–52.

Bhattacharya, R., and Banerjee, B., 1989a, "Influences of Large Amplitudes, Shear Deformation, and Rotatory Inertia on Axisymmetric Vibrations of Moderately Thick Circular Plates: A New Approach," *J. Sound and Vibration*, Vol. 133, No. 1, pp. 185–188.

Bhattacharya, R., and Banerjee, B., 1989b, "Influences of Large Amplitudes, Transverse Shear Deformation, and Rotatory Inertia on Free Lateral Vibrations of Transversely Isotropic Plates: A New Approach," *Int. J. Non-linear Mechanics*, Vol. 24, No. 3, pp. 159–164.

Loura, P. A., and Sahady, P., 1969, *Journal of Engineering Mechanics*, ASCE, New York, Feb.

Wu, C. I., and Vinson, J. R., 1969, "Influences of Large Amplitudes, Transverse Shear Deformation, and Rotatory Inertia on Lateral Vibrations of Transversely Isotropic Plates," ASME Journal of Applied Mechanics, Vol. 36, pp. 254–260.

Appendix

E = Young's modulus
ν = Poisson's ratio
Gc = shear modulus
ρ = mass density
λ = material constant
Ao = amplitude of oscillations
$D = \dfrac{Eh^3}{12(1-\nu^2)}$ = flexural rigidity
α = coupling parameter
u, v, w = displacements in x, y, and z-directions, respectively
$\tau(t)$ = nonlinear time-dependent function
$C_p = \left[\dfrac{E}{\rho(1-\nu^2)}\right]^{1/2}$ = speed of wave propagation along the surface of the plate

$$\xi = r.e^{i\theta},\ \bar{\xi} = r.e^{-i\theta}$$

$$Z_1 = \frac{dz}{d\xi},\ Z_{\bar{1}} = \frac{d\bar{z}}{d\bar{\xi}},\ Z_2 = \frac{d^2z}{d\xi^2},\ Z_{\bar{2}} = \frac{d^2\bar{z}}{d\bar{\xi}},\ \text{etc.}$$

$$W_{,1} = \frac{\partial\omega}{\partial\xi},\ W_{,1\bar{1}} = \frac{\partial^2\omega}{\partial\xi\partial\bar{\xi}},\ W_{,2\bar{2}} = \frac{\partial^4\omega}{\partial\xi^2\delta\bar{\xi}^2},\ \text{etc.}$$

$$A_1 = W_{,2\bar{2}}(Z_1)^3(Z_{\bar{1}})^3 - W_{,2\bar{1}}Z_{\bar{2}}(Z_{\bar{1}})^2(Z_1)^3 - W_{,\bar{2}1}Z_2(Z_1)^2(Z_{\bar{1}})^3 + W_{,1\bar{1}}Z_2Z_{\bar{2}}(Z_1)^2(Z_{\bar{1}})^2$$

$$A_2 = W_{,3\bar{1}}(Z_{\bar{1}})^4(Z_1)^2 - 3W_{,2\bar{1}}Z_2(Z_{\bar{1}})^4Z_1 + 3W_{,1\bar{1}}(Z_2)^2(Z_{\bar{1}})^4 - W_{,1\bar{1}}Z_3(Z_{\bar{1}})^4Z_1$$

$$A_3 = W_{,1}W_{\bar{1}}[W_{,2\bar{2}}(Z_{\bar{1}})^2(Z_1)^2 - W_{,2\bar{1}}Z_{\bar{2}}Z_{\bar{1}}(Z_1)^2 - W_{,1\bar{2}}Z_2Z_1(Z_{\bar{1}})^2 + W_{,1\bar{1}}Z_2Z_{\bar{2}}Z_1Z_{\bar{1}}$$

$$A_4 = (W_{,1\bar{1}})^3(Z_1)^2(Z_{\bar{1}})^2$$

$$A_5 = (W_{,1})^2(Z_1)^2[W_{,\bar{3}1}(Z_{\bar{1}})^2 - 3W_{,\bar{2}1}Z_{\bar{2}}Z_{\bar{1}} + 3W_{,1\bar{1}}(Z_{\bar{2}})^2 - W_{,1\bar{1}}Z_{\bar{3}}Z_{\bar{1}}]$$

$$A_6 = W_{,\bar{1}}[W_{,\bar{2}1}W_{,2}(Z_1)^2(Z_{\bar{1}})^2 - W_{,1\bar{1}}W_{,2}Z_{\bar{2}}(Z_1)^2Z_{\bar{1}} - W_{,\bar{2}1}W_{,1}Z_2(Z_{\bar{1}})^2Z_1 + W_{,1\bar{1}}W_{,1}Z_2Z_{\bar{2}}Z_1Z_{\bar{1}}]$$

$$A_7 = W_{,\bar{1}}(Z_{\bar{1}})^2[W_{,2\bar{1}}W_{,1\bar{1}}(Z_1)^2 - (W_{,1\bar{1}})^2Z_2Z_1]$$

$$A_8 = W_{,1\bar{1}}[W_{,2}W_{,\bar{2}}(Z_1)^2(Z_{\bar{1}})^2 - W_{,\bar{2}}W_{,1}Z_2Z_1(Z_{\bar{1}})^2 - W_{,2}W_{,\bar{1}}Z_{\bar{2}}(Z_1)^2Z_{\bar{1}} + W_{,1}W_{,\bar{1}}Z_{\bar{2}}Z_2Z_1Z_{\bar{1}}]$$

$$A_9 = W_{,2}W_{,1}(Z_1)^2[W_{,\bar{3}}(Z_{\bar{1}})^2 - 3W_{,\bar{2}}Z_{\bar{2}}Z_{\bar{1}} + 3W_{,\bar{1}}(Z_{\bar{2}})^2 - W_{,\bar{1}}Z_{\bar{3}}Z_{\bar{1}}] + (W_{,1})^2Z_2Z_1[-W_{,\bar{3}}(Z_{\bar{1}})^2 + 3W_{,\bar{2}}Z_{\bar{2}}Z_{\bar{1}} - 3W_{,\bar{1}}(Z_{\bar{2}})^2 + W_{,\bar{1}}Z_{\bar{3}}Z_{\bar{1}}]$$

$$A_{10} = W_{,1\bar{1}}(Z_1)^4(Z_{\bar{1}})^4$$

$$A_{11} = (Z_{\bar{1}})^5[W_{,2}(Z_1)^3 - W_{,1}Z_2(Z_1)^2]$$

$$A_{12} = W_{,1\bar{1}}W_{,1}W_{\bar{1}}(Z_1)^3(Z_{\bar{1}})^3$$

$$A_{13} = W_{,\bar{2}}(W_{,1})^2(Z_{\bar{1}})^3(Z_1)^3$$

$$A_{14} = W_{,\bar{1}}(W_{,1})^2(Z_{\bar{2}})(Z_{\bar{1}})^2(Z_1)^3$$

$$A_{15} = W(Z_1)^5(Z_{\bar{1}})^5$$

Journal of Applied Mechanics

Book Reviews

Boundary Elements: An Introductory Course, by C. A. Brebbia and J. Dominguez. McGraw-Hill, New York, 1989. 293 pages.

REVIEWED BY J. L. TASSOULAS[1]

During the last 15 years, the boundary integral equation method, also known widely as the "boundary element method" (hereafter referred to as the "method") has seen increasing use in a variety of engineering problems. The present book is an introduction to the subject. Both C. A. Brebbia and J. Dominguez have been active in teaching and research on the method.

In a brief introduction, the authors enumerate the main advantages of the method over the (more) popular finite element method. The (biased) comparison of the two methods points out that mesh generation is easier when using boundary elements, demonstrates, by means of an example, the superior performance of the method in problems involving stress concentrations and claims that the use of boundary elements is the only practical approach towards the solution of problems posed on infinite domains. Concluding the introduction, it is stated that the objective of the book is "to provide a simple and up-to-date introduction to the method" so as to help increase its popularity among engineers. There is also a suggestion that the method be taught at both the undergraduate and graduate levels. The book is, however, intended for use in a first course on the method.

Chapter 1 introduces the method as a weighted-residual technique. The discussion is based on boundary value problems in one dimension. Other weighted-residual techniques are outlined as well. Also, the Poisson equation in two dimensions is processed as a weighted residual so as to establish the analogy between one and more than one dimensions.

The formulation of boundary elements for problems governed by the Laplace and Poisson equations is presented in detail in Chapter 2. Constant, linear quadratic, and higher-order elements are described for two-dimensional problems. A number of computer programs are included so as to demonstrate the implementation of the elements. The use of "discontinuous" elements is suggested to overcome difficulties that may arise at boundary corners. This is followed by a quick look at boundary elements for three-dimensional problems, while the use of boundary element subregions and the formulation of axisymmetric boundary elements are discussed briefly.

Boundary elements for elastostatics are considered in Chapters 3 and 4. Two computer programs are supplied: one with constant elements and another with quadratic elements, both for two-dimensional problems. Examples of use of the programs are included while other problems to which the programs can be applied are suggested as exercises.

Finally, in Chapter 5, other topics are covered rather superficially: coupling of boundary elements with finite elements, boundary elements for fracture mechanics, and the use of the method in steady-state elastodynamics.

References are given after each chapter and in Appendix C. The interested reader will be able to locate the rest of the literature on the method through these references.

The book appears to fulfill its promise of a "simple and up-to-date introduction to the method." Perhaps, missing from the book are even brief discussions of the use of the method in other types of problems of engineering interest; e.g., eigenvalue problems, transient dynamics problems, and continuum mechanics problems involving various nonlinearities. The reader may think that the absence of some of these topics from the book suggests that the method is not particularly suitable for such problems. In any case, for the purposes of an introductory course on the method, the book is worth consideration.

[1]Department of Civil Engineering, The Univ. of Texas, Austin, Texas 78712-1076.

The Fokker-Planck Equation: Methods of Solution and Application, 2nd ed., by H. Risken. Springer-Verlag, New York.

REVIEWED BY T. K. CAUGHEY[2]

This is the second edition (in paperback) of the author's excellent 1984 book on the Fokker-Planck Equation, applications, and methods of solution. With the exception of the correction of some misprints in the first edition and the addition of a short review of recent developments, the book is essentially unchanged from the first edition. Professor Risken has made substantial contributions to the application and solution of the Fokker-Planck Equation in laser physics, diffusion in periodic potentials, and other noise-related areas, and has written an excellent survey of such methods. The first edition has been very well received, and the new paperback edition should reach an even wider audience.

[2]Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91109.