Transactions of the ASME®

Technical Editor, LEON M. KEER

The Technological Institute
Northwestern University
Evanston, IL 60201

APPLIED MECHANICS DIVISION

Chairman, THOMAS L. GEERS
Secretary, S. LEIBOVICH
Associate Editors,
D. B. BOGY
R. M. CHRISTENSEN
R. J. CLIFTON
R. L. HUSTON
W. D. IWAN
L. M. KEER
W. G. KNAUSS
F. A. LECKIE
J. T. C. LIU
R. M. McMEEKING
FRANCIS C. MOON
A. K. NOOR
J. W. NUNZIATO
R. H. PLAUT
J. G. SIMMONDS
K. R. SREENIVASAN
Z. WARHAFT
L. T. WHEELER

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

Members-at-Large
J. T. COKONIS
M. FRANKE
M. KUTZ
F. LANDIS
J. R. LLOYD
T. C. MIN
R. E. NICKELL
R. E. REDER
R. ROCKE
F. W. SCHMIDT
W. O. WINER

President, R. ROSENBERG
Exec. Dir.
D. L. BELDEN
Treasurer, ROBERT A. BENNETT

PUBLISHING STAFF
Mng. Dir., Publ., JOS. SANSONE
Managing Editor,
CORNELIA MONAHAN
Production Editor, REMO SALTA
Prod. Asst., MARISOL ANDINO

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) for $115 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Second class postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to The Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.
CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.

PRICES: To members, $27.00, annually; to nonmembers, $115.00.
Add $12.00 for postage to countries outside the United States and Canada.

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc.

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 55 • NUMBER 1 • MARCH 1988

TECHNICAL PAPERS

(Contents continued on Inside Back Cover)

CONTENTS (Continued)

BRIEF NOTES

DISCUSSION

BOOK REVIEWS

Representation of Mechanical Behavior in the Presence of Changing Internal Structure

E. T. Onat
Department of Mechanical Engineering,
Yale University,
New Haven, CT 06520

F. A. Leckie
Department of Theoretical and
Applied Mechanics,
University of Illinois at
Urbana-Champaign,
Urbana, IL 61801

The paper is concerned with the representation of the relationship that exists, for a given material and temperature and for small deformations, between histories of applied stress and the observed strain and the accompanying changes in internal structure of the material. Emphasis is given to creep damage in metals as a vehicle for illustration of the main ideas introduced in the paper. In particular, the role played by irreducible even rank tensors in the representation of internal structure is discussed and clarified. The restrictions placed by thermodynamics on constitutive equations are considered and the use of potentials in these equations is examined and criticized.

1 Introduction

Deformation of most materials of interest (metals, rocks, granular media, composites) is accompanied by changes in their internal structure. For instance, as a metal deforms at high temperature, these structural changes may take place: Holes may nucleate and grow within the grain or on the grain boundaries, precipitates may coarsen or dissolve, a dominant dislocation structure may be destroyed, or the microstructure may alter in a manner yet to be defined. All these changes are referred to as "creep damage." This damage usually accelerates the creep-rate, which in turn accelerates the damage rate until rupture occurs. The shape of the tertiary portion of the creep curve is a consequence of the coupling of the creep and damage rates.

One of the early contributions to the multiaxial stress creep of metals was made by Johnson et al. (1962) who observed that the rupture life of copper is dependent on the maximum tensile stress while that of a precipitate hardened aluminum is dependent on the effective shear stress. Optical observation revealed that in the case of copper voids had grown on the grain boundaries whose planes were perpendicular to the maximum stress direction. In the case of the aluminum alloy there was no evidence of voids, in spite of the fact that the creep curves had a pronounced tertiary portion which suggests evidence of some form of damage. This distinction in creep behavior can lead to significant differences in component properties. For example, Bridgeman notches in the aluminum alloy demonstrate strong notch-strengthening characteristics while for those in copper no such strengthening is evident (Leckie and Hayhurst, 1974).

The paper is concerned with the representation of the relationship that exists, for a given material and temperature and for small deformations, between histories of applied stress and the observed strain, and the accompanying internal happenings, such as the ones mentioned above.

We consider representations based on state variables. The presence of deformation induced anisotropy implies that the state variables must measure the relevant aspects of not only internal structure but also its orientation (Section 4). It is known that (irreducible) tensorial state variables constitute a natural tool for the representation of internal structure and its orientation. They also provide a convenient measure of the degree of anisotropy present in the material. In Section 5 we show that tensorial variables could arise from the statistics of internal structure.

Dissipative behavior exhibited by materials beyond their elastic range is the subject of Section 6 where we discuss briefly the thermodynamic basis of dissipativity and then consider the constraints imposed by dissipativity on constitutive equations.

We examine the use of potentials in the construction of constitutive equations. We exhibit examples of physically realizable dissipative materials that do not accept a representation based on potentials. The last section offers examples of constitutive equations appropriate to creep of metals. We observe that in some cases experimental evidence does not support the use of potentials in constitutive equations.

2 Representation of Mechanical Behavior

We consider now the task of providing a mathematical description or representation of the observed mechanical behavior and the accompanying changes in internal structure. For simplicity we confine the discussion to the case of small, isothermal deformations. Moreover, we assume that nonlocal effects are absent so that deformations may be taken to be homogeneous without loss of generality.

We assume that before the onset of deformation, the

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 1, 1986; final revision June 20, 1987.

material has remained at *constant temperature* θ_o and under *zero stress* for a sufficiently long time so that it is in thermal and mechanical equilibrium at time $t = 0$, zero time coinciding with the start of the loading process.

We let σ and ϵ denote, respectively, the Cauchy stress and the infinitesimal strain tensor. We assume that phenomenological evidence on mechanical behavior comes from tests where a time dependent stress $\sigma(t)$ is applied to a virgin specimen on the time interval $[0, T]$. Moreover, the testing method and the stress intensities used are such that the *shape changes* and *rotations* that the material experiences are small. Thus the response of the material is the (infinitesimal) strain $\epsilon(t)$ observed on $[0, T]$.

We are looking for an algorithm that will act on the stress histories of interest on $[0, T]$ and will produce, to within an acceptable error, the observed strain responses and the relevant aspects of the accompanying changes in internal structure.

Experience has shown that the representations based on the notion of state and on the differential equations that govern the evolution of state have definite advantages over other methods of representation (i.e., integral representations).

We imagine that the changing internal state of a deforming material element can be represented by the pair (S, σ). Here $\sigma(t)$ is the current stress carried by the material and $S(t)$ stands for the n parameters that measure those aspects of inelastic changes in internal structure that have taken place on $(0, t)$ and are relevant to future behavior of the material. Thus it is assumed that the future behavior of the material will depend (to within a given approximation) only on the values of $n + 6$ parameters at the time of interest and on the stresses applied beyond this time.

We shall be interested in materials that accept a representation of the following type for their mechanical behavior:

$$\frac{d\epsilon}{dt} = f(S(t), \sigma(t)) + D(S(t))\frac{d\sigma(t)}{dt} \tag{2.1a}$$

$$\frac{dS}{dt} = g(S(t), \sigma(t)) \tag{2.1b}$$

$$f{:}\Sigma \to T_2^s,\ D{:}T_2^s \to T_2^s,\ g{:}\Sigma \to R^n$$

Here Σ is the state set composed of all pairs (S, σ) encountered in all tests of interest and is a subset of $R^n \times T_2^s$ where T_2^s is the space of symmetric second rank tensors. The "size" of Σ depends on the class of stress histories employed in experiments. According to (2.1*a*), the current strain rate is composed of the rate of creep strain

$$\dot{\epsilon}_c = f(S, \sigma)$$

which depends upon the present internal state (S, σ) and of the elastic rate of strain

$$\dot{\epsilon}_e = D(S)\dot{\sigma}$$

where D is the current elastic compliance of the material that may depend upon the *inelastic state* S. The second equation (2.1*b*) constitutes the law of evolution for the inelastic state. The absence of $\dot{\sigma}$ on the right-hand side of (2.1*b*) merits attention and comment. This absence implies that a sudden increase in the stress level will leave S unchanged. This, however, would not be the case, for instance, if the sudden closure or extension of cracks plays an important role in the mechanical behavior. Nevertheless we shall be interested in materials for which (2.1*b*) holds.

Let S_o denote the initial internal state. Since the material is in thermo-mechanical equilibrium before the onset of stressing, S must retain its value S_o and no strains should develop in the absence of stress. Thus the constitutive functions f and g must have the property

$$f(S_o, 0) = 0,\ g(S_o, 0) = 0. \tag{2.2}$$

The representation (2.1) could be deduced, in principle, from the applied stress histories and the observed strain responses (c.f., Geary and Onat, 1974; Onat, 1966) and a precise meaning could be given to $S(t)$ in terms of an equivalence relation on stress histories on $[0, t]$.

Some of the parameters that comprise $S(t)$ will contain information on internal damage that occurred on the time interval $[0, t]$. We may assume, following and generalizing the ideas of Kachanov (1958), that failure due to accumulation of internal damage will occur when the trajectory of the state point (S, σ) intersects certain parts of the boundary $\partial\Sigma$ of the state set Σ.

In metals an important main component of S are the parameters q that provide information on the distribution of dislocations in the form of averages over the bulk of the solid. These parameters and the current stress σ have a controlling influence on the rate of creep, especially during the early stages of deformation. Internal damage has a more local character. It may be confined to grain boundaries in the form of voids, to the vicinity of precipitates as zones of decohesion or local cracking, to intense pileups of dislocations that act as nuclei to explosive void growth during the final stages of tertiary creep. One might need several damage parameters (d) to describe these events. During a creep experiment the evolution of q and d may not take place simultaneously. In some materials q may grow at first, and higher q values may then initiate the growth of various kinds of internal damage described by d.

Finally we note that the representation (2.1) is only a starting point for work on constitutive equations. For a given material and for a given set of purposes one needs to construct an explicit representation.

However, considerations of material symmetry and of thermodynamics endow equations (2.1) with additional structure without the need of being explicit about the material. The paper is devoted to the study of this additional structure.

3 Considerations of Material Symmetry. Tensorial State Variables

We assume that the material is isotropic in its stress free initial state $(S_o, 0)$. But as the material deforms it usually loses some or all of its symmetries. This might come about, for instance, by enhanced void formation on grain boundaries perpendicular to the direction of maximum tensile stress.

The n parameters associated with the internal state S must, therefore, measure the relevant aspects of the *internal structure* and its *orientation*. To be more explicit about this point let us consider two experiments conducted with identical virgin specimens (Fig. 1). Both specimens carry a constant stress with the same principal values, say $\sigma_1 > 0$, $\sigma_2 < 0$, $\sigma_3 = 0$ but in the experiment (*ii*) the σ_1 direction makes an angle of θ with the previous one. Microscopy upon the removal of stress may reveal a pattern of voids along the grain boundaries as shown in Fig. 1(*a*) and Fig. 1(*b*).

Since the material is initially isotropic the void patterns in two tests will (statistically) differ from each other only by the rotation θ. The question arises: If the parameters $S = (q_1, \ldots, q_n)$ are assigned to the stress free element a, then what parameters $S' = (q_1', \ldots, q_n')$ should be given to specimen b? Clearly the element a when rotated by an amount θ will be indistinguishable from element b from the point of view of future mechanical behavior. Nevertheless, the elements a and b will, in general, behave differently under identical future stress histories (because of the difference in the orientations of their identical internal structure) and hence S' will have to be different from S.

Mandel (1973) has argued that the orientation of internal structure should be measured by means of a triad of vectors attached to the material element. However, this idea has many

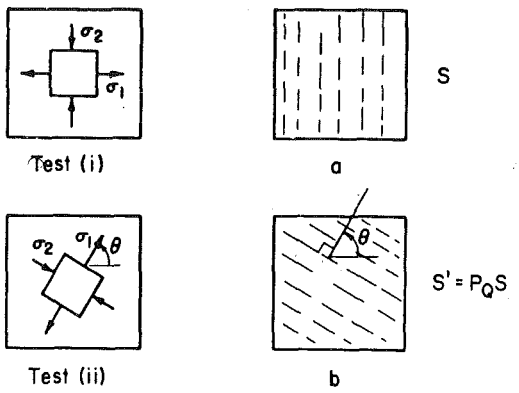


Fig. 1 State and orientation S of the material

drawbacks. Experience has shown that it is advantageous to consider the internal structure and its orientation as a single entity and to use tensorial state variables for the representation of this entity (cf Fardshisheh and Onat, 1973; Geary and Onat, 1974; Leckie and Onat, 1981; Murakami and Ohno, 1980; Chaboche, 1982).

We extend the ideas associated with Fig. 1 and make the following table which is concerned with behavior observed in pairs of closely related experiments:

Table 1

	Input	Output	Internal Structure
(*i*)	$\sigma(\tau)$	$\epsilon(\tau)$	$S(\tau)$
(*ii*)	$Q\sigma(\tau)Q^T$	$Q\epsilon(\tau)Q^T$	$S'=P_QS(\tau)$

The stress history in experiment (*ii*) differs from the one in (*i*) by a constant rotation Q of respective internal forces. The second column in Table 1 follows from isotropy of the virgin material.

Reflection will show that the internal state S' in (*ii*) will depend only on S and on the rotation $Q \epsilon 0^+(3)$ where $0^+(3)$ is the group of rigid body rotations:

$$S' = \hat{f}(S, Q) \tag{3.1}$$

For a fixed Q this function defines a transformation on the space $\hat{\Sigma}$ of distinct stress-free states. We emphasize this by rewriting equation (3.1) as

$$S' = P_Q S \quad P_Q : \hat{\Sigma} \to \hat{\Sigma} \tag{3.2}$$

It is then easily seen that the set G of all such transformations as Q ranges over $0^+(3)$

$$G = \{P_Q : Q \epsilon 0^+(3)\}$$

constitute a *group* of transformations on $\hat{\Sigma}$.

This observation and the fact that σ and ϵ are even rank, together with the assumption that $\hat{\Sigma}$ is a compact manifold imply (cf Geary and Onat, 1974) that S may be thought of as composed of *irreducible even* rank tensors

$$S = (s_1, \ldots, s_m) \tag{3.3}$$

and that

$$P_Q S = (p_Q s_1, \ldots, p_Q s_m)$$

where p_Q are linear tensor transformations appropriate to the rank of tensor on which they operate. For instance, if α is a second rank tensor, then in components

$$(p_Q\alpha)_{ij} = Q_{ip}Q_{jp}\alpha_{pq} \quad i, j, p, q = 1,2,3$$

where repeated indices imply summation.

It may be useful to comment briefly on irreducible tensors that constitute building blocks for the representation of internal structure and texture of materials. There are three types of irreducible second-rank tensors. These are (*i*) isotropic: $\alpha_{ij} = a\delta_{ij}$ where a is a scalar, (*ii*) antisymmetric: $\alpha_{ij} + \alpha_{ji} = 0$, and (*iii*) symmetric and traceless (deviatoric): $\alpha_{ij} = \alpha_{ji}$, $\alpha_{ii} = 0$. What distinguishes these tensors are their symmetries. For the study of symmetries one considers the equation

$$P_Q\alpha = \alpha. \; Q \epsilon 0^+(3) \tag{3.4}$$

Then the rigid body rotations that satisfy (3.4) constitute the symmetry group of α. For isotropic second-rank tensors, (3.4) is satisfied by any $Q \epsilon 0^+(3)$ and these tensors are equivalent to scalars. Each antisymmetric second rank tensor is associated with an axial vector and the symmetry group of such a tensor is composed of all rotations about the axis. On the other hand, a symmetric *traceless* second rank tensor remains unchanged under 180 deg rotations about its principal directions. A higher rank irreducible tensor has no symmetries generically, but it is capable of exhibiting a unique collection of symmetries (for brief remarks on this question, cf Onat, 1986).

Now returning to representations of Type (2.1), we think S as composed of a set of irreducible tensors (cf equations (3.3)). Since the virgin state S_o is isotropic $P_QS_o = S_o$ for all $Q \epsilon 0^+(3)$. In view of the above comments on symmetries of irreducible tensors, this implies that all state variables other than scalars must vanish at $t = 0$. Without loss of generality we can assign zero values to scalar state variables at $t = 0$ so that

$$s_1 = 0, \ldots, s_m = 0 \text{ at } t = 0.$$

It is also worth noting that the several of the s_i's may be of the same rank and type.

The law of evolution (2.1*b*) for the inelastic state now takes the more explicit form

$$\dot{s}_i = g_i(s_1, \ldots, s_m; \sigma), \quad g_i : S \to T^{(i)}, \quad i = 1, \ldots, m \tag{3.5}$$

where $T^{(i)}$ is the space of tensors of the rank and type of s_i.

If the constitutive law (2.1) is to satisfy the requirements of Table 1, then in view of equations (3.3) the constitutive functions must be subject to the invariance requirements listed below:

$$\begin{aligned} &Qf(s_j,\sigma)Q^T = f(p_Qs_j, Q\sigma Q^T) \\ &p_QD(s_j) = D(p_Qs_j) \qquad i, j = 1, \ldots m, Q \epsilon 0^+(3). \\ &p_Qg_i(s_j,\sigma) = g_i(p_Qs_j, Q\sigma Q^T). \end{aligned} \tag{3.6}$$

An advantage of the tensorial state variables is that they contain information on the symmetries of a deforming element. Indeed if

$$P_QS = S \tag{3.7}$$

for some $Q \epsilon 0^+(3)$, then it follows from previous considerations that the rigid body rotation Q applied to the stress-free but previously deformed element in S will produce an element which cannot be distinguished from the one in S as far as future mechanical behavior is concerned. Perhaps more importantly the quantity

$$\max\{\|P_QS - S\|, Q \epsilon 0^+(3)\}$$

where the double bars denote a suitable norm will be a measure of anisotropy of the material element S.

4 Representation of Internal Damage Caused by Void Growth on Grain Boundaries

In the previous section we established the tensorial nature of state variables. Our considerations were based on general properties and we avoided any discussion of special cases. In this section we consider internal damage that occurs during tertiary creep of polycrystalline solids to show that tensorial state variables arise naturally from the statistics of internal happenings that accompany deformation.

We consider "homogeneous" deformations of the material

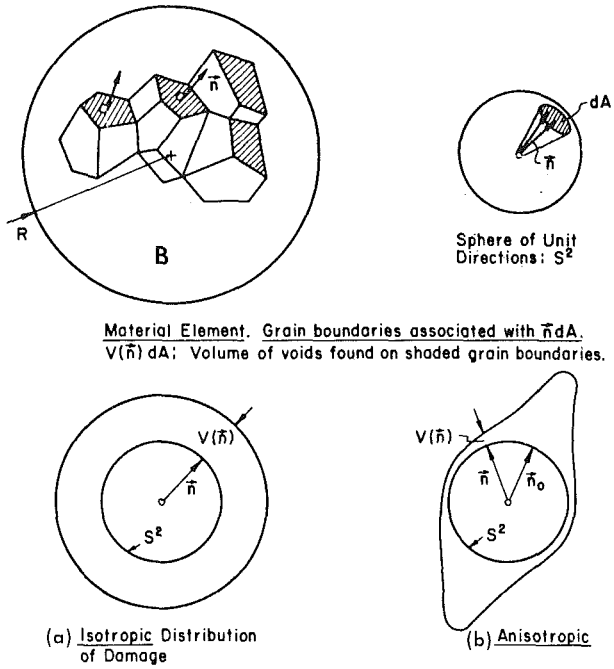


Fig. 2 Orientation of damage

(this does not mean, however, that the deformation is homogeneous in the scale of grains). At a given time we focus our attention on the material points that occupy a spherical region B of unit volume, Fig. 2. This spherical element contains a very large number of grains. We assume that the grain boundaries are planar. We are interested in the mathematical representation of the distribution of cavities among grain boundaries that lie within this element. With this purpose in mind we consider another sphere. This one has unit radius and it depicts directions in space by the position vectors of its surface points (which we denote by S^2). An infinitesimal area dA on S^2 about the end point of $\vec{n}$ will represent a bundle of directions about $\vec{n}$ (cf Fig. 2).

There are two aspects to the problem. The first one is concerned with the statistics of the polycrystalline aggregate within the element B: the distributions of grain size, orientation of grain boundaries, and the orientation of grain lattice may be of interest. Although these distributions remain essentially constant in the present case of small deformation and rotations, they may play a role in the formation of voids. Here we shall assume that no anisotropy is present in these distributions. However, we accept the possibility that voids may arrange themselves on grain boundaries in an anisotropic manner.

We may start the study of this possibility by considering the grain boundaries within B that are orthogonal to a direction present in the generic vector bundle $(\vec{n}, dA)$ shown in Fig. 2. Thus we concentrate our attention on grain boundaries that are nearly perpendicular to $\vec{n}$. A few such boundaries are emphasized in the Figure.

We let $V(\vec{n})dA$ denote the total volume of voids (per unit volume of the material) on grain boundaries perpendicular to vectors of the bundle associated with $\vec{n}$ and dA.

We have here a real value function V defined on the unit sphere S^2:

$$V: S^2 \to R \tag{4.1}$$

that describes the directional dependence of the void distribution amongst grains. Observe that by its definition V has the property

$$V(\vec{n}) = V(-\vec{n}) \tag{4.2}$$

Moreover,

$$\int_{S^2} V(\vec{n})\,dA = 2V_T \tag{4.3}$$

where V_T is the total volume of voids within the material sphere B of unit volume.

In Fig. 2 we show the graphs of two extreme distributions. The first distribution may be called isotropic because the voids are distributed "equally" between grain boundaries of all orientations. In the second distribution the voids are mostly found on grain boundaries perpendicular to the direction $\vec{n}_o$.

Functions which are defined on the unit sphere are common in many branches of physics. Mathematical representation of these functions is a well studied problem (cf Gelfand et al., 1963). A Fourier series based on spherical functions is a standard tool of representation for functions (4.1). Here we consider a modified version of this representation (Onat, 1984). The present version has satisfactory properties of transformation under rigid body rotations of the material element. This representation also provides a simple example in which irreducible even rank tensors appear as "state" variables. The modified Fourier representation of $V(\vec{n})$ has the following form

$$V(\vec{n}) = V_o \cdot 1 + V_{ij} f_{ij}(\vec{n}) + V_{ijk\ell} f_{ijk\ell}(\vec{n}) + \ldots$$
$$i, j, k, \ell, \ldots = 1,2,3 \tag{4.4}$$
$$\vec{n} \in S^2$$

where the usual summation convention is used and 1, f_{ij}, $f_{ijk\ell}$, . . . are the basis functions and V_o, V_{ij}, $V_{ijk\ell}$, . . . are the Fourier coefficients. The basis functions are linear combinations of certain families of spherical functions. They can be defined directly as follows

$$f_{ij}(\vec{n}) = n_i n_j - \frac{1}{3}\delta_{ij},$$

$$f_{ijk\ell}(\vec{n}) = n_i n_j n_k n_\ell - \frac{1}{7}(\delta_{ij}n_k n_\ell + \delta_{ik}n_j n_\ell + \delta_{i\ell}n_j n_k + \delta_{jk}n_k n_\ell \tag{4.5}$$
$$+ \delta_{j\ell}n_i n_k + \delta_{k\ell}n_i n_j) + \frac{1}{5\times 7}(\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}),$$
$$\ldots$$

where (n_1, n_2, n_3) are the components of the *unit* vector $\vec{n}$ in a rectangular basis.

We see that the basis functions are created by the even symmetric products $(n_i n_j)$, $(n_i n_j n_k n_\ell)$, . . . of the components of the unit vector $\vec{n}$. Additional terms in each set of definitions are so chosen that the functions f_{ij}, $f_{ijk\ell}$, . . . are completely symmetric with respect to their indices and they are traceless.

Thus, for example, as it is easily seen from the first of equations (4.5)

$$f_{ij}(\vec{n}) = f_{ji}(\vec{n}) \text{ and } f_{11}(\vec{n}) + f_{22}(\vec{n}) + f_{33}(\vec{n}) = 0 \tag{4.6}$$

It follows then that there are only five independent functions in the set $\{f_{11}(\vec{n}), f_{22}(\vec{n}), \ldots, f_{23}(\vec{n})\}$. Similarly we conclude that there are 9 independent functions in the set $\{f_{ijk\ell}(\vec{n})\}$. It has also been shown that the basis functions sets 1, $\{f_{ij}\}$, $\{f_{ijk\ell}\}$, . . . are orthogonal to each other. This means that

$$\int_{S^2} 1 \cdot f_{ij}\,dA = \int_{S^2} 1 \cdot f_{ijk\ell}\,dA = \int_{S^2} f_{pq} \cdot f_{ijk\ell}\,dA = \ldots = 0 \tag{4.7}$$

The expansion coefficients V_{ij}, $V_{ijk\ell}$, . . . are also completely symmetric and traceless with respect to their indices and they are expressed as inner products of $V(\vec{n})$ and the basis functions:

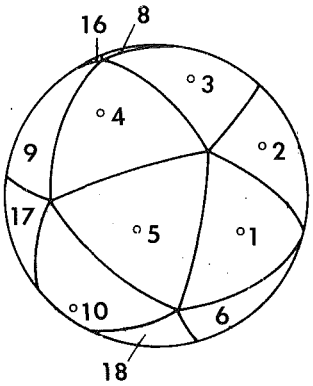


Fig. 3 Regular tesselation of sphere of directions with 20 cells

$$V_o = \frac{1}{4\pi}\int_{S^2} V(\vec{n})dA = \frac{V_T}{2\pi}$$

$$V_{ij} = \frac{1}{4\pi}\,\frac{3\times 5}{2}\int_{S^2} V(\vec{n})f_{ij}(\vec{n})dA,$$

$$V_{ijk\ell} = \frac{1}{4\pi}\,\frac{3\times 5\times 7\times 9}{2\times 3\times 4}\int_{S^2} V(\vec{n})f_{ijk\ell}(\vec{n})dA, \qquad (4.8)$$

. . .

We now show that the expansion coefficients transform as tensors in the sense of equation (4.11)). For this purpose let us return to the tests of Table 1. Let test (*i*) give rise to void distribution $V(\vec{n})$. Let $P_Q V(\vec{n})$ be the void distributions observed in the test (*ii*). Reflection will show that

$$P_Q V(\vec{n}) = V(Q^{-1}\vec{n}) = V(Q^T\vec{n}) \qquad (4.9)$$

where $Q^{-1} = Q^T$ is the transpose or the inverse of $Q \in 0^+(3)$.

It is easily seen using equations (4.5), (4.8), and (4.9) that the expansion coefficients of the distribution $P_Q V$ are obtained from those of V by tensor transformations:

$$V: \quad S^2 \to R = (V_o, V_{ij}, \ldots)$$

$$P_Q V: \quad S^2 \to R = (V_o, Q_{ip}Q_{jq}V_{pq}, \ldots) \qquad (4.10)$$

It must also be emphasized that these even rank tensors are completely symmetric and traceless and hence they are irreducible.

We have just shown that the directional distribution of void volume among the grain boundaries can be represented by a sequence of irreducible even rank tensors $(V_o, V_{ij}, V_{ijk\ell}, \ldots)$. When there is no bias in this distribution (cf Fig. 2(*a*)) then the coefficients V_{ij}, $V_{ijk\ell}$, . . . vanish. On the other hand, when voids are concentrated on grain boundaries perpendicular to a particular direction $\vec{n}_o$, (cf Fig. 2(*b*)), all Fourier coefficients will be nonzero.

It is clear that the future mechanical behavior of the element of interest will in general depend on the void distribution defined by this sequence. It may be that in some cases only the first term of the sequence (which is proportional to the total void volume found in the element) will be of importance from the point of view of mechanical behavior. On the other hand, the effective moduli of a polycrystal which carries a dilute distribution of voids are controlled by the first three "terms" of the sequence (i.e., V_o, V_{ij}, $V_{ijk\ell}$) (cf Onat, 1984).

We close this Section with some remarks on the history of the representation of internal structure and texture of materials. Hilliard (1962) was amongst the first to advocate the use of distribution functions such as $V(\vec{n})$ in the study of anisotropy (see also Bunge, 1969). The use of tensors in the representation of distribution functions appears in the work of Kuksenko and Tamuzs (1981), Leckie and Onat (1981), Onat (1982, 1984), and Kanatani (1984).

To end these brief historical remarks we consider (spherical) histograms which are the oldest tool of representation. For purposes of illustration we use a regular icosohedron to subdivide (i.e., to tessellate) the surface of the sphere of directions (Fig. 3). Let the unit vectors $\vec{n}_\alpha$ define the midpoints of the triangular elements of the tesellation ($\alpha = 1, \ldots, 20$). Because of central symmetry and for the numbering used in the figure we have

$$\vec{n}_1 + \vec{n}_{11} = \ldots = \vec{n}_{10} + \vec{n}_{20} = 0 \qquad (4.11)$$

Now we let V_α denote the void volume found on grain boundaries perpendicular to unit vectors whose ends lie on the αth triangle of tessellation. Clearly

$$V_1 = V_{11}, \ldots, V_{10} = V_{20}. \qquad (4.12)$$

The set $(V_1, \ldots, V_{20})$, which is an approximate representation of $V(\vec{n})$, is called a histogram.

Several remarks about (spherical) histograms may be in order. Histograms are particularly suitable for the representation of strongly anisotropic distributions: the location and the strength of a peak can be described by the simple histogram (0, 0, . . . , V_α, . . . , 0, 0). Whereas in Fourier analysis an accurate description of a peak would require the calculation of a large number of Fourier coefficients.

In histograms accuracy depends on the fineness of the tessellation. But a fine tessellation of the unit sphere inevitably involves uneven subdivisions (cf Horn, 1986), and geodesics of Buckminster Fuller). It is also worth noting that no tessellation has more symmetries than the symmetries of a regular icosohedron. This point gains importance if one wants to use a histogram as a state variable. To clarify this point we recall that if the stress history σ on [0, t] gives rise to the distribution $V(\vec{n})$ then the stress history $Q\sigma Q^T$ where Q is a fixed rotation will give rise to $V'(\vec{n})$ which is obtained from the previous one by a rotation:

$$V'(\vec{n}) = V(Q^T\vec{n})$$

Let $\{V'_\alpha\}$ be the histogram of $V'(\vec{n})$ based on the same regular icosohedron. We raise the questions: What can we say about $\{V'_\alpha\}$? Could we claim that

$$\{V'_\alpha\} = P_Q\{V_B\} \qquad (4.13)$$

where P_Q is a Q-activated map on R^{20}?

Reflection will show that equation (4.13) holds *only* if $Q \in g_{ico}$ where g_{ico} is the symmetry group of the regular icosohedron under rigid body rotations. It is then clear that the invariance requirements (4.6) placed on constitutive equations need to be reconsidered if one wishes to use $\{V_\alpha\}$ as a state variable, because $P_Q\{V_\alpha\}$ is defined only for $Q \in g_{ico}$ (cf Section 6).

5 Thermodynamic Constraints on Laws of Evolution

Dissipative behavior exhibited by materials beyond their elastic range is the subject of this section, where we discuss briefly the thermodynamic basis of dissipativity and then consider the constraints imposed by dissipativity on constitutive equations (3.1).

We recall that the virgin material is in thermo-mechanical equilibrium at temperature θ_o and carries no stress and we are interested in small isothermal and homogeneous deformations and rotations of the material.

In studying time dependent material behavior which involves large irreversible changes in internal structure as in the present case, one must raise the question as to whether it is permissible to use concepts such as temperature and free energy which are defined only for systems in equilibrium. Here when we speak of *isothermal homogeneous* deforma-

tions we mean that the test sample is in thermal contact with a heat bath at constant temperature θ_o and its deformation is homogeneous only over length scales comparable to sample size; as Zener (1948) pointed out, in polycrystalline aggregates a "homogeneous" deformation of the sample is accompanied by fluctuations in strain and stress from grain to grain. This, in general, creates fluctuations in temperature within the sample (which in turn, give rise to dissipation). However, here we shall be interested in relatively slow deformations. For such deformations inhomogeneities in temperature would be small and θ_o could be regarded as the temperature of the sample. We encounter more delicate questions when we consider a little later the advisability of using the concepts of free energy and entropy. Fortunately, in the present case of isothermal deformations the issue of dissipativity can be discussed without recourse to entropy and with the help of observable histories of stress and strain and the associate concepts such as the notions of "minimum work to reach a state" (cf Willems, 1972) and "maximum recoverable work from a given state" (Breuer and Onat, 1964).

Thus we consider the work $\overset{t}{\underset{o}{W}}(\sigma)$ needed to deform, *homogeneously* and slowly, a *unit* volume of virgin material on the time interval $[o, t]$ by applying the stress history σ:

$$\overset{t}{\underset{o}{W}}(\sigma) = \int_o^t \sigma_{ij}(\tau)\dot{\epsilon}_{ij}(\tau)d\tau \tag{5.1}$$

It is a consequence of the first and second laws of thermodynamics that

$$-M \leq \overset{t}{\underset{o}{W}}(\sigma) \tag{5.2}$$

for all stress histories of interest where the nonnegative quantity M has a fixed finite value for the given material. The above inequality states that the work done by the *material* of unit volume must be bounded from above by $M \geq 0$. If this were not the case, then by the first law (and noting that the internal energy per unit volume is finite) unbounded quantities of heat drawn from a single heat bath at temperature θ_o would be converted into useful work and the second law would be violated.

Although most of the following results hold for the general case of $M \geq 0$ we shall consider the case $M = 0$ which covers all structural materials. Materials for which $M = 0$ is said to be *passive*. If $\overset{t}{\underset{o}{W}}(\sigma > 0$ for a stress history that takes the material element back to its original shape at time t, then we say that the passive material is *dissipative*.

For passive materials we combine expressions (5.1) and (5.2) with (2.1*a*) to obtain

$$0 \leq \overset{t}{\underset{o}{W}}(\sigma) = \int_o^t [\sigma_{ij}(\tau)f_{ij}(S(\tau), \sigma(\tau)) + \sigma_{ij}(\tau)D_{ijk\ell}(S(\tau))\dot{\sigma}_{k\ell}(\tau)]d\tau \tag{5.3}$$

where for a given stress history $\sigma(\tau)$, $S(\tau)$ follows from the solution of the evolution law (2.1*b*).

It is clear then that the constitutive function f, g, and D must be subject to constraints, if equation (5.3) were to be satisfied for any stress history and time t of interest. We shall say that the constitutive functions f, g, and D are thermodynamically acceptable if they satisfy these constraints. In order to reveal these constraints we may introduce the notion of minimum work to reach a state $(S, \sigma) \in \Sigma \subset R^n$.

Since the state space Σ is finite dimensional there exists, in general, an infinite number of stress histories $\{\sigma_i(\tau);\ \tau\epsilon[o, T_i]\}$ that bring the virgin material element to the state (S, σ). Each such history would require the expenditure of work $\overset{Ti}{\underset{o}{W}}(\sigma_i)$. On the other hand, for dissipative materials $\overset{Ti}{\underset{o}{W}}(\sigma_i)$ is bounded from below by zero (cf equation (5.3)) and hence there exists[1] a stress history in the above set of histories, denoted by σ_* and defined on $[o, T_*]$ that requires least work to reach the state (S, σ). Clearly for a given material the least work $\overset{T_*}{\underset{o}{W}}(\sigma_*)$ is a function ϕ of the state (S, σ). Note that

$$\phi(S, \sigma) \geq 0 \tag{5.4}$$

and we shall assume that the equality signs holds only for the virgin stress free state $(S_o, 0)$. An implication of this assumption is that the virgin material is compressible.

Recalling the discussion concerning Table 1 of Section 3, we easily establish that the function ϕ must satisfy the invariance requirement

$$\phi(P_Q S, Q\sigma Q^T) = \phi(S, \sigma) \text{ for all } Q\epsilon 0^+(3),\ (S, \sigma) \in \Sigma \tag{5.5}$$

It follows from the definition of $\phi(S, \sigma)$ that

$$\int_{t_1}^{t_2} \sigma_{ij}(\tau)\dot{\epsilon}(\tau)d\tau \geq \phi(S_2, \sigma_2) - \phi(S_1, \sigma_1) \tag{5.6}$$

where t_1 and t_2 denote two consecutive times in a given deformation process and (S_1, σ_1) and (S_2, σ_2) are the respective states of the material.

When $\phi(S, \sigma)$ is a smooth function, (5.6) implies that

$$\sigma_{ij}\dot{\epsilon}_{ij} \geq \dot{\phi} = \frac{\partial \phi}{\partial S} \cdot \dot{S} + \frac{\partial \phi}{\partial \sigma} \cdot \dot{\sigma} \tag{5.7}$$

where $\dot{\sigma}$ and $\dot{\epsilon}$ are the stress and the strain rate experienced by the material at a given time and $\dot{\phi}$ is the rate of change of ϕ along the corresponding trajectory in Σ defined by the constitutive law. Combining equation (5.7) with the constitutive law (2.1) we obtain

$$\sigma_{ij}f_{ij}(S, \sigma) \geq \frac{\partial \phi}{\partial S} \cdot g(S, \sigma) + \left(\frac{\partial \phi}{\partial \sigma_{k\ell}} - \sigma_{ij}D_{ijk\ell}(S)\right)\dot{\sigma}_{k\ell} \tag{5.8}$$

where the dot denotes the inner product in R^n.

The above inequality is valid for any stress history and hence for any $\dot{\sigma}_{k\ell}$. This implies, in the absence of bounds on $\dot{\sigma}_{k\ell}$, that

$$\frac{\partial \phi(S, \sigma)}{\partial \sigma_{k\ell}} = \sigma_{ij}D_{ijk\ell}(S). \tag{5.9}$$

It follows from equation (5.9) and the assumed smoothness of the function $\phi(S, \sigma)$ that

$$\phi(S, \sigma) = \frac{1}{2} D_{ijk\ell}(S)\sigma_{ij}\sigma_{k\ell} + \hat{\phi}(S) \tag{5.10}$$

and the compliance D must be assigned symmetries appropriate to a conservative elastic material,

$$D_{ijk\ell}(S) = D_{k\ell ij}(S). \tag{5.11}$$

By previous assumption (cf (5.4)) $\hat{\phi}(S) > 0$ when $S \neq S_o$. If we further assume that stress-free elements remain compressible then D must be positive definite.

It also follows from (5.8) that

$$\sigma_{ij}f_{ij}(S, \sigma) - \frac{\partial \phi(S, \sigma)}{\partial S} \cdot g(S, \sigma) \geq 0 \tag{5.12}$$

where

$$\frac{\partial \phi}{\partial S} \cdot g = \sum_{i=1}^{m} \frac{\partial \phi}{\partial s_i} \cdot g_i.$$

Expressions (5.10), (5.11), and (5.12) are important consequences of the thermodynamic inequality (5.2) for the class of materials considered in the paper.

It is instructive to write the inequality (5.12) in a different form. For this purpose introduce the fourth rank tensor C of the elastic moduli

[1] It is assumed that the space of admissible stress histories is suitably large, and that T_* may be infinite.

$$C(S) = D^{-1}(S) \tag{5.13}$$

where C and D are considered as linear maps on T_2^s and D^{-1} exists because D is positive definite in the present case. Now using equations (5.9) and (5.11) we can write (5.12) in the following form

$$\frac{\partial \phi}{\partial \sigma} \cdot (-Cf) + \frac{\partial \phi}{\partial S} \cdot g \leq 0 \tag{5.14}$$

It is easily seen from equations (2.1) that $-Cf$ measures the rate of stress during a relaxation episode with $\dot{\epsilon} = 0$. Thus (5.14) states that ϕ cannot increase during a relaxation test and ϕ is a *Lyapunov function* for the vector field with components $(-Cf, g)$ in Σ.

Several points need to be made about the consequences of thermodynamic inequality (5.2). It is known that there are, in general, an infinite number of functions of state that can serve as ϕ in (5.12) or (5.14). Two of these are (*i*) the minimum work per unit volume needed to reach the state (S, σ) which we now denote by $\phi_2(S, \sigma)$, and (*ii*) the maximum work $\phi_1(S, \sigma)$ that can be recovered from the state (S, σ).

To clarify this and other points let us consider the simple case of uniaxial stressing of a linear viscoelastic solid. The time dependent relaxation modulus $\mathcal{E}(t)$ of the solid is assumed to be composed of a sum of decreasing exponentials:

$$\sigma(t) = \int_o^t \mathcal{E}(t-\tau)\dot{\epsilon}(\tau)d\tau, \quad \mathcal{E}(t) = \sum_{i=1}^{n} C_i e^{-t/T_i} \tag{5.15}$$

where σ and ϵ are the uniaxial stress and the corresponding strain. Here C_i and T_i are positive constants that, in principle, be determined by a relaxation experiment. Now we take for simplicity $n = 2$ although the conclusions reached below will be independent of n. The integral representation (6.15) with $n = 2$ is equivalent to a representation of the type (2.1)

$$\dot{\epsilon} = \frac{\sigma - cq}{\nu} + \frac{\sigma}{\mu} + \frac{\dot{\sigma}}{E} \tag{5.16}$$

$$\dot{q} = \frac{\sigma - cq}{\nu}$$

where q replaces S and the four positive constants E, c, ν, and μ follow uniquely from the four constants C_i and T_i. The above constitutive equations can be realized by the 4-element model shown in Fig. 4.

Next we look for a function $\phi(q, \sigma)$ that satisfies the inequality (5.12). Equation (5.10) suggests that we take

$$\phi(q, \sigma) = \frac{1}{2} dq^2 + \frac{\sigma^2}{2E}$$

where d is a positive constant. In view of equations (5.16), (5.12) now becomes

$$\frac{1}{\nu}(\sigma - cq)(\sigma - dq) + \frac{\sigma^2}{2\mu} \geq 0 \tag{5.17}$$

It will be seen that (5.17) will be satisfied for all values σ and q if $d \in [d_1, d_2]$ where the positive constants d_1 and d_2 depend upon $(c, \mu/\nu)$ and c is contained in the interval $[d_1, d_2]$. Thus the thermodynamic inequality (5.12) is satisfied by an infinity of ϕ functions. It is known that the choices of $d = d_1$ and $d = d_2$ correspond, respectively, to functions ϕ_1 and ϕ_2 introduced earlier. When $d = c$, ϕ measures the strain energy ψ_m stored in the springs of the model shown in Fig. 4.

$$\psi_m = \frac{1}{2} cq^2 + \frac{\sigma^2}{2E}. \tag{5.18}$$

When the temperature fluctuations that accompany a time dependent deformation are negligible, then the free energy and entropy of the model at temperature θ_o can be defined as the sum of the free energies and entropies of its elastic elements. Thus the strain energy ψ_m will be identified with the free energy of the model. On the other hand, the internal energy e_m of the model will have the form

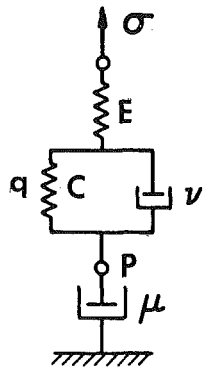


Fig. 4 The 4-element model with modifications: (*i*) mass at *P*; (*ii*) $\mu = \mu(q)$

$$e_m = \frac{1}{2} cq^2 + \frac{1}{2} \frac{\sigma^2}{E} + a_1\theta_o q + a_2\theta_o \frac{\sigma}{E} \tag{5.19}$$

where a_1 and a_2 are positive constants.

The question arises as to whether the thermodynamic quantities defined for the model have any relevance to the solid under consideration. A careful answer to this study would require, amongst other things, calorimetric measurements for the determination of the internal energy e of the solid. It would then be possible to compare the experimentally observed nonlinear dependence of e on σ and q with the one given in equation (5.19).

If the above comparison is satisfactory, then it would mean that the free energy ψ of a linear viscoelastic solid at temperature θ_o would follow from the knowledge of its mechanical behavior. However, in the case of nonlinear viscoelastic solids, such as the ones considered in this paper, one could not hope to even this degree of success because for such materials one cannot in general, construct a unique model (composed of well-defined viscous and elastic elements) from the knowledge of mechanical behavior. Hence our hesitation to work with the free energy ψ which cannot be determined by experiments; the nearest quantity which is amenable to experimentation is the internal energy e but this requires calorimetric measurements. One may hope that ψ could still be calculated with the help of a model, but in the case of nonlinear solids one would need a deeper insight into physics of the situation for the construction of such a model (cf Cocks and Ponter, 1985).

Fortunately in the present case there is no need to work with the free energy. We have shown that if the constitutive functions f, g, and D are thermo-dynamically acceptable (i.e., if the inequality (5.3) is satisfied for all stress histories of interest) then $D(S)$ has the property (5.11) and there exists functions $\phi(S, \sigma)$ such as ϕ_1 and ϕ_2 which are of the form (5.10) and fulfill the constraints (5.3) and (5.4) such that the fundamental inequalities (5.12) or (5.14) are satisfied.

The converse of this result is useful in practice and states that if $D(S)$ has the property (5.11) and if f and g satisfy the inequalities (5.12) or (5.14) where ϕ is a function of the form (5.10) and obeys the constraints (5.4) and (5.5), then f, g, and D are thermodynamically admissible.

The Ω-Form of Constitutive Equations. There is a convenient method of constructing constitutive equations of the type (2.1) which requires the use of the Helmholtz free energy ψ and a potential Ω in such a manner that the thermodynamic requirement (5.2) is satisfied automatically (cf Rice, 1971, and Chaboche, 1985). Apart from the difficulty that, in general, ψ is not available from experiments, the Ω-form of constitutive

equations has the drawback that it establishes too strong a connection between the rate of creep deformation and the rate of change of internal state (cf Onat, 1986) and it may not have the support of experiments in cases where creep damage is important.

To clarify these points we give a brief account of the method. Later we shall look at it with a critical eye. The starting point of the method is the Helmholtz free energy ψ which in the present case depends on (S, σ) and can serve as a ϕ function so that ψ has the form given in (5.10) and the inequality (5.12) is satisfied when ϕ is replaced by ψ.

Next, one defines the thermodynamic force Σ as follows:

$$\Sigma=\frac{\partial\psi}{\partial S}(S, \sigma) \text{ or more explicitly } \Sigma_i=\frac{\partial\psi}{\partial s_i}(s_j, \sigma). \tag{5.20}$$

If the above relation which for a fixed σ takes S to Σ is one-to-one, then S can be thought as a function of Σ and σ

$$S=h(\Sigma, \sigma) \tag{5.21}$$

and the constitutive functions f and g may be regarded as functions of (Σ, σ).

Next a nonnegative-valued function $\Omega(\Sigma, \sigma)$ is introduced such that: (*i*) Ω = const. > 0 surfaces in the (Σ, σ)-space are *convex* and *nested* and contain the origin (0, 0).

It is then postulated that $\Omega(\Sigma, \sigma)$ generates the constitutive functions f and g as follows:

$$\dot{\epsilon}_c=f(S, \sigma)=\hat{f}(\Sigma, \sigma)=\frac{\partial\Omega}{\partial\sigma}(\Sigma, \sigma) \tag{5.22}$$

and

$$\dot{S}=g(S, \sigma)=\hat{g}(\Sigma, \sigma)=-\frac{\partial\Omega}{\partial\Sigma}(\Sigma, \sigma)$$

We observe immediately that in view of equation (5.20), (5.12) becomes

$$\sigma\cdot\frac{\partial\Omega}{\partial\sigma}+\Sigma\cdot\frac{\partial\Omega}{\partial\Sigma}\geq 0$$

and it is automatically satisfied because of the property (*i*) assigned to the $\Omega = C$ surfaces.

The advantage of the approach is that the choice of two scalar-valued functions of state vis. $\psi(S, \sigma)$ and $\Omega(\Sigma, \sigma)$ defines the mechanical behavior of the material completely without violating thermodynamics.

A feature of the Ω form (i.e., equation (5.22)) is the strong connection that it establishes between the creep strain rate and the evoluation of state.

In reading some recent work on constitutive equations one may get the impression that the Ω form of constitutive equations is a consequence of thermodynamics. However, it is known that the thermodynamic requirement (5.2) (and we cannot think of any other thermodynamic constraints for the problem considered here) does not imply that the constitutive equations of the type (5.1) can always be put in the Ω form (cf Onat, 1986). Next we inquire whether there are additional physical considerations that lend validity to the Ω form. This inquiry leads us to construct physically realizable models that do not accept an Ω form. Thus it is unlikely that the Ω form can be defended on general physical grounds. One can only hope that for some purposes the Ω form may constitute an acceptable convenient tool of representation.

It would seem that the origins of the Ω form are to be found in linear viscoelasticity. We know that the linear viscoelastic solid defined by (5.16) accepts an Ω form with $2\Omega(\Sigma, \sigma)$ measuring the current rate of energy dissipation in the dashpots of the associated model composed of springs and dashpots (cf Fig. 4). It is worth observing that a model composed of springs, dashpots and *masses* (which is dissipative) does not accept a proper Ω form.

On the other hand, a simple nonlinear model which exhibits tertiary creep but does not accept an Ω form can be constructed as follows:

We start with the model shown in Fig. 4 but now we assume that the viscosity μ of the lower dashpot depends upon the elongation q of the midspring:

$$\mu=\mu(q)>0$$

say,

$$\mu=\mu_1+\mu_2 e^{-q/q_o} \tag{5.23}$$

where μ_1, μ_2, and $q_o > 0$. (Reflection will show that with a suitable linkage mechanism one can control the size of an opening in the lower dashpot by the length change of the midspring and this would make this dashpot viscosity to depend upon q).

The behavior of the modified model will be controlled by equations (5.16) and (5.23). It will be seen that the modified model is dissipative and it will exhibit tertiary creep, but it will not accept an Ω form:

It follows from equation (5.10) that we must look for a ψ function of the form

$$\psi(\sigma, q)=\frac{\sigma^2}{2E}+\hat{\psi}(q) \tag{5.24}$$

Thus $\dfrac{\partial\psi}{\partial q}=\dfrac{\partial\hat{\psi}}{\partial q}=\Sigma$ which we assume has the inverse $q = h(\Sigma)$.

The first of equations (5.22) and (5.16) then yield

$$\Omega=\frac{1}{2\nu}(\sigma-c\,h(\Sigma))^2+\frac{\sigma^2}{2\mu(h(\Sigma))}+k(\Sigma) \tag{5.25}$$

But the second equation in (5.22) requires that

$$\dot{q}=-\frac{\partial\Omega}{\partial\Sigma} \tag{5.26}$$

If equation (5.26) is to be consistent with the second equation (5.16) then μ cannot depend on q. Hence the result that the model does not accept a representation in the Ω form. In the next Section we shall exhibit constitutive equations which have the support of experiments but also do not accept an Ω form.

6 Constitutive Equations

Constitutive equations of the type (2.1) are now proposed which describe the growth rates for deformation and creep damage in copper-like and aluminum alloy-type materials. These materials are selected because they apparently exhibit extremes of behavior. The behavior of many metals appear to fall between these two extremes, so that weighted combinations of the suggested constitutive equations are likely to provide a suitable description of particular materials.

6.1 Copper-Like Materials. Experimental observations (Johnson, 1962; Hayhurst, 1972) and the results of the micromechanics of high temperature grain boundary void growth (Cocks and Ashby, 1982) suggest that the damage growth on grain boundaries with unit normal $\vec{n}$ takes place only when the normal stress σ_n acting on the plane is positive and may depend on the damage already present there. This suggests that we may use the following law of evolution for the now time-dependent void distribution density $V(\vec{n}, t)$ introduced in Section 4.

$$\frac{d}{dt}V(\vec{n}, t)=\begin{cases} f^{(\sigma_n(t),\, V(\vec{n}, t))} & \sigma_n\geq 0 \\ 0 & \sigma_n<0 \end{cases} \tag{6.1a}$$

where

$$\sigma_n(t)=\sigma_{ij}(t)n_i n_j, \tag{6.1b}$$

and f is a positive valued function and σ_{ij} are the components of the stress carried by the element.

The above evolution law suggests that a histogram representation of $V(\vec{n}, t)$ would be particularly suitable for this material. Thus, if one uses a regular icosohedron for the tessellation of the sphere of directions, then the internal damage will be defined (approximately) by the ten parameters $V_1, \ldots, V_{10}$ where V_α measures the void volume found on the grain boundaries (of a material sphere of unit volume) whose normals lie within the bouquet of normals of the αth face of icosohedron (cf Fig. 5 and equations (4.12)). The evoluation law (6.1) may now be replaced by the approximate one:

$$\frac{dV_\alpha(t)}{dt} = f(\sigma_{n_\alpha}(t),\ V_\alpha(t)) \geq 0, \quad \alpha = 1, \ldots, 10 \qquad (6.2)$$

$f = 0$ when $\sigma_{n_\alpha} \leq 0$, where

$$\sigma_{n_\alpha}(t) = \sigma_{ij}(t) n_i^\alpha n_j^\alpha$$

and n_i^α are the components of the unit vector $\vec{n}_\alpha$ that defines the midpoint of the αth face of the icosohedron.

The state and orientation of the material will be defined by the pair (S, σ) and now we must specify the components of S. Some of these components will measure relevant aspects of distribution of dislocations (which are the main cause of permanent deformations), and some others, such as $\{V_\alpha\}$, measure relevant aspects of damage distribution. If we assume that the times over which transient phenomena associated with the rearrangement of dislocations (that are caused by changes in load) take place are much smaller than rupture time, then in the presence of slowly varying loads the internal state of the material will depend only on the stress carried by the element and on the voids present within. We shall assume, in addition, that the creep rate depends only on the set $\{V_\alpha\}$.

We recognize that other aspects of void distribution (size and shape of voids, distance between voids, etc.) could have an influence on creep rate. However, it is hoped that for the materials, under study each set $\{V_\alpha\}$ will be associated with void distributions that would not differ from each other substantially.

Under the above assumptions we have

$$(S, \sigma) = (V_1, \ldots, V_{10}, \sigma)$$

and we can now go to the constitutive law (2.1). The evolution law (6.2) will now replace equation (2.1*b*). A proposed expression for the strain rate (cf equation (2.1*a*)) is

$$\dot{\epsilon}_{ij} = \dot{V}_T \delta_{ij} + f_1(J_2,\ tr\sigma;\ V_T,\ \Sigma V_\alpha \tau_\alpha) s_{ij} + D_{ijk\ell}\dot{\sigma}_{k\ell} \qquad (6.3)$$

where V_T measures the total void volume found in a material sphere of unit volume:

$$V_T = V_1 + V_2 + \ldots + V_{10}$$

and

$$\dot{V}_T = \sum_{\alpha=1}^{10} f(\sigma_{n_\alpha},\ V_\alpha)$$

In equation (6.3), J_2 is the second invariant of the deviatoric stress $\{s_{ij}\}$ and $\tau_1, \ldots, \tau_{10}$ are, respectively, shear stresses found on planes perpendicular to $\vec{n}_1, \ldots, \vec{n}_{10}$, and f_1 is a positive valued function of its arguments. The elastic compliance $D_{ijk\ell}$ is positive definite and symmetric. Experimental evidence suggests that, while measurable, the effect of damage on elastic properties is, usually, not large so that $D_{ijk\ell}$ will be considered as constant.

The first term in equation (6.3) measures the contribution of internal voids to macroscopic volume change, the second term represents a damage enhanced creep deformation; in the presence of damage f_1 depends not only on J_2, as would be the case in classical creep, but also on $tr\sigma$ and on the present void distribution. Although experimental evidence by Trampzynski et al. (1981) suggests that creep rate is dependent only on the total damage V_T and not on its directional distribution we included $\Sigma V_\alpha \tau_\alpha$ in the arguments of f_1 as an example of a quantity which would make the creep rate depend on the orientation of the void distribution. It will be noticed that this last argument of f_1 is so chosen that the requirements of Table 1 are satisfied for rigid body rotations belonging to the isotropy group g_{ico} of the icosohedron.

To check whether the proposed constitutive equations give rise to dissipative behavior, let us calculate the rate of work of stresses from equation (6.3)

$$\sigma_{ij}\dot{\epsilon}_{ij} = \Big(\sum_{\alpha=1}^{10} f(\sigma_{n_\alpha},\ V_\alpha)\Big) tr\sigma + 2f_1 J_2 + D_{ijk\ell}\sigma_{ij}\dot{\sigma}_{k\ell} \qquad (6.4)$$

Observe that the second term and the time integral of the last term on $[0, T]$ are nonnegative. The first term is also nonnegative if all the principal stresses are either positive or negative. (in the latter case $\sigma_{n_\alpha} < 0$ and hence $f(\sigma_{n_\alpha}, V) = 0$). However, when the principal stresses are not of the same sign, then the first term in equation (6.4) may be negative. But in such a case when $|tr\sigma|$ is large so shall be J_2 so that if f_1 dominates[2] f, as we assume to be the case here, then the sum of the first two terms in equation (6.4) will be nonnegative and hence dissipativity will be assured.

We now show that the constitutive equations (6.2) and (6.3), which is supported by experiments on copper, cannot accept an Ω form.

In view of (6.10) and other previous observations, the free energy ψ for the material under consideration will be of the form

$$\psi(S, \sigma) = \frac{1}{2} D_{ijk\ell}\sigma_{ij}\sigma_{k\ell} + \hat{\psi}(V_1, \ldots, V_{10})$$

Since D is assumed to be a constant thermodynamic, forces will be defined by

$$\Sigma_\alpha = \frac{\partial \hat{\psi}(V_1, \ldots, V_{10})}{\partial V_\alpha}, \quad \alpha = 1, \ldots, 10.$$

Comparing now the first of equations (5.22) with (6.3) we find

$$\dot{\epsilon}_{ij}^c = \frac{\partial \Omega}{\partial \sigma_j}(\Sigma_1, \ldots, \Sigma_{10};\ \sigma_{k\ell}) = \sum_{\alpha=1}^{10} f(\sigma_{n_\alpha},\ V_\alpha)\delta_{ij} + f_1 s_{ij}$$

which implies that the σ-dependence of Ω must be through the invariants $tr\sigma$ and J_2:

$$\Omega = \hat{\Omega}(\Sigma_\alpha,\ tr\sigma,\ J_2)$$

Now the second of equations (5.22) yields

$$\dot{V}_\alpha = -\frac{\partial \hat{\Omega}}{\partial \Sigma_\alpha}(\Sigma,\ tr\sigma,\ J_2)$$

But it is clear that the above expression is not *consistent* with equation (6.2) where the dependence of $\dot{V}_\alpha$ on σ is highly directional.

6.2 Aluminum-Like Materials. Experiments by Johnson et al. (1962) and Trampzynski et al. (1981) on aluminum alloys suggests that the damage is a scalar and that the damage growth rate is dependent on the second invariant J_2 of the deviatoric stress tensor. Then satisfactory growth equations are

$$\dot{V}_o = f(V_o, J_2)$$

$$\dot{\epsilon}_{ij} = \hat{f}(J_2, V_o) s_{ij} + D_{ijk\ell}\dot{\sigma}_{k\ell}$$

and the passivity condition (6.21) is satisfied if $\hat{f}$ is positive

[2] A more precise statement of this argument requires that quantities of interest are nondimensionalized.

valued, and D_{ijkl} do not change appreciably during deformation (cf Lemaitre, 1984; Hayhurst et al., 1984).

Acknowledgments

The authors wish to acknowledge the support of the National Science Foundation.

References

Breuer, S., and Onat, E. T., 1964, "On Recoverable Work in Linear Visco-Elasticity," *J. Appl. Math. and Phys.*, Vol. 15, pp. 12–21.

Bunge, Hans, J., 1969, "Mathematische Methoden der Texturanalysis," Akademie Verlag, Berlin.

Chaboche, J. L., 1982, "Le Concept de Constrainte Effective Appliqué à l'Élasticité et à la Viscoplasticité en Présence d'un Endommagement Anisotrope," *Colloque Internationaux du CNRS*, No. 295, pp. 738–759.

Chaboche, J. L., 1985, "On the Constitutive Equations of Materials Under Monontonic or Cyclic Loadings," Rech. Aérosp. 183-5, pp. 31–43.

Cocks, A. C. F., and Ashby, M. F., 1982, "On Creep Fracture by Void Growth," *Progress in Materials Science*, Vol. 27, pp. 189–244.

Cocks, A. C. F., and Ponter, A. R. S., 1985, "Constitutive Equations for the Plastic Deformation of Solids," Report 85-1, Dept. of Engineering, University of Leicester.

Fardshisheh, F., and Onat, E. T., 1973, "Representation of Elastoplastic Behavior by Means of State Variables," *Problems of Plasticity*, Noordhoff, Leyden, pp. 89–115.

Gel-fand, I. M., Minlos, R. A., and Ya Shapiro, 1963, *Representations of the Rotation and Lorentz Groups and Their Applications*, Pergamon Press, Oxford, pp. 36–47.

Geary, J. A., and Onat, E. T., 1974, "Representations of Non-linear Hereditary Mechanical Behavior," Oak Ridge National Laboratory Report, ORNL-TM-4525.

Hayhurst, D. R., Dimmer, P. R., and Morrison, C. J., 1984, "Development of Continuum Damage in the Creep Rupture of Notched Bars," *Phil. Trans. R. Soc. Lond.*, Vol. A 311, pp. 103–129.

Hilliard, J. E., 1962, *Trans. Met. Soc. AIME*, Vol. 274, p. 1202.

Horn, B. K. P., 1986, *Robot Vision*, M.I.T. Press, Cambridge, MA, p. 379.

Johnson, A. E., and Henderson, J., 1962, "Complex-Stress, Creep, Relaxation and Fracture of Metallic Alloys," *HMSO*, Edinburgh.

Kachanov, L. M., 1958, "Time of the Fracture Process under Creep Conditions," *Izv. Akad. Nauk. SSSR. Otd. Tech. Nauk*, Vol. 8, p. 26.

Kanatani, Ken-Ichi, 1984, "Stereological Determination of Structural Anisotropy," *Int. J. Engng. Sci.*, Vol. 22, pp. 531–546.

Kuksenko, V. W., and Tamuzs, V. P., 1981, *Fracture Micromechanics of Polymer Materials*, Martinus Nijhoff, The Hague, pp. 202–204.

Leckie, F. A., and Hayhurst, D. R., 1974, "On Creep Rupture in Structures," *Proc. Roy Soc.*, Vol. A340, pp. 324–347.

Leckie, F. A., and Onat, E. T., 1981, "Tensorial Nature of Damage Measuring Internal Variables," *Proc. IUTAM Symp. Physical Nonlinearities in Structures*, Senlis, Springer, pp. 140–155.

Lemaitre, J., 1984, "How to Use Damage Mechanics?" *Nuclear Engineering and Design*, Vol. 80, pp. 233–245.

Mandel, J., 1973, "Director Vectors and Constitutive Equations for Plastic and Visco-Plastic Media," *Problems of Plasticity*, Noordhoff, Leyden, pp. 135–143.

Murakami, S., and Ohno, N., 1980, "A Continuum Theory of Creep and Creep Damage," *3rd IUTAM Symp. on Creep in Structures*, Leicester.

Onat, E. T., 1966, "The Notion of State and its Implications in Thermodynamics of Inelastic Solids," *Proc. IUTAM Symposium, Vienna*, pp. 292–314.

Onat, E. T., 1966, "Description of Mechanical Behavior of Inelastic Solids," *Proc. Fifth U. S. National Congress of Applied Mechanics*, pp. 421–434.

Onat, E. T., 1982, "Representation of Inelastic Behavior in the Presence of Anisotropy and of Finite Deformations," *Recent Advances in Creep and Fracture of Engineering Materials and Structures*, B. Wilshire and D. R. J. Owen, eds., Pineridge Press, Swansea, pp. 231–264.

Onat, E. T., 1984, "Effective Properties of Elastic Materials that Contain Penny Shaped Voids," *Int. J. Engng. Sci.*, Vol. 22, pp. 1013–1021.

Onat, E. T., 1986, "Representation of Mechanical Behavior in the Presence of Internal Damage," *Engineering Fracture Mechanics*, Vol. 25, pp. 605–614.

Rice, J. R., 1971, "Inelastic Constitutive Relation for Solids: An Internal Variable Theory and its Application to Metal Plasticity," *J. Mech. Phys. Solids*, Vol. 19, pp. 433–455.

Trampczynski, W. A., Hayhurst, D. R., and Leckie, F. A., 1981, "Creep Rupture of Copper and Aluminum Under Nonproportional Loading," *J. Mech. Phys. Solids*, Vol. 29, pp. 353–374.

Willems, J. C., 1972, "Dissipative Dynamical Systems Part I: General Theory," *Arch. Rational Mech. Analysis*, Vol. 45, pp. 321–351.

Zener, C., 1948, *Elasticity and Anelasticity of Metals*, Univ. of Chicago Press, Chicago, Illinois.

On the Validity of Saint Venant's Principle in Finite Strain Plasticity

D. Durban[1]

W. J. Stronge

Department of Engineering,
University of Cambridge,
Cambridge CB2 1PZ, England

Diffusion of effects from a local disturbance in a homogeneous stress field is analyzed within the framework of continuum plasticity. An axial rate of decay for the variability of the effects of a nonuniform disturbance imposed on one end of a long circular cylinder is determined as a function of the axial stress (or stretch). The analysis considers constitutive equations corresponding to incompressible, finite strain versions of J_2 flow and deformation theories. Both theories result in effects that oscillate and decay exponentially with distance from the imposed disturbance; the rate of decay decreases as the uniaxial tension increases. Deformation theory predicts a larger rate of decay than flow theory except within a small range of stress near the ultimate (necking) load; at necking the rate of decay vanishes.

1 Introduction

This paper presents an elementary investigation of the validity of Saint Venant's principle in *finite strain plasticity*. Most existing studies of this principle are concerned with elastic solids (Horgan and Knowles, 1983). An early exception (Hill, 1950) is a comparison between the exact slip-line solution and Prandtl's cycloid solution to the compression of a block between rough plates. That analysis is for a rigid/perfectly-plastic material where the governing system is hyperbolic. The pressure difference between the two solutions shows an oscillatory decay with distance from the edge which "may be regarded as the plastic analogue of Saint Venant's principle in elasticity." Basically, this principle applies to materials and geometries in which a local disturbance is diffused, e.g., elliptic boundary value problems. In continuum plasticity, however, we will show that a characteristic length for diffusion increases without limit as the tension stress approaches necking.

The setting of this analysis is that of a semi-infinite circular cylinder in a state of uniform axial tension. A small rate-type disturbance that varies radially is suddenly imposed on one end. This axisymmetric disturbance is required to provide continuous plastic loading throughout the body. The constitutive equations employed in our analysis are incompressible, finite strain versions of the J_2 flow and deformation theories of plasticity. Within these theories, the uniaxial stress-strain relations are arbitrary except that the resulting field equations should remain elliptic. However, specific examples have been calculated using the Ramberg-Osgood hardening relation.

This formulation results in an incremental (rate) boundary value problem. An eigenfunction expansion for the effects of the end disturbance is examined for eigenvalues having a smallest real part. These give a lower bound on the axial rate of decay for effects emanating from the local disturbance.

A main conclusion that emerges from this study is that the decay rate decreases as the effective stress increases until, at the ultimate (necking) load, there is no decay. Furthermore, a comparison of flow and deformation theories shows that these material representations have the same rate of decay at low stress levels (elastic range) and at stresses very near the necking stress. These findings are supported by a simple asymptotic expansion in the vicinity of necking; there, contact is also made with the corresponding bifurcation analysis of Hutchinson and Miles (1974). Throughout most of the range between the elastic and necking stresses, deformation theory has a larger rate of decay than flow theory.

2 Mathematical Formulation

Consider an elastoplastic, incompressible, semi-infinite circular cylinder (Fig. 1) in a state of uniform axial tension $\sigma_z = \sigma$. Assume that an instantaneous, axially-symmetric but radially nonuniform, surface disturbance is induced over the end $z = 0$. Put differently, we superpose an axisymmetric but radially nonuniform incremental boundary field on the homogeneous uniaxial tensile stress. The radial profile and the kind of incremental disturbance (i.e., whether traction rates or velocities) is arbitrary, except that continuous plastic loading is required everywhere within the body. The cylindrical surface at radius $r = R$ is assumed to remain stress free.

Our main purpose is to investigate the rate of decay of the local nonuniformity, imposed at the ends, as a function of the tensile stress σ. To this end, the resulting nonhomogeneous rate boundary value problem will be approached by the standard eigenfunction expansion method. The corresponding

[1]Permanent address: Department of Aeronautical Engineering, Technion, Haifa 32000, Israel.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May 12, 1987; final revision, August 15, 1987.

Paper No. 88-WA/APM-12.

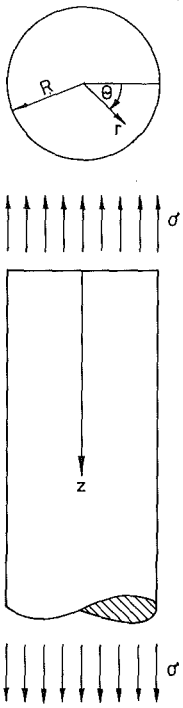


Fig. 1 Cylinder with uniform Cauchy stress $\sigma_z = \sigma$ at current radius R

analysis in linear elasticity (Little and Childs, 1967) predicts an exponential decay of the form $\exp(-\zeta z)$ where, for incompressible materials, the smallest decay rate is given by

$$\zeta R = 2.811 \pm i\ 1.340 \tag{2.1}$$

Denoting by (u,w) the radial and axial components of the perturbed velocity field within the body, we have the strain rate components

$$\epsilon_r = u_{,r} \quad \epsilon_\theta = \frac{u}{r} \quad \epsilon_z = w_{,z} \quad \gamma_{rz} = \frac{1}{2}(u_{,z} + w_{,r}) \tag{2.2}$$

The incompressibility constraint $\epsilon_r + \epsilon_\theta + \epsilon_z = 0$ is satisfied by introducing a stream function ϕ so that

$$u = \frac{1}{r}\phi_{,z} \quad w = -\frac{1}{r}\phi_{,r} \tag{2.3}$$

Material behavior is modelled here by two versions of the finite strain isotropic hardening J_2 theory of plasticity. First, we have flow theory in a reduced form for incompressible materials:

$$\overset{\nabla}{\mathbf{S}} = \frac{2}{3} E\,\mathbf{D} - \frac{E - E_t}{\sigma_e^2}\,\mathbf{S}\,\mathbf{S}\cdot\cdot\mathbf{D} \tag{2.4}$$

where $\mathbf{S}$ is the stress deviator, $\overset{\nabla}{\mathbf{S}}$ is its Jaumann rate, $\mathbf{D}$ is the strain rate tensor, E is the elastic modulus, $\sigma_e^2 = (3/2)\mathbf{S}\cdot\cdot\mathbf{S}$ is the effective stress, and E_t is the tangent modulus, a known function of σ_e. Observing that the only nonvanishing strain-rates are equations (2.2), we find from equation (2.4) and from $\sigma_e \equiv \sigma$ the constitutive relations in terms of nondimensional stress-rate components $\overset{\nabla}{\Sigma}_{ij} = 3\overset{\nabla}{\sigma}_{ij}/2E$,

$$\overset{\nabla}{\Sigma}_r = \dot{\pi} + \epsilon_r + (1 - \eta_t)\epsilon_z/2 \tag{2.5a}$$

$$\overset{\nabla}{\Sigma}_\theta = \dot{\pi} + \epsilon_\theta + (1 - \eta_t)\epsilon_z/2 \tag{2.5b}$$

$$\overset{\nabla}{\Sigma}_z = \dot{\pi} + \eta_t \epsilon_z \tag{2.5c}$$

$$\overset{\nabla}{\Sigma}_{rz} = \gamma_{rz} \tag{2.5d}$$

where $\eta_t = E_t/E$, the hydrostatic tension $\pi = (\Sigma_r + \Sigma_\theta + \Sigma_z)/3$, and the superposed dot stands for the usual derivative with respect to time. Notice that the deviatoric stress $\mathbf{S}$ and the Cauchy stress σ are related by $\mathbf{S} = \sigma - (2E\pi/3)\mathbf{I}$ where $\mathbf{I}$ is the second order unit tensor.

We also consider the incompressible deformation theory which has the form:

$$\overset{\nabla}{\mathbf{S}} = \frac{2}{3} E_s\mathbf{D} - \frac{E_s - E_t}{\sigma_e^2}\,\mathbf{S}\,\mathbf{S}\cdot\cdot\mathbf{D} + \frac{2}{3} E_s\mathbf{Q}\cdot\cdot\mathbf{D} \tag{2.6a}$$

where E_s is the secant modulus (a known function of σ_e) and $\mathbf{Q}$ is the fourth order tensor determined by Hill's (1978) principal axis method (Bassani et al., 1980). When decomposed on the principal axes of strain $\mathbf{Q}$ has the nonvanishing components

$$Q_{ijij} = Q_{ijji} = \frac{1}{2}\left(\frac{\lambda_i^2 + \lambda_j^2}{\lambda_i^2 - \lambda_j^2}\ln\frac{\lambda_i}{\lambda_j} - 1\right)$$
$$\text{no sum, for all } i \neq j \tag{2.6b}$$

where λ_i are the principal stretches. For our problem, equations (2.6) give the four scalar relations

$$\overset{\nabla}{\Sigma}_r = \dot{\pi} + \eta_s[\epsilon_r + (1 - \xi)\epsilon_z/2] \tag{2.7a}$$

$$\overset{\nabla}{\Sigma}_\theta = \dot{\pi} + \eta_s[\epsilon_\theta + (1 - \xi)\epsilon_z/2] \tag{2.7b}$$

$$\overset{\nabla}{\Sigma}_z = \dot{\pi} + \eta_s\,\xi\,\epsilon_z \tag{2.7c}$$

$$\overset{\nabla}{\Sigma}_{rz} = \eta_s\,\Lambda\,\gamma_{rz} \tag{2.7d}$$

where $\eta_s = E_s/E$ and $\xi = E_t/E_s$. The influence of finite axial stretch λ on the shear strain-rate is contained in the term Λ where

$$\Lambda = \frac{\lambda^3 + 1}{\lambda^3 - 1}\,ln(\lambda^{3/2}) \tag{2.8}$$

As is well known, the flow theory (2.4)–(2.5) represents a hypoelastic solid, while the deformation theory (2.6)–(2.7) models a hyperelastic solid; i.e., the latter represents stress fields that are derivable from strain energy functions (as long as no unloading occurs).

A useful observation in this context is that relations (2.5) can be recovered from relations (2.7) with the transformation

$$\eta_s \to 1 \quad \xi \to \eta_t \quad \Lambda \to 1 \tag{2.9}$$

It is also worth mentioning that a different (hypoelastic) version of the deformation theory (2.6), due to Stören and Rice (1975), can be obtained by omitting the $\mathbf{Q}$ term from equation (2.6*a*). The resulting constitutive relations of this particular nonholonomic model are identical to relations (2.7) except that the shear relation (2.7*d*) has

$$\Lambda = 1 \tag{2.10}$$

Turning to equilibrium requirements, this formulation employs the rate form of the equilibrium equation

$$\nabla\cdot(\overset{\nabla}{\sigma} - \mathbf{D}\cdot\sigma + \sigma\cdot\mathbf{W} + \mathbf{I}\cdot\cdot\mathbf{D}\,\sigma) = \mathbf{0} \tag{2.11}$$

where ∇ is the gradient operator, σ is the Cauchy stress tensor, and $\mathbf{W}$ is the spin tensor defined by

$$\mathbf{W} = (\nabla\mathbf{V} - \mathbf{V}\nabla)/2 \tag{2.12}$$

$\mathbf{V}$ being the velocity vector. For the axisymmetric velocity field assumed in this problem, the spin tensor has just one pair of nonzero components, namely

$$W_{rz} = -W_{zr} = (w_{,r} - u_{,z})/2 \tag{2.13}$$

referred to the orthonormal cylindrical triad $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z)$. For incompressible materials the trace $\mathbf{I}\cdot\cdot\mathbf{D}$ vanishes identically so, with the aid of equation (2.13), we obtain the two equilibrium equations

$$\overset{\nabla}{\Sigma}_{r,r} + (1/r)(\overset{\nabla}{\Sigma}_r - \overset{\nabla}{\Sigma}_\theta) + (\overset{\nabla}{\Sigma}_{rz} - (3/2)\,\bar{\sigma}W_{rz})_{,z} = 0 \tag{2.14}$$

$$(\overset{\nabla}{\Sigma}_z - (3/2)\ \bar{\sigma}\ \epsilon_z)_{,z} + (\overset{\nabla}{\Sigma}_{rz} - (3/2)\ \bar{\sigma}\gamma_{rz})_{,r} + (1/r)$$

$$\times(\overset{\nabla}{\Sigma}_{rz} - (3/2)\ \bar{\sigma}\gamma_{rz}) = 0 \tag{2.15}$$

where $\bar{\sigma} = \sigma/E$.

Finally, the stress-free boundary condition for the cylindrical surface is, in the rate form,

$$\mathbf{e}_r \cdot (\overset{\nabla}{\boldsymbol{\sigma}} - \mathbf{D}\cdot\boldsymbol{\sigma} + \boldsymbol{\sigma}\cdot\mathbf{W} + \epsilon_r\boldsymbol{\sigma}) = \mathbf{0} \quad \text{for } r = R \tag{2.16}$$

Using equations (2.2), (2.13), and (2.16) we get two conditions for stress on the cylindrical surface:

$$\overset{\nabla}{\Sigma}_r = 0 \qquad \text{for } r = R \tag{2.17}$$

$$\overset{\nabla}{\Sigma}_{rz} - (3/2)\ \bar{\sigma}\ \gamma_{rz} = 0 \qquad \text{for } r = R \tag{2.18}$$

Equations (2.14)–(2.15), as well as conditions (2.17)–(2.18), are of course independent of any particular choice of constitutive relations. However, for the material models employed in this study it is permissible, in view of equations (2.5*d*) and (2.7*d*), to replace equation (2.18) by the simpler condition

$$\gamma_{rz} = 0 \qquad \text{for } r = R \tag{2.19}$$

To sum up, the differential equations (2.14)–(2.15) and boundary conditions (2.17)–(2.19), when combined with the constitutive relations (2.5), or (2.7), and the kinematical relations (2.2)–(2.3), present a linear homogeneous system for $\dot{\pi}$ and ϕ. That system generates the eigenfunctions which in turn can be used to solve the nonhomogeneous rate boundary value problem for the prescribed disturbance at $z = 0$.

Continuous plastic loading at any material point within the cylinder is assured as long as $\mathbf{S}\cdot\cdot\mathbf{D} > 0$. With the simple tension ground state this condition is reduced to $\sigma\epsilon_z > 0$ or simply

$$\epsilon_z > 0 \tag{2.20}$$

at any material point. It will be shown later that restriction (2.20) is satisfied when the induced edge disturbance has a sufficiently large uniform component in the axial direction.

3 Eigenfunction Solution

For the sake of simplicity we shall demonstrate the method of generating the eigenfunctions for the flow theory, with the analogous results for the deformation theory presented later *in extensio*. After substituting the constitutive relations (2.5) into the rate equilibrium equations (2.14)–(2.15), we obtain two equations which can be rewritten, in the form

$$\dot{\pi} = -\left(\eta_t - \frac{3}{2}\ \bar{\sigma}\right)\epsilon_z - \left(1 - \frac{3}{2}\ \bar{\sigma}\right)\int \frac{(r\gamma_{rz})_{,r}}{r}\, dz \tag{3.1}$$

$$[\epsilon_{r,r} + (1/2)(1-\eta_t)\epsilon_{z,r} + (1/r)(\epsilon_r - \epsilon_\theta) + \gamma_{rz,z}]_{,z}$$
$$- (\eta_t\epsilon_{z,z} + \gamma_{rz,r} + (1/r)\ \gamma_{rz})_{,r} - (3/2)\ \bar{\sigma}W_{rz,zz}$$
$$+ \frac{3}{2}\ \bar{\sigma}\left(\epsilon_{z,z} + \gamma_{rz,r} + \frac{\gamma_{rz}}{r}\right)_{,r} = 0 \tag{3.2}$$

The last equation becomes, with the aid of equations (2.2)–(2.3), a linear homogeneous equation for the stream function ϕ. Anticipating an exponential decay of the nonuniform part of the end disturbance we expand ϕ in the series

$$\phi = -(1/2)\ \epsilon_o r^2 z + A_1 r f_1(r) e^{-\zeta_1 z} + A_2 r f_2(r) e^{-\zeta_2 z} + \ \ldots \tag{3.3}$$

where ϵ_o is the uniform part of the axial strain rate, A_1, A_2, ... are constants, $f_1(r), f_2(r), \ldots$ are the (suitably normalized) eigenfunctions, and ζ_1, ζ_2, ... are the associated eigenvalues. Each of these eigenvalues is, in general, a complex number with a positive real part that governs the axial decay rate. It is understood that the eigenfunctions in equation (3.3) are ordered in increasing magnitude of the real part of the eigenvalues, viz.,

$$\mathrm{Re}\{\zeta_1\} < \mathrm{Re}\{\zeta_2\} < \mathrm{Re}\{\zeta_3\} < \ \ldots \tag{3.4}$$

Of course ζ_1 is of particular interest since this eigenvalue indicates the lower bound on the axial rate of decay for the effects of the imposed disturbance. Note that the axial strain rate that follows from equation (3.3) is, by equations (2.2)–(2.3),

$$\epsilon_z = \epsilon_o + A_1\zeta_1 \frac{(rf_1)'}{r} e^{-\zeta_1 z} + A_2\zeta_2 \frac{(rf_2)'}{r} e^{-\zeta_2 z} + \ \ldots \tag{3.5}$$

where $(\ )' = \dfrac{d(\ )}{dr}$.

Irrespective of the kind of edge disturbance induced, one can always impose a sufficiently high ϵ_o so as to comply with the plastic loading restriction (2.20). This rate formulation for plastic deformation considers either stress-rate or velocity disturbances imposed on the end of the cylinder; these disturbances vary radially. Since we're only concerned with the decay in the cross-sectional variation of the eigenfunction solution for the stress-rates or velocities induced in cylinder, a uniform strain-rate ϵ_o can be superposed on the disturbance. This influences the location of the loading-unloading interface in the cylinder but it has no effect on the homogeneous stress state.

For a representative term of equation (3.3), written as $\phi = rf(r)e^{-\zeta z}$, we obtain the following ordinary differential equaitons from (3.2) after some algebraic manoeuvres

$$(\mathfrak{L} - k_1^2)(\mathfrak{L} - k_2^2)f = 0 \tag{3.6}$$

where operator $\mathfrak{L}$ is defined by

$$\mathfrak{L}(\bullet) = \frac{d^2(\bullet)}{dr^2} + \frac{1}{r}\frac{d(\bullet)}{dr} - \frac{(\bullet)}{r^2} \tag{3.7}$$

and $$k_1^2 = \left(\frac{1 - 3\eta_t + is}{2 - 3\bar{\sigma}}\right)\zeta^2 \quad k_2^2 = \left(\frac{1 - 3\eta_t - is}{2 - 3\bar{\sigma}}\right)\zeta^2 \tag{3.8}$$

$$s = \sqrt{3(1 + 3\eta_t)(1 - \eta_t) - 9\bar{\sigma}^2} \tag{3.9}$$

Equation (3.6) has four independent solutions, expressed by the modified Bessel functions, $I_1(k_1 r)$, $I_1(k_2 r)$, $K_1(k_1 r)$, $K_1(k_2 r)$. The K_1 functions, however, are unbounded at $r = 0$; therefore, they are not applicable for the cylinder problem. The appropriate solution of equation (3.6) is thus reduced to

$$f = C_1 I_1(k_1 r) + C_2 I_1(k_2 r) \tag{3.10}$$

where C_1 and C_2 are constants.

Compliance with boundary condition (2.19) gives, with the aid of equations (2.2)–(2.3),

$$C_1(\zeta^2 - k_1^2)I_1(k_1 R) + C_2(\zeta^2 - k_2^2)I_1(k_2 R) = 0 \tag{3.11}$$

The other boundary condition (2.17) follows from equations (2.5*a*) and (3.1), using again equations (2.2)–(2.3) as well as equation (3.11) and Bessel function identities,

$$C_1[(1 - (3/2)\ \bar{\sigma})(\alpha_2 k_1 R)\zeta^2 I_o(k_1 R) - 2k_1^2 I_1(k_1 R)]$$
$$+ C_2[(1 - (3/2)\ \bar{\sigma})(\alpha_1 k_2 R)\zeta^2 I_o(k_2 R) - 2k_2^2 I_1(k_2 R)] = 0 \tag{3.12}$$

with $$\alpha_1 = 1 - (k_1/\zeta)^2 \quad \alpha_2 = 1 - (k_2/\zeta)^2 \tag{3.13}$$

A nontrivial solution of the linear homogeneous algebraic system (3.11)–(3.12) exists only for discrete eigenvalues ζ where the determinant of that system vanishes. This requirement leads to the transcendental eigenvalue equation

$$(1 - (3/2)\ \bar{\sigma})[(\alpha_2^2 k_1 R) I_o(k_1 R) I_1(k_2 R)$$
$$- (\alpha_1^2 k_2 R) I_o(k_2 R) I_1(k_1 R)]$$
$$+ 2(\alpha_1 - \alpha_2) I_1(k_1 R) I_1(k_2 R) = 0 \tag{3.14}$$

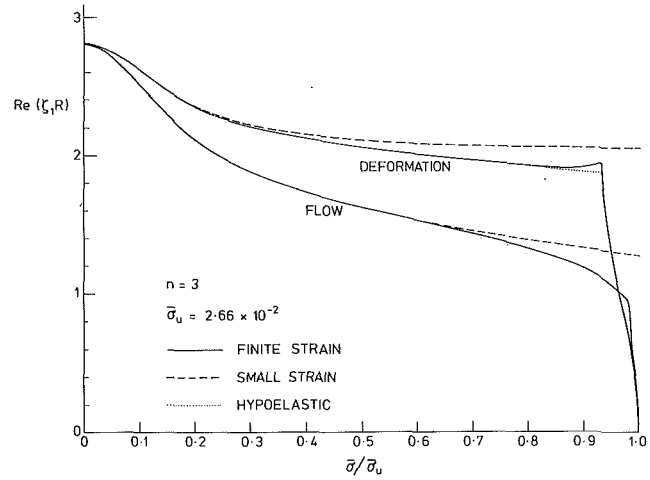


Fig. 2 Variation of the real part of the first eigenvalue with the applied uniform tension for $n = 3$

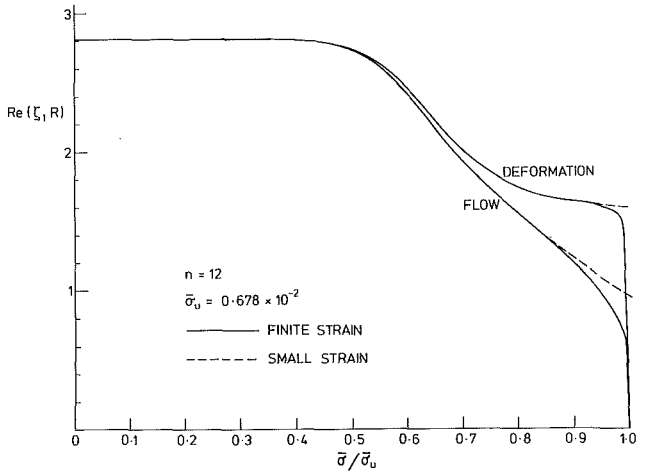


Fig. 3 Variation of the real part of the first eigenvalue with the applied uniform tension for $n = 12$

The roots (eigenvalues) of equation (3.14) can be determined, by resorting to available numerical methods, with any given uniaxial hardening characteristic. Once the eigenvalues have been found it is possible to obtain the associated eigenfunctions from equation (3.10), with the ratio C_1/C_2 determined by one of the two equations (3.11)–(3.12). The proof that each eignefunction produces a self-equilibrating end system is straightforward and follows the standard proof of linear elasticity.

The eigenfunction analysis for the deformation theory (2.6)–(2.7) proceeds along similar lines only with a more complicated algebra. Differential equation (3.6) remains valid with k_1^2, k_2^2 given by

$$k_1^2 = \left(\frac{2\Lambda - 1 - 3\xi + is}{2\Lambda - 3\bar{\sigma}/\eta_s}\right)\zeta^2 \quad k_2^2 = \left(\frac{2\Lambda - 1 - 3\xi - is}{2\Lambda - 3\bar{\sigma}/\eta_s}\right)\zeta^2 \tag{3.15}$$

$$s = \sqrt{(1+3\xi)(4\Lambda - 1 - 3\xi) - (3\bar{\sigma}/\eta_s)^2} \tag{3.16}$$

The corresponding transcendental equation for the eigenvalues reads

$$\left(\Lambda - \frac{3\bar{\sigma}}{2\eta_s}\right)[(\alpha_2^2 k_1 R) I_o(k_1 R) I_1(k_2 R) - (\alpha_1^2 k_2 R) I_o(k_2 R) I_1(k_1 R)] + 2(\alpha_1 - \alpha_2) I_1(k_1 R) I_1(k_2 R) = 0 \tag{3.17}$$

where (α_1, α_2) depend on (k_1, k_2) as in equations (3.13). Note

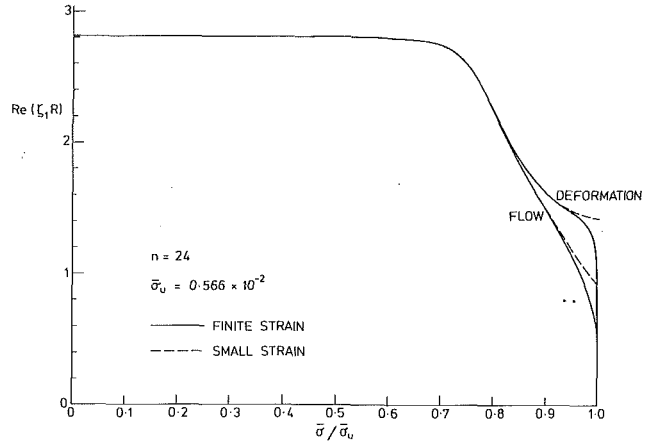


Fig. 4 Variation of the real part of the first eigenvalue with the applied tension for $n = 24$

that with transformation (2.9) we recover from equation (3.17) the analogous flow theory equation (3.14).

In passing it is worth mentioning that the cylinder radius R is the current radius of the deformed configuration. Its relation to the original undeformed radius R_o follows from the incompressibility constraint as

$$R = R_o \lambda^{-1/2} \tag{3.18}$$

4 Numerical Results and Discussion

The eigenvalue equations (3.14) and (3.17) were solved numerically for the Ramberg-Osgood relation

$$\ell n \lambda = \bar{\sigma} + (3/7)\, e_o^{1-n} \bar{\sigma}^n \tag{4.1}$$

with a characteristic strain $e_o = 0.005$ and several different values of the hardening parameter n. The necking stress of material (4.1) can be obtained from the known ultimate load condition $\bar{\sigma} = \eta_I$, for an incompressible solid; this is approximately

$$\bar{\sigma}_u \approx \left(\frac{7}{3n}\right)^{1/n} e_o^{\frac{n-1}{n}} \tag{4.2}$$

Figures 2–4 and Figs. 5–6 show the dependence of the real and imaginary parts, respectively, of the first eigenvalue on the applied tension. At low levels of stress there is complete agreement with the linear elastic result (2.1) over the entire range of n. The real part of ζ_1 decreases as $\bar{\sigma}$ increases with a steep descent towards zero as the necking stress $\bar{\sigma}_u$ is approached (Figs. 2–4). The only exception to this monotonic decrease is the small teat (Fig. 2) obtained from the deformation theory at low values of n. In contrast, the imaginary part displays a nonmonotonic variation with increasing effective stress; the imaginary part vanishes somewhat before the necking stress is attained (Fig. 5–6). Note also that the difference in the results obtained from the two theories is getting smaller as n increases.

A simple asymptotic expansion near the ultimate stress can be provided in support of these numerical findings. Anticipating that $|\zeta_1| \to 0$ as $\bar{\sigma} \to \bar{\sigma}_u$ and using the standard Bessel function expansions

$$I_o(Z) = 1 + Z^2/4 + \ldots, \quad I_1(Z) = Z/2 + Z^3/16 + \ldots \tag{4.3}$$

we examine the asymptotic solution of equation (3.17), for small $|\zeta_1 R|$, as

$$(\zeta_1 R)^2 \sim \frac{8\left(\xi - \frac{\bar{\sigma}}{\eta_s}\right)\left(\Lambda - \frac{3\bar{\sigma}}{2\eta_s}\right)}{\left(\Lambda - \frac{3\bar{\sigma}}{2\eta_s}\right)\frac{\bar{\sigma}}{\eta_s} - (2\Lambda - 1 - 3\xi)\left(\xi - \frac{\bar{\sigma}}{\eta_s}\right)} \tag{4.4}$$

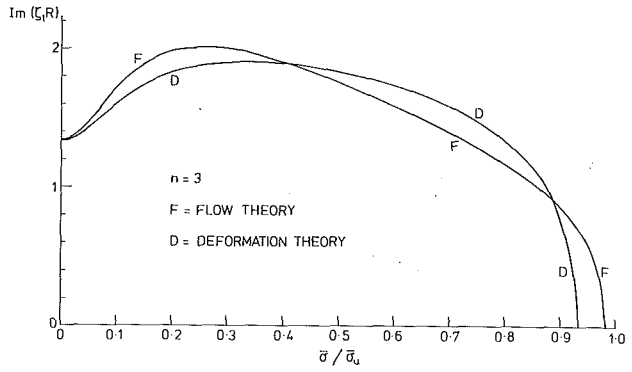


Fig. 5 Variation of the imaginary part of the first eigenvalue with the applied tension. Finite strain analysis for $n = 3$

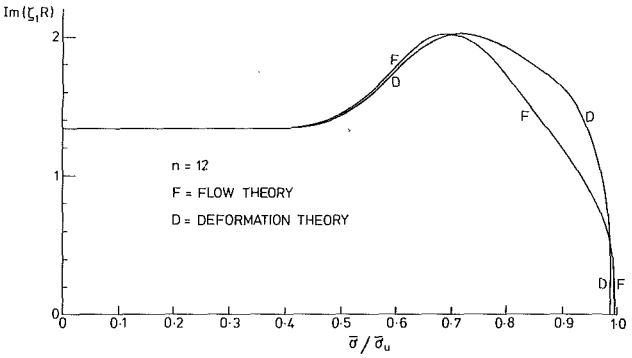


Fig. 6 Variation of the imaginary part of the first eigenvalue with the applied tension. Finite strain analysis for $n = 12$

Near the necking point it is permissible to approximate equation (4.1) by only the power law hardening term. Relation (4.4) is then reduced essentially to

$$\zeta_1 R \sim \sqrt{8n(1-\bar{\sigma}/\bar{\sigma}_u)} \tag{4.5}$$

for both the deformation and flow theories (2.9). This asymptotic result confirms that ζ_1 becomes a purely real number as the necking stress is approached.

The present analysis indicates that the axial decay rate for a disturbance at the end of the cylinder decreases as the axial stretch or effective stress increases. As the stress approaches $\bar{\sigma}_u$ where necking occurs, the decay rate approaches zero. Effects of the nonuniform disturbance decay in an oscillatory manner except in the vicinity of $\bar{\sigma}_u$ where the first eigenvalue is a real number. These conclusions hold for both theories. Sample calculations with the hypoelastic theory (2.10) reveal essentially the same behavior as with the finite strain deformation theory unless n becomes small (Fig. 2).[2] It is worth mentioning that beyond the ultimate load, for $\bar{\sigma} > \bar{\sigma}_u$, purely imaginary eigenvalues become possible, as has been shown in the bifurcation analysis by Hutchinson and Miles (1974). Their asymptotic expression for the bifurcation stress of a very long cylinder (op. cit. equation (5.3)) reads

$$\bar{\sigma} - \eta_t \sim \frac{\gamma^2}{8}\,\bar{\sigma} \tag{4.6}$$

where $\gamma = k\pi R/L$, L is the length of the cylinder and k is an odd integer. The associated oscillatory bifurcation mode, in the axial direction, is of the form $\sin(\gamma z/R)$. It is now easily verified, with the aid of the transformation $\zeta R \to i\gamma$, that (4.4) becomes asymptotically identical with (4.6) when the stress $\bar{\sigma}$ is slightly larger than $\bar{\sigma}_u$.

Calculations for higher eigenvalues of equations (3.14) and (3.17) were also made for a range of values of the strain hardening parameter n. Here again there was complete agreement with the linear elastic results of Little and Childs (1967) at low levels of stress (Tables 1–2). However, no special behavior of these higher eigenvalues has been detected in the neighborhood of the necking stress.

[2] In the limit of an elastic/perfectly-plastic material ($n \to \infty$), necking and the possibility of diffusionless behavior are associated with any plastic deformation.

Table 1 First two eigenvalues for $n = 3$

	FLOW THEORY		DEFORMATION THEORY		
$\bar{\sigma}/\bar{\sigma}_u$	$\zeta_1 R$	$\zeta_2 R$	$\zeta_1 R$	$\zeta_2 R$	$(\ell n\ \lambda)/e_o$
0	2.81 ± i 1.34	6.10 ± i 1.63	2.81 ± i 1.34	6.10 ± i 1.63	0.0
0.1	2.50 ± i 1.71	5.38 ± i 3.09	2.60 ± i 1.60	5.60 ± i 2.71	0.60
0.3	1.88 ± i 2.00	4.03 ± i 4.54	2.21 ± i 1.90	4.73 ± i 3.86	3.35
0.5	1.62 ± i 1.76	3.57 ± i 4.76	2.07 ± i 1.84	4.39 ± i 3.94	10.78
0.7	1.43 ± i 1.40	3.39 ± i 4.80	1.87 ± i 1.58	4.12 ± i 3.71	25.99
0.9	1.18 ± i 0.86	3.31 ± i 4.79	1.92 ± i 0.82	3.77 ± i 3.18	52.10
0.99	0.54 ± i 0.00	3.28 ± i 4.77	0.50 ± i 0.00	3.50 ± i 2.65	68.24

Table 2 First two eigenvalues for $n = 12$

	FLOW THEORY		DEFORMATION THEORY		
$\bar{\sigma}/\bar{\sigma}_u$	$\zeta_1 R$	$\zeta_2 R$	$\zeta_1 R$	$\zeta_2 R$	$(\ell n\ \lambda)/e_o$
0	2.81 ± i 1.34	6.10 ± i 1.63	2.81 ± i 1.34	6.10 ± i 1.63	0.0
0.1	2.81 ± i 1.34	6.09 ± i 1.64	2.81 ± i 1.34	6.09 ± i 1.64	0.14
0.3	2.81 ± i 1.34	6.08 ± i 1.64	2.81 ± i 1.34	6.08 ± i 1.63	0.41
0.5	2.73 ± i 1.44	5.88 ± i 2.12	2.73 ± i 1.43	5.87 ± i 2.08	0.68
0.7	1.92 ± i 2.02	4.11 ± i 4.51	2.01 ± i 2.01	4.29 ± i 4.35	1.18
0.9	1.18 ± i 1.20	3.27 ± i 4.93	1.64 ± i 1.63	3.57 ± i 4.66	5.90
0.99	0.68 ± i 0.48	3.21 ± i 4.94	1.34 ± i 0.00	3.46 ± i 4.35	16.01

The results for the eigenvalues presented in this study fall within the elliptic range of the governing partial differential equations. Ellipticity of the system is maintained as long as s from equation (3.16) remains real. That condition can be restated as (Bassani et al., 1980)

$$\bar{\sigma}^2 < (\eta_s + 3\eta_t)(4\eta_s\Lambda - \eta_s - 3\eta_t)/9 \tag{4.7}$$

for the deformation theory, and as

$$\bar{\sigma}^2 < (1 + 3\eta_t)(1 - \eta_t)/3 \tag{4.8}$$

for the flow theory. These conditions are satisfied for both finite strain models over the range of deformation considered here. For the hypoelastic material (2.4)–(2.5) with a pure power hardening law, restriction (4.7) becomes

$$(\ell n\ \lambda)^2 < \frac{(n+3)(n-1)}{3n^2} \tag{4.9}$$

in complete agreement with Needleman and Rice (1978, equation (49)).

The role of changes in geometry can be assessed by comparing the finite strain solutions with the corresponding results of the small strain analysis. The Jaumann rate in equations (2.5) and (2.7) is then replaced by the simple time derivative and Λ in equation (2.7*d*) is identically equal to unity. Likewise, all coupled terms in the equilibrium equations (2.14)–(2.15), and boundary condition (2.16), are omitted. The resulting eigenvalue equation for the small strain deformation theory is then obtained directly from equations (3.15)–(3.17) with $\bar{\sigma} = 0$ and $\Lambda = 1$. For the small strain flow theory, the eigenvalue equation follows with the further transformation $\xi \to \eta_t$. Numerical results for the first eigenvalue for the small strain models are in fairly good agreement (Figs. 2–4) with the finite strain results except near the ultimate tension stress.

A final comment is related to *elastic unloading* from a current stress σ. Among the constitutive equations used here only the flow theory (2.4)–(2.5) is capable of properly modelling the resulting rate boundary value problem with $\eta_t \equiv 1$. Then if the uniform component ϵ_o in equation (3.5) is sufficiently large and negative to ensure elastic unloading at every material point, the axial decay rate is given essentially by the linear

elastic solution (2.1). The elastic decay rate is always larger than the decay rate for continuing plastic deformation at the same stress σ; the largest difference between these decay rates occurs near the necking load.

Acknowledgment

D. Durban wishes to express his thanks for the kind hospitality of the Cambridge University Engineering Department during the sabbatical year 1986/7.

Part of this work has been supported by the V.P.R. – Aeronautical Engineering Research Fund, Technion, Haifa.

References

Bassani, J. L., Durban, D., and Hutchinson, J. W., 1980, "Bifurcations at a Spherical Hole in an Infinite Elastoplastic Medium," *Math. Proc. Camb. Phil. Soc.*, Vol. 87, pp. 339–356.

Hill, R., 1950, *The Mathematical Theory of Plasticity*, Oxford, p. 233.

Hill, R., 1978, "Aspects of Invariance in Solid Mechanics," *Advances in Applied Mechanics*, Vol. 18, pp. 1–75.

Horgan, C. O., and Knowles, J. K., 1983, "Recent Developments Concerning Saint-Venant's Principle," *Advances in Applied Mechanics*, Vol. 23, pp. 179–269.

Hutchinson, J. W., and Miles, J. P., 1974, "Bifurcation Analysis of the Onset of Necking in an Elastic/Plastic Cylinder Under Uniaxial Tension," *J. Mech. Phys. Solids*, Vol. 22, pp. 61–71.

Little, R. W., and Childs, S. B., 1967, "Elastostatic Boundary Region Problem in Solid Cylinders," *Quart. Appl. Mech.*, Vol. 25, pp. 261–274.

Needleman, A., and Rice, J. R., 1978, "Limits on Ductility Set by Plastic Flow Localization," *Mechanics of Sheet Metal Forming*, D. P. Koistinen and N. M. Wang, eds., Plenum Press, pp. 237–265.

Stören, S., and Rice, J. R., 1975, "Localized Necking in Thin Sheets," *J. Mech. Phys. Solids*, Vol. 23, pp. 421–441.

Micromechanical Aspects of Isotropic Granular Assemblies With Linear Contact Interactions

R. J. Bathurst
Associate Professor,
Civil Engineering Department,
Royal Military College of Canada,
Kingston, Ontario, Canada K7K 5L0

L. Rothenburg
Associate Professor,
Civil Engineering Department,
University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1

The paper presents a micromechanical analysis of plane granular assemblies of discs with a range of diameters, and interacting according to linear contact force-interparticle compliance relationships. Contacts are assumed to be fixed and indestructible. Macroscopically, the system is described in terms of a two-dimensional analogue of generalized Hooke's law. Explicit expressions for elastic constants in terms of microstructure are derived for dense isotropic assemblies. It is shown that Poisson's ratio for dense systems depends on the ratio of tangential to normal contact stiffnesses. The derived expression for Poisson's ratio is verified by numerically simulating plane assemblies comprising 1000 particles. The effect of density on Poisson's ratio is investigated using numerical simulations. The theory of dense plane systems is extended to dense three-dimensional systems comprising spheres. Finally, it is shown that Poisson's result $\nu = 1/4$ is recovered for spherical particles with central interactions.

Introduction

This study is concerned with the micromechanics of two-dimensional random isotropic assemblies of discs. The paper is limited to a detailed analysis of microscopic processes in the simplest system of this class comprising particles with a fixed system of indestructible contacts. Particles interact according to linear contact force-interparticle compliance relationships at the contacts.

From a macroscopic point of view, deformation properties of the system are described by a two-dimensional analogue of generalized Hooke's law and emphasis is placed on the relationship between macroscopic elastic parameters (E, ν) and characteristics of microstructure such as interparticle stiffness and contact density.

Numerical simulations reported in this paper show that despite the apparent simplicity of the considered systems, their behavior is complex at the microscopic level. Results of numerical simulations are used in the current study to both guide and verify analytical developments which link microscopic and macroscopic descriptors.

Results of this study show that the principal element of complexity is local interparticle rotations which, nevertheless, become negligible in very dense systems. For these systems, an explicit relationship between Poisson's ratio and the ratio of linear contact stiffness components is derived. The latter development is presented for both plane and three-dimensional systems and Poisson's result $\nu = 1/4$ (e.g., Love, 1926) is recovered for spherical particles with central interactions.

Description of Microstructure

General. Microchemical studies of granular materials require introduction of some unique physical concepts and have necessarily evolved a terminology specific to the discipline. In the soil mechanics literature the term *fabric* has been used extensively as a generic term to describe the geometry of particle packing (microstructure). In this section, characteristics of fabric relevant to mechanical description of two-dimensional assemblies of discs are introduced. Similar characteristics can be introduced for three-dimensional systems.

The assemblies under study are assumed to comprise essentially rigid particles which are joined together at indestructible compliant point contacts.

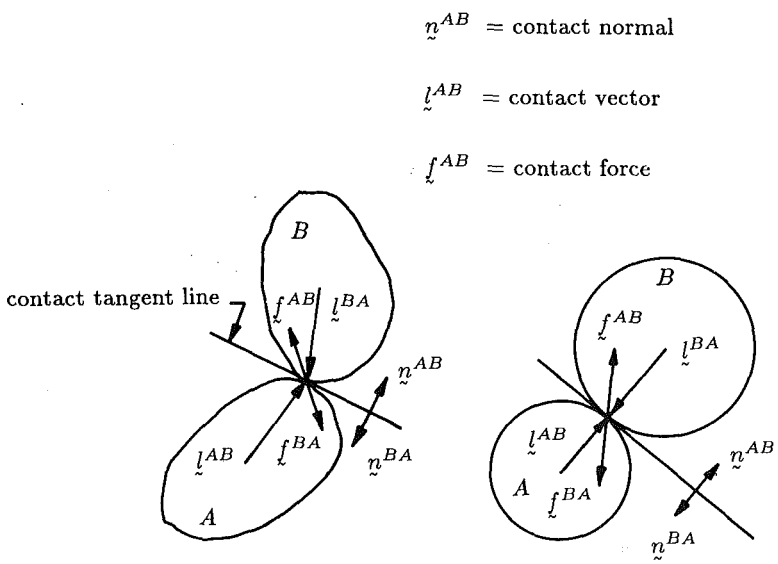
An individual particle at static equilibrium may be in contact with several neighbors. The number of contacts per particle is called the *coordination number* of the particle. Clearly each *physical* contact contributes two *contacts* to the assembly. The *average coordination number*, γ of the assembly is:

$$\gamma = \frac{M_V}{N} \tag{1}$$

Here M_V represents the total number of contacts within the assembly volume and N, the total number of particles.

Coordination number introduced above is an incomplete description of particle packing as it carries no information on relative particle orientations. This aspect of microstructure is often described by particle *contact normals* where a contact

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 2, 1987; final revision August 19, 1987.



a) Non-circular Particle b) Disc-shaped Particle

Fig. 1 Contact normals, contact vectors, and contact forces

normal $\mathbf{n}^c$ is the exterior directed normal to the tangent plane at the point of contact between particles. The relative frequency of contacts with different orientations of normals can be described in terms of a *contact orientation* distribution $E(\theta)$ such that $M_V E(\theta)\,\Delta\theta$ is the number of contacts with normals between θ and $\theta+\Delta\theta$ (Horne, 1965).

Description of particles of arbitrary shape requires identification of contacts in terms of *contacts vectors* (Rothenburg and Selvadurai, 1981*a*). For each contact, contact vector $\mathbf{l}^c$ is directed from the mass center of a particle to a point of contact with a neighbor. The orientational distribution of contact vector lengths implicitly contains information related to particle shape. Contact normals and contact vectors are illustrated in Fig. 1. For spherical particles (or discs) the direction of contact vectors is coincident with the direction of contact normals and the length of contact vectors is equal to particle radius.

Microstructure of Isotropic Assemblies of Discs. In the current study, contact normals are assumed to be distributed homogeneously through a two-dimensional system consisting of a very large number of *discs*.

Rothenburg (1980) has proposed that for two-dimensional assemblies of discs, $E(\theta)$ can be represented by a truncated even Fourier series of the form:

$$E(\theta)=\frac{1}{2\pi}\{1+a\cos 2(\theta-\theta_a)+b\cos 4(\theta-\theta_b)\} \tag{2}$$

Expression (2) satisfies the condition $E(\theta)=E(\theta-\pi)$ for assemblies of discs and when integrated over the limits $0\le\theta\le 2\pi$ gives:

$$\int_0^{2\pi} E(\theta)d\theta=1 \tag{3}$$

Terms a and b are called *coefficients of anisotropy* and define frequencies of contact normals in *directions of anisotropy* θ_a and θ_b. It should be noted, however, that expression (2), in the strictest sense, describes an infinite, statistically homogeneous, assembly where the normalized distribution function $E(\theta)$ can be continuous. For any large but finite system of particles this relationship is a useful approximation.

In this paper, theoretical developments are restricted to *isotropic* assemblies (i.e., systems with $a=0$, $b=0$). Under these conditions:

$$E(\theta)=\frac{1}{2\pi} \tag{4}$$

Lack of orientational bias in numerically simulated assemblies was checked by calculating parameters of anisotropy from complete information on orientation of contact normals. The technique used to carry out these calculations has been described by Bathurst (1985).

In the current investigation, theoretical developments are also restricted to assemblies with a narrow range of disc sizes in which there is no bias between particle size and direction of interparticle contacts. Hence:

$$\bar{l}^c(\theta)=\bar{l}_o \tag{5}$$

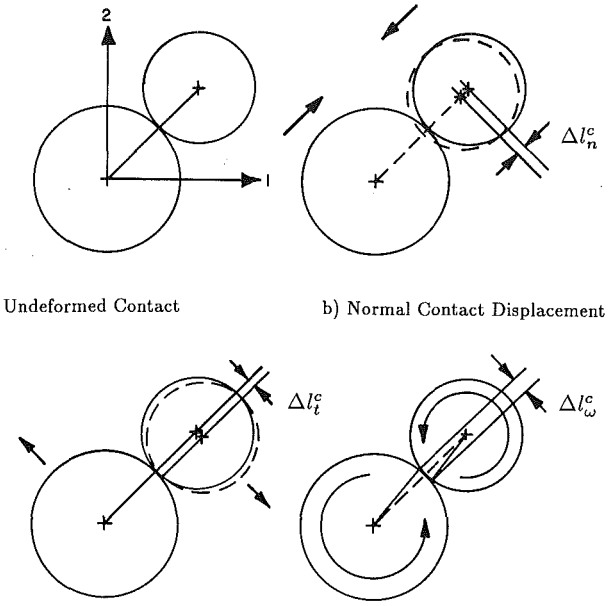
Here $\bar{l}_o$ represents the *average* contact length taken from all assembly contacts.

Theoretical Developments

Average Stress Tensor from Averages of Contact Forces. An average stress tensor in terms of the summation of discrete contact forces and fabric can be expressed as:

$$\bar{\sigma}_{ij}=\frac{1}{V}\sum_{c\in V} f_i^c l_j^c \qquad i,j=1,2 \tag{6}$$

Terms f_i^c and l_j^c refer to scalar components of contact forces $\mathbf{f}^c$ and contact vectors $\mathbf{l}^c$ at contact locations (refer to Fig. 1). Equivalent expressions for three-dimensional idealized granular assemblies have been reported by Weber (1966), Dantu (1968), Rothenburg (1980), Christoffersen et al. (1981), and Bathurst (1985). Rothenburg (1980) and Rothenburg and Selvadurai (1981a) have proposed that expression (6) is a useful approximation to the stress tensor of continuum mechanics for granular assemblies comprising a large but finite number of particles. This equivalency can be understood by considering sums of force-contact vector components for many subregions of a given assembly volume. Quantities calculated from equation (6) would be expected to fluctuate from subvolume to subvolume. However, as the subdomains



a) Undeformed Contact b) Normal Contact Displacement

c) Translational Shear Displacement d) Rotational Shear Displacement

Fig. 2 Contact displacement components

increase in volume and number of particles within a homogeneous system, these fluctuations can be expected to become smaller and smaller. This tendency to a single representative *average stress tensor* is assured by the composition of the function where each term makes a small contribution to $f_i^c l_j^c / V$. For finite but large particle systems, the average stress tensor from discrete information is an accurate analogue to the stress tensor of continuum mechanics and in the following text they are assumed equivalent (i.e., $\sigma_{ij} = \bar{\sigma}_{ij}$).

Unfortunately, calculation of the average stress tensor using relationship (6) requires exact knowledge of contact forces and contact vector terms for all particles. Equivalent more manageable expressions can be developed by considering certain *averages* of grouped discrete information in a similar manner to the approach adopted in the previous section.

If contacts are grouped within a finite number of orientational class intervals, then group averages $\overline{f_i^c l_j^c}(\theta_g)$ can be calculated. The stress tensor relationship (6) can now be rewritten as:

$$\sigma_{ij} = \frac{M_V}{V} \sum_{\theta_g} \overline{f_i^c l_j^c}(\theta) E(\theta) \Delta\theta \tag{7}$$

Here a normalized discontinuous $E(\theta)$ is used to describe the orientational distribution of contact normals. Assuming an assembly with $\lim_{V\to\infty}$, $\lim_{M_V\to\infty}$ and $\lim_{\Delta\theta\to 0}$, relation (7) can be expressed in integral form as:

$$\sigma_{ij} = \frac{M_V}{V} \int_0^{2\pi} \overline{f_i^c l_j^c}(\theta) E(\theta) d\theta \tag{8}$$

If *isotropic* assemblies are considered, possible correlations between $\bar{l}_j^c(\theta)$ and $\bar{f}_i^c(\theta)$ are not a concern and the stress tensor expression is simplified to:

$$\sigma_{ij} = \frac{m_v \bar{l}_o}{2\pi} \int_0^{2\pi} \bar{f}_i^c(\theta) n_j(\theta) d\theta \tag{9}$$

Here, the term $m_v = M_V/V$ is introduced for brevity and is used to denote contact density with respect to assembly area. The above expression forms the basis of a constitutive relationship once the link between contact forces and strain is established.

Relationship Between Average Contact Forces and Strain Tensor. The link to average contact displacements can be made through a contact force-displacement law. A linear contact model offers mathematical simplicity and can be expressed as follows:

$$f_n^c = k_n \left(\frac{\Delta l_n^c}{l} \right)$$

$$f_s^c = k_s \left(\frac{\Delta l_t^c}{l} + \frac{\Delta l_\omega^c}{l} \right) \tag{10}$$

Here l is the distance between particle centers in contact (i.e., the sum of contact vector lengths at a contact); $\Delta l_n^c / l$ is the relative normal displacement between particle centers; $(\Delta l_t^c / l + \Delta l_\omega^c / l)$ represents relative tangential displacement at a contact and consists of two terms describing relative translational displacement between particle centers and relative rotation. These terms are illustrated in Fig. 2. Parameters k_n and k_s in equations (10) refer to normal and tangential (shear) contact stiffnesses and f_n^c, f_s^c the associated contact force components. A positive value for f_s^c signifies a contact shear force which tends to rotate a disc in a counterclockwise direction.

Further theoretical developments are simplified if rotations Δl_ω^c can be neglected. The results of numerical simulations presented later in this paper show that this simplification is valid for dense systems. For these systems, the development of contact forces is entirely due to relative displacement components between particle centers. Considering that expression (9) for the stress tensor involves only *averages* of forces of similar orientations, it is reasonable to equate the latter to quantities describing *average* displacement components for similarly oriented contacts. It is useful at this point to introduce terms describing relative normal and tangential (shear) interparticle displacement components such as $\delta_n^c(\theta)$ and $\delta_t^c(\theta)$ averaged over groups of contacts with similar orientations:

$$\bar{\delta}_n^c(\theta) = \left(\frac{\overline{\Delta l_n^c(\theta)}}{l}\right)$$

$$\bar{\delta}_t^c(\theta) = \left(\frac{\overline{\Delta l_t^c(\theta)}}{l}\right) \tag{11}$$

Averages of forces with similar orientations can now be written as:

$$\bar{f}_n^c(\theta) = k_n \bar{\delta}_n^c(\theta)$$

$$\bar{f}_s^c(\theta) = k_s \bar{\delta}_t^c(\theta) \tag{12}$$

In order to link $\bar{\delta}_n^c(\theta)$ and $\bar{\delta}_t^c(\theta)$ with the strain tensor, it is convenient to resort to properties of the strain tensor of continuum mechanics as follows: In a uniformly strained continuum, a vector **L** connecting two arbitrary points is transformed into vectors $\mathbf{L}+\Delta\mathbf{L}$ in such a manner that $\Delta L_i = \epsilon_{ij} L_j$. Relative displacements in the direction normal and tangential to vector **L** can be calculated as follows:

$$\frac{\Delta L_n}{L} = \epsilon_{ij} n_i n_j$$

$$\frac{\Delta L_t}{L} = \epsilon_{ij} t_i n_j \quad i, \quad j = 1, 2 \tag{13}$$

Here **n** and **t** are coincident and orthogonal to **L**, respectively, and are defined by $\mathbf{n} = (\cos\theta, \sin\theta)$ and $\mathbf{t} = (-\sin\theta, \cos\theta)$. Expressions (13) cannot be applied on a scale comparable to the size of *grains* which physically constitute a *continuum*. However, if this is nevertheless done, it can expected that relations (13) will hold true when expressed as averages taken over an ensemble of similarly oriented points. When such points correspond to centers of particles forming contacts with similar orientations, it is reasonable to expect that:

$$\bar{\delta}_n^c(\theta) = \zeta(\epsilon_{ij} n_i n_j)$$

$$\bar{\delta}_t^c(\theta) = \zeta(\epsilon_{ij} t_i n_j) \tag{14}$$

where ζ is a constant. More detailed analysis presented by Rothenburg (1980) suggests that $\zeta < 1$. Assumptions which lead to equations (13) and link microscopic averages with similar characteristics calculated on the basis of rules of continuum mechanics are quite common (e.g., Batchelor and O'Brien, 1977). In the present paper the above relationships are verified directly on the basis of numerical simulations described later in this paper.

Stress-Strain Relationship. If expressions for average normal and tangential contact forces (12) are combined with equations (14) and the resulting expression for the average contact force vector is introduced into equation (9), then the following stress-strain relationship can be recovered:

$$\sigma_{ij} = A_{ijkl}\epsilon_{kl} \quad i, j, k, l = 1, 2 \tag{15}$$

where:

$$A_{ijkl} = \frac{\zeta k_n \bar{l}_o m_v}{2\pi} \int_0^{2\pi} \{n_i n_j n_k n_l + \lambda t_i n_j t_k n_l\} d\theta \tag{16}$$

In these expressions, parameter λ is introduced as the ratio of tangential to normal contact stiffness (i.e., $\lambda = k_s/k_n$).

Direct calculation of integrals defining components of A_{ijkl} results in Hooke's law for two-dimensional isotropic material with bulk and shear moduli as follows:

$$K = \frac{m_v \bar{l}_o k_n \zeta}{4}, \quad G = \frac{m_v \bar{l}_o k_n \zeta}{4}\left(\frac{1+\lambda}{2}\right) \tag{17}$$

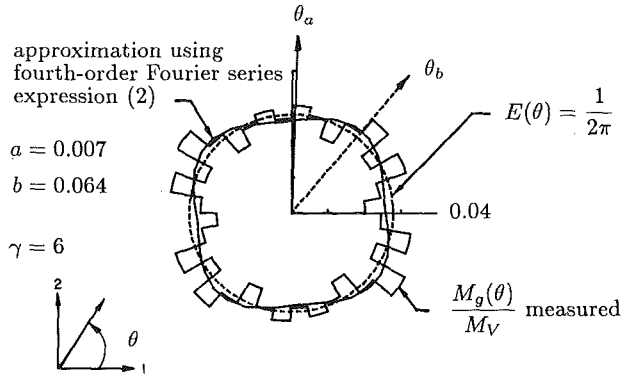


Fig. 3 Normalized distribution of contact orientations for 1000 disc assembly

Although the above moduli contain parameter ζ whose value is unspecified, Poisson's ratio is independent of ζ and depends on the ratio of tangential to normal contact stiffness according to:

$$\nu = \frac{1-\lambda}{3+\lambda} \tag{18}$$

Numerical Simulation of Disc Assemblies

General. In the current investigation, numerical simulation of disc assemblies comprising 1000 particles was undertaken to verify fundamental relationships proposed in the preceding text. A principal advantage of numerical simulation is that it allows all *microscopic* information to be extracted from the assemblies under study. In addition, the influence of micromechanical properties such as stiffness ratio λ can be assessed more readily from these experiments than from comparable physical models (i.e., photo-elastic disc assemblies).

Numerical simulations were carried out using a program which is a modified version of the program BALL originally reported by Strack and Cundall (1978) and used by them to investigate the micromechanical behavior of *cohesionless* disc assemblies. Major modifications involved changes to internal bookkeeping to take advantage of specific computer hardware and elimination of data updating algorithms made possible by numerical assemblies comprising *fixed* contacts. The program implements a time-finite-difference scheme which solves the system of equations modelling a dynamic transient mechanical system. The mechanical system can be imagined as a network of lumped-mass-dashpot elements in which linear springs connect disc-shaped masses. Although the system is dynamic, the transient state approaches a static equilibrium condition if loading rates at the sample boundaries are kept low enough that inertial forces are always a small fraction of the average contact forces acting through the assembly. Kinetic energy is dissipated through the introduction of artificial damping, without which, the approximation to a static equilibrium condition would not be achieved.

Numerical tests were carried out from initial (undeformed) assemblies with near-isotropic microstructure. The disc radii in these tests fell within a narrow size range $0.78 \le r/\bar{r}_o \le 1.29$. Prior to loading, the coordination number of an assembly could be modified by searching out *near* contacts or deleting selected contacts in a random manner. Qualitatively this is equivalent to introducing small distortions in disc geometry such that interparticle contacts are created or lost while maintaining coincidence of contact normals and contact vectors.

The resulting microstructure for a typical assembly in the current study is illustrated by the polar histograms plotted in

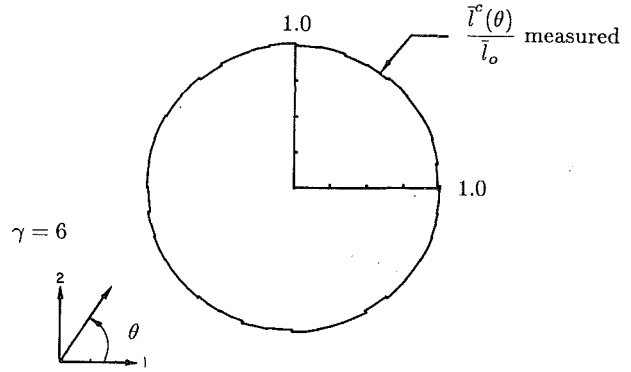


Fig. 4 Normalized distribution of contact lengths for 1000 disc assembly

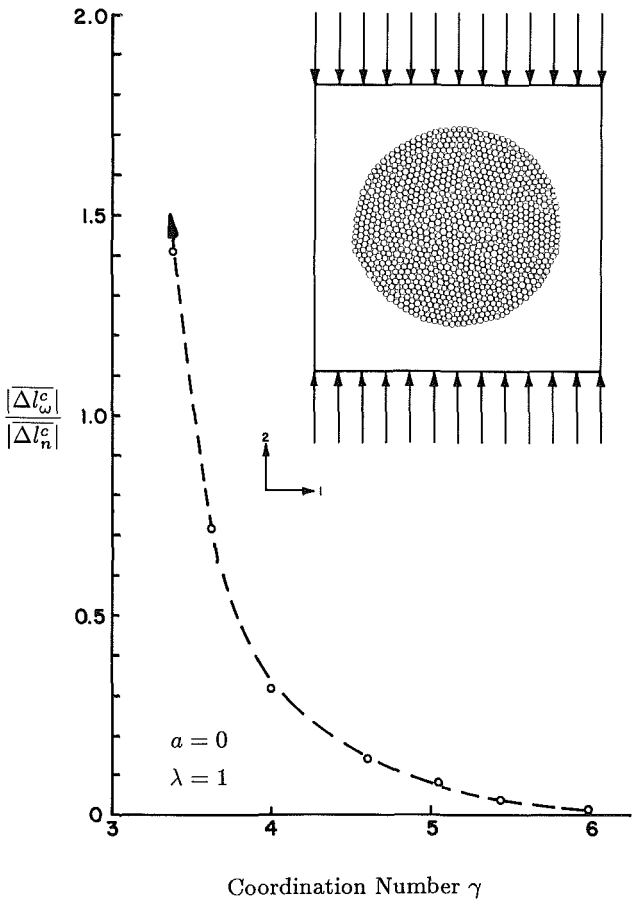


Fig. 5 Influence of assembly coordination number on average rotational contact displacements

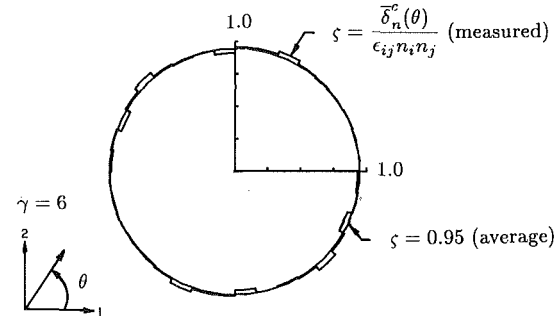


Fig. 6(a) Stiffness reduction coefficient ς from distribution of average normal contact displacements $\bar{\delta}_n^c(\theta)$

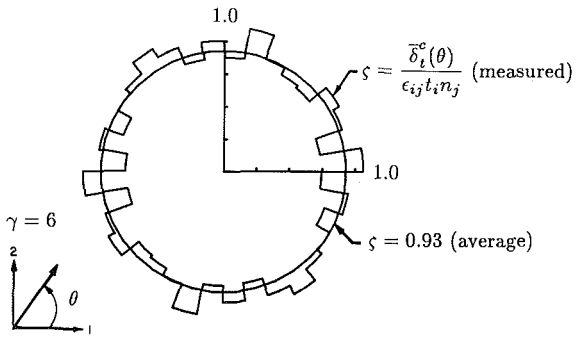


Fig. 6(b) Stiffness reduction coefficient ς from distribution of average tangential contact displacements $\bar{\delta}_t^c(\theta)$

Figs. 3 and 4. Figure 3 shows that the assembly is essentially isotropic with respect to the *second-order* distribution of contact normals (i.e., $a \approx 0$). Some anisotropy in higher-order microstructure is evident from the figure and can be quantified by the fourth-order coefficient of anisotropy as $b = 0.064$. Nevertheless, Bathurst (1985) has shown from the results of similar numerical experiments on *cohesionless* disc assemblies that coefficient terms greater than order two in equation (2) do not significantly influence stress quantities when *anisotropic* distributions for $E(\theta)$ are considered in equations (7) and (8). Isotropic microstructure with respect to the distribution of contact lengths is clearly evident from Fig. 4.

Test Program. A series of numerical simulations were undertaken to verify assumptions (14) and Poisson's ratio expression in terms of interparticle stiffness ratio (18).

Assemblies comprising 1000 discs were subjected to biaxial compression by imposing at the sample boundary discrete forces approximating the stress state $\sigma_{22} > 0$, $\sigma_{11} = 0$ and $\sigma_{12} = \sigma_{21} = 0$. Under these conditions Poisson's ratio could be calculated directly by measuring the resulting principal strain ratio.

Disc interactions in this investigation were controlled by the linear force-displacement laws given in expressions (10). The ratio of interparticle stiffnesses was kept constant for all contacts but was varied between tests over the range $0 \leq \lambda \leq 1$. A stiffness ratio of unity represents a lower limit on the ratio of tangential to normal compliances for elastic spheres in contact according to Mindlin (1949). Truly elastic spheres or discs interact in a nonlinear manner but a linear spring model is useful for verification of theoretical concepts. It is expected that nonlinear contact interactions would lead to other qualitative effects. This topic is currently under investigation by the authors.

Test Results. The influence of coordination number (i.e., system density) on shear displacements generated through particle rotations can be appreciated from Fig. 5. The figure shows that the relative magnitude of $|\overline{\Delta l_\omega^c}|$ for all contacts increases dramatically for assemblies as $\gamma \to 3$. Plane assemblies with coordination number lower than 3 cannot generally be maintained in static equilibrium. For assemblies with a coordination number close to 3, average particle rotations are large, reflecting the freedom afforded interparticle deformations by low system density. Conversely, the magnitude of average particle rotations reduces to zero as the maximum coordination number of 6 is approached. In this case, particles are constrained to the point that their rotations virtually disappear.

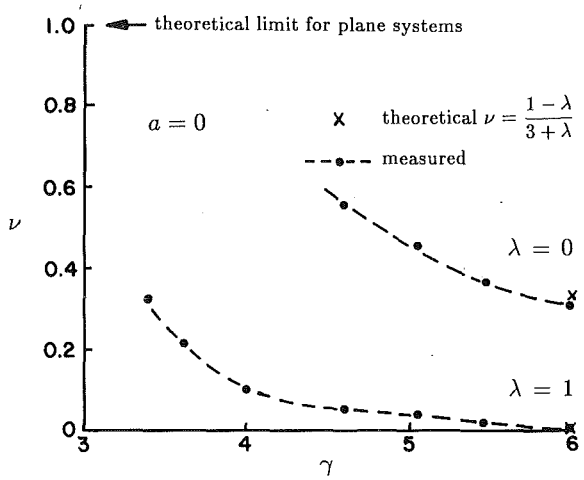


Fig. 7 Measured Poisson's ratio ν versus coordination number γ

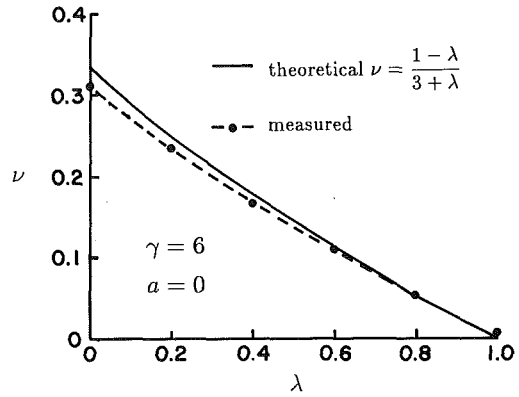


Fig. 8 Comparison of measured and theoretical values of Poisson's ration ν against stiffness ratio λ

A major concept that allows stress-strain relationships for the considered systems to be developed is obtained in equations (14). It is through this relationship that the link between average interparticle displacements and macroscopic strain is made. Relationship (14) implies that ζ is a direction-independent constant. Figures 6(*a*) and 6(*b*) plot ratios $\bar{\delta}_n^c(\theta)/(\epsilon_{ij}n_in_j)$ and $\bar{\delta}_t^c(\theta)/(\epsilon_{ij}t_in_j)$ as polar histograms. It is clear from the figures that these ratios are direction-independent and nearly equal for both normal and tangential displacements and hence verify the fundamental assumption contained in expression (14).

The Poisson's ratio expression (18) predicts values of 0 and 1/3 for stiffness ratios 1 and 0, respectively, provided particle rotations are prohibited. The deviation from predicted values for assemblies with $3 \le \gamma \le 6$ is shown in Fig. 7. It should be noted that the predicted values of Poisson's ratio are generally lower than measured values from numerical simulations. The discrepancy is observed to increase with the magnitude of particle rotations. This component of shear deformations was essentially neglected in theoretical developments. Theoretical expressions, therefore, overpredict actual shear stiffness while correctly predicting bulk modulus. In general, this situation leads to an underestimate of Poisson's ratio. Nevertheless, Fig. 7 shows that as $\gamma \to 6$, theoretically predicted values for Poisson's ratio emerge. The data on this figure for $\lambda = 0$ is restricted to tests with $\gamma \ge 4.5$ corresponding to the range of stable numerical results. For less dense systems, the number of particles in unstable configurations was great enough to prevent the entire assembly from approaching static equilibrium within a reasonable number of calculation cycles.

The results of a series of tests with $\gamma = 6$ and $0 \le \lambda \le 1$ are given in Fig. 8. The data show that relationship (18) gives a reasonable estimate of Poisson's ratio for these systems.

Implications to Three-Dimensional Systems

Numerical simulation of two-dimensional assemblies of discs can be thought of as an analogue to idealized assemblies of *spheres* having variable radius and interacting through linear compliant *fixed* contacts. Unfortunately, numerical simulation of these systems is prohibitively expensive for assemblies having a statistically meaningful number of particles. Nevertheless, the theoretical developments leading to the Poisson's ratio expression for two-dimensional systems are analogous to the approach which can be adopted to arrive at a similar expression for dense three-dimensional assemblies. For example, if three-dimensional assemblies of spheres with *isotropic* fabric are considered, then $E(\Omega) = 1/4\pi$, $\bar{l}^c(\Omega) = \bar{l}_o$ and the stress tensor expression become:

$$\sigma_{ij} = \frac{m_v \bar{l}_o}{4\pi} \int_\Omega \bar{f}_i^c(\Omega) n_j(\Omega) d\Omega \qquad i,j = 1, 2, 3 \tag{19}$$

Here, $d\Omega = \sin\beta d\beta d\theta$ corresponding to the unit spherical coordinate system with $0 \le \beta \le \pi$ and $0 \le \theta \le 2\pi$. Components of the normal vector **n** are related to the unit spherical coordinate system according to:

$$\begin{aligned} n_1 &= \sin\beta\sin\theta \\ n_2 &= \cos\beta \\ n_3 &= \sin\beta\cos\theta \end{aligned} \tag{20}$$

Rothenburg and Selvadurai (1981b) have shown that if particle rotations are negligible, average contact force components in relation (19) can be equated to the strain tensor according to:

$$\bar{f}_i^c(\Omega) = \zeta k_n \{\lambda \epsilon_{il} n_l + (1-\lambda)(\epsilon_{kl} n_l n_k) n_i\} \quad i,k,l = 1,2,3 \tag{21}$$

Substitution of equations (20) and (21) into equation (19) leads to an expression for Poisson's ratio as follows:

$$\nu = \frac{1-\lambda}{4+\lambda} \tag{22}$$

Based on experience from numerical simulation of two-dimensional systems, it is reasonable to expect that this relationship is valid for dense assemblies of spheres with a fixed system of contacts and linear contact interactions. If these assemblies are restricted to a narrow range of particle sizes, then dense assemblies would correspond to systems with $\gamma \to 12$. Equation (22) is in agreement with the value of $\nu = 1/4$ established by Poisson (e.g., Love, 1926) for random assemblies of spheres with central interactions (i.e., $\lambda = 0$).

Acknowledgments

The authors would like to acknowledge the contribution of J. DiPietrantonio, who drafted many of the figures.

References

Batchelor, G. K., and O'Brien, R. W., 1977, "Thermal or Electrical Conduction Through a Granular Material," *Proc. Royal Soc. London, A355,* pp. 313–333.

Bathurst, R. J., 1985, "A Study of Stress and Anisotropy in Idealized Granular Assemblies," Ph.D. Thesis, Civil Engineering Dept., Queen's University at Kingston, Kingston, Ontario, Canada.

Christoffersen, J., Mehrabadi, M. M., and Nemat-Nasser, S., 1981, "A

Micromechanical Description of Granular Material Behaviour," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 339–344.

Dantu, P., 1958, "Etude Statistique des Forces Intergranulaires Dans un Milieu Pulvérulent," *Geotéchnique*, Vol. 18, pp. 50–55.

Horne, M. R., 1965, "The Behavior of an Assembly of Rotund, Rigid, Cohesionless Particles I and II," *Proc. Royal Soc. of London*, Vol. 286, pp. 62–97.

Love, A. E. H., 1926, *Mathematical Theory of Elasticity*, Oxford University Press, Oxford

Mindlin, R. D., 1949, "Compliance of Elastic Bodies in Contact," ASME JOURNAL OF APPLIED MECHANICS, Vol. 16, No. 3, pp. 259–268.

Rothenburg, L., 1980, "Micromechanics of Idealized Granular Systems," Ph.D. Thesis, Civil Engineering Dept., Carleton University, Ottawa, Ontario, Canada.

Rothenburg, L., and Selvadurai, A. P. S., 1981a, "A Micromechanical Definition of the Cauchy Stress Tensor for Particulate Media," *Proc. Int. Symposium on the Mechanical Behavior of Structured Media*, A.P.S. Selvadurai, ed., Ottawa, Ontario, Canada.

Rothenburg, L., and Selvadurai, A. P. S., 1981b, "Micromechanical Aspects of Random Assemblies of Spheres with Linear Contact Interactions," *Eighth Canandian Congress of Applied Mechanics*, University of Moncton, Moncton, pp. 217–218.

Strack, O. D. L., and Cundall, P. A., 1978, "The Distinct Element Method as a Tool for Research in Granular Media, *NSF Report ENG 76-20711, PART I*, Dept. Civil and Mining Engineering, University of Minnesota.

Weber, J., 1966, "Recherches Concernant les Contraintes Integranulaires dans les Milieux Pulvérulents," *Bull. de Liais. Ponts et Chaussées*, No. 20.

A Microcrack Model of Dilatancy in Brittle Materials

S. Nemat-Nasser
Professor of Applied Mechanics and Engineering Sciences,
University of California, San Diego,
La Jolla, CA 92093
Fellow ASME

M. Obata
Post Doctoral Research Fellow,
Department of Applied Mechanics and Engineering Sciences,
University of California, San Diego,
La Jolla, CA 92093

For a solid containing preexisting flaws, overall nonlinear constitutive relations are developed on the basis of a model which endows a preexisting flaw with frictional and cohesive resistance, and which includes nucleation and growth of tension cracks at the preexisting flaw, as it deforms under the action of an overall compressive load. The preexisting flaws may be randomly distributed or may have an initial preferential distribution. They may be of varying sizes and orientations. Even when the flaws are randomly distributed, their preferential activation, and the nucleation and growth of tension cracks at preferential flaws, render the overall response of the solid highly anisotropic. As a first step toward a more complete constitutive micromechanical modeling, a dilute distribution of preexisting flaws is assumed, rate constitutive relations are developed for loading and unloading, which include hysteresis, dilatancy, and other characteristics observed experimentally in rocks, ceramics, concrete, and similar brittle materials. A number of illustrative examples are worked out, and the results are compared to relevant experimental observations.

1 Introduction

Ceramics, rocks, concrete, and similar brittle materials have a highly nonlinear and complex overall response to applied loads. This includes load-induced anisotropy, hysteresis, dilatancy, and strongly path-dependent stress-strain relations. Such complex behavior, to a great extent, stems from the activation of microdefects which are commonly present in these materials. The defects may be preexisting cracks, cavities, soft or hard inclusions, etc. These defects serve as stress concentrators and, therefore, they locally change the state of stress, leading to the formation of tension cracks, even under overall compressive loads. It has been demonstrated experimentally and supported by theoretical models that tension cracks of this kind nucleate and grow in a highly preferential manner and, therefore, even when the microdefects are randomly distributed, so that initially the solid is basically isotropic, the response of the solid becomes highly anisotropic and stress-path-dependent, as loading continues (see Horii and Nemat-Nasser, 1982). The failure modes in compression of solids of this kind have been extensively studied and modeled analytically as well as illustrated by model experiments (Brace and Bombolakis, 1963; McClintock and Walsh, 1963; Hoek and Bieniawski, 1965; Scholz, 1968; Scholz and Kranz, 1974; Zoback and Byerlee, 1975; Holcomb, 1978; Holzhausen, 1978; Paterson, 1978; Kachanov, 1982; Moss and Gupta, 1982; Nemat-Nasser and Horii, 1982; Kranz, 1983; and Horii and Nemat-Nasser, 1985a, 1986). In particular, it has been shown by Nemat-Nasser and Horii (1982) and Horii and Nemat-Nasser (1985a, 1986) that axial splitting, often observed under uniaxial compression, can be explained in terms of a model which considers an isolated preexisting thin straight flaw endowed with frictional and cohesive resistance, which nucleates tension cracks at its tips under the action of axial compression. The model of a sliding crack with frictional resistance (but without cohesive resistance) has been criticized by some authors (e.g., Holcomb and Stevens, 1980; Janach and Guex, 1980; Costin, 1983) on the grounds that actual electron microscopic observations do not seem to support the presence of such cracks, but rather show a complex pattern of axial tension cracks emanating from a variety of sources. There are, however, a number of reasons which seem to support the usefulness of the basic model for the micromechanical constitutive formulation of the inelastic response and the failure modes of materials of this kind. These include: (1) The criticism of the sliding crack model has been based on *qualitative* aspects of microscopic observations. (2) Recent data by Sondergeld and Estey (1982) and Yanagidani et al. (1985) suggest that the focal mechanism for stress induced acoustic emission involves double couples and hence, shear motion. (3) Recent quantitative stereologic evaluation of SEM analysis of rocks by Wong (1985) shows that the results obtained from the sliding crack model are not in conflict with SEM observations. Furthermore, recognizing the limitations of the model (e.g., two-dimensionality), Wong comments that the agreement between the theoretical prediction and the quantitative microscopy results is quite remarkable. (4) Recently, Scholz et al. (1986) have reexamined Bridgman's ring experiment by subjecting a fully jacketed ring of pyrex glass tightly fitted over a polished and hardened steel rod, to overall hydrostatic pressure, and observed isolated axial

[1] Currently Assistant Professor, Department of Civil Engineering, Nagoya Institute of Technology, Nagoya, Japan.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 17, 1986; final revision August 21, 1987.

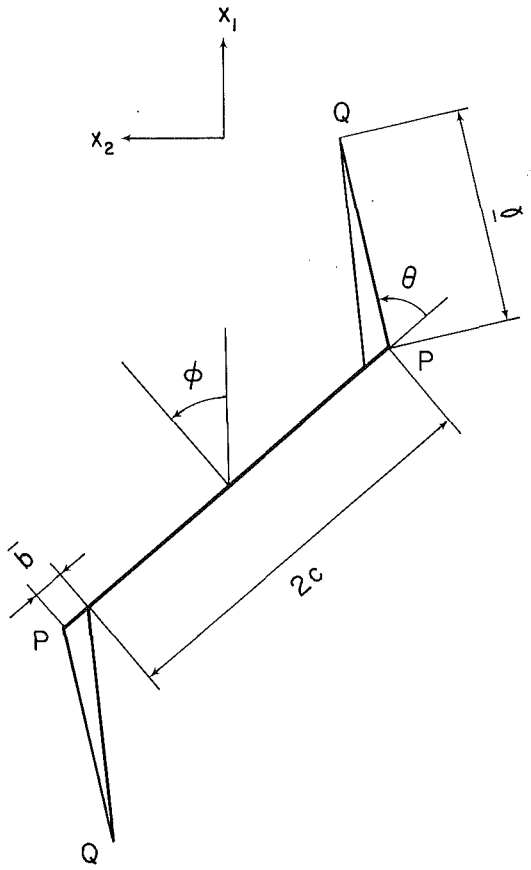


Fig. 1 An isolated preexisting flaw PP' with tension cracks $P'Q'$ and PQ

cracks growing in a stable manner from interior tiny flaws of no greater than 20 μm. Analysis based on Nemat-Nasser and Horii's (1982) sliding crack formulation showed that under the prevailing all-around compressive stresses in this experiment, microflaws of around 10 μm are sufficient to initiate such axial cracks. Scholz et al. (1986) conclude that "axial cracks can be initiated by tiny flaws and grow stably for long distances under an overall compressive stress state. Thus shear cracks or other types of stress concentrations would not necessarily be expected to be prominent in SEM studies. In a heterogeneous material such as rock, where stress concentrations like those we observed at the ring-rod interface, can be expected to occur commonly at grain boundaries, it is not surprising that axial cracking predominates." In view of these observations, and considering the resolution limit of the SEM (about 0.02 μm), criticism of the sliding crack model based on the qualitative examination of SEM results seems to require reevaluation.

Notwithstanding the above comments, we do not intend to suggest the sliding crack mechanism as the only source of axial cracking. Indeed, model studies (Nemat-Nasser, 1985) have shown microcracking under axial compression from inclusions and cavities of various geometric shapes. What we intend to do is to consider the sliding crack model as a representative of a thin flaw which can deform plastically (and hence has cohesive resistance) and can slide (and hence has frictional resistance), and in a systematic and rigorous manner, examine the results. These results then can be used to study the merits or shortcomings of the basic mode.

To this end, we shall consider a solid which contains a dilute distribution of such ideal flaws. The flaws may be of varying sizes and orientations which may have either random or preferential distributions. The matrix is assumed to be isotropically elastic. Inelasticity and anisotropy develop because of the frictional and cohesive resistance of the preexisting flaws, and because of the formation and growth of tension cracks at the tips of the flaws. During unloading, the effects of frictional and cohesive locking of the flaws, their backsliding upon further unloading, the partial closure of tension cracks, and finally the residual strains are included. All calculations are made for a two-dimensional model. Once the basic parameters of the model are defined, then the problem is analyzed consistently and systematically, without introducing additional assumptions. This allows us to make a judgment on the merits of the model from the obtained results which are then a true reflection of the basic model's assumptions.

2 Formulation

We consider a micromechanical modeling of the inelastic response of brittle materials such as ceramics and rocks at relatively low temperatures and under moderate confining pressures, so that: (1) rate effects can be ignored, and (2) microcracking and inelastic slip at preexisting flaws can be regarded as the basic micromechanisms giving rise to inelasticity; for rather detailed discussions and extensive lists of relevant references, see Paterson (1978), Krantz (1983), and Horii and Nemat-Nasser (1986). The basic building block of the model is the mechanism shown in Fig. 1. It consists of an isolated preexisting flaw PP' and tension cracks PQ and $P'Q'$ embedded in an isotropically elastic unbounded solid. Under the action of farfield compressive stresses, the preexisting flaw PP' undergoes frictional sliding which may be accompanied by cohesive (due to possible plastic deformation of the flaw) resistance as well as, possibly, dilatancy normal to PP' due to the deformation of the flaw and the possible presence of asperities. The relative deformation of the two faces of the flaw results in the creation of high tensile fields at the flaw's tips and leads to the formation of tension cracks which grow toward the direction of maximum compression. The model of frictional sliding of preexisting cracks (without cohesion or dilatancy) which leads to the formation of tension cracks at its tips was originally suggested by Brace and Bombolakis (1963) and has been studied both by model experiments and analytically by a number of investigators (McClintock and Walsh, 1963; Holzhauzen, 1978; Moss and Gupta, 1982; Kachanov, 1982; Nemat-Nasser and Horii, 1982; Horii and Nemat-Nasser, 1985a, b, 1986; and Steif, 1984). Although micromechanical observations on actual rocks have not clearly identified preexisting cracks as the basic source of inelasticity in rocks, recent analysis by Wong (1985) seems to suggest that the model is indeed viable and does capture both qualitatively and quantitatively some of the main features of the problem.

In this section we shall calculate the inelastic strain due to: (1) slip along the preexisting flaw, (2) dilatancy normal to the flaw, and (3) the associated out-of-plane tension cracks at the tips of the flaw. The dilatancy normal to the flaw will be assumed to be very small, but its effect on creating tension cracks at the tips of the flaw is significant.

The total strain at each state depends on the parameters that define the basic model shown in Fig. 1. The values of these parameters are highly stress-path-dependent. Hence, an incremental solution is necessary. Using the computed strain tensor, we formulate a rate-constitutive relation, using the equilibrium conditions for a representative flaw and the associated tension cracks.

Throughout this paper we shall use a fixed rectangular Cartesian coordinate system, x_1, x_2, as the background frame of reference. To simplify notation, we use the following direct representation:

$$\mathbf{a}:\mathbf{b}=a_{ij}b_{ij},$$
$$\mathbf{L}:\mathbf{b}=L_{ijkl}b_{kl}, \tag{2.1}$$
$$\mathbf{a}\otimes\mathbf{b}=a_{ij}b_{kl},\ i,j,k,l=1,2,$$

where repeated indices are summed.

2.1 Description of Strain. The total strain consists of an elastic strain due to the elastic deformation of the matrix, plus an inelastic strain due to slip and dilatancy of the preexisting flaws, accompanied by inelastic strain due to microcracking. In the sequel all physical quantities with the dimension of length are normalized by dividing by half of the preexisting flaw length, c. For example, if $\bar{b}$ is the slip (assumed constant) along flaw PP', we use $b = \bar{b}/c$ to represent the dimensionless slip. Similarly, if $\bar{d}$ is the displacement (assumed constant over the flaw length) of the flaw normal to its face, we use $d = \bar{d}/c$.

Let

$$f=c^2N \tag{2.2}$$

represent the density of the flaws of length $2c$ and orientation ϕ, where N is the number of such flaws per unit area. The strain due only to slip b of these flaws then is

$$\epsilon_b=2fb\mathbf{p}_0, \tag{2.3a}$$

where

$$\mathbf{p}_0=\frac{1}{2}\begin{bmatrix}-\sin 2\phi & \cos 2\phi\\ \cos 2\phi & \sin 2\phi\end{bmatrix}. \tag{2.3b}$$

Similarly, the strain due to dilatancy d normal to PP' is given by

$$\epsilon_d=2fd\mathbf{p}_1 \tag{2.4a}$$

where

$$\mathbf{p}_1=\begin{bmatrix}\cos^2\phi & \cos\phi\sin\phi\\ \cos\phi\sin\phi & \sin^2\phi\end{bmatrix}; \tag{2.4b}$$

see Appendix A. Since the sliding-induced opening of a preexisting flaw normal to the surface must be bounded, we shall assume that this opening becomes saturated and hence use

$$d=d_0(1-e^{-\gamma b}). \tag{2.5}$$

Note that b (and hence d) is assumed to be uniform along the flaw. This simplification is reasonable once tension cracks are suitably long.

Slip b and dilatancy d introduce normal and tangential concentrated gaps of values ($b\sin\theta + d\cos\theta$) and ($b\cos\theta - d\sin\theta$), at points P and P', for cracks PQ and $P'Q'$, of common length ℓ, and orientation θ with respect to the preexisting flaw PP'. The strain associated with these gaps can be computed and is given by

$$\epsilon_c=fb\ell\mathbf{q}_0+fd\ell\mathbf{q}_1, \tag{2.6a}$$

where

$$\mathbf{q}_0=\frac{1}{2}\begin{bmatrix}-2\sin\phi\cos(\theta+\phi) & \cos(\theta+2\phi)\\ \cos(\theta+2\phi) & 2\cos\phi\sin(\theta+\phi)\end{bmatrix}, \tag{2.6b}$$

$$\mathbf{q}_1=\frac{1}{2}\begin{bmatrix}2\cos\phi\cos(\theta+\phi) & \sin(\theta+2\phi)\\ \sin(\theta+2\phi) & 2\sin\phi\sin(\theta+\phi)\end{bmatrix}; \tag{2.6c}$$

see Appendix B.

The applied stresses change the crack opening displacement. This is not included in expression (2.6a). Under the condition of a positive Mode I stress intensity factor at Q and Q' (crack closure will be considered later on), i.e., for $K_I > 0$, the strain

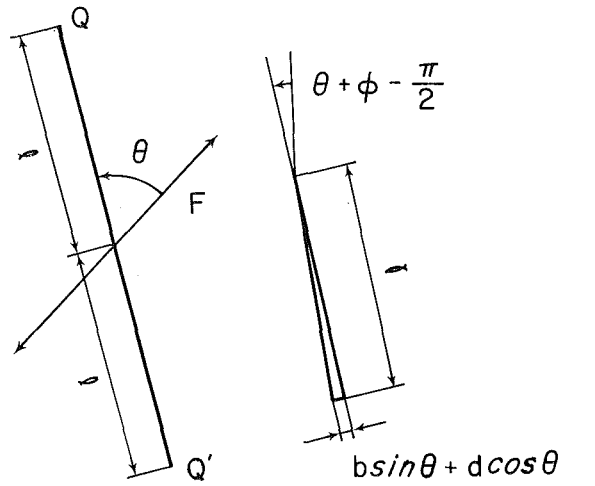


Fig. 2 (*a*) An equivalent isolated crack with a pair of collinear concentrated forces applied at its center. (*b*) An equivalent isolated crack with concentrated gap at its end.

due to the change in crack opening caused by the applied loads can be computed and is given as follows:

$$\epsilon_a=f\frac{1-\nu}{8\mu}\pi\ell^2[4(\sigma:\alpha)\alpha+(\sigma:\beta)\beta], \tag{2.7a}$$

where

$$\alpha=\begin{bmatrix}\cos^2(\theta+\phi) & \cos(\theta+\phi)\sin(\theta+\phi)\\ \sin(\theta+\phi)\cos(\theta+\phi) & \sin^2(\theta+\phi)\end{bmatrix}, \tag{2.7b}$$

$$\beta=\frac{\partial\alpha}{\partial\theta}=\begin{bmatrix}-\sin 2(\theta+\phi) & \cos 2(\theta+\phi)\\ \cos 2(\theta+\phi) & \sin 2(\theta+\phi)\end{bmatrix}; \tag{2.7c}$$

see Appendix C. In equation (2.7a), μ is the shear modulus, ν is Poisson's ratio, and σ is the stress tensor.

With $\mathbf{L}$ denoting the elastic modulus tensor of the matrix material, the total strain tensor is now given by

$$\begin{aligned}\epsilon=&\mathbf{L}^{-1}:\sigma\\ &+2fb\mathbf{p}_0+2fd\mathbf{p}_1\\ &+fb\ell\mathbf{q}_0+fd\ell\mathbf{q}_1\\ &+f\frac{1-\nu}{8\mu}\pi\ell^2[4\sigma:\alpha)\alpha+(\sigma:\beta)\beta].\end{aligned} \tag{2.8}$$

2.2 Rate Constitutive Relations. To complete the formulation, we must relate the kinematic parameters b, d, ℓ, and θ to the applied stress σ. We shall do this by calculating the opening mode stress intensity factor, K_I, at crack tips Q and Q' in two different ways: (1) by considering the applied loads and the condition of slip across the flaw (see Fig. 2(a)), as discussed by Horii and Nemat-Nasser (1986), and (2) by calculating the stress intensity factor in the presence of the applied stress and the imposed concentrated gaps at P and P' (see Fig. 2(b)).

As has been shown and verified by numerical examples by Horii and Nemat-Nasser (1986), the Mode I stress intensity factor at Q and Q' of cracks PQ and $P'Q'$ can be computed with very good accuracy by considering an equivalent crack of length 2ℓ, subjected to a pair of collinear concentrated forces, F, as well as the applied overall stresses, as shown in Fig. 2. In this representation, F denotes the driving force transmitted across the preexisting flaw; it lies along the flaw and equals in magnitude the resultant force transmitted across the flaw.

Following the procedure outlined by Horii and Nemat-Nasser (1986), we obtain

$$K_{\mathrm{I}}=\frac{2}{\sqrt{\pi(\ell+\ell^*)}}(\boldsymbol{\sigma}:\mathbf{p}_2-\tau_c)\sin\theta+\boldsymbol{\sigma}:\boldsymbol{\alpha}\sqrt{\pi\ell}, \tag{2.9a}$$

where

$$\mathbf{p}_2=\mathbf{p}_0+\bar{\mu}\mathbf{p}_1,\ \ell^*=0.27,\ \tau_c=\tau_c^0 e^{-\eta|b|},\ |b|\equiv\int_0^t|\dot{b}|\,dt. \tag{2.9b-d}$$

Here, $\bar{\mu}$ is the coefficient of friction, τ_c is the value of cohesion at the current value of the total absolute value of slip (we shall assume that cohesion is an exponentially decaying function of the *accumulated* slip with decaying factor η defined by equation (2.9*d*)), and τ_c^0 is the initial cohesion. The parameter ℓ^* is introduced in such a manner that equation (2.9*a*) is valid even at crack initiation, i.e., at $\ell = 0^+$. Since the opening mode stress intensity factor for a closed sliding crack is maximum in a direction which makes an angle $\theta_c = 0.39\pi$ with the crack orientation, ℓ^* is obtained by equating the stress intensity factor given by equation (2.9*a*) at $\ell = 0$ and $\theta = \theta_c$ to that maximum value. In the present application the effect of tension cracks are important once they become relatively long. Hence, whether or not ℓ^* is included in equation (2.9*a*) does not affect the results, although its presence renders K_{I} non-singular at $\ell = 0$.

The Mode I stress intensity factor associated with the gaps and in the presence of applied stresses, on the other hand, is given by

$$K_{\mathrm{I}}'=\frac{\mu}{1-\nu}\frac{1}{\sqrt{2\pi(\ell+\ell^{**})}}(b\sin\theta+d\cos\theta)+\boldsymbol{\sigma}:\boldsymbol{\alpha}\sqrt{\frac{\pi\ell}{2}}, \tag{2.10a}$$

where

$$\ell^{**}=\frac{\pi^2}{32}\ell^*; \tag{2.10b}$$

see Appendix D.

To obtain a relation between b and $\boldsymbol{\sigma}$, we require

$$K_{\mathrm{I}}=K_{\mathrm{I}}' \text{ when } \boldsymbol{\sigma}:\mathbf{p}_2-\tau_c\geq 0. \tag{2.11}$$

Furthermore, the crack length ℓ is obtained such that

$$K_{\mathrm{I}}=K_c, \tag{2.12}$$

where K_c is the critical value of the stress intensity factor in the opening mode. θ is determined so as to maximize K_{I} for a given length ℓ.

It should be noted that model experiments (Nemat-Nasser and Horii, 1982) and exact analysis (Horii and Nemat-Nasser, 1983, 1985a) show that tension cracks emanate from the flaw tips at an angle of about 70 deg with respect to the orientation of the flaw, and curve toward the direction of the maximum compression. The model shown in Fig. 1 substitutes the curved crack by an *equivalent* straight one; the equivalence is in the sense that the orientation θ of the *equivalent straight crack* is computed at each incremental loading by maximizing K_{I} with respect to θ.

Because of the complex and nonlinear relations that exist among various parameters which defined the basic model shown in Fig. 1, the overall stress-strain relation will be highly nonlinear and stress-path-dependent. It is, therefore, necessary to formulate the problem incrementally. To this end we will express $\dot{\ell}$, $\dot{\theta}$, and $\dot{b}$, in terms of $\dot{\boldsymbol{\sigma}}$, for given loading regimes. These quantities can be written symbolically as

$$\dot{\ell}=\mathbf{A}:\dot{\boldsymbol{\sigma}},\ \dot{\theta}=\mathbf{B}:\dot{\boldsymbol{\sigma}},\ \dot{b}=\mathbf{C}:\dot{\boldsymbol{\sigma}}. \tag{2.13a-c}$$

We will calculate the second-order tensors **A**, **B**, and **C** for prescribed loading regimes. Note that, in view of assumption (2.5), we have

$$\dot{d}=\zeta(d_0-d)\dot{b}. \tag{2.14}$$

With equations (2.13) and (2.14), and upon time differentiation of equation (2.8), we obtain the following rate constitutive equation:

$$\begin{aligned}\dot{\boldsymbol{\epsilon}}&=\mathbf{D}:\dot{\boldsymbol{\sigma}},\\ \mathbf{D}&=\mathbf{L}^{-1}\\ &+f\left\{b\mathbf{q}_0+d\mathbf{q}_1+\frac{1-\nu}{4\mu}\pi\ell[4(\boldsymbol{\sigma}:\boldsymbol{\alpha})\boldsymbol{\alpha}+(\boldsymbol{\sigma}:\boldsymbol{\beta})\boldsymbol{\beta}]\right\}\otimes\mathbf{A}\\ &+f\left\{b\ell\mathbf{r}_0+d\ell\mathbf{r}_1+\frac{1-\nu}{8\mu}\pi\ell^2[4(\boldsymbol{\sigma}:\boldsymbol{\beta})\boldsymbol{\alpha}+4(\boldsymbol{\sigma}:\boldsymbol{\alpha})\boldsymbol{\beta}\right.\\ &\quad\left.+(\boldsymbol{\sigma}:\boldsymbol{\gamma})\boldsymbol{\beta}+(\boldsymbol{\sigma}:\boldsymbol{\beta})\boldsymbol{\gamma}]\right\}\otimes\mathbf{B}\\ &+f\{2\mathbf{p}_0+\ell\mathbf{q}_0+\zeta(d_0-d)(2\mathbf{p}_1+\ell\mathbf{q}_1)\}\otimes\mathbf{C}\\ &+f\frac{1-\nu}{8\mu}\pi\ell^2[4\boldsymbol{\alpha}\otimes\boldsymbol{\alpha}+\boldsymbol{\beta}\otimes\boldsymbol{\beta}],\end{aligned} \tag{2.15a}$$

where

$$\boldsymbol{\gamma}=\frac{\partial\boldsymbol{\beta}}{\partial\theta},\ \mathbf{r}_0=\frac{\partial\mathbf{q}_0}{\partial\theta},\ \mathbf{r}_1=\frac{\partial\mathbf{q}_1}{\partial\theta}. \tag{2.15b-d}$$

2.3 Specific Cases During Loading. In what follows, we consider various loading cases separately.

(*i*) *Sliding Occurs Without Crack Growth.* This happens when the driving shear stress on the flaw exceeds the frictional and cohesive resistance of the flaw, i.e., when $\boldsymbol{\sigma}:\mathbf{p}_2 - \tau_c \geq 0$, while $K_{\mathrm{I}} < K_c$. From condition $K_{\mathrm{I}} = K_{\mathrm{I}}'$, we obtain, by time differentiation,

$$\left[\frac{\partial K_{\mathrm{I}}}{\partial b}-\frac{\partial K_{\mathrm{I}}'}{\partial b}\right]\dot{b}=(\mathbf{b}_1-\mathbf{b}_2):\dot{\boldsymbol{\sigma}}, \tag{2.16}$$

where $\mathbf{b}_1$ and $\mathbf{b}_2$ are the matrix coefficients of $\dot{\boldsymbol{\sigma}}$, obtained from equation (2.9*a*) and from equation (2.10*a*), respectively. Thus we have

$$\dot{b}=\frac{1}{\dfrac{\partial K_{\mathrm{I}}}{\partial b}-\dfrac{\partial K_{\mathrm{I}}'}{\partial b}}(\mathbf{b}_1-\mathbf{b}_2):\dot{\boldsymbol{\sigma}}, \tag{2.17}$$

and, hence,

$$\mathbf{C}=\frac{1}{\dfrac{\partial K_{\mathrm{I}}}{\partial b}-\dfrac{\partial K_{\mathrm{I}}'}{\partial b}}(\mathbf{b}_1-\mathbf{b}_2),\ \mathbf{A}=\mathbf{0},\ \mathbf{B}=\mathbf{0}. \tag{2.18a-c}$$

(*ii*) *Sliding Occurs With Crack Growth.* This happens when $K_{\mathrm{I}} = K_c$ and $\boldsymbol{\sigma}:\mathbf{p}_2 - \tau_c \geq 0$. The problem can be stated as follows: Maximize K_{I} with respect to θ and b under the condition that $K_{\mathrm{I}} = K_{\mathrm{I}}'$. With the aid of a Lagrangian multiplier, we arrive at the following conditions:

$$K_{\mathrm{I}}=K_{\mathrm{I}}'=K_c,\qquad \frac{\partial K_{\mathrm{I}}}{\partial\theta}\frac{\partial K_{\mathrm{I}}'}{\partial b}-\frac{\partial K_{\mathrm{I}}}{\partial b}\frac{\partial K_{\mathrm{I}}'}{\partial\theta}\equiv J=0. \tag{2.19a,b}$$

Time differentiation yields,

$$\frac{\partial K_{\mathrm{I}}}{\partial\ell}\dot{\ell}+\frac{\partial K_{\mathrm{I}}}{\partial\theta}\dot{\theta}+\frac{\partial K_{\mathrm{I}}}{\partial b}\dot{b}=\mathbf{b}_1:\dot{\boldsymbol{\sigma}},$$

$$\frac{\partial K_I'}{\partial \ell}\dot{\ell}+\frac{\partial K_I'}{\partial \theta}\dot{\theta}+\frac{\partial K_I'}{\partial b}\dot{b}=\mathbf{b}_2:\dot{\sigma}, \qquad (2.20a\text{–}c)$$

$$\frac{\partial J}{\partial \ell}\dot{\ell}+\frac{\partial J}{\partial \theta}\dot{\theta}+\frac{\partial J}{\partial b}\dot{b}=\mathbf{b}_3:\dot{\sigma}.$$

These linear equations yield $\dot{\ell}$, $\dot{\theta}$, and $\dot{b}$. The solution is in the form of equations (2.13).

2.4 Specific Cases During Unloading. After a continued loading regime, we consider unloading by reducing the applied stresses. Unloading at specific flaws may also take place when, for example, we change the load path. Our analysis of unloading includes all cases when the net shear traction, $\sigma:\mathbf{p}_2 - \tau_c$, transmitted across a preexisting flaw decreases for whatever reason.

(*i*) *Crack Grows Without Sliding.* Upon unloading, cracks may continue to grow, as the lateral compression is reduced but, because of an initial locking of the flaw, the maximum gaps at P and P' remain unchanged. The conditions for crack growth then are

$$K_I'=K_c,\ \frac{\partial K_I'}{\partial \theta}=0. \qquad (2.21a,b)$$

Equation (2.21*b*) ensures that cracks grow in a direction which maximizes K_I'. Time differentiation now yields

$$\frac{\partial K_I'}{\partial \ell}\dot{\ell}+\frac{\partial K_I'}{\partial b}\dot{\theta}=\mathbf{b}_2:\dot{\sigma},$$

$$\frac{\partial}{\partial \ell}\left(\frac{\partial K_I'}{\partial \theta}\right)\dot{\ell}+\frac{\partial}{\partial \theta}\left(\frac{\partial K_I'}{\partial \theta}\right)\dot{\theta}=\frac{\partial}{\partial \theta}\mathbf{b}_2:\dot{\sigma}. \qquad (2.23a\text{–}b)$$

From these we have $\dot{\ell}$ and $\dot{\theta}$ in the form of equations (2.13). Since b remains fixed, in this case $\dot{b}=0$, and hence $\mathbf{C}=\mathbf{0}$.

(*ii*) *Backsliding Occurs Without Crack Closure.* When the residual shear stress on the preexisting flaw overcomes the cohesive and frictional resistance, backsliding of the flaw may take place. Since the frictional and cohesive resistance on the preexisting flaw acts against the direction of the relative displacement, K_I in equation (2.9*a*) must be redefined for the unloading case by

$$K_I=\frac{2}{\sqrt{\pi(\ell+\ell^*)}}(\sigma:\mathbf{p}_2'+\tau_c)\sin\theta+\sigma:\alpha\sqrt{\pi\ell}, \qquad (2.24a)$$

$$\mathbf{p}_2'=\mathbf{p}_0-\bar{\mu}\mathbf{p}_1. \qquad (2.24b)$$

Before backsliding takes place, K_I of equation (2.24*a*) may formally be larger than the *actual* stress intensity factor K_I' obtained on the basis of the applied overall loads and the existing gaps at P and P'. This is similar to the loading case with $\sigma:\mathbf{p}_2 - \tau_c \le 0$, so that no slip occurs. The quantity K_I, defined by equation (2.24*a*), attains its usual physical meaning only when backsliding occurs, or is just about to occur. Therefore, the condition for the initiation of backsliding is given by

$$K_I \le K_I'; \qquad (2.25)$$

see Appendix E. During backsliding, however, we require that

$$K_I=K_I'. \qquad (2.26)$$

Upon time differentiation at fixed ℓ and θ we obtain an equation similar to equation (2.17) for $\dot{b}$.

(*iii*) *Crack Closure Without Backsliding.* When crack closure occurs, the strain tensor is given by

$$\epsilon=\mathbf{L}^{-1}:\sigma + 2fb\mathbf{p}_0+2fd\mathbf{p}_1 - f\frac{1}{2\pi}\frac{\mu}{1-\nu}\frac{1}{(\sigma:\alpha)}(b\sin\theta+d\cos\theta)^2\alpha + f\frac{\ell_m}{2}(b\cos\theta-d\sin\theta)\beta + f\frac{1-\nu}{8\mu}\pi\ell_m^2(\sigma:\beta)\beta, \qquad (2.27)$$

where ℓ_m denotes the maximum crack length (see Appendix F). From the time differentiation of equation (2.27) and equations (2.13), we obtain the rate constitutive equation for this special case in the following explicit form:

$$\dot{\epsilon}=\mathbf{D}:\dot{\sigma},$$

$$\mathbf{D}=\mathbf{L}^{-1} + f[2\mathbf{p}_0+2\zeta(d_0-d)\mathbf{p}_1 - \frac{1}{\pi}\frac{\mu}{1-\nu}\frac{1}{\sigma:\alpha}(\sin\theta + \zeta(d_0-d)\cos\theta)(b\sin\theta+d\cos\theta)\alpha + \frac{\ell_m}{2}\{\cos\theta-\zeta(d_0-d)\sin\theta\}\beta]\otimes\mathbf{C} + f\left\{\frac{1}{2\pi}\frac{\mu}{1-\nu}\frac{1}{(\sigma:\alpha)^2}(b\sin\theta+d\cos\theta)^2\alpha\otimes\alpha + \frac{1-\nu}{8\mu}\pi\ell_m^2\beta\otimes\beta\right\}. \qquad (2.28)$$

Note that θ is fixed in this case.

The condition for crack closure without backsliding of the flaw is

$$K_I'=0, \qquad (2.29)$$

whose time differentiating at constant θ and b yields

$$\dot{\ell}=(\mathbf{b}_2:\dot{\sigma})/\left(\frac{\partial K_I'}{\partial \ell}\right)=\mathbf{A}:\dot{\sigma},\ \mathbf{B}=\mathbf{C}=\mathbf{0}. \qquad (2.30a\text{–}c)$$

(*iv*) *Crack Closure with Backsliding.* The conditions are

$$K_I=K_I'=0. \qquad (2.32a,b)$$

Since θ remains constant during crack closure, only $\dot{\ell}$ and $\dot{b}$ must be expressed in terms of $\dot{\sigma}$. This is done by time differentiation of $K_I=0$ and $K_I'=0$ at constant θ and by solving the resulting linear equations for $\dot{\ell}$ and $\dot{b}$. $\mathbf{B}=\mathbf{0}$ because θ remains constant.

2.4 Comments on the Model. The model presented above includes a number of features which have been suggested by several researchers in the past, e.g., Brace and Bombolakis (1963), Zoback and Byerlee (1975), Kachanov (1982), and Moss and Gupta (1982). In particular, Moss and Gupta have obtained nonlinear stress-strain and dilatancy relations based on a sliding preexisting crack with tension cracks (wing cracks) emanating from its tips. These authors include frictional and cohesive effects, and calculate the inelastic strain associated with the slip of the preexisting crack as well as of the tension cracks. In their model, however, the orientation of the tension cracks (wings) relative to the preexisting crack is fixed throughout the entire loading history, and must be prescribed. Moreover, in their model it is necessary to prescribe an initial length and crack opening displacement for the wings. In the model presented here in our work, the flaws at which tension cracks nucleate, the orientation and the length of the tension cracks, their opening displacement, and whether they continue

to grow, become dormant, or actually close, are all obtained as part of the solution to the problem and, indeed, change during the course of loading and unlolading, in a rather complex manner. Indeed, in our model the tension cracks may continue to grow at a state of unloading, because of the locking of the preexisting flaw, as illustrated in the examples given in Section 4. Furthermore, we have attempted to render the model self-consistent in the sense that the stress intensity factors at the tips of the tension cracks are obtained on the basis of the existing gap (created due to the sliding of the preexisting flaw) and in the presence of farfield loads, and these are then made to correspond to the stress intensity factors calculated on the basis of the applied loads and the effective total force transmitted across the preexisting flaw. In the model of Moss and Gupta, the stress intensity factors are calculated only on the basis of the applied loads and the effective force across the preexisting crack. We note that if the preexisting flaw is locked, then the stress intensity factors must be computed on the basis of the existing gaps and the applied loads, and these stress intensity factors may not be equal to the ones calculated on the basis of the applied loads and the effective force across the preexisting flaw, when the flaw remains dormant.

3 Average Quantities

So far, the problem has been formulated for flaws of fixed orientation ϕ and size $2c$. When a dilute distribution of preexisting flaws is involved, the interaction effects may be neglected, and the overall strains and strain rates may be estimated by a simple average of the individual contributions. To this end let $N = N(\phi)$ be the number of flaws of orientation ϕ and size $2c$ per unit area. Hence, $f = c^2N$ is given as a function of ϕ. The average strain, $\bar{\epsilon}$, then may be computed from

$$\bar{\epsilon} = \frac{1}{2\pi}\int_0^{2\pi} \epsilon(\phi, f(\phi))d\phi, \tag{3.1}$$

where $\epsilon(\phi, f(\phi))$ is given, in our analysis, by equation (2.8) or by equation (2.27) with f properly weighted. A similar expression can be written for the average strain rate. When a finite number of specific orientations of preexisting flaws is involved, the integral in equation (3.1) reduces to a finite sum,

$$\bar{\epsilon} = \frac{1}{M}\sum_{\alpha=1}^{M} \epsilon(\phi_\alpha, f(\phi_\alpha)), \tag{3.2}$$

where M is the total number of involved orientations.

In the formulation presented above and in the preceding section, no attempt is made to include interaction among neighboring flaws and their associated cracks. There are a number of possible ways that this interaction may be included. Perhaps the most effective way would be to estimate the stress intensity factors at the crack tips, by considering several interacting flaws or an infinite row of such flaws, in the manner discussed by Horii and Nemat-Nasser (1983, 1985a, b, 1986). This poses an extremely complex mathematical problem where, although solvable as shown by Horii and Nemat-Nasser (1985b), the resulting equations are so involved that they do not lend themselves to simple analytical calculations of the kind given in the present paper.

An alternative approach is to consider an averaging technique similar to the self-consistent scheme, Budiansky (1965) and Hill (1965). This would require the solution of the problem of out-of-plane crack growth from a preexisting flaw embedded in a linearly elastic but anisotropic material, which currently is not available. Therefore, any improvement on the simple averaging procedure given by equation (3.1) seems to require either extensive and complicated computations beyond the scope of the present work or it must be based on additional simplifying assumptions, which is not our intention at this time.

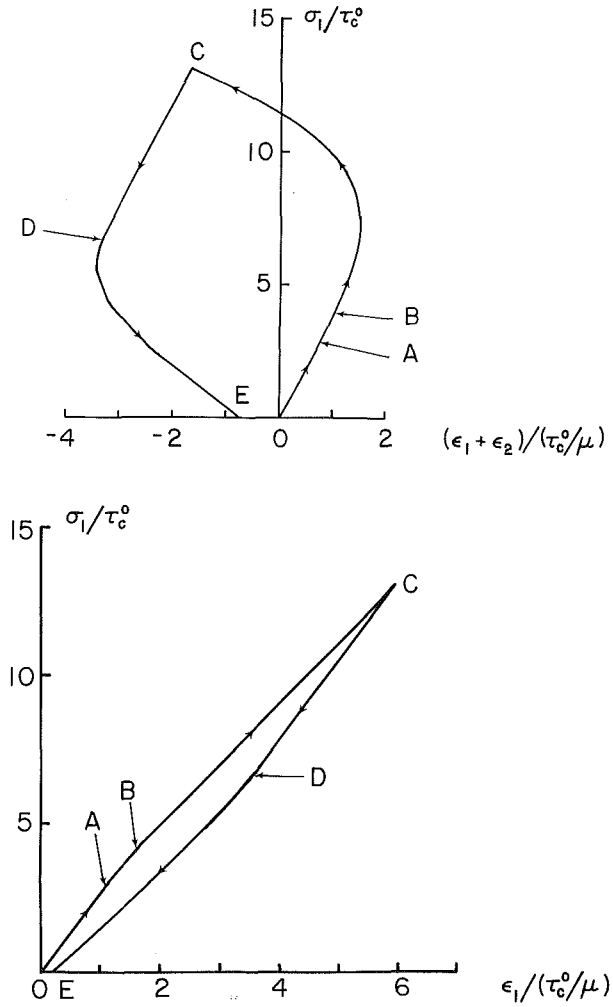


Fig. 3 *(a)* **Volumetric strain versus differential stress curve for a loading and unloading loop. *(b)* Axial strain versus stress curve for a loading and unloading loop; $f = 0.3$, $M_0 = 10$, $\nu = 0.25$, $\bar{\mu} = 0.3$, $K_c = 1.0$, $d_0 = 0.05$, and $\eta = 0.05$.**

In the following section we shall assume a dilute distribution of preexisting flaws, use the simple averaging given by equation (3.2), and present a number of illustrative examples.

4 Results and Discussions

In this section, we shall illustrate the predictive capability of the model presented in the preceding section, in terms of several numerical examples, and discuss the results.

The basic parameters of the model are:

(1) $f = c^2N$ which represents the density and the size of the flaws. It can be given any suitable distribution in terms of the flaw orientation ϕ. In our examples we shall assume it to be constant.

(2) M_0 which is the number of flaw orientations between 0 deg and 90 deg. In our examples we shall assume an isotropic initial flaw distribution, and therefore use finite, equally spaced orientations between 0 deg and 90 deg.

(3) $K_c = \bar{K}_c/(\tau_c^0\sqrt{c})$ which is the dimensionless critical value of the stress intensity factor. Since typical values of fracture toughness, $\bar{K}_c$, for rocks are of the order of $10^5 \sim 10^6$ $Nm^{-3/2}$, taking a typical flaw length to be of the order of 10^{-4}m, and the cohesive stress to be of the order of 10^7 Nm^{-2}, we estimate K_c to be of the order of 1.

(4) d_0 which represents the limiting maximum flaw expansion (dilatancy) normal to its face (see equation (2.5)). It is

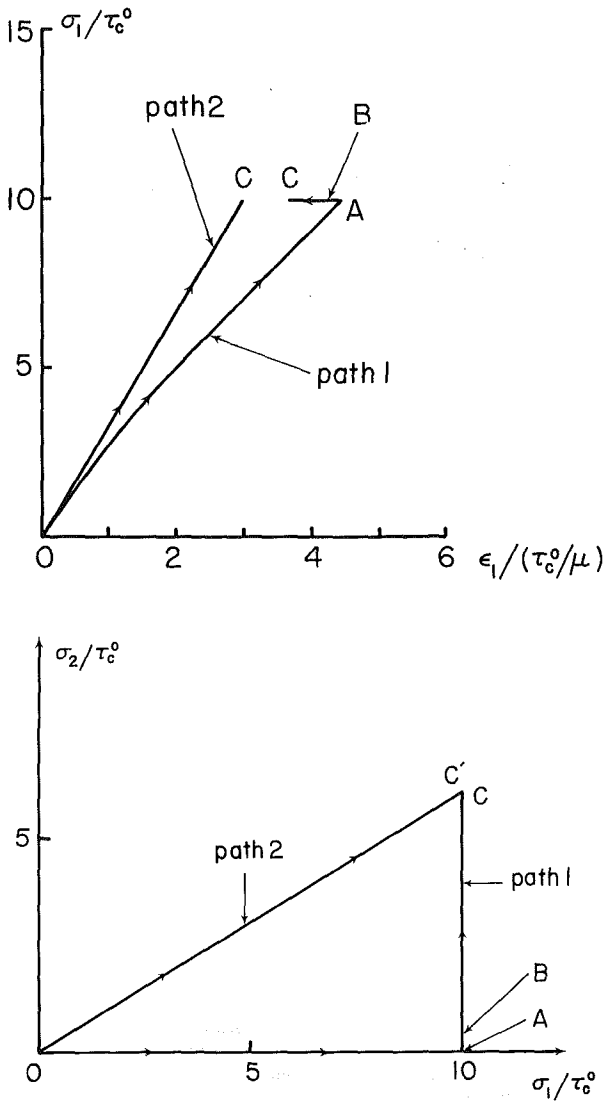


Fig. 4 (a) Axial strain versus differential stress for different loading paths: $f = 0.3$, $M_0 = 10$, $\nu = 0.25$, $\bar{\mu} = 0.3$, $d_0 = 0$, $K_c = 1.0$, and $\eta = 0.05$; (b) loading paths

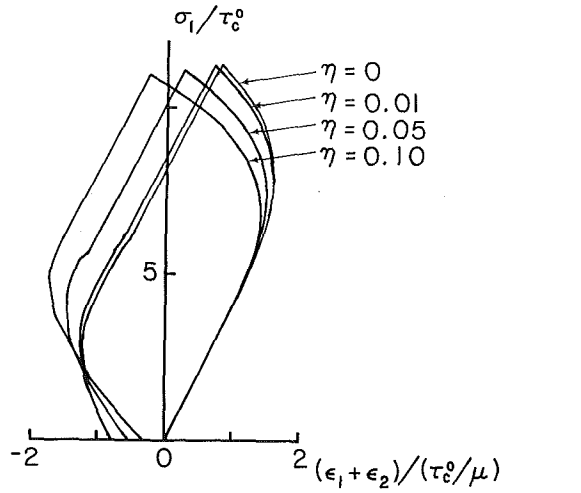


Fig. 5 Volumetric strain versus differential stress curves for various η; $f = 0.3$, $M_0 = 10$, $\nu = 0.25$, $\bar{\mu} = 0.3$, $d_0 = 0$, and $K_c = 1.0$.

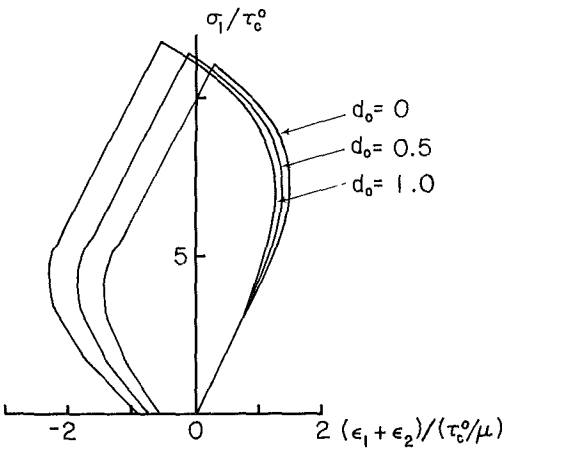


Fig. 6 Volumetric strain versus differential stress curves for various d_0; $f = 0.3$, $M_0 = 10$, $\nu = 0.25$, $\bar{\mu} = 0.3$, $K_c = 1.0$, and $\eta = 0.05$

nondimensionalized using half the flaw length c. It is also *normalized* by dividing by τ_c^0/μ. Hence, $d_0 = 1$, for example, represents an actual maximum possible flaw expansion of 0.5 percent of half the flaw length c when $\tau_c^0/\mu = 5.0 \times 10^{-3}$.

(5) ζ which is the coefficient of the exponential saturation of the flaw dilatancy with respect to the flaw slip. Since we normalize the slip b (and the dilatancy d) by measuring it in units of τ_c^0/μ, ζ (and η in item 6 below) is measured in units of μ/τ_c^0 which is of the order of $10^2 \sim 10^3$, a rather large dimensionless number. For all examples, we use $\zeta = 1$; hence, full flaw expansion is attained at flaw slips of less than 1 percent of the flaw length.

(6) η which is the coefficient of the exponential decay of the rate of cohesive resistance of the preexisting flaw.

(7) $\bar{\mu}$ which is the coefficient of internal friction possibly having a value between 0 and 1.

(8) ν which is the Poisson ratio of the elastic matrix. In actual cases, ν should be interpreted as the average overall Poisson ratio of a solid with microdefects. For the examples displayed in Figs. 3–6, we have used $f = 0.3$, $M_0 = 10$, $\nu = 0.25$, $\zeta = 1.0$, and, unless otherwise indicated explicitly, $K_c = 1.0$, $d_0 = 0.05$, $\eta = 0.05$, and $\bar{\mu} = 0.30$. In these figures, stresses are normalized by τ_c^0 and strains by (τ_c^0/μ).

Figures 3(*a*) and 3(*b*) show typical responses in uniaxial compression. In Fig. 3(*a*) the normalized overall volumetric strain and in Fig. 3(*b*) the normalized overall axial compressive strain, are plotted on the horizontal axis in terms of the normalized compression as ordinate. As is common in the rock mechanics literature, compression and contraction are viewed as positive. Up to point A, the overall response is linearly elastic, because the local stresses acting on the preexisting flaws are not as yet large enough to activate these flaws. At point A frictional sliding initiates on some flaws, and at point B cracks nucleate at some suitably oriented flaws. Upon further loading, tension cracks begin to grow, leading to substantial dilatancy which quickly offsets the elastic volumetric contraction, leading to overall volumetric expansion as the axial compressive load is further increased (see Fig. 3(*a*)).

The stress-strain relation shown in Fig. 3(*b*) remains monotonic in uniaxial loading, and closely resembles the experimentally observed behavior of rocks (see, e.g., Brace and Bombolakis, 1963). In the presence of lateral tension, however, the curve would terminate at a point when a most critical tension crack begins to grow in the axial direction in an unstable manner, leading to axial splitting (see Nemat-Nasser and Horii, 1982). In the presence of lateral compression the stress-strain curve remains monotonic and no failure is predicted by this model because the interaction effects among neighboring flaws are not incorporated in the present model. As has been shown by Horii and Nemat-Nasser (1985a, 1986),

Table 1 Instantaneous compliance at various loading and unloading states

σ_{11}/τ_c^0	$\epsilon_{11}/(\tau_c^0/\mu)$	μD_{1111}	μD_{1122}	μD_{2222}	μD_{2211}	μD_{1212}
0*	0	0.375	−0.125	0.375	−0.125	0.25
3.68	1.42	0.421	−0.212	0.462	−0.171	0.257
7.46	3.21	0.483	−0.652	3.775	−0.510	0.331
13.1**	5.99	0.391	−0.391	27.6	−1.18	2.368
5.28†	2.95	0.472	−1.03	14.2	−0.648	2.657

* Initial elastic loading
** Upon unloading
† Unloading with backsliding

the interaction among a row of suitably oriented preexisting flaws can lead to the formation of a fault, as tension cracks nucleate at these flaws and grow in an unstable manner. The analytical computations associated with such a model of faulting are rather complex and, therefore, considerable simplification is necessary before such a failure model can be incorporated into a micromechanically-based constitutive relation of the kind developed in the preceding section. The present constitutive model, therefore, is not suitable for such loading regimes.

At point C in Figs. 3(a) and 3(b), unloading has been initiated. The response from C to D is linearly elastic. However, the material is slightly more compliant, since microcracking has led to certain stiffness degradation. The difference between the elastic moduli at initial loading and at unloading from point C is rather small in the axial direction, but it is considerable for lateral tension (but not for lateral compression). This is because during axial compression, cracks grow essentially in the axial direction, and such cracks do not influence much the axial stiffness of the material, but they do have a considerable weakening effect for lateral tension.

The instantaneous compliances at several states during the course of a monotonic axial compression and of unloading are given in Table 1. As is seen, strong anisotropy develops due to slip and microcracking.

From point C to point D in Figs. 3(a) and 3(b), all the flaws are locked and all the tension cracks remain open. Backsliding of some preexisting flaws initiates at point D. In addition, some tension cracks grow before becoming dormant. Considerable volumetric contraction begins to occur after this state. In this example, no crack closure occurs during the entire unloading process which leads to point E corresponding to some permanent total volumetric expansion. This type of response has been inferred by Scholz and Kranz (1974) on the basis of experimental observations.

In Figs. 4(a) and 4(b), we show examples of loading-path-dependency. Figure 4(a) shows the stress-strain curves for two different loading paths defined in Fig. 4(b). On Path I, uniaxial compression is applied until point A, producing slip and microcracks at some preexisting flaws. Then, with axial compression held fixed, the lateral compressive stress is applied up to point C. During this regime, no further slip occurs on any flaws because the driving shear stresses on the flaws actually decrease, but not enough to initiate backsliding. At point B, closure of some tension cracks begins.

Path II, on the other hand, defines a proportional loading directly to point C'. Unlike for Path I loading, for Path II no slip is activated. The response, therefore, is linearly elastic. The microstructures of the material for points C and C' are quite different.

Our model includes a number of important parameters, and we now proceed to examine their effects on the model's predictions. These parameters are listed at the beginning of this section. Except for ζ, η, and d_0, they all have rather clear physical significance and therefore, their values may be estimated on the basis of experimental observations. The

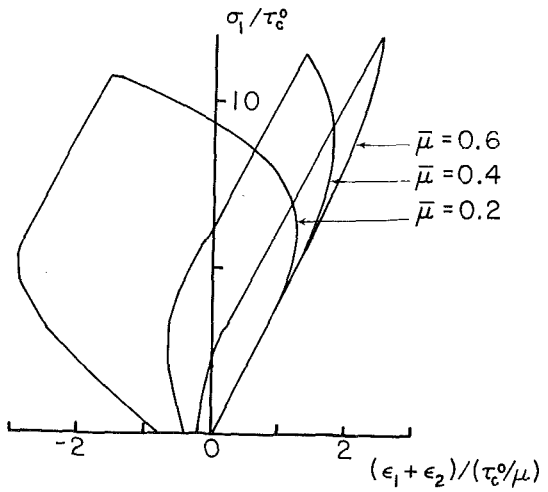


Fig. 7 Volumetric strain versus differential stress curves for various $\bar{\mu}$; $f = 0.3$, $M_0 = 10$, $\nu = 0.25$, $K_c = 1.0$, $d_0 = 0.05$, and $\mu = 0.05$

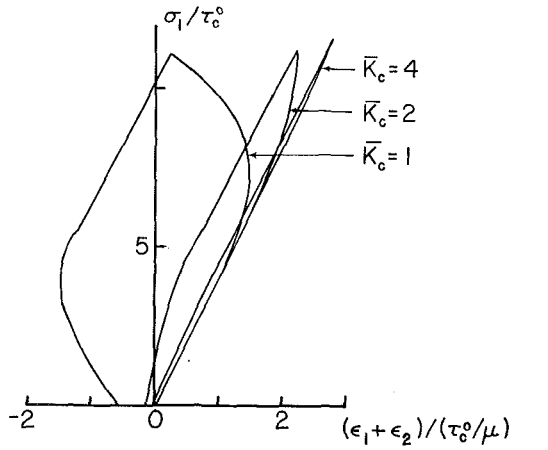


Fig. 8 Volumetric strain versus differential stress curves for various K_c; $f = 0.3$, $M_0 = 10$, $\nu = 0.25$, $\bar{\mu} = 0.3$, $d_0 = 0.05$, and $\eta = 0.05$

parameter d_0 represents maximum expansion (due to asperities or other effects) that the flaw itself can undergo because of sliding. The parameters ζ and η represent the rate at which flaw expansion saturates and the rate at which the cohesive resistance of the flaw decays, both measured per unit rate of sliding. These parameters have a rather significant influence on the overall volumetric response predicted by the model.

Figures 5 and 6 show the effect of η and of d_0 in uniaxial compression, followed by unloading. As pointed out before, when $(\tau_c^0/\mu) = 5.0 \times 10^{-3}$, $d_0 = 1$ represents an actual maximum limiting flaw expansion normal to its surface of only 0.5 percent of half the total flaw length, and hence it is small. Nevertheless, it has a rather significant influence on the response predicted by the model.

The influence of internal friction $\bar{\mu}$ and the normalized fracture toughness K_c on the dilatancy in uniaxial compression followed by unloading, is depicted in Figs. 7 and 8, respectively. As one would expect, a greater dilatancy results at lower frictional resistance of the flaws and lower fracture toughness of the matrix material. It should be noted that there is a considerable difference between the effects of the frictional parameter $\bar{\mu}$ and the normalized fracture toughness K_c on the overall response: whereas the friction coefficient $\bar{\mu}$ affects the flaw response during the entire loading and unloading regime, K_c's influence emerges only with tension

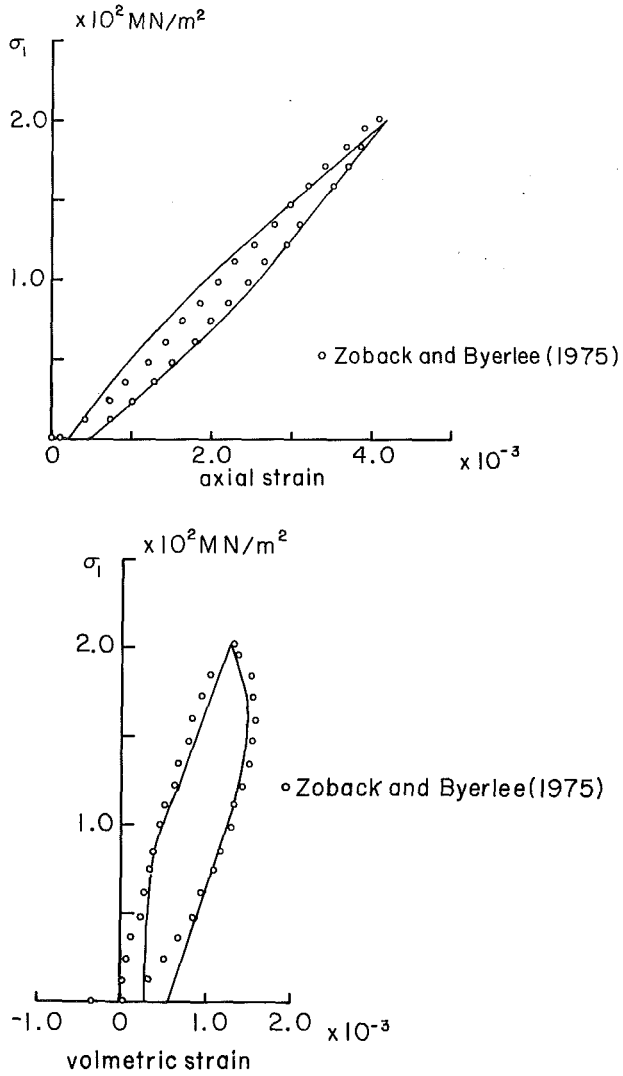


Fig. 9 Comparison with experimental data reported by Zoback and Byerlee (1975): (a) Volumetric strain versus differential stress curve. (b) Axial strain versus differential stress curve; $f = 0.35$, $M_0 = 10$, $\nu = 0.35$, $\bar{\mu} = 0.1$, $K_c = 2.1$, $d_0 = 0$, $\eta = 0$. $\mu = 2.1 \times 10^{10}$ Nm^{-2}, and $\tau_c^0 = 2.1 \times 10^7$ Nm^{-2}.

crack initiation. Note also that the normalized K_c involves the flaw size $2c$ and the cohesive stress τ_c^0, as well as the actual fracture toughness of the surrounding material. Therefore, while the order of magnitude of K_c can be estimated on the basis of experimental observations, its actual value involves quantities that are difficult to measure. For example, the fracture toughness should actually reflect the behavior of aggregates of crystals involving a variety of preexisting defects in a rock or ceramic sample.

As a final illustration, we seek to fit the model to experimental results reported by Zoback and Byerlee (1975). Since actual rocks contain preexisting open cracks and pores which tend to close upon application of loads, the initial stress-strain relations often reported are considerably softer than the subsequent response. We have not included this in our calculations, although it can easily be done, as detailed by Horii and Nemat-Nasser (1982). Therefore, we do not seek to match the initial portion of the stress-strain curves, but only compare their overall shapes. To offset the initial soft response, we simply shift the model's prediction, as shown in Figs. 9(*a*) and 9(*b*).

Since the model is a two-dimensional one, whereas the experimental results are for uniaxial compression of a cylindrical sample, a certain amount of adjustment in the model parameters is required. The reported flaw size $2(c)$ is of the order $4 \sim 8 \times 10^{-4}$, and taking $f = Nc^3/V \sim 0.2$, for the two-dimensional case we choose $f = (0.2)^{2/3} \sim 0.34$. The Poisson ratio and the shear modulus are estimated to be $\nu = 0.35$ and $\mu = 2.1 \times 10^4$ MNm^{-2}, based on the data reported by Zoback and Byerlee (1975). In addition, we take $\bar{K}_c = 0.7$ MNm$^{-3/2}$, $\tau_c^0 = 21$ MNm^{-2}, and set $2c = 5.0 \times 10^{-4}$m (therefore, $K_c = 2.1$). The coefficient of friction is taken as $\bar{\mu} = 0.1$, which is rather small, but may be justified on the ground that many of the flaws seem to have been initially open and, therefore, their frictional resistance ought to be small. Considering the simplicity of the model, the fit shown in Figs. 9(*a*) and 9(*b*) is rather good. Nevertheless, in view of the number of free parameters involved, the comparison should be regarded as an illustration and not necessarily a direct support of the viability of the model.

In summary, it may be concluded that the micromechanical model consisting of preexisting flaws, with frictional and cohesive resistance to sliding, and capable of nucleating tension cracks, seems to capture a number of important features observed in the response of brittle materials with microdefects. The model presented here does not include the important interaction effects among neighboring flaws. Judging from results reported by Horii and Nemat-Nasser (1986) on the failure in compression of materials of this kind under moderate and large confining pressures, we feel that constitutive modeling based on the mechanisms considered by these authors should prove to be effective.

Acknowledgment

This work was sponsored in part by the Air Force Office of Scientific Research, under AFOSR Grant No. AFOSR-86-0035, and in part by the Defense Advanced Research Projects Agency—Office of Naval Resarch (DARPA order No. 5213, amendment 2—ONR Contract No. N00014-85-C-0853) under contract with S-Cubed Division of Maxwell Labs, Inc.

References

Brace, W. F., and Bombolakis, E. G., 1963, "A Note on Brittle Crack Growth in Compression," *Journal of Geophysical Research*, Vol. 68, pp. 3709-3713.

Budiansky, B., 1965, "On the Elastic Moduli of Some Heterogeneous Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 223-227.

Costin, L., 1983, "A Microcrack Model for Deformation of Brittle Rock," *Journal of Geophysical Research*, Vol. 88, pp. 9485-9492.

Hill, R., 1965, " A Self-Consistent Mechanics of Composite Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 213-222.

Hoek, E., and Bieniawski, Z. T., 1965, "Brittle Fracture Propagation in Rock Under Compression," *International Journal of Fracture Mechanics*, Vol. 1, pp. 137-155.

Holcomb, D. J., 1978, "A Quantitative Model of Dilatancy in Dry Rock and its Application to Westerly Granite," *Journal of Geophysical Research*, Vol. 83, pp. 4941-4950.

Holcomb, D. J., and Stevens, J. L., 1980, "The Reversible Griffith Crack: A Viable Model for Dilatancy," *Journal of Geophysical Research*, Vol. 85, pp. 7101-7107.

Holzhausen, G. R., 1978, "Sheet Structure in Rock and Some Related Problems in Rock Mechanics," Ph.D. Thesis, Stanford University, Stanford, CA.

Horii, H., and Nemat-Nasser, S., 1982, "Overall Moduli of Solids With Microcracks: Load-Induced Anisotropy," *Journal of the Mechanics and Physics of Solids*, Vol. 31, pp. 155-171.

Horii, H., and Nemat-Nasser, S., 1983, "Estimate of Stress Intensity Factors for Interacting Cracks," *Advances in Aerospace Structures, Materials and Dynamics, AD-06*, U. Yuceoglu, R. L., Sierakowski, and D. A. Glasgow, eds., ASME, New York, pp. 111-117.

Horii, H., and Nemat-Nasser, S., 1985a, "Compression-Induced Microcrack Growth in Brittle Solids: Axial Splitting and Shear Failure," *Journal of Geophysical Research*, Vol. 90, pp. 3105-3225.

Horii, H., and Nemat-Nasser, S., 1985b, "Elastic Fields of Interacting Inhomogeneities," *International Journal of Solids and Structures*, Vol. 21, pp. 731-745.

Horii, H., and Nemat-Nasser, S., 1986, "Brittle Failure in Compression: Splitting, Faulting, and Brittle-Ductile Transition," *Philosophical Transactions, Royal Society London*, Vol. 319, pp. 337–374.

Janach, W., and Guex, L. H., 1980, "In-Plane Propagation of Shear Microcracks in Brittle Rocks Under Triaxial Compression," *Journal of Geophysical Research*, Vol. 85, pp. 2543–2553.

Kachanov, M. L., 1982, "A Microcrack Model of Rock Inelasticity, Part II: Propagation of Microcracks," *Mechanics of Materials*, Vol. 1, pp. 29–41.

Kranz, R. L., 1983, "Microcracks in Rocks: A Review," *Tectonophysics*, Vol. 100, pp. 449–480.

McClintock, F. A., and Walsh, J. B., 1962, "Friction on Griffith Cracks in Rocks Under Pressure," *Proceedings of the Fourth U.S. National Congress of Applied Mechanics*, ASME, New York, pp. 1015–1021.

Moss, W. C., And Gupta, Y. M., 1982, "A Constitutive Model Describing Dilatancy and Failure in Brittle Rock," *Journal of Geophysical Research*, Vol. 87, pp. 2985–2998.

Muskhelishvili, N. I., 1953, *Singular Integral Equations*, Noordhoff, Groningen, Holland.

Nemat-Nasser, S., 1985, Discussion of "Geometric Probability Approach to the Characterization and Analysis of Microcracking in Rocks," *Mechanics of Materials*, Vol. 4, pp. 277–281.

Nemat-Nasser, S., and Horii, H., 1982, "Compression-Induced Nonplanar Crack Extension with Application to Splitting, Exfoliation, and Rockburst," *Journal of Geophysical Research*, Vol. 87, pp. 6805–6821.

Paterson, M. S., 1978, *Experimental Rock Deformation—The Brittle Field*, Springer-Verlag, Berlin-Heidelberg-New York.

Scholz, C. H., 1968, "Experimental Study of the Fracturing Process in Brittle Rock," *Journal of Geophysical Research*, Vol. 73, pp. 1447–1454.

Scholz, C. H., Boitnott, G., and Nemat-Nasser, S., 1986, "The Bridgman Ring Paradox Revisited," *Pure and Applied Geophysics*, Vol. 124, pp. 587–599.

Scholz, C. H., and Kranz, R., 1974, "Notes on Dilatancy Recovery," *Journal of Geophysical Research*, Vol. 79, pp. 2132–2135.

Sondergeld, C. H., and Estey, L. H., 1982, "Source Mechanism and Microfracturing During Uniaxial Cycling of Rock," *Pure and Applied Geophysics*, Vol. 120, p. 151.

Steif, P. S., 1984, "Crack Extension Under Compressive Loading," *Engineering Fracture Mechanics*, Vol. 20, pp. 463–473.

Wong, T. F., 1985, "Geometric Probability Approach to the Characterization and Analysis of Microcracking in Rocks," *Mechanics of Materials*, Vol. 4, pp. 261–276.

Yanagidani, T., Ehara, S., Nishizawa, O., Kusunose, K., and Terada, M., 1985, "Localization of Dilatancy in Westerly Granite Under Constant Uniaxial Stress," *Journal of Geophysical Research*, Vol. 90, pp. 6840–6858.

Zoback, M. D., and Byerlee, J. D., 1975, "The Effect of Cyclic Differential Stress on Dilatancy in Westerly Granite Under Uniaxial and Triaxial Conditions," *Journal of Geophysical Research*, Vol. 80, pp. 1526–1530.

APPENDIX A

Strains Due to Crack Opening

The strain due to the crack opening displacement is given by

$$\epsilon_{ij} = \frac{1}{2}\frac{N}{V}\int_{-c}^{c} \{n_i[u_j] + n_j[u_i]\}\,ds \tag{A.1}$$

where V is the volume (with unit thickness), N is the number of cracks in this volume, $\mathbf{n}$ is a unit vector normal to a crack surface, and $[u_i]$ is the ith component of the crack opening displacement. From this, the inelastic strain due to crack slip is computed as follows. In a local coordinate system with the crack on the x_1 axis,

$$\epsilon_{12} = \epsilon_{21} = \frac{N}{V}\int_0^c [u_2]dx = \frac{N}{V}\,\bar{b}c = fb,$$
$$\epsilon_{11} = \epsilon_{22} = 0,\ b = \bar{b}/c. \tag{A.2}$$

Transformation to the global coordinate system is made with the following matrix:

$$\mathbf{T} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}, \tag{A.3}$$

where ϕ is the angle between the local and global x_1 axis. By use of this matrix, we have

$$\epsilon_b = 2fb\,\frac{1}{2}\begin{bmatrix} -\sin 2\phi & \cos 2\phi \\ \cos 2\phi & \sin 2\phi \end{bmatrix}, \tag{A.4}$$

in the global coordinate system.

For the Mode I crack displacement, we have

$$\epsilon_{11} = 2\,\frac{N}{V}\int_0^c [u_1]dx = 2fd, \tag{A.5}$$

in the local coordinates, and

$$\epsilon_d = 2fd\begin{bmatrix} \cos^2\phi & \cos\phi\sin\phi \\ \cos\phi\sin\phi & \sin^2\phi \end{bmatrix}, \tag{A.6}$$

in the global coordinates.

APPENDIX B

Strains Due to a Dislocated Crack

First, we examine the Mode I crack displacement. Consider an array of edge dislocations continuously distributed along $-a < x_2 < a$, $x_1 = 0$ (see Fig. B.1). The corresponding normal stress is (see e.g., Muskhelishvili, 1953).

$$\sigma_{11}(x_2,0) = -\frac{1}{2\pi}\,\frac{\mu}{1-\nu}\int_{-a}^{a}\frac{B_1(\xi)}{\xi - x_2}\,d\xi, \tag{B.1}$$

and the displacement in the x_1 direction on the x_2 axis is given by

$$u_1(0,x_2) \equiv g(x_2) = -\int_{-\infty}^{x_2} B_1(\xi)\,d\xi, \tag{B.2}$$

where $B_1(x_2)$ is the dislocation density. Since crack surfaces are traction-free, with the farfield stresses applied, the following integral equation defines the dislocation density:

$$-\frac{1}{2\pi}\,\frac{\mu}{1-\nu}\int_{-a}^{a}\frac{B_1(\xi)}{\xi - x_2}\,d\xi + p(x_2) = 0,\ -a < x_2 < a, \tag{B.3}$$

where $p(x_2)$ is the traction on the x_2 axis due to farfield stress.

A solution of equation (B.3) singular at $x_2 = \pm a$ is given by (e.g., Muskhelishvili, 1953)

$$B_1(x_2) = -\frac{1}{\pi^2}\,\frac{1}{\sqrt{a^2 - x_2^2}}\left\{2\pi\,\frac{1-\nu}{\mu}\int_{-a}^{a}\frac{\sqrt{a^2-\xi^2}}{\xi - x_2}\,p(\xi)\,d\xi + C\right\}, \tag{B.4}$$

where C is a constant yet to be determined by another boundary condition. In this case, the farfield stress is zero but the crack is dislocated at $x_2 = a$,

$$p(x_2) \equiv 0,\ [u_1](a) = g_1. \tag{B.5}$$

From equations (B.2), (B.4), and (B.5), we have

$$B_1(x_2) = -\frac{g_1}{\pi}\,\frac{1}{\sqrt{a^2 - x_2^2}}. \tag{B.6}$$

Note that the condition $[u_1] = 0$ for $-a > x_2$ is automatically satisfied by this array of dislocations. Substituting equation (B.6) into (B.1), we obtain the normal stresses on $x_1 = 0$,

$$\sigma_{11}(x_2,0) = -\frac{g_1}{2\pi}\,\frac{\mu}{1-\nu}\,\frac{\mathrm{sgn}(x_2)}{\sqrt{x_2^2 - a^2}}\,H(|x_2| - a), \tag{B.7}$$

where $H(x)$ is the Heaviside step function. The stress intensity factor at $x_2 = -a$ now is

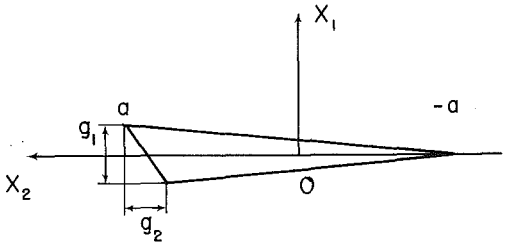


Fig. B.1

$$K_I = \frac{g_1}{2}\frac{\mu}{1-\nu}\frac{1}{\sqrt{\pi a}}. \tag{B.8}$$

From equations (B.2), (B.5), and (A.1), we compute the strain components in the local coordinate system as follows:

$$\epsilon_{11} = g_1 a, \qquad \epsilon_{12} = \epsilon_{21} = \epsilon_{22} = 0. \tag{B.9}$$

Note that both g_1 and a are dimensionless (see Appendix A).

Now consider the Mode II crack displacement. If the concentrated gap at $x_2 = a$ is denoted by g_2, we obtain

$$\epsilon_{12} = \epsilon_{21} = \frac{1}{2} g_2 a. \tag{B.10}$$

In our particular problem,

$$g_1 = b\sin\theta + d\cos\theta, \qquad g_2 = b\cos\theta - d\sin\theta, \qquad a = \frac{\ell}{2}. \tag{B.11}$$

Thus, considering that we have a pair of dislocated cracks, the strain components in the local coordinate system are given by

$$\epsilon_c = f\frac{b\ell}{2}\begin{bmatrix} 2\sin\theta\cos\theta & \\ \cos\theta & 0 \end{bmatrix} + f\frac{d\ell}{2}\begin{bmatrix} 2\cos\theta & -\sin\theta \\ -\sin\theta & 0 \end{bmatrix}. \tag{B.12}$$

The transformation to the global coordinate system is made with

$$\mathbf{T} = \begin{bmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}, \tag{B.13}$$

and we obtain

$$\epsilon_c = fb\ell\frac{1}{2}\begin{bmatrix} -2\sin\phi\cos(\theta+\phi) & \cos(\theta+2\phi) \\ \cos(\theta+2\phi) & 2\cos\phi\sin(\theta+\phi) \end{bmatrix} \tag{B.14}$$

$$+ fd\ell\frac{1}{2}\begin{bmatrix} 2\cos\phi\cos(\theta+\phi) & \sin(\theta+2\phi) \\ \sin(\theta+2\phi) & 2\sin\phi\sin(\theta+\phi) \end{bmatrix},$$

in the global coordinate system.

APPENDIX C

Strains Due to Crack Displacement Caused by Applied Stresses

Under the farfield stresses T and S, the crack opening displacements are given by

$$[u_1] = 2T\frac{1-\nu}{\mu}\sqrt{(\ell/2)^2 - x_2^2}, \qquad [u_2] = 2S\frac{1-\nu}{\mu}\sqrt{(\ell/2)^2 - x_2^2}. \tag{C.1}$$

In our particular problem, $T = \boldsymbol{\sigma}:\boldsymbol{\alpha}$ and $S = \boldsymbol{\sigma}:\boldsymbol{\beta}/2$. Therefore, after integrating, the strains in the local coordinate system are expressed as

$$\epsilon_a = f\frac{1-\nu}{8\mu}\pi\ell^2\begin{bmatrix} 4\boldsymbol{\sigma}:\boldsymbol{\alpha} & \boldsymbol{\sigma}:\boldsymbol{\beta} \\ \boldsymbol{\sigma}:\boldsymbol{\beta} & 0 \end{bmatrix}. \tag{C.2}$$

Using equation (B.11), we have the expression in the global coordinate system,

$$\epsilon_a = f\frac{1-\nu}{8\mu}\pi\ell^2[4(\boldsymbol{\sigma}:\boldsymbol{\alpha})\boldsymbol{\alpha} + (\boldsymbol{\sigma}:\boldsymbol{\beta})\boldsymbol{\beta}]. \tag{C.3}$$

APPENDIX D

Stress Intensity Factor of a Dislocated Crack

Consider the case where the crack opening displacement normal to the flaw is zero. By use of equation (B.8), the Mode I stress intensity factor of a dislocated crack of length ℓ is obtained to be

$$K_I' = \frac{\mu}{1-\nu}\frac{b\sin\theta}{\sqrt{2\pi\ell}} + \boldsymbol{\sigma}:\boldsymbol{\alpha}\sqrt{\frac{\pi\ell}{2}}. \tag{D.1}$$

Since this expression is singular when $\ell = 0$, we replace it by the following nonsingular form

$$K_I' = \frac{\mu}{1-\nu}\frac{1}{\sqrt{2\pi(\ell+\ell^{**})}}b\sin\theta + \boldsymbol{\sigma}:\boldsymbol{\alpha}\sqrt{\frac{\pi\ell}{2}}, \tag{D.2}$$

and seek to estimate ℓ^{**} such that at $\ell = 0^+$, $K_I' = K_I$. When $\ell = 0$, the average slip of the preexisting flaw equals the average Mode II crack displacement, and is given by

$$\bar{b} = \frac{1}{2c}\int_{-c}^{c} 2(\boldsymbol{\sigma}:\mathbf{p}_2 - \tau_c)\frac{1-\nu}{\mu}\sqrt{c^2 - x^2}\,dx = \frac{1}{2}\frac{1-\nu}{\mu}\pi c(\boldsymbol{\sigma}:\mathbf{p}_2 - \tau_c). \tag{D.3}$$

On the other hand, when $\ell = 0$, equation (2.9a) reduces to

$$K_I = \frac{2}{\sqrt{\pi\ell^*}}(\boldsymbol{\sigma}:\mathbf{p}_2 - \tau_c)\sin\theta, \tag{D.4}$$

which is maximum at $\theta = \pi/2$. From equations (D.2), (D.3), and (D.4) with $\theta = \pi/2$, ℓ^{**} is determined to be

$$\ell^{**} = \frac{\pi^2}{32}\ell^*. \tag{D.5}$$

We simply generalize this result when dilatancy d is included. The final form of K_I' then is

$$K_I' = \frac{\mu}{1-\nu}\frac{1}{\sqrt{2\pi(\ell+\ell^{**})}}(b\sin\theta + d\cos\theta) + \boldsymbol{\sigma}:\boldsymbol{\alpha}\sqrt{\frac{\pi\ell}{2}}. \tag{D.6}$$

APPENDIX E

Stress Intensity Factor in Unloading

Equation (2.25) with K_I defined by equation (2.24a), may be explained in the following way. At the end of a loading regime when the applied farfield stresses are $\boldsymbol{\sigma}_{\max}$, the residual stress on the preexisting flaw is given by

$$\tau_m = \boldsymbol{\sigma}_{\max}:\mathbf{p}_2 - \tau_c. \tag{E.1}$$

Here, τ_m is the maximum shear stress experienced by the flaw during this loading regime. Backsliding is considered to occur when this residual stress overcomes the frictional and cohesive resistance of the flaw, i.e., when

$$\tau_m - \boldsymbol{\sigma}:\mathbf{p}_0 \geq \tau_c - \bar{\mu}\boldsymbol{\sigma}:\mathbf{p}_1. \tag{E.2}$$

At the end of a loading regime, we have

$$K_I = K_I' = \frac{2}{\sqrt{\pi(\ell+\ell^*)}}\tau_m\sin\theta + \boldsymbol{\sigma}_{\max}:\boldsymbol{\alpha}\sqrt{\pi\ell}. \tag{E.3}$$

Therefore, at the onset of backsliding, the equality in expression (E.2) yields

$$K_I' \simeq KI_I. \tag{E.4}$$

The essential difference between equations (2.24*a*) and (2.10*a*) is due to the terms associated with the farfield applied stress, i.e., $\sigma{:}\alpha\sqrt{(\pi\ell/2)}$ and $\sigma{:}\alpha\sqrt{(\pi\ell)}$. This difference stems from the fact that the computation of the stress intensity factor in the presence of a gap is for a crack of length ℓ, whereas that associated with the force F, is for a crack of length 2ℓ. Because tension cracks grow in the maximum compressive stress direction, this difference is quite small. Even for extreme cases, e.g., just after crack nucleation, ℓ is small and hence this term is very small. We may, therefore, use the expression (2.25) and (2.26) with K_I defined by equation (2.24*a*), in order to ensure a consistent formulation without loss of accuracy.

APPENDIX F

Partially Closed Dislocated Crack

We consider the case shown in Fig. *B*.1. In this case, the integral equation corresponding to equation (*B*.3) becomes

$$\int_{-a}^{a} \frac{B_1(\xi)}{\xi - x_2}\, d\xi = 2\pi \frac{1-\nu}{\mu} T. \tag{F.1}$$

If we solve this equation under the condition that σ_{11} is bounded at $x_2 = -a$ and unbounded at $x_2 = a$, we obtain (e.g., Muskhelishvili, 1953)

$$B_1(x) = 2T \frac{1-\nu}{\mu} \frac{\sqrt{a+x_2}}{\sqrt{a-x_2}}. \tag{F.2}$$

After integration of equation (*B*.2), the crack opening displacement is given by

$$[u_1](x_2) = -\frac{1-\nu}{\mu} 2T\left[a \sin^{-1}\frac{x_2}{a} + a\frac{\pi}{2} - \sqrt{a^2 - x_2^2}\right]. \tag{F.3}$$

The condition at $x_2 = a$ yields

$$a = -\frac{1}{2\pi}\frac{\mu}{1-\nu}\frac{g_1}{T}. \tag{F.4}$$

Apparently, this result is reasonable only when $T < 0$. In our particular problem,

$$g_1 = b \sin\theta + d\cos\theta,\ T = \sigma{:}\alpha. \tag{F.5}$$

Thus, integration of the crack opening displacement leads to the following expression for the strain component ϵ_{11} in the local coordinate system:

$$\epsilon_{11} = -f\frac{1}{2\pi}\frac{\mu}{1-\nu}\frac{1}{(\sigma{:}\alpha)}(b\sin\theta + d\cos\theta)^2. \tag{F.6}$$

The strain component ϵ_{12} requires no modification when we assume no friction on tension cracks. Considering the contribution by the farfield stresses as well as the dislocation array, we obtain

$$\epsilon_{12} = f\frac{\ell_m}{2}(b\cos\theta - d\sin\theta) + f\frac{1-\nu}{8\mu}\pi\ell_m^2(\sigma{:}\beta) \tag{F.7}$$

in the local coordinate system. The transformation to the global coordinate system is done by equation (*B*.13). In the global coordinate system, we hence have

$$\begin{aligned}\epsilon_c = &-f\frac{1}{2\pi}\frac{\mu}{1-\nu}\frac{1}{(\sigma{:}\alpha)}(b\sin\theta + d\cos\theta)^2\alpha \\ &+ f\frac{\ell_m}{2}(b\cos\theta - d\sin\theta)\beta \\ &+ f\frac{1-\nu}{8\mu}\pi\ell_m^2(\sigma{:}\beta)\beta.\end{aligned} \tag{F.8}$$

Numerical Studies of Dynamic Compaction of Inert and Energetic Granular Materials

M. R. Baer
Fluid and Thermal Sciences Department,
Sandia National Laboratories,
Albuquerque, NM 87185

In this paper a numerical study of dynamic compaction of granular materials is presented. A multiphase mixture model treating two-phase nonequilibrium flow is applied to describe dynamic compaction experiments. Compaction is modeled using a description of grain distortion driven by pressure differences. Various energetic and inert granular materials including HMX, Fluid A (Nitrocellulose), Melamine, and Teflon are studied for various piston-impact conditions and compared to existing data. Numerical analysis of a ramp-wave piston-impact experiment is also examined. The multiphase mixture model predicts reasonably well the observed steady compaction wave characteristics. Dynamic compaction experiments with detailed resolution are needed to better address the transient behavior of compaction waves in granular materials.

Introduction

It is now recognized that the onset of deflagration-to-detonation transition (DDT) in granular energetic materials is greatly influenced by the mechanical behavior of granular microstructure (Baer and Nunziato, 1986). Under strong confinement, high gas pressures produce flows with large drag forces and compaction concurrent with the compression of the granular reactant. This compaction precedes deflagration and alters the granular character of the reactant. Reduction of connected porosity decreases bed permeability and induces rapid pressurization of the evolved combustion gases. The combustion-compaction zone then acts like a piston driven onto the unreacted granular material, whereupon a shock wave forms and localized heating at individual "hot-spots" enhances the growth to detonation (Nunziato, 1984).

Several existing DDT models, that incorporate micromechanical models describing compaction behavior (e.g., Baer and Nunziato, 1986; Butler et al., 1982; Kooker and Anderson, 1981; Kim, 1984), have been greatly influenced by the work of Carroll and Holt (1972). In this study a multiphase description of granular flow is used with a rate equation for volume fraction (e.g., see Baer and Nunziato, 1986). The balance laws are posed with proper mixture constraints and admissible forms of the micromechanical aspects of multiphase combustion are suggested by thermodynamic arguments. Several inert and energetic granular materials are studied for various piston-impact conditions and compared to experimental investigations that have been conducted by Sandusky and Liddiard (1985). In these experiments, measurements of particle velocities, compaction wave speed, and compacted density were obtained and pressure was estimated using a jump condition across a "shock" compaction wave. Since the load transfer to the unreacted granular material is due to the action of a piston (as opposed to interphase drag effects) the focus of the numerical – experiment comparison centers on the mechanical response of the granular material.

In the sections to follow, a brief overview of the compaction experiments is provided and the analytical model is described. Appropriate parameters for the model are given for energetic and inert powders. A method-of-lines numerical method is employed in the solution of the multiphase flow equations and a piston boundary condition is accommodated using a coordinate transformation. In addition to detailed numerical solutions, steady wave solutions of the multiphase equations have been determined assuming only the solid phase governs compaction. This simplified approach yields qualitative estimates of compaction behavior.

Constant-velocity piston simulations are given for the conditions examined by Sandusky and Liddiard (1985). The evolution of the compaction wave is resolved to yield compaction wave speed and structure. The space and time variation of mixture pressure and volume fraction are determined and compared to experimental estimates. In these studies, good agreement between the model and Sandusky and Liddiard's experiment is seen which lends confidence in the numerical treatment of compaction described by the mixture approach. As a further test, a ramp-loading case was modeled for a granular Teflon column using initial conditions that were experimentally investigated by Sandusky and Liddiard. Calculations show a compaction wave evolving to a "shock" compaction wave at an interior location well removed from the piston/bed interface.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 20, 1986; final revision July 20, 1987.

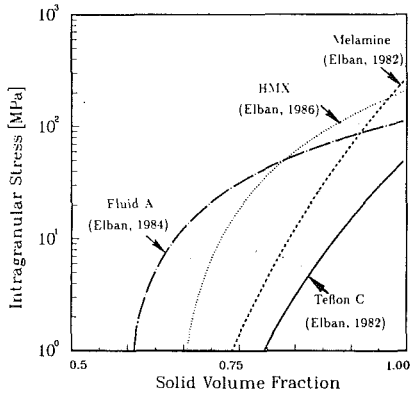


Fig. 1 Intragranular stress versus solid volume fraction for various granular materials as curve fit to Elban data (1982, 1984, 1986)

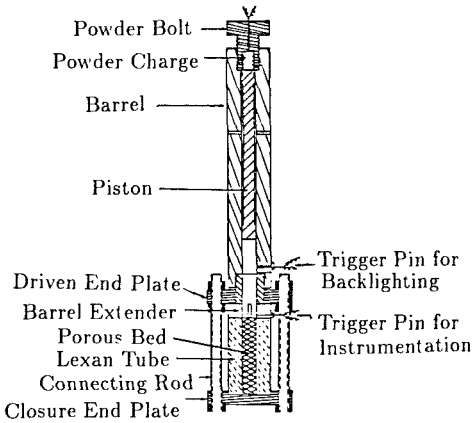


Fig. 2 Schematic of Sandusky and Liddiard (1985) piston-impact experiment

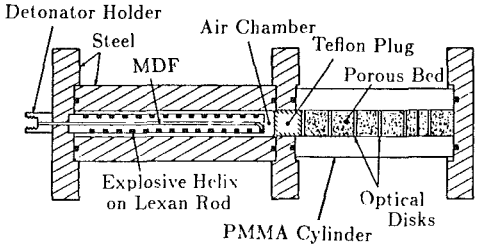


Fig. 3 Sandusky and Liddiard (1985) ramp-loading piston-impact experiment

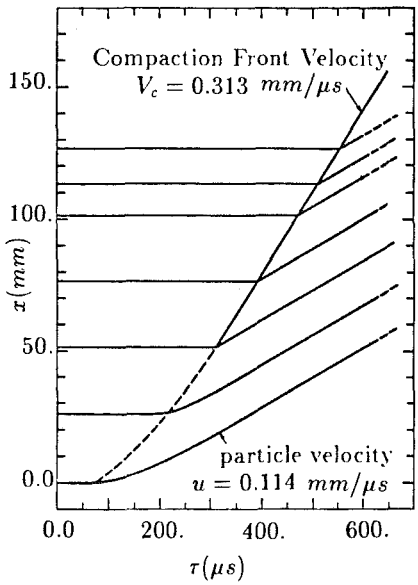


Fig. 4 Embedded disk trajectories of ramp-loading experiment in granular column of Teflon C

Experimental Studies of Compaction

Although there have been many experimental studies of the properties of pressed powders, much emphasis has centered on quasi-static mechanical behavior. Usually, the relationship between the local stress and the powder's configurational state (i.e., porosity, particle size, etc.) is sought. For example, measurements of the bed displacement are obtained for a quasi-statically mechanically loaded porous column. By varying the load, the resulting data can be correlated to a relationship such as the one posed by Herrmann (1969).

Specific to reactive granular materials, quasi-static compaction studies have been conducted by Elban (1982, 1984, 1986). In typical experiments, measurements of applied and transmitted loads are obtained with the required bed displacement. After correcting for wall effects, the intragranular stress, β_s, is determined as a function of solid volume fraction (or percent theoretical maximum density). Figure (1) shows intragranular stress for the inert and energetic materials used in this study.

During dynamic loading, rate effects of compaction may be significant. Carroll and Holt (1972) were among the first to recognize this importance and formulated a micromechanical model based on the collapse of a collection of spherical voids. Inertial effects of pore collapse were driven by pressure differences. Later, experimental studies of shock-loaded metal foams (Butcher, et al., 1974) indicated that a material viscosity rate dependence is appropriate to describe material collapse and the resulting analytical description is similar to the classical Rayleigh bubble equation (1917). For a given applied load, a state of pressure equilibrium is asymptotically attained when the difference in the phase pressures equals the intragranular stress. Furthermore, it has been shown that the generalized Carroll-Holt model is a thermodynamically consistent description of compaction (Nunziato and Walsh, 1980) and a simplified form is obtained when the individual void volume is small. This description is used in this study.

Recently, a series of dynamic compaction experiments have been conducted by Sandusky and Liddiard (1985) which provide an important data base to test the theoretical model for compaction behavior of granular materials. In particular, dynamic loads have been purposely chosen similar to those experienced during a DDT event.

Although Sandusky and Liddiard report various means for dynamically-loading granular columns, the piston-impact experiments are the main focus of this study (see Fig. (2)). Briefly, the expanding combustion gas from a powder charge accelerates a piston which is driven onto a porous bed confined in a tube. Tracer disks are placed at various locations within the bed and flash X-rays are used to determine particle velocities within the porous bed. From successive photographs, the x-t plane of the compaction wave front and the density change following compaction is deduced from the displacement of the embedded disks. Assuming the compaction wave has a "shock" wave structure, jump relationships are used to define the mixture pressure following compaction. In these experiments, consecutive disks were spaced at a distance of 0.64 cm; thus, only the global features of the wave motion were investigated. Unfortunately, the early transient wave development was not probed. The observed uniform displacement of the tracers suggested that the compaction wave had a structure, from the leading edge to its trailing edge, within the separation distance of the embedded disks.

Since it was observed that quasi-steady compaction waves quickly develop, a second type of piston-impact experiment (Sandusky and Liddiard, 1985) was considered in which a

ramp wave loads a porous column. Figure (3) shows a schematic of this type of experiment. The ramp loading is produced by a time-varying reactive burn that accelerates a piston. The particle velocity information is measured using high speed photography of backlit transparent embedded disks. The gradual increase in piston velocity produces compaction with a developing wave structure and the "shock" compaction wave appears within the interior of the porous bed. Figure (4) shows a representative x-t plot of the 2.5 cm separated tracers within a granular Teflon bed.

Theoretical Description

For the sake of brevity, only the final forms of the mixture equation are given; the interested reader can find a detailed description of the complete formulation in the reference by Baer and Nunziato (1986). Although energetic materials were used in Sandusky and Liddiard's experiments, the reactive terms of the mixture model are neglected and the simulations correspond to preignition events. Furthermore, the analysis is one-dimensional and treats a two-phase system (granular solid and air).

In the continuum mixture approach, each phase is assumed to occupy every point in the field and at every location we assign a single (but separate) thermodynamic state for each phase. Although an air component is included, the solid phase contains most of the mass, momentum, and energy of the system. For the sake of completeness, the interphase exchanges of momentum (drag) and energy (convective and conductive heat transfer) are included.[1] In this study, the effects of granular material strength are not examined and the effects of wall friction are not modeled.

In the following description we denote a particular phase with the subscript a (the solid is denoted as $a=s$ and the gas phase as $a=g$). Associated with each phase, we assign a material density, ρ_a, a volume fraction, ϕ_a, a particle velocity, v_a, pressure, p_a, temperature, T_a, and an internal energy, e_a. In the complete nonequilibrium treatment of fully compressible multiphase flow, the balance of mass, momentum and energy are given in the following form:

Mass

$$\frac{\partial}{\partial t}(\phi_a\rho_a)+\frac{\partial}{\partial x}(\phi_a\rho_a v_a)=0, \tag{1}$$

Momentum

$$\frac{\partial}{\partial t}(\phi_a\rho_a v_a)+\frac{\partial}{\partial x}(\phi_a P_a+\phi_a\rho_a v_a^2)=m_a^+, \tag{2}$$

Energy

$$\frac{\partial}{\partial t}(\phi_a\rho_a E_a)+\frac{\partial}{\partial x}((\phi_a\rho_a E_a+\phi_a P_a)v_a)=e_a^+ \tag{3}$$

where $E_a=e_a+v_a^2/2$ and the "+" superscript denotes a phase interaction term.

For the rate of change of solid volume fraction, a compaction model, derived from the entropy inequality, is used:

$$\frac{\partial\phi_s}{\partial t}+v_s\frac{\partial\phi_s}{\partial x}=\frac{f(\phi_s)}{\mu_c}(P_s-P_g-\beta_s) \tag{4}$$

and $\phi_s+\phi_g=1$ is the saturation volume constraint. The μ_c is the material viscosity, β_s is the intragranular stress of the granular bed (quasi-static experimental data) and $f(\phi_s)$ is a function that is constrained by the entropy inequality to be positive. In the current work $f(\phi_s)=\phi_s(1-\phi_s)$ is used which restricts the volume fraction rate when the solid or gas phase volume fractions approach zero.

[1] In these simulations, a numerical sensitivity study confirmed that these effects are small.

Table 1 Equation of state and thermophysical data

Material	ρ_s^o [g/cc]	Γ_s	c_v^s [erg/gK]	K_T [MPa]	N
HMX	1.90	1.1	1.50×10^7	1.35×10^4	9.8
Fluid A	1.62	1.5	1.03×10^7	3.98×10^3	7.4
Teflon C	2.30	1.1	1.00×10^7	5.67×10^3	7.3
Melamine	1.57	1.5	1.23×10^7	1.80×10^4	4.1

Consistent with early work, the interactions of momentum, m_a^+, and energy, e_a^+, are given by:

$$m_s^+=-m_g^+=P_g\frac{\partial\phi_s}{\partial x}-C_d(v_s-v_g), \tag{5}$$

$$e_s^+=-e_g^+=\left(P_g\frac{\partial\phi_s}{\partial x}-C_d(v_s-v_g)\right)v_s-h(T_s-T_g)$$
$$-(P_s-\beta_s)\frac{f(\phi_s)}{\mu_c}(P_s-P_g-\beta_s). \tag{6}$$

The drag, C_d, and convective heat transfer coefficient, h, are retained in the forms given in earlier work (Baer and Nunziato, 1986). Although these phase interactions are nonunique, these choices are motivated by prior experimental observations and are constrained by the thermodynamics.

In the experiments, observed pressures were less than 100 MPa; thus, an ideal-gas equation of state for air is appropriate. For the solid phase, a well established equation of state description is used (Sheffield et al., 1977):

$$P_s=\Gamma_s\rho_s(e_s+\Gamma_s\rho_s c_v^s T_s^o\left(\frac{1}{\rho_s^o}-\frac{1}{\rho_s}\right)-\frac{K_T}{\rho_s^o N(N-1)}\left[\left(\frac{\rho_s}{\rho_s^o}\right)^{N-1}$$
$$-(N-1)\left(1-\frac{\rho_s^o}{\rho_s}\right)-1\right])+\frac{K_T}{N}\left(\left(\frac{\rho_s}{\rho_s^o}\right)^N-1\right) \tag{7}$$

where the reference state T_s^o is taken as 300 K, ρ_s^o is the unstressed solid density, K_T is the bulk modulus, c_v^s is the specific heat, Γ_s is the Grüneisen constant, and N is the exponent for the density variation of the bulk modulus. The solid equation of state parameters are determined using unreacted shock Hugoniot data (Marsh, 1980) and the thermophysical data are taken from Gibbs and Popolato (1980) and Dobratz (1981). Table (1) shows equation of state data for the materials of interest in this study.

The only remaining parameter in this model is the material viscosity, μ_c. From a survey of prior work (i.e., see Kahasinov et al., 1981; Kooker, 1986; Maiden and Nutt, 1986), the value of μ_c is estimated to be 10^2-10^4 poise.

Numerical Method

In prior DDT calculations (Baer and Nunziato, 1986), the method-of-lines numerical technique was used to solve the strongly coupled multiphase flow equations. In this Eulerian approach, spatial derivatives are discretized using central differences and the resulting set of ordinary differential equations is solved using well-developed ODE software which can resolve a large set of highly stiff equations. This method is of particular interest since highly coupled phase interactions produces time/length scales which differ greatly from those of the hydrodynamics. To stabilize hydrodynamic calculations, a minimal amount of numerical viscosity is required to prevent numerical dispersion in shock wave solutions (Hyman, 1979). Additionally, an adaptive gridding scheme is used to further enhance the accuracy of the numerical solution (Baer et al., 1986).

To incorporate a piston boundary condition in this numerical solution, a coordinate transformation is defined:

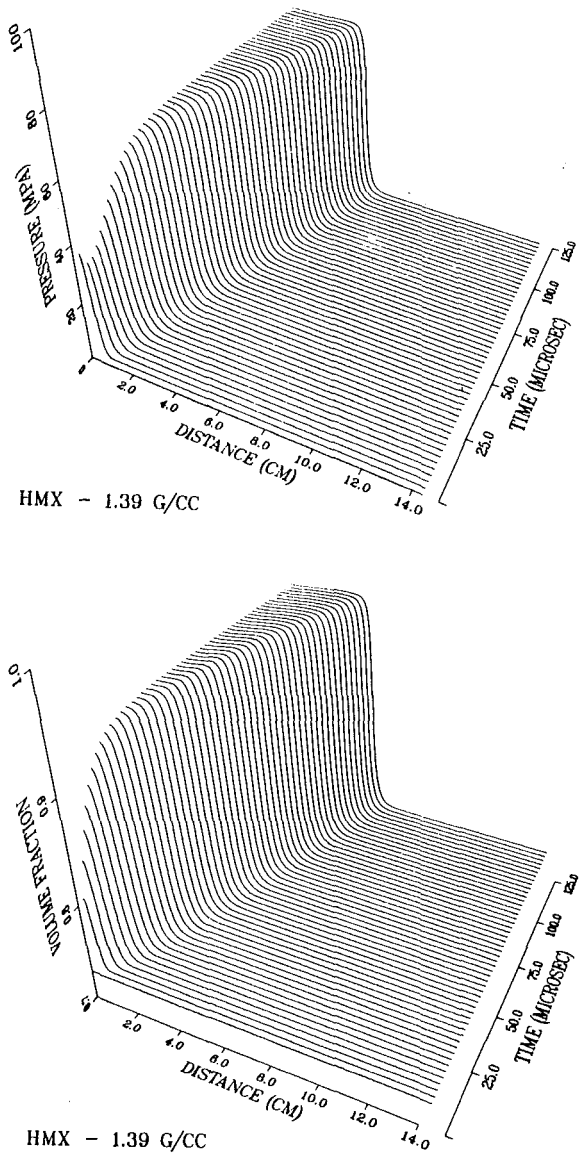


Fig. 5 Mixture pressure and solid volume fraction wave forms showing development and growth of compaction wave in a granular column of HMX impacted at 100 [m/s]

$$\eta = \frac{x - x_0(t)}{x_L(t) - x_0(t)}; \quad \eta \epsilon (0,1), \quad \tau = t; \quad \tau \geq 0 \tag{8}$$

where $x_L(t)$ and, $x_0(t)$ are, respectively, the right and left locations of the computation boundaries. With this transformation, boundary motion appears explicitly in the equations of motion and the computation domain is transformed to a fixed space. Thus, a conservation equation of the form:

$$\frac{\partial \rho}{\partial t} + \frac{\partial f}{\partial x} = S \tag{9}$$

transforms to:

$$\frac{\partial \rho}{\partial \tau} - \left(\frac{\dot{x}_0 + \eta(\dot{x}_L - \dot{x}_0)}{(x_L - x_0)}\right)\frac{\partial \rho}{\partial \eta} + \frac{1}{(x_L - x_0)}\frac{\partial f}{\partial \eta} = S. \tag{10}$$

A grid convective term appears in terms of the boundary velocities, $v_g = (\dot{x}_0, \dot{x}_L)$ at $\eta = (0,1)$, respectively. To recast the equation in conservative form, the bed length, $L = (x_L - x_0)$, is introduced and equation (10) becomes

$$\frac{\partial}{\partial \tau}(\rho L) + \frac{\partial}{\partial \eta}(f - (\dot{x}_0 + \eta \dot{L})\rho) = SL. \tag{11}$$

Table 2 Unstressed granular material data

Material	ϕ_s^o	particle size d_p [μm]
HMX, $C_4H_8N_8O_8$	0.73	870.
Fluid A (Nitrocellulose)	0.57	35.0
Teflon C	0.60	30.0
Melamine	0.65	50.0

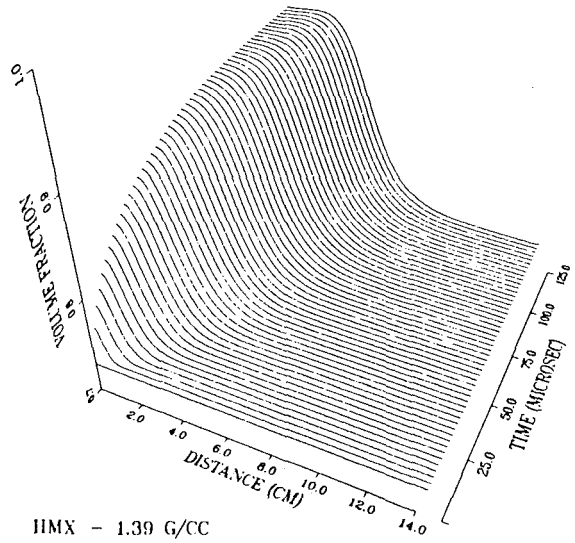


Fig. 6 Numerical solutions of solid volume fraction following impact using a material viscosity $\mu_c = 10^4$ [poise]

Numerical Calculations and Comparisons to Experiment

In this section, numerical results are compared with Sandusky and Liddiard's experimental observations. In these one-dimensional calculations, a 14.7 cm long porous column was represented using an adaptive computational scheme which varied from 50 to 200 nodes with eight levels of grid refinement (which significantly reduces the effects of numerical diffusion). The granular column was assumed to be composed of uniform-size particles and packed with a uniform initial solid volume fraction. Table (2) lists the initial states of the various energetic and inert materials considered in this study.

The first set of calculations/experimental results are presented for granular HMX. Figure (5) shows the time and space evolution of mixture pressure[2] and solid volume fraction for a typical case. Clearly, the induced wave motion has shock character and steady wave motion is seen to evolve following early time development.

To resolve the wave structure, marker particles were incorporated in the calculations to follow the fractional change of solid volume fraction across the compaction front. It was observed that a uniform compaction wave thickness was produced as the compaction wave traverses the porous column. Because a constant compaction wave thickness, δ_c, was predicted in the current work, the location where 50 percent of the change in solid volume fraction (across the compaction wave front) was followed and used in the determination of the compaction wave speed, V_c.

Material viscosity has an important influence on compaction wave structure; however, it has a negligible effect on the steady wave speed. Figure (6) shows the time evolution of volume fraction for $\mu_c = 10^4$ [poise]. As the material viscosity is increased, a thick compaction wave front is predicted. In the

[2]Mixture pressure is defined as:
$P_m = \phi_s P_s + \phi_g P_g + (\phi_s \rho_s v_s^2 + \phi_g \rho_g v_g^2) - (\phi_s \rho_s v_s + \phi_g \rho_g v_g)^2 / (\phi_s \rho_s + \phi_g \rho_g)$.

experiments, uniform displacement of the embedded disks was observed, suggesting that the compaction wave thickness is less than the disk spacing. The use of a material viscosity $\mu_c > 10^3$ [poise] produces relatively thick compaction waves which violates this observation. Thus, an upper limit on material viscosity is suggested from these experiments.

Shown in Figs. (7)–(9) are comparisons of predicted mixture pressure, compaction wave speed, and the percent theoretical maximum density of the compacted region to those observed by Sandusky and Liddiard for Class D HMX ($\phi_s^o = 0.73$) for various piston velocities, u_p. The experimental data are denoted by filled dots. Clearly, the mixture model predictions correspond well with the experimental observations. Table (3) provides detailed numerical-experimental results for HMX.

Sandusky and Liddiard (1985) report compaction studies for other energetic and inert materials. These experiments were also numerically modeled and predictions versus experimental observations of these studies are shown in Figs. (10)–(12). Agreement of wave speed, mixture pressure and the percent theoretical maximum density corresponds to a line with a slope of 1. Again, for all materials examined at the various loading conditions, good agreement with the experimental results is observed.

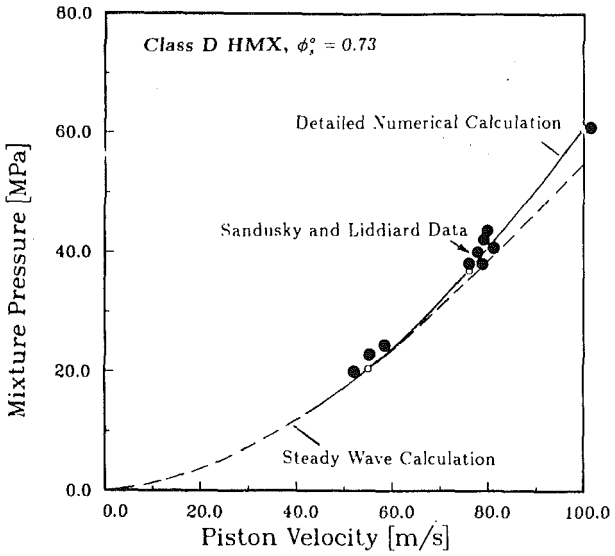


Fig. 7 Comparison of numerical and experimental results for mixture pressure following compaction produced at various piston impact velocities. The "experimental pressures" are deduced from a jump condition using experimental wave speed and compacted density.

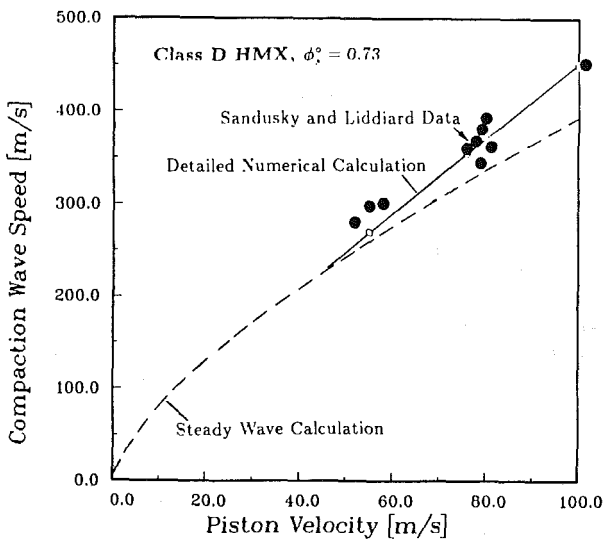


Fig. 8 Comparison of numerical and experimental results for compaction wave speed for various piston impact velocities

Steady-Wave Analysis

Since the detailed numerical calculations indicated that compaction waves evolve to produce steady-wave motion, a simplified analysis has been examined. Steady compaction wave behavior has been previously investigated by Kooker (1986), however, compaction was treated as a discontinuous

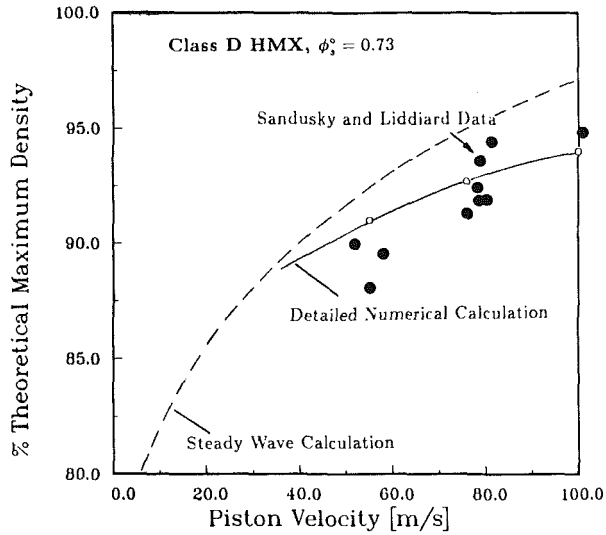


Fig. 9 Comparison of numerical and experimental results for percent theoretical maximum density of granular material following compaction for various piston impact velocities

Table 3 Numerical-experimental data for HMX

HMX ($\phi_s^o = 0.73$) Calculations						Experiment			
u_p [m/s]	μ_c [poise]	V_c [m/s]	δ_c [cm]	P_m [MPa]	%TMD	u_p [m/s]	V_c [m/s]	P_m [MPa]	%TMD
100.	10^2	440.	0.04	60.0	94.7	101.	432.	62.1	95.2
	10^3	450.	0.45	60.5	94.0				
	10^4	400.	3.19	55.0	93.0				
						80.	393.	43.7	91.8
						79.	382.	42.1	92.3
						78.	368.	40.0	92.7
76.0	10^3	355.	0.51	37.0	92.7	81.	364.	41.0	93.9
						79.	347.	38.1	94.5
						76.	359.	37.9	92.6
55.0	10^3	270.	0.53	20.5	91.0	58.	304.	24.6	90.2
						55.	300.	23.	89.4
						52.	280.	20.3	89.6

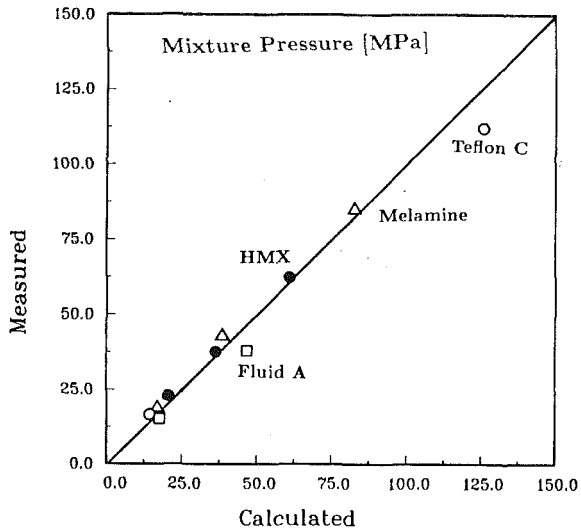


Fig. 10 Comparison of numerical and experimental results for mixture pressure following compaction of various inert and energetic materials. The "experimental pressures" are deduced from a jump condition using experimental wave speed and compacted density.

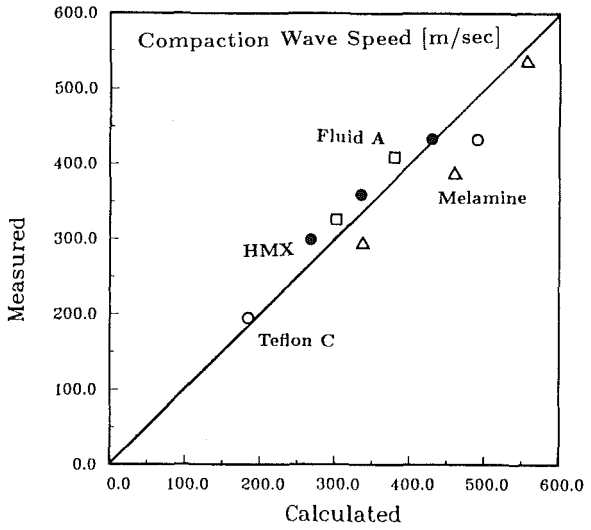


Fig. 11 Comparison of numerical and experimental results for compaction wave speed of various inert and energetic materials

wave.[3] In this study, a steady-wave analysis is examined which allows compaction with finite wave structure.

Since the detailed numerical solution also confirmed that gas influences are minor, only the solid-phase equations are used. In the solid-phase energy equation, the compaction work is included as given by the last term in equation (6). To simplify notation, the density variations in the solid equation of state (7) are collected into the quantity denoted as $G(\rho_s)$:

$$P_s = \Gamma_s \rho_s e_s + G(\rho_s). \tag{12}$$

Next, a wave coordinate is introduced as $\zeta = x - V_c t$ and the steady equations of motion transform to:

$$\frac{d}{d\zeta}(\phi_s \rho_s (v_s - V_c)) = 0, \tag{13}$$

$$\frac{d}{d\zeta}(\phi_s P_s + \phi_s \rho_s v_s (v_s - V_c)) = 0, \tag{14}$$

[3] Additionally, Kooker treats gas phase effects in his compaction shock analysis.

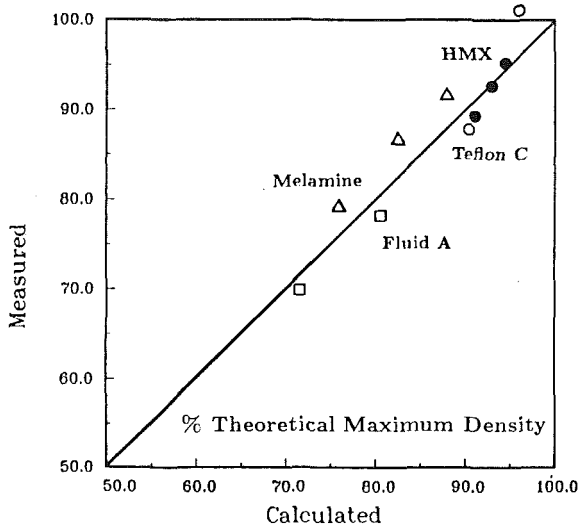


Fig. 12 Comparison of numerical and experimental results for percent theoretical maximum density following compaction for various materials

$$\frac{d}{d\zeta}(\phi_s \rho_s E_s (v_s - V_c) + \phi_s P_s v_s)$$

$$= -(P_s - \beta_s)(v_s - V_c)\frac{d\phi_s}{d\zeta}, \tag{15}$$

and

$$(v_s - V_c)\frac{d\phi_s}{d\zeta} = \frac{f(\phi_s)}{\mu_c}(P_s - \beta_s). \tag{16}$$

With the use of the equation of state (12), the energy equation (15) is recast into the following form:

$$\frac{d}{d\zeta}\left(e_s \rho_s^{-\Gamma_s} - \int_{\rho_s^o}^{\rho_s} \frac{G(\rho)}{\rho^{(2+\Gamma_N)}} d\rho\right)$$

$$= \rho_s^{-(\Gamma_s+1)} \frac{d}{d\zeta}\left(\int_{\phi_N^o}^{\phi_N} \frac{\beta_s(\phi)}{\phi} d\phi\right). \tag{17}$$

For the conditions of interest, compressibility effects are small and the material density on the right side of equation (17) is approximated as a constant. Thus, equations (13), (14), and (17) can be integrated corresponding to the condition that the granular material ahead of the compaction wave is undisturbed (or P_s, e_s, $v_s = 0$; $\phi_s = \phi_s^o$; $\rho_s = \rho_s^o$);

$$\phi_s \rho_s (V_c - v_s) = \phi_s^o \rho_s^o V_c, \tag{18}$$

$$\phi_s P_s + \phi_s \rho_s v_s (v_s - V_c) = 0, \tag{19}$$

$$e_s = \rho_s^{\Gamma_s} \int_{\rho_s^o}^{\rho_s} \frac{G(\rho)}{\rho^{(2+\Gamma_s)}} d\rho + \frac{1}{\rho_s} \int_{\phi_s^o}^{\phi_s} \frac{\beta_s(\phi)}{\phi} d\phi. \tag{20}$$

From equation (16), the variation in solid volume fraction is seen to be asymptotic to the condition, $P_s = \beta_s(\phi_s)$. Furthermore, at the piston interface ($\zeta \to -\infty$) the particle velocity must approach the piston velocity. With these conditions, compaction wave speed is obtained by combining equation (12) and equations (18)–(20). After determining the wave speed, the wave profile is then determined by an ordinary differential equation (16). Thus, the compaction wave speed and the asymptotic states following compaction can be resolved independent of the compaction wave structure and material viscosity.

As representative steady wave calculations, the constant-velocity impact experiments for granular HMX are examined using the model inputs of the detailed numerical study. The steady-wave solutions for compaction wave speed, the solid volume fraction following compaction, and pressure for

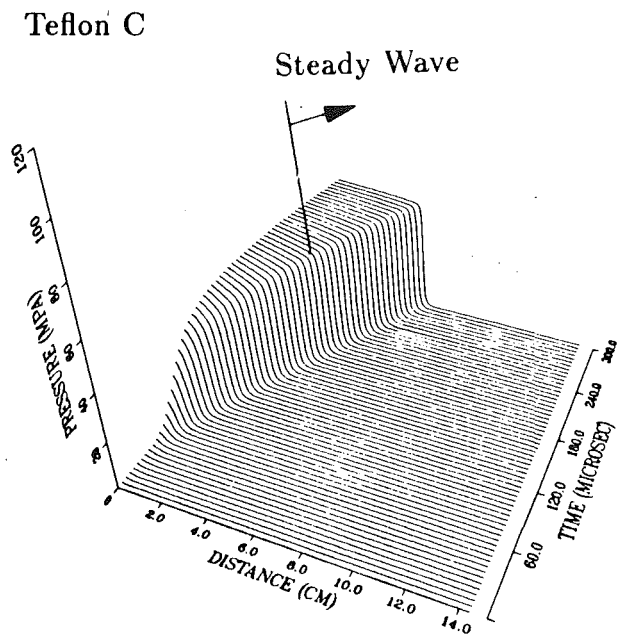


Fig. 13 Mixture pressure wave development in granular Teflon column predicted for the ramp-wave experiment

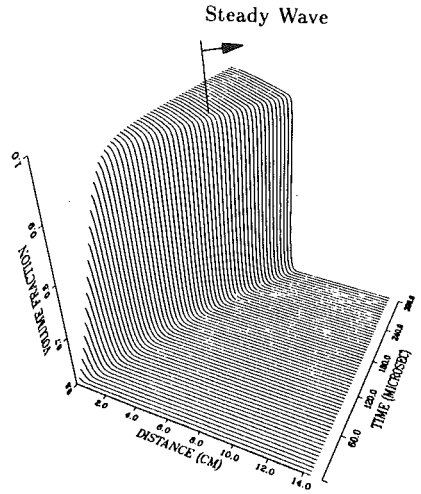


Fig. 14 Solid volume fraction development in granular Teflon column predicted for the ramp-wave experiment.

various piston velocity conditions are shown in Figs. (7)–(9) and compared to the detailed model and experimental observation. A reasonable estimate of compaction behavior is obtained using the steady-wave analysis. Some offset to the detailed model and the experimental results are seen and can be traced to the neglect of any gas-phase influence in the steady-wave analysis. In particular, the asymptotic state following compaction requires $P_s = \beta_s + P_g$. Future studies will address effects of the gas-phase in steady compaction wave behavior.

A Time-Varying Piston Impact Study

From the constant-velocity impact studies, it is seen that the transient nature of compaction occurs during early time wave development. Unfortunately, the existing experimental studies lack the resolution in time and space to monitor the evolution of a compaction wave. In an attempt to probe early time wave growth, Sandusky and Liddiard report a single experiment in which a time-varying piston (ramp) load is imparted on a Teflon C granular column. Measurements of the compaction wave motion were obtained using high-speed photography of backlit transparent tracers at 2.5 [cm] intervals along the axial length of the granular column.

In a numerical simulation, the experimental time-varying piston velocity was used as a model input. Wave characteristics of mixture pressure and solid volume fraction are shown in Figs. (13) and (14) using a material viscosity of $\mu_c = 10^3$ [poise]. As observed by Sandusky and Liddiard, the early wave growth eventually forms a compaction shock approximately 4 [cm] from the bed-piston interface. When the material viscosity exceeds $\mu_c > 10^4$ [poise] a thick compaction wave evolves and the onset to steady wave motion is significantly delayed, contrary to experimental observation.

Summary and Conclusions

In this study, numerical predictions of piston-impact experimental compaction studies have been addressed. Steady compaction shock waves are predicted for constant-velocity piston impacts. Wave speed, mixture stress, and the jump in solid volume fraction across the front compare reasonably well with experimental observations. A variety of materials and loading conditions have been studied and in each case the multiphase hydrodynamics is shown to be predicted reasonably well using a compaction rate law that is driven by pressure differences.

In this compaction study, the only varied parameter is the material viscosity. This granular property has been shown to be of importance in determining early wave development and in establishing the structure of the compaction front. Steady-wave analysis has been investigated which suggests that the rate dependence on the solid volume fraction does not significantly influence the compaction wave speed or the global wave characteristics. Qualitative estimates of compaction wave behavior can be obtained using the simplified steady-wave model.

Although Sandusky and Liddiard's piston-impact experiments provide an important data base for model comparison, sufficient details are lacking to determine the material viscosity. It is suggested that transient pressure measurements near the piston interface or spatial-resolved measurements of the compaction density across the wave front would lead to this determination.

An additional study of a ramp-wave compaction experiment has also been successfully described by the multiphase mixture model and a bound on the material viscosity has been implied. Detailed measurements of the early time wave development would provide additional information regarding the compaction wave characteristics.

Acknowledgments

I gratefully acknowledge the many fruitful discussions with J. W. Nunziato (SNLA) and D. Kooker (BRL). I also thank R. E. Benner (SNLA) for his assistance with the adaptive gridding used in the numerical model. Finally, I thank the reviewers of this paper for the suggested modifications and corrections of this work.

This work was performed at Sandia National Laboratories, supported by the US Department of Energy under contract number DE-AC04-76DP00789.

References

Baer, M. R., and Nunziato, J. W., 1986, "A Two-Phase Mixture Theory for the Deflagration-to-Detonation Transition (DDT) in Reactive Granular Materials," *Int. J. of Multiphase Flow*, Vol. 12, No. 6, pp. 861–889.

Baer, M. R., Benner, R. E., Gross, R. J., and Nunziato, J. W., 1986, "Modeling and Computation of Deflagration to Detonation Transition (DDT) in Reactive Granular Materials," *Lectures in Applied Mathematics*, Vol. 24, pp. 479–498.

Butcher, B. M., Carroll, M. M., and Holt, A. C., 1974, "Shock-Wave Compaction of Porous Aluminum," *J. of Applied Physics*, Vol. 45, pp. 3864–3875.

Butler, P. B., Lembeck, M. F., and Krier, H., 1982, "Modeling of Shock Development and Transition to Detonation Initiated by Burning in Porous Propellant Beds," *Combustion and Flame*, Vol. 46, pp. 75–93.

Carroll, M. M., and Holt, A., 1972, "Static and Dynamic Pore-Collapse for Ductile Porous Materials," *J. of Applied Physics*, Vol. 43, pp. 1627–1635.

Dobratz, B. M., 1981. "LLNL Explosives Handbook Properties of Chemical Explosives and Chemical Simulants," Lawrence Livermore National Laboratory, UCRL-52997.

Elban, W. L., Gross, S. B., Kim, K., and Bernecker, R. R., 1982, "Quasi-Static Compaction Studies for DDT Investigations: Inert Materials," Naval Surface Weapon Center, NSWC TR 81-113.

Elban, W. L., 1984, "Quasi-Static Compaction Studies for DDT Investigations: Ball Propellants," *Propellants, Explosives and Pyrotechnics*, Vol. 9, pp. 119–129.

Elban, W. L., and Chiarito, M. A., 1986, "Quasi-Static Compaction Study of Coarse HMX Explosive," *Powder Technology*, Vol. 46, pp. 181–193.

Gibbs, T. R., and Popolato, A., 1980, *LASL Explosive Property Data*, University of California Press, Berkeley, CA.

Herrmann, W., 1969, "Constitutive Equation for the Dynamic Compaction of Ductile Porous Materials," *J. Appl. Phys.*, Vol. 40, pp. 2490–2499.

Hyman, J. M., 1979, "A Method of Lines Approach to the Numerical Solution of Conservation Laws," Rept. LA-VR 79-837, Los Alamos National Laboratory.

Khasainov, B. A., Borisov, A. A., Ermolaev, B. S., and Korotkov, A. I., 1981, "Two Phase Visco-Plastic Model of Shock Initiation of Detonation in High Density Pressed Explosives," *Seventh Symposium (International) on Detonation*, Naval Surface Weapons Center, NSWC MP 82-334, pp. 435–447.

Kim, K., 1984, "Numerical Simulation of Convective Combustion of Ball Powders in Strong Confinement," *AIAA*, Vol. 22, pp. 793–796.

Kooker, D. E., and Anderson, R. D., 1981, "A Mechanism for the Burning Rate of High Density, Porous Energetic Materials," *Seventh Symposium (International) on Detonation*, pp. 198–215.

Kooker, D. E., 1986, "A Numerical Study of Compaction Waves in Class D HMX," JANNAF Propulsion System Hazards Meeting, March 3–7, Monterey, CA.

Marsh, S. P., ed., 1980, *LASL Shock Hugoniot Data*, University of California Press, Berkeley, CA.

Maiden, D. E., and Nutt, G. L., 1986, "A Hot-Spot Model for Calculating the Threshold for Shock Initiation of Pyrotechnic Mixtures," *Proceedings of the 11th International Pyrotechnics Seminar*, pp. 813–826.

Nunziato, J. W., 1984, "Initiation and Growth-to-Detonation in Reactive Mixtures," *Shock Waves in Condensed Matter – 1983*, J. R. Asay, R. A. Graham, G. K. Straub, eds., Elsevier, New York, pp. 581–588.

Nunziato, J. W., and Walsh, E. K., 1980, "On Ideal Multiphase Mixtures with Chemical Reactions and Diffusion," *Arch. Rational Mech. Anal.*, Vol. 73, pp. 285–311.

Rayleigh, J. W. S., 1917, "On the Pressure Developed in a Liquid During the Collapse of a Spherical Cavity," *Phil. Mag.*, Vol. 34, pp. 94–98.

Sandusky, H. W., and Liddiard, T. P., 1985, "Dynamic Compaction of Porous Beds," Naval Surface Weapons Center, NSWC TR 83-246.

Sandusky, H. W., and Bernecker, R. R., 1985, "Compressive Reaction in Porous Beds of Energetic Materials," *Eighth Symposium (International) on Detonation*, pp. 631–640.

Sheffield, S. A., Hayes, D. B., and Mitchell, D. E., 1977, "An Equation of State and Chemical Kinetics for Hexanitrostilbene (HNS) Explosive," *Sixth Symposium (International) on Detonation*, pp. 748–754.

A Nonlinear Analysis of an Equilibrium Craze: Part I—Problem Formulation and Solution

T. Ungsuwarungsri
Graduate Research Assistant.

W. G. Knauss
Professor of Aeronautics and Applied Mechanics.

Graduate Aeronautical Laboratories, California Institute of Technology, Pasadena, CA 91125

This study investigates the effect of nonlinear cohesive forces on crack growth with the special problem of craze mechanics in mind. The work is presented in two parts. In the first and present one, we develop a numerical method for determining the equilibrium shape of a craze in an infinite elastic plane whose fibrils exhibit very general nonlinear force-displacement (P-V) behavior, including strain softening characteristics. The second part of this study deals with the numerical simulation of craze and crack growth (Ungsuwarungsri and Knauss, 1986).[1] The problem formulation is based on the superposition of the relevant elasticity Green's function and the solution for the resulting nonlinear problem is effected by using Picard's successive approximation scheme. Both field equilibrium and the Barenblatt condition for vanishing stress and strain singularities are satisfied simultaneously, rendering the craze tip profile cusp-like. The formulation allows the stress distribution profile and the corresponding P-V relation to be computed from experimentally measured craze/crack displacement contours; it also allows the computation of the craze or crack/craze profile if the P-V relation, far-field load, and craze or crack size are specified. Numerical investigations indicate that only certain classes of the fibril P-V relations are consistent with measured craze profiles. In addition, it is found that for a given P-V relation, nontrivial solutions exist only for certain ranges of craze lengths.

1 Introduction

When glassy polymers are subjected to extensional strains larger than about 0.2 percent, they begin to exhibit local "plastic" deformation. Two distinct processes of plastic deformations have been observed, one being mainly a shearing phenomenon such as diffuse shear yielding and localized shear band formation, and the other being caused by cavitation and void growth commonly called "crazing." The latter process is characterized by a loss of intermolecular cohesion, molecular flow, and orientation as well as a significant decrease in local density (Bucknall, 1977). In this work, we shall concern ourselves only with phenomena related to crazing.

The subject of crazing has been extensively studied by both material (polymer) scientists and applied mechanicians. We shall discuss briefly the pertinent physical concepts and the motivations leading to the present investigation.

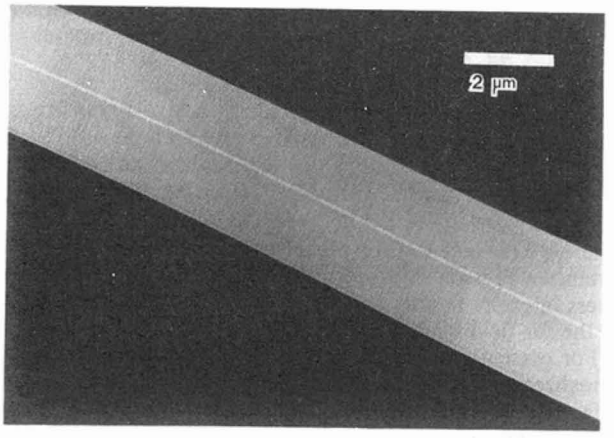
It has been recognized for some time that fracture in thermoplastic polymers is usually preceded by the formation of crazes. Crazes are formed in the bulk polymer along planes normal to the direction of maximum principal tensile stress (Sternstein et al., 1968); they tend to nucleate at microdefects or inclusions and grow by the formation of fibrils from the bulk polymer (Argon and Salama, 1977; Lauterwasser and Kramer, 1979) and thicken through a combination of two distinct mechanisms, namely through "surface drawing" and "creep." In surface drawing, the new polymer is drawn from the bulk into the craze of in the form of fibrils. During creep existing fibrils are simply elongated, mostly locally around the weak portions. Although "creep" implicitly signifies a time dependent process, we will exclude time consideration in this work.[2] For our purpose it is sufficient to recognize that the interplay of these two mechanisms gives rise to an effective force-displacement (P-V) behavior which we consider to be initially prescribed. Here "force" represents the cohesive stress exerted by the fibril on the craze/bulk polymer boundary. "Displacements" refers to the net displacement of the craze boundary (see Section 2.1 for details).

[1] The unabridged version of these two papers (Ungsuwarungsri and Knauss, 1985) is given in the reference list.

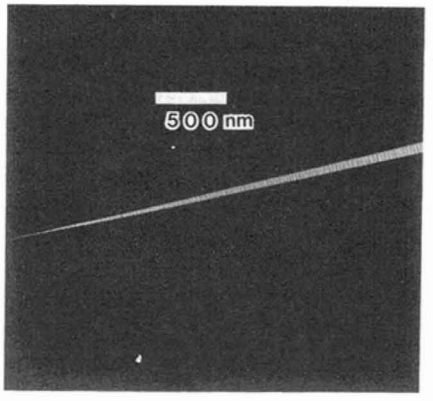
Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 21, 1986; final revision August 25, 1987.

[2] For an analysis accounting for these two basic mechanisms, see the paper by Kramer and Hart (1984).



(a)



(b)

Fig. 1 Electron micrographs of a craze (*a*) and its tip region (*b*). Courtesy of L. Berger and E. J. Kramer

Electron microscopy shows that the typical craze thickness is on the order of 1 μm and the length to thickness (aspect) ratio is approximately 10^2 to 10^3. Crazes generally consist of a network of fibrils between 10 to 40 nm in diameter, interspersed with voids of similar dimensions. The transition boundary between a craze and the bulk polymer where surface drawing takes place is very thin, i.e., less than 2 nm. Craze fibrils have been shown to exhibit considerable strength (Spurr and Niegisch, 1962; Kambour and Kopp, 1969). Therefore, crazing could, under certain conditions, substantially enhance the toughness of bulk polymers. However the presence of crazes does not necessarily enhance the strength of the overall structure. For more in-depth discussions on microstructures and properties of crazes, we refer to the reader to the references at the end of this paper. Figure 1 shows a typical craze and its tip region.[3]

While it would be appropriate to review in some detail the analytical approaches used to model the crazing phenomenon, we refrain from doing so in the interest of brevity. Instead we list the authors of relevant contributions. Papers pertinent to nonsingular stresses at tips of cracks or crazes were contributed by Prandtl (1933), Dugdale (1960), and Barenblatt (1962); applications of cohesive force models were published by Goodier and Field (1963), Goodier and Kanninen (1966), and Rice (1966). Models with cohesive forces that depend on craze or crack-opening displacements have been contributed with simultaneous restrictions on nonsingular stress fields by Anderson and Bergkvist (1970) and Knauss (1974, 1976), and with allowance for singular stresses at the tip by Atkinson (1970), Walton and Weitsman (1984), and Weitsman (1986), where the latter two publications included a small-scale correction to the stress field of the Barenblatt-Dugdale type. In addition, analyses strongly motivated and/or supported by experiments were offered by Knight (1965), Verheulpen-Heymans and Bauwens (1976), Lauterwasser and Kramer (1979), Wang and Kramer (1982), and Donald and Kramer (1982), as well as Beran (1982).

In most of these studies the main objective was the determination of the stress distribution along the craze boundary from experimental displacement profiles. A major purpose of the present study is, however, to address the somewhat more demanding problem of determining both the stress and displacement profiles along the craze or crack-plus-craze given a force-extension relation for the fibrils such that a Barenblatt-Dugdale condition is also satisfied simultaneously. Upon examining the above references it becomes evident that a computationally efficient method for this purpose is still needed when nonlinear force-displacement relations (p-v relations) for the fibrils are involved. The primary objective of this work is to present such an analysis and to use it in the study of various aspects of the mechanics of craze and crack growth in thermoplastic materials.

In the following sections, we discuss the model for representing a craze. The idealizations made to model the problem are explained in Section 2.1. The mathematical formulation including nondimensionalization and discretization for final implementation then follow in Sections 2.2 and 2.3. In Section 3, two alternate methods for computing stress distribution profiles from craze displacement contours are presented for the case of a full craze and for a craze with a central crack. Because of their simplicity, these methods are both computationally more efficient and inherently more accurate (for the same degree of discretization) than those used in the past by other investigators. The convergence characteristics of the solution technique are briefly explored and the general applicability and performance of the numerical schemes are assessed in Section 4.1. Some characteristics of the fibril P-V relations that are inconsistent with realistic craze profiles are then examined and the existence of solutions is addressed briefly in Section 4.2. The question of the uniqueness of solutions is raised but not addressed rigorously.

2 Problem Formulations and Solutions

In this section, we reduce the physical problem to a mathematically manageable one through suitable approximations. The resulting equations are then nondimensionalized and solved through a numerical algorithm based on Picard's iteration.

2.1 Geometric and Mathematical Idealization. We consider first the geometry of the problem as depicted in Fig. 2(*a*). A craze of length C with an (internal) central crack of length A is embedded in an infinite plane subjected to a remote and uniform tensile stress Σ_∞ applied normal to the major axis of the craze. X, Y are the coordinates and $W(X)$ denotes the Y-coordinate of the craze contour. We let $P(X)$ represent the cohesive force (stress) distribution exerted by the fibrils.

As shown by Lauterwasser and Kramer (1979), a craze is formed from a primordial craze, characterized by a contour $W_o(X)$: this contour contains the bulk polymer material which is subsequently transformed into craze fibrils through the surface drawing process. Thus the dashed contour $W_o(X)$ in Fig. 2(*b*) is displaced in the drawing process to the solid contour $W(X)$ due to the application of Σ_∞. Since we know from experiments that $W_o(X)$ is much smaller than C and that the slope, $W_o'(X)$, is also very small everywhere, we can recast the

[3] The micrographs were taken by L. Berger and made available to us by Professor E. J. Kramer.

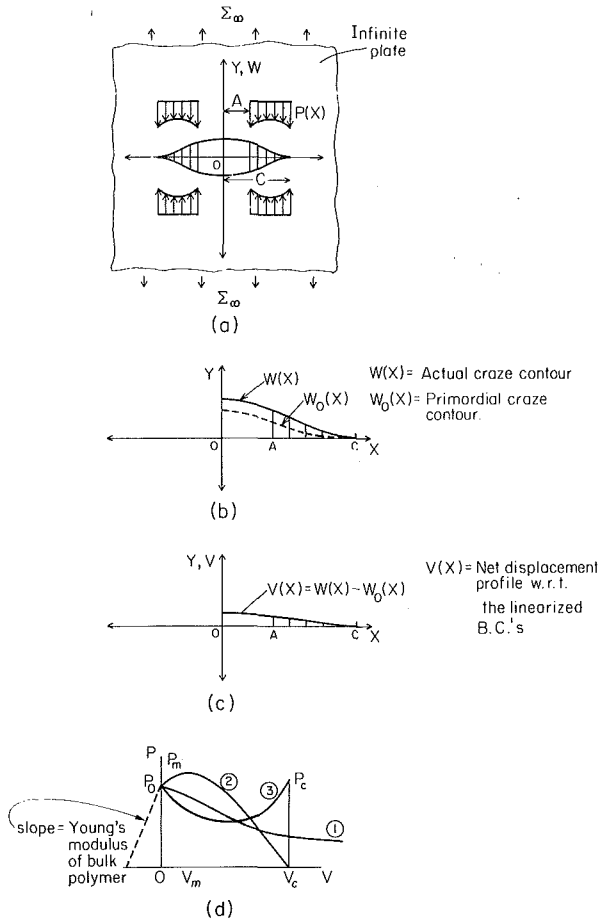


Fig. 2 (a) Problem geometry definition; (b) actual craze contour with primordial craze contour; (c) linearized geometry: net displacement profile; (d) model for craze fibril p-v relation

geometry of the problem as shown in Fig. 2(*c*) where $V(X)$ is the net displacement of the craze profile caused by the applied stress as measured from $W_o(X)$ so that $V(X)=W(X)-W_o(X)$. Cast in this light, the force-displacement (P-V) relation may be represented schematically as shown in Fig. 2(*d*). Keeping in mind that $V<0$ corresponds to the primordial craze under small strain prior to fibrillation, the dashed portion of the P-V curve is assumed linear with a slope equal to Young's modulus, E, for the bulk polymer. At $V=0$, the threshold (or fibrillation) stress P_o is reached and from then on the P-V relation is nonlinear. Three P-V relations are illustrated, curve #1 is representative of a craze that does not contain a crack. We shall later refer to such a craze as a "full craze." Curve #2 allows for continued hardening before softening to failure (fibril degradation and rupture) at V_c. Curve #3 is typical of a craze with a central crack; this rehardening behavior may be understood in the context of the midrib behavior (Lauterwasser and Kramer, 1979)[4].

2.2 The Mathematical Formulation. By superposing the Green's functions obtained through Westergaard's stress function method (Tada et al., 1973), the solutions to the problem can be written as[5]

$$V(X)=\left(\frac{1+\eta}{4\mu}\right)\frac{1}{\pi}\int_0^C (\Sigma_\infty - P(T))\log\left|\frac{\sqrt{C^2-X^2}+\sqrt{C^2-T^2}}{\sqrt{C^2-X^2}-\sqrt{C^2-T^2}}\right|dT \tag{2.1}$$

$$K_I=2\sqrt{\frac{C}{\pi}}\left[\int_o^C \frac{\Sigma_\infty - P(T)}{\sqrt{C^2-T^2}}dT\right] \tag{2.2}$$

where $\eta=(3-\nu/1+\nu)$ for plane stress and $(3-4\nu)$ for plane-strain; ν = Poisson's ratio; μ = shear modulus. K_I is the Mode I stress intensity factor, and $V(X)$ is the net displacement due to the far-field loading and $P(t)$ the fibril cohesive stress.

For presentation of the results, it is helpful to nondimensionalize pertinent parameters. We let a craze length C_o be the reference length. For later reference when craze growth is involved we append the suffix "0" to denote a stationary craze. The stresses are scaled by $(4\mu/1+\nu)$; in this fashion, the Poisson's ratio is absorbed for both cases of plane stress and plane strain. In short,

$$C_{ref}=C_o$$

$$\Sigma_{ref}=\frac{E}{2} \quad \text{for plane stress}$$

$$=\frac{E}{2(1-\nu^2)} \quad \text{for plane strain}$$

We shall use lower case letters for dimensionless quantities and define

$$C\equiv\frac{C}{C_o},\quad a\equiv\frac{A}{C_o},\quad x\equiv\frac{X}{C_o},\quad y\equiv\frac{Y}{C_o}$$

$$w\equiv\frac{W}{C_o},\quad w_o\equiv\frac{W_o}{C_o},\quad v\equiv\frac{V}{C_o},\quad \tau\equiv\frac{T}{C_o}$$

and, $\sigma_\infty\equiv\dfrac{\Sigma_\infty}{\Sigma_{ref}},\quad p\equiv\dfrac{P}{\Sigma_{ref}}$

Thus, in dimensionless forms, equations (2.1) and (2.2) are

$$v(x)=\int_o^c [\sigma_\infty - p(\tau)]m(c,x,\tau)d\tau \tag{2.3}$$

$$k_I=2\sqrt{\frac{c}{\pi}}\left[\int_o^c \frac{\sigma_\infty - p(\tau)}{\sqrt{c^2-\tau^2}}d\tau\right] \tag{2.4}$$

where the kernel in equations (2.3) is given by

$$m(c,x,\tau)=\frac{1}{\pi}\log\left|\frac{\sqrt{c^2-x^2}+\sqrt{c^2-\tau^2}}{\sqrt{c^2-x^2}-\sqrt{c^2-\tau^2}}\right| \tag{2.5}$$

Note that here k_I is dimensionless. Integrating equation (2.3) yields

$$v(x)=\sigma_\infty\sqrt{c^2-x^2}-\int_o^c m(c,x,\tau)p(\tau)d\tau \tag{2.6}$$

For the stress at the craze tip to be finite, or equivalently, for the craze tip to close smoothly, i.e., $v'(c)=0$, it is necessary and sufficient that k_I vanish. Then equation (2.4) gives

$$\sigma_\infty=\frac{2}{\pi}\int_o^c \frac{p(\tau)}{\sqrt{c^2-\tau^2}}d\tau \tag{2.7}$$

Substituting σ_∞ from equations (2.7) into (2.6), one obtains

$$v(x)=\int_o^c h(c,x,\tau)p(\tau)d\tau \tag{2.8}$$

where

$$h(c,x,\tau)=\frac{2}{\pi}\sqrt{\frac{c^2-x^2}{c^2-\tau^2}}-m(c,x,\tau) \tag{2.9}$$

[4]The midrib is formed initially at the craze tip due to fibril drawing at high stress. As the craze grows, the midrib is unloaded as it becomes part of the mature (developed) craze. The introduction of a crack in the craze causes the stress at the crack tip to rise, so that the midrib is reloaded and therefore must sustain higher stress.

[5]Huang (1985) independently employed an identical formulation in studying the size of the crack tip cohesive zone but based on an interatomic potential.

The formulation is completed by specifying the fibril force-displacement behavior[6]

$$p(x) = p[v(x)] \tag{2.10}$$

Equations (2.8) and (2.10) must be solved simultaneously. This is accomplished by using Picard's method of successive approximation outlined below[7]

We start with an initial guess of the form,

$$v^o(x) = \sigma_\infty^o (c-x)^r \tag{2.11}$$

where σ_∞^o and r are to be appropriately chosen; other forms of $v^o(x)$ are also acceptable.

The kth approximation is given by

$$v^k(x) = \int_o^c h(c,x,\tau) p[v^{k-1}(\tau)] d\tau; \quad k = 1,2, \ldots \tag{2.12}$$

We consider that a solution has been attained when $v^k(x)$ differs from $v^{k-1}(x)$ at *all* (discretized) points by 0.1 percent or less. The convergence of the scheme is found to be relatively insensitive to the choice of the initial guess of σ_∞^o and r (see Section 3.1).

It is worth noting that in the case where a crack is present as determined by the critical crack tip opening displacement (CTOD) criterion,[8] the crack length is *not fixed a priori* but is obtained as part of the solution. We next discretize the equations for numerical implementation.

2.3 Discretization and the Resulting Algorithm. Because the kernels $m(c,x,\tau)$ and $h(c,x,\tau)$ defined by equations (2.5) and (2.9) can be evaluated in closed form, the integration of these "singular" integrals can be calculated very accurately without requiring excessive subdivisions of the domain.[9] We discretize the domain of integration into n equal intervals of length Δx, where

$$\Delta x = \Delta\tau = c/n$$

and x_i, t_i, v_i, p_i are evaluated at the midpoint of each interval. With this in mind, an integral $I(x)$ can be approximated as[10]

$$I(x_i) \equiv \int_o^c f(c,x_i,\tau) p(\tau) d\tau \equiv \bar{f}_j(c,x_i) p_j = \bar{f}_{ij} p_j \tag{2.13}$$

where

$$\bar{f}_{ij} \equiv \bar{f}_j(c,x_i) \equiv \int_{\tau_j - \Delta\tau/2}^{\tau_j + \Delta\tau/2} f(c,x_i,\tau) d\tau \tag{2.14}$$

Integrals of the form (2.14) in this study are easily evaluated in closed forms; specifically, from equations (2.5) and (2.6), one obtains

$$v_i = \sigma_\infty \sqrt{c^2 - x_i^2} - \bar{m}_{ij} p_j \tag{2.15}$$

where

$$\bar{m}_{ij} \equiv \frac{1}{\pi}\left[\tau \log\left|\frac{\sqrt{c^2-x_i^2}+\sqrt{c^2-\tau^2}}{\sqrt{c^2-x_i^2}-\sqrt{c^2-\tau^2}}\right| - 2\sqrt{c^2-x_i^2}\cos^{-1}\left(\frac{\tau}{c}\right) + x_i \log\left|\frac{x_i\sqrt{c^2-\tau^2}-\tau\sqrt{c^2-x_i^2}}{x_i\sqrt{c^2-\tau^2}+\tau\sqrt{c^2-x_i^2}}\right|\right]_{\tau=\tau_j-\Delta\tau/2}^{\tau=\tau_j+\Delta\tau/2} \tag{2.16}$$

Similarly, equation (2.7) may be written as

$$\sigma_\infty = \int_o^c l(c,\tau) p(\tau) d\tau = \bar{l}_j p_j \tag{2.17}$$

where $l(c,\tau) = \frac{2}{\pi}\frac{1}{\sqrt{c^2-\tau^2}}$ and

$$\bar{l}_j \equiv \frac{2}{\pi}\left[\sin^{-1}\left(\frac{\tau}{c}\right)\right]_{\tau=\tau_j-\Delta\tau/2}^{\tau=\tau_j+\Delta\tau/2} \tag{2.18}$$

In view of equations (2.8) and (2.9), we obtain in compact form,

$$v_i \equiv v(x_i) \equiv \bar{h}_{ij} p_j \tag{2.19}$$

where

$$\bar{h}_{ij} \equiv \sqrt{c^2 - x_i^2}\ \bar{l}_j - \bar{m}_{ij} \tag{2.20}$$[11]

It turns out from equations (2.16), (2.18), and (2.20) that $\bar{l}_j$ is independent of c while $\bar{m}_{ij}$ and $\bar{h}_{ij}$ are homogeneous in c. These properties prove useful in the study of craze and crack growth in the second part of this work (Ungsuwarungsri and Knauss, 1986). In concise form, one has

$$\bar{l}_j = \bar{l}_j(c,n) = \bar{l}_j(1,n) \tag{2.21}$$

$$\bar{m}_{ij} = \bar{m}_{ij}(c,n) = c \cdot \bar{m}_{ij}(1,n) \tag{2.22}$$

$$\bar{h}_{ij} = \bar{h}_{ij}(c,n) = c \cdot \bar{h}_{ij}(1,n) \tag{2.23}$$

Thus, for a given degree of discretization n, $\bar{l}_j$, $\bar{m}_{ij}$, and $\bar{h}_{ij}$ only need to be computed once for $c = 1$ and stored for later use. In the simulation of craze and crack growth (Ungsuwarungsri and Knauss, 1986) c increases continually, thus the properties exhibited in equations (2.21)–(2.23) reduce the computation time substantially. Note that $\bar{l}_j$, $\bar{m}_{ij}$, and $\bar{h}_{ij}$ are independent of the p-v relation considered.

The algorithm can now be summarized. Given a craze length c and p-v relation, we seek solutions in the form of $p(x)$, $v(x)$, and σ_∞ according to the following procedure:

Prescribed p-v relation : $p(x_i) = p_i = p(v(x_i)) = p(v_i)$ (2.24*a*)

Initial guess : $v^o(x_i) = v_i^o = \sigma_\infty^o (c - x_i)^r$ (2.24*b*)

kth iteration : $v_i^k = v^k(x_i) = \bar{h}_{ij}(c,n) p(v_j^{k-1})$ (2.24*c*)

$$p_i^k = p(v_i^k) \tag{2.24d}$$

$$\sigma_\infty^k = \bar{l}_j p_j^k \tag{2.24e}$$

Convergence is considered achieved when the maximal change for any v_j is less than 0.1 percent between two successive iterations.

3 Schemes for Determining Stress Profiles From Displacement Profiles

As mentioned in the introduction, for the case of a full craze, the stress distribution profile $p(x)$ may be obtained from the measured displacement profile $v(x)$ and σ_∞ using the Fourier transform method (Sneddon, 1951). However, for a craze with a central crack, only the displacement profile in the craze zone is needed to determine both $p(x)$ and σ_∞ using a distributed dislocation scheme (Wang and Kramer, 1982). Here we present two alternative methods for determining the stress profiles based on the formulation in the previous section; one method is suitable for a full craze, the other for a craze containing a central crack. Both schemes have been fully tested and found to be numerically more efficient than the methods previously proposed.

First we discuss the method suitable for the problem of a full craze. Given $v(x_i)$ and σ_∞, equation (2.15) can be rewritten as

[6] The more general case, $p(x) = p[v(x),x]$ can be handled within the framework of our formulation using the same solution scheme. A p-v relation with explicit dependence on x to account for the growth history is a more realistic candidate; however, due to lack of any such data, we confine our attention to the form (2.10).

[7] As it turns out, this simple scheme is much more efficient than the Newton iterative scheme which was also implemented and yielded the same results.

[8] See Ungsuwarungsri and Knauss (1986) for further discussion.

[9] For contributions about the singular points, the Cauchy Principal values of the integrals are used.

[10] Repeated indices imply summation from 1 to n.

[11] Note that the matrix $\bar{h}_{ij}$ is singular, i.e., given a displacement profile $v(x)$, the cohesive stress profile can only be determined to within an additive constant. This, however, does not affect the validity of our algorithm.

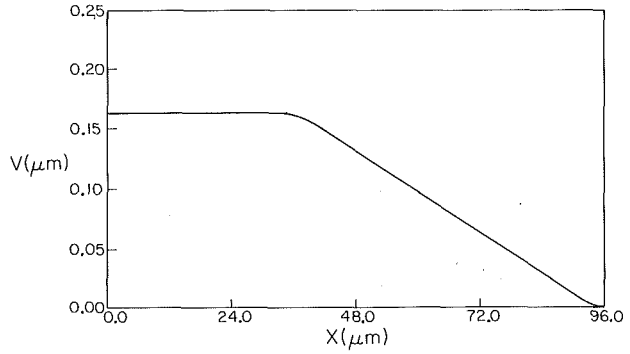


Fig. 3 Experimental craze contour from Lauterwasser and Kramer (1979)

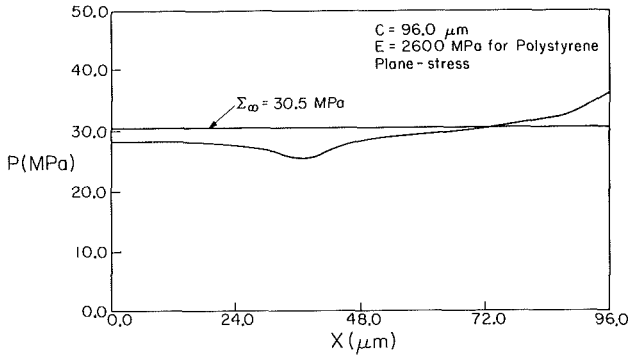


Fig. 4 Stress distribution profile derived from *V(X)* in Fig. 3 using the method in Section 3. This stress distribution guarantees a nonsingular stress at the tip of the craze.

$$\tilde{m}_{ij}p_j = (\sigma_\infty\sqrt{c^2 - x_i^2} - v_i) \tag{3.1}$$

Equation (3.1) is a system of n linear equations from which p_j can be determined either by Gaussian elimination or by inverting $\tilde{m}_{ij}$. In the latter case, equations (3.1) and (2.22) gives

$$p_j = \tilde{m}_{ij}^{-1}(c,n)(\sigma_\infty\sqrt{c^2 - x_i^2} - v_i) \tag{3.2}$$

where

$$\tilde{m}_{ij}^{-1}(c,n) = \tilde{m}_{ij}^{-1}(1,n)/c \tag{3.3}$$

which only needs to be computed once and stored for later use.

For demonstration purposes, this scheme is applied to $v(x)$ and σ_∞ taken from Lauterwasser and Kramer (1979), represented in Fig. 3. The resulting cohesive force distribution $p(x)$ shown in Fig. 4 and that derived by Lauterwasser and Kramer (1979) through the Fourier Transfer method are virtually identical. The p-v relation obtained by combining Figs. 3 and 4 is plotted in Fig. 5. Substitution of the p-v relation into the stress finiteness condition (2.7) or (2.17) demonstrates that it is also satisfied, i.e., no singular stresses at the craze tip are present.

This method is, of course, also applicable to a craze with an internal crack. However, since the measured displacement profiles in the crack region are usually not reliable on a point-by-point basis (Wang and Kramer, 1982), an alternative scheme using only the displacement profile in the craze zone as boundary data is preferable. Such a method has been presented by Wang and Kramer (1982); however, the method requires differentiation of the experimentally measured displacement profile which incurs additional error. We present here a scheme which requires as inputs only the displacement v_j in the craze zone in the crack tip location. The outputs are the stress profile $p(x)$, the applied stress σ_∞, and the displacement in the crack region.

Let i_t be such that the fibril cohesive force at the ith station is

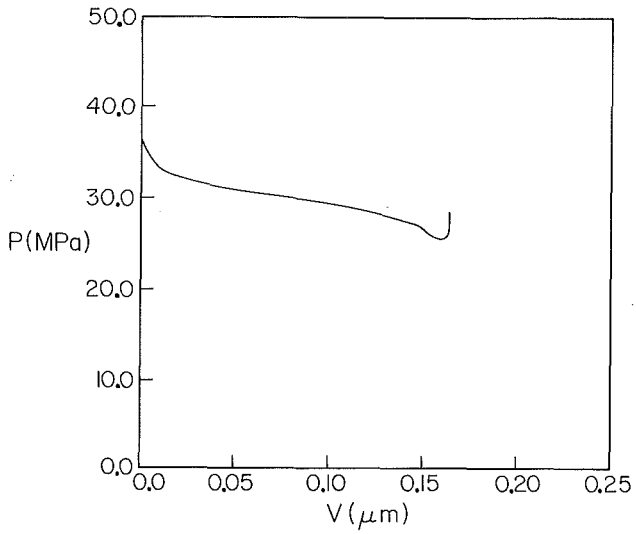


Fig. 5 The p-v relation derived from Figs. 3 and 4

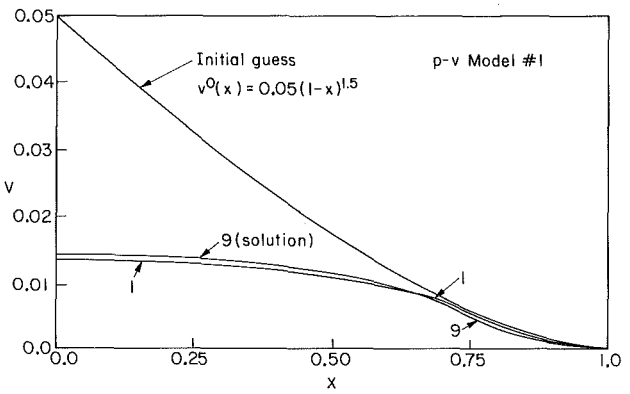


Fig. 6 Convergence illustration of v versus x ($\sigma_\infty^0 = 0.05$, $r = 1.5$). Iterations #2-8 fall in between iterations #1 and 9.

$$\begin{aligned} p_i &= 0 \quad \text{for} \quad i < i_t \\ p_i &\neq 0 \quad \text{for} \quad i \geq i_t \end{aligned} \tag{3.4}$$

Given the displacement in the craze zone, v_i for $i \geq i_t$, we wish to find σ_∞ and p_i for $i \geq i_t$ and v_i for $i < i_t$. Consider equation (2.19),

$$v_i = \bar{h}_{ij}p_j, \quad i,j = 1,2, \ldots ,n \tag{2.19}$$

In light of equations (3.4), (2.19) reduces to

$$v_l = 0; \quad l = 1,2, \ldots ,i_t - 1 \tag{3.5a}$$

$$v_l = \bar{h}_{lk}p_k; \quad l,k = i_t,\ i_t + 1, \ldots ,n \tag{3.5b}$$

Equation (3.5b) can be solved for p_k, $k = i_t, \ldots ,n$ by Gaussian elimination. Then by equations (2.17) and (2.19)

$$\sigma_\infty = \bar{l}_j p_j = \bar{l}_k p_k, \quad k = i_t, \ldots ,n \tag{3.6}$$

$$v_i = \bar{h}_{ik}p_k, \quad i = 1, \ldots ,i_t - 1 \text{ and } k = i_t, \ldots ,n \tag{3.7}$$

This completes the presentation of the two alternative schemes for computing $p(x)$ from $v(x)$.

4 The Characteristics of the Algorithm and the p-v Relation

We report next on the convergence characteristics of the algorithm and attempt to give some broad criteria on its applicability and effectiveness. Then two models of p-v relations are studied with the aim of determining how the existence of a solution depends on the nonlinearity of the p-v relation. The findings indicate that there are restrictions on the characters of

the p-v relations if realistic craze profiles are to result, i.e., profiles with nonnegative (noninterpenetrating) displacements at all points.

4.1 The Convergence of the Algorithm. Having discussed the algorithm and the convergence criterion in Sections 2.2 and 2.3, we consider the characteristics of the algorithm as applied to specific problems. In Figs. 6, we show how the solution $v(x)$ for a (half-cosine) p-v relation[12] converges for an initial guess of the form

$$v^o(x) = \sigma^o_\infty (c-x)^r$$

where in this example, the initial data are $\sigma^o_\infty = 0.05$ and $r = 1.5$. Nine iterations are required for convergence in this case. (Each iteration takes about 1.2 CPU seconds on a VAX 11-780.) The solution for σ_∞ is 0.0176. Again, a displacement tolerance of 0.1 percent is used as the convergence criterion. For $n = 200$, we find that checking for convergence at only 20 or 100 equi-spaced points give identical results to checking at all 200 points. This is so because the convergence is quite uniform for most p-v relations. Throughout this work, we conservatively use 100 checkpoints for $n = 200$.[13]

With regard to the general applicability of the algorithm, we find that the initial guess does not need to be very close to the correct solution for the scheme to work as illustrated in Fig. 6. It is found that the exponent "r" in equation (2.24b) for the initial guess $v^o(x)$ may range from 0.5 to 2.0, but $r = 1.5$ seems to work best in most instances. When there is no solution, such as those situations to be discussed later in Section 4.2, the algorithm quickly diverges for a wide range of initial guesses.[14] For p-v relations which soften initially and then reharden such as the one studied in Part 2 of this work, the convergence could be very slow and 30–70 iterations may be required depending on the initial guess.

In cases where the algorithm did not converge, the Newton iterative scheme was tried but did not yield convergence either. Per iteration, the Newton scheme takes about 250 times the CPU time needed for Picard's method. For all cases studied, Picard's iterative scheme yielded the same results as the Newton method.

4.2 The Characteristics of Admissible p-v Relations. We examine next how the various characteristics of a given nonlinear p-v relation effect the existence of a solution. This issue is important since little is known regarding the fibril p-v behavior that is consistent with observed craze displacement profiles.[15] As it turns out, one cannot arbitrarily choose p-v relations and expect physically acceptable results; we find interpenetration of the crack craze surfaces and singular stresses unacceptable.

In the present study, we find that there are certain restrictions on the p-v relations for the algorithm to converge. In the unabridged version of this work (Ungsuwarungsri and Knauss, 1985), heuristic arguments, similar to the one presented in connection with the finding (1) below, were used to demonstrate that the proposed algorithm is capable of seeking out a solution for a given p-v relation and craze length, if such a solution exists. Otherwise, our numerical studies show it to be most likely that algorithm divergence indicates the absence of a solution. The question of uniqueness is then addressed, but only a brief summary of the findings is included here and the interested reader is encouraged to consult the unabridged work for more detail.

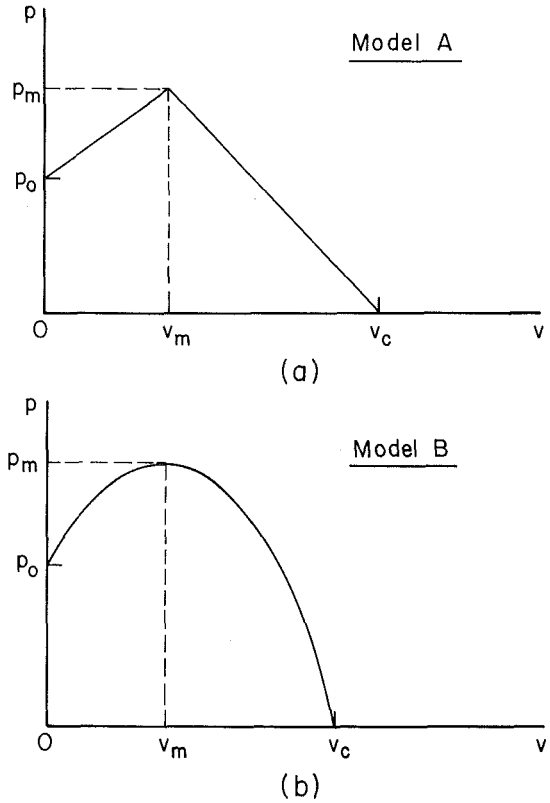


Fig. 7 Two p-v relations used in studying characteristics of "admissible" p-v relations: (*a*) piecewise-linear model; (*b*) cosine model

Consider the two families of p-v relations depicted in Fig. 7.[16] Both models are designed so as to allow the relative amount of hardening and softening in the p-v behavior to be adjusted by varying a few parameters. We summarize the results from a more extensive study documented elsewhere (Ungsuwarungsri and Knauss, 1985) to draw the following conclusions:

(1) *p-v Relations With Hardening and Softening.* It is found that only a limited amount of hardening is permissible for a solution to exist. To demonstrate this finding, consider the following heuristic argument in connection with Figs. 8(*a*)–8(*e*). Assume first that a reasonable craze profile as shown in Fig. 8(*a*) is the result of prescribing a p-v relation depicted in Fig. 8(*b*). Then, without considering continuum mechanics, one deduces from 8(*a*) and 8(*b*) the cohesive force distribution in Fig. 8(*c*). However, on mechanical grounds, i.e., using equation (2.19), this $p(x)$ gives rise to a displacement profile $\tilde{v}(x)$ as shown in Fig. 8(*d*) which is not equal to that in Fig. 8(*a*) and exhibits interpenetration in the craze tip region. Moreover, the p-$\tilde{v}$ relation deduced then from Figs. 8(*c*) and 8(*d*) and shown in Fig. 8(*e*) bears little resemblance to the initial p-v relation of Fig. 8(*b*). Thus, through this somewhat roundabout argument, we conclude that when the prescribed p-v relation hardens by an "appreciable" amount prior to softening (as in Fig. 8(*b*)) and our solution algorithm diverges (i.e., cannot find a solution) for a wide range of initial guesses, there is no solution for the p-v relation prescribed. This line of argument is helpful in ascribing the

[12] Specifically, the p-v relation in Fig. 7(*b*) with $p_o = p_m = 0.0393$, $v_m = 0.0$, and $v_c = 0.008$ is used.

[13] We tried $n = 50$, 100, 200, 300, 400, 500 and found $n = 200$ to be the best compromise between the integration accuracy (within 0.0001 percent and the computing time.

[14] For detail, see Section 3.2 of Ungsuwarungsri and Knauss (1985).

[15] E. Smith (1974) and Knauss (1974, 1976) presented some admissible p-v relations for the crack tip cohesive zone.

[16] Note that in both models, the fibril stress vanishes for displacements greater than a critical value v_c; we thus deal with a critical crack tip opening displacement (CTOD) criterion which can be justified on physical grounds as is discussed at length in Part II of this work (Ungsuwarungsri and Knauss, 1986).

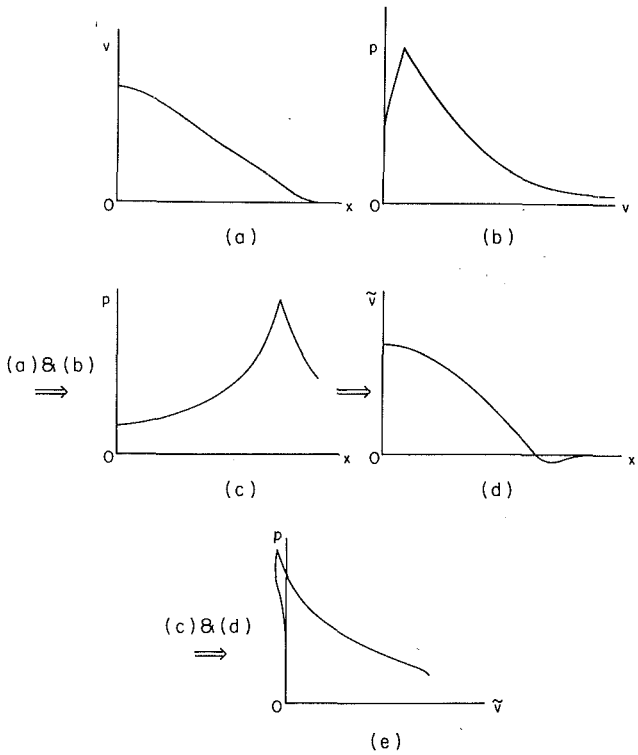


Fig. 8 Demonstration of the nonexistence of solution for prescribed p-v relations such as the one in Fig. (*b*). See discussion in Section 4.2.

nonexistence of a solution when our algorithm fails to produce one.

(2) *Monotonically Softening p-v Relations* ($v_m = 0$). We observe that when v_c exceeds certain values, only the trivial solution[17] results, otherwise a nontrivial solution exists. A heuristic argument similar to the one presented in finding (1) above is given in the unabridged work (Ungsuwarungsri and Knauss, 1985).

We do not address the question of the uniqueness of solutions rigorously. As is the case in most nonlinear problems, uniqueness is difficult to establish. Nevertheless, based on the extensive experience acquired during the course of this study, we are confident that for a given craze length and σ_∞, the solution found by using Picard's algorithm is unique. However, we find that for a given craze length, there may be two different values of σ_∞ and two corresponding sets of $p(x)$ and $v(x)$ which form the following solutions: The solution with the higher σ_∞ corresponds to a full craze or a craze with a very short internal crack, while the other one with the lower σ_∞ is invariably the solution of a craze with a "longer" central crack. These aspects are discussed in more detail in the second part of this study (Ungsuwarungsri and Knauss, 1986).

5 Concluding Remarks

We have shown that the problem of a craze (with or without an internal crack) located in an infinite plane under craze-normal loading can be solved exactly within numerical limitations. The formulation and solution scheme are simple and straightforward. The implemented Picard's algorithm is effective and much more economical than Newton's method. The formulation can also be used, with definite advantage over the schemes which are presently available, to determine stress distribution profiles and the p-v relations from experimental displacement profiles. Certain characteristics of the fibril force-displacement (p-v) relation which are incompatible with observed craze profiles are pointed out. The methodology developed here is applied in Part II of this study (Ungsuwarungri and Knauss, 1986) in simulating craze and crack growth, yielding an additional set of informative findings.

[17]The solution corresponding to a fully closed craze, i.e., $v(x) = 0$ at all points or, simply, no craze exists. It follows from equation (2.8) that since $\int_0^c h(c,x\tau)d\tau = 0$, $v(x) \equiv 0$ is a possible solution.

Acknowledgments

This work was performed as part of an initiation study into nonlinear crack tip mechanics; it was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-84-0254 with the technical contact being Capt. David Glasgow. Furthermore, support from E. I. Dupont de Nemours & Company and NASA under grant NAG-1-474 is gratefully acknowledged. In the past, Drs. K. Palaniswamy and K. Ravi-Chandar have contributed towards the formulation of the problem. We thank Professor E. J. Kramer for the helpful discussions regarding recent experimental findings and for providing us with the micrographs in Fig. 1. The first author is also grateful to his colleagues P. Washabaugh and S. Krishnaswamy for their helpful comments and discussions.

References

Andersson, H., and Bergkvist, H., 1970, "Analysis of a Nonlinear Crack Model," *Journal of the Mechanics and Physics of Solids*, Vol. 18, pp. 1–28.

Argon, A. S., Cohen, R. E., Gebizlioglu, O. S., and Schwier, C. E., 1983, "Crazing in Block Copolymers and Blends," *Advances in Polymer Science*, Vol. 53-53, H. H. Kausch, ed., Springer-Verlag Berlin, Heidelberg, pp. 275–334.

Argon, A. S., and Salama, M. M. 1977, "Growth of Crazes in Glassy Polymers," *Philosophical Magazine*, Vol. 36, pp. 1217–1234.

Atkinson, C., 1970, "An Iterative Scheme for Solving Problems Relating to Cracks Opening Under a Displacement Dependent Internal Stress," *International Journal of Fracture Mechanics*, Vol. 6, pp. 193–198.

Barenblatt, G. I., 1962, "The Mathematical Theory of Equilibrium Cracks in Brittle Fracture," in *Advances in Applied Mechanics*, Vol. VII, Academic Press, New York, pp. 59–129.

Bevan, L., 1982, "Finite Element and Boundary-Element Analysis of Craze Micromechanics," *Journal of Applied Polymer Science*, Vol. 27, pp. 4263–4272.

Bucknall, C. B., 1977, *Toughened Plastic*, Applied Sciences Publication, London, pp. 136–181.

Donald, A. M., and Kramer, E. J., 1982, "The Entanglement Network and Craze Micromechanics in Glassy Polymers," *Journal of Polymer Science*, Polymer Physics Edition, Vol. 20, pp. 1129–1141.

Dugdale, D. S., 1960, "Yielding of Steel Sheets Containing Slits," *Journal of the Mechanics and Physics of Solids*, Vol. 8, pp. 100–104.

Goodier, J. N., and Field, F. A., 1963, "Plastic Energy Dissipation in Crack Propagation," *Fracture of Solids*, D. C. Drucker and J. J. Gilman, ed., Wiley (Interscience), New York, pp. 103–118.

Goodier, J. N., and Kanninen, M., 1966, "Crack Propagation in a Continuum Model with Nonlinear Atomic Separation Laws," Technical Report No. 165, Div. of Eng. Mech., Stanford University.

Huang, N. C., 1985, "On the Size of the Cohesive Zone at the Crack Tip," ASME Journal of Applied Mechanics, Vol. 52, pp. 490–492.

Hull, D., 1973, "The Microstructure and Properties of Crazes," *Deformation and Fracture of High Polymers*, Kausch, Hassell, and Jaffe, eds., Plenum Press, pp. 171–189.

Kambour, R. P., and Kopp, R. W., 1969, "Cyclic Stress-Strain Behavior of the Dry Polycarbonate Craze," *Journal of Polymer Science, A-2*, Vol. 7, pp. 183–200.

Knauss, W. G., 1974, "Crack Propagation in Viscoelastic Diaspastically Non-Simple Solids: A Progess Report," *Int. Union of Theoretical and Appl. Mech., Mech. of Viscoelastic Media and Bodies Symposium*, Gothenburg, Sweden, J. Hult, ed., pp. 263–278. A more extensive version can be found in *Strength and Structure of Solid Materials*, a Joint Japan-USA Seminar, 1976, H. Miyamoto et al., eds., Noordhoff International Publishing, Leyden, pp. 389–416.

Knight, A. C., 1965, "Stress Crazing of Transparent Plastics: Computed Stresses at a Nonvoid Craze Mark," *Journal of Polymer Science*, Part A, Vol. 3, pp. 1845–1857.

Kramer, E. J., 1983, *Advances in Polymer Science*, Vol. 52-53, H. H. Kausch, ed., Springer-Verlag Berlin, Heidelberg, Chapter 1.

Kramer, E. J., and Hart, E. W., 1984, "Theory of Slow, Steady State Crack Growth in Polymer Glasses," *Polymer*, Vol. 25, pp. 1667–1678.

Lauterwasser, B. D., and Kramer, E. J., 1979, "Microscopic Mechanisms of Craze Growth and Fracture," *Philosophical Magazine A*, Vol. 39, pp. 469–495.

Rice, J. R., 1966, "Plastic Yielding at Crack Tip," *Proceedings of the 1st International Conference on Fracture*, Sendai, Japan, T. Yokobori et al., eds., Vol. 1, Japanese Society for Strength and Fracture of Materials, Tokyo, pp. 283–308.

Smith, E., 1974, "The Structure in the Vicinity of a Crack Tip: A General Theory Based on the Cohesive Zone Model," *Engineering Fracture Mechanics*, Vol. 6, pp. 213–222.

Sneddon, I. N., 1951, *Fourier Transforms*, McGraw-Hill, New York, p. 426.

Spurr, O. K., and Niegisch, W. D., 1962, "Stress Crazing of Some Amorphous Thermoplastics," *Journal of Applied Polymer Science*, Vol. 6, pp. 585–599.

Sternstein, S. S., Ongchin, L., and Silverman, A., 1968, "Inhomogeneous Deformation and Yielding of Glasslike High Polymers," *Applied Polymer Symposia*, No. 7, pp. 175–199.

Sun, B. N., and Hsiao, C. C., 1985, "Nonlinear Quasi-fracture Behavior of Polymers," *Journal of Applied Physics*, Vol. 57, No. 2, pp. 170–174.

Tada, H., Paris, P. C., and Irwin, G. R., 1973, *The Stress Analysis of Cracks Handbook*, Del Research Corporation, Hellertown, PA, p. 5.10.

Ungsuwarungsri, T., and Knauss, W. G., 1985, "A Nonlinear Analysis of an Equilibrium Craze in an Infinite Medium Subjected to Symmetrical Loading," GALCIT Report 85-15, California Institute of Technology, Pasadena, CA.

Ungsuwarungsri, T., and Knauss, W. G., 1988, "A Nonlinear Analysis of an Equilibrium Craze, Part II: Simulations of Craze and Crack Growth," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, No. 1, pp. 52–58.

Verheulpen-Heymans, N., and Bauwens, J. C., 1976, "Effect of Stress and Temperature on Dry Craze Growth Kinetics During Low-Stress Creep of Polycarbonate," Part 1 and 2, *Journal of Material Science*, Vol. 11, pp. 1–6 and 7–16.

Walton, J. R., and Weitsman, Y., 1984, "Deformations and Stress Intensities due to a Craze in an Extended Elastic Material," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, No. 1, pp. 84–92.

Wang, W. V., and Kramer, E. J., 1982, "A Distributed Dislocation Stress Analysis for Crazes and Plastic Zones at Crack Tips," *Journal of Material Science*, Vol. 17, pp. 2013–2026.

Weitsman, Y., 1986, "Nonlinear Analysis of Crazes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 97–102.

A Nonlinear Analysis of an Equilibrium Craze: Part II—Simulations of Craze and Crack Growth

T. Ungsuwarungsri
Graduate Research Assistant.

W. G. Knauss
Professor of Aeronautics and Applied Mechanics.

Graduate Aeronautical Laboratories, California Institute of Technology, Pasadena, CA 91125

In this study we investigate the effects of nonlinear fibril behavior on the mechanics of craze and crack growth. The effect of strain-softening cohesive material on crack stability is of particular interest and is examined via a craze and crack model developed in the first part of this work where the formulation and solution of the problem are discussed.[1] In this second part, quasi-static growth of a craze with a central crack is analyzed for different nonlinear force-displacement (p-v) relations for the craze fibrils. A "critical crack tip opening displacement" (CTOD), or more precisely, "critical fibril extension" is employed as the criterion for fracture. The p-v relation is further assumed to be invariant with respect to the craze and crack lengths. The results are compared with the Dugdale model; the craze zone size and the energy dissipation rate approach asymptotic values in the limit of long cracks. The problem of craze growth from a precut crack under increasing far-field loading is then studied. In the case where the p-v relation is monotonically softening, the crack can start to grow in an unstable manner before the crack tip opening displacement reaches its critical value.

1 Introduction

In Part I of this study (Ungsuwarungsri and Knauss, 1986), the mathematical and experimental treatment of craze formation were summarized. We developed there an analytical and computationally efficient scheme to obtain the stress distribution and displacement profiles of a craze/crack for a prescribed craze length and an arbitrarily nonlinear force-displacement (p-v) behavior of the craze fibrils. The solution scheme developed in the first part is applied here to investigate how the nonlinear fibril p-v characteristics affect the mechanics of craze and crack growth. The critical crack tip opening displacement (CTOD) is employed in a natural way as a fracture criterion. The applicability of this CTOD criterion to crazes containing cracks is based on experimental evidence and will be discussed in the next section. It is assumed that under quasi-static conditions the fibril p-v relation does not change its character as the craze and crack grow. This assumption is made for lack of experimental data and because of inadequate understanding of the fibril behavior at the present time. In a later section, we examine the problem of a craze initiating from an existing (precut) crack with special interest on the effect of "softening" fibril behavior on crack stability.

2 The Two Basic Assumptions

We examine first the "crack tip opening displacement" criterion (CTOD) and the "invariance" assumption regarding the fibril p-v behavior as the craze and crack grow in the light of existing experimental results.

2.1 The CTOD Criterion. When a central crack is present within a craze, the crack may or may not propagate as the far-field load σ_∞ is increased. If the crack does not propagate, it is clear that the fibril at the crack tip can sustain additional growth (lengthening) though not necessily higher stress or strain. This is the consequence of two physical mechanisms that operate simultaneously, namely, surface drawing and fibril creep.[2] If surface drawing alone operates, fibrils will never break and the crack will never advance. This possibility is contrary to experimental observations since cracks do grow within crazes. As can be explained in terms of craze and fibril microstructures, a crack that propagates quasi-statically through a craze either runs right along the midrib where the fibrils are weakest or propagates in a "patch" or "mackeral"

[1] Here all quantities are presented in their nondimensionalized forms as defined in the first part of this study (Ungsuwarungsri and Knauss, 1986). For additional detail, see the unabridged version of this work (Ungsuwarungsri and Knauss, 1985).

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 21, 1986; final revision August 25, 1987.

[2] Surface drawing is the process by which bulk polymer is converted (drawn) into fibrillar craze material. Fibril creep is simply the stretching of existing fibrils without drawing additional polymer material into the craze from the bulk.

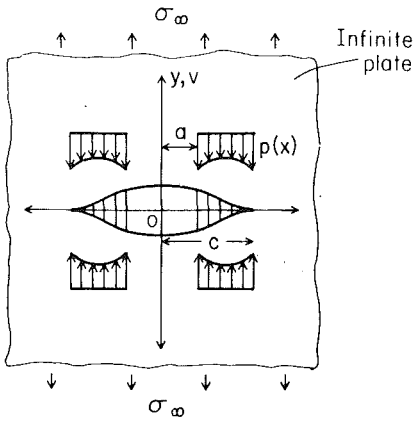


Fig. 1 Problem definitions

pattern (Hull, 1973; Lauterwasser and Kramer, 1979). In the first situation, the heightened stress at the crack tip causes the previously unloaded midrib to stretch until the local extension ratio exceeds the maximum sustainable by the bundles of polymeric molecule chains that constitute the fibril, thus leading to fracture. In the latter event, usually observed at higher crack velocities (Hull, 1973), the high crack tip stress causes fibrils at the craze-polymer interface to be drawn at high local strains (Lauterwasser and Kramer, 1979). One observes then that the crack, in this case, prefers to run along either interface where the fibrils are locally weaker than at the midrib, creating the mackeral pattern. In either case, the above situations support the use of a critical crack tip opening displacement (or a critical fibril extension criterion). This is also confirmed by experiments performed by Döll et al. (1979) and Schinker and Döll (1979) who found that the CTOD is relatively constant for a wide range of crack velocities. Specifically, these authors showed that the maximum craze widths at the crack tip (critical CTOD or v_c in our notation), measured in PMMA compact tension specimens for crack speeds ranging from 10^{-8} to 20 mm/s, are essentially constant with a standard deviation of about 8 percent.[3] Their findings suggest that the fibrils spanning the tip of the crack can only sustain a limited, relative, displacement of their ends. An interpretation of the experimental result in terms of molecular behavior is given by Döll et al. (1979) but such extensive discussions are beyond the scope of this presentation. For our purpose, it suffices to assume that the surface drawing mechanism is somehow constrained at the crack tip and creep invariably gives rise to fibril breakage at a roughly constant CTOD.[4] We assume in this study, therefore, that for a given polymer a critical CTOD exists and is constant.

2.2 Comments on the Invariance of the p-v Relation. To study craze and/or crack growth, one needs to know whether the p-v relation changes as propagation proceeds and if so, how. The question can only be answered by experiments. Since no experimental data on this particular topic is available, we feel free to assume further that the p-v relation is "invariant" with respect to both the craze and the crack lengths for quasi-static propagation under a "controlled" environment.[5] As it turns out, on the basis of the subsequent studies, this assumption can only be valid over a limited range of craze and crack lengths (see Section 3).

[3]In private communication, Professor Kramer indicated that under a low crack speed of about 1 micron/s the critical CTOD increases substantially. This casts some doubt on the validity of the number 10^{-8} mm/s given by Döll et al. (1979) and Shinker and Döll (1979). However, for the speed range between 10^{-3} and 20 mm/s, the constancy of the CTOD (for PMMA) appears acceptable. See the recent work of Kramer and Hart (1984) for further discussion on this point.

[4]It is worthwhile to note that cracks are observed to form in "full" crazes at the polymer/craze interface and not at the midrib. The locations of the first-formed crack in a full craze are also not necessarily at the center part of the craze where the fibrils are longest.

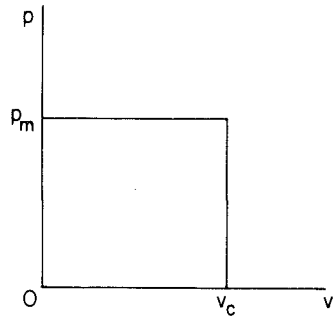


Fig. 2 The Barenblatt-Dugdale model

3 Numerical Simulations of Craze and Crack Propagation

We proceed now to apply the solution scheme developed in Part I of this work (Ungsuwarungsri and Knauss, 1986) to study how the nonlinear fibril (cohesive) p-v characteristics affect the mechanics of craze and crack growth. The geometry of the problem[6] is depicted in Fig. 1. The craze length is denoted by c, the crack length by a, and the vertical boundary displacement of the craze by v. The remotely applied stress is σ_∞ and the fibril cohesive force is represented by p[7]. Given a craze length and a p-v relation for the fibrils, the problem is to find $p(x), v(x)$, and σ_∞ such that both equilibrium and the smooth closure (Barenblatt's) condition are satisfied simultaneously.[8] The solution scheme has been presented in Part I.

Quasi-static craze and crack growth is simulated by varying the craze length c continuously. Based on the discussion in Section 2, we assume the fibril force-displacement relation of the form

$$p(x) = p[v(x)] \tag{3.1}$$

The effect of several p-v relations will be examined. We start by first reviewing the Barenblatt-Dugdale model (Dugdale, 1960; Barenblatt, 1962) shown in Fig. 2.

3.1 The Barenblatt-Dugdale Model. This simplest model assumes constant cohesive (fibril) stress throughout the cohesive (craze) zone as denoted by p_m in Fig. 2. The problem has been analyzed by Rice (1966) and Goodier and Field (1963); we summarize below the pertinent results for reference.[9] They are obtained directly from the equations presented in Section 2 of Ungsuwarungsri and Knauss (1986). The Barenblatt condition requires

$$\sigma_\infty = \frac{2}{\pi} p_m \cos^{-1}\left(\frac{a}{c}\right) \tag{3.2}$$

and the displacement $v(x)$ is

[5]Temperature, pressure, moisture content, chemical environment (such as the presence of alcohol or any chemical agents possibly affecting fibril drawing and creep) are all considered to remain constant.

[6]We consider here the planar problem, i.e., plane stress or plane strain.

[7]The mathematical modelling of the problem has been described in the first part of this study.

[8]In this paper, the (net) stress intensity factor is always zero and thus does not enter the discussion as a pertinent parameter. The external loading is completely characterized by σ_∞ which is related to the "loading" stress intensity factor by $k_I = \sigma_\infty \sqrt{\pi c}$. Note that the stress states at both the crack tip and the craze tip are determined by the nonlinearity of the fibril p-v relation rather than the "usual" K field.

[9]Following Part I, all equations and parameters are in dimensionless forms.

$$v(x)=\frac{p_m}{\pi}\left[a\log\left|\frac{\sqrt{c^2-x^2}+\sqrt{c^2-a^2}}{\sqrt{c^2-x^2}-\sqrt{c^2-a^2}}\right|\right.$$

$$\left.+x\log\left|\frac{x\sqrt{c^2-a^2}-a\sqrt{c^2-x^2}}{x\sqrt{c^2-a^2}+a\sqrt{c^2-x^2}}\right|\right] \tag{3.3}$$

At $x=a$, v equals v_c the critical CTOD; thus applying L'Hospital's rule, one obtains

$$v(a)=v_c=\frac{2p_m}{\pi}a\log\left(\frac{c}{a}\right) \tag{3.4}$$

Let ω denote the size of the craze (yield) zone, i.e.,

$$\omega\equiv c-a \tag{3.5}$$

Furthermore, let γ be the "fibril fracture energy" (or the "surface energy") which is the work expended in pulling a fibril from $v=0$ to $v=v_c$. For the Barenblatt-Dugdale model, we have simply

$$\gamma\equiv p_m v_c \tag{3.6}$$

By analyzing the data for six polymers in the study by Donald and Kramer (1982) we find that the dimensionless parameters p_m and v_c have values in the ranges of

$$0.01\le p_m\le 0.1$$

$$0.001\le v_c\le 0.01$$

We shall use these numbers later as guides for bounding the various parameters in the p-v relations to be studied so that the results of our analyses will be quantitatively meaningful.

We examine next the energy release rate of the system. Let U represent the elastic energy of the system. The energy release rate G is defined in terms of the potential energy Π as

$$G=-\frac{\partial\Pi}{\partial a} \tag{3.7}$$

Consider the fixed-grip case so that no external work is done during the course of crack propagation. Then equation (3.7) reduces to

$$G=-\frac{\partial U}{\partial a} \tag{3.8}$$

Now U is given by the elastic energy stored initially in the plate (planar body), U_p, minus the energy U_c which is required to form a crack of length $2a$ and two craze zones of length ω each. We have then

$$U=U_p-U_c \tag{3.9}$$

$$U_p=\frac{1}{4}\sigma_\infty^2 d^2 \tag{3.10}$$

where we let d denote the dimension of this "infinite" square plate (i.e., $d>>c$).[10] U_c is given by

$$U_c=4\left[\gamma a+\int_a^c p_m v(x)dx\right] \tag{3.11}$$

where the factor of 4 accounts for the four quadrants. Substituting $v(x)$ from equation (3.3) and making use of equations (3.2) and (3.4), one obtains after some algebra

$$U_c=4a\left[\gamma+\frac{p_m}{2}\{\sigma_\infty\sqrt{c^2-a^2}-2v_c\}\right] \tag{3.12}$$

Using equations (3.9) and (3.10) in (3.8) yields

$$G=-\frac{1}{2}\sigma_\infty d^2\left(\frac{\partial\sigma_\infty}{\partial a}\right)+\frac{\partial U_c}{\partial a} \tag{3.13}$$

At this point an important question enters: What is the criterion for crack propagation (fracture)? Goodier and Field (1963) assumed that the crack grows in a self-similar manner, such that

$$\frac{\partial c}{\partial a}=\frac{c}{a} \tag{3.14}$$

This condition is the consequence of assuming σ_∞ to be constant as the crack and craze grows. Under these conditions, $\partial\sigma_\infty/\partial a$ vanishes and G becomes, in view of equation (3.13), $\partial U_c/\partial a$, which is the "plastic" (cohesive) work rate. However, the crack tip opening displacement $v(a)=v_c$ as given by equation (3.4) must then be a linear function of a. Thus the CTOD in the model presented by Goodier and Field model increases as the crack grows. This behavior is inconsistent with experimental observations as discussed in Section 2.

In the present model, the critical CTOD, v_c, is kept constant; this condition in turn requires that σ_∞ must decrease as a increases (see Fig. 7) in order to maintain the quasi-static condition

$$\left.\frac{\partial\sigma_\infty}{\partial a}\right|_{v_c}<0 \tag{3.15}$$

Viewing equation (3.13) in this light, it is evident that G becomes unbounded if d is unbounded, no matter how small $\sigma_\infty(\partial\sigma_\infty/\partial a)$ may be, since $\partial U_c/\partial a$ is always finite.

Therefore, a crack, the tip displacement of which has reached the critical value, is always unstable in this geometry.[11] To this result we only need to add the reminder that, for a clamped infinite (very large) plate, a small enlargement of an internal crack would hardly reduce the far-field σ_∞ sufficiently to keep the crack and craze growing in a quasi-static manner. Therefore, once a crack (with or without a cohesive zone) starts growing, and the fracture criterion such as a critical CTOD is satisfied, catastrophic failure ensues. We shall see later that the converse is not true.

The plastic work rate, $\partial U_c/\partial a$, calculated using the CTOD criterion, differs markedly from the result obtained by Goodier and Field under the assumption of self-similar growth (see Ungsuwarungsri and Knauss, 1985, and Goodier and Field, 1963, for details). Based on the present criterion, the plastic work rate increases as a increases and approaches the constant value 4γ rapidly. Physically, for a short crack, one has a relatively large craze zone ω; to advance the crack by an amount Δa, very little energy is needed since c increases more slowly than a (i.e., the craze zone ω shrinks). For longer cracks, the size of the craze zone tends toward a constant value and thus the craze zone simply translates as the crack propagates. In this latter case, the dissipated energy is thus 4γ per unit crack advance. This is demonstrated later in Fig. 8 together with results for other nonlinear p-v models.

3.2 General p-v Models. Next we examine six different p-v relations labeled #1 through #6 in Fig. 3 and compare the results to the Dugdale model labeled D. Model #6 approximates the p-v behavior of craze fibrils as it is known today (Wang and Kramer, 1982; Donald and Kramer, 1982). As a check, we employ this model to simulate craze and crack growth. The resulting craze contours and stress distribution profiles are depicted in Figs. 4 and 5, respectively. They do closely resemble the experimental $v(x)$ and $p(x)$ obtained by Wang and Kramer (1982), and Donald and Kramer (1982).

To compare the results for all these different p-v relations, it is desirable to use the Dugdale model as a benchmark (here we use the model labeled D in Fig. 3 with $p_m=0.025$ and $v_c=0.008$). All the p-v models are designed so that the area

[10] Note that the factor 1/4 in place of the normally expected factor of 1/2 is a result of the particular nondimensionalization scheme employed.

[11] If the tip displacement is less than v_c, by definition the crack does not grow as the tip fibril has not yet been broken. For softening p-v relations, however, instability may occur before V_c is reached (see Section 4).

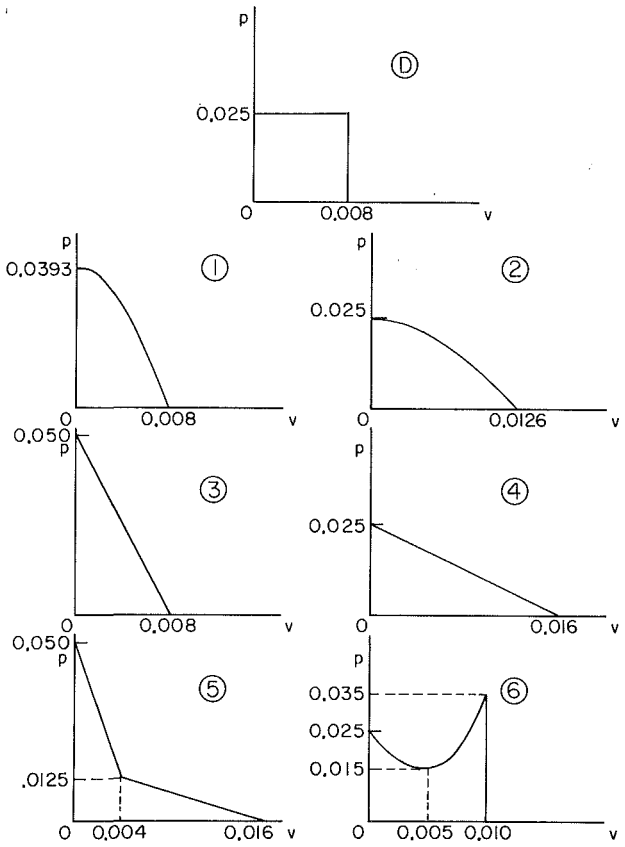


Fig. 3 p-v models used in the nonlinear craze and crack propagation studies in Sections 3 and 4. All models have the same γ.

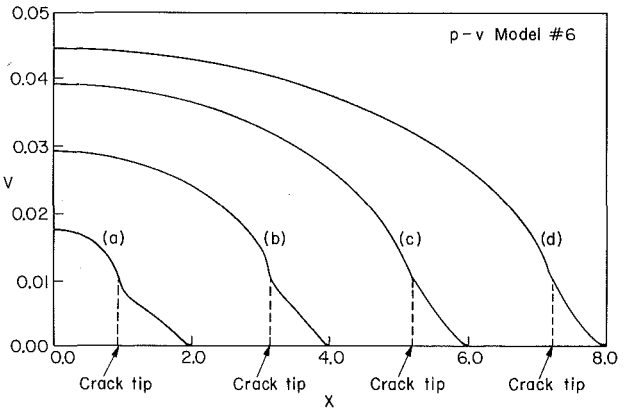


Fig. 4 Craze and crack propagation simulations for p-v model #6: displacement profiles

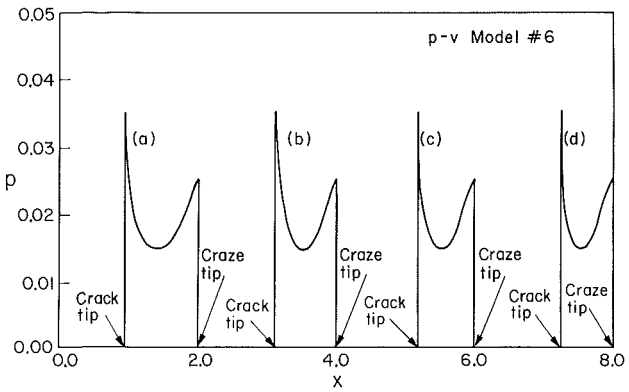


Fig. 5 Craze and crack propagation simulations for p-v model #6: stress distribution profiles

under each p-v curve, γ, is constant and equal 2×10^{-4}, the same as that for the Dugdale model employed.

The numerical procedures have been presented in Section 2.3 of Part 1 (Ungsuwarungsri and Knauss, 1986). The craze length is varied from $c=0.0$ to $c=10.0$, the increment Δc ranges from 0.01 to 0.02. Smaller increments for Δc are necessary to capture the behavior of the system for short crack lengths, particularly in calculating the rate of plastic work, $\partial U_c/\partial a$.[12] During the course of these craze growth simulations, it was found that convergence is generally attained in fewer iterations if $v_i(c)$ is used as the initial guess of $v_i(c+\Delta c)$. (This is not the case for p-v model #6, however.) In order not to confuse the presentation of the results, the trivial solution which occurs at short craze lengths and corresponds to a closed craze $(v(x)=0)$ is not included. Also, for the Barenblatt-Dugdale model D, results that are "nonphysical" in the sense of a crack growing in a shrinking craze, are ommitted. Interested readers are referred to the unabridged version of this study (Ungsuwarungsri and Knauss, 1985) for more details.

In Fig. 6, the craze zone ω is plotted as a function of the crack length a. The craze zone is seen to decrease monotonically and to tend toward constant asymptotic values as a increases for all p-v models.[13] Thus in the limit of long cracks, the craze zone is simply shifted as the crack and craze grow. Note that ω is primarily dependent on v_c but also depends weakly on the character of the p-v relation.

In Fig. 7, the far-field stress σ_∞ normalized by $p_m=0.025$ is

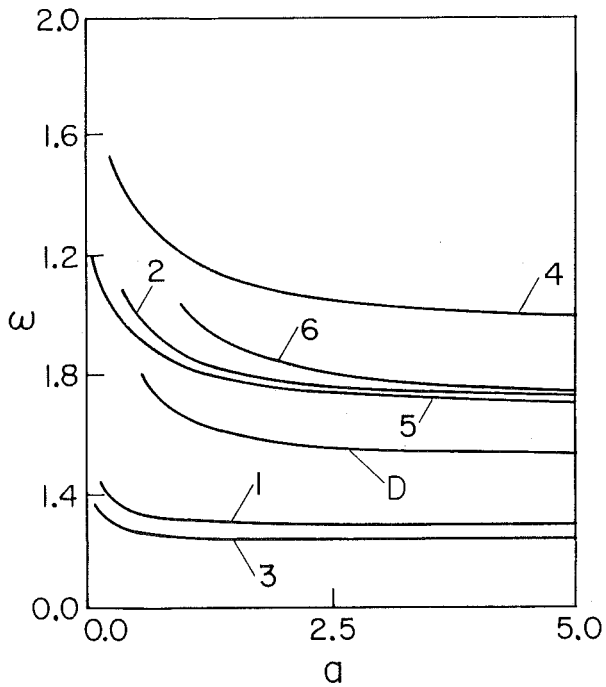


Fig. 6 The craze zone ω as a function of a for various nonlinear p-v models (p-v models as indicated)

plotted versus c for the various p-v models. The remotely applied stress decreases monotonically as the craze/crack grows, which is necessary to maintain quasi-static growth. This confirms the conclusion reached in the previous section from the energy release rate analysis that once the critical CTOD is

[12] U_c is obtained by integrating numerically

$$U_c=4\left[\gamma a+\int_a^c ev(v(x))dx\right] \text{ where } ev(v)=\int_o^v p(v')dv'$$

[13] For the Barenblatt-Dugdale model, this can be analytically shown to be the case.

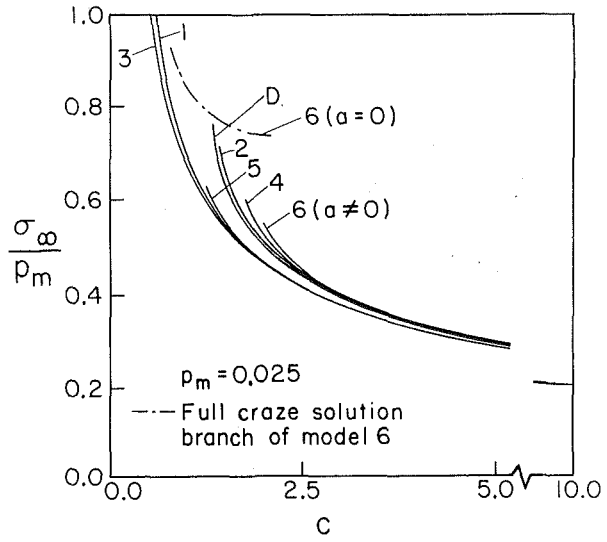


Fig. 7 σ_∞/p_m **as a function of c for various nonlinear p-v models (p-v models as indicated)**

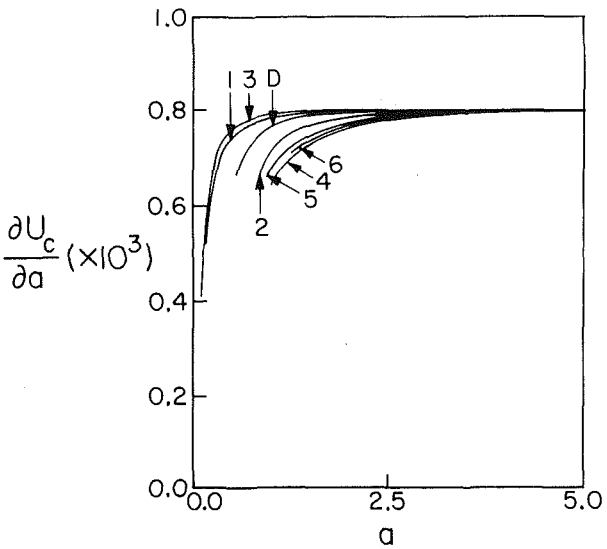


Fig. 8 The rate of plastic work dissipation $\partial U_c/\partial a$ **as a function of** a **for various nonlinear p-v models (p-v models as indicated)**

reached the crack will grow unstably if the remotely applied stress is maintained constant. An interesting feature is exhibited by the dash-dot curve which represents a "full craze" ($a=0$) solution branch and is present only in the p-v model #6 (among all models studied here). This unusual behavior merits further elaboration.

One finds that for a craze length $c<0.8$, the algorithm converges to the trivial solution in which the craze is closed ($v(x)=0$) at all points (see Fig. 13). For c greater than 0.8, there exists a "full craze" solution regime up to $c=2.2$. At $c=2.2$, the solution shows that $v(0)$ is approximately 0.006 and thus $p(x)$ has a minimum at $v(x)=0.005$ (see Fig. 3). For c larger than 2.2, there is no full craze solution (the algorithm "jumps" and converges to solutions with cracks, i.e., $a\neq 0$). This behavior can again be explained using a heuristic argument, see Section 3.2 of the unabridged work (Ungsuwarungsri and Knauss, 1985). Note that for $1.87<c<2.2$, the solutions are not unique, i.e., for a given c in this range, we either have a full craze solution at a higher σ_∞ or a crack/craze solution at a lower σ_∞. Whether this relatively narrow range of nonuniqueness is the result of numerics or is real has not been explored further.

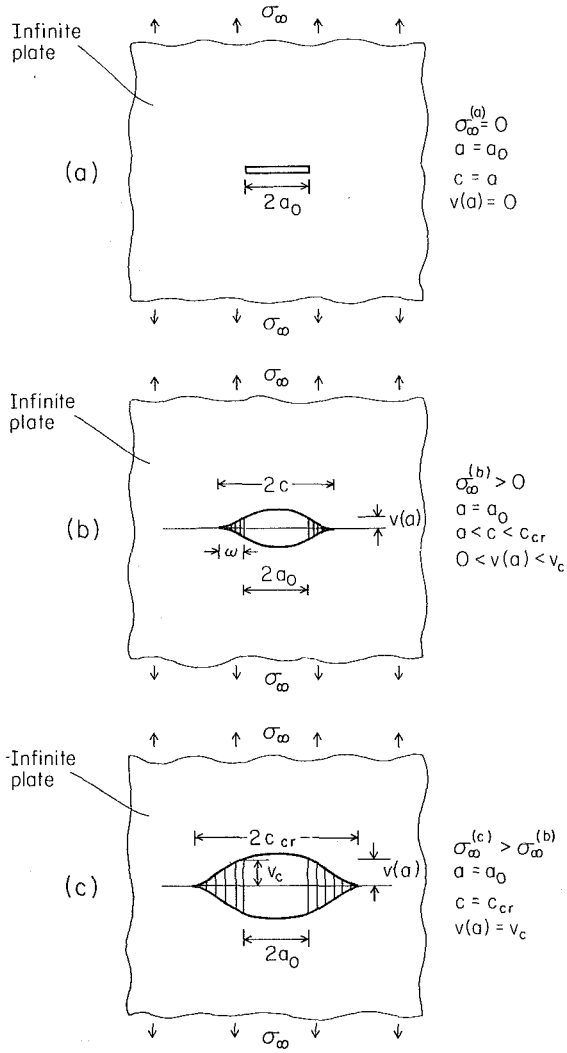


Fig. 9 Craze growth initiating from a precut crack: (a) initial precut crack, no far-field loading; (b) increasing far-field loading, no crack growth; (c) critical CTOD reached, crack growth imminent

The above result, which indicates the absence of solutions for the prescribed p-v relation (#6) over some ranges of craze and/or crack lengths may seem, at first sight, contradictory to our physical intuition. One would think that crazes and/or cracks of different sizes should exist in a continuous size "spectrum." This seeming contradition is resolved by recognizing that, most probably, the p-v relation is not "invariant" with respect to all values of the crack and craze lengths as initially assumed.[14] More elaborate modelling of the fibril p-v characteristic is therefore needed (see Section 5 for further discussion).

In Fig. 8, the rate of "plastic" work, $\partial U_c/\partial a$, is plotted as a function of the crack length a. It approaches the asymptotic limit 4γ fairly quickly for all p-v models studied.[15]

4 Craze Growth Initiating From a Precut Crack

In this section, we consider the problem of a craze growing from a precut slit as shown in Fig. 9. Under the assumption of a CTOD criterion, one would expect the crack to start advancing when the remotely applied stress σ_∞ is high enough to

[14] Note that this conclusion is based solely on mechanics analysis.

[15] The curve representing p-v model #6 has been smoothed out for $a>2.5$. For this particular p-v relation, to get accurate v_c in the craze zone (whose size is = 0.8 for $a>2.5$), finer discretization is required to capture the steep rise and fall of the stress distribution in the craze zone.

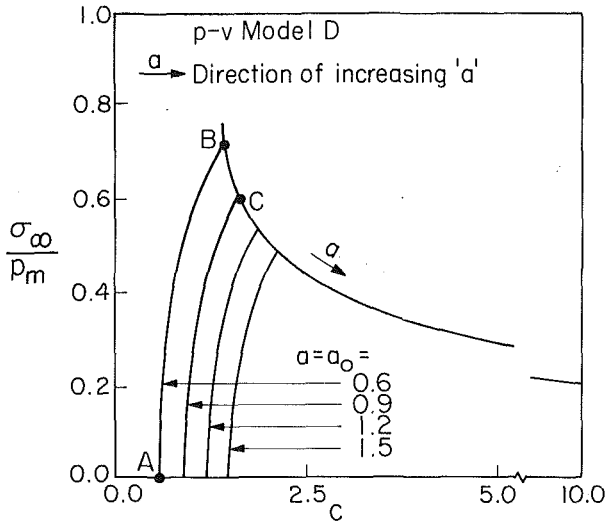


Fig. 10 σ_∞/p_m as a function of c for $a_0 = 0.6$, 0.9, 1.2, and 1.5 for the Barenblatt-Dugdale model *D*

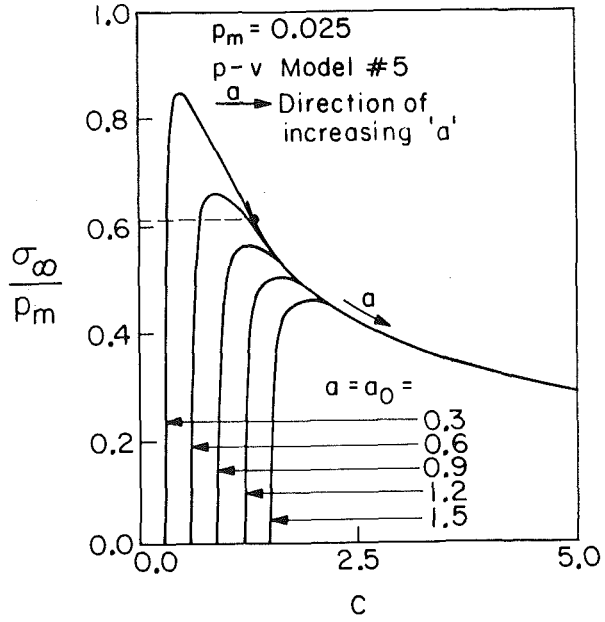


Fig. 11 σ_∞/p_m as a function of c for $a_0 = 0.3$, 0.6, 0.9, 1.2, and 1.5 for p-v model #5

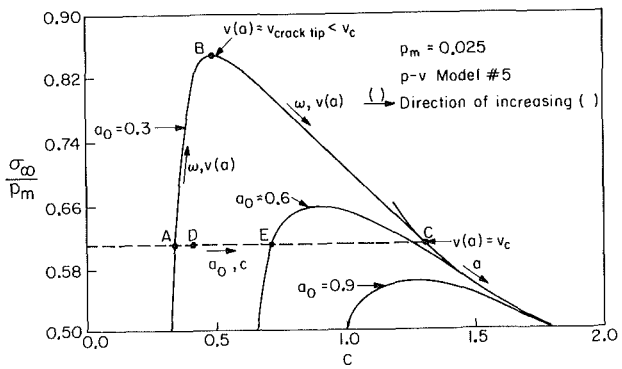


Fig. 12 σ_∞/p_m as a function of c for $a_0 = 0.3$, 0.6, and 0.9 for p-v model #5 (enlargement of a part of Fig. 11)

make the crack tip displacement $v(a)$ reach the value v_c. Thereafter the crack and craze would grow unstably as discussed in the previous section.

However, it is conceivable that for certain p-v relations, σ_∞ may reach its maximum value before the crack tip opening displacement reaches its critical value v_c, thus resulting in instability *before* the CTOD criterion is satisfied. We shall show that this is indeed the case for p-v relations that soften monotonically up to fracture (fibril breakage). To start, let us first study the Barenblatt-Dugdale model for later reference.

4.1 The Barenblatt-Dugdale Model. We start with a precut crack of length $2a_o$. Initially the craze length $c = a_o$, i.e., $\omega = 0$. We then increase σ_∞ gradually until $v(a)$ reaches v_c. By equation (3.4) when $v(a) = v_c$, the "critical" craze length c_{cr} is

$$c_{cr} = a_o \exp\left(\frac{\pi v_c}{2p_m a_o}\right) \tag{4.1}$$

Hence for $a_o \le c < c_{cr}$, the remotely applied stress σ_∞ can be obtained from equation (3.2), i.e.

$$\sigma_\infty = \frac{2p_m}{\pi}\cos^{-1}\left(\frac{a_o}{c}\right), \quad a_o \le c < c_{cr} \tag{4.2}$$

Based on equations (4.1) and (4.2), we can plot the applied stress σ_∞ as a function of c for constant $a = a_o$ as shown in Fig. 10. Four initial crack lengths are considered ($a_o = 0.6$, 0.9, 1.2, and 1.5). The curve representing quasi-static crack and craze growth referred to subsequently as the "master" curve) and on which the CTOD always attains its critical value is also identified. On the "master" curve, the craze and crack growth is unstable in the sense that σ_∞ must be reduced continually in order to maintain quasi-static propagation.

Along the curves on which $a = a_o$, the CTOD $v(a)$ gradually grows as the remotely applied stress σ_∞ is increased, e.g., from point A to B; it then reaches its critical value v_c at B. Thus for the Barenblatt-Dugdale model, σ_∞ increases monotonically in the process and instability can occur only when the "master" curve is reached.

4.2 General p-v Models. Next let us consider the same problem as in Section (4.1) using p-v relations #1 through #6 of Fig. 3. In contrast to the Barenblatt-Dugdale model, we find that the p-v relations that soften monotonically up to the point of fracture (models #1 through #5) instability may occur before the critical CTOD is reached at the crack tip. Briefly, the numerical analysis proceeds in the following sequence:

(*i*) Start with a value a_o which corresponds to a c_{cr} that lies on the master curve.

(*ii*) Increase c gradually from a_o to c_{cr} in such way that the crack tip falls exactly on a discretized point x_i. This avoids unnecessary interpolation.

(*iii*) Find the solution, pre-specifying that $p(x)$ vanish for $x \le a_o$. The solution would automatically have $v(a) < v_c$ as long as $c < c_{cr}$.

For illustrative purposes we present first the result for model #5 as being representative of the class of monotonically "softening" p-v models (#1 through #5). More complete results are given in the unabridged work (Ungsuwarungsri and Knauss, 1985). In Figs. 11 and 12 (enlargement of a part of Fig. 11), one observes that, for $a_o = 0.3$ to 1.5, instabilities occur before the "master" curve is reached.

Consider a stress-controlled experiment. As the far-field stress σ_∞ is increased from zero to point B (Fig. 12), the CTOD, i.e., $v(a)$, approaches the critical CTOD, v_c. At B, any further increase in σ_∞ will cause the crack and craze to grow unstably since quasi-static growth will no longer be maintained at that stress level. If one reduces σ_∞ instead, keeping in mind that the craze length c cannot decrease while $a = a_o$ remains constant, point C could then be reached quasi-statically. Therefore, at point B, any small positive perturbation in σ_∞ or in a_o would cause the crack and craze to grow in an unstable manner. Note that the solution cannot "jump"

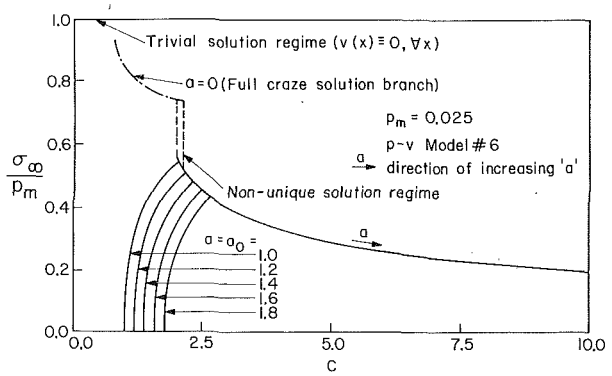


Fig. 13 σ_∞/p_m **as a function of c for** $a_o = 1.0, 1.2, 1.4, 1.6$, **and 1.8 for p-v model #6**

spontaneously from point A to point C because more work has to be done to further develop the craze zone and this is achieved by following the path ABC as discussed.

To explore this instability further, consider perturbing the system at point A by "cutting" fibrils at the crack tip making a_o increase infinitesimally. Does this cause catastrophic failure? As depicted in Fig. 12, the solution will then move from point A to a nearby point D which corresponds to $a_o = 0.3^+$ and stop. Thus the craze and crack at point A is stable since a small perturbation in a_o does not lead to total failure. Note in passing that the unstable behavior is more pronounced for the shorter initial crack lengths.

We turn next to the result for p-v model #6 which rehardens after the initial softening and fails abruptly as in the Barenblatt-Dugdale model (see Fig. 13). One observes that the unstable character experienced in softening p-v models as clearly absent which closely resembles the result for the Barenblatt-Dugdale model shown in Fig. 10. These results indicate that in cases where craze fibrils fail in an abrupt manner (as opposed to gradual softening to rupture), instability does not take place until the CTOD reaches its critical value v_c.

5 Concluding Remarks

In this second part of the study, we have examined the quasi-static growth of cracks and crazes by considering a number of nonlinear fibril p-v models in comparison with the Barenblatt-Dugdale model. The energy dissipation rate consistent with the critical CTOD criterion is used for these materials. We also considered the case of craze growth starting from a precut crack; under these circumstances instabilities are observed to occur before the critical CTOD is reached at the crack tip for p-v models that soften monotonically to fracture.

As discussed earlier, not enough is known about the p-v characteristics of the craze fibrils. Further investigation into the interplay between the surface drawing and creep mechanics should shed light on how the p-v behavior of fibrils changes as the craze and crack grow; a first attempt along this line has been undertaken by Kramer and Hart (1984). Experiments such as those performed by Kramer and his coworkers need to be carried out for a wider range of craze and crack lengths to determine to what extent the fibril p-v relation is invariant with respect to c and a.

Acknowledgments

This work was performed as part of an initiation study into nonlinear crack tip mechanics; it was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-84-0254 with the technical contact being Capt. David Glasgow. Also support from E.I. Dupont de Nemours and Company and NASA under grant NAG-1-474 is gratefully acknowledged. We thank Professor E. J. Kramer for the helpful discussions on the physical aspects of the problem. The first author is also grateful to his colleagues Mr. Washabaugh and Mr. Krishnaswamy for their many helpful comments and discussions.

References

Barenblatt, G. I., 1962, "The Mathematical Theory of Equilibrium Cracks in Brittle Fracture," *Advances in Applied Mechanics*, Vol. VII, Academic Press, New York, pp. 59–129.

Döll, W., Schinker, M. G., and Koenczoel, L., 1979, "A Time Independent Fracture Criterion for PMMA?," *International Journal of Fracture*, Vol. 15, pp. R145–149.

Donald, A. M., and Kramer, E. J., 1982, "The Entanglement Network and Craze Micromechanics in Glassy Polymers," *Journal of Polymer Science*, Polymer Physics Edition, Vol. 20, pp. 1129–1141.

Dugdale, D. S., 1960, "Yielding of Steel Sheets Containing Slits," *Journal of the Mechanics of Physics and Solids*, Vol. 8, pp. 100–104.

Goodier, J. N., and Field, F. A., 1963, "Plastic Energy Dissipation in Crack Propagation," *Fracture of Solids*, D. C. Drucker and J. J. Gilman, eds., Wiley (Interscience), New York, pp. 103–118.

Hull, D., 1973, "The Microstructure and Properties of Crazes," *Deformation and Fracture of High Polymers*, Kausch, Hassell, and Jaffe, eds., Plenum Press, New York, pp. 171–189.

Kramer, E. J., and Hart, E. W., 1984, "Theory of Slow, Steady State Crack Growth in Polymer Glasses," *Polymer*, Vol. 25, pp. 1667–1678.

Lauterwasser, B. D., and Kramer, E. J., 1979, "Microscopic Mechanisms of Craze Growth and Fracture," *Philosophical Magazine A.*, Vol. 39, 4, pp. 469–495.

Rice, J. R., 1966, "Plastic Yielding at Crack Tip," *Proceedings of the 1st International Conference on Fracture*, Sendai, Japan, T. Yokobori et al., eds., Vol. I, Japanese Soc. for Strength and Fracture of Materials, Tokyo, pp. 283–308.

Schinker, M. G., and Döll, W., 1979, "Interference Optical Measurements of Large Deformations at the Tip of a Running Crack in a Glassy Thermoplastic," *Physics Conference Series No. 47*, J. Harding, ed., Inst. of Physics, Bristol, England, pp. 224–232.

Ungsuwarungsri, T., and Knauss, W. G., 1986, "A Nonlinear Analysis of an Equilibrium Craze: Part I – Problem Formulations and Solutions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, No. 1, pp. 44–51.

Wang, W. V., and Kramer, E. J., 1982, "A Distributed Dislocation Stress Analysis for Crazes and Plastic Zones at Crack Tips," *Journal of Materials Science*, Vol. 17, pp. 2013–2026.

Continuum Damage Mechanics: Part I—General Concepts

J. L. Chaboche
Office National d'Etudes et de Recherches Aérospatiales
F-92320 Châtillon Cedex, France

Continuum Damage Mechanics (C.D.M.) has developed continuously since the early works of Kachanov and Rabotnov. It constitutes a practical tool to take into account the various damaging processes in materials and structures at a macroscopic continuum level. The main basic features of C.D.M. are considered in the first part together with its present capabilities, including damage definitions and measures, and its incorporation into a thermodynamic general framework. Practical damage growth equations will be reviewed in the second part of the paper.

1 Introduction

In the classical sense, Continuum Mechanics allows us to describe the heterogeneous microprocesses involved during the straining of materials and structures at the macroscale; elastic and plastic strains and the corresponding hardening effects are generaly accepted to be represented by global continuum variables, even if microdefects, such as grains, subgrains, dislocations, and cells, are not really homogeneous. The damaging processes correspond to localizations and accumulations of the strains and are considerably more irreversible. Due to the larger scale to be considered, continuum concepts are more difficult to introduce. Since the early works of Kachanov (1958) and Rabotnov (1969), who considered the creep of metals, the concept of macroscopic damage variables has developed markedly. The distributed defects in materials and structures not only lead to crack initiation and final fracture, but also induce progressive material deterioration (material damage) which can be measured through the decrease of strength, stiffness, toughness, stability, and residual life.

The new concepts initiated the development of Continuum Damage Mechanics (C.D.M.) (Hult, 1979; Chaboche, 1981; Krajcinovic, 1984). They are supported by the general framework of thermodynamics of irreversible processes and offer complementary possibilities to Fracture Mechanics.

The aim of the present paper is to review some general features of C.D.M. and to summarize its main possibilities, considering successively the following aspects:

(*a*) Definitions and measures of damage, including the possibility of describing the microstructural damaged state in terms of appropriate mechanical variables

(*b*) Description of the mechanical behavior of the damaged material. This can be studied in the framework of thermodynamics and can include the influence of damage anisotropy

(*c*) Formulation of equations governing the evolution of these damage variables (see Part II).

2 Damage Measures and Definitions

The first step in developing a damage theory concerns the definition of the damage variable. Obviously damage is not directly measurable as strain or plastic strain. In the present section, we consider different ways of defining the damage internal variable through indirect measurement procedures. In fact, such measurements are not always practicable but furnish conceptual definitions. Let us note that interpretation of each measure in terms of a damage variable requires a corresponding model.

2.1 The Problem of Crack Initiation. Before any damage theory can be developed, it is necessary to know precisely what we mean by the ultimate state of the damage processes; under the present development of C.D.M., this final state corresponds generally to macroscopic crack initiation, that is the "breaking up" of the continuum volume element. A large degree of arbitrariness is present in the definition of crack initiation, especially in fatigue where the behavior of newly nucleated cracks and short cracks shows various complex interactions with the microstructure (see the schematic view of Fig. 1(*a*) and Jeal, 1985).

In fact the consideration of a macroscopic crack in the framework of Fracture Mechanics supposes a defect sufficiently large as compared to the microscopic heterogeneities (grains, subgrains, other defects and microcracks. . .). The main macroscopic crack is assumed to be developed through several grains, in order to show a sufficient macroscopic homogeneity, in size, geometry and direction, leading to a possible treatment through the Fracture Mechanics concepts (see schematic illustration by Fig. 1).

2.2 Damage Measures Through the Remaining Life. From an engineer's point of view, the main objective of a damage theory is to allow predictions of the lifetime of a structure. Then the remaining life concept is a natural way to define damage. The most conventional definition for such a damage parameter is the life ratio, N/N_F in fatigue, where N and N_F

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 10, 1986; final revision May 27, 1987.

represent, respectively (for a given loading condition), the present number of cycles already applied and the total number of cycles to crack initiation (or failure). In this case the damage theory corresponds to the linear Miner's rule. For instance, the remaining life at the second level, N_2, after damaging to N_1/N_{F_1} at a first level, is:

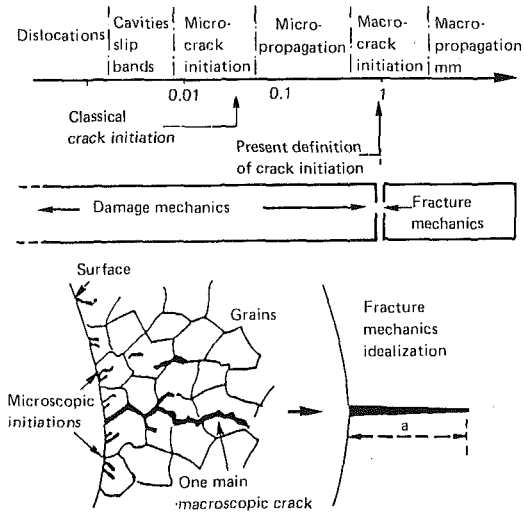


Fig. 1 (*a*) Schematics of the fatigue crack growth; (*b*) illustration of a macroscopic crack initiation concept

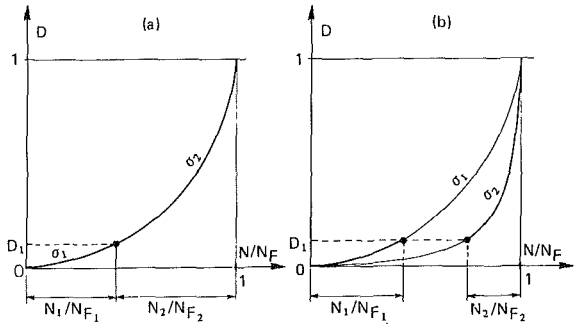


Fig. 2 Schematic of damage evolution curves as deduced from remaining life measures

$$\frac{N_2}{N_{F_2}} = 1 - D = 1 - \frac{N_1}{N_{F_1}} \tag{1}$$

More generally, the remaining life concept does not necessarily lead to the linear rule; after a certain damaging process, the present damage is measured by performing a "measure test," with a fixed loading under which the nominal life (for an initially undamaged specimen) is N_{F_2}. If the measured remaining life is N_2, the damage after the initial damaging process is:

$$D = 1 - \frac{N_2}{N_{F_2}} \tag{2}$$

The remaining life measurements provide evidence of interesting properties of damage. For instance, in fatigue there is not a unique damage evolution curve as a function of the life ratio N/N_F, but a dependency on the applied loading (Fig. 2) (Manson, 1979; Chaboche, 1974). This leads to the conclusion that damage growth equations have to show unseparability, of the damage and loading variables (Krempl, 1977). However, such measurements are not sufficient to completely fix the values of the damage. As shown by Chaboche (1980), a one-to-one mapping changes the damage value without changing any remaining life result. Let us note that the remaining life measures are also practicable for creep damage (Woodford, 1973).

2.3 Damage Measures from the Microstructure. A second natural way to define a damage variable is to observe and quantify the irreversible defects: intergranular cavities in creep, surface microcracks in fatigue, dimensions of cavities in ductile fracture. Such measurements have already been used in many situations (see Dyson and McLean, 1977; Levaillant and Pineau, 1982). Some difficulties arise when interpreting these results:

(*a*) Such measures are destructive, which limit their use to observe the development of damage.

(*b*) Defects are difficult to observe during the first phase of the damaging processes. Moreover, the initial state is not easily characterized.

(*c*) The quantification has to be done in terms of macroscopic variables which are usable in structure computation. Then, in each a case, convenient hypotheses have to be considered and a particular model developed.

Interpretation in terms of mechanical parameters can be

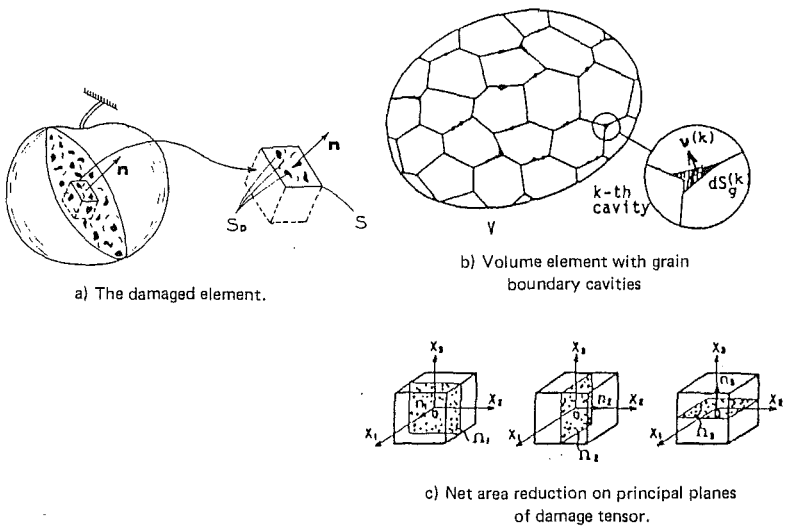


Fig. 3 Net area reduction: (*a*) The damaged element; (*b*) volume element with grain boundary cavities; (*c*) net area reduction on principal planes of damage tensor

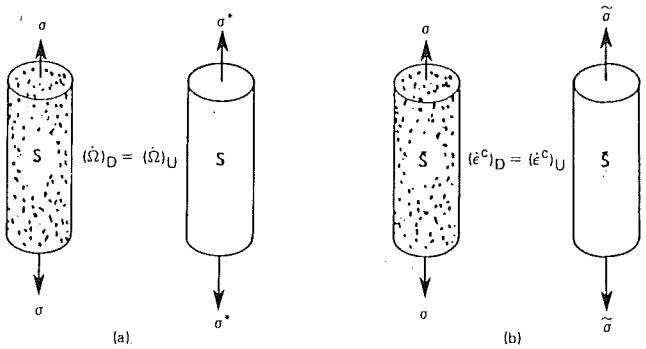


Fig. 4 (*a*) Net stress tensor for damage growth; (*b*) effective stress tensor

made by means of a net area reduction (Fig. 3(*a*)). By definition, the damage variable D_n associated with direction n is:

$$D_n = \frac{S_D}{S} = \frac{S - S^*}{S} \tag{3}$$

$D_n = 0$ corresponds to the undamaged state, $D_n = D_c$, a critical value, corresponds to the rupture of the element ($0.2 < D_c < 0.8$ for metals).

The corresponding net stress concept has been developed by Murakami and Ohno (1980) in the case of creep damage of polycrystalline metals. The microstructural changes can be characterized mainly by nucleation and growth of various cavities on grain boundaries. Assuming that the principal effect consists of a net area reduction due to the distribution of cavities, the state of material damage may be described by a second rank symmetric tensor (Murakami, 1983) (Fig. 3(*b*)), generalizing equation (3) as:

$$\Omega = \frac{3}{S_g(V)} \sum_{k=1}^{N} \int_V \vec{v}^{(k)} \otimes \vec{v}^{(k)} dS_g^{(k)} \tag{4}$$

where $dS_g^{(k)}$ and $\vec{v}^{(k)}$ are, respectively, the area of a given boundary occupied by the kth cavity in volume V and the unit normal vector to it. Figure 4(*a*) indicates schematically that some equivalence is supposed between the present damaged material and the undamaged one (in terms of damage rate) to define the net stress $\boldsymbol{\sigma}^*$.

The above definition takes into account the directional nature of damage in a natural way. Other methods can be used to introduce the anisotropy of damage. An attractive one, developed in a framework similar to the slip theory of plasticity, is to use a family of vectors (Krajcinovic, 1983) each of them being associated with to a typical direction of microcracks.

2.4 Damage Measures Through Physical Parameters and the Effective Stress Concept. The influence of damage on physical quantities can be measured and used to define properly the damage variable:

- Density change (Jonas and Baudelet, 1977), which can be interpreted as a damage variable in ductile failure;
- resistivity change (Cailletaud et al., 1980) which, through a convenient model, leads to very similar damage measures to that for mechanical parameters (see below);
- acoustic emission, change in sound velocity. . . ;
- change in the fatigue limit (Bui-Quoc et al., 1971) which can also be interpreted in terms of the remaining life;
- change in the mechanical behavior of the material, interpreted through the effective stress concept:

"A damaged volume of material under the applied stress σ shows the same strain response as the undamaged one submitted to the effective stress $\tilde{\sigma}$." If the damage D represents the

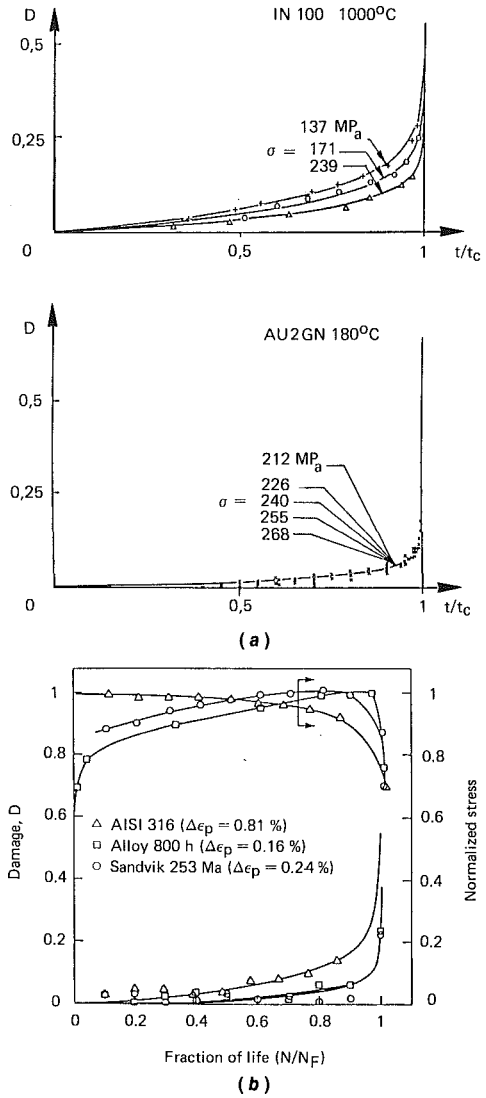


Fig. 5 (*a*) Creep damage measures—alloys IN 100 at 1000°C and AU2GN at 180°C; (*b*) damage accumulation curves as measured from the change in elastic response with corresponding cyclic stress response

loss of effective area taking into account decohesions and local stress concentrations, one can write:

$$\tilde{\sigma} = \sigma \frac{S}{\tilde{S}} = \frac{\sigma}{1 - D} \tag{5}$$

This definition through the gross behavior of the material is supported by the results of homogenization techniques (Duvaut, 1976), considering periodic arrays of defects.

Damage measures through the effective stress concept have been performed in several situations. Let us mention:

- The case of ductile rupture (Lemaître, 1985), which measures the change of the elastic modulus. From elasticity equations for both damaged and undamaged material

$$\sigma = \tilde{E}\,\epsilon_e \quad \tilde{\sigma} = E\,\epsilon_e$$

one obtains:

$$\tilde{\sigma} = \frac{E}{\tilde{E}}\,\sigma = \frac{\sigma}{1 - D} \rightarrow D = 1 - \frac{\tilde{E}}{E} \tag{6}$$

- Brittle creep damage (Lemaître and Chaboche, 1978), using the power law to describe secondary creep, which can be considered as an undamaged state:

$$\dot{\epsilon}_s = \left(\frac{\sigma}{\lambda}\right)^N \tag{7}$$

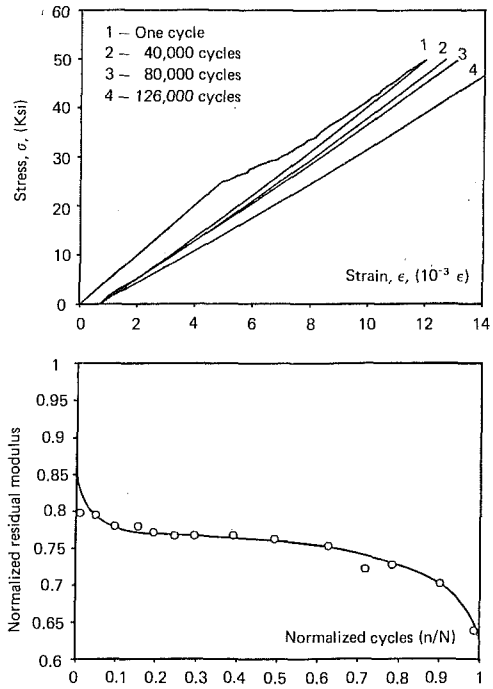


Fig. 6 Fatigue of $[0/90_4]_S$ graphite/epoxy specimens, σ_{max} = 345 MPa (50 ksi), R = 0.1, from Charewicz and Daniel (1985). (a) stress-strain curves at various stages of the fatigue life; (b) normalized residual elastic modulus.

Several tests give the exponent N. Damage follows easily from the measurement of strain rate during tertiary creep and the effective stress concept (5):

$$D = 1 - \left(\frac{\dot{\epsilon}}{\dot{\epsilon}_s}\right)^{1/N} \tag{8}$$

Figure 5(*a*) shows the case of the superalloy IN 100 (Lemaître and Chaboche, 1978).

- In the case of fatigue damage, the measures are more difficult to interpret, due to its particular localization (often near the surface) and to the superposition of cyclic hardening or softening processes (which cannot be directly considered as damage processes!). In fact, the effective stress concept (Chaboche, 1974) allows some useful correlations, both through elastic and inelastic behaviors, as shown in Fig. 5(*b*) taken from Plumtree and Nilsson (1986). In the three materials, showing different cyclic responses, the damage measured through elastic modulus changes (6) agrees fairly well with the decreasing part of the peak stress evolution. Such measurements also correlates well with quantitative microcrack evaluations (Cailletaud and Levaillant, 1984; Hua and Socie, 1984).
- Damage in composites develops continuously and grows by various mechanisms at the microscale (as in fiber debonding, matrix microcracking, delamination). The case of fatigue is illustrated in Fig. 6 for graphite/epoxy laminates (from Charewicz and Daniel, 1985). One observes simultaneously the change in the elastic modulus on the first cycle (due to partial debonding), its progressive decrease during fatigue (Fig. 6(*b*)), accompanying microcracking of the matrix with some stabilization, and the final delamination giving rise to rupture. Let us note some similarities with the case of metals, with much more pronounced effects. Charewicz and Daniel (1985) show clearly the correspondence between the decrease of elastic modulus, the decrease of residual strength, and the remaining life in fatigue.

This last example shows the adequacy of Damage Mechanics to describe completely different kinds of materials.

Table 1

Observable variable	Internal variables	Associated variables
Elastic strain tensor ϵ_e		Stress tensor σ
Temperature T		Entropy S
	Isotropic hardening r	Size increase of yield surface R
	Damage D	Damage strain energy release rate Y

Talreja (1985) gives further theoretical developments on the case of composites.

Let us note that damage can be active or passive (Krajcinovic and Fonseka, 1981). When the defects, especially microcracks, are closed, they do not affect the macroscopic behavior. The damage state is not eliminated but has to be considered as passive; this can be the case under compression, for instance. The theory developed by Krajcinovic and Fonseka (1981) allows a natural treatment of this effect, which is particularly important for materials like concrete. Ladevèze and Lemaître (1984) gave also some developments around this point.

Another important feature of damage is its probabilistic nature (at the microscale). When considering the residual strength of the material, that leads to statistical definitions of the damage variables (Kracjcinovic and Silva, 1982; Chrzanowski, 1976).

3 Thermodynamic Aspects

The present developments are based on a thermodynamic theory of irreversible processes with internal state variables (Sidoroff, 1975; Germain et al., 1983). The presentation is limited here to the simple case of isotropic hardening within the small strain hypothesis and to an isotropic damage. The extension to kinematic hardening is well known (Halphen and Nguyen, 1975), and theories with anisotropic damage evolution have been developed by Chaboche (1978) and Lemaître and Chaboche (1985). Generalization to finite strain can be found in Rousselier (1980).

The chosen state variable are given in Table 1 for the present case. The plastic strain tensor is defined from the total strain tensor by $\epsilon_p = \epsilon - \epsilon_e$ and the accumulated plastic strain by $\dot{p} = (2/3\ \dot{\epsilon}_p : \dot{\epsilon}_p)^{1/2}$. In the case of a nondamaging material and neglecting the time recovery effects, it can be demonstrated (Lemaître and Chaboche, 1985) that r reduces to the accumulated plastic strain p (in the case of a generalized associative flow rule).

3.1 Thermodynamic Potential. The specific free energy Ψ, taken as the thermodynamic potential in which elasticity and plasticity are uncoupled, gives the law of thermoelasticity coupled with damage:

$$\Psi = \Psi_e(\epsilon_e, T, D) + \Psi_p(T, r) \tag{9}$$

As proposed by Chaboche (1977), the damaged elastic behavior is described through a strain equivalence and, referring to the effective stress concept:

$$\Psi_e = \frac{1}{2\rho}(1-D)\Lambda : \epsilon_e : \epsilon_e \tag{10}$$

The stress is defined as:

$$\sigma = \rho\,\frac{\partial \Psi}{\partial \epsilon_e} = (1-D)\Lambda : \epsilon_e \quad \text{or} \quad \sigma = \tilde{\Lambda} : \epsilon_e \tag{11}$$

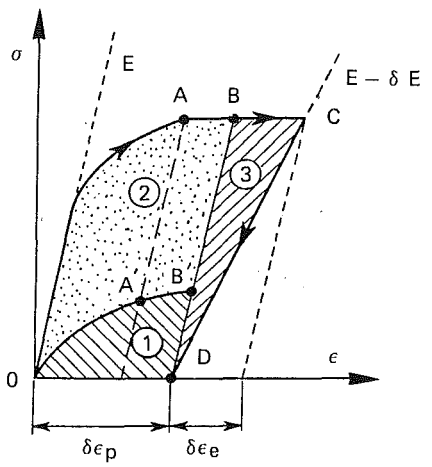


Fig. 7 Schematic of the dissipation during plastic flow and damage growth

The thermodynamic forces Y associated with D and R associated with r are defined as:

$$Y=\rho\frac{\partial\Psi}{\partial D}=-\frac{1}{2}\Lambda:\epsilon_e:\epsilon_e \qquad R=\rho\frac{\partial\Psi}{\partial r} \tag{12}$$

Let us note that D includes all the damaging effects, the density ρ being considered as constant. In the finite strain case (Rousselier, 1980), the change of ρ (due to the growth of cavities) can be used as a damage variable. Moreover, in the present case, complete separability of the hardening and damage processes is assumed, Ψ_p does not depend on D.

W_e being the density of elastic strain energy defined by $dW_e = \sigma: d\epsilon_e$, the expression for Y shows that (Chaboche, 1977):

$$-Y=\frac{W_e}{1-D}=\frac{1}{2}\frac{dW_e}{dD} \text{ at constant } \sigma \text{ and Temperature} \tag{13}$$

Then, the variable $-Y$ can be considered as the elastic strain energy release rate associated with a unit damage growth. The analogy with Fracture Mechanics concepts is clear. $-Y$ may be calculated as a function of the hydrostatic stress $\sigma_H = 1/3\, Tr(\sigma)$ and the Von Mises equivalent stress $\sigma_{eq} = (3/2\ \sigma': \sigma')^{1/2}$ where σ' is the stress deviator (Lemaître, 1984):

$$-Y=\frac{\sigma_{eq}^2}{2E(1-D)^2}\left[\frac{2}{3}(1+\nu)+3(1-2\nu)\left(\frac{\sigma_H}{\sigma_{eq}}\right)^2\right] \tag{14}$$

The relation (13) defining $-Y$ as the elastic strain energy release rate and the above dissipation aspects can be illustrated in Fig. 7, showing the different parts of dissipation during the plasticity and rupture processes (Lemaître and Chaboche, 1985). The curve $OA'B'$ represents the evolution of hardening during plastic flow OAB. Parts AB and BC correspond, respectively, to the plastic flow and the elastic strain increase during the damaging process (schematized at constant stress). The total dissipated energy separates into: (1) the energy stored in the system (hardening); (2) the heat dissipated energy; (3) the energy released by the system during the damaging process $-Y\,\delta D$, eventually converted into heat.

The fact that energy stored in the material is the work done above the initial yield limit is a consequence of some simplifying assumptions on the isotropic hardening. In fact, additional energies are converted into heat, which can be described by using the superposition of a nonlinear kinematic hardening as in Lemaître and Chaboche (1985) and specific choices for the part Ψ_p of the free energy.

3.2 Dissipation. The second principle of thermodynamics imposes that the intrinsic dissipation has to be positive:

$$\sigma:\dot{\epsilon}_p - R\,\dot{r} - Y\,\dot{D}>0 \tag{15}$$

Hardening and damage being uncoupled, it is sufficient to assume:

$$-Y\,\dot{D}>0 \tag{16}$$

As $-Y$ is a quadratic function, this leads to $\dot{D} > 0$.

The rupture criterion "$-Y = |Y| = Y_c \rightarrow$ crack initiation" corresponds to an elastic energy criterion. It may be written in terms of D through the one-dimensional rupture stress σ_R:

$$Y_c=\frac{\sigma_R^2}{2E(1-D_c)^2} \rightarrow D_c=1-\frac{\sigma_R}{(2EY_c)^{1/2}} \tag{17}$$

Many experiments have shown that: $0.2 < D_c < 0.8$, which allows $(1 - D_c)^x$ to be neglected with regard to 1 when x is much greater than 1.

The potential of dissipation is a scalar convex function of flux variables ($\dot{\epsilon}_p$, $\dot{r}$, $\dot{D}$ and the heat flux q) or their dual variables (by means of the Legendre-Fenchel transform), the state variables acting as parameters (Germain et al., 1983):

$$\phi(\sigma, R, Y; \epsilon_e, T, p, D)$$

It gives the constitutive equations for the evolution of dissipative variables, written here as:

$$\dot{\epsilon}_p=\frac{\partial\phi^*}{\partial\sigma} \qquad \dot{r}=-\frac{\partial\phi^*}{\partial R} \qquad \dot{D}=\frac{\partial\phi^*}{\partial Y} \tag{18}$$

If ϕ^* is a convex function of $-Y$, the damage dissipation (16) is automatically positive (Chaboche, 1977).

In the case of time independent plasticity and isotropic hardening, the plastic flow can be particularized with the Von Mises plastic potential (Lemaître and Chaboche, 1985):

$$f(\sigma, R, D)=\tilde{\sigma}_{eq}-R-k<0 \tag{19}$$

where $\tilde{\sigma}_{eq}$ is the equivalent effective stress, here $\sigma_{eq}/(1-D)$.

It follows then from the normality rule (σ' is the stress deviator):

$$\dot{\epsilon}_p=\lambda\frac{\partial f}{\partial\sigma}=\frac{3}{2}\frac{\lambda}{1-D}\frac{\sigma'}{\sigma_{eq}} \tag{20}$$

$$\dot{r}=\lambda\frac{\partial f}{\partial R}=\lambda=\left(\frac{2}{3}\dot{\epsilon}_p:\dot{\epsilon}_p\right)^{1/2}=(1-D)\dot{p}$$

The plastic multiplier is determined from the consistency condition $\dot{f} = 0$. When damage is zero, the state variable reduces to the accumulated plastic strain p.

In fact the above normality rule is a sufficient but not a necessary condition (Onat, 1985). In some cases it appears as too restrictive, especially for materials or rupture conditions where the energetic rupture criterion (17) does not allow a correct description, as is the case in creep. The only condition to be verified is $\dot{D} > 0$ (because $-Y$ is always positive).

The present theory can be generalized to the case of large strains, including some modifications. In the theory developed by Rousselier (1980), two damage parameters are used, the first one corresponds to the density change and acts on the elastic behavior, the second one follows from an explicit dependency between the plastic potential and the hydrostatic stress and generalizes in some way the Rice and Tracey (1969) formula for the growth of spherical voids under high triaxiality.

4 Conclusion

The general concepts of Damage Mechanics have been reviewed, considering the measures and definitions of damage variables and their incorporation into a general thermodynamic framework. One of the main features of C.D.M. is to take into account the coupling effects between damaging processes and stress-strain behavior.

Different damage growth equations can be specified and

particularized, including the description of creep and fatigue processes, ductile damage, and brittle damage. This is the subject of the second part of this review (Chaboche, 1988), which will also consider the possibilities of C.D.M. in the lifetime and crack prediction techniques.

References

Bui-Quoc, T., Dubuc, J., Bazergui, A., Biron, A., 1971, "Cumulative Fatigue Damage under Stress Controlled Conditions," *J. of Basic Engng. Trans. ASME*, pp. 691-698.

Cailletaud, G., Policella, H., and Baudin, G., 19808, "Mesure de Déformation et d'Endommagement par Méthode Électrique," *La Recherche Aérospatiale*, No. 1980-1.

Cailletaud, G., and Levaillant, C., 1984, "Creep-Fatigue Life Prediction: What about Initiation?" *Nuclear Engineering and Design*, No. 83, pp. 279-292.

Chaboche, J. L., 1974, "Une Loi Différentielle d'Endommagement de Fatigue avec Cumulation non Linéaire." *Revue Française de Mécanique*, No. 50-51, english translation in "Annales de l'ITBTP," HS 39, 1977.

Chaboche, J. L., 1977, "Sur l'utilisation des Variables d'État Interne pour la Description du Comportement Viscoplastique et de la Rupture par Endommagement," *Symp. Franco-Polonais de Rhéologie et Mécanique*, Cracovie.

Chaboche, J. L., 1978, "Description Thermodynamique et Phénoménologique de la Viscoplasticité Cyclique avec Endommagement," Thèse Univ. Paris VI et Publication ONERA, No. 1978-3.

Chaboche, J. L., 1980, "Lifetime Predictions and Cumulative Damage Under High Temperature Conditions," *Int. Symp. on Low Cycle Fatigue and Life Prediction*, Firminy, France, ASTM STP 770.

Chaboche, J. L., 1981, "Continuous Damage Mechanics: A Tool to Describe Phenomena Before Crack Initiation," *Nuclear Engineering and Design*, Vol. 64, pp. 233-247.

Chaboche, J. L., 1988, "Continuum Damage Mechanics—Part II: Damage Growth, Crack Initiation, and Crack Growth," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, No. 1, pp. 65-72.

Charewicz, A., and Daniel, I. M., 1985, "Fatigue Damage Mechanisms and Residual Properties of Graphite/Epoxy Laminates," *Symp. IUTAM on Mechanics of Damage and Fatigue*, Haifa, Israel.

Chrzanowski, M., 1976, "Use of the Damage Concept in Describing Creep-Fatigue Interaction Under Prescribed Stress," *Int. J. Mech. Sci.*, Vol. 18, pp. 69-73.

Duvaut, C., 1976, "Analyse Fonctionnelle—Mécanique des Milieux Continus-Homogénéisation," *Theoretical and Applied Mechanics*, North-Holland, Amsterdam.

Dyson, V. D., and McLean, D., 1977, "Creep of Nimonic 80A in Torsion and in Tension," *Met. Sci.*, Vol. 2, 37.

Germain, P., Nguyen, Q. S., and Suquet, P., 1983, "Continuum Thermodynamics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 1010-1020.

Halphen, B., and Nguyen, S., 1975, Sur les Matériaux Standard Généralisés," *J. de Mécanique*, Vol. 14, No. 1, pp. 39-63.

Hua, C. T., and Socie, D. F., 1984, "Fatigue Damage in 1045 Steel Under Constant Amplitude Biaxial Loading," *Fatigue Engng. Mater. Struct.*, Vol. 7, No. 3, pp. 165-179.

Hult, J., 1979, "Continuum Damage Mechanics—Capabilities Limitations and Promises," *Mechanisms of Deformation and Fracture*, Pergamon, Oxford, pp. 233-347.

Jeal, R. H., 1985, "Damage Tolerance Concepts for Critical Engine Components," *60th SMP-AGARD Meeting*, San Antonio, AGARD-CP-393.

Jonas, J. J., and Baudelet, B., 1977, "Effect of Crack and Cavity Generation on Tensile Stability," *Acta Metallurgica*, Vol. 25, Pergamon Press, eds., pp. 43-50.

Kachanov, L. M., 1958, "Time of the Rupture Process Under Creep Conditions," *Izv. Akad. Nauk. SSR*, Otd Tekh. Nauk No. 8, pp. 26-31.

Kracjcinovic, D., 1983, "Creep of Structures—A Continuous Damage Mechanics Approach," *J. Struct. Mech.*, Vol. 11, pp. 1-11.

Krajcinovic, D., 1984, "Continuum Damage Mechanics," *Applied Mechanics Review*, Vol. 37, No. 1.

Krajcinovic, D., and Fonseka, G. U., 1981, "The Continuous Damage Theory of Brittle Materials," Parts 1 and 2, ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 809-824.

Krajcinovic, D., and Silva, M. A. G., 1982, "Statistical Aspects of the Continuous Damage Theory," *Int. J. Solids Structures*, Vol. 18, No. 7, pp. 551-562.

Krempl, E., 1977, "On Phenomenological Failure Laws for Metals under Repeated and Sustained Loading (Fatigue and Creep)," *Conf. on Environmental Degradation of Engineering Materials*, Blacksburg, VA.

Ladeveze, P., and Lemaitre, J., 1984, "Damage Effective Stress in Quasi-Unilateral Condition," *IUTAM Congress*, Lyngby, Denmark.

Lemaitre, J., 1984, "How to Use Damage Mechanics," *Nuclear Engng. and Design*, Vol. 80, pp. 233-245.

Lemaitre, J., 1985, "A Continuum Damage Mechanics Model for Ductile Fracture," ASME *J. Engng. Mat. and Technology*, Vol. 197, pp. 83-89.

Lemaitre, J., and Chaboche, J. L., 1978, "Aspect Phénomènologique de la Rupture par Endommagement," *J. de Mécanique Appliquée*, Vol. 2, No. 3, pp. 317-365.

Lemaitre, J., and Chaboche, J. L., 1985, "Mécanique des Matériaux Solides," Dunod, Paris.

Levaillant, C., and Pineau, A., 1980, "Assessment of High Temperature LCF and Life of Austenitic Stainless Steels by Using Intergranular Damage as a Correlating Parameter," *LCF and Life Prediction*, ASTM-STP 770, Amzallag, Leis, Rabbe, eds., 1982, pp. 169-193.

Manson, S. S., 1979, "Some Useful Concepts for the Designer in Treating Cumulative Damage at Elevated Temperature," *I.C.M. 3*, Cambridge, Vol. 1, pp. 13-45.

Murakami, S., 1983, "Notion of Continuum Damage Mechanics and its Application to Anisotropic Creep Damage Theory," *J. Eng. Mat. and Technology*, Vol. 105, p. 99.

Murakami, S., and Ohno, N., 1980, "A Continuum Theory of Creep and Creep Damage," *3rd IUTAM Symp. on Creep in Structures*, Leicester.

Onat, E. T., 1985, "Representation of Mechanical Behaviour in the Presence of Internal Damage," *IUTAM Symp. on Mechanics of Damage and Fatigue*, Haifa, Israel.

Plumtree, A., and Nilsson, J. O., 1986, "Damage Mechanics Applied to High Temperature Fatigue," *Int. Spring Meeting "Fatigue at High Temperature,"* Paris.

Rabotnov, Y. N., 1969, *Creep Problems in Structural Members*, North-Holland.

Rice, J. R., and Tracey, D. M., 1969, "On the Ductile Enlargement of Voids in Triaxial Stress Fields," *J. Mech. Phys. Solids*, p. 201.

Rousselier, G., 1980, "Finite Deformation Constitutive Relations Including Ductile Fracture Damage," *IUTAM Symp. on Three-Dimensional Constitutive Relations and Ductile Fracture*, Dourdan, Nemat-Nasser, eds., North-Holland Publ. Comp., pp. 331-355.

Sidoroff, F., 1975, "On the Formulation of Plasticity and Viscoplasticity with Internal Variables," *Arch. Mech.*, Poland, Vol. 27, No. 5-7, pp. 807-819.

Talreja, R., 1985, "A Continuum Mechanics Characterization of Damage in Composite Materials," *Proc. R. Soc. Lond.*, No. A 399, pp. 195-216.

Woodford, D. A., 1973, "A Critical Assessment of the Life Fraction Rule, for Creep Rupture under Nonsteady Stress or Temperature," *Int. Conf. on Creep Fatigue Elevated Temperature Applications*, Philadelphia, PA.

Continuum Damage Mechanics: Part II—Damage Growth, Crack Initiation, and Crack Growth

J. L. Chaboche
Office National d'Etudes et de Recherches Aérospatiales, F-92322 Châtillon Cedex, France

Continuum Damage Mechanics (CDM) allows the description of the influence of damage on the stress-strain behavior of materials. In the present part, some practical damage growth equations are reviewed for creep, fatigue, creep-fatigue interaction, ductile damage, and brittle damage. The capabilities of CDM to improve both the crack initiation and crack propagation predictive tools are then discussed. Particular attention is given to the new developments of the "local approaches to fracture."

1 Introduction

The general purpose of Continuum Damage Mechanics is to introduce the possibility of describing the coupling effects between damage processes and the stress-strain behavior of materials (Hult, 1979; Chaboche, 1981; Krajcinovic, 1984). Part I of the present paper (Chaboche, 1988) reviewed the main features of CDM and its incorporation in a general thermodynamic framework. The damage growth equations were written in a general form, by means of a dissipative potential.

The present part develops some useful and practical damage rate equations. Different damage variables are associated with different damage processes such as creep, fatigue, ductile and brittle damage. Various growth equations are considered and discussed on the basis of some examples on various kinds of materials. They can be incorporated in the general framework presented in Part I.

Applications of CDM and, more precisely, of the damage growth equations are possible in two domains:

(*a*) To improve the life prediction techniques, with the purpose of calculating damage growth and crack initiation in structures. Several domains can be considered: high cycle and low cycle fatigue, creep, and creep-fatigue.

(*b*) To improve the macroscopic crack growth calculation techniques, usually based on the Fracture Mechanics concepts. Including the damage processes explicitly allows additional possibilities of improving the "local approaches to fracture," which are now extensively developed. Sections 3.2 and 3.3 consider the respective capabilities of the two approaches.

2 Damage Growth Equations

2.1 Creep Damage. Continuum Damage Mechanics was developed first for the case of creep damage (Kachanov, 1958; Rabotnov, 1969). The Rabotnov-Kachanov equation can be considered as very classical. With D varying between 0 for the undamaged material and 1 for the rupture, under pure tensile stresses, the expressions reduce to:

$$dD = \left(\frac{\sigma}{A}\right)^r (1-D)^{-k} dt \tag{1}$$

where r, k, A are material and temperature dependent coefficients. Their determination is made from constant stress creep data, for which the integration from 0 to 1 gives the rupture time:

$$t_c = \frac{1}{k+1}\left(\frac{\sigma}{A}\right)^{-r} \tag{2}$$

while evolution of damage is given by:

$$D = 1 - \left(\frac{t}{t_c}\right)^{1/k+1} \tag{3}$$

The concept of effective stress, introduced in the secondary creep law (Norton's equation), allows the calculation of tertiary creep curves as well as the prediction of the change in creep ductility (Chaboche, 1978, 1984). Figure 1 gives an example:

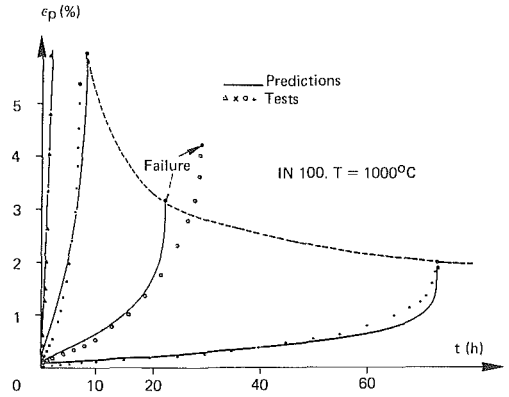


Fig. 1 Calculated and measured creep curves on superalloy IN 100. Prediction of creep ductility.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, October 10, 1986; final revision May 27, 1987.

$$\dot{\epsilon}_p = \left[\frac{\sigma}{K(1-D)}\right]^n \tag{4}$$

Under multiaxial stress conditions the damage equations can be generalized by describing isodamage surfaces (or isochronous surfaces) defined with proper stress invariants: (*a*) the octahedral shear stress $J_2(\sigma)$ related to the effects of shear; (*b*) the hydrostatic stress $J_1(\sigma) = Tr(\sigma) = \sigma$: **1**, which greatly affects the growth of the cavities; and (*c*) the maximum principal stress $J_0(\sigma) = \sigma_1$, which opens the microcracks and causes them to grow.

Following the method of Hayhurst (1972), the equivalent stress can be defined through a linear combination:

$$\chi(\sigma) = \alpha J_o(\sigma) + \beta J_1(\sigma) + (1-\alpha-\beta)J_2(\sigma), \tag{5}$$

where α, β are coefficients dependent on the material and temperature. The time to failure under a fixed multiaxial stress is expressed as:

$$t_c = \frac{1}{k+1}\left\langle\frac{\chi(\sigma)}{A}\right\rangle^{-r} \tag{6}$$

The creep damage equations can be considered as taking into account the evolution of microstructural defects in an indirect manner. Works by the material scientists have shown that the increase of damage results from a combination of two mechanisms, the nucleation and growth of cavities (Greenwood, 1973; Dyson, 1979). It is possible to make some connections between the equations obtained from material science and the more macroscopic ones developed in the framework of CDM (Hayhurst, 1983).

For the multiaxial stress criterion, the physical interpretations often lead to a product form:

$$\chi(\sigma) = [J_o(\sigma)]^\alpha [J_2(\sigma)]^{1-\alpha} \tag{7}$$

which can describe adequately the isochronous surfaces in a large domain (Hayhurst, 1983; Delobelle, 1985). However, the form (7) automatically gives no damage under pure compression, which is not perfectly true in polycrystalline metals, as shown by two level creep tests (Policella et al., 1982). Let us mention a possible different form for the damage growth equation, using the creep strain instead of time (Constesti, 1986):

$$dD = A[J_o(\sigma)]^\alpha \epsilon_{eq}^\beta d\epsilon_{eq} \tag{8}$$

such a form was deduced from cavity measurements on notched specimens. It is possible to show that equations (1) and (4) are equivalent to equation (8) through the following one-to-one mapping (Contesti, 1986):

$$D \rightarrow [1-(1-D)^{k-n+1}]^{\beta+1} \tag{9}$$

2.2 Fatigue Damage. In the case of fatigue several aspects have to be considered:

(*a*) the existence of microinitiation and micropropagation stages;

(*b*) the nonlinear-cumulative effects for two-level tests or block-program loading conditions;

(*c*) the existence of a fatigue limit, but its marked decrease after prior damage;

(*d*) the effect of mean-stress either for the fatigue limit or for the S-N curves.

Fatigue damage accumulation models have been considered by Marco and Starkey (1954), Manson (1979), and Chaboche (1974). A common general form is obtained with:

$$dD = D^{\alpha(\sigma_M,\bar{\sigma})}\left[\frac{\sigma_M-\bar{\sigma}}{M(\bar{\sigma})}\right]^\beta dN \tag{10}$$

where σ_M and $\bar{\sigma}$ are, respectively, the maximum and mean stresses. Several choices for α (Chaudonneret and Chaboche, 1986) lead to the rules considered by Manson (1979), Subramanyan (1976), Hashin and Laird (1980). The key of the nonlinear effect is the dependency of α on σ_M and $\bar{\sigma}$ which, after integrating for a two-level test, gives:

$$\frac{N_2}{N_{F_2}} = 1 - \left(\frac{N_1}{N_{F_1}}\right)^{\frac{1-\alpha_2}{1-\alpha_1}} \tag{11}$$

The function $M(\bar{\sigma})$ is deduced from a linear dependency between $\bar{\sigma}$ and the fatigue limit.

This cumulative damage equation allows a very good description of multilevel fatigue tests (Chaboche, 1974). In a certain sense it includes in a continuous way the microinitiation and micropropagation phases as discussed by Manson (1979), Cailletaud and Levaillant (1984), and Miller and Zachariah (1977). Moreover, by a convenient variable change, equation (10) can be incorporated into Continuum Damage Mechanics, with the effective stress concept, as shown by damage measurements described by Lemaître and Chaboche (1978, 1985). The microcrack measurements made by Cailletaud and Levaillant (1984), Hua and Socie (1984), and Socie et al. (1983), for instance, then show the possible equivalence between: (*a*) the definition by the effective stress concept; (*b*) the definition in terms of the remaining life concepts (subjacent in equation (10)); and (*c*) the quantification of physical damage, in terms of microcracking.

In the case of Low-Cycle Fatigue, the conventional parametrization of the life is written in terms of the plastic (or total) strain range. Provided the existence of a one-to-one relation between σ_M and $\Delta\epsilon_p$ (the cyclic curve) equation (10) can still be used and contains independently the influence of mean-stress.

The generalization to multiaxial loading conditions is a difficult problem. At least two parameters have to be considered: (*a*) an equivalent shear-stress amplitude, and (*b*) a mean (or maximum) hydrostatic stress. Experiments near the fatigue limit show the independency in the mean shear-stress and in the hydrostatic range (Dang Van, 1973). A possible form to generalize equation (10) has been proposed by Chaboche (1978). Additional studies are required to generalize such models in the case of nonproportional loading conditions.

2.3 Creep-Fatigue Interaction. One advantage of the CDM approach in creep and fatigue is to allow a natural way to predict creep-fatigue interaction (Chrzanowski, 1976). The simplest hypothesis consists in a direct summation of creep and fatigue damages, which leads to (Chaboche, 1980):

$$dD = f_c(\sigma, D)dt + f_F(\sigma_M, \bar{\sigma}, D)dN \tag{12}$$

where f_c is, for instance, deduced from equation (1) and f_F from equation (10). These two functions can be determined from pure tensile creep tests and pure fatigue tests (high frequency). The conditions at low-frequency or with hold times are then predicted by integrating numerically equation (12). This approach has given good results for several materials (Lemaître and Chaboche, 1985; Plumtree and Lemaître, 1979; Del Puglia and Vitale, 1982).

Let us note that the additive hypothesis does not correspond to the direct addition of physical damages of different natures. For instance, microcracks and cavities are not added. Only their mechanical effects are added, in the framework of the effective stress concept. The mechanical effects are obtained by the above mentioned one-to-one mappings between the physical damage (cracks or cavities) and the corresponding macroscopic variable D (see Sections 2.1 and 2.2).

2.4 Oxidation-Fatigue-Creep Interactions. At high temperature, the influence of time is often increased by the superposition of oxidation processes (Rezai-Aria and Rémy, 1986). The oxidation may enhance both the fatigue and creep mechanisms, contributing to both damage nucleation (in the form of surface cracks, preferential grain boundary oxidation,

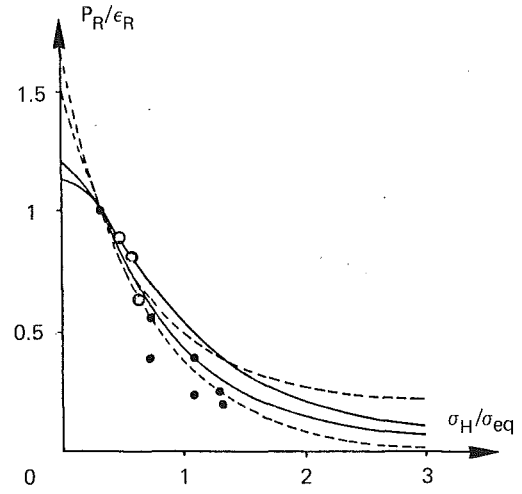


Fig. 2 Influence of triaxiality on fracture strain. • •: A508 steel; ○ ○: H steel; = = = = =: domain covered by McClintock/Rice-Tracey models; ═══: domain covered by the present model.

or internal voids) and damage growth (crack tip deterioration, void growth from particles).

Such additional factors have been considered as acting on the fatigue damage process in terms of a fatigue damage model including both an initiation and a propagation period (Cailletaud and Levaillant, 1984). The incorporation of oxidation in the creep damage equations can be considered as implicit (on a conceptual point of view).

2.5 Ductile Plastic Damage. Ductile damage in metals is essentially the initiation and growth of cavities due to large deformations. Experiments of ductile rupture show that the dissipative potential ϕ^* can be expressed, in the framework of time-independent plasticity (Lemaître and Chaboche, 1985):

$$\phi^* = \tilde{\sigma}_{eq} - R - k + \frac{S}{2}\frac{1}{1-D}\left(-\frac{Y}{S}\right)^2 \tag{13}$$

$$\dot{\epsilon}_p = \lambda \frac{\partial \phi}{\partial \boldsymbol{\sigma}}$$

$$\dot{r} = -\lambda \frac{\partial \phi^*}{\partial R} = \left(\frac{2}{3}\,\dot{\epsilon}_p : \dot{\epsilon}_p\right)^{1/2}(1-D) = \dot{p}(1-D) \tag{14}$$

$$\dot{D} = -\lambda \frac{\partial \phi^*}{\partial Y} = \frac{-Y}{S}\,\dot{p} \tag{15}$$

Using the relation (14) of Part I (Chaboche, 1987a), between σ_{eq} and $-Y$, and a particular choice for hardening:

$$\sigma_{eq} = R + k + K\dot{p}^{1/M} \tag{16}$$

K and M being coefficients, which are material dependent, one obtains the following differential constitutive equation for ductile plastic damage (Lemaître, 1985):

$$\dot{D} = \frac{K^2}{2ES}\left[\frac{2}{3}(1+\nu) + 3(1-2\nu)\left(\frac{\sigma_H}{\sigma_{eq}}\right)^2\right]p^{2/M}\dot{p} \tag{17}$$

In the case of radial loading, when the principal directions of stresses do not vary, the triaxiality ratio σ_H/σ_{eq} is constant and this expression may be integrated using the conditions:

$p < p_D$ (damage threshold) $\rightarrow D = 0$

$p = p_R$ (strain to rupture) $\rightarrow D = D_c$.

Neglecting elastic strain in the calculation of p and using $p_D/p_R = \epsilon_D/\epsilon_R$, the equation for damage evolution may be written in terms of the one-dimensional threshold ϵ_D and one-dimensional strain to rupture ϵ_R:

$$D = \frac{D_c}{\epsilon_R - \epsilon_D}\left\langle p\left[\frac{2}{3}(1+\nu) + 3(1-2\nu)\left(\frac{\sigma_H}{\sigma_{eq}}\right)^2\right] - \epsilon_D\right\rangle \tag{18}$$

Identification of such a model consists in the quantitative evaluation of the coefficients ϵ_D, ϵ_R, and D_c (Poisson's ratio being known from elasticity), which can be done from experimental measurements such as those mentioned in Section 2.4 of Part I.

As shown in Fig. 2, the present model gives the influence of triaxiality on the strain to failure (Lemaître, 1985). It compares fairly well with the McClintock (1968) or Rice and Tracey (1969) models and can be used to predict the fracture limits of metal forming (Lemaître, 1984).

Let us note that the modified thermodynamic framework developed by Rousselier (1981) leads to similar results. It fits directly the Rice and Tracey model for a special choice of free energy. This theory allows a correct prediction of plastic instabilities and can be used to predict ductile fracture by means of local approaches (Rousselier et al., 1985; Rousselier, 1986). Let us mention also some work based on homogenization concepts which, in the case of ductile fracture, gives information about the damage evolution equations (Suquet, 1982; Dragon and Chihab, 1985).

2.6 Brittle Damage and Elastic Behavior. The failure of some brittle materials can be treated simply by considering the coupling between damage and the elastic behavior. Such theories have been developed for concrete by Krajcinovic and Fonseka (1981), Mazars (1985), and Marigo (1985). The possibilities can be illustrated by means of a very simple damage growth equation (Lemaître and Chaboche, 1985), written here for pure tension:

$$dD = \begin{cases} \left(\dfrac{\epsilon}{\epsilon_o}\right)^s d\epsilon & \text{if } \epsilon = \xi \text{ and } d\epsilon = d\xi > 0 \\ 0 & \text{if } \epsilon < \xi \text{ or } d\epsilon < 0 \end{cases} \tag{19}$$

where ϵ_o and s are material constants and ξ is a variable threshold ($d\xi = d\epsilon$ when $\epsilon = \xi$). With the initial conditions $D = \xi = \epsilon = 0$, integrating (19) and using the elastic behavior (equation (11) of Part I) leads to:

$$D = \left(\frac{\epsilon}{\epsilon_R}\right)^{s+1} \qquad \sigma = E\,\epsilon\left[1 - \left(\frac{\epsilon}{\epsilon_R}\right)^{s+1}\right] \tag{20}$$

where $\epsilon_R = [(s+1)\,\epsilon_o^s]^{1/s+1}$ is the rupture strain, when $D = 1$.

Figure 3 illustrates the stress-strain evolution obtained with $s = 2$ and the corresponding load-unload behavior for the damaged elastic material. As shown in Fig. 3(*b*), such a simple model correctly predicts the observed behavior of concrete under pure compression (from Krajcinovic and Fonseka, 1981).

2.6 Damage Anisotropy. The directional nature of damage is clear in many situations. In creep, for instance, microstructural observations have identified two classes of metallic materials (Leckie and Hayhurst, 1984; Hayhurst, 1983): (*a*) for copper, cavitation takes place on grain boundaries more or less perpendicular to the maximum principal stress, and (*b*) for aluminum alloys, grain boundary cavitation is much more isotropically distributed.

Clearly these observations relate to the anisotropy of the damage rate equation. In Continuum Damage Mechanics the problem of anisotropy concerns also the effect of damage on the constitutive behavior. Several levels of theories can then be considered, where the damage variable is considered as a scalar function (Leckie and Onat, 1980), vectors (Krajcinovic and Fonseka, 1981; Krajcinovic and Silva, 1982), or second (Kachanov, 1980) or fourth-order tensors (Chaboche, 1979; Chaboche, 1981). See also Krajcinovic (1984) for a more complete review.

The scalar measure of damage anisotropy is defined (in the creep range) by considering the grain boundaries orthogonal to the direction n and the fraction of boundaries which are cavitated. Then, to each direction n is associated a scalar valued function $V(n)$. It was shown by Leckie and Onat (1980) that a tensorial decomposition of such elementary damage measure leads to even-order tensors.

However, if one considers the damage as produced by small parallel cracks, it is possible to associate a vector to each crack direction (Krajcinovic and Fonseka, 1981). Such a theory, similar to the slip theory of plasticity, can be developed with a similar thermodynamic framework as the one presented in Part I, introducing, for instance, a vector as the dual variable to the present damage variable (instead of the scalar Y or of the fourth-order tensor $\mathbf{Y}$ defined by Lemaître and Chaboche, 1985). This kind of theory is very attractive and gives the direct distinction between active damage (open microcracks) and passive damage (closed microcracks). The main difficulty is the number of independent systems which have to be considered in a general case, but several useful applications have been obtained for concrete under special loading cases (Krajcinovic and Fonseka, 1981).

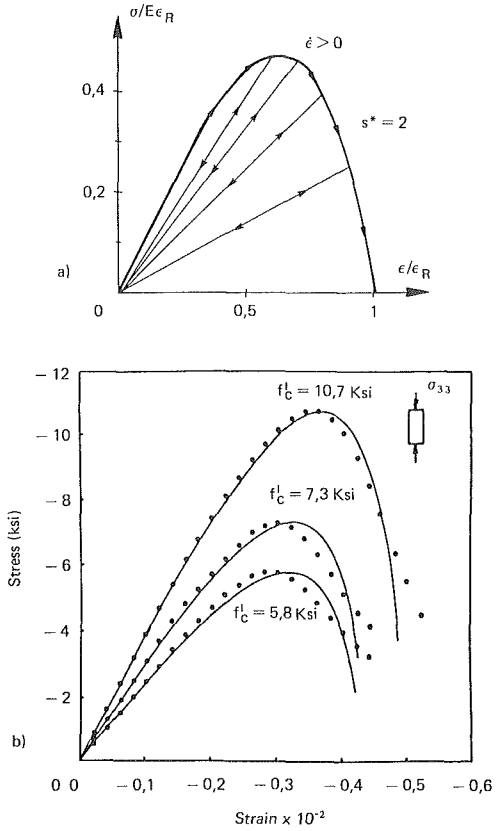


Fig. 3 *(a)* **Stress-strain curve with the elastic behavior coupled with damage; (*b*) example of concrete under compression (from Krajcinovic and Fonseka, 1981)**

Another method of representing the actual state of damage is to use a second-order symmetric damage tensor (Kachanov, 1980) as suggested by the relation (4) of Part I. The theory developed by Murakami and Ohno (1980) introduces the net stress concept in the damage growth equation and an effective stress tensor, through a fourth-order transformation, in the constitutive equation. The anisotropy of the damage growth equations is a linear combination of the isotropic case and a pure anisotropic one.

The damage state can be represented directly as a nonsymmetrical fourth-order tensor, in the framework of the effective stress concept for the constitutive behavior (Chaboche, 1979). Damage measurement follows directly from the methods mentioned in Section 2.4 of Part I. As in the previous theory the damage growth equation uses a linear combination of the isotropic case and a simple anisotropic one. Systematic comparisons between these two theories have shown their similarities and their capabilities to describe actual anisotropic damage states (Murakami and Imaizumi, 1982; Chaboche, 1984).

3 Life Prediction in Structures

3.1 Prediction of Crack Initiation. The life calculations of structures at high temperature incorporate two aspects treated independently or successively: (*a*) the macrocrack initiation is the important aspect for the design methodologies, and (*b*) the crack propagation prediction is used for design and maintenance through the concept of "Damage Tolerance."

Relative to the first step, the *classical approach* considers successively the calculation of the stress-strain state at each point of the structure and the prediction of crack initiation from the obtained stress or strain fields. In the case of Low-Cycle-Fatigue such calculations involve the full cyclic inelastic

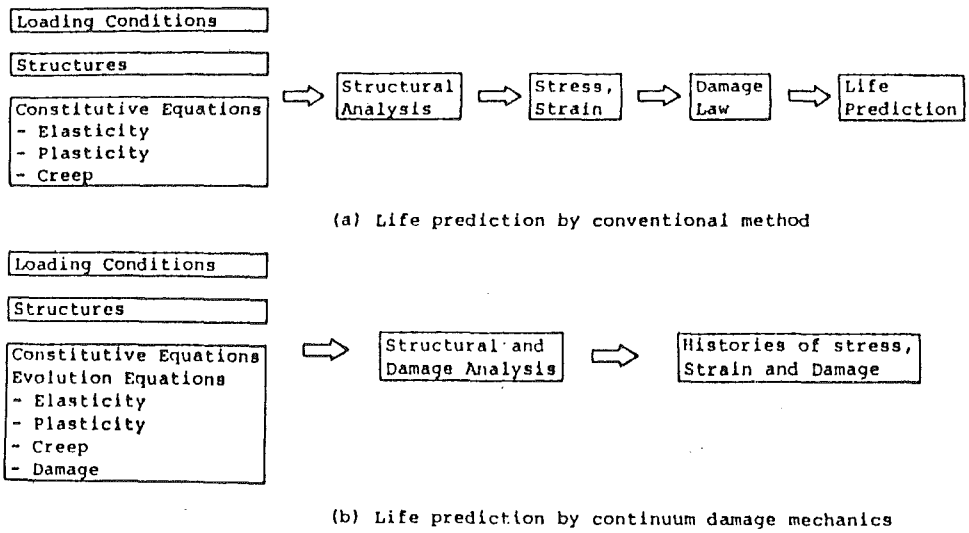


Fig. 4 Comparison between life predictions by conventional method (*a*) and that of continuum damage mechanics (*b*)

analysis to obtain relevant cyclic stabilized values. The difficulties associated with the stabilization of the solution were underlined elsewhere (Cailletaud and Chaboche, 1986).

One of the problems encountered in practical applications is the problem of fatigue damage nucleation and growth from initial stress concentrations. The macroscopic stress/strain gradients play an important role as compared to the microlevel, and the total volume of highly stressed material could be different. This difficulty is generally overcome by one of the two following ways: (*i*) taking into account the macro and micro-plasticity which induce stress redistribution, and (*ii*) introducing a critical distance, depending on the material.

Future approaches will consider the coupling phenomena between the progressive deterioration of the material and its stress-strain behavior, in the framework of Continuum Damage Mechanics, as shown schematically in Fig. 4, taken from Marquis et al. (1981). The existence of a stabilized cycle in the structure is no longer possible and the physical processes at each point of the structure have to be followed by the mathematical model during the whole life, that is for thousands of cycles.

Application of these concepts in fatigue, or more specifically in High-Temperature-Low-Cycle-Fatigue with hold periods, is not possible at the present time for actual structures, even with the most powerful computers. However, the increase of computer capacity and rapidity, the decrease of computing time cost, and the research of special algorithms to perform "cycle jumps" could allow some practical possibilities. Methods have been examined by Savalle and Culié (1978) and applied to the case of the tension-compression specimens.

In the creep situation, on the contrary, the coupling effect has been already considered in several applications (Marquis et al., 1981; Hayhurst, 1983) and permits the improvements of life predictions.

3.2 Crack Propagation by Fracture Mechanics Concepts. The classical approach to predict crack growth and fracture in structural components is based on Fracture Mechanics concepts. Crack growth laws and failure criteria are based on global parameters such as the stress intensity factor K, the elastic strain energy release rate G, the contour integral J. In many cases Linear Elastic Fracture Mechanics constitutes a practical tool to predict crack growth, especially in a situation of non-dissipative "small scale yielding" materials. There is no doubt about its utility and worthiness in the past and in the future.

However, in some situations this global approach presents some difficulties, even inconsistencies, especially when material nonlinearities play an important role, for example in ductile fracture or creep crack growth. Three developments for high temperature problems are briefly discussed below.

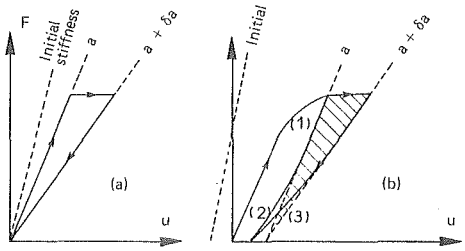
3.2.1 Creep-Fatigue Interaction for Brittle Materials. Analysis of crack propagation in the framework of linear fracture mechanics is possible under high temperature fatigue, even under conditions of creep-fatigue interaction, at least for materials with a low ductility. The case of IN 100, treated by Policella and Lesne (1986), is very representative. Under constant temperature conditions the following results have been obtained:

(1) The stress intensity factor K, calculated elastically or measured through the elastic strain energy release rate G (change in the measured stiffness of the specimen) is a useful parameter to correlate crack growth.

(2) Pure fatigue crack propagation (5 Hz frequency) is well correlated by a simple power law.

(3) Pure creep crack growth is described by a similar equation, using time instead of cycle.

(4) Fatigue with creep hold periods can be predicted from the addition of the two preceding contributions (with time integration of the creep term).



(1) → Plastic flow during loading.
(2) → Actual unloadings before and after crack growth.
(3) → Fictitious elastic unloadings.

Fig. 5 Schematics of the elastic energy release rate: (*a*) elastic brittle solid; (*b*) elasto-plastic brittle solid

In the present case, the material has a low ductility but, due to the high temperature (1000°C) there are viscoplastic strains in the whole specimen. Linear Fracture Mechanics continues to apply because viscoplastic strain concentrates more at the crack tip. The concept of an elastic energy release, necessary to propagate the crack is justified by comparing the stiffness before and after the crack increment (Lemaître and Chaboche, 1985). In such an extension of Linear Fracture Mechanics to creeping materials, the stiffness is considered during an infinitely rapid unloading, inducing elastic changes in the whole structure (only compressive flow at the crack tip); by this way the conditions appear similar to that of "small scale yielding." Figure 5 illustrates schematically the definition of the elastic strain energy release (assuming here that the crack propagates at the maximum load). Let us note the consistency of using an "elastic" global parameter such as this with general thermodynamic concepts (Germain et al., 1983).

3.2.2 Application Under Thermal Fatigue. In thermal fatigue, the main loads applied to the structure are thermal ones. Linear Fracture Mechanics can still be used, defining properly the elastic energy release rate (De Lorenzi, 1982), especially for materials with a low ductility.

However, direct application of Linear Elastic Fracture Mechanics is incorrect. It must be applied *after* a preliminary elasto-plastic analysis (or elasto-viscoplastic) of the structure, without crack, in order to correctly take into account the global stress redistributions (external loading redistribution) induced by inelastic flow.

The example of the plate submitted to cyclic thermal gradients (Lesne and Chaboche, 1984), without any external mechanical loads, illustrates clearly the point that, due to compressive stresses in the heated region, where the crack propagates, the stress intensity factor would be negative if calculated directly with the elastic fields.

The procedure detailed by Policella and Lesne (1986) leads to a possible predictive way, using Linear Fracture Mechanics for a crack which has grown in the initially plastified structure (initial plastic strain field is obtained through the complete inelastic analysis of the structure without the crack). In this way, the crack tip plasticity is not included in the analysis, which clearly corresponds to the "small scale yielding" assumption and justifies the use of some crack growth equations determined with similar conditions (Policella and Lesne, 1986).

3.2.3 Application to Ductile Materials. For more ductile materials, difficulties are encountered in the application of the Linear Fracture Mechanics; there is no longer a one-to-one correlation between K and the crack growth rate, especially under creep conditions. Nonlinear Fracture Mechanics introduces different global parameters, such as J in ductile rup-

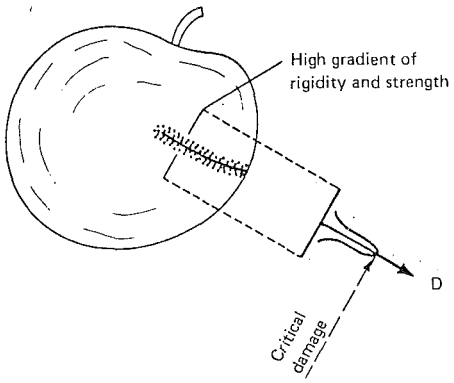


Fig. 6 Crack as a damaged zone

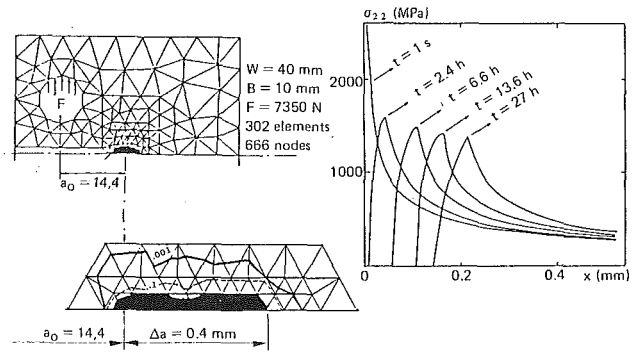


Fig. 7 Creep crack growth prediction by a local approach for a CT specimen in INCO 718 at 600°C

ture, ΔJ in fatigue (Low-Cycle-Fatigue), C^* in creep. These parameters are justified only for special inelastic constitutive behavior and the correlations are not significantly better (Bensussan et al., 1985).

An alternative global approach can be developed to treat ductile materials; continue to use a linear elastic fracture parameter such as the stress intensity factor (or the elastic energy release rate), extending the concept of "small scale yielding" for fictitious elastic unloadings, as shown above (Lemaître and Chaboche, 1985). This variable can be considered as the thermodynamic force associated to the crack length (Germain et al., 1983). In other words, its definition has nothing to do with the actual behavior of the material; K (or G) plays the role of the stress in classical plasticity. The influence of nonlinear behavior, of history effects, and of material processes at the crack tip will then be specified through the use of additional internal variables (as variable threshold parameters used by Pellas et al., 1977; Baudin and Robert, 1984) and additional growth equations:

$$\begin{aligned} da &= f(K, K_s, \ldots)dt \\ dK_s &= g(K, K_s, \ldots)dt \end{aligned} \tag{21}$$

In the case of fatigue crack growth it was shown by Lemaître and Chaboche (1985) that the dissipation is proportional to the effective stress intensity $K_{\text{eff}} = K_{\max} - K_s$, where K_s corresponds to the crack opening, chosen here as a growth rate threshold.

Let us remark that the present approach, considering only the linear elastic energy release rate, is a theoretical concept. It does not suppose the real measurement of the compliance, but considers simply the stiffness change induced by the geometry change when the crack grows (neglecting the possible influence of the overall inelastic deformation on the compliance). In other words, the elastic unloadings have to be considered as fictitious, allowing to define the same driving force as in the commonly accepted small scale yielding. Moreover, the crack closure effects, during tension or compression unloading periods, are taken into account by the phenomenological parameter K_s in equation (21).

3.3 Local Approaches to Fracture. As mentioned above some difficulties are encountered in applying global concepts of Fracture Mechanics to some complex situations including short cracks, history effects (overloads, warm-prestress, etc.), ductile fracture, and creep crack growth.

Even if some correlations or parametrizations have been found satisfactory in many cases, alternative methods, called "local approaches," have been under development for several years (Janson and Hult, 1977). They consider the actual behavior at the crack tip, trying to calculate as precisely as possible the local stress and strain fields and the corresponding deterioration. Local failure criterion allows a crack increment to be predicted, and eventually the crack instability.

In fact, two levels can be considered for local approaches:

- Numerical methods using discrete crack increments: One uses the techniques of node release to produce the crack growth when some critical value of a physical quantity is reached at some critical distance ahead of the crack tip (Newman, 1974; Hinnerirhs, 1980; Anquez, 1983; Devaux and Mottet, 1984). This physical quantity can be an equivalent stress or strain, energy, or some measure of local damage. In such situations the prediction of crack growth using this method will be mesh dependent (Newman, 1974), but the local mesh size is fixed by means of statistical considerations (Rousselier, 1986; Mudry, 1986). Simplified approaches can be developed, using analytical stress-strain fields near the crack tip (Maas and Pineau, 1985).
- Use of Continuum Damage Mechanics, including the progressive deterioration and the corresponding stiffness reduction. The crack (damaged zone) is then taken as the locus of material points with no rigidity, where damage has reached its critical value (Fig. 6). Conceptually, due to the stress redistribution associated with viscoplasticity and strength decrease, there is no need to introduce critical distance or node release techniques for crack growth simulation. After some maximum, the stresses in the plastic zone decrease continuously as damage increases, approaching the crack tip. This approach was studied first by Hayhurst et al. (1975) in the creep domain and is now under a large development in ductile fracture (Rousselier et al., 1985; Rousselier, 1986), creep crack growth (Saanouni et al., 1985; Hayhurst, 1985), fatigue (Ben Allal et al., 1984), and rupture of concrete (Mazars, 1985).

Among many others (Lemaître, 1985), several applications of local approaches have been done:

- From a theoretical point of view, the use of CDM with a discontinuous damage growth law (sudden change from 0 to 1) solves Rice's paradox, giving consistency to the energy dissipated at the crack tip. Moreover, such simplified damage laws allow analytical solutions for simple structures (Bui and Ehrlacher, 1985; Bui et al., 1986).
- The application of the first level of local approaches to fatigue can be considered now as classical (Newman, 1974; Anquez, 1983). For example, it gives correct predictions of crack opening or crack closure, even under complex loading conditions (multi levels, overloads, underloads, etc.).
- Fracture of brittle materials such as concrete has been simulated in various ways (Mazars, 1985; Bazant and Pijaudier-Cabot, 1987) and for several kinds of components, using continuum damage concepts and taking into account the coupling between damage and the elastic behavior, as mentioned in Section 2.6.
- In the case of ductile fracture two levels have been applied (Devaux and Mottet, 1984; Rousselier et al., 1985; Rousselier,

1986), introducing or not the notion of damage. The local mesh size is fixed through "critical distance" considerations or a statistics of defects in the volume element. Good predictions have been obtained with the two methods, comparing notched axisymmetric specimens, cracked specimens, and compact tension specimens.

• Creep crack growth has also been predicted using local approaches of both types (Saanouni et al., 1985; Hayhurst, 1985; Walker and Wilson, 1984; Tvergaard, 1986). In the case of using CDM concepts the main difficulties are associated with the time integration of the very stiff differential equations (Walker and Wilson, 1984). Special automatic time stepping has been developed for points approaching the rupture condition ($D = 1$). Figure 7 shows an application to the creep crack growth in a C.T. specimen in Inconel 718 at 600°C (Saanouni et al., 1985).

Such local approaches of fracture are very attractive, due to the possibility of implementing more physical rupture criteria and damage processes at the crack tip. They have already given good results in laboratory simulations and are beginning to be applied at an industrial level, especially in the case of cleavage and ductile fracture (in Devaux et al., 1986, for instance). Examples where local approaches have a benefit over classical fracture mechanics have been presented by Rousselier et al. (1985), Bensussan et al. (1985), Mudry (1985), and Devaux et al. (1986).

Several problems delay a more general use:

• *The cost of calculations.* However, the difference isn't so large if an inelastic analysis has also to be performed to apply Fracture Mechanics. Simplified procedures, using some analytical fields, could offer an intermediate, more practical, way.

• *The dependence on the finite element modeling, which seems a common feature of the various local approaches.* Both the "crack width", the crack growth rate, and the failure load depend on the chosen mesh size, even when using the complete elastic-plastic-damage coupling as by Saanouni et al. (1985). From a practical point of view the local mesh size has to be fixed in every application after checking a particular one. Another attractive way may be to introduce a nonlocal definition for damage growth, as already considered by Rousselier et al. (1985), Bazant and Pijaudier-Cabot (1987), Billardon (1986), and Saanouni and Lesne (1987).

4 Conclusion

Some practical damage rate equations have been briefly presented and discussed, in the framework of Continuum Damage Mechanics. Applications were shown on various kinds of materials, concerning: creep processes, fatigue processes, creep-fatigue interaction, ductile plastic damage of metals, and brittle damage, especially in concrete. In each case, uniaxial and multiaxial equations have been presented. Such damage growth theories can be used to predict crack initiation in structural components, as shown recently in various applications. Moreover, the use of CDM allows some interesting additional possibilities in the framework of local approaches to fracture.

This rational approach of damage proceeds as an extension of Continuum Mechanics. It allows many possibilities, introducing, for instance, some conncections between microstructure measurements and mechanical parameters. Moreover, this theory finds applications in many different situations (various loading conditions, various materials, various physical processes, etc.).

In that it concerns life predictions in structures, Damage Mechanics allows one to take into account the coupling effects between deterioration processes and mechanical behavior. It will certainly be developed further and extensively used in the future as a complementary tool, between Continuum Mechanics and Fracture Mechanics.

References

Anquez, L., 1983, "Elastoplastic Crack Propagation (Fatigue and Failure)," *La Recherche Aérospatiale 1983-2.*

Baudin, G., and Robert, M., 1984, "Crack Growth Life-time Prediction Under Aeronautical Type Loading," *5th European Conf. on Fracture*, Lisbon.

Bazant, Z. P., and Pijaudier-Cabot, G., 1987, "Modeling of Distributed Damage by Non-Local Continuum with Local Strain," *4th Int. Conf. on Numerical Methods in Fracture Mechanics*, San Antonio, Texas, A. R. Luxmoore et al., eds., Pineridge Press, Swansea, pp. 411-432.

Ben Allal, A., Billardon, R., and Marquis, D., 1984, "Prévision de l'Amorçage et de la Propagation des Fissures par la Mécanique de l'Endommagement," Journées Int. de Printemps, "Amorçage des Fissures sous Sollicitations Complexes," Paris.

Bensussan, P., Maas, E., Pelloux, R., and Pineau, A., 1985, "Creep Crack Initiation and Propagation: Fracture Mechanics and Local Approach," *5th Int. Seminar on Inelastic Analysis and Life Prediction in High Temperature Environment*, Paris.

Beziat, J., Diboine, A., Levaillant, C., and Pineau, A., 1983, "Creep Damage in a 316 S.S. Under Triaxial Stresses," *4th Int. Seminar on Inelastic Analysis and Life Prediction in High Temperature Environment*, Chicago, IL.

Billardon, R., 1986, "Fully Coupled Strain and Damage Finite Element Analysis of Ductile Fracture," *Int. Sem. on Local Approaches of Fracture*, Moret-sur-Loing, France.

Bui, H. D., and Ehrlacher, A., 1981, "Propagation of Damage in Elastic and Plastic Solids," *ICF 6*, Vol. 2, Cannes.

Bui, H. D., Dang Van, K., and De Langre, E., 1986, "A Simplified Analysis of Creep Crack Growth Using Local Approach," *Int. Seminar on Local Approaches of Fracture*, Moret-sur-Loing, France.

Cailletaud, G., and Levaillant, C., 1984, "Creep-Fatigue Life Prediction: What About Initiation?", *Nuclear Engng. and Design*, No. 83, pp. 279-292.

Cailletaud, G., and Chaboche, J. L., 1986, "On the Calculation of Structures in Cyclic Plasticity or Viscoplasticity," *Computers and Structures*, Vol. 23, No. 1, pp. 23-31.

Chaboche, J. L., 1974, "Une Loi Différentielle d'Endommagement de Fatigue avec Cumulation non Linéaire," *Revue Française de Mécanique*, No. 50-51, English translation in "Annales de l'IBTP," HS 39, 1977.

Chaboche, J. L., 1972, "Description Thermodynamique et Phénomènologique de la Viscoplasticité Cyclique avec Endommagement," Thèse Univ. Paris VI, publication ONERA, N 1978-3.

Chaboche, J. L., 1979, "Le Concept de Contrainte Effective Appliqué à l'Élasticité et à la Viscoplasticité en Présence d'un Endommagement Anisotrope," Col. *EUROMECH 115*, Grenoble 1979, Editions du CNRS, 1982.

Chaboche, J. L., 1980, "Lifetime Predictions and Cumulative Damage under High Temperature Conditions," *Int. Symp. on Low Cycle Fatigue and Life Prediction*, Firminy, France, ASTM STP 770, 1982.

Chaboche, J. L., 1981, "Continuous Damage Mechanics. A Tool to Describe Phenomena Before Crack Initiation," *Nuclear Engng. and Design*, Vol. 64, pp. 233-247.

Chaboche, J. L., 1984, "Anisotropic Creep Damage in the Framework of Continuum Damage Mechanics," *Nuclear Engng. and Design*, Vol. 79, pp. 309-319.

Chaboche, J. L., 1988, "Continuum Damage Mechanics: Part I—General Concepts," ASME Journal of Applied Mechanics, Vol. 55, No. 1, pp. 59-64.

Chaudonneret, M., and Chaboche, J. L., 1986, "Fatigue Life Prediction of Notched Specimens," *Int. Conf. on Fatigue of Engineering Materials and Structures*, Sheffield.

Chrzanowski, M., 1976, "Use of the Damage Concept in Describing Creep-fatigue Interaction Under Prescribed Stress," *Int. J. Mech. Sci.*, Vol. 18, pp. 69-73.

Contesti, E., 1986, "Endommagement en Fluage: Expérience et Modélisation," Réunion GRECO Grandes Déformations et Endommagement—GIS Rupture à Chaud, Aussois, France.

Dang Van, K., 1973, "Sur la Résistance à la Fatigue des Métaux," *Sciences et Techniques de l'Armement*, Vol. 47, 3ème Fasc.

Del Puglia, A., and Vitale, E., 1982, "Damage Concept in Creep Fatigue: Current Theories and Applications," Proc. *ASME-PVP Conf.*, Orlando, FL.

Delobelle, P., 1985, "Etude en Contraintes Biaxiales des Lois de Comportement d'un Acier Inoxydable du type 17-21 SPH—Modélisation et Identification—Introduction de l'Endommagement, Cas de l'INCONEL 718," Thèse de Doctorat d'Etat, Besançon.

De Lorenzi, H. G., 1982, "On the Energy Release Rate and the J-Integral for 3-D Crack Configurations," *Int. J. of Fracture*, Vol. 19, pp. 183-193.

Devaux, J. C., and Mottet, G., 1984, "Déchirure Ductile des Aciers Faiblement Alliés: Modèles Numériques," Rapport N 79-057, Framatone.

Devaux, J. C., Saillard, P., and Pelissier-Tanon, A., 1986, "Calcul des Circonstances de Rupture d'une Virole Épaisse Préfissurée Soumise à un Choc Froid à l'Azote Liquide," *Int. Seminar "Local Approaches of Fracture,"* Moret-sur-Loing, France.

Dragon, A., and Chihab, A., 1985, "Quantifying of Ductile Fracture Damage Evolution by Homogenization Approach," *SMIRT 8*, Bruxelles.

Dyson, B. F., 1979, "Constrained Cavity Growth, Its Use in Quantifying Recent Creep Fracture Experiments," *Canadian Met. Quart.*, Vol. 18, p. 31.

Germain, P., Nguyen, Q. S., and Suquet, P., 1983, "Continuum Thermodynamics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 1010–1020.

Greenwood, G., 1973, "Creep Life and Ductility," Int. Congress on Metal, Cambridge, 1973, in *Microstructure and the Design of Alloys*, Vol. 2, p. 91.

Hashin, Z., and Laird, C., 1980, "Cumulative Damage under Two Level Cycling: Some Theoretical Predictions and Test Data," *Fatigue of Engineering Materials and Structures*, Vol. 2, pp. 345–350.

Hayhurst, D. R., 1972, "Creep Rupture under Multi-axial State of Stress," *J. Mech. Phys. Solids*, Vol. 20, No. 6, pp. 381–390.

Hayhurst, D. R., Dimmer, P. R., and Chernuka, M. W., 1975, "Estimates of the Creep Rupture Lifetime of Structures Using the Finite Element Method," *J. Mech. Phys. of Solids*, Vol. 23, p. 335.

Hayhurst, D. R., 1983, "On the Role of Creep Continuum Damage in Structural Mechanics," *Engineering Approaches to High Temperature Design*. Wilshire, Owen, eds., Pineridge Press, Swansea.

Hayhurst, D. R., 1985, "On the Use of Creep Damage Constitutive Equations to Predict the Behavior of Cracked Specimens Under Constant Applied Loads," *5th Int. Seminar on Inelastic Analysis and Life Prediction in High Temperature Environment*, Paris.

Hinnerirhs, T. D., 1980, "Viscoplastic and Creep Crack Growth Analysis by the Finite Element Method," Ph.D. Thesis, Air Force Institute of Technology.

Hua, C. T., and Socie, D. F., 1984, "Fatigue Damage in the 1045 Steel under Constant Amplitude Biaxial Loading," *Fatigue Eng. Mater. Struc.*, Vol. 7, No. 3, pp. 165–179.

Hult, J., 1979, "Continuum Damage Mechanics—Capabilities, Limitations and Promises," *Mechanisms of Deformation and Fracture*, Pergamon, Oxford, pp. 233–247.

Janson, J., and Hult, J., 1977, "Fracture Mechanics and Damage Mechanics a Combined Approach," *J. Mécanique Appliquée*, Vol. 1, No. 1.

Kachanov, L. M., 1958, "Time of the Rupture Process under Creep Conditions," *Izv. Akad. Nauk*, SSR, Otd Tekh. Naul, No. 8, pp. 26–31.

Kachanov, M., 1980, "Continuum Model of Medium with Cracks," ASCE *J. of the Engineering Mechanics Division*, Vol. 106, No. EM5, p. 1039–1051.

Krajcinovic, D., 1984, "Continuum Damage Mechanics, *Applied Mechanics Reviews*, Vol. 37, No. 1.

Krajcinovic, D., and Fonseka, G. U., 1981, "The Continuous Damage Theory of Brittle Materials," Parts 1 and 2, ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 809–824.

Krajcinovic, D., and Silva, M. A. G., 1982, "Statistical Aspects of the Continuous Damage Theory," *Int. J. Solids Structures*, Vol. 18, No. 7, pp. 551–562.

Leckie, F. A., and Hayhurst, D. R., 1974, "Creep Rupture of Structures," *Proc. Royal Soc.*, London, Vol. 340, pp. 323–347.

Leckie, F. A., and Onat, E. T., 1980, "Tensorial Nature of Damage Measuring Internal Variables," *IUTAM Symp. on Physical Non-linearities in Structural Analysis*, Senlis, France.

Lemaître, J., 1984, "How to Use Damage Mechanics," *Nuclear Engng. and Design*, Vol. 80, pp. 233–245.

Lemaître, J., 1985, "A Continuum Damage Mechanics Model for Ductile Fracture," ASME *J. Engng. Mat. and Technology*, Vol. 107, pp. 83–89.

Lemaître, J., 1985, "Local Approach of Fracture," *Symp. on Mechanics of Damage and Fatigue*, Haifa, Israel.

Lemaître, J., and Chaboche, J. L., 1978, "Aspect Phénomènologique de la Rupture par Endommagement," *J. de Mécanique Appliquée*, Vol. 2, No. 3, pp. 317–365.

Lemaître, J., and Chaboche, J. L., 1985, *"Mécanique des Matériaux Solides,"* Dunod, Paris.

Lesne, P. M., and Chaboche, J. L., 1984, "Prediction of Crack Initiation Under Thermal Fatigue and Creep," *2nd Int. Conf. on Fatigue Thresholds. Fatigue 84*, Birmingham.

Maas, E., and Pineau, A., 1985, "Creep Crack Growth Behavior of type 316 L Steel," *Engng. Fracture Mechanics*, Vol. 22, No. 2, pp. 307–325.

McClintock, F., 1968, "A Criterion for Ductile Fracture by the Growth of Holes," ASME JOURNAL OF APPLIED MECHANICS.

Manson, S. S., 1979, "Some Useful Concepts for the Designer in Treating Cumulative Damage at Elevated Temperature," *I.C.M. 3*, Cambridge, Vol. 1, pp. 13–45.

Marco, S. M., and Starkey, W. L., 1954, "A Concept of Fatigue Damage," *Trans. ASME*, Vol. 76, No. 4, pp. 627–632.

Marigo, J. J., 1985, "Modelling of Brittle and Fatigue Damage for Elastic Material by Growth of Microvoids," *Engng. Fracture Mechanics*, Vol. 21, No. 4, pp. 861–874.

Marquis, D., Billardon, R., Benallal, A., and Grolade, D., 1981, "Prediction of Crack Initiation and Growth in Structures by Continuum Damage Theory," EUROMECH Colloquium 147, *Damage Mechanics*, Cachan, France.

Mazars, J., 1985, "A Description of Micro and Macro Scale Damage of Concrete Structures," *Symp. on Mechanics of Damage and Fatigue*, Haifa, Israel.

Miller, K. J., and Zachariah, K. P., 1977, "Cumulative Damage Laws for Fatigue Crack Initiation and Stage 1 Propagation," *Journal of Strain Analysis*, Vol. 12, No. 4, pp. 262, 270.

Mudry, F., 1986, "A Local Approach to Cleavage Fracture," *Int. Seminar on Local Approaches of Fracture*, Moret-sur-Loing, France.

Murakami, S., and Ohno, N., 1980, "A Continuum Theory of Creep and Creep Damage," *3rd IUTAM. Symp. on Creep in Structures*, Leicester.

Murakami, S., and Imaizumi, T., 1982, "Mechanical Description of Creep Damage Stage and its Experimental Verification," *J. de Mécanique Théorique et Appliquée*, Vol. 1, No. 5, pp. 743–761.

Newman, Jr., J. C., 1974, "Finite Element Analysis of Fatigue Crack Propagation Including the Effects of Crack Closure," Ph.D. Thesis, V.P.I., Blacksburg, VA.

Pellas, J., Baudin, G., and Robert, M., 1977, "Fatigue Crack Growth Model Prediction with Two Differential Constitutive Equations," *ICF4 Proceedings*, Waterloo, pp. 1353–1360.

Plumtree, A., and Lemaitre, J., 1979, "Application of Damage Concepts to Predict Creep-Fatigue Failures," *ASME. Pressure Vessels and Piping Conf.*, Montreal.

Policella, H., Paulmier, P., and Savalle, S., 1982, "Effet de la Contrainte Moyenne en Fatigue Vibratoire à Haute Température. Résultats Expérimentaux sur IN 100 à 900°C," R. T. ONERA, No. 49/1765.

Policella, H., and Lesne, P. M., 1986, "Amorçage et Propagation de Fissures en Fatigue Thermique. Calculs de Prévision," *International Spring Meeting of Fatigue at High Temperature*, Paris.

Rabotnov, Y. N., 1969, "*Creep Problem in Structural Members*," North-Holland.

Rezai-Aria, F., and Remy, L., 1986, "Prédiction de Durée de Vie en Fatigue Thermique dans un Alliage à Base de Cobalt," *Int. Spring Meeting of Fatigue at High Temperature*, Paris.

Rice, J. R., and Tracey, D. M., 1969, "On the Ductile Enlargement of Voids in Triaxial Stress Fields," *J. Mech. Phys. Solids*, Vol. 17, p. 201.

Rousselier, G., 1980, "Finite Deformation Constitutive Relations Including Ductile Fracture Damage," *IUTAM Symp. on Three-Dimensional Constitutive Relations and Ductile Fracture*, Dourdan, Nemat-Nasser, eds., North-Holland Publ. Comp., 1981, pp. 331–355.

Rousselier, G., 1986, "Les Modèles de Rupture Ductile et Leurs Possibilités Actuelles dans le Cadre de l'Approche Locale de la Rupture," *Séminaire International Approches Locales de la Rupture*, Moret-sur-Loing, France.

Rousselier, G., Devaux, J. C., and Mottet, G., 1985, "Ductile Initiation and Crack Growth in Tensile Specimens—Application of Continuum Damage Mechanics," *SMIRT 8*, Brussels.

Saanouni, K., Chaboche, J. L., and Bathias, C., 1985, "On the Creep Crack Growth Prediction by a Local Approach," *IUTAM Symp. on Mechanics of Damage and Fatigue, Haifa*, also in *J. of Engineering Fracture Mechanics*, Vol. 25, No. 5/6, 1986, pp. 677–691.

Saanouni, K., and Lesne, P. M., 1987, "Non-Local Damage Model for Creep Crack Growth Prediction," *Int. Sem. "High Temperature Fracture Mechanisms and Mechanics*, Dourdan, France.

Savalle, S., and Culie, J. P., 1978, "Méthodes de Calcul Associées aux Lois de Comportement Cyclique et d'Endommagement," *La Recherche Aérospatiale*, No. 1978-5.

Socie, D. F., Fash, J. W., and Leckie, F. A., 1983, "A Continuum Damage Model for Fatigue Analysis of Cast Iron," *ASME Conf. in Life Prediction*, Albany, New York.

Subramanyan, S., 1976, "A Cumulative Damage Rule Based on the Knee Point of the S-N Curve," *ASME J. of Engng. Mat. and Technology*, pp. 316–321.

Suquet, P., 1982, "Plasticité et Homogénéisation," Thèse Doctorat d'Etat, Univ. Paris VI.

Tvergaard, V., 1986, "Analysis of Creep Crack Growth by Grain Boundary Cavitation," *Int. J. of Fracture*, Vol. 31, pp. 183–209.

Walker, K. P., and Wilson, D. A., 1984, "Constitutive Modelling of Engine Materials," PWA Report FR 17911 (AFWAL-TR-84-4073).

Plastic Analysis of Ice Contact Problems

R. P. Nordgren
Shell Development Co.,
Houston, TX 77001
Fellow ASME

The plane problem of a rigid stamp on a truncated wedge of rigid/perfectly plastic material is solved by the method of characteristics for a generalized von Mises yield function. Quadratic dependence of shear strength on mean normal stress is included in the yield function. For the corresponding three-dimensional contact problem upper bounds on contact force are obtained by limit analysis. Both crushing and splitting failure modes are treated. Numerical results are presented in dimensionless form for application to the prediction of contact stress of ice on arctic structures with sloping sides.

Introduction

The interaction of floating ice features with offshore structures is of current interest in arctic engineering. Arctic structures must be designed to resist local contact pressure as well as total force from the impinging ice. Structures with both vertical and sloping sides are of interest. The ice features of most concern are multi-year floes with embedded pressure ridges.

Ice is an extremely complex material whose mechanical behavior exhibits anisotropy, inhomogeneity, rate dependency, temperature dependency and scale effects. Certain aspects of ice behavior can be studied with a plasticity model in which shear strength depends on mean normal stress in a manner similar to plasticity models for soils and rocks. Upon idealizing ice as a perfectly plastic material, several contact problems have been treated by the methods of limit analysis by Ralston (1978), Reinecke and Remer (1979), Reinecke (1980), and others. The results of these investigations are directly applicable to the prediction of contact pressures of ice on arctic structures with *vertical* sides.

The present paper is directed toward predicting ice contact pressures on arctic structures with *sloping* sides, i.e., on cones and pyramids. The ice is idealized as an isotropic, rigid/perfectly plastic material with shear strength quadratically dependent on mean normal stress. As in previous works the structure is assumed to be rigid.

Contact problems have received a great deal of attention in the applied mechanics literature. An introduction to the subject is provided in the recent book by Johnson (1985) and a number of plastic contact problems are treated in Chen's (1975) book. References to the literature are given in these books.

We first consider the plane contact problem of a rigid indenter acting on a rigid/perfectly plastic truncated wedge of ice. Solutions for the stress field in the plastic region are obtained by the classical method of characteristics for both plane stress and plane strain. A discontinuous solution is required in certain cases. These solutions complement solutions by Prandtl (1921) and Prager (1948) for the Tresca yield function and by Shield (1954) for the Mohr-Coulomb yield function.

Next we consider the three-dimensional problem of a truncated wedge of ice in contact with a rigid indenter over a rectangular region. An upper bound on the contact force is determined from the upper-bound theorem of limit analysis for a localized crushing mode of failure. For an indenter of infinite length the upper bound is in good agreement with the rigid/plastic solution for plane strain over a broad range of wedge angles. The upper bound on contact force also provides an estimate of average contact pressure.

In addition to the crushing mode of failure, the ice feature may fail by splitting across its thickness. This failure mode also is treated here by the upper-bound theorem of limit analysis. The splitting force bound is lower than the crushing force bound for sufficiently high ratios of indenter width to ice thickness in many cases. Other modes of failure for ice features also are possible. For example, bending failure of a pressure ridge riding up a conical structure has been analyzed by Winkler and Nordgren (1986).

Numerical results for the plane and three-dimensional contact problems are presented in dimensionless form to illustrate the effects of various parameters of interest in the ice contact problem. Contact pressure is ratioed on uniaxial compressive strength which can be determined from a standard unconfined compression test. Compression tests under confining pressure are required to determine the other two strength parameters in the proposed quadratic yield function. Such tests have been carried out by Jones (1978, 1982) on polycrystalline freshwater ice and by Cox et al. (1985) on sea ice.

Numerical results for the plane problem indicate that dependency of contact pressure on ice-structure contact angle, coefficient of friction, compressive strength, and the parameters which govern the dependency of shear strength on mean normal stress. Contact pressures are considerably lower for plane stress than for plane strain in general. This fact is technically significant in connection with lateral confinement

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 21, 1986; final revision September 25, 1987.

conditions and possible cracking of the ice feature across its thickness during contact with the structure.

In the three-dimensional contact problem the effect of the aspect ratio of the contact area is significant. The splitting mode of failure also depends strongly on the thickness of the ice feature. Discussion of the applicability of the numerical resuls to ice contact problems is given at the conclusion of this paper.

Theory of Plasticity

The theory of plasticity is developed in various texts including Hill (1950), Prager and Hodge (1951), Koiter (1960), and Geiringer (1973). In this section the basic equations of the theory are recorded and a quadratic yield function is introduced. We then review some results needed for limit analysis.

Basic Equations. In the usual Cartesian tensor notation, σ_{ij} is the stress tensor and e_{ij} is the strain tensor. In the theory of plasticity the strain tensor is the sum of an elastic part e'_{ij} and a plastic part e''_{ij}. The elastic strain is given by generalized Hooke's law for an elastic/plastic material or it is zero for a rigid/plastic material. For a perfectly plastic material a yield function $f(\sigma_{ij})$ is introduced and the plastic strain rate is given by the flow rule

$$\mathring{e}''_{ij} = \lambda \frac{\partial f}{\partial \sigma_{ij}}, \tag{1}$$

where

$$\lambda = 0 \quad \text{if } f<0 \text{ or if } f=0 \text{ and } \mathring{f}<0,$$
$$\lambda \geq 0 \quad \text{if } f=0 \text{ and } \mathring{f}=0.$$

Further considerations are necessary for singular yield functions which are not used here.

We consider an isotropic, elastic/perfectly plastic material with the quadratic yield function proposed by Reinicke and Remer (1978) in the form

$$f = aJ_2 + bI_1 + cI_1^2 - 1, \quad a>0, \quad b>0, \quad c>0, \tag{2}$$

where a, b, c are material constants and the stress invariants are defined by

$$J_2 = \tfrac{1}{2} s_{ij} s_{ij}, \quad I_1 = \sigma_{kk}, \tag{3}$$

with the deviatoric stress tensor s_{ij} defined as

$$s_{ij} = \sigma_{ij} - \tfrac{1}{3}\delta_{ij}\sigma_{kk}. \tag{4}$$

The yield condition ($f=0$) from equation (2) represents an ellipse in the I_1, $J_2^{1/2}$ plane. The quadratic dependence of the yield function on I_1 with $c>0$ accounts for the nonlinear increase and peak in shear strength with confining stress observed in triaxial tests by Jones (1978, 1982) on polycrystalline fresh-water ice and by Cox et al. (1985) on sea ice. Typical data of Jones is shown in Fig. 1 along with the fitted yield function (2) for this case.

In the limiting case $c=0$, equation (2) corresponds to the yield function proposed by Murrell (1963) for rocks. For $b = c = 0$, equation (2) reduces to the classical von Mises yield function for metals. For $c < 0$, the yield condition ($f=0$) with equation (2) represents a hyperbola in the I_1, $J_2^{1/2}$ plane. This case may be of interest for other materials but will not be considered here. Further, the yield function (2) can be made equivalent to the linear yield function in I_1 and $J_2^{1/2}$ proposed by Drucker and Prager (1952), but this again requires $c<0$ and will not be considered here.

Limit Analysis. In this paper we will apply the upper-bound collapse theorem of limit analysis which has been reviewed by Koiter (1960). This theorem states that an elastic/perfectly plastic body cannot support a given system of loads if an admissible collapse mechanism $\mathring{e}^*_{ij}$ exists for which

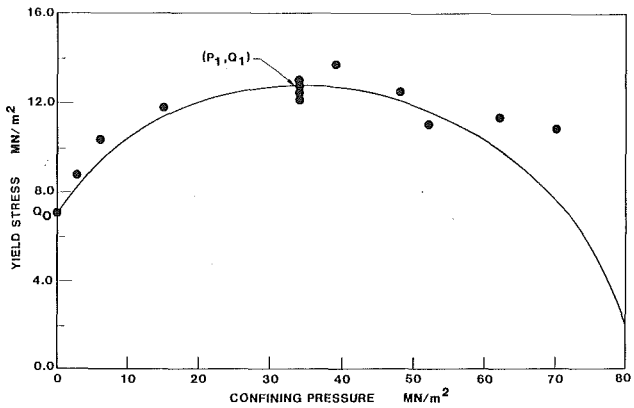


Fig. 1 Quadratic yield function (2) fit to Jones' (1982) triaxial test data for polycrystalline fresh-water ice at a strain rate of 5.4 × 10^{-4} s^{-1}

the work rate of the given loads exceeds the rate of plastic energy dissipation in the body, i.e., for which (in the absence of body force)

$$\int_{S_\sigma} \sigma_i v_i^* dS > \int_V F(\mathring{e}^*) dV + \int_{S_D} F_s(\delta \mathring{e}^*) dS, \tag{5}$$

where σ_i is the given stress vector on the external surface S_σ of the body, S_D is an internal surface of velocity discontinuity, and the dissipation functions F and F_s are defined by

$$F(\mathring{e}^*) = \sigma_{ij}(\mathring{e}^*)\mathring{e}^*_{ij}, \quad F_s(\delta \mathring{e}^*) = \sigma_{ij}(\delta \mathring{e}^*)\delta \mathring{e}^*_{ij}. \tag{6}$$

To be admissible e^*_{ij} and δe^*_{ij} must satisfy the strain rate-velocity relations

$$\mathring{e}^*_{ij} = \tfrac{1}{2}(v^*_{i,j} + v^*_{j,i}), \quad \delta \mathring{e}^*_{ij} = \tfrac{1}{2}(\delta v_i^* n_j + \delta v_j^* n_i), \tag{7}$$

where v_i^* is the velocity field (in the collapse mechanism) and δv_i^* is the velocity discontinuity on S_D with unit normal n_i. The second relation (7) follows from work-energy considerations across S_D. Also, $\mathring{e}^*_{ij}$ and $\delta \mathring{e}^*_{ij}$ must satisfy any restrictions imposed by the flow rule (1), e.g., $\mathring{e}^*_{ij} = 0$ for the von Mises yield function. Further, F and F_s must be positive definite. In the Appendix we consider these requirements for yield functions of the form $f(I_1, J_2)$ of which equation (2) is a special case. For the quadratic yield function (2) we find that the dissipation functions (6) can be written in terms of strain rate as

$$F(\mathring{e}) = \frac{1}{6c}\left\{(b^2+4c)^{1/2}\left[18\frac{c}{a}\,\mathring{\mathcal{E}}_{ij}\mathring{\mathcal{E}}_{ij} + (\mathring{e}_{kk})^2\right]^{1/2} - b\mathring{e}_{kk}\right\}, \tag{8}$$

$$F_s(\delta \mathring{e}) = \frac{1}{6c}\left\{(b^2+4c)^{1/2}\left[\delta v_n^2 + 3\frac{c}{a}\,(3\delta v_t^2 + 4\delta v_n^2)\right]^{1/2} - b\delta v_n\right\}, \tag{9}$$

where the deviatoric strain tensor is defined by

$$\mathcal{E}_{ij} = e_{ij} - \tfrac{1}{3}\delta_{ij} e_{kk}, \tag{10}$$

and δv_n and δv_t are magnitudes of the normal and tangential components of δv_i on S_D. The restrictions imposed by the flow rule and the positive definiteness of F and F_s are satisfied provided that

$$-b/2c < I_1 < 0, \quad a > 3b^2, \quad \mathring{e}_{kk} > 0, \quad \delta v_n > 0, \tag{11}$$

as shown in the Appendix. The first inequality on I_1 restricts the stress state to the right-half of the yield ellipse in the I_1, $J_2^{1/2}$ plane. The inequality on a and b is satisfied by a wide margin for ice. The last two inequalities in expression (11) are a conse-

quence of the yield function (2) and the flow rule (1) for the right-half of the yield ellipse. The expressions (8) and (9) for the dissipation functions agree with those given by Reinicke and Remer (1978) without the restrictions (11). It should be noted that further restrictions are necessary to assure the positive definiteness of F and F_s for the case $c < 0$.

Rigid/Plastic Solutions for Plane Wedges

Method of Characteristics. We begin with a brief review of the method of characteristics for the solution of plane problems for a rigid/perfectly plastic material. The formulation essentially follows that given by Geiringer (1973) where details can be found.

In plastic regions the yield function is satisfied by giving the inplane principal stresses the parameterization $\sigma_1(s)$ and $\sigma_2(s)$. The plane components of stress are related to the principal stresses by

$$\sigma_{11}=\sigma_1\cos^2\theta+\sigma_2\sin^2\theta,\quad \sigma_{22}=\sigma_1\sin^2\theta+\sigma_2\cos^2\theta,$$
$$\sigma_{12}=(\sigma_1-\sigma_2)\sin\theta\cos\theta, \tag{12}$$

where θ is the angle between the x_1 axis and the direction of σ_1. Upon introducing curvilinear coordinates u and v along the principal directions of stress and regarding s and θ as functions of u and v, the equations of equilibrium take the form

$$\frac{\partial s}{\partial u}=g(s)\frac{\partial\theta}{\partial v},\quad \frac{\partial s}{\partial v}=h(s)\frac{\partial\theta}{\partial u}, \tag{13}$$

where

$$g(s)=(\sigma_2-\sigma_1)/\sigma_1',\quad h(s)=(\sigma_2-\sigma_1)/\sigma_2',$$

in which prime denotes differentiation with respect to s. The characteristic curves C^+ and C^- associated with equations (13) are determined by

$$\frac{dv}{du}=\left\{\begin{array}{l}+\tan\alpha \text{ on } C^+\\ -\tan\alpha \text{ on } C^-\end{array}\right\},\quad \tan\alpha=(g/h)^{1/2},\quad gh>0, \tag{14}$$

where α is the angle between the characteristic curves and the u direction. The restriction $gh>0$ is necessary for the equations to be hyperbolic. Compatibility equations along the characteristics reduce to

$$S-\Gamma\theta=\text{const. on } C^+,\quad S+\Gamma\theta=\text{const. on } C^-, \tag{15}$$

where

$$S=-\int(gh)^{-1/2}ds,\quad \Gamma\equiv\operatorname{sgn}(g)=\operatorname{sgn}(h).$$

In contrast to Geiringer (1973) we have not assumed that $g>0$, an assumption which is not convenient in the present application. Further, it is known that the strain rate-velocity equations (7) and the flow rule (1) admit characteristic curves that coincide with the stress characteristics (14).

The development of an analytical solution to a specific plane problem by the method of characteristics involves the combination of various basic solutions to satisfy the boundary conditions of the problem. The boundary between the plastic zone and the rigid remainder of the body must be a characteristic curve. The basic solutions to be used here are the constant stress solution and the centered wave solution. In the constant stress solution both characteristics are straight lines and the parameter s is constant. In the centered wave solution one family of characteristics, say C^+, consists of radial straight lines from a point. In polar coordinates r and ψ, the C^- characteristics are found to be determined by (Geiringer, 1973)

$$r\frac{d\psi}{dr}=-\tan 2\alpha, \tag{16}$$

and

$$S(s)=2\eta_o+(\psi-\alpha)\Gamma \quad \text{on } C^-, \tag{17}$$

where η_o is an arbitrary constant. When α is constant, integration of equation (16) gives logarithmic spirals for the C^- characteristics.

The foregoing formulation of the method of characteristics will be applied for the quadratic yield function (2) for plane stress and plane strain.

Plane Stress. For plane stress ($\sigma_{33}=0$) the yield condition ($f=0$) from equation (2) is satisfied by the parameterization

$$\left\{\begin{array}{l}\sigma_1\\ \sigma_2\end{array}\right\}=\tfrac{1}{2}[\mp A\cos s+B\sin s+C], \tag{18}$$

where

$$A=\frac{2}{\sqrt{a}}\left[1+\frac{3b^2}{(a+12c)}\right]^{1/2},\quad B=A\left[\frac{3a}{a+12c}\right]^{1/2},$$
$$C=\frac{-6b}{(a+12c)}.$$

By equations (13) and (14) we have

$$\tan\alpha=[B\cos s-A\sin s]^{1/2}[B\cos s+A\sin s]^{-1/2},$$
$$g=2/(B/A+\tan s),\quad h=2/(B/A-\tan s), \tag{19}$$

and

$$g>0,\quad h>0,\quad gh>0 \quad \text{if} \quad |\tan s|<B/A, \tag{20}$$

which restricts the states of stress for which characteristics exist. By equations (15)

$$S(s)=\tfrac{1}{2}\tan^{-1}\left\{\frac{AD}{B}\sin s\right\}$$
$$-\tfrac{1}{2}\left[1+\frac{B^2}{A^2}\right]^{1/2}\tan^{-1}\left\{\left[1+\frac{B^2}{A^2}\right]^{1/2}D\sin s\right\}, \tag{21}$$

where

$$D=[\cos^2 s-(A/B)^2\sin^2 s]^{-1/2}.$$

The foregoing results for plane stress remain valid for the case $c=0$.

Plane Strain. The plane strain condition for a rigid/plastic material ($\dot{e}_{33}=0$) and the flow rule (1) give

$$\sigma_3=\sigma_{33}=[(a-6c)(\sigma_{11}+\sigma_{22})-3b]/(2a+6c), \tag{22}$$

which can be used to eliminate σ_3 from the yield function (2). Then, $f=0$ is satisfied by the parameterization (18) with

$$A=\left[\frac{b^2}{ac}+\frac{4}{a}\right]^{1/2},\quad B=\frac{A}{3}\left[3+\frac{a}{c}\right]^{1/2},$$
$$C=-\frac{b}{3c}. \tag{23}$$

For the limiting case $c=0$, equations (23) are singular and equation (18) must be replaced by a new parameterization. A quadratic form in s for σ_1 and σ_2 has been used but details will not be given here.

Wedge Compression. We consider a truncated, rigid/plastic wedge acted on by a flat rigid indenter as shown in Fig. 2. The indenter compresses the wedge with contact pressure p and imparts a shear stress τ. The wedge may be unsymmetric with different side angles β_1 and β_2 (Fig. 2). Guided by previous solutions, we assume that a plastic region $ABCDO$ forms while the remainder of the wedge is rigid. Another form of solution is possible in which plastic regions extend to both faces of the wedge. However, the present solution with a single plastic region leads to a lower value for p in general. The single

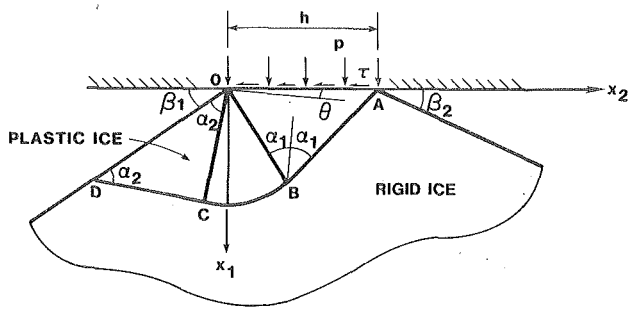


Fig. 2 Plane rigid/plastic ice wedge under a rigid indenter

plastic zone could extend to the right rather than the left in which case the solution can be obtained from the solution for Fig. 2 by considering the mirror image of Fig. 2. In order for the notation in the mirror image solution to conform to the original Fig. 2, β_2 and β_1 must be interchanged and the sign of the shear traction τ must be reversed in the solution. Either the β_1 solution (Fig. 2) or the β_2 solution (mirror image of Fig. 2) can be chosen so as to minimize the contact pressure p. For $\tau \geqq 0$ we find that the β_1 solution minimizes p if $\beta_1 \geqq \beta_2$, as can be seen from the details of the solution. As usual for rigid/plastic solutions we do not determine a stress field in the rigid region. Thus we have only an upper bound on the plastic limit load. Also, we have not determined the velocity field in the plastic region although it appears possible to do so.

With reference to rectangular Cartesian coordinates x_1 and x_2 (Fig. 2), the region ABO is in a state of constant stress with principal stresses

$$\sigma_1(s_1) \equiv \hat{\sigma}_1, \quad \sigma_2(s_1) \equiv \hat{\sigma}_2, \quad \text{and } \theta(s_1) \equiv -\phi. \tag{24}$$

The stress boundary condition on AO requires that

$$\begin{aligned} \hat{\sigma}_1 \cos^2\phi + \hat{\sigma}_2 \sin^2\phi &= -p, \\ (\hat{\sigma}_2 - \hat{\sigma}_1)\sin\,\phi\,\cos\,\phi &= \tau. \end{aligned} \tag{25}$$

Then, assuming that p and τ are related by the Coulomb friction law

$$\tau = \mu p, \tag{26}$$

equations (25) lead to

$$\begin{aligned} \phi &= \tfrac{1}{2}\xi - \tfrac{1}{2}\sin^{-1}[(\sin\xi)\,(\hat{\sigma}_1 + \hat{\sigma}_2)/(\hat{\sigma}_2 - \hat{\sigma}_1)], \\ \xi &\equiv \tan^{-1}\mu, \end{aligned} \tag{27}$$

which determines ϕ in terms of $\hat{\sigma}_1$ and $\hat{\sigma}_2$, i.e., in terms of s_1.

The region CDO also is a region of constant stress with DO stress free which, with equations (12), leads to the principal stresses

$$\sigma_1(s_2) = -\sigma_{12}(s_2)[\tan\beta + \operatorname{ctn}\beta], \tag{28a}$$

$$\sigma_2(s_2) = 0. \tag{28b}$$

The stress parameter s_2 is determined by equation (28b).

The region BCO has the centered wave solution with equation (17) holding on the C^- curves. On the C^+ lines OB and OC (Fig. 2)

$$\psi = \alpha(s_1) - \phi \quad \text{on } OB, \quad \psi = \alpha(s_2) + \beta - \pi/2 \quad \text{on } OC, \tag{29}$$

and then equation (17) gives

$$S(s_2) - S(s_1) = \pi/2 - \beta - \Gamma\phi. \tag{30}$$

With s_2 known from equation (28b), equation (30) determines s_1 and then equations (25) and (27) determine the contact pressure p. Equation (30) can be solved numerically for s_1 for a specific yield function with known function $S(s)$.

A limitation on the foregoing solution is the requirement that the region BCO exists, i.e.,

$$\phi + \beta + \alpha(s_2) < \alpha(s_1) + \pi/2. \tag{31}$$

This inequality can be checked once s_1 and s_2 are determined. If equation (31) is violated then the discontinuous solution of the next subsection will apply.

There is a further limitation on the foregoing solution, namely that the characteristics remain real (14). The requirement $gh > 0$ must be checked for specific yield functions. It turns out to be a problem only for plane stress where the state of stress in the region ABO approaches the limit point $\sigma_1'(s) = 0$ as β_1 decreases to the critical value β^*. By analogy with other solutions for plane stress in plates, e.g., Taylor (1948), the author speculates that a solution can be obtained for $\beta_1 < \beta^*$ by allowing the plate to thicken near the indenter OA. Such a solution is beyond the scope of the present investigation. However, we believe that the contact stresses in the thickened region will not exceed the values obtained in the present solution for constant thickness.

Wedge Compression – Discontinuous Solution. When the continuous solution violates (31), the region BOC (Fig. 2) vanishes and there is a discontinuity in stress along the line OB. The solution for the discontinuous case was given by Prager (1948) for the von Mises yield function and by Shield (1954) for the Mohr-Coulomb yield function. We follow the same method of solution for the three-parameter yield function (2).

In the regions AOB and BOC (Fig. 2 with $B \equiv C$) the constant state of stress is again represented by equations (24) and (28), respectively. For equilibrium the stress vector across OB must be continuous which leads to the requirements

$$\begin{aligned} \delta\sigma_{22} - \delta\sigma_{11} + (\delta\sigma_{11} + \delta\sigma_{22})\cos 2\gamma &= 0, \\ 2\delta\sigma_{12} - (\delta\sigma_{22} + \delta\sigma_{11})\sin 2\gamma &= 0, \end{aligned} \tag{32}$$

where γ is the angle between OB and the normal to OA, and

$$\delta\sigma_{ij} = \sigma_{ij}(s_2) - \sigma_{ij}(s_1).$$

Then, by equations (12), (24), (28) and (32), we obtain the following continuity requirements on the principal stresses:

$$\begin{aligned} [\delta\sigma_1 + \delta\sigma_2]\sin 2\gamma &= [\sigma_2(s_2) - \sigma_1(s_2)]\sin 2(\beta_1 + \phi), \\ [\delta\sigma_1 + \delta\sigma_2]\cos 2\gamma &= \sigma_2(s_1) - \sigma_1(s_1) \\ &\quad + [\sigma_2(s_2) - \sigma_1(s_2)]\cos 2(\beta_1 + \phi). \end{aligned} \tag{33}$$

With s_2 determined by equation (28b) and ϕ by equation (27), equations (33) serves to determine s_1 and γ. The solution for s_1 can be obtained by numerical iteration for specific yield functions. Then the contact pressure p follows from equations (25) and (27) as in the continuous solution. Numerical results are presented later in this paper.

Upper Bounds for Rectangular Indenters

When the indenter is rectangular the contact problem is three-dimensional in nature. In this case an upper bound on the contact force can be obtained from the upper-bound theorem of limit analysis. Two failure modes will be considered here, namely, a crushing mode where the deformation is localized, and a splitting mode where the ice feature fails across its thickness. Before treating three-dimensional crushing, we obtain an upper bound for the plane contact problem for comparison with the rigid/plastic solution of the previous section. The plane limit analysis also serves as a basis for the three-dimensional analysis.

Plane Crushing. For application of the upper-bound theorem to the plane crushing problem, we employ a velocity field that is constant in the regions ABD and BCD of the x_1, x_2 plane in Fig. 3. The remainder of the ice is taken to be rigid and thus its shape is arbitrary for this analysis. In ABD and BCD, respectively, the velocity vector is taken as

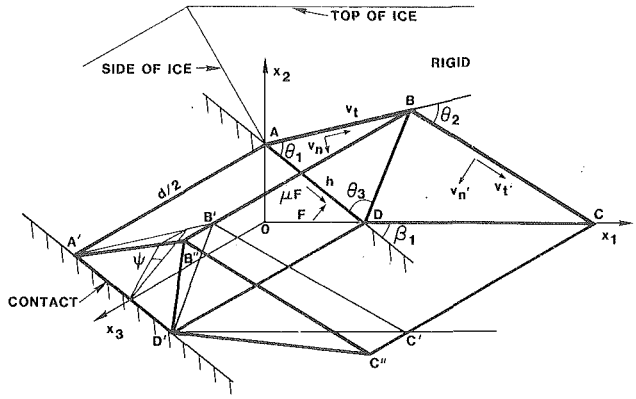


Fig. 3 Three-dimensional crushing of an ice wedge by a rigid indenter. Velocity field for limit analysis.

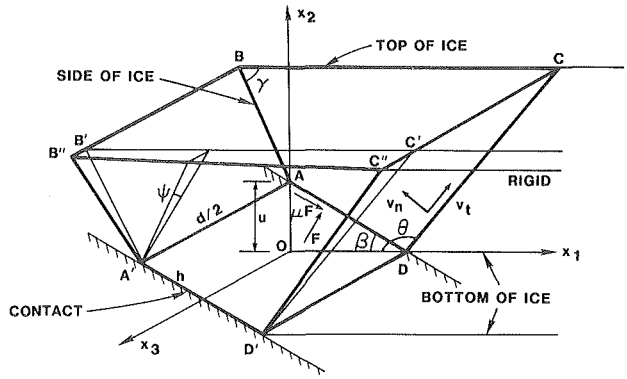


Fig. 5 Three-dimensional splitting of an ice wedge by a rigid indenter. Velocity field for limit analysis.

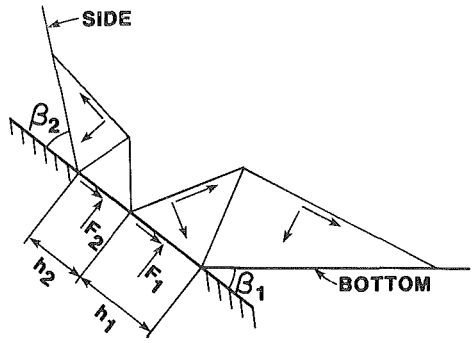


Fig. 4 Alternate velocity field for three-dimensional crushing

$$\mathbf{v}_{ABD} = [v_t \cos\theta_1 + v_n \sin\theta_1]\mathbf{e}_1 + [v_t \sin\theta_1 - v_n \cos\theta_1]\mathbf{e}_2,$$

$$\mathbf{v}_{BCD} = [v_t' \cos\theta_{12} + v_n' \sin\theta_{12}]\mathbf{e}_1 + [v_t' \sin\theta_{12} - v_n' \cos\theta_{12}]\mathbf{e}_2,$$

$$\theta_{12} = \theta_1 - \theta_2. \tag{34}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are unit base vectors and the constant velocity components v_t, v_n, v_t', v_n' and the angles θ_1, θ_2 are shown in Fig. 3. The velocity discontinuities are v_t, v_n, on AB; v_t', v_n' on BC; and they are easily determined in terms of v_t, v_n, v_t', and v_n' on DB. The discontinuity-surface dissipation function F_s on AB, BC, and DB can be evaluated from equation (9) and is denoted by F_{AB}, F_{BC}, and F_{DB}, respectively. The rate of work on DA by the normal force F and tangential force μF (Fig. 3) is readily calculated and then the upper-bound theorem (5) gives

$$F \le \frac{h[F_{AB}\sin\theta_3 + F_{BC}\sin\theta_1 \sin(\beta_1+\theta_3)/\sin(\beta_1+\theta_2-\theta_1) + F_{DB}\sin\theta_1]}{[v_t \sin\theta_1 - v_n \cos\theta_1 + \mu(v_t \cos\theta_1 + v_n \sin\theta_1)]\sin(\theta_1+\theta_3)}, \tag{35}$$

where h denotes the indenter width DA. The upper bound (35) can be minimzed with respect to v_t, v_t', v_n' θ_1, θ_2, and θ_3 while setting $v_n = 1$. The minimization has been carried out numerically by digital computer.

On dividing F from minimization of (35) by the indenter width h we have an upper bound on the average contact pressure. This upper bound is found to be within a few percent of the contact pressure from the rigid/plastic solution of the previous section for 15 deg $\le \beta \le$ 90 deg in general. This agreement increases confidence in the upper-bound solution for the chosen velocity field.

Three-Dimensional Crushing. The foregoing treatment of plane crushing can be extended to an indenter of finite length d using the same failure mechanism (Fig. 3) where there is symmetry about the x_1, x_2 plane. The planes $ABB''A'$, $BCC''B''$, and $A'B''C''D'$ (and there images about $x_3 = 0$) separate the plastic region from the rigid ice feature. A slight reduction in the upper bound is obtained in some cases by rotating the planes $A'B'C'D'$ about $A'D'$ through the angle ψ (Fig. 3). Dissipation on the plane $A'B''C''D'$ can be determined from the velocity field (34) and added to the numerator in (35) which is first multiplied by d as is the denominator. The determination of the dissipation on the rotated face $A'B''C''D'$ is straightforward but lengthy and the details will not be given here. Minimization of the upper bound is carried out with respect to ψ and the variables of the plane analysis of the preceding subsection. Numerical results are presented in the next section of this paper.

The upper bound for three-dimensional crushing may be reduced in some cases by considering a second set of velocity fields extending upward as shown in the x_1, x_2 plane of Fig. 4. The analysis for this second set is of the same form as the first set since it is its mirror image. The widths h_1 and h_2 (Fig. 4) can be adjusted to minimize the total force $F_1 + F_2$.

Three-Dimensional Splitting. The splitting mode of failure is shown in Fig. 5 where there is symmetry about the x_1, x_2 plane. The planes $DCC''D'$ and $A'B''C''D'$ (and their images about $x_3 = 0$) separate the plastic region from the rigid ice sheet. The velocity vector in the plastic region $ABCDA'B''C''D'$ is given by

$$\mathbf{v} = -[v_t \cos\theta + v_n \sin\theta]\mathbf{e}_1 + [v_t \sin\theta - v_n \cos\theta]\mathbf{e}_2. \tag{36}$$

Again the energy dissipation on $DCC''D'$ and $A'B''C''D'$ can be calculated along with the work done by F and μF on $ADD'A'$. Then an upper bound on the contact force F follows from expression (5). The analysis is similar to that for the crushing mode and details will not be given here. Numerical minimization is carried out with respect to v_t/v_n, θ, and ψ as before.

Combined Crushing and Splitting. In some cases the previous upper bounds can be reduced by combining the failure mechanisms of crushing and splitting as shown in Fig. 6. Here again the widths h_1 and h_2 can be adjusted to minimize the total force $F_1 + F_2$.

Numerical Results

Limited numerical results will be given here to illustrate the effect of parameters of interest in ice contact problems. It is convenient to use strength parameters determined from the results of confined (conventional) triaxial tests on cylindrical specimens as shown in Fig. 1. Namely, we use the unconfined

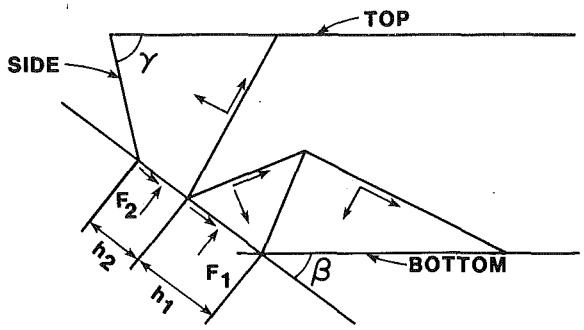


Fig. 6 Velocity field for combined crushing and splitting

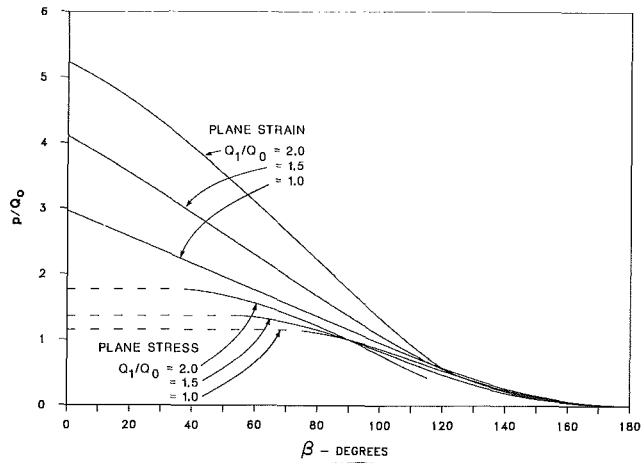


Fig. 7 Contact pressure versus contact angle for plane crushing of a rigid/perfectly plastic wedge for $P_1/Q_0 = 2$, $\mu = 0$, and varied Q_1/Q_0

compressive strength Q_0, the maximum yield stress Q_1 and the pressure P_1 at this maximum, where yield stress is defined as axial stress minus confining pressure. The parameters a, b, and c in the yield function (2) are easily determined in terms of Q_0, Q_1, and P_1. In presenting results the calculated contact pressure is ratioed on Q_0 with Q_1/Q_0 and P_1/Q_0 as parameters. The results presented correspond to the rising portion of the quadratic yield function in Fig. 1 for which the first inequality of (11) is satisfied.

From the plane rigid/plastic analysis the effects of Q_1/Q_0, P_1/Q_0 and contact angle β (β_1 or β_2 in Fig. 2) on contact pressure p are illustrated in Figs. 7 and 8. Here p/Q_0 is plotted against the contact angle β for both plane strain and plane stress without friction ($\mu = 0$). The discontinuous solution applies for $\beta > 90$ deg (approximately) in the plane strain case. As seen from Figs. 7 and 8, p is strongly dependent on β as might be expected from previous results for other yield functions. The dependence of p on Q_1/Q_0 (Fig. 7) is much stronger than on P_1/Q_0 (Fig. 8). The plane stress case has much lower contact pressure than the corresponding plane strain case for $\beta < 90$ deg. The dashed lines for plane stress represent the limiting stress imposed by the yield condition as discussed in the foregoing derivation after expression (31). Finite thickening of the plane-stress plate is expected for β in the range of these dashed lines. The reduction in contact pressure for plane stress has practical implications with regard to confinement conditions and possible vertical fracture through the thickness during contact of the ice feature with an arctic structure.

The significant effect of the coefficient of friction on contact pressure is illustrated in Fig. 9, again from the plane rigid/plastic analysis. The sign of μ is positive for shear stress τ in the direction shown in Fig. 2. For applications to arctic structures the positive sign for μ will usually apply as the ice tends to move up the structure during contact. The negative value for μ may apply to the case where the plastic zones extend to the right in Fig. 2 and β is identified with β_2. It should be noted that there is considerable uncertainty on the value of μ for ice on steel, concrete, or structural coatings.

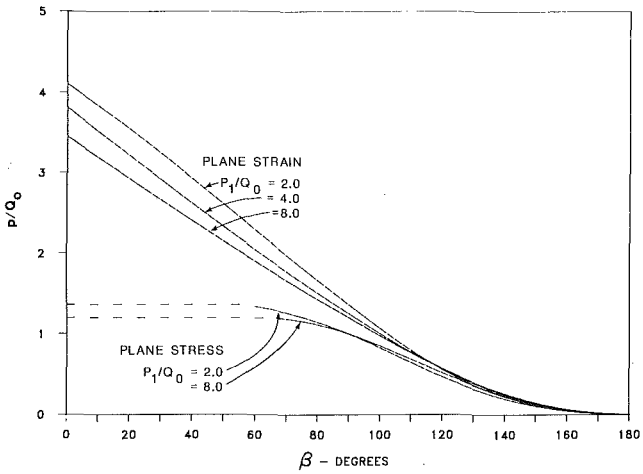


Fig. 8 Contact pressure versus contact angle for plane crushing of a rigid/perfectly plastic wedge for $Q_1/Q_0 = 1.5$, $\mu = 0$, and varied P_1/Q_0

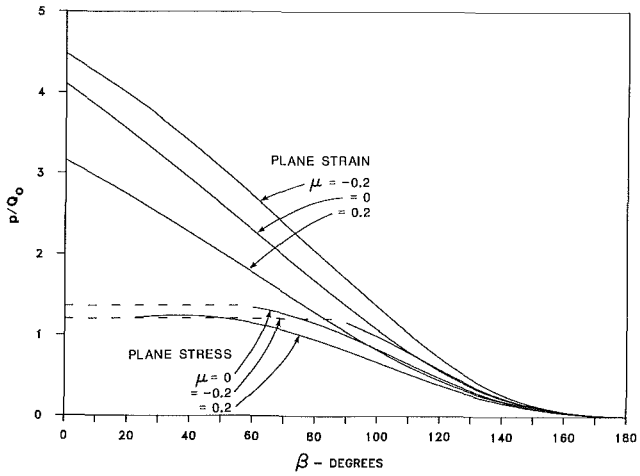


Fig. 9 Contact pressure versus contact angle for plane crushing of a rigid/perfectly plastic wedge for $Q_1/Q_0 = 1.5$, $P_1/Q_0 = 2$, and varied μ

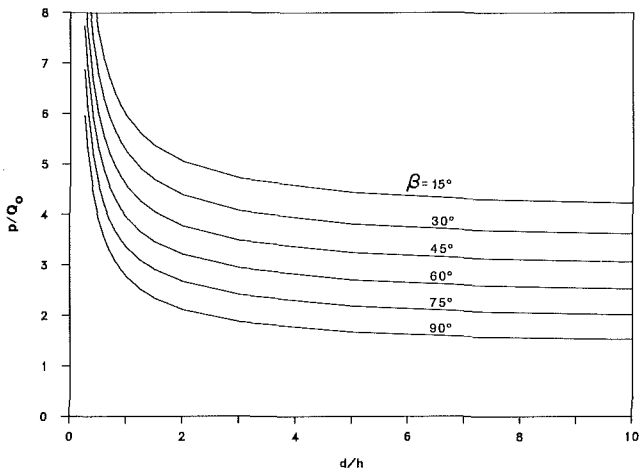


Fig. 10 Contact pressure (upper bound) versus aspect ratio for three-dimensional crushing of a perfectly plastic wedge for $Q_1/Q_0 = 1.5$, $P_1/Q_0 = 2$, $\mu = 0$, $\gamma = 90$ deg, and varied contact angle

Limit analysis results for three-dimensional crushing of a truncated wedge (Fig. 3) are shown in Fig. 10. The average

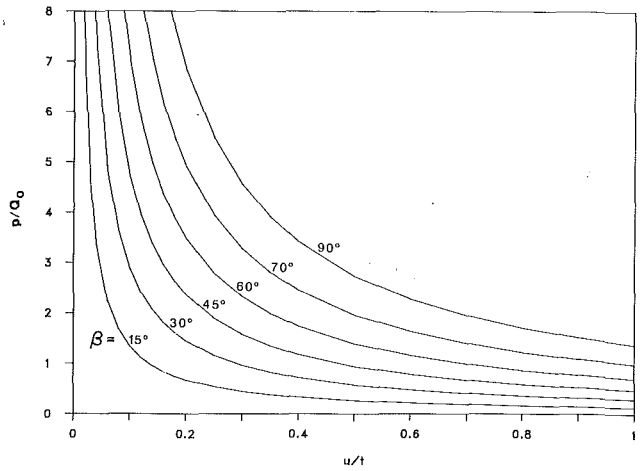


Fig. 11 Contact pressure (upper bound) versus indenter width parameter (Fig. 5) for plane splitting of a perfectly plastic wedge for $Q_1/Q_0 = 1.5$, $P_1/Q_0 = 2$, $\mu = 0$, and varied contact angle

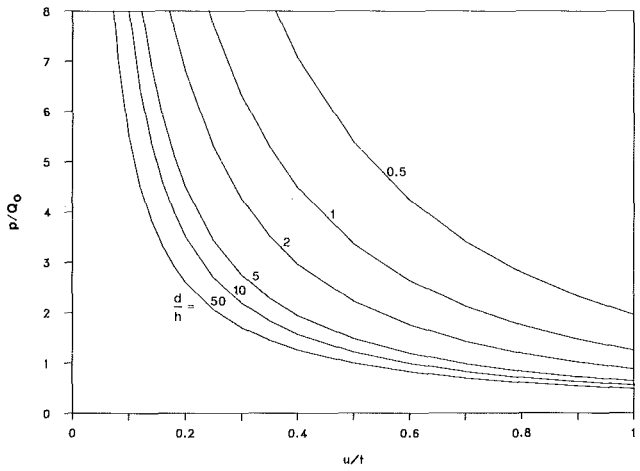


Fig. 12 Contact pressure (upper bound) versus indenter width parameter (Fig. 5) for three-dimensional splitting of a perfectly plastic wedge for $\beta = 45$ deg, $\gamma = 90$ deg, $Q_1/Q_0 = 1.5$, $P_1/Q_0 = 2$, $\mu = 0$ and varied aspect ratio

contact pressure p/Q_0 obtained from the upper bound on contact force is plotted versus the aspect ratio d/h of the contact area for various values of the contact angle β (β_1 in Fig. 3). As in the plane case, p increases with decreasing β. Furthermore, p increases sharply as d/h decreases below 2. At a value of d/h somewhat less than 1 it is expected that there exists an alternate failure mechanism which gives a lower upper bound than that shown in Fig. 10. This was found to be the case in the limit analysis of Reinicke and Remer (1978) for $\beta = 90$ deg. The failure mechanism in their case was plane strain in horizontal planes and a similar mechanism is expected for $\beta < 90$ deg. An estimate of the contact pressure in this new mechanism can be obtained from the plane-strain, rigid/plastic solution for $\beta = 0$. This solution is now applied in the plane normal to the contact plane and the vertical plane rather than in the vertical plane. For example, from Fig. 7 for $Q_1/Q_0 = 1.5$ and $\beta = 0$ the estimated maximum value of p/Q_0 is 4.1. Guided by the upper-bound theorem and the geometry of Fig. 3, we speculate that the actual value of p is larger than 4.1 and it increases with decreasing β.

The results of Fig. 10 also can be applied to the velocity field of Fig. 4 with h and β interpreted as h_2 and β_2 as well as h_1 and β_1. Then h_1/h_2 can be adjusted to minimize the total force $F_1 + F_2$.

Results for limit analysis of splitting are shown in Figs. 11 and 12. The thickness of the ice sheet t enters the problem for the splitting mode. In the special case of plane strain ($d/h \gg 1$), Fig. 11 shows p/Q_0 versus the fraction u/t of sheet thickness penetrated for a range of slope angles β (Fig. 5). The contact pressure decreases with increasing u/t and decreasing β. The effect of aspect ratio d/h is illustrated in Fig. 12 for $\beta = 45$ deg. Contact pressure increases substantially with decreasing aspect ratio. A similar set of curves can be calculated for other values of β. A combination of splitting and crushing modes as shown in Fig. 6 can be considered using Figs. 10 and 12, again adjusting h_1/h_2 to minimize the total force.

From the results of Figs. 10 and 12 one can determine if a transition from crushing to splitting failure occurs as u/t increases. The extended failure modes of Figs. 4 and 6 can also be considered in the transition study. In some cases the transition from crushing to splitting failure furnishes a maximum contact area on which the higher contact stress from crushing acts.

Conclusions

We have considered plane and three-dimensional crushing and splitting failure of ice features in contact with a rigid indenter using rigid/perfectly plastic solutions and the upper-bound theorem of limit analysis. Numerical results have been presented to illustrate the effects of various parameters on the contact pressure. These results are intended for application to the prediction of local contact stress of ice features on offshore arctic structures with vertical and sloping sides. More extensive numerical results can be generated for application to specific arctic structures.

In applications it must be kept in mind that plasticity theory may not provide a completely adequate model for the complex material behavior of sea ice. The use of laboratory-measured strength parameters in the yield function is open to question and corrections may be necessary to account for rate effects, scale effects, anisotropy and inhomogeneity. Further, the noted decrease in contact stress from plane strain to plane stress indicates sensitivity to lateral confinement conditions during contact. To help resolve such questions, large-scale experimental results and field measurements would be valuable.

In addition to the crushing and splitting failure modes analyzed here, other failure modes are possible for ice features on arctic structures, e.g., bending or cracking modes of failure. Also, the ice may fail away from the structure. Such failure modes could limit contact stress to values below those given by the present results for crushing and splitting failure. The other failure modes seem more likely to occur as the contact area increases in size. Thus, the present results appear to be most applicable to relatively small contact areas.

Further theoretical investigation is possible on several related topics of technical interest. The problem of plane *finite* deformation of a symmetric wedge under a rigid stamp was solved by Hill (1950) for the Tresca yield function. This analysis might be extended to other yield functions and to the unsymmetric wedge. Solutions to contact problems for a viscoplastic material are of interest since the strength of ice depends on the strain rate. Solutions for an anisotropic plastic or viscoplastic material also would be useful. Application of the lower-bound collapse theorem of limit analysis may be helpful in studying overall failure modes of the ice feature during contact. The role of fracture phenomena in ice contact problems certainly merits study.

References

Chen, W.-F., 1975, *Limit Analysis and Soil Plasticity*, Elsevier Scientific Publishing Co., Amsterdam.

Cox, G. F. N., Richter-Menge, J. A., Weeks, W. F., and Mellor, M., 1985, "A Summary of the Strength and Modulus of Ice Samples from Multi-Year Pressure Ridges," ASME *Journal of Energy Resources Technology*, Vol. 107, pp. 93-102.

de Boer, R., 1987, "Failure Conditions for Brittle and Granular Materials," *Quart. of Applied Math.*, Vol. 44, pp. 71-79.

Drucker, D. C., and Prager, W., 1952, "Soil Mechanics and Plastic Analysis," *Quart. of Applied Math.*, Vol. 10, pp. 157-165.

Geiringer, H., 1973, "Ideal Plasticity," in *Handbuch der Physik*, Vol. VIa/3, C. Truesdell, ed., Springer-Verlag, Berlin, pp. 403-533.

Hill, R., 1950, *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford.

Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge.

Jones, S. J., 1978, "Triaxial Testing of Polycrystalline Ice," *Proceedings of the Third International Congress on Permafrost*, Vol. 1, Edmonton, pp. 670-674.

Jones, S. J., 1982, "The Confined Compressive Strength of Polycrystalline Ice," *J. Glaciology*, Vol. 28, pp. 171-177.

Koiter, W. T., 1960, "General Theorems for Elastic-Plastic Solids," in *Progress in Solid Mechanics*, Vol. 1, I. N. Sneddon and R. Hill, eds., North-Holland Publishing Co., Amsterdam, pp. 167-221.

Murrell, S. A. F., 1963, "A Criterion for Brittle Fracture of Rocks and Concrete under Triaxial Stress and the Effect of Pore Pressure on the Criterion," *Proceedings of the Fifth Rock Mechanics Symposium*, in *Rock Mechanics*, C. Fairhurst, ed., University of Minnesota, pp. 563-577.

Prager, W., 1948, "Discontinuous Solutions in the Theory of Plasticity," in *Courant Anniversary Volume*, Interscience Publishers, New York, pp. 289-299.

Prager, W., and Hodge, P. G., Jr., 1951, *Theory of Perfectly Plastic Solids*, Wiley, New York.

Prandtl, L., 1921, "Uber die Eindringungsfestigkeit (Harte) plastischer Baustoffe und die Festigkeit von Schneiden," *Z. Angew. Math. Mech.*, Vol. 1, pp. 15-20.

Ralston, T. D., 1978, "An Analysis of Ice Sheet Indentation," *Proceedings of the Fourth IAHR Symposium on Ice Problems*, Luleå, Sweden, pp. 13-21.

Reinicke, K. M., 1980, "Analytical Approach for the Determination of Ice Forces Using Plasticity Theory," *IUTAM Symposium on Physics and Mechanics of Ice*, Copenhagen, Aug. 1979, Springer-Verlag, Berlin.

Reinicke, K. M., and Remer, R., 1978, "A Procedure for the Determination of Ice Forces – Illustrated for Polycrystalline Ice," *Proceedings of the Fourth IAHR Symposium on Ice Problems*, Luleå, Sweden, pp. 217-234.

Shield, R. T., 1954, "Stress and Velocity Fields in Soil Mechanics," *J. Math. Phys.*, Vol. 33, pp. 144-156.

Taylor, G. I., 1948, "The Formation and Enlargement of a Circular Hole in a Thin Plastic Sheet," *Quart. J. of Mech. Appl. Math.*, Vol. 1, pp. 103-124.

Winkler, M. M., and Nordgren, R. P., 1986, "Ice Ridge Ride-Up Forces on Conical Structures," *Proceedings of the Sixth IAHR Symposium on Ice Problems*, Iowa City, pp. 171-183.

APPENDIX

A derivation is given here for the elimination of the stress tensor from the dissipation functions F and F_s in the upper-bound theorem (5) to obtain (8) and (9) with the restrictions (11). The derivation is carried as far as possible for the class of yield functions of the form $f(I_1, J_2)$ which contains (2) as a special case. The final stages of the derivation are carried out for the yield function (2). It should be noted that (2) cannot be written in the form of a homogeneous function of the stress tensor. Thus, Euler's theorem cannot be used in the derivation as was done by Drucker and Prager (1952) and de Boer (1987).

For $f(I_1, J_2)$ the flow rule (1) with the definitions (3) gives the plastic strain rate tensor as

$$\mathring{e}_{ij} = \lambda\left[\frac{\partial f}{\partial I_1}\,\delta_{ij} + \frac{\partial f}{\partial J_2}\,s_{ij}\right], \tag{A-1}$$

where the double prime on e_{ij} is dropped for simplicity. From (A-1) and (3) we obtain the identity

$$9\mathring{\mathcal{E}}_{ij}\mathring{\mathcal{E}}_{ij}\left(\frac{\partial f}{\partial I_1}\right)^2 = 2(\mathring{e}_{kk})^2 J_2\left(\frac{\partial f}{\partial J_2}\right)^2, \tag{A-2}$$

where $\mathcal{E}_{ij}$ is the plastic deviatoric strain tensor defined by (10). By (A-1) and the yield condition ($f=0$), the volumetric dissipation function (6) becomes

$$F = \tfrac{1}{3}\mathring{e}_{kk}\left[I_1 - 2J_2\,\frac{dI_1}{dJ_2}\right], \tag{A-3}$$

where the J_2 derivative is determined from $f(I_1, J_2) = 0$. Since F is positive definite, both $\mathring{e}_{kk}$ and the quantity in brackets in (A-3) must have the same sign. Since J_2 decreases with increasing I_1 for cases of greatest practical interest, attention is confined to the case

$$\mathring{e}_{kk} \geqq 0 \quad \text{and} \quad I_1 - 2J_2\,\frac{dI_1}{dJ_2} > 0. \tag{A-4}$$

The other case of opposite inequalities from (A-4) also could be treated. In view of (A-4) the positive sign is chosen on taking the square root of both sides of (A-2) which allows (A-3) to be written as

$$F = \tfrac{1}{3}I_1\mathring{e}_{kk} + (2J_2\mathring{\mathcal{E}}_{ij}\mathring{\mathcal{E}}_{ij})^{1/2}. \tag{A-5}$$

In order to eliminate the stress invariants from (A-5), one must solve (A-2) and $f=0$ for I_1 and J_2. This task cannot be carried out explicitly for the general case $f(I_1, J_2)$.

For the quadratic yield function (2) an explicit solution for I_1 and J_2 can be obtained which allows (A-5), after algebraic manipulations, to be written as equation (8). Further, in this case the second restriction (A-4) reduces to

$$-b/2c < I_1 < 2/b. \tag{A-6}$$

As noted previously, the yield condition from equation (2) is an ellipse in the I_1, $J_2^{1/2}$ plane for $c>0$. The first inequality of (A-6) is satisfied in the right half of the ellipse which is of practical interest for ice. The second inequality of (A-6) is satisfied automatically from the yield condition. The hyperbolic case $c<0$ may be of interest for other materials and could be investigated further.

A development similar to the foregoing can be carried out for the dissipation function F_s on surfaces of velocity discontinuity. The deviatoric component of the strain jump is

$$\delta\mathcal{E}_{ij} = \delta e_{ij} - \tfrac{1}{3}\delta e_{kk}\delta_{ij}, \tag{A-7}$$

and it follows from equation (7) that

$$\delta\mathring{\mathcal{E}}_{ij}\delta\mathring{\mathcal{E}}_{ij} = \tfrac{1}{2}|\delta v_t|^2 + \tfrac{2}{3}|\delta v_n|^2. \tag{A-8}$$

With the aid of (A-8) and the flow rule for $\delta\mathring{e}_{ij}$ the identity corresponding to (A-2) is

$$\frac{9}{4}\,\delta v_t^2\left(\frac{\partial f}{\partial I_1}\right)^2 = \delta v_n^2\left[J_2\left(\frac{\partial f}{\partial J_2}\right)^2 - 3\left(\frac{\partial f}{\partial I_1}\right)^2\right]. \tag{A-9}$$

The quantity in brackets in (A-9) must be positive which is a further restriction on the yield function f. With the restriction (A-4) on I_1 and $\delta v_n > 0$, upon taking the positive sign for the square root of (A-9), the surface dissipation function (6) becomes

$$F_s = \tfrac{1}{3}I_1\delta v_n + \left[J_2\left(\delta v_t^2 + \frac{4}{3}\,\delta v_n^2\right)\right]^{1/2}. \tag{A-10}$$

Again, to eliminate I_1 and J_2 from (A-10) one must solve (A-9) and $f = 0$ for I_1 and J_2. For the quadratic yield function (2) the elimination of I_1 and J_2 can be carried out explicitly and (A-10) reduces to equation (9). Also for (2), the restriction from (A-9) becomes

$$J_2 > \frac{3(b^2+4c)}{a(a+12c)}, \tag{A-11}$$

which is satisfied if (sufficient condition)

$$a > 3b^2 \quad \text{and} \quad -b/2c < I_1 < 0. \tag{A-12}$$

The restrictions derived in this Appendix are combined into expression (11).

Some History-Dependent Problems for Dissimilar Cylinders With Finite Friction

A. Azarkhin
Product Engineering Division,
Alcoa Laboratories,
Alcoa Center, PA 15069
Mem. ASME

This paper is concerned with the problem of contact between two dissimilar elastic cylinders. The cylinders are pressed together by a compressive force, P, *and loaded by a tangential force,* H, *which is transmitted in the form of a distribution of frictional tractions at the interface. We assume that Coulomb's law of friction holds at the interface. We develop a numerical algorithm which is sufficiently general to deal with an arbitrary history of normal and tangential loading,* P(t), H(t) *and which retains the full description of the coupling between the normal and tangential tractions. The proposed algorithm gives good numerical accuracy in comparison with previous analytical and numerical solutions of a variety of problems for normal and tangential loading of two contacting cylinders with friction. It is also found to be very efficient in the use of computer time. Results for some previously unsolved problems involving the full coupling between normal and tangential contact tractions indicate that such coupling can have a significant effect on the solution, particularly in regard to the extent of slip and stick zones within the contact area.*

Introduction

This paper is concerned with the problem of contact between two dissimilar elastic cylinders as shown in Fig. 1. The cylinders are pressed together by a compressive force, P, and loaded by a tangential force, H, which is transmitted in the form of a distribution of frictional tractions at the interface.

We assume that Coulomb's law of friction holds at the interface, i.e., that frictional tractions in regions of relative tangential motion (slip) are proportional to the local normal pressure and oppose the relative motion, while in regions of no relative motion (stick), the magnitude of the frictional traction is lower than that which is required to produce slip. The Coulomb law is a dissipative mechanism and it causes the problem to depend upon the history of loading in a manner analogous to plastic deformation. Indeed, there are many similarities between the two processes, including the tendency of frictional contact problems to exhibit residual stress and shakedown in cyclic loading (Comminou and Barber, 1983). To give a complete definition of the problem, it is, therefore, necessary to specify the loading history, through defining P and H as functions of time, t. We note, however, that the problem is considered throughout to be quasi-static in the sense that wave propagation and other inertia effects are neglected. In other words, time enters the problem solely as a parameter defining the sequence of loading.

Similar Materials. A problem which illustrates this behavior is the solution by Cataneo (1938) and Mindlin (1949) for two cylinders of similar materials which are first loaded by a normal force, P, and then by a tangential load, H, P remaining constant. For similar materials, the normal tractions are unaffected by tangential loading and follow the classical Hertzian semi-elliptic distribution. The contact area, therefore, depends only upon P and remains constant while the tangential force is applied, but frictional slip in the direction of H occurs in two symmetric regions adjacent to the two edges of the contact area. The size of the slip regions increases with H until slip occurs throughout the contact area at the limiting value, H

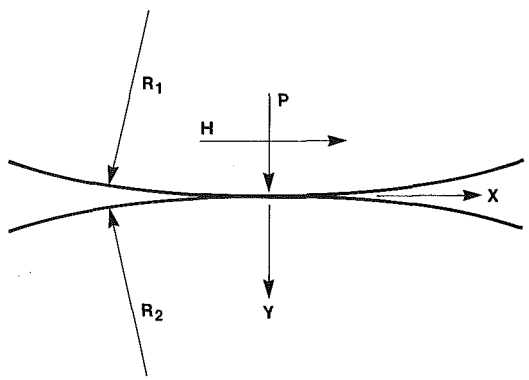


Fig. 1 Model of two contacting cylinders

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 7, 1987; final revision, July 20, 1987.

Paper No. 88-WA/APM-11.

$= \mu P$, at which point rigid body sliding or "gross slip" occurs between two cylinders.

Mindlin and Deresiewicz (1953) have investigated the related problem of two contacting spheres under more complex loading and unloading histories and document the range of frictional traction distribution which can be obtained. Similar results have also been used to investigate the problem of oblique impact between an elastic sphere and a plane (Maw et al., 1976).

A special case of some interest concerns the steady rolling contact of two cylinders with constant normal and tangential forces. Each point on the surfaces of the cylinders experiences a sequence of events, starting with initial contact at the leading edge of the contact region, passing through the region of maximum pressure and eventually emerging at the trailing edge. This loading sequence differs from that in Mindlin's static contact problem and results in a different solution, first obtained by Carter (1926). Stick is found at the leading edge of the contact area, but there is a slip region adjacent to the trailing edge. This solution finds application in the transmission of traction between rollers in machinery and in the acceleration of a wheeled vehicle on a plane surface or track.

Kalker (1970, 1971) addressed the transient question in which two cylinders are first loaded as in Mindlin's problem and are then caused to roll against each other at constant P and H. A numerical algorithm was used and the results show that the frictional traction distribution tends rapidly to the steady rolling solution.

Dissimilar Materials. The problem is considerably more complex when the two cylinders are of different materials, since in this case purely normal contact tractions tend to produce relative tangential displacement and, hence, the normal and tangential boundary value problems become coupled.

The normal loading of two dissimilar bodies with *full adhesion* was solved by Mossakovskii (1954, 1963) using an incremental approach. Goodman (1962) developed an approximate solution to the same problem, in which the influence of frictional tractions on the normal pressure distribution was neglected. Spence (1968) has since shown that the stress field at various stages of the loading history must be self-similar and, hence, a nonincremental solution can be developed.

All of these solutions predict an infinite ratio between tangential and normal tractions at the edge of the contact area and are, therefore, physically unrealistic, since some slip must occur if the coefficient of friction is finite. The corresponding problem with finite friction has been treated by Spence (1973, 1975) for both contacting cylinders and spheres. For the case of two cylinders, there is a central region of stick and symmetrically disposed slip at the edge of the contact area. However, in contrast to Mindlin's problem (see above), slip in the two regions occurs in opposite directions. All these results relate to purely normal loading ($H=0$). The corresponding problem with tangential loading has not yet been solved.

If the normal force, P, is increased to a maximum value and then reduced, a different pattern of stick and slip zones is to be anticipated during the unloading phase. Turner (1979) gives a solution to the unloading problem for an elastic half-space indented by a flat rigid punch. For this case, slip always occurs at the edge of the contact area during loading, but during unloading this region sticks until a certain proportion of the maximum load has been removed.

Bentall and Johnson (1967) developed a numerical solution for the steady rolling of two dissimilar cylinders under normal loading, in which the pressure distribution was represented in a piece-wise linear form. They found three slip zones – one at each edge of the contact area and one internal zone. Another numerical algorithm due to Kalker (1970, 1971) uses a variational formulation for the transient problem with both normal and tangential loading. However, Kalker's solution is approximate in that it neglects the influence of tangential traction on the normal contact pressure, i.e., the Hertzian pressure distribution is assumed at all times.

In this paper, we develop a numerical algorithm which is sufficiently general to deal with an arbitrary history of normal and tangential loading, $P(t)$, $H(t)$ and which retains the full description of the coupling between the normal and tangential tractions. The method used follows Bentall and Johnson (1967) in representing the normal and tangential traction distributions in piece-wise linear form and using the governing physical inequalities iteratively to determine the extents of contact, stick, and slip zones. All of the problems cited above fall under this general description and can, therefore, be used to test the resulting algorithm. The algorithm will then be applied to some further problems of practical interest in a subsequent paper.

Governing Equations

We use the Green's function representation to formulate the governing equations in terms of the unknown pressure and shearing tractions at the interface. The normal and horizontal displacements at the surface, caused by a vertical unity force applied at the origin are, respectively (Johnson, 1985, pp. 15–18),

$$v = \frac{-2(1-\nu^2)}{\pi E} \log|x| + \text{constant} \tag{1}$$

$$u = -\operatorname{sgn}(x) \frac{(1-2\nu)(1+\nu)}{2E} + \text{constant} \tag{2}$$

where y is directed into the body. The corresponding results for a horizontal unit force are:

$$v = \operatorname{sgn}(x) \frac{(1-2\nu)(1+\nu)}{2E} + \text{constant} \tag{3}$$

$$u = \frac{2(1-\nu^2)}{\pi E} \log|x| + \text{constant} \tag{4}$$

The constants in equations (1)–(4) denote rigid body displacements. Equations (1) to (4) can be used to write the following system of governing equations:

$$\frac{2}{\pi}\left(\frac{1-\nu_1^2}{E_1}+\frac{1-\nu_2^2}{E_2}\right)\int_{(AUB)} p(x_1)\log|x-x_1|\,dx_1 + \left[\frac{(1-2\nu_2)(1+\nu_2)}{2E_2}-\frac{(1-2\nu_1(1+\nu_1)}{2E_1}\right]\times$$

$$\int_{(AUB)} \tau(x_1)\operatorname{sgn}(x_1-x)\,dx_1 = \frac{x^2}{2}\left(\frac{1}{R_1}+\frac{1}{R_2}\right)-c_1 \qquad X\epsilon AUB \tag{5}$$

$$\int_{(A)} \left[\frac{(1-2\nu_2)(1+\nu_2)}{2E_2}-\frac{(1-2\nu_1)(1+\nu_1)}{2E_1}\right]\operatorname{sgn}(x_1-x)p(x_1)\,dx_1 - \frac{2}{\pi}\int_{(AUB)}\left[\frac{1-\nu_2^2}{E_2}+\frac{1-\nu_1^2}{E_1}\right]\tau(x_1)\log|x-x_1|\,dx_1 + F(x)+c_2=0; \quad X\epsilon A \tag{6}$$

$$\tau = -\mu p \operatorname{sgn}\left(\frac{du_1}{dt}-\frac{du_2}{dt}\right) \quad X\epsilon B \tag{7}$$

In these equations, p is the normal contact pressure, τ contact shearing traction, A denotes the area of stick contact, B is area of slip. Equation (5) states that there is no mismatch of the

surfaces in the vertical direction in the area of contact. In particular, the first integral gives elastic displacements caused by the pressure, the second term shows the contribution of the shear traction. The right-hand side of the equation takes into account curvatures of both cylinders and a relative rigid body displacement in the vertical direction. Similarly, in equation (6), it is stated that the horizontal displacements of the two rolls are the same in the no-slip contact area. The last term in this equation is the relative rigid body displacement in the x direction. This is formally accommodated by the term $F(x)$ in equation (6), which shows the relative slip of the two surfaces, accumulated in the process of deformation. Finally, equation (7) gives the well-known relationship (Coulomb law) between the pressure and shear in the slip contact area. Additional equations and constraints following from boundary conditions are:

$$\int_{(AUB)} p(x)dx = P \tag{8}$$

$$\int_{(AUB)} \tau(x)dx = H \tag{9}$$

$$p(x) \geq 0 \tag{10}$$

$$\frac{x^2}{2R} - c_1 - v_2 - v_1 \geq 0; \qquad AUB \tag{11}$$

Equations (8) and (9) are equilibrium equations. It is stated, respectively, in equations (10) and (11) that the normal traction is compressive in the area of contact, and that there is no interpenetration of the two cylinders in the region of separation. The effective radius, R, in equation (11) is defined by:

$$1/R = 1/R_2 + 1/R_1 \tag{12}$$

Equations (5) to (11) determine the solution to the problem, including the extents of the stick, slip and separation zones, and the constants c_1, c_2.

The solution to the problem is history-dependent and hence the local displacements and tractions as well as the applied loads are functions of time, t. This dependence is implied, but not explicitly indicated throughout the subsequent development. However, it is assumed that the rate of change of load is sufficiently slow for inertia effects to be neglected. Thus, the problem is treated as a sequence of quasi-static solutions to the incremental loading problem, the initial conditions for which depend upon the previous history.

Dimensionless Formulation

To reduce the number of parameters we write:

$$x = A_0 x^* \tag{13}$$

$$p = p^* \frac{P}{A_0} \tag{14}$$

$$\tau = \tau^* \frac{P}{A_0} \tag{15}$$

$$u = u^* A_0 \tag{16}$$

$$v = v^* A_0 \tag{17}$$

where dimensionless quantities are denoted by asterisks and A_0 denotes the semiwidth of the contact area which would be obtained in the absence of friction, i.e.,

$$A_0^2 = \frac{4PR}{\pi E} \tag{18}$$

and effective Young's modulus, E, is given by

$$\frac{1}{E} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \tag{19}$$

Then equations (6) and (7) after dropping asterisks, become:

$$\int_{AUB} p \log|x - x_1| dx_1 + \frac{\pi}{2}\beta \times \int_{AUB} \tau(x_1)\mathrm{sgn}(x_1 - x)dx_1 = x^2 - c_1; \quad X \epsilon AUB \tag{20}$$

$$\frac{\pi}{2}\beta \int_{AUB} \mathrm{sgn}(x_1 - x)p(x_1) + \int_{AUB} \tau(x_1)\log|x_1 - x|dx_1 + f(x) = c_2; \quad X \epsilon A \tag{21}$$

where Dundurs' constant [18]:

$$\beta = \frac{E_1(1-2\nu_2)(1+\nu_2) - E_2(1-2\nu_1)(1+\nu_1)}{2[E_2(1-\nu_1^2) + E_1(1-\nu_2^2)]} \tag{22}$$

and dimensionless relative slip of the surfaces is defined by

$$f(x^*) = \frac{\pi E_2 E_1 F(x)}{2P[E_1(1-\nu_2^2) + E_2(1_1\nu^2)]} \tag{23}$$

Equations (7) and (10) do not change, and equations (8), (9), and (11) can be written as:

$$\int_{AUB} p(x)dx = 1 \tag{24}$$

$$\int_{AUB} \tau(x)dx = H/P \tag{25}$$

$$x^2 - c_1 - v_1 - v_2 \geq 0 \tag{26}$$

We can see that generally the solution to the problem depends on three dimensionless parameters β, μ, H/P and the function $f(x)$ describing the history of slip. In some particular cases the number of parameters can be reduced even more. For example, for the steady state with $H = 0$ there are two parameters. If, in addition, the materials of the two cylinders are similar, $\beta = 0$ and the solution depends on μ alone. If the coefficient of friction is high enough to prevent slip, equation (7) is not used and there are no parameters at all, which means that all solutions can be obtained from one dimensionless solution. There are no parameters either when $\mu = 0$ (classical Hertzian problem).

Numerical Implementation

For numerical implementation the pressure and shear tractions are represented by superposition of a set of polygons with unknown values at apexes (nodes), following Bentall and Johnson (1967). These values are found from the solution of the system of linear algebraic equations following from integral equation (5), and from equations (6) or (7) depending, respectively, on slip or stick conditions. Equations are written for all nodes inside the contact area. We note that there are two more unknowns than there are equations, because of the constants C_1 and C_2. Two more equations come from equations (8) and (9). The problem is nonlinear in that the area of contact is unknown in advance. Nor it is known in which portion of the contact area there is slip. This is found by iteration from the solution of the system of governing equations. Initially we guess some area of contact, using for example Hertzian solution, and solve equations (20) and (21). If negative pressure is found at some nodes, we assume separation at these nodes. If overlapping of the two cylinders is found at some other nodes, we introduce contact there, and repeat the calculations. We then verify slip or no-slip conditions as follows. First, we assume the extent of slip and no-slip areas everywhere. Then we check to see if the condition:

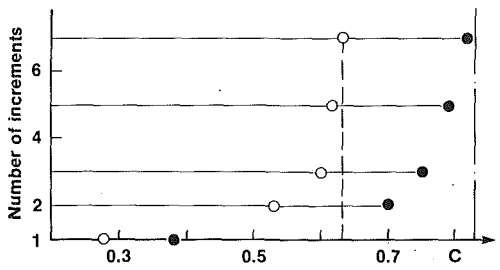


Fig. 2 Extent of the no-slip area as function of number of load increments: circles—assuming Hertzian normal pressure distribution; dark circles—with coupling between normal and tangential tractions; vertical lines—results of Spence (1975) with and without coupling

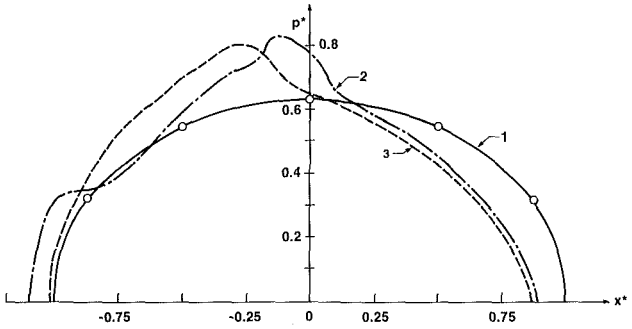


Fig. 3 Dimensionless normal pressure distribution for contact of two dissimilar cylinders, $\mu = 0.6$, $H/P = 0.3$: 1—$\beta = 0$, subsequent application of vertical and horizontal forces; 2—$\beta = 0.5$, subsequent application of vertical and horizontal forces; 3—$\beta = 0.5$, simultaneous growth of both forces

$$|\tau| \leq \mu P \tag{27}$$

is violated at some nodes. If it is, we replace one algebraic equation following from equation (6) by equation (7) with the direction of the shearing tractions taken from the previous iteration, and repeat the procedure. If, on the other hand, we assume certain signs of τ at some points of the slip area, and the results show that the relative displacements of the two surfaces do not agree with this assumption, we change the slip to stick conditions in the next iteration. Then we again iterate on contact-separation conditions, etc. Since the problem is discretized, we enforce all inequalities in a finite number of steps—and the accuracy is then determined by the grid size alone.

The history of slip, represented by the function $f(x)$ in equation (21), enters the numerical scheme as follows. For the first step of loading, there is no previously accumulated relative tangential displacements of the surfaces and $f(x) = 0$. For all the other steps of loading, the function $f(x)$ is updated on the basis of the results for the current step, namely, relative slip of the surfaces for the current step is added to $f(x)$.

For rolling problems, governing equations are written in the stationary coordinate system while points move. Let us assume, for the sake of simplicity, that the step is chosen so that rolls' surfaces move one node per increment. Then updating consists of adding the slip increments, as described above, and moving information by one node. Efficiency of the algorithm is discussed in the next section.

Examples

Problems with history dependence must be formulated incrementally and this strictly requires an infinite number of infinitely small steps. However, in the numerical solution, a relatively small number of steps is usually sufficient to give acceptable accuracy. To illustrate this, we consider Spence's

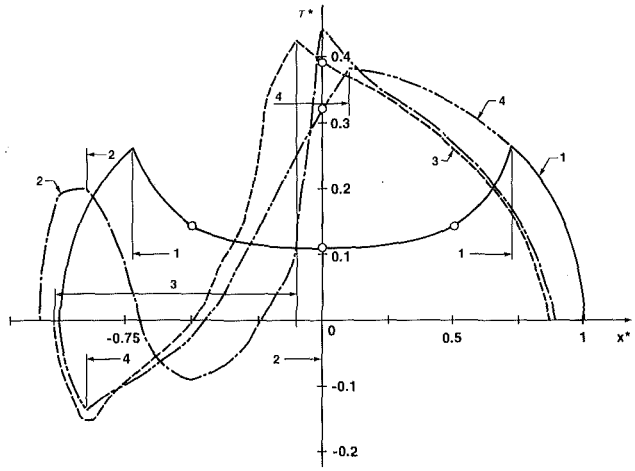


Fig. 4 Dimensionless shearing pressure distribution for contact of two dissimilar cylinders, $\mu = 0.6$, $H/P = 0.3$: 1—$\beta = 0$, subsequent application of vertical and horizontal forces; 3—$\beta = 0.5$, simultaneous growth of both forces; 4—$\beta = 0.5$, simultaneous growth of both forces, influence on shear on normal pressure is neglected

problem (1975) of two dissimilar cylinders pressed together by a vertical force which increases gradually from zero to its final value, P. Notice that here we use an incremental formulation of the problem, although Spence was able to avoid this by making use of self-similarity.

Figure 2 shows the proportion, c, of the contact area which is in a state of stick, as a function of the number of increments used to describe the growth of the normal force, P. The results relate to the case where $\beta = 0.5$ and $\mu = 0.6$, for which the exact solution (Spence, 1975) gives $c = 0.83$. The figure also shows the corresponding results which follow from the "Goodman approximation" (Goodman, 1962) of assuming that the normal contact pressure distribution is always Hertzian, for which the analytic solution gives $c = 0.63$. In both cases, we see that seven or more loading increments are sufficient to give good numerical accuracy. Results for different values of β and μ confirm this conclusion, except for high values of the ratio μ/β, for which a larger number of increments is required.

The second example concerns the contact of two cylinders loaded by both vertical and horizontal forces, contact pressure and shear traction being presented in Figs. 3 and 4, respectively. In all cases, the coefficient of friction, $\mu = 0.6$ and $H/P = 0.3$.

Curve 1 corresponds to the limiting case, $\beta = 0$, treated analytically by Cattaneo (1938) and Mindlin (1949) and the results of this analysis are shown for a few points for comparison.

When the materials are dissimilar, coupling between normal and tangential effects is proportional to Dundurs' constant, β, which cannot exceed the value 0.5. We, therefore, present results for this limiting value to identify the greatest effect the dissimilarity in materials can have. Curve 2 in Figs. 3 and 4 corresponds to the same loading condition as in Cattaneo's problem—i.e., normal force, P, is applied first, after which the horizontal force is applied gradually keeping P constant. By contrast, curve 3 corresponds to the case where the two forces are increased simultaneously, maintaining a constant ratio of $H/P = 0.3$. Notice that the differences between these curves shows the significant effect of loading history on the traction distributions and also on the extent of the stick and slip zones (indicated by arrows in Fig. 4). The behavior of this system under various loading conditions will be discussed at greater length in a subsequent paper.

Figure 3 also shows that the coupling causes the normal contact pressure to differ significantly from the Hertzian distribu-

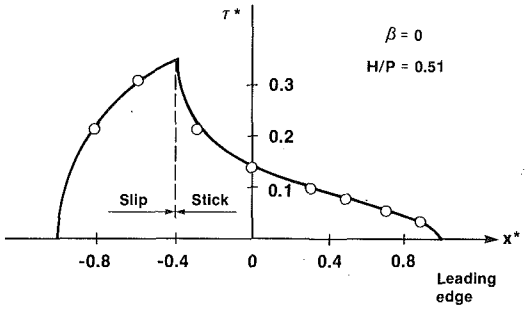


Fig. 5 Shear stress distribution for rolling of two similar materials. Comparison with Carter and Kalker: solid lines—present results and Kalker's (1971); circles—Carter's (1926).

tion (curve 1), thus casting doubt on the accuracy of solutions which neglect this coupling. For comparison, we show in Fig. 4 (curve 4), the shear tractions obtained for the problem of curve 3 when this coupling term is neglected.

We now proceed with some examples of rolling contact. Steady state rolling of two cylinders of similar materials was investigated by Carter (1926) and independently by others (Fromm, 1927; Föeppl, 1947; Chartet, 1947; Poritsky, 1950). The Carter results are compared with the present results in Fig. 5. Kalker's solution (1971), and ours obtained with twenty nodes to test the program, are practically identical. Notice that in the present algorithm, the steady state solution is obtained as the limiting behavior of a transient problem in which the forces P, H are first applied, and rolling is then permitted to occur by small increments until the traction distributions reach steady state.

Steady state rolling of cylinders of different materials without traction ($H = 0$) was studied by Bentall and Johnson (1967) and by Kalker (1971) omitting the coupling term with β in equation (20). Note that with this approximation, the solution depends only upon the ratio β/μ. An example of the shear traction distribution is shown in Fig. 6. All three results, including ours with the same assumptions, are very close. Removal of this assumption produces noticeable changes in the shear traction and the normal pressure, as shown by the dotted curves in Figs. 6 and 7, respectively.

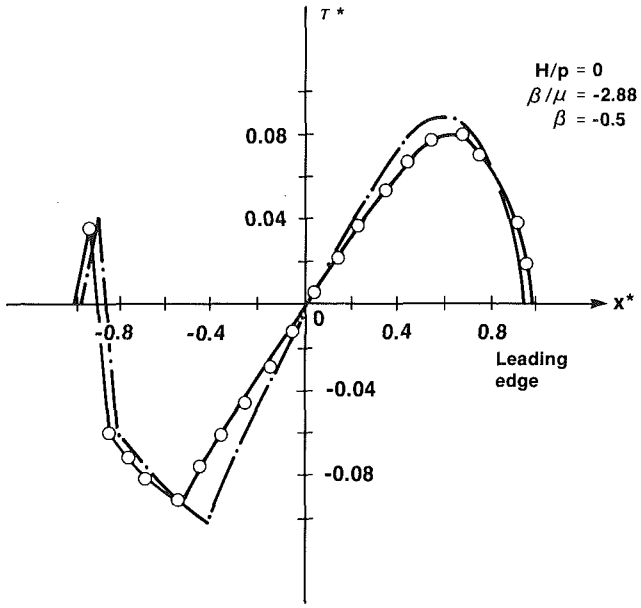


Fig. 6 Shear stress distribution for rolling of two dissimilar materials; $H/P = 0$, $\beta/\mu = -2.88$. Comparison with Bental-Johnson and Kalker: solid line—Bentall-Johnson's; circles—present results without coupling; broken line—present results with coupling

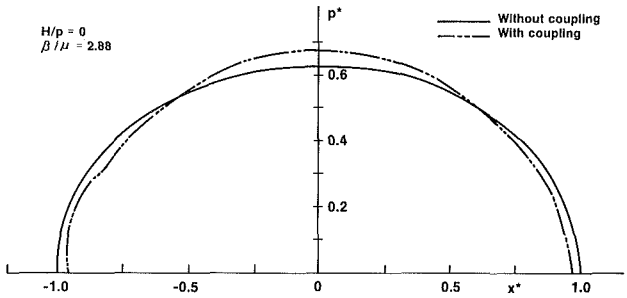


Fig. 7 Normal pressure distribution for rolling of two dissimilar materials; $H/P = 0$, $\beta/\mu = -2.88$

Conclusions

The above examples demonstrate that the proposed algorithm gives good numerical accuracy in comparison with previous analytical and numerical solutions of a wide variety of problems for normal and tangential loading of two contacting cylinders with friction. It is also found to be very efficient in the use of computer time.

The topics of this presentation are history-dependent problems, coupling effect between normal and tangential contact tractions, the extent of slip and stick area, and rolling problems. But some other related problems are covered by the algorithm, e.g., sliding problems (Hills and Sackfield, 1985; Ahmadi et al., 1983). In this case, the problem is simplified since the shearing contact stresses are then defined by the friction law $|\tau| = \mu p$ rather than by equation (6). The developed computer program does this automatically if one prescribes $H = \mu P$.

Results for some previously unsolved problems involving the full coupling between normal and tangential contact tractions indicate that such coupling can have a significant effect on the solution, particularly in regard to the extent of slip and stick zones within the contact area. This is a matter of some importance in many practical systems, where microslip can be the cause of local wear damage, and it will be investigated in more detail in a subsequent paper.

Acknowledgments

The suggestions made by J. R. Barber significantly improved this manuscript.

References

Ahmadi, H., Keer, L. M., and Mura, T., 1983, "Non-Hertzian Contact Stress Analysis for an Elastic Half-Space – Normal and Sliding Contact," *International Journal Solids Structures*, Vol. 19, No. 4, pp. 357–373.

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. 389–409.

Carter, F. W., 1926, "On the Action of a Locomotive Driving Wheel," *Proceedings of the Royal Society*, Series A, Vol. 112, pp. 151–157.

Cataneo, C., 1938, "Sul Contatto di Due Corpi Elastici: Distribuzione Locale Degli Sforzi," *Rendiconti dell' Accademia Nazionale dei Lincei*, Vol. 27, Series 6, pp. 342–348.

Chartet, A., 1947, "Propertiétés Générales de Contacts de Roulement," *Comptes. rend. Acad. Sci.*, Vol. 225, pp. 986.

Comminou, M., and Barber, J. R., 1983, "Frictional Slip Between a Layer and a Substrate Due to a Periodic Surface Force," *Int. J. Solids Structures*, Vol. 19, pp. 533–539.

Dundurs, J., 1965, "Discussion," ASME JOURNAL OF APPLIED MECHANICS, No. 3, pp. 650–652.

Föeppl, L., 1947, "Die Strange Lösung die Rollende Reibung," München.

Fromm, H., 1927, "Berechnung des Schulpfes beim Rollen Deformierbaren Scheiben," *ZAMP*, Vol. 7.

Goodman, L. E., 1962, "Contact Stress Analysis of Normally Loaded Rough Spheres," Series E, ASME JOURNAL OF APPLIED MECHANICS, Vol. 29, pp. 512–522.

Hills, D. A., and Sackfield, A., 1985, "Sliding Contact Between Dissimilar Elastic Cylinders," ASME *Journal of Tribology*, Vol. 107, pp. 463–466.

Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press.

Kalker, J. J., 1970, "Transient Phenomena in Two Elastic Cylinders Rolling Over Each Other with Dry Friction," Series E, ASME JOURNAL OF APPLIED MECHANICS, Vol. 37, pp. 677–688.

Kalker, J. J., 1971, "A Minimum Principle for the Law of Dry Friction, with Application to Elastic Cylinders in Rolling Contact," Series E, ASME JOURNAL OF APPLIED MECHANICS, Vol. 38, pp. 875–887.

Maw, N., Barber, J. R., and Fawcett, J. N., 1976, "The Oblique Impact of Elastic Spheres," *Wear*, Vol. 38, pp. 101–114.

Mindlin, R. D., 1949, "Compliance of Elastic Bodies in Contact," Series E, ASME JOURNAL OF APPLIED MECHANICS, Vol. 16, pp. 259–268.

Mindlin, R. D., and Deresiewicz, H., 1953, "Elastic Spheres in Contact Under Varying Oblique Forces," ASME JOURNAL OF APPLIED MECHANICS, Vol. 75, pp. 327–344.

Mossakovski, V. I., 1954, "The Fundamental General Problem of Theory of Elasticity for Half Space with a Circular Line Separating the Boundary Conditions," *PMM*, Vol. 18, pp. 187–202. (in Russian)

Mossakovski, V. I., 1963, "Compression of Elastic Bodies Under Conditions of Adhesion," *PMM*, Vol. 27, pp. 418–427.

Poritsky, H., 1950, "Stresses and Deflections of Cylindrical Bodies in Contact," Series E, ASME JOURNAL OF APPLIED MECHANICS, Vol. 17, pp. 191.

Spence, D. A., 1968, "Self-Similar Solutions to Adhesive Contact Problems with Incremental Loading," *Proceedings of the Royal Society*, A305, pp. 55–80.

Spence, D. A., 1973, "An Eigenvalue Problem for Elastic Contact with Finite Friction," *Proceedings Cambridge Philosophical Society*, Vol. 73, pp. 249–268.

Spence, R. D., 1975, "The Hertz Contact Problem with Finite Friction," *J. Elasticity*, Vol. 5, pp. 297–319.

Turner, J. R., 1979, "The Frictional Unloading Problem on Linear Elastic Half-Space," *J. Institute Math. and its Applications*, p. 439.

Thermoelastic Solutions for Thermal Distributions Moving Over Half Space Surfaces and Application to the Moving Heat Source

M. D. Bryant
Mechanical and Aerospace Engineering,
North Carolina State University,
Raleigh, NC 27695-7910

A method is developed for obtaining fundamental thermal and thermoelastic solutions for thermal distributions moving over the surface of an elastic half space. This method uses the concept of a moving temperature wave along with a novel form of an exponential Fourier transform. The technique is developed and then demonstrated on the example of a moving heat source. Exact results are matched with results from Carslaw and Jaeger (1959) and Barber (1984).

Introduction

When two contacting bodies slide against one another, frictional power dissipation creates thermal fields in the bodies that result in thermoelastic stresses and displacements. Resulting effects of these thermal and thermoelastic fields have been documented in conference proceedings (Burton, 1980; Dow, 1982; Kennedy, 1985) and papers (Kennedy, 1984; Barber, 1969; Ju and Huang, 1982; Huang and Ju, 1985). Among these effects are thermoelastic instability and heat checking.

Estimation of the thermal and thermoelastic fields associated with friction often uses moving thermal distributions (such as heat sources) over a half space surface. Methods of calculating these fields are diverse.

Integral transforms have been the focus of the references (Huang and Ju, 1982, 1985; Ling and Mow, 1965; Mow and Cheng, 1967; Ling, 1973). Here it is often necessary to assume a large Peclet number $\mathrm{Pe} = c\ell/k$, where c is the speed of the thermal patch, ℓ is a characteristic length, and k is the thermal diffusivity of the affected body. In addition to this limitation, inversion of the resulting Fourier transforms can be tedious and difficult, often requiring numerical evaluation of infinite integrals (Huang and Ju, 1982).

Burton and colleagues (Dow and Burton, 1972, 1973; Burton et al., 1973; Kilaparti and Burton, 1978; Heckmann and Burton, 1980; Tseng and Burton, 1982; Oh et al., 1983) have used the concept of moving sinusoidal temperature waves in conjunction with thermoelastic displacement potentials (Timoshenko and Goodier, 1970). This approach postulates moving periodic thermal distributions and is particularly useful for seals (Burton et al., 1973; Kilaparti and Burton, 1978) where periodicity may apply; however, evaluation of the related Fourier series can be cumbersome, especially if the series is slowly convergent.

A third approach conceived by Barber et al. (Barber and Martin-Moran, 1982; Barber, 1984; Azarkhin and Barber, 1985; Hills and Barber, 1985) uses principles of Carslaw and Jaeger (1959) applied to thermoelasticity to develop stresses and displacements. Barber (1984) has successfully used these ideas to obtain exact solutions for the tangential stress and displacements on the surface of a half space, given a moving heat source. Although valid for all values of Pe, subsurface quantities remain unsolved. A final approach (Kennedy, 1980; Kennedy and Ling, 1974) uses finite elements to estimate the thermoelastic field.

The purpose of this paper is to derive a method for obtaining fundamental thermal and thermoelastic solutions for two-dimensional thermal distributions moving on the surface of an elastic half space. This method uses Burton's concept of a moving temperature wave; however, instead of using a Fourier series an integral transform (related to an exponential Fourier transform) is taken over the spatial frequency. No assumptions regarding Pe are made, except that speeds c are negligible compared to sound speeds of the material.

The method is then applied to a moving heat source to obtain surface and subsurface quantities. Temperatures and thermoelastic quantities are then matched with results of Carslaw and Jaeger (1959) and Barber (1984).

The "novel" Fourier transform described here differs from the approach used by Ling and associates (Ling and Mow, 1965; Mow and Cheng, 1967; Ling, 1973) in which Fourier transforms were applied over spatial coordinates. The ap-

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 17, 1987; final revision October 7, 1987.

proach in this paper alters the definition of the Fourier transform pair to render the problem more tractable.

The class of problems to be considered is shown in Fig. 1. A thermal distribution moves in the x direction with velocity $c > 0$ over the surface of an isotropic elastic half space $y \geq 0$. All thermal and thermoelastic quantities must decay for large y; in addition the surface $y = 0$ must be free of mechanical tractions.

Basic Formulations

Thermal and Elastic Potentials. Solutions to the two dimensional Fourier heat conduction equation

$$\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}=\frac{1}{k}\frac{\partial T}{\partial t}, \quad |x|<\infty,\ y>0 \tag{1}$$

with boundary conditions

$$T\to 0 \text{ as } y\to\infty \tag{2}$$

were constructed (Dow and Burton, 1972; Tseng and Burton, 1982) for moving thermal distributions in the form of sinusoidal temperature waves

$$T_s = T_o e^{-by}\sin(\kappa x' + ay). \tag{3}$$

In this treatise T_o is a constant, $\kappa \geq 0$ is the spatial frequency, and

$$x' = x - ct \tag{4}$$

has wave speed $c > 0$. For equation (3) to solve equation (1), the parameters $a \geq 0$ and $b \geq 0$ must obey the relations

$$b^2 - a^2 = \kappa^2 \tag{5a}$$

$$2ab = \beta\kappa, \tag{5b}$$

where $\beta = c/k$. Equations (5) can be solved to give

$$a=\left(\frac{\kappa}{2}\right)^{1/2}[-\kappa+\sqrt{\kappa^2+\beta^2}]^{1/2} \tag{6a}$$

$$b=\left(\frac{\kappa}{2}\right)^{1/2}[\kappa+\sqrt{\kappa^2+\beta^2}]^{1/2} \tag{6b}$$

where the positive branch of the square root must be chosen to satisfy (2).

Associated with equation (3) is the stress function (Oh et al., 1983)

$$\chi_s = \Psi - \Phi \tag{7}$$

that yields the stresses

$$\sigma_y = \frac{\partial^2 \chi_s}{\partial x^2} \tag{8a}$$

$$\tau_{xy} = -\frac{\partial^2 \chi_s}{\partial x \partial y} \tag{8b}$$

$$\sigma_x = \frac{\partial^2 \chi_s}{\partial y^2} \tag{8c}$$

The function

$$\Psi=\frac{E\alpha\, T_o}{\nu_2\kappa\beta}\left\{\left[\frac{a}{\kappa}\sin(\kappa x')+\frac{b}{\kappa}\cos(\kappa x')\right]e^{-\kappa y} - \cos(\kappa x' + ay)e^{-by}\right\} \tag{9}$$

is a thermoelastic displacement potential determined in Tseng and Burton (1982) that satisfies

$$\frac{\partial^2 \lambda}{\partial x^2}+\frac{\partial^2 \lambda}{\partial y^2}=\frac{1+\nu}{\nu_2}\alpha\, T, \tag{10a}$$

with

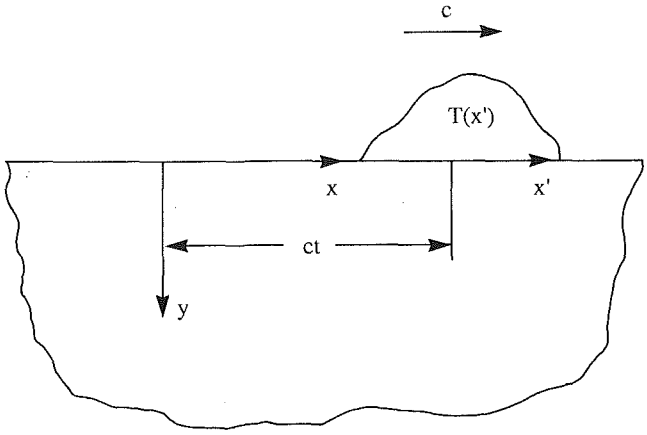


Fig. 1 A thermal distribution moves in the positive x direction with speed c across the surface of a conducting elastic half space y > 0

$$\Psi = -\frac{E}{1+\nu}\lambda. \tag{10b}$$

The parameter

$$\nu_2 = \begin{cases} 1-\nu & \text{for plane strain} \\ 1 & \text{for plane stress,} \end{cases} \tag{11}$$

E is Young's Modulus, ν is Poisson's ratio, and α represents the coefficient of thermal expansion.

The Airy stress function

$$\Phi=\frac{E\alpha T_o}{\nu_2\beta\kappa^2}[a\sin(\kappa x')+(b-\kappa)\cos(\kappa x')](1+\kappa y)e^{-\kappa y} \tag{12}$$

determined by Tseng and Burton (1982), causes

$$\sigma_y = \tau_{xy} = 0 \quad \text{for} \quad y = 0 \tag{13a,b}$$

Since

$$\chi_s(x',y)\to 0 \quad \text{as} \quad y\to 0, \tag{13c}$$

the stresses (8) also decay.

If $(x' + \pi/2\kappa)$ is substituted for x' in equations (3), (7), (9), and (12), then these equations can be phase shifted to yield corresponding results T_c and χ_c for a cosine wave. With the aid of $i = \sqrt{-1}$, the complex sums

$$T_\kappa = T_c + i\, T_s$$

$$\chi_\kappa = \chi_c + i\, \chi_s$$

can then be constructed as

$$T_\kappa(x',y) = T_o e^{-y(b-ai)+i\kappa x'} \tag{14}$$

$$\chi_\kappa(x',y)=\frac{E\alpha T_o}{\nu_2\beta}\left\{\left[\frac{i-(a+bi-i\kappa)y}{\kappa}\right]e^{-\kappa(y-ix')} - \frac{i}{\kappa}e^{-y(b-ai)+i\kappa x'}\right\}. \tag{15}$$

Equations (14) and (15) represent the thermal and thermoelastic solutions for a complex harmonic temperature wave having spatial frequency $\kappa \geq 0$. Integration of T_κ and χ_κ over all κ with $T_o = T_o(\kappa)$ gives

$$T(x',y) = \int_0^\infty T_o(\kappa)e^{-y(b-ai)+i\kappa x'}\,d\kappa \tag{16}$$

$$\chi(x',y)=\frac{E\alpha}{\nu_2\beta}\int_0^\infty T_o(\kappa)\left\{\frac{[i-(a+bi-i\kappa)y]}{\kappa}e^{-\kappa(y-ix')} - \frac{i}{\kappa}e^{-y(b-ai)+i\kappa x'}\right\}d\kappa \tag{17}$$

which satisfy the problems engendered by equations (1) and (2), and (10) and (13).

Using the definitions

$$L = L_1 + y(L_2 + L_3) \tag{18a}$$

$$L_1 = i\int_0^\infty \frac{T_o(\kappa)}{\kappa} e^{-\kappa(y-ix')} d\kappa \tag{18b}$$

$$L_2 = -i\int_0^\infty T_o(\kappa)\frac{(b-ai)}{\kappa} e^{-\kappa(y-ix')} d\kappa \tag{18c}$$

$$L_3 = i\int_0^\infty T_o(\kappa) e^{-\kappa(y-ix')} d\kappa \tag{18d}$$

and

$$M = \int_0^\infty T_o(\kappa)\left(\frac{-i}{\kappa}\right) e^{-y(b-ai)+i\kappa x'} d\kappa \tag{19}$$

equation (17) can be written as

$$\chi(x',y) = \frac{E\alpha}{\nu_2\beta}\{L + M\} \tag{20}$$

The function χ, given by equation (20), is the sum of two integrals. The first integral L, a biharmonic function given by equations (18), can be computed as a Laplace transform over κ since $y > 0$; here transform tables such as Gradshteyn and Ryzhik (1980) and Oberhettinger and Badii (1973) can be used. The second integral M has the property

$$\frac{\partial M}{\partial x'} = T(x',y); \tag{21a}$$

also

$$\Psi = \frac{E\alpha}{\nu_2\beta} M \tag{21b}$$

satisfies equations (10) and is thus a thermoelastic displacement potential that corresponds to temperatures $T(x', y)$.

Stresses. The stresses can be determined by substituting χ in equation (20) for χ_s in equations (8). Determination of σ_y and τ_{xy} are straightforward and can be written as

$$\sigma_y = \frac{E\alpha}{\nu_2\beta}\left\{\frac{\partial^2 L}{\partial x^2} + \frac{\partial}{\partial x} T(x',y)\right\} \tag{22a}$$

$$\tau_{xy} = -\frac{E\alpha}{\nu_2\beta}\left\{\frac{\partial^2 L}{\partial x \partial y} + \frac{\partial}{\partial y} T(x',y)\right\} \tag{22b}$$

where equation (4) was used for chain rule differentiation. The stress σ_x requires the determination of $\partial^2 M/\partial y^2$; use of equations (10) and (21) simplifies the calculation giving the result

$$\sigma_x = \frac{E\alpha}{\nu_2\beta}\left\{\frac{\partial^2}{\partial y^2} L - \beta\, T(x',y) - \frac{\partial}{\partial x} T(x',y)\right\} \tag{22c}$$

Note that equation (4) was again used for differentiation.

Displacements. The x and y directional displacements u and v can be calculated as

$$u = u_L - \frac{1+\nu}{\nu_2}\frac{\alpha}{\beta}\frac{\partial M}{\partial x} \tag{23a}$$

$$v = v_L - \frac{1+\nu}{\nu_2}\frac{\alpha}{\beta}\frac{\partial M}{\partial y} \tag{23b}$$

The components u_L and v_L arise from the biharmonic function L and can be determined to within rigid body motions by integrating Hooke's laws for strains $\epsilon_x = \partial u/\partial x$ and $\epsilon_y = \partial v/\partial y$. The other terms arise because M is a thermoelastic displacement potential.

Performance of the aforementioned operations convert equations (23) to

$$u = \frac{1+\nu}{\nu_2}\frac{\alpha}{\beta}\left\{(1-\nu_1)i\left[2L_2 + L_3 + y\frac{\partial}{\partial y}(L_2 + L_3)\right] - \nu_1\frac{\partial}{\partial x} L - T\right\} \tag{24a}$$

$$v = \frac{1+\nu}{\nu_2}\frac{\alpha}{\beta}\left\{(1-\nu_1)\left[L_2 + 2L_3 - y\frac{\partial}{\partial y}(L_2 + L_3)\right] - \nu_1\frac{\partial}{\partial y} L - M_y\right\}. \tag{24b}$$

In equations (24)

$$\nu_1 = \begin{cases} \nu & \text{for plane strain} \\ \dfrac{\nu}{1+\nu} & \text{for plane stress} \end{cases} \tag{25}$$

and

$$M_y = i\int_0^\infty T_o(\kappa)\frac{(b-ai)}{\kappa} e^{-y(b-ai)+i\kappa x'} d\kappa. \tag{26}$$

Thermal Boundary Conditions on $y = 0$. On the half space surface $y = 0$ the normal component q of the heat flux $\mathbf{q} = -K \nabla T|_{y=0}$ and the temperature $T(x', 0)$ are given as

$$q(x') = K\int_0^\infty T_o(\kappa)(b-ai) e^{i\kappa x'} d\kappa \tag{27a}$$

$$T(x',0) = \int_0^\infty T_o(\kappa) e^{i\kappa x'} d\kappa. \tag{27b}$$

Equations (27) are similar to Fourier transforms; using

$$T_o(\kappa)K(b-ai) = \frac{1}{\pi}\int_{-\infty}^\infty q(x') e^{-i\kappa x'} dx' \tag{28a}$$

$$T_o(\kappa) = \frac{1}{\pi}\int_{-\infty}^\infty T(x',0) e^{-i\kappa x'} dx' \tag{28b}$$

as "inverse" transforms, it can be shown that the real parts (Re) of $q(x')$ and $T(x', 0)$ can be recovered through the forward and "inverse" transformation processes if the imaginary parts of $q(x')$ and $T(x', 0)$ are zero. Thus for the rest of this paper it will be assumed that both $q(x')$ and $T(x', y)$ are real.

General Problem Solution

The linearity of the previous equations insures that if either $q(x')$ or $T(x', 0)$ are specified on $y = 0$, then the corresponding stresses (σ_y, τ_{xy}, σ_x), the displacements (u, v), and the temperature T can be determined by taking the real parts (Re) of equations (22), (24), and (16), respectively.

If the temperature $T(x', y)$ is known *a priori* by some other method, then the stresses and the displacement u can be computed by taking the real parts of equations (22) and (24*a*).

Moving Heat Source. To demonstrate the technique and to show how temperatures $T(x', y)$ and integral M_y may be calculated via contour integration, the solution for a heat source moving with velocity c on the insulated half space surface $y = 0$ will be determined.

The boundary condition on $y = 0$ for the moving line (point) heat source having strength Q can be expressed as

$$q(x') = Q\delta(x'), \tag{29}$$

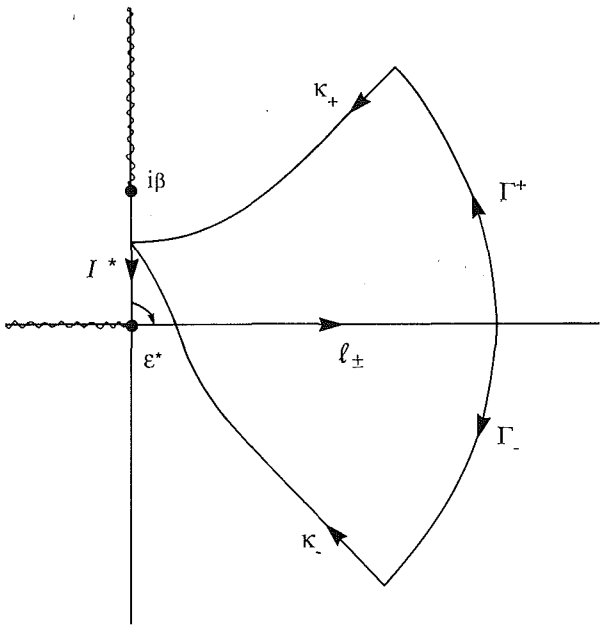


Fig. 2 The closed contours C_+ and C_- are shown in the complex κ plane. On the curves κ_+ and κ_-, κ is real.

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

Substitution of equation (29) into equation (28*a*) leads to the result

$$T_o(\kappa) = \frac{Q}{K\pi} \frac{1}{(b-ai)}. \tag{30}$$

Equation (30) can now be substituted into equations (16), (18), and (26). Before doing this, the factor $(b - ai) = [(b - ai)^2]^{1/2}$ will be written in the equivalent forms

$$(b-ai) = [\kappa(\kappa - i\beta)]^{1/2} \tag{31a}$$

$$b-ai = \left(\frac{\kappa}{2}\right)^{1/2} \left\{\left[\kappa+\sqrt{\kappa^2+\beta^2}\right]^{1/2} - i\left[-\kappa+\sqrt{\kappa^2+\beta^2}\right]^{1/2}\right\} \tag{31b}$$

using equations (5) and (6), respectively. Note that in equations (31), the positive branch of the square root must always be used to satisfy equation (2).

If equation (30) is substituted into equation (16), and if equation (31*a*) is substituted into the result, then the temperature can be expressed as

$$T(x',y) = \mathrm{Re}\left\{\frac{Q}{\pi K}\int_0^\infty \frac{e^{-y\sqrt{\kappa(\kappa-i\beta)}+i\kappa x'}}{\sqrt{\kappa(\kappa-i\beta)}}\, d\kappa\right\}. \tag{32}$$

The integral of equation (32) will be simplified using contour integration over the closed curves

$$C_\pm = \ell_\pm U \Gamma_\pm U \kappa_\pm U I^* U \epsilon^* \tag{33}$$

shown in Fig. 2. The subscripts $\pm$ refer to integration around the upper closed curve C_+ when $x' > 0$, and integration around the lower closed curve C_- when $x' < 0$. The symbol U means set union. The integrand of equation (32) is analytic inside $C_\pm$. Branch points exist at $\kappa = 0$ and $\kappa = i\beta$; corresponding branch cuts are placed on the negative real axis and positive imaginary axis, respectively. The branch of $\sqrt{\kappa(\kappa-i\beta)}$ that is positive on the positive real axis was used to satisfy equation (2).

Contributions to equations (32) arising from $\Gamma_\pm$ and ϵ^* vanish in the limit. Over the curves $\kappa_\pm$ the argument of the exponential becomes the real, negative variable $-s$, where

$$s = y\sqrt{\kappa(\kappa-i\beta)} - i\,\kappa\,x' \tag{34}$$

has the range of values

$$s_o \le s < \infty, \tag{35}$$

with

$$s_o = \frac{\beta}{2}[x' + \sqrt{(x')^2+y^2}]. \tag{36}$$

Over the contours $\kappa_\pm$, κ assumes the values

$$\kappa(s) = \frac{i[\beta y^2 + 2x's] + y\sqrt{4s^2 - 4\beta x's - \beta^2 y^2}}{2[(x')^2+y^2]}. \tag{37}$$

The contours $\kappa_\pm$ both intersect the positive imaginary axis at the same point

$$\kappa(s_o) = \frac{i\,\beta}{2}\left\{1 + \frac{x'}{\sqrt{(x')^2+y^2}}\right\} < i\beta \tag{38}$$

for $y > 0$.

On the contour I^*, where $\kappa = i\mu$ and $0 < \epsilon \le \mu \le [\kappa(s_o)/i]$, the contribution I_μ to equation (32) is purely imaginary. Since the real part of the integral is desired, I_μ contributes nothing to $T(x', y)$.

Use of the Cauchy integral theorem along with the above information yields

$$T(x',y) = \frac{Q}{\pi K}\int_{s_o}^\infty \frac{2e^{-s}ds}{\sqrt{4s^2 - 4\beta x's - \beta^2 y^2}}, \tag{39}$$

which can be transformed to the form

$$T(x',y) = \frac{Q}{\pi K}\, e^{-s_o}\int_0^\infty \frac{e^{-s}ds}{\sqrt{s(s+2s_o-\beta x')}}. \tag{40}$$

The integral of equation (40) can be evaluated using integral tables (Oberhettinger and Badii, 1973) to yield the final result

$$T(x',y) = \frac{Q}{\pi K}\, T_\delta(x',y) \tag{41a}$$

$$T_\delta(x',y) = e^{-\frac{\beta x'}{2}} K_o\left(\frac{\beta}{2}\sqrt{(x')^2+y^2},\right) \tag{41b}$$

where K_o is the modified Bessel's function of the second kind of order zero. The result of equations (41) is identical—to within a factor of 2—to that contained in Carslaw and Jaeger (1959), Section 10.7, when the directions of heat source motion are made equivalent. The factor of 2 represents the fact that in Carslaw and Jaeger (1959) the solution was developed for an infinite space where heat $Q/2$ flows into each of two symmetric semispaces.

In equations (41), the subscript δ indicates that $T_\delta(x', y)$ pertains to the moving δ distribution of heat.

Determination of Integral L. With the use of equations (30) and (31), the integrals L_1, L_2, and L_3 of equations (18*b*, *c*, *d*) can be written as

$$L_1 = \frac{Q}{\pi K}\frac{i}{2^{1/2}}\int_0^\infty \kappa^{-1}\frac{e^{-\kappa(y-ix')}}{\kappa^{1/2}\sqrt{\kappa^2+\beta^2}}$$

$$\left\{\left[\kappa+\sqrt{\kappa^2+\beta^2}\right]^{1/2} + i\left[-\kappa+\sqrt{\kappa^2+\beta^2}\right]^{1/2}\right\} d\kappa \tag{42a}$$

$$L_2 = -\frac{Q}{\pi K}\, i\int_0^\infty \frac{e}{\kappa}^{-\kappa(y-ix')}\, d\kappa \tag{42b}$$

$$L_3 = \frac{Q}{\pi K}\frac{i}{2^{1/2}}\int_0^\infty \frac{e^{-\kappa(y-ix')}}{\kappa^{1/2}\sqrt{\kappa^2+\beta^2}}$$

$$\left\{\left[\kappa+\sqrt{\kappa^2+\beta^2}\right]^{1/2} + i\left[-\kappa+\sqrt{\kappa^2+\beta^2}\right]^{1/2}\right\} d\kappa \tag{42c}$$

These integrals can be solved as Laplace transforms with the aid of Oberhettinger and Badii (1973) to produce

$$L_1 = -\frac{Q}{2K}\int_{y-ix'}^{\infty} e^{\frac{-i\beta s}{2}} H_o^{(1)}\left(\frac{\beta s}{2}\right) ds \tag{43a}$$

$$L_2 = -\frac{Qi}{\pi K}\int_{y-ix'}^{\infty} \frac{ds}{s} = \frac{Qi}{\pi K} \log s \Big|_{\infty}^{y-ix'} \tag{43b}$$

$$L_3 = -\frac{Q}{2K} e^{\frac{-i\beta}{2}(y-ix')} H_o^{(1)}\left\{\frac{\beta}{2}(y-ix')\right\}. \tag{43c}$$

Here the nth order Hankel function of the first kind

$$H_n^{(1)}(s) = J_n(s) + i\, Y_n(s) \tag{44}$$

depends on the nth order Bessel functions of the first and second kind. Also note that although L_2 is unbounded, its derivatives exist.

Stresses. Substitution of equations (41), (43), and (18*a*) into equations (22) along with taking the real part (Re) gives the stresses

$$\sigma_y = \frac{E\alpha Q}{\pi \nu_2 \beta K} \operatorname{Re}\{S_1 + S_2 + T_{\delta x}\} \tag{45a}$$

$$\tau_{xy} = -\frac{E\alpha Q}{\pi \nu_2 \beta K} \operatorname{Re}\{iS_1 + S_3 + T_{\delta y}\} \tag{45b}$$

$$\sigma_x = \frac{E\alpha Q}{\pi \nu_2 \beta K} \operatorname{Re}\{-S_1 + S_2 + 2iS_3 - \beta T_\delta(x',y) - T_{\delta x}\} \tag{45c}$$

where

$$S_1 = y\left\{\frac{i(y+ix')^2}{(x'^2+y^2)^2} + \frac{\beta\pi}{4} e^{-i\frac{\beta}{2}(y-ix')}\left\{-\beta H_o^{(1)}\left[\frac{\beta}{2}(y-ix')\right] + \left[i\beta + \frac{y+ix'}{x'^2+y^2}\right] H_1^{(1)}\left[\frac{\beta}{2}(y-ix')\right]\right\}\right\} \tag{46a}$$

$$S_2 = \frac{i\beta\pi}{4} e^{-i\frac{\beta}{2}(y-ix')}\left\{H_o^{(1)}\left[\frac{\beta}{2}(y-ix')\right] - i\, H_1^{(1)}\left[\frac{\beta}{2}(y-ix')\right]\right\} \tag{46b}$$

$$S_3 = \frac{y+ix'}{(x'^2+y^2)} \tag{46c}$$

$$T_{\delta x} = -\frac{\beta}{2} e^{-\frac{\beta x'}{2}}\left\{K_o\left[\frac{\beta}{2}\sqrt{(x')^2+y^2}\right] + \frac{x'}{\sqrt{(x'^2)+y^2}} K_1\left[\frac{\beta}{2}\sqrt{(x')^2+y^2}\right]\right\} \tag{47a}$$

$$T_{\delta y} = -\frac{\beta}{2} e^{-\frac{\beta x'}{2}} \frac{y}{\sqrt{(x')^2+y^2}} K_1\left[\frac{\beta}{2}\sqrt{(x')^2+y^2}\right]. \tag{47b}$$

For $y = 0$ equations (45*a*) and (45*b*) yield $\sigma_y = \tau_{xy} = 0$. Equation (45*c*) for $y = 0$ yields a result identical to Barber's (1984) equation (38) for plane strain conditions, provided that one uses the identities (Watson, 1966)

$$K_o(z) = -\frac{\pi}{2} i\, H_o^{(1)}(iz) \tag{48a}$$

$$K_1(z) = -\frac{\pi}{2} H_1^{(1)}(iz). \tag{48b}$$

Here $-\pi < arg\ (z) < 0$, so that the Laplace transform of equation (42*c*) exists, and

$$K_o(-r) = K_o(r) + i\pi\, I_o(r) \tag{48c}$$

$$K_1(-r) = -K_1(r) + i\pi\, I_1(r). \tag{48d}$$

In equations (48*c*) and (48*d*), $r > 0$. The functions $I_o(r)$ and $I_1(r)$ are modified Bessel functions of the first kind.

Displacements. The displacements may be obtained via equations (24), (41), and (43) as

$$u = \frac{1+\nu}{\nu_2}\frac{\alpha}{\beta}\frac{Q}{\pi K} \operatorname{Re}\{(1-\nu_1)i[2U_2 + U_3 + y\, U_{23}] - \nu_1 U_x - T_\delta\} \tag{49a}$$

$$v = \frac{1+\nu}{\nu_2}\frac{\alpha}{\beta}\frac{Q}{\pi K} \operatorname{Re}\{(1-\nu_1)[U_2 + 2U_3 - y\, U_{23}] - \nu_1 U_y - M_y^*\} \tag{49b}$$

with

$$U_2 = i\left\{\log\frac{\beta}{2}\sqrt{(x')^2+y^2} + i \arctan(-x'/y) + A\right\} \tag{50a}$$

$$U_3 = -\frac{\pi}{2} e^{-i\frac{\beta}{2}(y-ix')} H_o^{(1)}\left[\frac{\beta}{2}(y-ix')\right] \tag{50b}$$

$$U_{23} = S_2 + i\, S_3 \tag{50c}$$

$$U_x = i\, U_3 - i\, y\, U_{23} \tag{50d}$$

$$U_y = U_2 + y\, U_{23} \tag{50e}$$

and

$$M_y^* = \arccos\left(\frac{x'}{\sqrt{(x')^2+y^2}}\right) - \frac{\pi}{2} + G(x',y). \tag{51}$$

In equation (50*a*), A is an arbitrary constant representing rigid body translations; A contains the unbounded constant part of L_2 given by (43*b*). The expression for M_y^* given by equation (51) is derived in Appendix 1, and the function $G(x', y)$ is also given in Appendix 1 as equation (*A*1.5).

To within rigid body motions under plane strain conditions, the displacements of equations (49) for $y = 0$ are identical to Barber's (1984) surface displacements, equations (35) to (37) in that reference. Here for $x' > 0$ the imaginary part of equation (48*c*) becomes important when evaluating (49*b*).

Discussion and Conclusions

A method for obtaining fundamental thermal and thermoelastic solutions for thermal distributions moving over the surface of an elastic half space has been developed. This method recasts the problem into that of determining the potential functions of equations (16), (18), and (19). Computation of and relations between temperature (16), stresses (22), and displacements (24) follow.

Using this method, the thermal and thermoelastic field due to a moving heat source was determined in closed form. This solution, when integrated as a Green's function, could result in exact solutions to other moving thermal distributions or could provide the basis for numerical solution methods based upon numerical integration. These numerical integration methods would likely be simpler and more manageable than numerical transform methods (Huang and Ju, 1982).

Acknowledgments

Research support from the National Science Foundation grant number MEA-8410524 (Tribology program) is gratefully acknowledged.

References

Azarkhin, A., and Barber, J. R., 1985, "Transient Thermoelastic Contact Problem of Two Sliding Half-Planes," *Wear*, Vol. 102, pp. 1–13.

Barber, J. R., 1969, "Thermoelastic Instabilities in the Sliding of Conforming Solids," *Proc. Roy. Soc., Series A*, Vol. 312, pp. 381–394.

Barber, J. R., 1984, "Thermoelastic Displacements and Stresses Due to a Heat Source Moving Over the Surface of a Half Plane," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 636–640.

Barber, J. R., and Martin-Moran, C. J., 1982, "Green's Functions for Transient Thermoelastic Contact Problems for the Half-Plane," *Wear*, Vol. 79, pp. 11–19.

Burton, R. A., 1980, *Thermal Deformation in Frictionally Heated Systems*, Elsevier Sequoia S. A., New York.

Burton, R. A., Nerlikar, V., and Kilaparti, S. R., 1973, "Thermoelastic Instability in a Seal-Like Configuration," *Wear*, Vol. 24, pp. 177–188.

Carslaw, H. S., and Jaeger, J. C., 1959, *Conduction of Heat in Solids*, 2nd Ed., Oxford University Press, New York, pp. 266–270.

Dow, T. A., 1982, *Thermomechanical Effects in Wear*, Elsevier Sequoia S. A., New York.

Dow, T. A., and Burton, R. A., 1972, "Thermoelastic Instability of Sliding Contact in the Absence of Wear," *Wear*, Vol. 19, pp. 315–328.

Dow, T. A., and Burton, R. A., 1973, "The Role of Wear in the Initiation of Thermoelastic Instabilities of Rubbing Contact," ASME *Journal of Lubrication Technology*, Vol. 95, pp. 71–75.

Gradshteyn, I. S., and Ryzhik, I. M., 1980, *Tables of Integrals, Series, and Products*, Academic Press, New York, pp. 316, 573.

Heckmann, S. R., and Burton, R. A., 1980, "Surface Displacements for High Speed Rubs," *Wear*, Vol. 59, pp. 61–77.

Hills, D. A., and Barber, J. R., 1985, "Steady Motion of an Insulating Rigid Flat-Ended Punch Over a Thermally Conducting Half-Plane," *Wear*, Vol. 102, pp. 15–22.

Huang, J. H., and Ju, F. D., 1985, "Thermomechanical Cracking Due to Moving Frictional Loads," *Wear*, Vol. 102, pp. 81–104.

Ju, F. D., and Huang, J. H., 1982, "Heat Checking in the Contact Zone of a Bearing Seal (A Two-Dimensional Model of a Single Moving Asperity)," *Wear*, Vol. 79, pp. 107–118.

Kennedy, F. E., 1980, "Thermomechanical Phenomena in High Speed Rubbing," *Wear*, Vol. 59, pp. 149–163.

Kennedy, F. E., 1982, "Thermocracking of a Mechanical Face Seal," *Wear*, Vol. 79, pp. 21–36.

Kennedy, F. E., 1984, "Thermal and Thermomechanical Effects in Dry Sliding," *Wear*, Vol. 100, pp. 453–476.

Kennedy, F. E., 1985, *Thermomechanical Effects in Sliding Systems*, Elsevier Sequoia S. A., New York.

Kennedy, F. E., and Ling, F. F., 1974, "A Thermal, Thermoelastic, and Wear Simulation of a High-Energy Sliding Contact Problem," ASME *Journal of Lubrication Technology*, Vol. 96, pp. 497–508.

Kilparti, S. R., and Burton, R. A., 1978, "The Thermoelastic Patch Contact Problem for Large Peclet Number," ASME *Journal of Lubrication Technology*, Vol. 100, pp. 65–69.

Ling, F. F., 1973, *Surface Mechanics*, Wiley, New York.

Ling, F. F., and Mow, V. C., 1965, "Surface Displacements of a Convective Elastic Half-Space Under an Arbitrarily Distributed Fast-Moving Heat Source," ASME *Journal of Basic Engineering*, Vol. 87, pp. 729–734.

Luke, Y. L., 1962, *Integrals of Bessel Functions*, McGraw-Hill, New York, pp. 240–241.

Mow, V. C., and Cheng, H. S., 1967, "Thermal Stresses in an Elastic Half Space Associated with an Arbitrarily Distributed Moving Heat Source," *Z. Angew. Math. Phys.*, Vol. 18, pp. 500–507.

Oberhettinger, F., and Badii, L., 1973, *Tables of Laplace Transforms*, Springer-Verlag, New York, pp. 4, 12, 20.

Oh, Seik, Bryant, M. D., and Burton, R. A., 1983, "The Complete Stress Field Produced by Frictional Heating of a Thin Conductor Sliding Against an Insulator," *Journal of Thermal Stresses*, Vol. 6, pp. 307–321.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw-Hill, New York, pp. 476–478.

Tseng, M., and Burton, R. A., 1982, "Thermal Stress in a Two-Dimensional (Plane Stress) Half Space for a Moving Heat Input," *Wear*, Vol. 79, pp. 1–9.

Watson, G. N., 1966, *A Treatise on the Theory of Bessel Functions*, 2nd Ed., Cambridge University Press, New York, pp. 78, 80, 388.

APPENDIX 1

Calculation of M_y. The integral M_y obtained through substitution of equation (30) into equation (26) will be determined via contour integration around the contour $C_\pm$ defined by equation (33). As before for similar reasons, no contributions arise from $\Gamma_\pm$ and I^*. Integration over ϵ^* extracts the contribution

$$I_\epsilon = -\frac{Q}{\pi K}\frac{\pi}{2}, \tag{A1.1}$$

and integration over $\kappa_\pm$ yields

$$I_\kappa = \frac{Q}{\pi K}\beta y\int_{s_o}^{\infty}\frac{e^{-s}\,ds}{s\sqrt{4s^2-4\beta x' s-\beta^2 y^2}}. \tag{A1.2}$$

The integral I_κ of equation (A1.2) also contained an imaginary part which was dropped (only real parts are needed).

Equation (A1.2) can be transformed to the equivalent forms

$$I_\kappa = \frac{\beta Q}{2\pi K}e^{-s_o}y\int_0^{\infty}\frac{e^{-s}\,ds}{(s+s_o)\sqrt{s(s+2s_o-\beta x')}} \tag{A1.3a}$$

$$= \frac{\beta Q}{2\pi K}y\int_1^{\infty}e^{-\frac{\beta}{2}x's}K_o\left(\frac{\beta}{2}\sqrt{(x')^2+y^2}s\right)ds. \tag{A1.3b}$$

Equation (A1.3*a*) can be interpreted as a Laplace transform the result of which is equation (A1.3*b*); this transform was determined using the results of equations (40) and (41) to construct a first order ordinary differential equation in the integral of equation (A1.3*a*).

The integral of equation (A1.3*b*) can be split into two integrals having limits (0, ∞) and (0, 1). Using Watson (1966) for the first integral gives

$$I_\kappa = \frac{Q}{\pi K}\left\{\arccos\left(\frac{x'}{\sqrt{(x')^2+y^2}}\right)\right.$$

$$\left.-\frac{\beta y}{2}\int_0^1 e^{-\frac{\beta}{2}x's}K_o\left(\frac{\beta}{2}\sqrt{(x')^2+y^2}s\right)ds\right\} \tag{A1.4}$$

After the variable transformation $\tau = \beta/2\sqrt{(x')^2+y^2}\,s$, the term within the curly brackets of equation (A1.4) that contains the integral was expressed as a power series via results of Luke (1962). With the aid of these results, this term becomes

$$G(x',y) = -\frac{y}{\sqrt{(x')^2+y^2}}\left\{2e^{-\frac{\beta}{2}x'}\sum_{n=0}^{\infty}U_n(\cos\theta)\right.$$

$$\left[\sum_{m=0}^{\infty}\frac{\left(\frac{\beta}{4}\sqrt{(x')^2+y^2}\right)^{2m+n+1}}{m!(m+n+1)!}[\psi(m+n+2)-\psi(1)]\right.$$

$$\left.\left.-\left(\gamma+\log\frac{\beta}{4}\sqrt{(x')^2+y^2}\right)I_{n+1}\left(\frac{\beta}{2}\sqrt{(x')^2+y^2}\right)\right]\right\} \tag{A1.5}$$

In equation (A1.5) the functions

$$U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta} \tag{A1.6a}$$

$$\cos\theta = \frac{x'}{\sqrt{(x')^2+y^2}} \tag{A1.6b}$$

$$\psi(m+1) = -\gamma+\sum_{j=1}^{m}\frac{1}{j} \tag{A1.7}$$

with Euler's constant $\gamma = 0.57721566490\ldots$.
The modified Bessel function $I_n(s)$ is given in Watson (1962).

Laser-Induced Heating of a Multilayered Medium Resting on a Half-Space: Part I—Stationary Source

R. Kant
IBM Research,
Almaden Research Center,
San Jose, CA 95120-6099
Assoc. Mem. ASME

Laser induced heating of a multilayered medium resting on a homogeneous half-space is considered. The transient heat transfer equation is solved by employing the Laplace transform in the time domain and the Hankel transform in the space domain (r direction). Numerical inversion of the Laplace transform is obtained by using a technique developed by Crump. For the time intervals of interest, inversion of the Hankel transform is obtained by the Simpson rule. Application to magneto-optical recording is discussed.

1 Introduction

Due to increased use of composites, heat transfer in layered media has been of interest for quite some time. Numerous approaches have been taken to study heat transfer problems in composites, such as the sophisticated continuum mixture theory (Nayfeh, 1978), on one hand, and simple procedures which approximate heterogeneous medium to an "equivalent" (Christensen, 1979) medium, on the other. These theories, even though useful in some applications, impose enough constraints to preclude their application to problems of practical interest – one of which is discussed here.

With the introduction of the laser in production technology, the need for solutions which involve moving heat source and include phase change arose. In laser or conventional arc welding, moving source problems have been solved in the spatial description – that is, an observer moves with the source and it is assumed that the observer does not notice any change as the source moves along. In other words, only the convective part in the spatial derivative is retained (Rosenthal, 1964). Based on this approach, Modest and Abakians (1985) have solved a phase change problem in a semi-infinite body.

The importance of numerical solutions of this problem must also be pointed out. Tsai and Rubinsky (1984) and Salcudean et al. (1986) have attempted these problems. Grief uses explicit and implicit finite difference methods and Rubinsky uses a so-called "front tracking," finite element method. These solutions are very useful. As a word of caution, we might add that due to very fine layered structure of the medium used in optical recording, these methods consume excessive cpu times even on very fast machines, like IBM's 3081 processor. This problem is also attempted by Iravani and Wickramsinghe (1985). In a nearly similar treatment, these authors assume that the source term is harmonic in time and consequently the temperature at any plane parallel to $z=0$ is harmonic. This simplifies the problem considerably. It is difficult to comment on the applicability of their method to the problem under consideration. However, it can be pointed out that the frequency of the material response to a harmonic thermal excitation will necessarily be different, depending upon the material characteristics and configuration. Therefore, in the opinion of this author, the assumption is restrictive and has been removed. Furthermore, the results of Iravani and Wickramsinghe (1986), especially Fig. 10, indicate cooling process after the source is turned off is discontinuous in the sense that it is slower in the beginning and faster later on. It can be argued on the physical grounds that (Newton's law of cooling) initial cooling rate must be faster than later. Results presented here do not indicate such behavior.

Keeping this in view, we have developed an analysis for use in optical recording problems which do not involve the phase change. The purpose of this study is to show, at least qualitatively, how the medium will behave in the event of the phase change should occur and precisely, when the phase change does not occur.

2 Statement of the Problem

We consider a medium consisting of $n-1$ layers resting on a homogeneous half-space. One of these layers is called the "active layer." A known amount of heat is deposited on the plane $z_N=0$ (see below for the definition of coordinate systems) by a laser pulse of duration T at time zero. The problem is to find a temperature field $u_i(r,t,z)$ satisfying the equation

$$\frac{\partial^2 u_i}{\partial r^2}+\frac{1}{r}\frac{\partial u_i}{\partial r}+\frac{\partial^2 u_i}{\partial z^2}=\frac{1}{\alpha_i}\frac{\partial u_i}{\partial t}$$

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

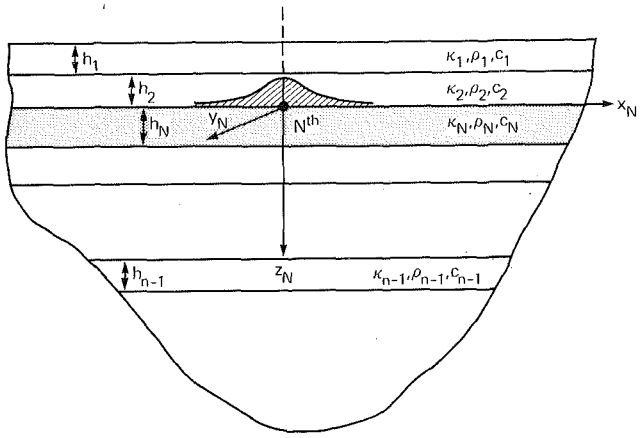


Fig. 1 Geometry of the problem and the coordinate systems

$$\alpha_i = \frac{k_i}{\rho_i c_i} \quad i = 1, \ldots, n \tag{1}$$

and the following initial, boundary continuity and regularity conditions

$$u_i(r,z,0) = 0 \tag{2}$$

$$k_1 \frac{\partial u_1}{\partial z} = 0 \quad \text{for all } \theta \text{ and } z_1 = 0 \tag{3}$$

$$\left.\begin{aligned} & u_i = u_{i+1} \\ & i = 1, \ldots, n-1 \\ & k_i \frac{\partial u_i}{\partial z} = k_{i+1} \frac{\partial u_{i+1}}{\partial z} \\ & i = 1, \ldots, n-1 \text{ and } \neq N \end{aligned}\right\} \tag{4}$$

$$u_n \to 0 \quad \text{as} \quad z \to \infty \tag{5}$$

and when $i = N$

$$k_{N-1} \frac{\partial u_{N-1}}{\partial z} - k_N \frac{\partial u_N}{\partial z} - q = 0 \tag{6}$$

where N designates the active layer number and q characterizes the source strength.

$$q = \begin{cases} p_0 e^{-\frac{r^2}{2c_r^2}} & 0 \le t \le T \\ 0 & t > T \end{cases} . \tag{7}$$

In equation (7), p_0 is the flux at the center of the laser beam and c_r is the effective laser radius. And finally, k_i, ρ_i and c_i are the conductivity, density, and specific heat of the material of the layer i.

In equations (1) through (7) and elsewhere, it is implied that each layer has its own coordinate system, embedded in it in such a way that the positive z direction points into the half-space and the plane formed by $r-\theta$ directions coincides with the surface of the layer whose z coordinate is assigned the value $z = 0$. In each layer, the value of z is ≥ 0. For the sake of clarity, the subscript i on the variables r, θ, z, referring to the coordinate system attached to each layer is dropped (see Fig. 1).

Equations (4) are continuity conditions for temperature and flux, respectively, at layer interfaces. In view of the comment above on the coordinate systems used in the problem, it is implied that these continuity conditions are evaluated at appropriate values of z coordinate. Equation (3) imply that the top layer also acts as an insulator and no heat escapes through the top surface. The assumption is made on physical grounds for the problem under consideration; however, the mathematical analysis presented here can accommodate any other boundary condition. Finally, equation (6) describes the conditions at the interface N, where a known amount of heat is deposited. It expresses the flux continuity condition at this interface and also accounts for the amount of heat added per unit area. It is worth mentioning here that in problems of this nature (optical recording), the incident energy is partially absorbed by all or some layers depending upon the coefficient of absorption of a given layer. In the case when partial absorption occurs at more than one interface, an appropriate number of the flux continuity equations (4) are replaced by an equation of type (6).

3 The Method

Application of integral transforms reduces the boundary and initial value problems governed by partial differential equations to ordinary differential equations. We apply the Laplace transforms in the time variable and the Hankel transform in the r variable to equations (1)–(7). These transform pairs are defined by

$$\left.\begin{aligned} \mathcal{L}\{f(t)\} &= F(s) = \int_0^\infty f(t) e^{-st} dt \\ \mathcal{L}^{-1}\{F(s)\} &= f(t) = \int_{c-i\infty}^{c+i\infty} F(s) e^{st} ds \end{aligned}\right\} \tag{8}$$

where c is chosen such that all the singularities, if any, are to the left of c, and

$$\left.\begin{aligned} \mathcal{H}\{g(r)\} &= \bar{g}(p) = \int_0^\infty g(r) J_0(pr) r dr \\ \mathcal{H}^{-1}\{\bar{g}(p)\} &= g(r) = \int_0^\infty \bar{g}(p) J_0(pr) p dp \end{aligned}\right\} . \tag{9}$$

We also define by $\bar{u}_i(p,t,z)$ and $U_i(p,s)$, the quantities

$$\left.\begin{aligned} \bar{u}_i(p,t,z) &= \mathcal{H}\{u_i(r,t,z)\} \\ U_i(p,s,z) &= \mathcal{L}\{\bar{u}_i(p,t,z)\} . \end{aligned}\right\} \tag{10}$$

Application of the above mentioned transform procedures to equations (1)–(7) and the use of the initial condition (2) yeilds

$$\frac{d^2 U_i}{dz^2} - \lambda_i^2 U_i = 0 \tag{11}$$

where

$$\lambda_i^2 = (p^2 + \alpha_i^{-1} s) \tag{12}$$

and in equations (2)–(6), u_i is replaced by U_i. The quantity q in equation (6) is replaced by Q, which is defined by (Watson, 1952):

$$\begin{aligned} Q &= \int_0^T \left[p_0 \int_0^\infty e^{-\frac{r^2}{2c_r^2}} J_0(pr) r dr \right] e^{-st} dt \\ &= p_0 \frac{[1 - e^{-sT}]}{s} c_r^2 e^{-\frac{1}{4} p^2 c_r^2} . \end{aligned} \tag{13}$$

4 The Solution

The solution of equation (11) is

$$U_i = A_i(p,s) e^{\lambda_i z} + B_i(p,s) e^{-\lambda_i z} \quad i = 1, \ldots, n \tag{14}$$

Equations (2)–(7) are $2n$ conditions for $2n$ unknowns A_i, B_i. Application of these conditions to solutions (14) results in $2n$ linear algebraic equations for functions A_i, B_i.

4.1 Inversion of Integral Transforms.[1] The solutions (14) have to be inverted to calculate the temperature values. For inversion of Laplace transform, we employ the technique developed by Crump (1976). According to this method, the inverse of a transform $U_i(p,s,z)$ is given by

$$\bar{u}_i(t,p,z) \simeq \frac{e^{at}}{\tau}\left[\frac{1}{2}\mathrm{Re}U_i(a) + \sum_{k=1}^{\infty}\left\{\mathrm{Re}U_i(a+k\pi i,p,z)\cos\frac{k\pi t}{\tau} - \mathrm{Im}U_i(a+k\pi i,p,z)\sin\frac{k\pi t}{\tau}\right\}\right] \tag{15}$$

where

$$\tau > t.$$

In equation (15), Re and Im, respectively, represent real and imaginary parts. This expression is valid for time interval $0 \le t \le \tau$. The accuracy of the inversion is determined by the expression

$$e^{-a\tau} = \epsilon \tag{16}$$

where ϵ is a desired small number. Finally, the temperature is obtained by

$$u_2(t,r,z) = \int_0^{\infty} \bar{u}_2(t,p,z) J_0(pr) p dp. \tag{17}$$

The integral (17) is evaluated by using the Simpson rule. This is justified by the fact that for time intervals of interest, the function $\bar{u}_2(s,p,z)$ is dominated by $e^{-\frac{1}{4}p^2c_r^2}$ factor and that remainder of $\bar{u}_2$ varies slowly with p. In Laplace transforms, $t \to 0$ corresponds to $s \to \infty$ and $t \to \infty$ corresponds to $s \to 0$.

5 An Example

The procedure outlined above can be employed to solve the problem of heat transfer in a n-layered medium resting on a half-space. Clearly, when the number of layers is large, one may have to employ a numerical scheme of some sort to solve the system of equations obtained after application of equations (2)–(7) to (14). For a few layers, one can solve these equations analytically. We have chosen a two-layered medium on a half-space to demonstrate the applicability of the solution procedure. The algebraic details of this example are simple, yet it retains all the essential features of the problem.

Application of the above mentioned procedure results into the following six equations

$$\begin{aligned}
&A_1 - B_1 = 0,\\
&A_1 e^{\lambda_1 h_1} + B_1 e^{-\lambda_1 h_1} = A_2 + B_2,\\
&A_1 e^{\lambda_1 h_1} - B_1 e^{-\lambda_1 h_1} = p_{21}(A_2 - B_2) + Q/k_1\lambda_1,\\
&A_2 e^{\lambda_2 h_2} + B_2 e^{-\lambda_2 h_2} = B_3,\\
&A_2 e^{\lambda_2 h_2} - B_2 e^{-\lambda_2 h_2}, = -p_{32} B_3,\\
&A_3 = 0,
\end{aligned} \tag{18}$$

where

$$p_{ij} = \frac{k_i \lambda_i}{k_j \lambda_j}. \tag{19}$$

The solution to equation (18) is:

$$A_1 = B_1 = \frac{1}{2}\left\{\left[\left(\frac{1-p_{32}}{1+p_{32}}\right)(1-p_{21})e^{-2\lambda_2 h_2} + (1+p_{21})\right]B_2 - Q/\lambda_1 k_1\right\}e^{\lambda_1 h_1},$$

$$A_2 = B_2\left(\frac{1-p_{32}}{1+p_{32}}\right)e^{-2\lambda_2 h_2}, \quad B_3 = 2B_2\frac{e^{-\lambda_2 h_2}}{(1+p_{32})} \text{ and} \tag{20}$$

$$B_2 = \frac{-Q\cosh\lambda_1 h_1}{\left\{\lambda_1 k_1 e^{-2\lambda_2 h_2}\left(\frac{1-p_{32}}{1+p_{32}}\right)(p_{21}\cosh\lambda_1 h_1 - \sinh\lambda_1 h_1) - (p_{21}\cosh\lambda_1 k_1 + \sinh\lambda_1 h_1)\right\}}. \tag{20}$$

In this example problem, the second layer is designated as the active layer. The solution for the active layer is

$$U_2(p,s,z) = A_2(p,s)e^{z_2\sqrt{p^2+\alpha_2 s}} + B_2(p,s)e^{-z_2\sqrt{p^2+\alpha_2 s}}. \tag{21}$$

To evaluate the temperature distribution throughout the layer, equation (21) must be inverted. However, if the temperature at the plane $z = 0$ (at interface) is desired, the inversion integration is somewhat simplified.

Equation (21) can be viewed as a product of two functions, one of which is $Q(p,s)$. The inverse of $Q(p,s)$ is known (7), thus it is advantageous to use the convolution theorem for the Laplace transforms. If $Q(p,s)$ and $V(p,s,z)$ are two functions and we wish to calculate the inverse of $Q(p,s) \cdot V(p,s,z)$, then (see Miles, 1971)

$$\mathcal{L}^{-1}[Q(p,s)\cdot V(p,s,z) = \int_0^t q(p,u)v(p,t-u,z)du. \tag{22}$$

In equation (21), q and v, respectively, are the inverses of Q and V. If we define $V(p,s,z)$ by

$$Q(p,s)V_{(p,s,z)} = A_2(p,s)e^{z_2\sqrt{p^2+\alpha_2 s}} + B_2(p,s)e^{-z_2\sqrt{p^2+\alpha_2 s}}. \tag{23}$$

and use expression (15) to define the inverse of $V(p,s,z)$ and then use equation (22) to obtain the Laplace inversion of $U_2(p,s,z)$, we get

$$\begin{aligned}
\bar{u}_2(p,t,z) = \sum_{k=0}^{\infty} \frac{p_0 c_r^2 e^{-0.25p^2c_r^2}}{(a^2 + \bar{\delta}_{0k}b^2)(1+\delta_{0k})} \cdot [&\mathrm{Re}V \cdot \{e^{at}(a\cos bt + b\sin bt) - a\}\\
&- \bar{\delta}_{0k}\mathrm{Im}V \cdot \{b + e^{at}(a\sin bt - b\cos bt)\}]\\
&\text{for} \quad t \le T, \text{ and}
\end{aligned} \tag{24}$$

$$\begin{aligned}
= \sum_{k=0}^{\infty} \frac{p_0 c_r^2 e^{-0.25p^2c_r^2}}{(a^2 + \bar{\delta}_{0k}b^2)(1+\delta_{0k})} \Big[&\mathrm{Re}V \cdot \Big\{e^{at}(a\cos bt + b\sin bt) - e^{a(t-T)}\{a\cos b(t-T) + b\sin b(t-T)\}\Big\}\\
&- \bar{\delta}_{0k}\mathrm{Im}V \cdot \Big\{e^{at}(a\sin bt - b\cos bt) - e^{a(t-T)} \cdot \{a\sin b(t-T) - b\cos b(t-T)\}\Big\}\Big]\\
&\text{for} \quad t > T
\end{aligned}$$

where

$$\begin{aligned}
&\mathrm{Re}V = \text{Real part of } \{V(a+\bar{\delta}_{0k}bi,z,p)\},\\
&\mathrm{Im}V = \text{Imaginary part of } \{V(a+\bar{\delta}_{0k}bi,z,p)\},\\
&b = k\pi/\tau,
\end{aligned} \tag{25}$$

δ_{0k} is the Kronecker delta and

$$\bar{\delta}_{0k} = 1 - \delta_{0k}.$$

Finally, $u_2(r,t,z)$ is obtained by substituting equation (24) in (17). This formally completes the solution. Below, we discuss results.

[1] Depending upon the desired accuracy of the solutions obtained, readers have a choice of employing various inversion techniques. See Krylov and Skoblya (1969), for example.

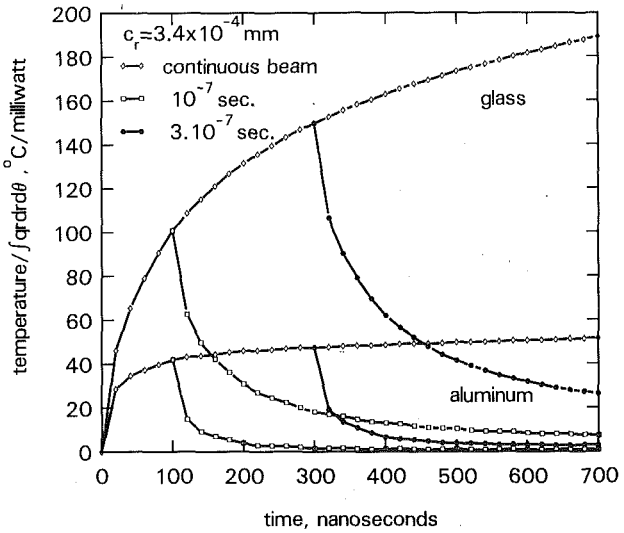


Fig. 2 Temperature at the center of the beam for laser pulses of various duration

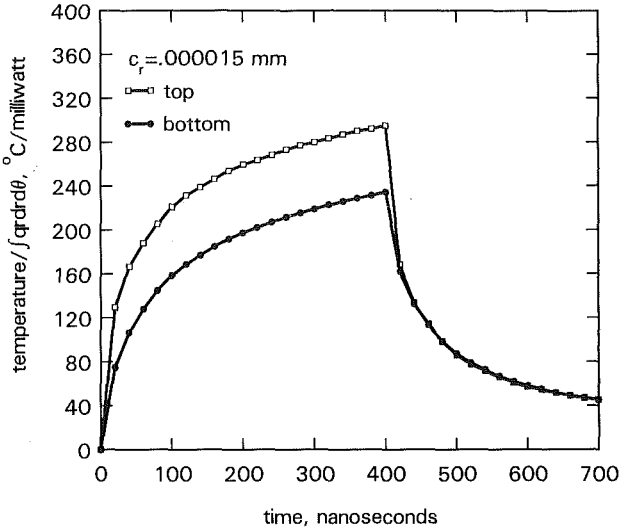


Fig. 4 Temperature at the top and the bottom of the active layer calculated at the center of the laser beam of 400 ns pulse (case 1)

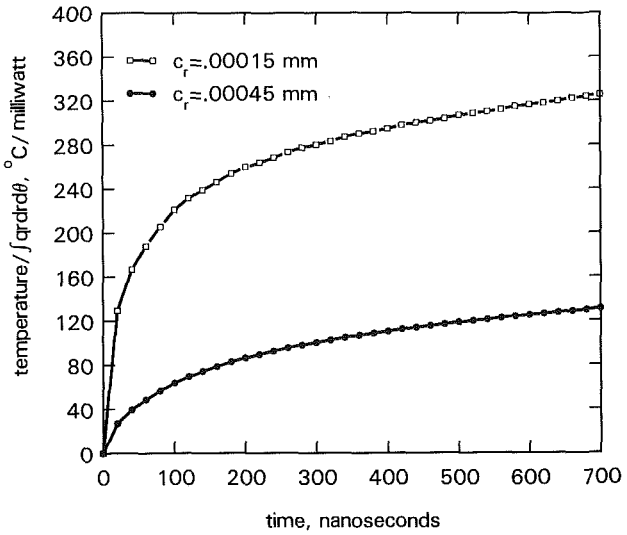


Fig. 3 Temperature at the center of the beam for various effective radii of the laser beam (case 1)

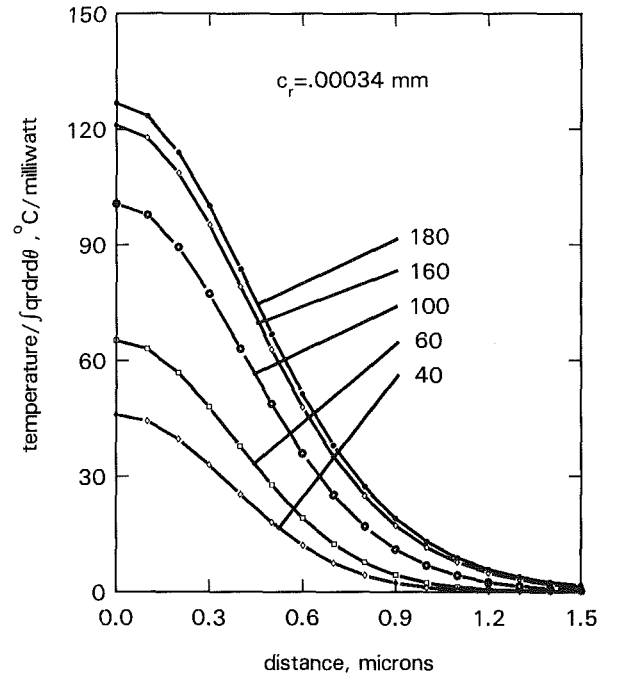


Fig. 5 Variation of temperature with distance from the center at various times (case 1)

6 Numerical Results and Discussion

In this section, we discuss results for a two-layered medium resting on a half-space. In Fig. 2, we show the temperature rise at the top surface of the active layer. The dashed lines show the case when the layers are resting on a glass medium and the solid lines for the case when they are on aluminum. Calculations of this type interest the investigators most in optical recording. First, as expected, the temperature at the center of the beam at the surface $z_N = 0$, for the case of glass half-space, is greater than the aluminum half-space. We also show the temperature in the active layer for laser pulse of various duration. We note that the temperature for the case of aluminum half-space falls to about 25 percent of its original value in the first 20 nanoseconds after the laser is turned off, whereas for the case of glass half-space, it falls to about 75 percent. Also, the use of the convolution integral (22) has enabled us to capture the time when the laser is turned off and the cooling begins. The geometry of the layers and material constants used in this example are given in Table 1.

In Fig. 3, we show the temperature rise for case 1 (see Table 1) for various effective diameters of the laser beam. The effective diameter (waist) is defined as radius where the intensity of the beam is inversely proportional to $\sqrt{e}$ (the exponential). In both these cases, the power of the beam is kept constant. Thus, the beam with lesser effective radius is more intense near the center. Therefore, we expect greater temperatures at the center of the beam for lesser effective radius. One of the purposes of such calculations is the need to demonstrate that the numerical procedures used here adapt to changes in various parameters in a stable fashion. It is also of interest to show the differences, if any, in the temperatures at the top and bottom layers of a layer. In Fig. 4, we do so for the active layer, case 1. This information is readily obtained by calculating U_3 in equation (21) in addition to U_2. This is not an additional computational task since equations (18) are solved in the very beginning. We note that even though the thickness of the active layers is very small, the temperature difference at the top and bottom could be large. We also note

Table 1

Layer No	$\rho\left(\frac{gm}{cm^3}\right)$	$c\left(\frac{Cal}{gm°C}\right)$	$k\left(\frac{Cal}{s°C\ cm}\right)$	h(Å)
1	2.2	0.18	0.0033	2000
2	2.34	0.048	0.0076	750
half-space case 1*	2.76	0.26	0.0036	∞
case 2†	2.96	0.23	3.34	∞

*Glass
†Aluminum

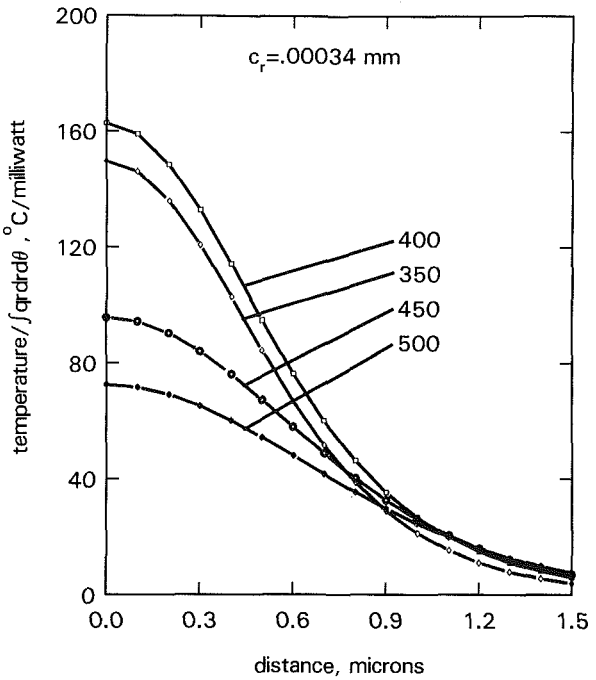


Fig. 6 Variation of temperature with distance from the center before and after the source has been turned off after 400 ns (case 1)

that the temperature of the top and bottom layers becomes identical very soon after the source has been turned off.

In Fig. 5, we show the temperature of the top surface of the active layer. We show that at a distance of 1.5 micron, the temperature is less than 10°C/milliwatt after 180 nanoseconds. Calculations of this nature are also important for optical recording since they determine the "spot" size – a critical parameter from the storage point of view. It should be remembered, however, that temperature at points, some distance away from the center, could be greater when the source is turned off before it approaches zero. This situation is shown in Fig. 6.

Acknowledgments

The author is thankful to Dr. V. Jipson for introducing this problem and to Dr. Kenneth L. Deckert for his continuous encouragement and helpful discussions. Both Drs. Jipson and Deckert are at IBM Research, Almaden Research Center, San Jose, CA. Thanks are also due to Ms. Sylvia Fujii of IBM Research Publications for preparing this manuscript.

References

Nayfeh, A. H., 1978, "Continuum Mixture Theory of Heat Transfer for Layered Medium," ASME JOURNAL OF APPLIED MECHANICS, Vol. 38.

Christensen, R. M., 1979, *Mechanics of Composite Materials*, Wiley, New York.

Rosenthal, D., 1946, *Transactions of the ASME*, Vol. 68, p. 849; also cited in Jacobs, *Heat Transfer*, Wiley, London.

Modest, and Abakians, 1985, "Evaporative Cutting of a Semi-Infinite Body with a Moving CW Laser," *Transactions of the ASME*, Paper No. 85-HT-25.

Tsai, H. L., and Runbinsky, B., 1984, "A Front Tracking Finite Element Study of Change of Phase Interface Stability During Solidification Processes in Liquids," *J. Cryst. Growth*, Vol. 70.

Salcudean, M., Choi, M., and Grief, R., 1986, "A Study of Heat Transfer during Arc Welding," *International J. Heat and Mass Transfer*, Vol. 29, No. 2.

Iravani, M. V., and Wickramsinghe, H. K., 1985, "Scattering Matrix Approach to Thermal Wave Propagation in Layered Structures," *J. Appl. Phys.*, Vol. 58, No. 1.

Crump, K. S., 1976, "Numerical Inversion of Laplace Transforms Using a Fourier Series Approximation," *J. Assoc. for Computing Industry*, Vol. 23, No. 1.

Krylov, V. I., and Skoblya, N. S., 1969, *Handbook of Numerical Inversion of Laplace Transforms*, translated from Russian, Israel Program for Scientific Translations, Jerusalem.

Miles, J., 1971, *Integral Transforms in Applied Mathematics*, Cambridge University Press, Cambridge.

Watson, G. N., 1952, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge.

Elastic Fracture Mechanics Concepts for Interfacial Cracks

J. R. Rice
Division of Applied Sciences,
Harvard University,
Cambridge, MA 02138

Elastic fracture mechanics concepts are reexamined for a crack on the interface between dissimilar solids. A derivation by function theory is given of the form of stress and displacement fields in the vicinity of the crack tip, equivalent to complete Williams expansions of both inner and outer (external to a nonlinear or contact zone) type. The complex stress intensity factor K *associated with an elastic interface crack, for which contact is ignored, is discussed and, specifically, its validity as a crack tip characterizing parameter is noted for cases of small scale nonlinear material behavior and/or small scale contact zones at the crack tip. That is, similar values of* K *for two cracked bodies then imply similar states as the crack tip, so that conditions for crack growth can be phrased in terms of* K *reaching a critical failure locus in a complex plane. The maintenance of a similar state at a crack tip under change of crack length is shown to require alteration of both the magnitude and phase angle of a combined tension and shear loading. Some possible definitions of stress intensity factors* K_I *and* K_{II} *of classical type associated with interface cracks are discussed. Also, the scaling of interface crack tip plastic zone size with load under small scale yielding conditions is found to deviate from classical scaling, in proportion to the square of the applied load level, and dependences of the field on distance from the tip and on load phase angle are found to be linked together.*

Introduction

There has been a resurgence of interest in the elastic interface crack problem, for which the characteristic oscillating stress singularity was determined by Williams (1959), and solutions to specific problems given by Cherepanov (1962) (see Cherepanov, 1979, pp. 625–630 and p. 808), England (1965), Erdogan (1965), and Rice and Sih (1965). Works by Park and Earmme (1986), Shih and Asaro (1988), and Hutchinson et al. (1987) provide examples of recent contributions.

Apparently the full form of the near tip field, in the sense of a complete Williams expansion, has not been given for the interface crack. That is, Williams (1959) gives eigenvalues λ (with stresses varying in proportion to r^λ) of the form $\lambda = n(\text{integer}) - 1/2 + i\epsilon$ for plane strain or plane stress. Here r is distance from the crack tip,

$$\epsilon = (1/2\pi)\ln[(\kappa_1/\mu_1 + 1/\mu_2)/(\kappa_2/\mu_2 + 1/\mu_1)], \qquad (1)$$

subscripts 1 and 2 refer to the materials in $y>0$ and $y<0$, respectively, as in Fig. 1, $\kappa = 3-4\nu$ for plane strain and $(3-\nu)/(1+\nu)$ for plane stress, ν = Poisson ratio, and μ = shear modulus. However, there are evidently other eigenvalues of form $\lambda = n$. The complete form of the near tip field is derived here in an analysis that is not based on the Williams (1959) product solution technique, but rather on an extension to the interface crack of an analysis by Rice (1968, pp. 214–215) based on analytic functions. The procedures follow those of Cherepanov (1962, 1979), England (1965), and Erdogan (1965) in solving specific interface crack problems.

The work also sheds light on how to interpret the elastic solutions discussed, in the spirit of linear elastic fracture mechanics procedures like those developed for cracked homogeneous solids, when there is a small zone of nonlinear material response and/or mechanical contact (Comninou, 1977a,b; Comninou and Schmueser, 1978) at the crack tip. The fact that the elastic interfacial crack solutions discussed here predict interpenetration of the crack walls near the tip is frequently taken as a reason to disregard them. However, it is

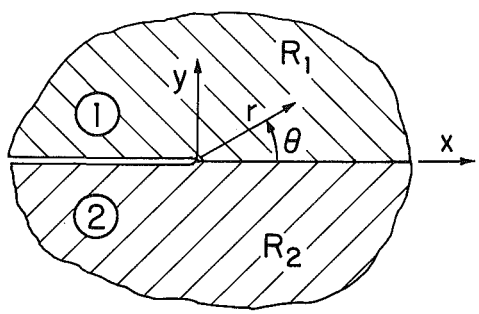


Fig. 1 Region near crack tip along bimaterial interface

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 3, 1987.

Paper No. 88-WA/APM-13.

explained that while the predicted interpenetration means the solutions must be wrong in detail on the scale of the contact zone, they do nevertheless provide a proper characterizing parameter for the near tip state in typical circumstances when that zone size is much smaller than crack length. The characterizing parameter is a complex stress intensity factor K, the same apart from a constant factor as that introduced by Sih and Rice (1964, Appendix) and Rice and Sih (1965), in which tensile and shear effects near the crack tip are intrinsically inseparable into analogues of classical mode I and mode II conditions. The inseparability is sometimes ignored (Malyshev and Salganik, 1965; Cherepanov, 1979). Such is shown to be suitable in limited circumstances but inappropriate, for sufficiently dissimilar materials, when loaded solids with substantially different crack sizes are compared.

Possible definitions of stress intensity fractors K_I and K_{II} of classical type (and physical units) for interfacial crack problems are discussed. Some of these characterize the near tip field just as does the complex K. Also, the nonclassical scaling of plastic zone size and stress field at the interface crack tip under small scale yielding is noted.

Near Tip Stress Field

The well-known Muskhelishvili representation of 2D elastic displacement (u) and stress (σ) fields in isotropic solids can be put in the form

$$2\mu(u_x+iu_y)=\kappa\phi(z)+(\bar{z}-z)\bar{\phi}'(\bar{z})-\bar{\Omega}(\bar{z}) \tag{2}$$

$$\sigma_{xx}+\sigma_{yy}=2[\phi'(z)+\bar{\phi}'(\bar{z})] \tag{3}$$

$$\sigma_{yy}-\sigma_{xx}+2i\sigma_{xy}=2[(\bar{z}-z)\phi''(z)-\phi'(z)+\Omega'(z)] \tag{4}$$

where $z=x+iy$, $\phi(z)$ and $\Omega(z)$ are analytic, $\phi'(z)=d\phi/dz$, and the overbar denotes complex conjugate. We seek the form of solution in some region $R(=R_1+R_2$, Fig. 1) surrounding a traction-free interface crack tip. Let ϕ_1, Ω_1, and ϕ_2, Ω_2 denote solutions in the two regions.

Since $(\sigma_{yy}-i\sigma_{xy})_1=(\sigma_{yy}-i\sigma_{xy})_2$ along all the interface,

$$\phi_1'(x)^+ +\bar{\Omega}_1'(x)^- =\phi_2'(x)^- +\bar{\Omega}_2'(x)^+ \tag{5}$$

everywhere along the x axis in R. Observing that since $\phi_1(z)$ and $\Omega_1(z)$ are analytic in R_1, $\bar{\phi}_1(z)$, and $\bar{\Omega}_1(z)$ are analytic in R_2, etc., this equation shows that

$$\phi_1'(z)-\bar{\Omega}_2'(z)=\phi_2'(z)-\bar{\Omega}_1'(z)=2g(z) \tag{6}$$

where $g(z)$ is analytic throughout R (including points along all the interface). The result also analytically continues the definition of $\phi_1'(z)-\bar{\Omega}_2'(z)$ to R_2 and of $\phi_2'(z)-\bar{\Omega}_1'(z)$ to R_1.

On the bonded portion of interface, $y=0$, $x>0$ in R, $(u_x+iu_y)_1=(u_x+iu_y)_2$. Thus, after differentiating both with respect to x,

$$[\kappa_1\phi_1'(x)^+ -\bar{\Omega}_1'(x)^-]/\mu_1=[\kappa_2\phi_2'(x)^- -\bar{\Omega}_2'(x)^+]/\mu_2 \tag{7}$$

on $x>0$. This equation allows analytic continuation of different linear combinations of $\phi'(z)$ and $\bar{\Omega}'(z)$ across the interface such that

$$(\kappa_1/\mu_1)\phi_1'(z)+(1/\mu_2)\bar{\Omega}_2'(z)$$

$$=(\kappa_2/\mu_2)\phi_2'(z)+(1/\mu_1)\bar{\Omega}_1'(z) \tag{8}$$

holds everywhere in R. Finally, on the cracked portion of the interface $y=0^+$, $x<0$ in R we set $(\sigma_{yy}-i\sigma_{xy})_1=0$ (by equation (5), this will also imply $(\sigma_{yy}-i\sigma_{xy})_2=0$ on $y=0^-$, $x<0$). Thus

$$\phi_1'(x)^+ +\bar{\Omega}_1'(x)^- =0 \tag{9}$$

on $x<0$ in R.

Now by using equations (6) and (8), we may express the various functions $\bar{\Omega}_1'(z)$, $\phi_2'(z)$, and $\bar{\Omega}_2'(z)$ in terms of $\phi_1'(z)$ and $g(z)$. The expressions are not written out explicitly here but when that for $\bar{\Omega}_1'(z)$ is substituted into equation (9) there results

$$(\kappa_2/\mu_2+1/\mu_1)\phi_1'(x)^+ +(\kappa_1/\mu_1+1/\mu_2)\phi_1'(x)^-$$

$$=2(\kappa_2+1)g(x)/\mu_2 \tag{10}$$

on $x<0$ in R. A homogeneous solution of this equation, exhibiting the strongest singularity compatible with bounded total strain energy, is provided by $\phi_1'(z)=z^{-1/2-i\epsilon}$, with ϵ defined by equation (1). Also, a particular solution is $\phi_1'(z)=2c_2g(z)/(c_1+c_2)$ where

$$c_1=(\kappa_1+1)/\mu_1,\ c_2=(\kappa_2+1)/\mu_2. \tag{11}$$

Thus the general solution for $\phi_1'(z)$ is

$$\phi_1'(z)=e^{-\pi\epsilon}z^{-1/2-i\epsilon}f(z)+2c_2g(z)/(c_1+c_2) \tag{12}$$

where $f(z)$, like $g(z)$, is also analytic everywhere in R. Further, by using the expressions discussed at the outset of this paragraph the other functions are given by

$$\Omega_1'(z)=e^{\pi\epsilon}z^{-1/2+i\epsilon}\bar{f}(z)-2c_2\bar{g}(z)/(c_1+c_2) \tag{13}$$

$$\phi_2'(z)=e^{\pi\epsilon}z^{-1/2-i\epsilon}f(z)+2c_1g(z)/(c_1+c_2) \tag{14}$$

$$\Omega_2'(z)=e^{-\pi\epsilon}z^{-1/2+i\epsilon}\bar{f}(z)-2c_1\bar{g}(z)/(c_1+c_2). \tag{15}$$

For a pure number ζ, $\zeta^{i\epsilon}$ may be evaluated as $e^{i\epsilon\ln\zeta}$.

Williams Expansion; Complex Stress Intensity Factor

A Williams type expansion of the near tip field is generated from equations (2)–(4) and (12)–(15) by writing f and g as local Taylor series expansions

$$f(z)=\sum_{n=0}^{\infty}a_nz^n,\ g(z)=\sum_{n=0}^{\infty}b_nz^n. \tag{16}$$

Then a_0 represents the strength of the crack tip singularity and may be written as

$$a_0=\bar{K}/2\sqrt{2\pi}\cosh(\pi\epsilon) \tag{17}$$

where K is a complex stress intensity factor which uniquely characterizes the singular field. It has been introduced, following Hutchinson et al. (1987), such that along the interface ahead of the crack tip

$$(\sigma_{yy}+i\sigma_{xy})_{\theta=0}=Kr^{i\epsilon}/\sqrt{2\pi r} \tag{18}$$

and along the crack faces

$$(u_y+iu_x)_{\theta=\pi}-(u_y+iu_x)_{\theta=-\pi}$$

$$=(c_1+c_2)Kr^{i\epsilon}\sqrt{r}/2\sqrt{2\pi}(1+2i\epsilon)\cosh(\pi\epsilon), \tag{19}$$

whereas the energy release per unit of new crack area is (e.g., Malyshev and Salganik, 1965)

$$G=(c_1+c_2)K\bar{K}/16\cosh^2(\pi\epsilon). \tag{20}$$

This K is related to the complex intensity factor k_1+ik_2 introduced by Sih and Rice (1964) and Rice and Sih (1965) by $K=(k_1+ik_2)\sqrt{\pi}\cosh(\pi\epsilon)$, and reduces to K_I+iK_{II} for a homogeneous solid ($c_1=c_2$, $\epsilon=0$). For the interface crack of length L subject to remotely uniform stresses σ_{yy}^{∞} and σ_{xy}^{∞} (Fig. 2)

$$K=(\sigma_{yy}^{\infty}+i\sigma_{xy}^{\infty})(1+2i\epsilon)L^{-i\epsilon}\sqrt{\pi L/2} \tag{21}$$

at the right-hand crack tip.

The coefficient b_0 in equations (16) represents a stress field of type σ_{xx} that is uniform but different in each of the two phases, in the manner discussed by Rice and Sih (1965). Further, equations (16) confirm that the full set of Williams eigenvalues have form $\lambda=n-1/2+i\epsilon$ and $\lambda=n$.

Nonlinear Material Response and Crack Surface Contact

Inevitably there will be a zone near the crack tip for which the present solution representation fails, whether due to material nonlinearity (e.g., a plastic zone) or to the well-known surface interpenetration predicted by equation (19), leading to Comninou contact zones. Let r_1 be the maximum radius of any such zone and let r_2 be the minimum radius from the crack tip of Fig. 1 to the specimen surface, or other crack tip, or place of load application. Assume that the excluded zone near the crack tip is small enough that $r_1 < r_2$; note also that r_2 will be the radius of convergence for the Taylor series in equations (16). Then the analysis presented above may be reviewed with R reinterpreted as the annular ring between r_1 and r_2. Every step is valid, so that the representations in equations (2)–(4) and (12)–(15) continue to hold with the understanding that $f(z)$ and $g(z)$ are analytic in the reinterpreted R, and thus have the Laurent representation

$$f(z) = \sum_{n=-\infty}^{+\infty} a_n z^n, \quad g(z) = \sum_{n=-\infty}^{+\infty} b_n z^n \tag{22}$$

on $r_1 < |z| < r_2$. The terms with negative n coincide to what, in the homogeneous case, has been called an outer Williams expansion (Rice, 1974) and, just as shown for that case, the term $b_{-1} = 0$ since there is no net force acting on the inner zone at the crack tip.

Small Scale Nonlinear or Contact Zones

For homogeneous materials, linear elastic fracture mechanics procedures (i.e., characterizing crack growth in terms of $K_I + iK_{II}$) are valid when the inevitable nonlinear zone at the crack tip is sufficiently small that, when r is rescaled on that zone size, the radius r_2 is effectively infinite. In that case the "small scale yielding" approach is valid: The actual crack problem is replaced by that of a semi-infinite crack in an infinite solid with asymptotic boundary condition that, at large r, the field approaches that of the standard Irwin-Williams elastic singularity of a strength characterized by $K_I + iK_{II}$ for the problem. Thus, even though nonlinearities cause the actual field to differ locally from the Irwin-Williams stress and deformation distributions, that actual field is uniquely characterized by $K_I + iK_{II}$, which provides the boundary conditions and hence determines the onset of crack growth.

The same remarks apply for the interface crack problem in terms of the complex intensity factor K. It uniquely characterizes the actual field in the small scale nonlinear or contact zone case, even though expressions such as (18) and (19) are not accurate within that zone. Conditions for onset of crack growth for the interface crack are thus properly phrased in terms of K for the small scale regime. Within that small scale nonlinear or contact zone description, the field for $|z| > r_1$ has the representation as in equations (2)–(4) and (12)–(15) with

$$f(z) = \sum_{n=-\infty}^{0} a_n z^n, \quad g(z) = \sum_{n=-\infty}^{-2} b_n z^n \tag{23}$$

where a_0 is specified in terms of K for the crack problem, as in (17). The remaining coefficients of the small scale solution are determined only as part of a full nonlinear analysis, e.g., as by Sham (1984) for the homogeneous elastic ideally plastic solid. They are all unique, if unknown, functions of K.

An improvement which, in the homogeneous material case, is known to significantly enlarge the range of load levels over which the small scale yielding analysis procedure gives accurate results is to include in the latter of equations (23), as a specified term, the term b_0 representing the crack-parallel near tip σ_{xx} stress of the elastic crack solution (Larsson and Carlsson, 1973; Rice, 1974).

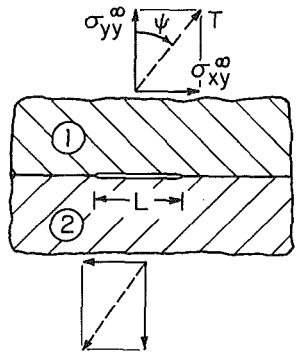


Fig. 2 Remote tension and shear loading of interface crack

Near Tip Contact Zone

An elementary estimate of the contact zone size, assuming that it is small compared to crack size, is given by finding the largest r for which the opening gap, $u_y(r,\pi) - u_y(r,-\pi)$, predicted by equations (19) vanishes, that is, for which

$$\mathrm{Re}[Kr^{i\epsilon}/(1+2i\epsilon)] = 0. \tag{24}$$

For the remotely loaded crack of Fig. 2, let

$$\sigma_{yy}^{\infty} + i\sigma_{xy}^{\infty} = Te^{i\psi} \tag{25}$$

where T is the magnitude of the traction vector on remote surface $y = \text{constant} >> L$, and where the phase angle ψ gives its direction. Thus, $\psi = 0$ corresponds to tension; $\psi = \pm\pi/2$ to shear in the $\pm x$ directions. In this case equation (24), with equations (21) and (25), becomes

$$\mathrm{Re}[e^{i\psi}(r/L)^{i\epsilon}] = \cos[\psi - \epsilon\,\ell n(L/r)] = 0 \tag{26}$$

Assume now that $\epsilon > 0$ (if not, we can just exchange the labels "1" and "2," which changes ϵ to $-\epsilon$, and then change ψ to $-\psi$ to describe the same physical problem but with $\epsilon > 0$). Let ψ lie in the range $-\pi/2 < \psi < +\pi/2$, so that some tensile component always acts. Then the contact zone size r_c is estimated by $\psi - \epsilon\ell n(L/r) = -\pi/2$, or

$$r_c = L\exp[-(\psi + \pi/2)/\epsilon]. \tag{27}$$

Since ϵ is typically small, r_c/L is a rapidly varying function of ψ and is very much smaller than unity over most of the range cited above, including, say, $-\pi/4 < \psi < \pi/2$. It will not remain small for any $\epsilon > 0$ when ψ approaches $-\pi/2$.

In general ϵ increases with increase of the stiffness ratio μ_2/μ_1. For example, if we take material "1" as cork (with $\nu_1 \approx 0$) and bond it to a stiff substrate like alumina (Al_2O_3) for "2," so that $\mu_1/\mu_2 \approx 0$, then ϵ has its largest feasible value (at least for solids with $\nu \geq 0$), namely, $\epsilon = 0.175$. Among the harder, nonpolymeric solids, a relatively extreme stiffness contrast is provided by fused silica (SiO_2) or soda lime glass for "1" and Al_2O_3 for "2," in which case $\epsilon = 0.075$. Representative values of ϵ are considerably lower for various "1"/"2" combinations of interest for practical metal and nonmetal interfaces. For example, Hutchinson et al. (1987) give $\epsilon = 0.039$ for Ti/Al_2O_3, 0.028 for Cu/Al_2O_3, 0.019 for Nb/Al_2O_3, 0.011 for Si/Cu, 0.005 for MgO/Ni, and 0.004 for Au/MgO based on elastic parameters that they tabulate.

If one adopts $r_c/L < 0.01$ as a suitable restriction on r_c so that the small scale contact zone concept may be applied, that is, so that the field may be regarded as being characterized by the complex K, then one requires that $\psi > -\pi/2 + 4.605\,\epsilon$. Thus it is required for validity of the linear elastic fracture mechanics approach outlined above that $\psi > -50$ deg when

$\epsilon = 0.15$, $\psi > -77$ deg when $\epsilon = 0.05$, and $\psi > -87$ deg when $\epsilon = 0.01$. These restrictions will generally be met in practical cases for which there is some nonnegligible tensile component of the loading relative to the crack.

Although a finite, if extremely small, contact zone is predicted by the procedure here for tensile loading, $\psi = 0$, nonlinear elastic analysis of a particular class of joined materials in plane stress conditions for that loading (Knowles and Sternberg, 1983) does not predict contact at the tip.

Coupled Size and Load Phase Effects

Suppose linear elastic fracture mechanics conditions are met so that crack growth conditions are controlled by the complex K. A remarkable result is that if we wish to duplicate the conditions near the tip of a crack of length L loaded, for example, in tension ($\psi = 0$), by testing another body with crack length $L' \neq L$, then that other body must be loaded by a combination of tension and shear. This differs from the ordinary fracture mechanics of homogeneous solids and it means that separate tensile and shear modes cannot be unambiguously defined.

Consider the geometry of Fig. 2 and rewrite K from equation (21) as

$$K = Te^{i\psi}(1 + 2i\epsilon)L^{-i\epsilon}\sqrt{\pi L/2}. \tag{28}$$

Thus if one is to duplicate, in a solid with crack length L', the same conditions near the tip as in a solid with crack length L loaded by traction T at angle ψ, then

$$T' e^{i\psi'}(L')^{-i\epsilon}\sqrt{L'} = Te^{i\psi}L^{-i\epsilon}\sqrt{L}. \tag{29}$$

Thus the traction T' must be altered according to the usual inverse square root dependence

$$T' = T\sqrt{L/L'}, \tag{30}$$

but the phase angle of the loading must be altered also, to

$$\psi' = \psi + \epsilon\, \ell n(L'/L). \tag{31}$$

Usually the phase angle change, $\psi' - \psi$, is small. For example, if one compares specimens for which the crack lengths differ by a factor of 10, the change in ψ is 2.3 ϵ, that is, 10 deg for $\epsilon = 0.075$ (the value for fused SiO_2/Al_2O_3) but only 1.4 deg for $\epsilon = 0.011$ (Si/Cu). More extreme changes in L such as comparison of a 50 μm defect in service with a 5 mm crack in a test specimen ($L'/L = 100$) changes ψ by 4.6 ϵ, and this may be nonnegligible for materials of significant difference in stiffness (e.g., 20 deg for $\epsilon = 0.075$, and 46 deg for the largest possible ϵ, 0.175).

Another perspective is in terms of a failure locus. Imagine that from mixed tension and shear loading tests on a solid as in Fig. 2 with crack length L a locus of stress states σ_{yy}^{∞}, σ_{xy}^{∞} at onset of crack growth have been determined. Such results define a curve in the σ_{yy}^{∞}, σ_{xy}^{∞} plane or, equivalently, define a polar plot of traction magnitude T versus load angle ψ. Equations (30) and (31) then show that when the crack length is increased to L', the failure locus both contracts self-similarly and rotates. Evidently the rotation will usually be small unless there are extreme changes in L and a large ϵ value.

Awkward as it may seem, the proper physical units of complex K, when one measures stress in MPa and length in m, are such that a particular value should be reported as

$$K = C\ MPa\ \sqrt{m}\ m^{-i\epsilon} \tag{32}$$

where C is a pure complex number. Also, if one changes the units for measuring length from m to mm, then not only is the magnitude of C increased by $\sqrt{1000} \approx 32$, but so also is its phase angle changed by $-\epsilon\, \ell n(1000) \approx -6.9\ \epsilon$ (i.e., by -20 deg when $\epsilon = 0.050$). Thus a value of K for which C is real in some system of units cannot properly be regarded as a "mode I" K, because C would not remain real with some other choice of units. One must conclude that tension and shear effects are inherently coupled near interface crack tips. Particular external loadings cannot be said without ambiguity to produce separate "mode I tension" or "mode II shear" conditions near the crack tip.

Once a particular system of units is chosen, the onset of crack growth may be characterized by C reaching a failure locus in a complex plane whose axes are the real and imaginary parts of C. Change of units will both self-similarly scale (as in conventional fracture mechanics) and also rotate the failure locus, so that the real and imaginary parts of C cannot be interpreted as "mode I" and "mode II" components.

It may be noted that a circular failure locus, whether in the σ_{yy}^{∞}, σ_{xy}^{∞} plane or the complex C plane, will be unaffected by rotation. Such a circular failure locus would correspond to a fixed energy release rate G at onset of crack growth (see equation (20)), irrespective of the phase angle of the applied loading.

Small ϵ and Possible Definitions of Stress Intensity Factors of Classical Type

Despite the intrinsically mixed tension and shear fields at the crack tip, and the related rotational effects discussed, for very many material combinations of interest ϵ will be very small, say, of order 0.01 to 0.03. In such cases the rotation angle, $\psi' - \psi$, associated with a factor of 10 change in L lies only between 1.3 deg and 4.0 deg. One is therefore inclined in such cases to neglect the rotation and seek a description without the complexities associated with complex K. Such has evidently been the motivation for proposals advanced by Malyshev and Salganik (1965) and Cherepanov (1979).

Observe that for any interfacial crack problem the complex K will have the form

$$K = \Lambda T\sqrt{L}L^{-i\epsilon} \tag{33}$$

where T is an applied traction loading, L is a relevant length describing the geometry (say, the shorter of crack length, uncracked ligament width, and distance from crack tip to a point of load application) and Λ is a complex number which depends on the phase angle of the applied loading, on ratios of the various other lengths to L and, in general, on ν_1, ν_2 and μ_1/μ_2. In describing any component of the singular near tip stress and displacement field, K always appears as the factor $Kr^{i\epsilon}$, i.e., as

$$Kr^{i\epsilon} = \Lambda T\sqrt{L}(r/L)^{i\epsilon} = \Lambda T\sqrt{L}e^{-i\epsilon\ell n(L/r)}. \tag{34}$$

Equation (18) for the stresses σ_{yy} and σ_{xy} ahead of the crack has the same form as for a homogeneous solid if we replace $Kr^{i\epsilon}$ by $K_I + iK_{II}$. So also does equation (19) for the crack surface displacements, under the same replacement, at least to the disregard of terms of order ϵ.

Thus to the extent that $Kr^{i\epsilon}$ is sensibly independent of r over some range of interest for the application of fracture mechanics methodology, we might choose a value (any value) of r in that range, say $\hat{r}$, and characterize the crack tip fields by intensity factors of classical type defined by

$$K_I + iK_{II} = K\hat{r}^{i\epsilon} = \Lambda T\sqrt{L}(\hat{r}/L)^{i\epsilon}. \tag{35}$$

How should $\hat{r}$ be chosen? It hardly matters within very broad limits, for small ϵ. For example, considering two choices of $\hat{r}$ that differ by a factor of 10, the two resulting values for $K_I + iK_{II}$ as defined above will have a ratio to one another of

$$(10)^{i\epsilon} = e^{2.3i\epsilon} = (0.9997 + 0.0230i) \text{ to } (0.9976 + 0.0689i) \tag{36}$$

for $\epsilon = 0.01$ to 0.03. These are sufficiently close to unity that factor of 10 precision in choosing $\hat{r}$ is more than adequate. Also, if one intends applications over a range of geometrical sizes so that L varies by a factor of order 10 or less, then the term $(\hat{r}/L)^{i\epsilon}$ in equation (35) will likewise vary by no more than a term of order $(10)^{i\epsilon}$, which has just been seen to differ negligibly from unity for the small ϵ values considered.

It is tempting to choose $\hat{r}$ as a fixed fraction of L, say, as $\hat{r}=L/50$ which gives

$$K_I+iK_{II}=\Lambda e^{-3.91i\epsilon}T\sqrt{L} \tag{37}$$

and, in fact, differs little numerically for small ϵ from what one gets by formally setting $\hat{r}=L$ (despite the asymptotic character of the expressions containing K) so that

$$K_I+iK_{II}=\Lambda T\sqrt{L}. \tag{38}$$

The latter choice is consistent with the discussion by Cherepanov (1979).

Of course, both of these choices for $\hat{r}$ are objectional in principle. They go against the spirit of elastic fracture mechanics where the intention (as in the earlier interpretation of complex K) is to define parameters which fully characterize the effects of load and geometry on the crack tip field. Extremely large changes in L, with T altered accordingly, so that $T\sqrt{L}$ is invariant, definitely affect the character of the near tip field (the rotation effect discussed earlier). This effect is lost in equations (37) and (38) since $\hat{r}$ has been scaled to L in those expressions and hence does itself change. If, however, one limits attention to more modest changes in L, on the order of a factor of 10 or less, and to material combinations giving small ϵ values (as seems typical), then either of equations (37) or (38) is suitable for operational use in fracture mechanics analysis within that context. For clarity, it is proposed that one should refer to these as classical stress intensity factors based on distance $\hat{r}=L/50$ or L, respectively.

Preferred Definitions of Stress Intensity Factors of Classical Type

It is preferable in principle to choose $\hat{r}$ in equation (35) as some material length, invariant to differences of crack size or other overall geometric dimensions in different applications. Such an $\hat{r}$ might be chosen as some large multiple of atomic dimensions or as a multiple of some less well defined "fracture process zone size" in certain specified test conditions. Upon reflection it may be realized that so long as we wish to use K_I and K_{II} as characterizing parameters, it really doesn't matter how we choose $\hat{r}$ so long as it is fixed for a given material combination, i.e., unaffected by changes in crack size or other geometric dimensions. Thus a suitable procedure is that one adopt a fixed $\hat{r}$ for all material combinations and that this be taken as $\hat{r}=1\mu$m so that

$$K_I+iK_{II}=K(\mu m)^{i\epsilon}=\Lambda T\sqrt{L}(\mu m/L)^{i\epsilon}. \tag{39}$$

The K_I and K_{II} so defined should be referred to as classical stress intensity factors based on distance $\hat{r}=1\mu$m.

Plainly to the extent that the complex K uniquely characterizes the crack tip field, so also do K_I and K_{II} based on $\hat{r}=1\mu$m (or on any other fixed $\hat{r}$), and this is so no matter how much L varies or however large or small is the value of ϵ. Essentially, the proposed definition of K_I+iK_{II} lets one avoid dealing with the awkwardly complex physical dimensions (MPa $\sqrt{m}\,m^{-i\epsilon}$) of complex K, in favor of those (MPa $\sqrt{m}$) of classical stress intensity factors, but retains all the features of complex K as a crack tip characterizing parameter. Its drawback is that it seems to carry the implication that "mode I" and "mode II" have unambiguous meaning.

Finally, one may also observe that if the crack did not quite lie on the interface but, rather, if its tip was in one of the two joined solids, then classical K_I and K_{II} could be defined at its tip. In this connection Hutchinson et al. (1987) have recently shown that if a crack lies parallel to the interface, in material "2" at $y=-h$, and if $h<<L$, then the classical K_I and K_{II} at its tip are given by

$$K_I+iK_{II}=qe^{i\phi}Kh^{i\epsilon}=q\Lambda e^{i\phi}T\sqrt{L}(h/L)^{i\epsilon}. \tag{40}$$

Here K is the complex intensity for the associated interface crack problem,

$$q=[(1+c_1/c_2)/2\cosh^2\pi\epsilon]^{1/2} \tag{41}$$

and ϕ is a function of ν_1, ν_2 and μ_1/μ_2 which, for typical material combinations like those enumerated in a previous section, satisfies (like ϵ) $|\phi|<<1$. For those cases of small ϵ and small ϕ one may write

$$K_I+iK_{II}\approx q\Lambda T\sqrt{L}(h/L)^{i\epsilon} \tag{42}$$

and, unless L/h is large in a logarithmic sense (i.e., so long as $\epsilon\ell n(L/h)$ is small compared to unity), this can be rewritten, as Hutchinson et al. note, as

$$K_I+iK_{II}\approx q\Lambda T\sqrt{L}. \tag{43}$$

The coincidence, apart from the factor q (or $qe^{i\phi}$) of expression (43) with (38), and of (40) and (42) with (35) and (39) is evident. In particular, the K_I+iK_{II} proposed in equation (39) to be associated with an interface crack coincides, apart from the factor $qe^{i\phi}(h/\mu m)^{i\epsilon}$ (which will often differ little from q), with the actual K_I+iK_{II} of equation (40) for an interface-parallel crack.

Small Scale Yielding at Bimaterial Crack Tips

In the small scale yielding problem, when one or both of the joined solids deforms plastically, the characteristic dimension of the plastic zone must depend only on the complex K and material properties. The properties include the yield strength σ_0 of the weaker of the two solids, the ratios of yield strengths, and dimensionless properties describing strain hardening and ratios of elastic constants. Let r_p be that characteristic dimension. It could represent the maximum radius of the plastic region or (if different) the radius along $\theta=0$.

Dimensional analysis shows that r_p must enter through the complex dimensionless combination

$$K/\sigma_0 r_p^{1/2-i\epsilon}=|\Lambda|(T/\sigma_0)\sqrt{L/r_p}\;e^{i\omega}(r_p/L)^{i\epsilon} \tag{44}$$

where, in the latter version, Λ in the generic form of complex K in equation (33) has been written as $|\Lambda|e^{i\omega}$. This is actually a pair of combinations (real and imaginary part or, more conveniently, amplitude and phase). Thus r_p must satisfy an equation of the form

$$r_p=\frac{K\bar{K}}{\sigma_0^2}f[\text{phase}(Kr_p^{i\epsilon})]=\Lambda\bar{\Lambda}\frac{T^2}{\sigma_0^2}Lf[\omega-\epsilon\ell n(L/r_p)] \tag{45}$$

where $f[\ldots]$ is a dimensionless function. This is, of course, an implicit equation for r_p. Furthermore, since the argument, $\omega-\epsilon\ell n(L/r_p)$, enters the dimensional analysis as the phase angle of a complex quantity, $f[\ldots]$ must have a periodic (with period 2π) dependence on that argument.

For the geometry of Fig. 2, $\Lambda=(1+2i\epsilon)e^{i\psi}\sqrt{\pi/2}$, and thus $\omega=\psi+\arctan(2\epsilon)$ where ψ is the phase angle of the remotely applied traction, $\sigma_{yy}^{\infty}+i\sigma_{xy}^{\infty}=Te^{i\psi}$. Thus in that case

$$r_p=\frac{\pi}{2}(1+4\epsilon^2)\frac{T^2}{\sigma_0^2}Lf[\psi+\arctan(2\epsilon)-\epsilon\ell n(L/r_p)]. \tag{46}$$

It is plausible that r_p should vary with the phase angle ψ of the applied traction, and hence that $f[\ldots]$ is not independent of its argument (i.e., not a constant). In such cases it is then required that for fixed ψ, r_p cannot have the classical small scale yielding proportionality, $r_p\propto T^2L/\sigma_0^2$, because f itself changes as r_p increases.

As an example, Shih and Asaro (1988) have recently done plane strain elastic-plastic numerical calculations for a bimaterial crack as in Fig. 2 with "2" being rigid, $\mu_2/\mu_1=\infty$, $(\sigma_0)_2/\sigma_0=\infty$, and $\nu_1=0.3$, in which case $\epsilon=0.0935$. Table 1 summarizes their results for the cases they report which are definitively in the small scale yielding range (either because $r_p<<1$ or because they used the small scale yielding formula-

Table 1

T/σ_0	ψ	r_p/L	$f[\ldots]$	$\omega-\epsilon\ell n(L/r_p)$
2×10^{-4}	0	2.8×10^{-8}	0.431	−1.44
2.236×10^{-1}	$-1.107(-63.4°)$	2.5×10^{-3}	0.308	−1.27
6×10^{-3}	0	1.6×10^{-5}	0.273	−0.85

tion for a semi-infinite crack, with the complex K for the geometry of Fig. 2 used in asymptotic boundary conditions). These cases are for a "deformation theory" treatment of plasticity in Ramberg-Osgood form with hardening exponent $n=3$, and r_p corresponds to the zone size along $\theta=0$. The loading is described in the first two columns, their finite element results for r_p/L in the third, and the implied value of $f[\ldots]$ in the fourth. The argument of $f[\ldots]$ is reported in the fifth column. Note that $f[\ldots]$ inferred from results with a strong negative shear component ($\psi=-63.4$ deg) fits between those for $\psi=0$ but with much smaller plastic zones, the results being ordered according to values of $\omega-\epsilon\ell n(L/r_p)$.

A similar dimensional analysis reveals the form of the stress field in the small scale yielding formulation. This must have the form

$$\sigma_{ij}=\sigma_0 g_{ij}[\sigma_0^2 r/K\bar{K},\theta,\text{phase}(Kr^{i\epsilon})]$$

$$=\sigma_0 g_{ij}[\sigma_0^2 r/\Lambda\bar{\Lambda}T^2L,\theta,\omega-\epsilon\ell n(L/r)] \quad (47)$$

where there is periodic dependence on the third group and, as earlier, the g_{ij} depend also on dimensionless material properties or property ratios.

Notes

Since preparation of this manuscript Symington (1987) has confirmed the expected result that the original Williams (1959) analysis leads to crack tip solutions with integer eigenvalues λ, represented here by the series for $g(z)$ in equation (16). Those λ did not appear as a conclusion of the Williams paper since he had evidently cancelled out a certain trigonometric factor, vanishing when $\lambda=$ integer, in evaluating the zeros of the 8 by 8 determinant in his analysis.

Also, in a recent pre-publication revision of their manuscript, Shih and Asaro (1988) have incorporated a dimensional analysis of the small scale yielding elastic-plastic problem similar to that in the last section of this paper. They simplify the presentation by noting that within the arbitrariness of arrangement of the results of a dimensional analysis, the r_p which appears within the argument of function f in equations (45) and (46) could equally be replaced by $K\bar{K}/\sigma_0^2=\Lambda\bar{\Lambda}T^2L/\sigma_0^2$, it then being understood that the function f is different from that here. This nicely removes the implicit nature of equations (45) and (46) for r_p, but likewise shows that classical scaling with T^2L/σ_0^2 cannot apply. Similarly, the r within the third argument of g_{ij} in equation (47) could be replaced also by $K\bar{K}/\sigma_0^2=\Lambda\bar{\Lambda}T^2L/\sigma_0^2$ with the understanding that different functions g_{ij} apply.

Acknowledgment

This study was supported by Office of Naval Research contract N00014-86-K-0753 through sub-agreement VB38639-0 from the University of California. The impetus to clarify the interpretation of the complex K came from discussions with A. G. Evans. That to determine the form of terms corresponding to the integer Williams eigenvalues came from queries by M. Symington and D. M. Parks. Discussions with J. W. Hutchinson and B. Budiansky have been helpful.

References

Cherepanov, G. P., 1962, "The Stress State in a Heterogeneous Plate With Slits," in Russian, *Izvestia AN SSSR, OTN, Mekhan. i Mashin.*, Vol. 1, pp. 131–137.

Cherepanov, G. P., 1979, *Mechanics of Brittle Fracture*, McGraw-Hill, New York.

Comninou, M., 1977a, "The Interface Crack," *J. Appl. Mech.*, Vol. 44, pp. 631–636.

Comninou, M., 1977b, "Interface Crack With Friction in the Contact Zone," *J. Appl. Mech.*, Vol. 44, pp. 780–781.

Comninou, M., and Schmueser, D., 1979, "The Interface Crack in a Combined Tension-Compression and Shear Field," *J. Appl. Mech.*, Vol. 46, pp. 345–348.

England, A. H., 1965, "A Crack Between Dissimilar Media," *J. Appl. Mech.*, Vol. 32, pp. 400–402.

Erdogan, F., 1965, "Stress Distribution in Bonded Dissimilar Materials With Cracks," *J. Appl. Mech.*, Vol. 32, pp. 403–410.

Hutchinson, J. W., Mear, M., and Rice, J. R., 1987, "Crack Paralleling an Interface Between Dissimilar Materials," *J. Appl. Mech.*, Vol. 54, pp. 828–832.

Knowles, J. K., and Sternberg, E., 1983, "Large Deformation Near a Tip of an Interface-Crack Between Two Neo-Hookian Sheets," *J. Elasticity*, Vol. 13, pp. 257–293.

Larsson, S. G., and Carlsson, A. J., 1973, "Influence of Nonsingular Stress Terms and Specimen Geometry on Small Scale Yielding at Crack Tips in Elastic-Plastic Materials," *J. Mech. Phys. Solids*, Vol. 22, pp. 263–277.

Malyshev, B. M., and Salganik, R. L., 1965, "The Strength of Adhesive Joints Using the Theory of Crack," *Inter. J. Frac. Mech.*, Vol. 1, pp. 114–128.

Park, J. H., and Earmme, Y. Y., 1986, "Application of Conservation Integrals to Interfacial Crack Problems," *Mech. Mater.*, Vol. 5, pp. 261–276.

Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," *J. Appl. Mech.*, Vol. 32, pp. 418–423.

Rice, J. R., 1968, "Mathematical Analysis in the Mechanics of Fracture," in *Fracture: An Advanced Treatise*, H. Liebowitz, ed., Vol. 2, pp. 191–311.

Rice, J. R., 1974, "Limitations to the Small Scale Yielding Approximation for Crack Tip Plasticity," *J. Mech. Phys. Solids*, Vol. 22, pp. 17–26.

Sham, T.-L., 1984, "A Finite-Element Study of the Asymptotic Near-Tip Fields for Mode I Plane-Strain Cracks Growing Stably in Elastic-Ideally Plastic Solids," in *Spec. Techn. Publ. 803*, Vol. 1, Amer. Soc. Testing Mater., pp. 52–79.

Shih, C. F., and Asaro, R., 1988, "Elastic-Plastic Analysis of Cracks on Bimaterial Interfaces; Part 1: Small Scale Yielding," *J. Appl. Mech.*, in press.

Sih, G. C., and Rice, J. R., 1964, "The Bending of Plates of Dissimilar Materials With Cracks," *J. Appl. Mech.*, Vol. 31, pp. 477–482.

Symington, M. F., 1987, "Eigenvalues for Interface Cracks in Linear Elasticity," *J. Appl. Mech.*, Vol. 54, pp. 973–974.

Williams, M. L., 1959, "The Stresses Around a Fault or Crack in Dissimilar Media," *Bull. Seismol. Soc. America*, Vol. 49, pp. 199–204.

Scattering by Multiple Crack Configurations

Ch. Zhang
Post-Doctoral Fellow

J. D. Achenbach
Fellow ASME

Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208

A system of boundary integral equations is presented which governs the crack-opening displacements for two-crack configurations. The integral equations are highly singular and they cannot be solved directly by numerical methods. By the approach of this paper the higher order singularities are, however, reduced to integrable singularities, and the integral equations are subsequently discretized and solved numerically. For several configurations numerical results have been obtained for scattered fields and for elastodynamic stress intensity factors. The scattered-field results are interpreted to apply for a partially closed crack as well as for two separate but neighboring cracks. The stress-intensity factors are intended to apply only to the case of separate cracks. The scattered-field results have relevance to the problem of detection and characterization of cracks in the field of quantitative nondestructive evaluation.

1 Introduction

Analytical and numerical results for ultrasonic wave scattering by a crack provide valuable quantitative information for methods to detect and characterize cracks. From the mathematical point of view an open crack is a surface which does not transmit tractions. The implementation of this statement in a mathematical formulation generally involves a number of idealizations. Thus, the faces of a perfect mathematical crack are taken as smooth. It is also assumed that they are infinitesimally close, even though it is supposed that the faces do not interact with each other.

Real cracks, particularly fatigue and stress corrosion cracks, have rough faces which may contact each other. Sometimes there is not a single crack, but rather a configuration of a principal crack and an adjoining sattelite crack, for example, a macrocrack and a neighboring microcrack. The question which has motivated this paper is to what extent these complicating features of real cracks or systems of cracks can be accounted for in an analysis.

The perfect mathematical crack model is acceptable for a real crack, provided that the latter's faces are slightly separated and that the length characterizing crack-face roughness is much smaller than the dominant wavelength of an incident pulse of ultrasonic wave motion. When the faces of a crack are in contact and the crack is actually partially closed, or when there are two neighboring cracks, the model must be adjusted. In this paper we assume that a partially closed crack can be represented by a configuration of two neighboring cracks. We consider the four configurations which are depicted in Fig. 1.

A crack with integrating crack faces may be a poor reflector of ultrasonic waves, and thus difficult to detect and characterize. To quantitatively investigate scattering by such a crack, Achenbach and Norris (1982) and Thompson and Fiedler (1984) have assumed that the interaction between the upper and lower faces of a crack can be described by appropriate relations between the tractions and displacements across the crack plane. Achenbach and Norris (1982) considered certain nonlinear relations which present different resistance to opening and closing and for which the resistance to sliding depends on the contact pressure. Thompson and Fiedler (1984) considered a linear spring constraint between the crack faces.

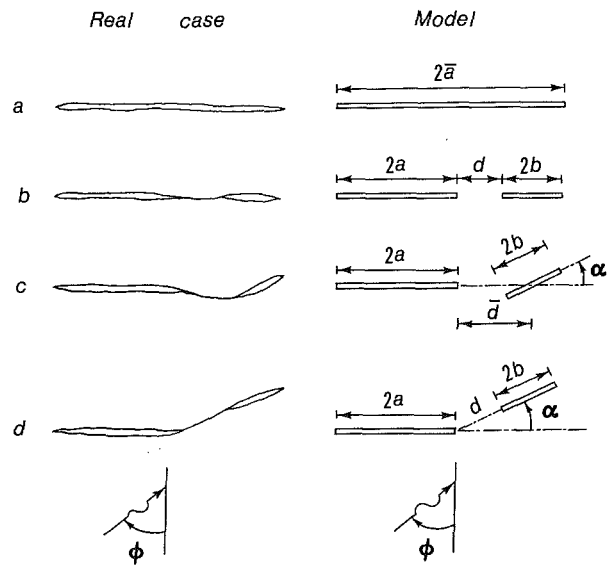


Fig. 1 Crack configurations

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 20, 1987; final revision August 26, 1987.

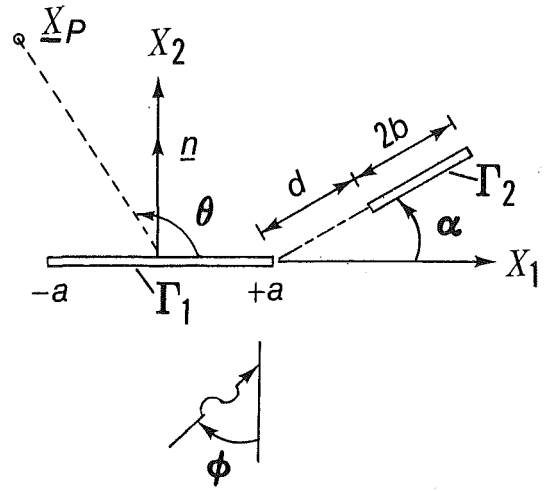


Fig. 2 Geometrical parameters

The mathematical formulation presented in this paper is two-dimensional, and for a state of plane strain. The usual displacement boundary integral equation formulation for scattering by volume scatterers disintegrates when the scatterer is reduced to a crack. Hence a traction boundary integral equation is used in this paper. The mathematical formulation for a single crack is well known (see, e.g., Budiansky and Rice, 1979). Here it is extended to a two-crack configuration, the result being a system of four coupled singular integral equations for the general case. These boundary integral equations are highly singular and they cannot be solved directly by numerical methods. To circumvent this difficulty a special approach is developed by which the higher order singularity has been reduced to an integrable singularity. Numerical results for scattered fields and for elastodynamic stress-intensity factors for several configurations have been obtained by the boundary element method. The scattered field results apply to a partially-closed crack as well as to two separate but neighboring cracks. The stress-intensity factors are intended to apply only to the case of separate cracks.

For the static case the interaction of a crack with a microcrack has been investigated by several authors: Chudnovsky and Kachanov (1983), Chudnovsky, et al. (1984), Kachanov (1985), Kachanov and Montagut (1986), Rubinstein (1985, 1986), and Rose (1986). The influence of the microstructure (such as voids, cracks) on the crack propagation has been analyzed by Hoagland et al. (1973). The scattering of horizontally polarized transverse waves by a system of the two cracks has been studied by Zhang (1986) and Gross and Zhang (1987). For collinear cracks certain in-plane cases have been considered by Itou (1978), who used integral transform methods, and Jain and Kanwal (1972), who used a low frequency perturbation method. Scattering of in-plane waves by a configuration of neighboring cracks has been studied for the special case of a planar array of periodically spaced cracks (see Angel and Achenbach, 1985).

2 The BIE and its Numerical Solution

We consider an unbounded, homogeneous, isotropic and linearly elastic solid, which contains a main crack and a neighboring microcrack. The geometrical configuration is shown in Fig. 2. The solid is in time-harmonic motion, but the term exp $(-i\omega t)$ will be suppressed. The stress equations of motion are given by (see Achenbach, 1973)

$$\sigma_{\alpha\beta,\beta}+\rho\omega^2 u_\alpha=0. \tag{1}$$

Here $\sigma_{\alpha\beta}$ defines the stress components, u_α denotes the displacement components, ρ is the mass density, and ω is the circular frequency. The stress components are related to the displacement gradients by Hooke's law

$$\sigma_{\alpha\beta}=\lambda\delta_{\alpha\beta}u_{\gamma,\gamma}+\mu(u_{\alpha,\beta}+u_{\beta,\alpha}), \tag{2}$$

where λ and μ are Lame's elastic constants. The boundary conditions on the faces of the cracks are

$$\sigma_{\alpha\beta}n_\beta=0, \tag{3}$$

where n_β defines the components of the normal vector of the crack faces.

The total fields, u_α, $\sigma_{\alpha\beta}$, induced by the interaction of the incident wave with the cracks can be written as the sum of the incident field, u^i_α, $\sigma^i_{\alpha\beta}$, and the scattered field, u^{sc}_α, $\sigma^{sc}_{\alpha\beta}$:

$$u_\alpha=u^i_\alpha+u^{sc}_\alpha,\quad \sigma_{\alpha\beta}=\sigma^i_{\alpha\beta}+\sigma^{sc}_{\alpha\beta}. \tag{4}$$

For crack problems the scattered displacement u^{sc}_γ can be expressed by the following representation integral (see Achenbach et al., 1982)

$$u^{sc}_\gamma(\mathbf{x}_p)=\int_\Gamma \sigma^G_{\alpha\beta\gamma}(\mathbf{x};\mathbf{x}_p)\Delta u_\alpha(\mathbf{x})n_\beta ds,\quad \mathbf{x}_p\notin\Gamma. \tag{5}$$

Here $\Gamma=\Gamma_1+\Gamma_2$, where Γ_1 and Γ_2 are the insonified sides of the cracks. Also $\mathbf{x}_p$ is the position vector of the observation point, $\sigma^G_{\alpha\beta\gamma}(\mathbf{x};\mathbf{x}_p)$ is the Green's function (see Appendix A), Δu_α is the displacement jump (crack opening displacement) across the cracks and n_β is the unit normal vector of Γ. In addition, the scattered field must fulfill the radiation condition.

Substitution of equation (5) in the constitutive equation (2) and use of the relations between surface tractions and the stress components,

$$f^{sc}_\alpha(\mathbf{x}_p)=\sigma^{sc}_{\alpha\beta}n_\beta\Big|_{\mathbf{x}_p}, \tag{6}$$

yields the representation formula for the traction at $\mathbf{x}_p$:

$$f^{sc}_\alpha(\mathbf{x}_p)=-n_\beta(\mathbf{x}_p)\int_\Gamma[\lambda\delta_{\alpha\beta}\sigma^G_{\delta\epsilon\gamma,\gamma}(\mathbf{x};\mathbf{x}_p)+\mu\sigma^G_{\delta\epsilon\alpha,\beta}(\mathbf{x};\mathbf{x}_p)+\mu\sigma^G_{\delta\epsilon\beta,\alpha}(\mathbf{x};\mathbf{x}_p)]\Delta u_\delta(\mathbf{x})n_\epsilon ds \tag{7}$$

We note here that the equation (7) becomes improper if the observation point $\mathbf{x}_p$ and the boundary point $\mathbf{x}$ coincide (see Schmerr, 1982). In this case the kernel function of equation (7) would be highly singular and nonintegrable (Appendix A). Hence we cannot take the limit $\mathbf{x}_p\to\Gamma$ at this point. In order to avoid this difficulty, we first divide Γ into N elements. Then equation (7) can be written in the following discretized form:

$$f^{sc}_\alpha(\mathbf{x}_p)=-n_\beta(\mathbf{x}_p)\sum_{j=1}^{N}\int_{s_j}^{s_{j+1}}[\lambda\delta_{\alpha\beta}\sigma^G_{\delta\epsilon\gamma,\gamma}(\mathbf{x};\mathbf{x}_p)+\mu\sigma^G_{\delta\epsilon\alpha,\beta}(\mathbf{x};\mathbf{x}_p)+\mu\sigma^G_{\delta\epsilon\beta,\alpha}(\mathbf{x};\mathbf{x}_p)]\Delta u_\delta(\mathbf{x})n_\epsilon ds, \tag{8}$$

where s_j denotes the first node and s_{j+1} denotes the second node of the jth element.

Because all three terms in the integrand of equation (8) have the same structure, we will consider only the following integral (the first term of equation (8))

$$I_j(\mathbf{x}_p)=\int_{s_j}^{s_{j+1}}\sigma^G_{\delta\epsilon\gamma,\gamma}(\mathbf{x};\mathbf{x}_p)\Delta u_\delta(\mathbf{x})n_\epsilon ds. \tag{9}$$

This equation can be rewritten by adding and subtracting the same terms

$$I_j(\mathbf{x}_p)=\int_{s_j}^{s_{j+1}}\sigma^G_{\delta\epsilon\gamma,\gamma}(\mathbf{x};\mathbf{x}_p)\Delta u_\delta(\mathbf{x})n_\epsilon ds-\int_{s_j}^{s_{j+1}}\sigma^G_{\delta\epsilon\gamma,\epsilon}(\mathbf{x};\mathbf{x}_p)\Delta u_\delta(\mathbf{x})n_\gamma ds+\int_{s_j}^{s_{j+1}}\sigma^G_{\delta\epsilon\gamma,\epsilon}(\mathbf{x};\mathbf{x}_p)\Delta u_\delta(\mathbf{x})n_\gamma ds. \tag{10}$$

The first two terms of equation (10) can be rewritten as (Appendix B)

$$\epsilon_{\gamma\epsilon}\sigma^G_{\epsilon\delta\gamma}\Delta u_\delta \Big|_{s_j}^{s_{j+1}} - \int_{s_j}^{s_{j+1}} \epsilon_{\gamma\epsilon}\epsilon_{\lambda\mu}\sigma^G_{\epsilon\delta\gamma}\Delta u_{\delta,\lambda} n_\mu ds \tag{11}$$

in which $\epsilon_{\gamma\epsilon}$ is the two-dimensional permutation tensor. Using the equation of motion for the Green's function

$$\sigma^G_{\delta\epsilon\gamma,\epsilon} + \rho\omega^2 u^G_{\delta\gamma} = 0, \quad \mathbf{x}_p \neq \mathbf{x}, \tag{12}$$

the third term of equation (10) can be rewritten as

$$\int_{s_j}^{s_{j+1}} \sigma^G_{\delta\epsilon\gamma,\epsilon}\Delta u_\delta n_\gamma ds = -\rho\omega^2 \int_{s_j}^{s_{j+1}} u^G_{\delta\gamma}\Delta u_\delta n_\gamma ds, \tag{13}$$

where $u^G_{\delta\gamma}$ is the Green's function for the displacement components (see Appendix A). By using equations (11) and (13), the integral I_j, equation (10) is now converted into

$$I_j(\mathbf{x}_p) = \epsilon_{\gamma\epsilon}\sigma^G_{\epsilon\delta\gamma}\Delta u_\delta \Big|_{s_j}^{s_{j+1}} - \int_{s_j}^{s_{j+1}} \epsilon_{\gamma\epsilon}\epsilon_{\lambda\mu}\sigma^G_{\epsilon\delta\gamma}\Delta u_{\delta,\lambda} n_\mu ds$$

$$-\rho\omega^2 \int_{s_j}^{s_{j+1}} u^G_{\delta\gamma}\Delta u_\delta n_\gamma ds. \tag{14}$$

It should be noted here that the singular terms of equation (14) at $\mathbf{x}_p = \mathbf{x}$ can be integrated analytically or numerically without difficulties as is usual in BIE formulation. Using the same idea described above for the second and the third term of equation (8), taking the limit $\mathbf{x}_p \to \Gamma$ and considering the boundary condition on the cracks,

$$f^i_\alpha(\mathbf{x}_P) + f^{sc}_\alpha(\mathbf{x}_P) = 0, \quad \mathbf{x}_P \epsilon \Gamma, \tag{15}$$

we obtain finally the following discretized BIE

$$f^i_\alpha(\mathbf{x}_P) = n^*_\beta(\mathbf{x}_P) \sum_{j=1}^{N} [H_{\alpha\beta\delta}(\mathbf{x};\mathbf{x}_p)\Delta u_\delta(\mathbf{x}) \Big|_{s_j}^{s_{j+1}}$$

$$+ \int_{s_j}^{s_{j+1}} J_{\alpha\beta\delta}(\mathbf{x};\mathbf{x}_p)\Delta u_\delta(\mathbf{x}) ds$$

$$+ \int_{s_j}^{s_{j+1}} H_{\alpha\beta\delta}(\mathbf{x};\mathbf{x}_p)\epsilon_{\lambda\mu}\Delta u_{\delta,\lambda} n_\mu ds], \quad \mathbf{x}_p \epsilon \Gamma. \tag{16}$$

In equation (16) $f^i_\alpha(\mathbf{x}_p)$ denotes the traction components on the crack faces due to the incident wave, and $H_{\alpha\beta\delta}$ and $J_{\alpha\beta\delta}$ are defined by

$$H_{\alpha\beta\delta} = \lambda\delta_{\alpha\beta}\epsilon_{\gamma\epsilon}\sigma^G_{\delta\epsilon\gamma} + \mu(\epsilon_{\beta\epsilon}\sigma^G_{\delta\epsilon\alpha} + \epsilon_{\alpha\epsilon}\sigma^G_{\delta\epsilon\beta}), \tag{17a}$$

$$J_{\alpha\beta\delta} = -\rho\omega^2[\lambda\delta_{\alpha\beta}u^G_{\delta\gamma}n_\gamma + \mu(u^G_{\delta\alpha}n_\beta + u^G_{\delta\beta}n_\alpha)]. \tag{17b}$$

If we choose suitable shape functions for Δu_δ and require that equation (16) is satisfied at the N collocation points on Γ, then we get a system of linear, algebraic equations which can be solved numerically.

In our analysis Δu_δ is taken to be constant over each element except for elements at crack tips. For these elements we use "crack tip elements," for example

$$\Delta u_\delta = C^\pm_\delta (a \mp x_1)^{1/2}, \tag{18}$$

to describe the proper behavior of Δu_δ at crack tips. The integration of regular terms in equation (16) is done numerically by using the Gaussian quadrature formula for constant elements, and by using the Gauss-Jacobian formula for crack-tip elements. The singular terms of equation (16) can be integrated numerically or analytically by expanding the kernels in series (see Kitahara, 1985).

3 Dynamic Stress Intensity Factors

In linear fracture mechanics the most relevant quantities are the stress intensity factors. They are related to the crack opening displacement Δu_δ by (see Erdogan, 1983)

$$\begin{Bmatrix} K^\pm_{\mathrm{I}} \\ K^\pm_{\mathrm{II}} \end{Bmatrix} = \frac{\mu\sqrt{2\pi}}{4(1-\nu)} \lim_{x_1 \to \pm a} \frac{1}{\sqrt{a \mp x_1}} \begin{Bmatrix} \Delta u_2(x_1) \\ \Delta u_1(x_1) \end{Bmatrix}$$

$$= \frac{\mu\sqrt{2\pi}}{4(1-\nu)} \begin{Bmatrix} C^\pm_2 \\ C^\pm_1 \end{Bmatrix} \tag{19}$$

where equation (18) has been used. Here "+" indicates the tip at $x_1 = a$ and "−" indicates the tip at $x_1 = -a$, where a is the half crack length (Fig. 2). Also, ν denotes Poisson's ratio.

If the crack opening displacement Δu_δ is calculated from equation (16), then the stress intensity factors can be determined by using equation (19). It should be mentioned here that the use of crack-tip elements is advisable for the determination of the K factors, because the constant elements cannot produce the proper local behavior of the crack opening displacement at crack tips.

4 The Scattered Far-Field

For $|\mathbf{x}_P| \to \infty$ the scattered field can be represented by (see Tan, 1975; Achenbach et al., 1982)

$$u^{sc}_\alpha(\mathbf{x}_P) \approx \frac{A^L_\alpha(\hat{\mathbf{x}}_P)}{(8\pi k_L|\mathbf{x}_P|)^{1/2}} \exp\left(ik_L|\mathbf{x}_P| + i\frac{\pi}{4}\right)$$

$$+ \frac{A^T_\alpha(\hat{\mathbf{x}}_P)}{(8\pi k_T|\mathbf{x}_P|)^{1/2}} \exp\left(ik_T|\mathbf{x}_P| + i\frac{\pi}{4}\right), \tag{20}$$

where $\hat{\mathbf{x}}_P$ is the unit vector along $\mathbf{x}_P$

$$\hat{\mathbf{x}}_P = \frac{1}{|\mathbf{x}_P|} \begin{Bmatrix} x_{P1} \\ x_{P2} \end{Bmatrix} = \begin{Bmatrix} \cos\theta \\ \sin\theta \end{Bmatrix}, \tag{21}$$

while k_L and k_T are the wavenumbers of longitudinal and transverse waves, respectively. The terms A^L_α and A^T_α are given by

$$A^L_\alpha = -\frac{ik_L}{(\lambda+2\mu)} \hat{x}_{P\alpha} \int_\Gamma [\lambda(\Delta u_\beta n_\beta) + 2\mu(\Delta u_\beta \hat{x}_{P\beta})(\hat{x}_{P\gamma} n_\gamma)]$$

$$\cdot \exp(-ik_L \hat{\mathbf{x}}_P \cdot \mathbf{x}) ds \tag{22}$$

$$A^T_\alpha = -ik_T \int_\Gamma [(\hat{x}_{P\beta} n_\beta)\Delta u_\alpha + (\hat{x}_{p\beta}\Delta u_\beta) n_\alpha - 2(\hat{x}_{P\beta}\Delta u_\beta) \cdot$$

$$(\hat{x}_{P\gamma} n_\gamma)\hat{x}_{P\alpha}] \exp(-ik_T \hat{\mathbf{x}}_P \cdot \mathbf{x}) ds. \tag{23}$$

The scattering cross section σ^{sc} of the cracks is defined as the ratio of the outward power flow of the scattered field and the intensity of the incident field. For an incident plane longitudinal wave of the form

$$u^i_\alpha = u^L_\alpha \exp(ik_L \hat{\boldsymbol{\beta}} \cdot \mathbf{x}), \tag{24}$$

the scattering cross section σ^{sc} is given by (see Tan, 1975; Achenbach et al., 1982)

$$\sigma^{sc} = \frac{1}{k_L} Im\left[\frac{u^{L*}_\alpha A^L_\alpha(\hat{\boldsymbol{\beta}})}{u^L_\alpha u^{L*}_\alpha}\right], \tag{25}$$

in which $\hat{\boldsymbol{\beta}} = (\sin\phi, \cos\phi)$ denotes the unit propagation vector of the incident wave, and u^{L*}_α denotes the complex conjugate of u^L_α. Similarly the scattering cross section for an incident plane transverse wave

$$u^i_\alpha = u^T_\alpha \exp(ik_T \hat{\boldsymbol{\beta}} \cdot \mathbf{x}) \tag{26}$$

takes the following form

$$\sigma^{sc} = \frac{1}{k_T} Im\left[\frac{u^{T*}_\alpha A^T_\alpha(\hat{\boldsymbol{\beta}})}{u^T_\alpha u^{T*}_\alpha}\right]. \tag{27}$$

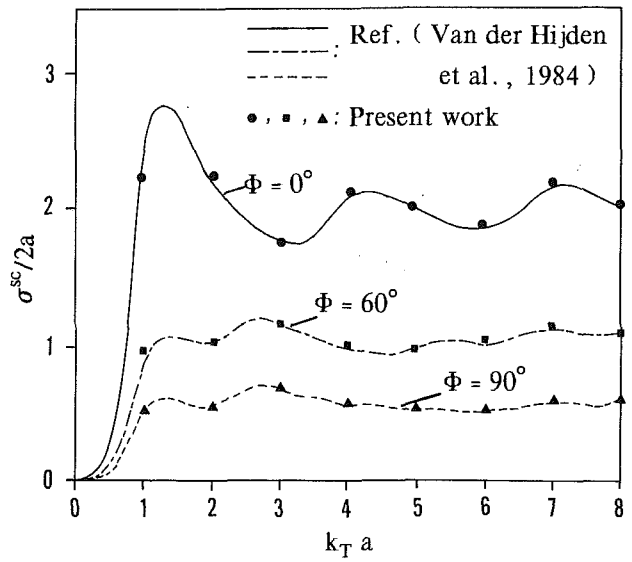


Fig. 3 Normalized scattering cross section for an incident plane *L* wave on a single crack of length 2*a*

5 Numerical Results

A number of geometrical parameters, as well as a material parameter and a frequency parameter, must be specified to carry out numerical calculations. The quantity $k_T a$, where k_T is the wavenumber of transverse waves ($k_T = \omega/c_T$) and a is the half-length of the main crack, was chosen as the frequency parameter. The geometrical parameters are: b/a, d/a, the angle of incidence, ϕ, and the angle defining the orientation of the satellite crack, α. Various values were selected for the frequency parameter and for the geometrical parameters. Only one material parameter, namely Poisson's ratio ν, enters the formulation. All calculations were carried out for $\nu = 1/3$. The main crack was discretized into 50 elements of the same length, while a proportional number of elements was used for the satellite crack. For the calculation of the scattered fields, the distance $|\mathbf{x}_P|$ to the point of observation was defined by $|\mathbf{x}_P|/a = 20$.

For the special cases $b/a = 0$ or $d/a \to \infty$, the configuration reduces to a single crack of half-length a. The case $b/a = 0$ was used to check the results obtained by the method of computation of this paper. For an incident longitudinal (L) wave the scattering cross section generated by the presence of a single crack is shown in Fig. 3. The results show excellent agreement with those obtained by a different method by Van der Hijden et al. (1984). Calculations for the normalized dynamic stress intensity factors agree also exceedingly well with those of Van der Hijden et al. (1984).

Figure 4(a) shows the normalized backscattered displacement $|u_2^{sc}|/|\mathbf{u}^i|$, versus b/a, for the configuration of two collinear cracks (see Fig. 1(b)), for a normally incident longitudinal wave, and for $d/a = 0.1$. For $k_T a = 1$, the influence of the neighboring crack is relatively small, in the sense that the increase as compared to $b/a = 0$ (a single crack) is less than 50 percent for values of b/a as large as unity. For $k_T a = 8$ the effect of the neighboring crack is much larger. The amplitude of the backscattered field doubles as b/a increases from 0 to 1. For both cases the effect of the satellite crack is negligible for $b/a < 0.1$. Also shown in Fig. 4(a) are the backscattered fields for a single crack of half-length $\bar{a} = a + b + d/2$ (see Fig. 1(a)). It is noted that these results show relatively little difference with the ones for the two-crack configuration. The results of Fig. 4(a) may be interpreted as showing that for normal incidence the amplitude of the backscattered field is just about the same for a partially closed crack as

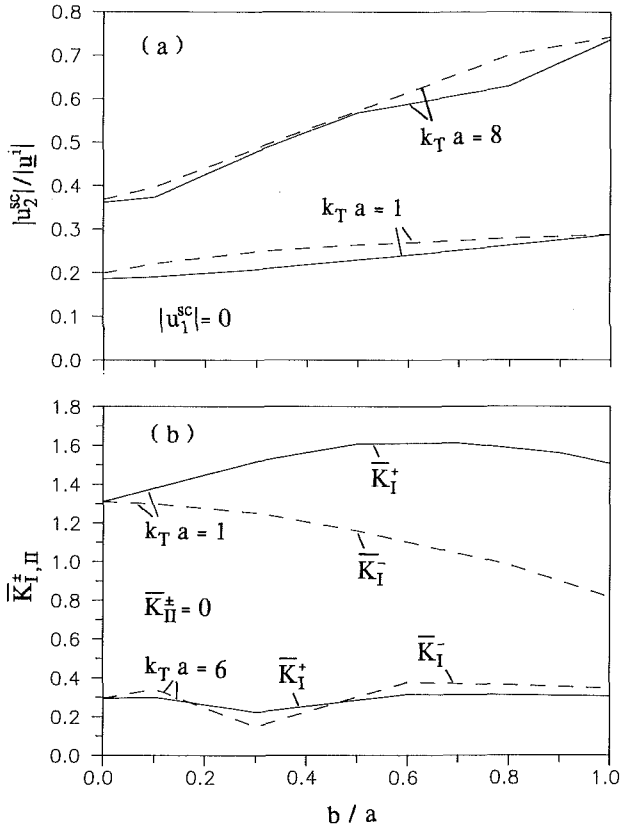


Fig. 4 (*a*) Back-scattered displacements for case *b* of Fig. 1: for normal incidence of a plane *L* wave, *d*/*a* = 0.1, two cracks (———), single crack of length 2$\bar{a}$ (·····); (*b*) Normalized dynamic stress intensity factors for case *b* of Fig. 1: normal incidence of a plane *L* wave, *d*/*a* = 0.1, $\bar{K}_{I,II}^{\pm} = |K_{I,II}^{\pm}|/|K_{I,II}^{sf}|$.

for a wide open crack. It appears that the more intense crack opening of the separate cracks compensates for the closure effect. Naturally this interpretation applies only when the tension corresponding to the incident wave field does not separate the contacting crack faces.

For the collinear two-crack configuration the normalized Mode I dynamic stress intensity factors at the tips of the main crack, induced by normal incidence of a longitudinal wave, are shown in Fig. 4(b). Due to symmetry with respect to $x_2 = 0$, the Mode II stress intensity factors are identically zero for this case. The numbers at $b/a = 0$ correspond to the stress-intensity factors for a single crack. With the presence of a satellite crack, the stress intensity factor is larger at the crack tip adjoining the satellite crack, but interestingly enough, the stress intensity factor at the crack tip away from the satellite crack is smaller, at least for $k_T a = 1$. As is also shown in Fig. 4(b), for $k_T a = 6$ a more ambiguous result is obtained. It is evident that the magnitudes of the dynamic stress intensity factors depend strongly on both b/a and $k_T a$.

The backscattered displacement field is shown versus d/a in Fig. 5(a). Here $b/a = 0.3$ and $k_T a = 1$. It is not surprising that the backscattered field decreases with increasing d/a even though the changes are small. Also shown is the corresponding curve for a single crack of half-length $\bar{a} = a + b + d/2$. The comparison of the two curves shows more clearly that a partially closed crack, with closure area d/a, may be expected to generate a smaller backscattered field.

The stress intensity factors at the tips of the main crack have been plotted versus d/a in Fig. 5(b). For the case of $k_T a = 1$, K_I^- increases as d/a increases. For both K_I^+ and K_I^-, the values for a single crack of half-length a are reached asymptotically as $d/a \to \infty$. It is surprising that, for $k_T a = 1$, K_I^- is

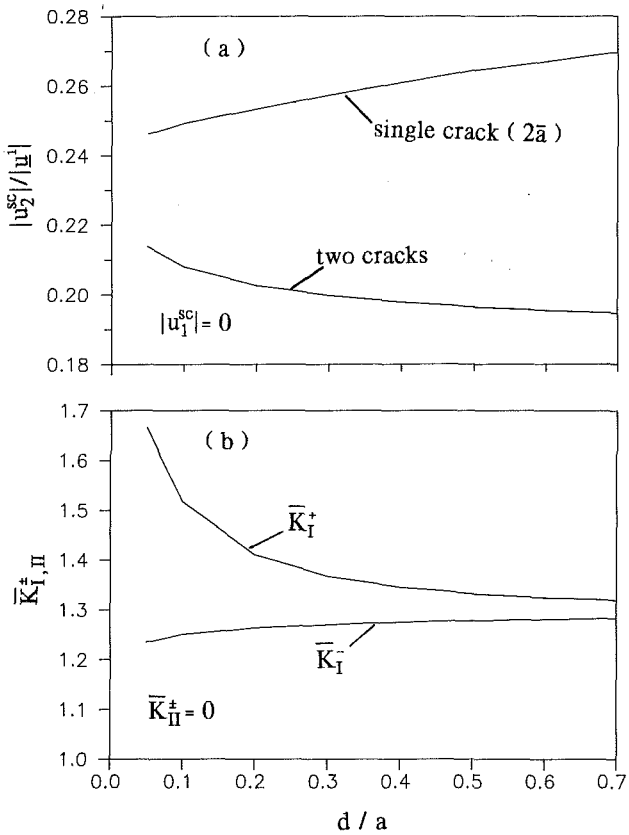


Fig. 5 (a) Back-scattered displacements for case *b* of Fig. 1: normal incidence of a plane *L* wave, $b/a = 0.3$, $k_T a = 1$; (b) Normalized dynamic stress intensity factors for case *b* of Fig. 1: normal incidence of a plane *L* wave, $b/a = 0.3$, $k_T a = 1$, $\bar{K}_{I,II}^{\pm} = |K_{I,II}^{\pm}|/|K_{I,II}^{st}|$.

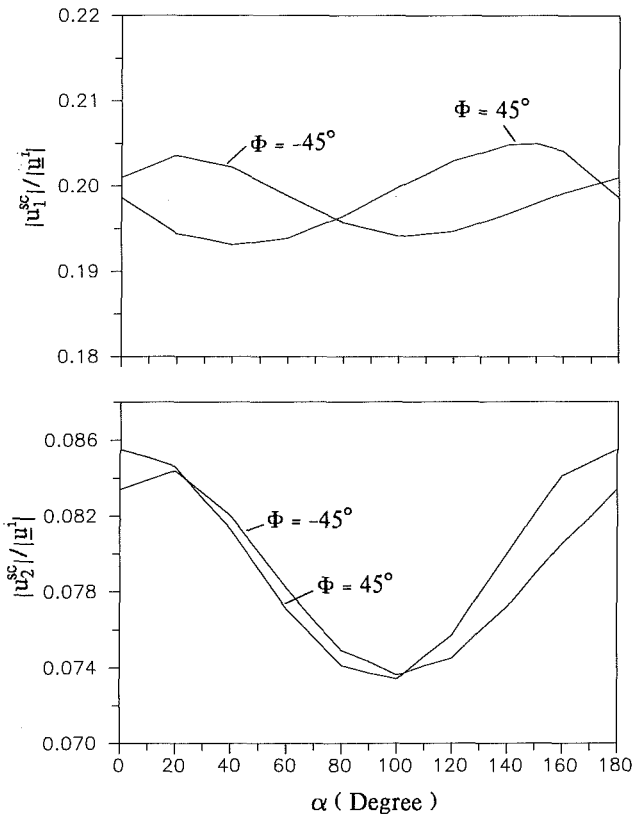


Fig. 6 Back-scattered displacements for case *c* of Fig. 1: incident plane *L* wave, $b/a = 0.2$, $\bar{d}/a = 0.3$, $k_T a = 1$

smaller than the value for a single crack of half-length a. For the static case K_I^- is always larger in the two-crack configuration than for the single crack case.

For a satellite crack whose center is on the extension of the crack line of the main crack, but which is arbitrarily oriented (see Fig. 1(*c*)), results for the backscattered fields are shown in Fig. 6. Results are presented for an incident L wave for two angles of incidence ($\phi = -45$ deg and $\phi = 45$ deg), while $b/a = 0.2$, $\bar{d}/a = 0.3$, and $k_T a = 1$. Both the horizontal and vertical backscattered displacements are shown versus the angle of inclination of the satellite crack, α. It is of interest to compare these results with the ones for two single cracks, of half-lengths a and $\bar{a} = a + (b + \bar{d})/2$, respectively. The back-scattered displacements for the crack of half-length a subjected to an incident L wave with $\phi = -45$ deg and $\phi = 45$ deg are $|u_1^{sc}|/|\mathbf{u}^i| = 0.19102$, $|u_2^{sc}|/|\mathbf{u}^i| = 0.08027$, and the corresponding results for the crack of half-length $\bar{a}$ are $|u_1^{sc}|/|u^i| = 0.23850$, $|u_2^{sc}|/|u^i| = 0.09989$. The larger crack of length $2\bar{a}$ produces a larger backscattered field than the configuration of two cracks of the same overall length. The smaller crack of length $2a$ produces a larger $|u_1^{sc}|/|\mathbf{u}^i|$ component than the two-crack configuration, while depending on how much the microcrack is inclined with respect to the main crack line, the $|u_1^{sc}|/|\mathbf{u}^i|$ component for a single crack can be larger or smaller than those for two cracks.

Figure 7 shows the dependence of the stress intensity factors on the orientation angle, α. The corresponding values for a single crack of length $2a$ under an incident L wave with $\phi = 45$ deg are: $|K_I^+|/|K_I^{st}| = 1.0647$, $|K_I^-|/|K_I^{st}| = 0.8640$, $|K_{II}^+|/|K_{II}^{st}| = 0.3078$, and $|K_{II}^-|/|K_{II}^{st}| = 0.2840$ (for $\phi = -45$ deg we only have to change the "+" and "−" subscripts). Comparison with results for two cracks (see Fig.

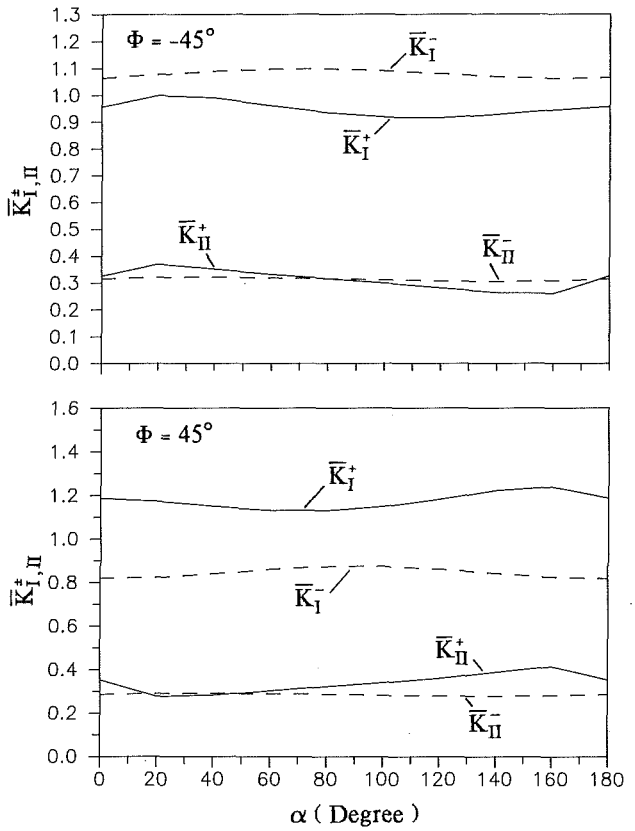


Fig. 7 Normalized dynamic stress intensity factors for case c of Fig. 1: incident plane *L* wave, $b/a = 0./2$, $\bar{d}/a = 0.3$, $k_T a = 1$, $\bar{K}_{I,II}^{\pm} = |K_{I,II}^{\pm}|/|K_{I,II}^{st}|$

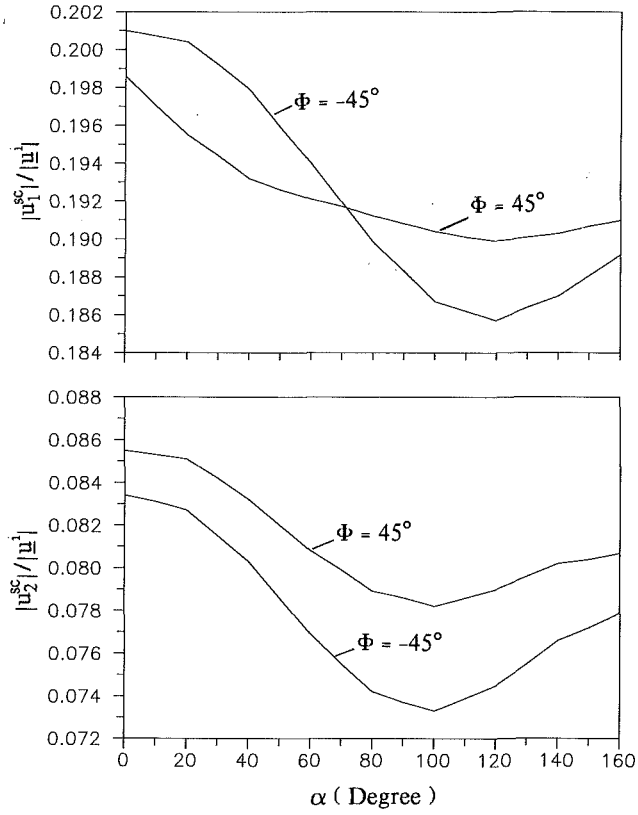


Fig. 8 Back-scattered displacements for case *d* of Fig. 1: incident plane *L* wave, *b/a* = 0.2, *d/a* = 0.1, $k_T a$ = 1

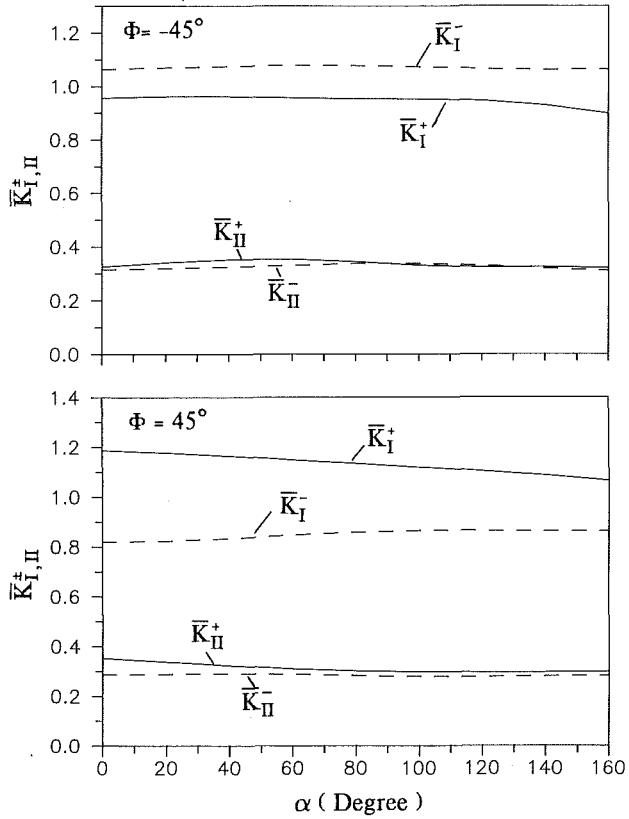


Fig. 9 Normalized dynamic stress intensity factors for case *d* of Fig. 1: incident plane *L* wave, *b/a* = 0.2, *d/a* = 0.1, $k_T a$ = 1, $\bar{K}_{I,II}^{\pm} = |K_{I,II}^{\pm}|/|K_{I,II}^{st}|$

7) shows that the K_I^+ factor of the main crack tip is always increased while K_{II}^+ can be increased or reduced. The tip of the main crack away from the satellite crack is less affected by the microcrack orientation and the $K_{I,II}^-$ factors are therefore almost constant.

Finally, Figs. 8 and 9 present some results for a two-crack configuration where the satellite crack is under a polar angle α (see Fig. 1(*d*)). As in the last case, the back-scattered displacements for the two crack configurations are always smaller than for a single crack of length $2\tilde{a}$, but depending on α these can be smaller or larger than those for a single crack of length $2a$. For a slightly inclined satellite crack (α small) the back-scattered field is larger than that for a single crack of length $2a$, but when the satellite crack is behind the right tip of the main crack ($\alpha > 90$ deg), then the back-scattered field becomes smaller. Because of the presence of the satellite crack, the K_I^+ factor is increased for all α values, while K_{II}^+ can be increased or reduced depending on the satellite crack orientation. Figure 9 shows that the influence of the polar angle α on the $K_{I,II}^-$ factors is very small for the case that is being considered.

Acknowledgment

This work was carried out in the course of research sponsored by the Office of Naval Research, Project N00014-85-K-0401.

References

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

Achenbach, J. D., Gautesen, A. K., and McMaken, H., 1982, *Ray Methods for Waves in Elastic Solids*, Pitman, Boston.

Achenbach, J. D., and Norris, A. N., 1982, "Loss of Specular Reflection Due to Nonlinear Crack-Face Interaction," *Journal of Nondestructive Evaluation*, Vol. 3, No. 4, pp. 229–239.

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.

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

Chudnovsky, A., and Kachanov, M., 1983, "Interaction of a Crack with a Field of Microcracks," *Int. J. Engng. Sci.*, Vol. 21, pp. 1009–1018.

Chudnovsky, A., Dolgopolsky, A., and Kachanov, M., 1984, "Elastic Interaction of a Crack with Microcracks," *Advances in Fracture Research*, Proc. of the Sixth Conf. on Fracture, New Delhi, India, Atluri, S. et al., eds., Vol. 2, pp. 825–833.

Erdogan, F., 1983, "Stress Intensity Factors," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 992–1002.

Gross, D., and Zhang, Ch., 1987, "Diffraction of SH-Waves by a System of Cracks, Solution by an Integral Equation Method," *Int. J. Solids Structures*, Vol. 24, pp. 41–49.

Hoagland, R. Hahn, G., and Rosenfield, A., 1973, "Influence of Microstructure on the Fracture Propagation in Rock," *Rock Mech.*, Vol. 5, pp. 77–106.

Itou, S., 1978, "Dynamic Stress Concentration around Two Coplanar Griffith Cracks in an Infinite Elastic Medium," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 803–806.

Jain, D. L., and Kanwal, R. P., 1972, "Diffraction of Elastic Waves by Two Coplanar Griffith Cracks in an Infinite Elastic Medium," *Int. J. Solids Structures*, Vol. 8, pp. 961–975.

Kachanov, M., 1985, "A Simple Technique of Stress Analysis in Elastic Solids with Many Cracks," *Int. J. Fracture*, Vol. 28, pp. R11–R19.

Kachanov, M., and Montagut, E., 1986, "Interaction of a Crack with Certain Microcrack Arrays," *Eng. Fract. Mech.*, Vol. 25, pp. 625–636.

Kitahara, M., 1985, *Boundary Integral Equation Methods in Eigenvalue Problems of Elastodynamics and Thin Plates*, Elsevier, Amsterdam.

Rose, L. R. F., 1986, "Microcrack Interaction with a Main Crack," *Int. J. Fracture*, Vol. 31, pp. 233–242.

Rubinstein, A. A., 1985, "Macrocrack Interaction with Semi-Infinite Microcrack Array," *Int. J. Fracture*, Vol. 27, pp. 113–119.

Rubinstein, A. A., 1986, "Macrocrack-Microdefect Interaction," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 505–510.

Schmerr, L. W., 1982, "The Scattering of Elastic Waves by Isolated Cracks Using a New Integral Equation Model," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 1, Thompson, D. O., and Chimenti, D. E., eds., Plenum Press, pp. 511–515.

Tan, T. H., 1975, "Diffraction Theory for Time-Harmonic Elastic Waves," Thesis, Delft University.

Thompson, R. B., and Fiedler, C. J., 1984, "The Effect of Crack Closure on Ultrasonic Scattering Measurements," *Review of Progress in QNDE*, 3A, Thompson, D. O., and Chimenti, D. E., eds., Plenum, New York, pp. 207–215.

Van den Hijden, J. H. M. T., and Neerhoff, F. L., 1984, "Scattering of Elastic Waves by a Plane Crack of Finite Width," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 646–651.

Zhang, Ch., 1986, "Die Anwendung von Integralgleichungsmethoden zur Lösung dynamischer Rissprobleme," Thesis, Technische Hochschule Darmstadt.

APPENDIX A

The Green's Function

The Green's function for the two-dimensional time-harmonic elastodynamic state is given by (see Tan, 1975; Achenbach et al., 1982)

$$\frac{4\rho\omega^2}{i}\,u^G_{\alpha\gamma}=\frac{r_\alpha r_\gamma}{r^2}[k_L^2H_o^{(1)}(k_Lr)-k_T^2H_o^{(1)}(k_Tr)]$$

$$+\frac{1}{r}\left(\delta_{\alpha\gamma}-\frac{2r_\alpha r_\gamma}{r^2}\right)[k_LH_1^{(1)}(k_Lr)$$

$$-k_TH_1^{(1)}(k_Tr)]+\delta_{\alpha\gamma}k_T^2H_o^{(1)}(k_Tr), \qquad (A.1)$$

where

$$r=|\mathbf{x}-\mathbf{x}_P|,\ r_\alpha=x_\alpha-x_{P\alpha}, \qquad (A.2)$$

and $H_o^{(1)}(\cdot)$ and $H_1^{(1)}(\cdot)$ denote Hankel functions of the first kind and zeroth and first orders, respectively. The function $u^G_{\alpha\gamma}(\mathbf{x};\mathbf{x}_P)$ denotes the displacement in the α-direction observed at position $\mathbf{x}$, due a unit force in the γ direction, applied at position $\mathbf{x}_P$. The corresponding components of the stress tensor follow from Hooke's law as

$$\frac{4\rho\omega^2}{i}\,\sigma^G_{\alpha\beta\gamma}=-\lambda k_L^3\delta_{\alpha\beta}\frac{r_\gamma}{r}H_1^{(1)}(k_Lr)$$

$$-\frac{\mu k_T^3}{r}(\delta_{\beta\gamma}r_\alpha+\delta_{\alpha\gamma}r_\beta)H_1^{(1)}(k_Tr)$$

$$+\frac{2\mu}{r^2}\left(\frac{4r_\alpha r_\beta r_\gamma}{r^2}-\delta_{\alpha\beta}r_\gamma-\delta_{\alpha\gamma}r_\beta-\delta_{\beta\gamma}r_\alpha\right)[k_T^2H_o^{(1)}(k_Tr)$$

$$-k_L^2H_o^{(1)}(k_Lr)]+\frac{4\mu}{r^3}\left(\delta_{\alpha\beta}r_\gamma+\delta_{\alpha\gamma}r_\beta+\delta_{\beta\gamma}r_\alpha\right.$$

$$\left.-\frac{4r_\alpha r_\beta r_\gamma}{r^2}\right)[k_TH_1^{(1)}(k_Tr)-k_LH_1^{(1)}(k_Lr)]+$$

$$+2\mu\frac{r_\alpha r_\beta r_\gamma}{r^3}[k_T^3H_1^{(1)}(k_Tr)-k_L^3H_1^{(1)}(k_Lr)]. \qquad (A.3)$$

It should be noted here that $u^G_{\alpha\beta}$ behaves as ℓnr while $\sigma^G_{\alpha\beta\gamma}$ behaves as $1/r$ when $r\to 0$. It can also easily be shown that $\sigma^G_{\alpha\beta\gamma,\delta}$ has a singularity of the type $1/r^2$ as $r\to 0$.

APPENDIX B

Derivation of Equation (11)

For convenience we use the abbreviation $\tilde{I}_j$ for the first two terms of equation (10):

$$\tilde{I}_j(\mathbf{x}_P)=\int_{s_j}^{s_{j+1}}[\sigma^G_{\delta\epsilon\gamma,\gamma}(\mathbf{x};\mathbf{x}_P)n_\epsilon-\sigma^G_{\delta\epsilon\gamma,\epsilon}(\mathbf{x};\mathbf{x}_P)n_\gamma]\Delta u_\delta(\mathbf{x})ds \qquad (B.1)$$

Using the following relation for the two-dimensional permutation tensor

$$\epsilon_{\gamma\epsilon}\epsilon_{\lambda\mu}=\delta_{\gamma\lambda}\delta_{\epsilon\mu}-\delta_{\gamma\mu}\delta_{\epsilon\lambda}, \qquad (B.2)$$

we can rewrite (B.1) as

$$\tilde{I}_j(\mathbf{x}_P)=\int_{s_j}^{s_{j+1}}\epsilon_{\gamma\epsilon}\epsilon_{\lambda\mu}\sigma^G_{\delta\epsilon\gamma,\lambda}\Delta u_\delta(\mathbf{x})ds. \qquad (B.3)$$

It can be shown that

$$\epsilon_{\lambda\mu}n_\mu ds=dx_\lambda. \qquad (B.4)$$

Partial integration of (B.3) with respect to x_λ by considering equation (B.4) yields

$$\tilde{I}_j(\mathbf{x}_P)=\int_{s_j}^{s_{j+1}}\epsilon_{\gamma\epsilon}\sigma^G_{\delta\epsilon\gamma,\lambda}\Delta u_\delta dx_\lambda$$

$$=\epsilon_{\gamma\epsilon}\sigma^G_{\delta\epsilon\gamma}\Delta u_\delta\Big|_{s_j}^{s_{j+1}}-\int_{s_j}^{s_{j+1}}\epsilon_{\gamma\epsilon}\epsilon_{\lambda\mu}\sigma^G_{\delta\epsilon\gamma}\Delta u_{\delta,\lambda}n_\mu ds. \qquad (B.5)$$

Initiation, Propagation, and Kinking of an Antiplane Crack

C. C. Ma
Department of Mechanical Engineering,
National Taiwan University,
Taipei, Taiwan 10764

P. Burgers
Hibbitt, Karlsson & Sorensen, Inc.,
Providence, RI 02906

An infinite linear elastic body containing a semi-infinite crack is loaded by a planar antiplane stress pulse parallel to the crack. The stress wave strikes the crack at time $t = 0$ *and at some arbitrary later time* t_f*, the crack begins to extend straight ahead with constant speed* v_o*. After some later* t_b*, the crack suddenly stops, then kinks and propagates with constant speed* v_c*, making an angle* δ *with the original crack. A superposition scheme is used to construct the exact full-field solution of the propagating crack. The full-field solution for stresses for the constant speed propagating crack with a delay time* t_f *is found to be the Mode III analog of Baker's problem in Mode I plus the stress pulse, and the displacement on the crack faces behind the moving crack tip is just the solution of Baker's problem when expressed in crack tip coordinates and is independent of the delay time* t_f*. When the crack suddenly stops, the stress field, which is radiated out from the stopped crack tip, corresponds to the stationary crack stress field of a crack whose crack tip has been at the stopped crack position for all time. The dynamic stress intensity factor at the kinked crack tip is then obtained by using a perturbation method. The region of the stress intensity factor controlled field is investigated for both stationary and propagating cracks. It is found that this region depends on the loading conditions and which stress components are considered. The region also depends on the crack tip speed and will contract as the crack tip speed increases.*

Introduction

Crack branching has been frequently observed in dynamic fracture research on brittle materials at high crack speeds. This branching complicates the evaluation of the energy release rate and stress intensity factor of the propagating crack for both experimental and theoretical work. One of the aims of current research on crack branching is to determine conditions which control the onset of bifurication or crack kinking.

The experimental work on dynamic fracture is very difficult to perform, since the whole fracture process finishes in a very short time and the field quantities need to be measured without influencing the process itself. In the study of dynamic crack growth phenomena, it is very important to have available as complete a description of the prevailing mechanical fields as possible. In the interpretation of experimental data, it is particularly important to have a detailed description of the dependence of the stress intensity factor on the applied loads and specimen geometry, because it is the local stress distribution which determines the instantaneous driving force for continued crack growth. The most successful methods for measurement of the dynamic stresss intensity factor are the optical methods of photoelasticity and shadow spots. Ravi-Chandar and Knauss (1984a,b,c,d) have recently performed some carefully controlled experiments with dynamic loading of a long crack and use the shadow-spot technique to measure the dynamic stress intensity factor during crack propagation.

The analytical work dealing with crack branching at an arbitrary angle with respect to the primary crack is rather limited and much is elastostatic in nature. Significant progress has been made recently in obtaining correct solutions in analyzing the problem of a semi-infinite crack which kinks (or bifurcates) under dynamic loading. Burgers and Dempsey (1982) solved the Mode III problem for bifurcation half-angles $= 0$, $\pi/2$ and Burgers (1982) solved the same problem numerically for all angles. Dempsey et al. (1982), obtained the analytical Mode III solution for crack kinking under stress wave loading and verified the numerical results in Burgers (1982). Dempsey et al. (1986) studied the problem of asymmetric crack bifurcation for antiplane strain under stress wave loading. The numerical method described in Burgers (1982) has been used by Burgers (1983) and Burgers and Dempsey (1984) to provide numerical solutions to the plane strain crack kinking and bifurcation problems. The approximate method for both Mode III and mixed Mode I-II crack kinking under stress wave loading was investigated by Achenbach et al. (1984). This approximate method is based on the fact that the elastodynamic crack kinking stress intensity factors are not significantly affected by the geometry of the corner; the kinking angle δ was used as the perturbation parameter.

In all the results mentioned above, each branching crack tip propagates at a constant velocity from the instant that either

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, July 22, 1986; final revision June 15, 1987.

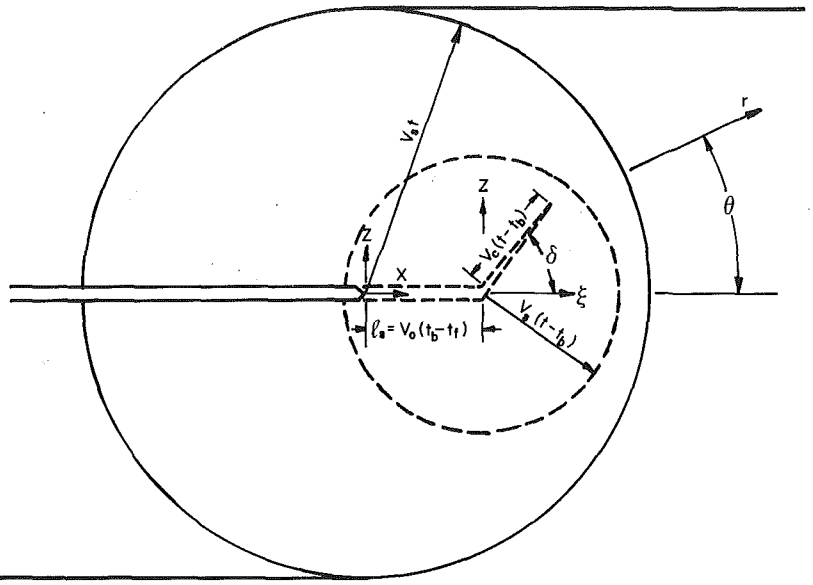


Fig. 1 Geometry of the wave fronts for a crack which kinks, subject to an incident stress wave

an incident plane stress wave strikes the semi-infinite crack tip or that the uniform loading is applied to the semi-infinite crack faces. The problems then have no characteristic length and give rise to self-similar solutions. More recently, a finite delay time has been included in the initiation of the nonplanar crack by Ma and Burgers (1986a), in which the approximate method in Achenbach et al. (1984) has been used for analyzing the antiplane strain case. The dynamic stress intensity factor for both incident planar stress wave loading and uniform loading applied on the semi-infinite crack faces is written in a simple closed form. As the delay time tends to zero, this closed-form solution reduces to the results obtained in Achenbach et al. (1984). Ma and Burgers (1986b) extended the delay time effect to the inplane case. For the delay time equal to zero, the results in Ma and Burgers (1986b) agree closely to Achenbach et al. (1984) up to a high kinking angle (70 deg). A surprising result in Ma and Burgers (1986b) is that this approximate method obtained a similar accuracy for Mode I stress intensity factor as the Mode III case, when compared to the numerical results in Burgers and Dempsey (1984) for mixed mode kinking. The Mode II stress intensity factor also had a very good agreement.

The results in Ma and Burgers (1986a, 1986b) indicate that if a maximum energy release rate is accepted as a crack propagation criterion, then for both the incident stress wave parellel to the original crack faces and uniform dynamic loading applied to the original crack faces, the crack will propagate straight ahead of the original crack for any delay time. The theoretical prediction for crack propagation straight ahead is also observed experimentally by Ravi-Chander and Knauss (1984a-d) for crack face loading and Ravi-Chandran (1986) for stress wave loading for the Mode I loading case. The analysis undertaken here is the extension of previous work (Ma and Burgers, 1986a, 1986b) in which the kinked crack initiates out from the original stationary crack tip. What we consider here is allowing the original crack to propagate straight ahead for some time and then kink. The geometry then corresponds more closely to what is observed experimentally, although the problem is solved in antiplane strain to simplify the mathematics.

We begin with constructing the dynamic stress fields around a crack propagating with constant velocity, and then calculating the complete stress fields around the crack when it stops suddenly. With the stopping crack stress fields available, the approximate method in Achenbach et al. (1984), Ma and Burgers (1986a, 1986b) may then be applied to calculate the dynamic stress intensity factor for the kinking crack. It is noted that some techniques used in this paper were first described in a series of papers by Freund (1972a, 1972b, 1973, 1974) which deal with the propagation of a crack along the original crack line.

Description of Problem and Solution Methodology

We consider a stress-free linear elastic homogeneous isotropic infinite medium that contains a stationary semi-infinite crack. This crack will be referred to as the original crack and lies along the negative x axis with the origin of the coordinate system at the crack tip. An incident stress wave parallel to the crack faces arrives at the semi-infinite crack faces at time $t=0$ so that the deformation occurs in antiplane strain. A short time later, at $t=t_f$, a new crack propagates out of the original semi-infinite crack in the positive x direction with constant velocity v_o less than the shear wave speed v_s. At time $t=t_b>t_f$, the propagating crack stops and instantaneously (or after some delay time) a kinked crack initiates out of the stopped crack tip, making an angle δ with the original crack and propagating with constant velocity v_c. The pattern of wavefronts and the position of the crack tip for $t>t_b$ is shown in Fig. 1.

In order to obtain the final results, the superposition of solutions plays a significant role. First, the diffraction field of the stationary crack due to the stress wave loading must be worked out. The stress distribution that is radiated out on the prospective fracture plane ahead of the crack must be negated by the propagating crack in order to extend. The complete full field stress solution around the crack tip, moving with constant velocity, can be expressed as the sum of the solutions of some simpler problems.

When the crack suddenly stops at $l_s=v_o(t_b-t_f)$, we find that the stationary crack solution is radiated out for *the full field* behind the shear wave speed and is exactly the stress distribution that a stationary crack tip would have if it had been at $x=l_s$ for all time. With this remarkable result at hand, all the results shown in Ma and Burgers (1986a) can be applied to this case with a small modification. The stopping crack solution was first given by Eshelby (1969), but the derivation here follows Freund (1972a, 1972b, 1973, 1974) and is possibly easier to follow.

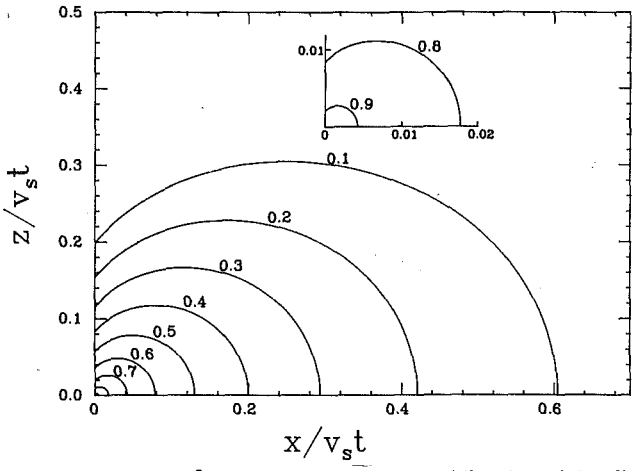


Fig. 2 The ratio of τ^s_{zy} for the exact solution and the stress intensity factor controlled field due to uniform loading on the crack faces

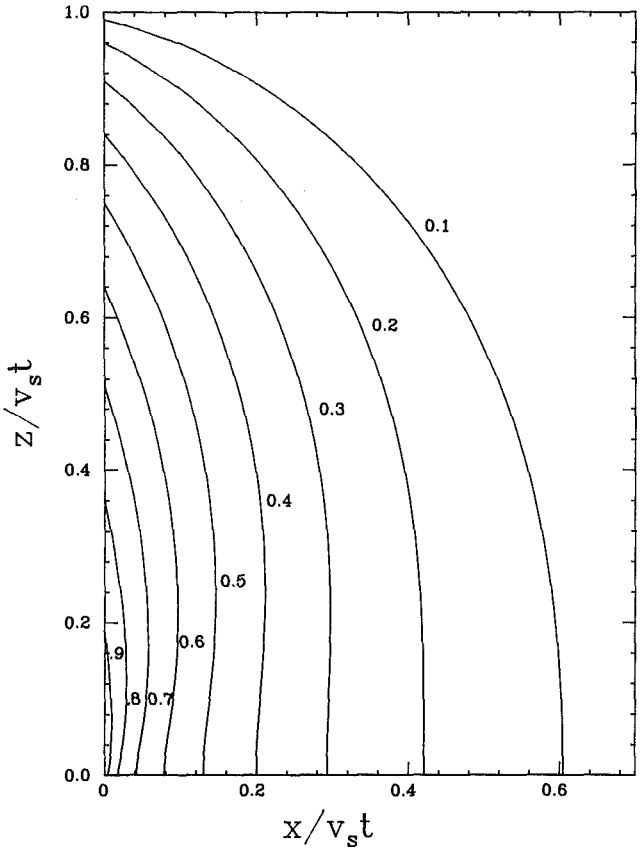


Fig. 3 The ratio of $\tau^s_{\theta y}$ for the exact solution and the stress intensity factor controlled field due to uniform loading on the crack faces

The Stress Field Solution for the Propagating Crack

In a stationary coordinations systems of x and z, two-dimensional antiplane wave motions are governed by

$$\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial z^2}=b^2\frac{\partial^2 w}{\partial t^2}, \tag{1.1}$$

where w is the out-of-plane displacement, and 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 stresses are

$$\tau_{xy}=\mu\frac{\partial w}{\partial x},\ \tau_{zy}=\mu\frac{\partial w}{\partial z}. \tag{1.2a,b}$$

We consider first the loading τ_o applied uniformly on the semi-infinite crack faces at time $t=0$. The full field solutions for the stresses can be written in a simple form; a derivation is given in Ma and Burgers (1986a)(although this is not the first derivation): for $\theta<\pi/2$

$$\tau^s_{zy} = \frac{\tau_o}{\pi}\left\{2\cos\frac{\theta}{2}\sqrt{\frac{t}{br}-1}-\tan^{-1}\sqrt{\frac{t/br-1}{1-\sin\theta}}\right.$$
$$\left.-\tan^{-1}\sqrt{\frac{t/br-1}{1+\sin\theta}}\right\}, \tag{1.3}$$

$$\tau^s_{xy} = -\frac{2\tau_o}{\pi}\sin\frac{\theta}{2}\sqrt{\frac{t}{br}-1}, \tag{1.4}$$

$$\tau^s_{\theta y} = \tau^s_{zy}\cos\theta-\tau^s_{xy}\sin\theta$$
$$= \frac{\tau_o}{\pi}\left\{2\cos\frac{\theta}{2}\sqrt{\frac{t}{br}-1}-\cos\theta\tan^{-1}\sqrt{\frac{t/br-1}{1-\sin\theta}}\right.$$
$$\left.-\cos\theta\tan^{-1}\sqrt{\frac{t/br-1}{1+\sin\theta}}\right\}. \tag{1.5}$$

In the limit $r\to 0$, the stresses become

$$\tau^s_{zy}=\tau^s_{\theta y}\approx\frac{2\tau_o}{\pi}\cos\frac{\theta}{2}\sqrt{\frac{t}{br}},\ \tau^s_{xy}\approx-\frac{2\tau_o}{\pi}\sin\frac{\theta}{2}\sqrt{\frac{t}{br}} \tag{1.6a,b}$$

The appearance of a stress singularity has implications from the point of view of fracture mechanics. It is generally assumed that the magnitude of the stress intensity factor, possibly as a function of crack velocity, is a criterion for extension of a crack. From an experimental point of view, measurements of the field quantities near the crack tip are used to obtain the stress intensity factor, such as in Ravi-Chander and Knauss (1984a-d) where the shadow-spot method is used. Whether the data points they have obtained from the experimental observation are within the stress intensity controlled field or not, plays an important role in the determination of the correct stress intensity factor.

In Fig. 2, the ratio for stress τ^s_{zy} evaluated from the actual stress (1.3) and the stress calculated from the stress intensity controlled field (1.6*a*) is plotted. If we accept that for a ratio up to 0.9, the actual stress is accurately described by the stress intensity factor controlled field, then, the stress intensity factor controlled field will be valid only for points close to the crack tip; that is, within a distance from the crack tip of 0.2 to 0.4 percent of the distance to the cylindrical shear wave front. For the stress τ^s_{xy}, this region is much larger—approximately 19 percent of the shear wave front. The ratio for the stress $\tau^s_{\theta y}$ is plotted in Fig. 3. The region of accuracy is greater in the z-direction than the x-direction, because for small θ, the stress $\tau^s_{\theta y}$ is controlled by τ^s_{zy} while for high θ by τ^s_{xy}.

For an incident stress wave loading, we assume a ratio of 1.1 defines the region where the stress intensity factor controlled field is suitably accurate. From Fig. 4, we find that this region for τ^s_{zy} will be up to 20 percent of the shear wave front. From the results shown above, one can see that the region where the stress intensity controlled field can accurately describe the stress state near the crack tip is a function of the loading conditions and components of stresses. The general features shown above for anti-plane strain are also observed for the in-plane case (Ma and Burgers, 1985).

In order for the crack to extend at constant speed v_o for $t>t_f$ in the direction of the fracture plane, the traction

distribution equal but opposite to the stress τ^s_{zy} on $\theta=0$ must be canceled:

$$(\tau^s_{zy})_o=\tau^s_{zy}(\theta=0)=\frac{2\tau_o}{\pi}\left\{\sqrt{\frac{t}{bx}-1}-\tan^{-1}\sqrt{\frac{t}{bx}-1}\right\}. \quad (1.7)$$

The solutions of stresses for constant speed crack propagation can be obtained by determining the elastic field for a concentrated force appearing through the moving crack-tip, and then building up the general stress distribution by superposition. It is observed that $(\tau^s_{zy})_o$ is a homogeneous function of degree zero in x and t; that is, it depends only on the ratio t/x. This means that any fixed stress level radiates out along the x-axis at a constant speed. With this feature of the stress distribution, it was first observed by Freund (1973) that the force element needed for constructing the solution can be expressed by the following boundary conditions

$$\tau^F_{zy}(x,0,t)=(mt+n)\Delta(x-ut)H(t) \text{ for } -\infty<x<v_ot, \quad (1.8)$$

$$w^F(x,0,t)=0 \text{ for } v_ot<x<\infty, \quad (1.9)$$

where Δ is the Dirac delta function and H is the Heaviside step function. m and n are arbitrary parameters which are independent of x and t. These concentrated force elements with magnitude n acting on the crack tip at $t=0$ when the crack begins to extend. For $t>0$, the crack tip moves in the positive x direction with speed v_o, while the concentrated forces which increase as a linear function of time move in the same direction with speed $u<v_o$.

The complete transformed solution is found by means of Laplace transforms and the Wiener-Hopf technique. The exact solution can then be found by inverting the Laplace transforms by de Hoop's method. The full field solution can be expressed as follows:

$$\tau^F_{zy}=-\frac{nh}{\pi\beta+(h)}Im\left[\frac{\beta+(\lambda_s)\partial\lambda_s/\partial t}{\lambda_s-h}\right]_{\lambda_s(t)}$$
$$+\frac{mh^2}{\pi}\int_{t_s}^{t}Im\left\{\beta+(\lambda_s)\left[\frac{1}{(\lambda_s-h)\beta+(h)}\right]_h\frac{\partial\lambda_s}{\partial s}\right\}ds \quad (1.10a)$$

$$\tau^F_{zy}=\frac{nh}{\pi\beta+(h)}Im\left\{\frac{\lambda_s\partial\lambda_s/\partial t}{(\lambda_s-h)\beta-(\lambda_s)}\right\}_{\lambda_s(t)}$$
$$-\frac{mh^2}{\pi}\int_{t_s}^{t}Im\left\{\frac{\lambda_s}{\beta-(\lambda_s)}\left[\frac{1}{(\lambda_s-h)\beta+(h)}\right]_h\frac{\partial\lambda_s}{\partial s}\right\}ds \quad (1.10b)$$

where

$$\lambda_s(s)=\frac{-(\xi s+b^2z^2/d)+iz\{s^2-b^2[(\xi+s/d)^2+z^2]\}^{1/2}}{\xi^2+(1-b^2/d^2)z^2},$$

$$t_s=\frac{b\{b\xi/d+[\xi^2+(1-b^2/d^2)z^2]^{1/2}\}}{(1-b^2/d^2)},$$

$$\xi=x-v_ot,\ d=\frac{1}{v_o},\ h=\frac{1}{v_o-u},$$

$$\beta+(h)=[b+h(1-b/d)]^{1/2},\ \beta-(h)=[b-h(1+b/d)]^{1/2}.$$

where the subscript h denotes differentiation with respect to h. The x coordinate is eliminated in favor of a new coordinate ξ which is always fixed with respect to the moving crack tip. Using the superposition scheme described by Freund (1973), the full field solution of the stress around the propagating crack tip due to the stress $(\tau^s_{zy})_0$ acting on the new crack faces can be constructed by choosing $m=-1$, $n=-t^*=-ht_f/d$, replacing t by $t-t^*$ and $\xi=x-v_ot$ by $\xi=x-v_o(t-t_f)$ in the force element solution (1.10) and integrating over the appropriate range of $u=x/t$. If we consider a point with the coordinates

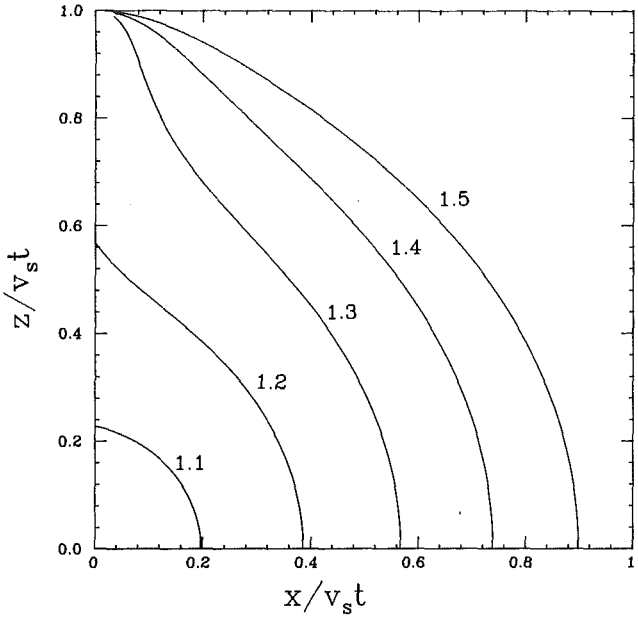


Fig. 4 The ratio of τ^s_{zy} for the exact solution and the stress intensity factor controlled field due to stress wave loading

(ξ,z), the range of u for which the stresses, due to the force element in equation (1.8), will reach this point after the crack tip begins to move at $t=t_f$ with speed v_o, is the range of x/t between $x=0$, $t=t_f$ and $x=x^*$, $t=t_f+x^*d$, where x^* is the distance the crack has propagated along the crack line from which a shear wave, emitted at the instant the crack tip reaches x^*, will just reach the point at (ξ,z) at time t. x^* satisfies the equation

$$\{[v_o(t-t_f)-x^*+\xi]^2+z^2\}^{1/2}=v_s(t-t_f-x^*d). \quad (1.11)$$

Solving for x^*, we get

$$x^*=v_o(t-t_f)-v_ot_s.$$

The range of u is, therefore, $0<u<x^*/(t-t_s)$ and the final result is

$$\tau^N_{zy}=\int_0^{x^*/(t-t_s)}\left\{\frac{t_fh^2}{\pi d\beta_+(h)}\mathrm{Im}\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial t}{\lambda_s-h}\right]_{\lambda_s(t-t^*)}\right.$$
$$\left.-\frac{h^2}{\pi}\int_{t_s}^{t-t^*}Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{(\lambda_s-h)\beta_+(h)}\right]_h ds\right\}(\tau^s_{zy})_o du$$
$$=-\frac{1}{\pi}\int_0^{x^*/(t-t_s)}h^2\left\{\int_{t_s}^{t-t^*}Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{(\lambda_s-h)\beta_+(h)}\right]ds\right\}_h\cdot$$
$$(\tau^s_{zy})_o du,\ t>t_f. \quad (1.12a)$$

$$\tau^N_{xy}=\frac{1}{\pi}\int_0^{x^*/(t-t_s)}h^2\left\{\int_{t_s}^{t-t^*}Im\left[\frac{\lambda_s\partial\lambda_s/\partial s}{(\lambda_s-h)\beta_+(h)\beta_-(\lambda_s)}\right]ds\right\}_h\cdot$$
$$(\tau^s_{zy})_o du,\ t>t_f. \quad (1.12b)$$

Since the complicated functions in equations (1.12) are written as functions of h, it is advantageous to replace the variable of integration in equations (1.12) by h. The range of integration on h is $d<h<d^*$

$$\tau^N_{zy}=-\frac{1}{\pi}\int_d^{d^*}\left\{\int_{t_s}^{t-t^*}Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{(\lambda_s-h)\beta_+(h)}\right]ds\right\}_h(\tau^s_{zy})_o dh, \quad (1.13)$$

where

$$d^* = \frac{(t-t_s)d}{t_f}.$$

We now go through the full analysis only for τ_{zy}^N; τ_{xy}^N can be obtained in a similar fashion.

At first sight, the integral in equation (1.13) is ideally suited for integration by parts, which will allow us to rewrite it into more simple expressions, each representing the solution of a more fundamental standard problem. However, we find that the function $(\tau_{zy}^s)_o$ has a square root singularity at $h=d$. If we apply the integration by parts, neither the integrated term nor the remaining integral will exist, even though the sum exists.

To get around this difficulty, the method given by Freund (1973) and applied very successfully by Ma and Burgers (1986a, 1986b) is used. If the lower limit of integration in equations (1.13) is replaced by $d+\epsilon$, $\epsilon << d$, it can be shown that those terms which are similar at $\epsilon=0$ exactly cancel each other. The desired result can then be obtained by taking the limit as $\epsilon \to 0$.

Integration by parts of (1.13) and making use of the explicit expression for $(\tau_{zy}^s)_o$ in equation (1.7) yields

$$\tau_{zy}^N = \lim_{\epsilon\to 0} -\frac{\tau_o}{\pi}\int_{d+\epsilon}^{d^*}\left\{\int_{t_s}^{t-t^*} Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{(\lambda_s-h)}\right]ds\right\}\cdot \frac{d^{3/2}}{\pi b^{1/2}(h-d)^{3/2}h}dh$$

$$+\frac{\tau_o}{\pi}\left(\frac{2d}{\pi b^{1/2}\epsilon^{1/2}}-1\right)\int_{t_s}^{t-t_f} Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{(\lambda_s-d)d^{1/2}}\right]ds. \qquad (1.14)$$

Evaluation of the real integral in the first term of equation (1.14) becomes possible by converting it into a line integral in the complex h plane and changing the contour of integration. The integrand has branch point singularities at $h=d,d^*$, and is single-valued in the plane cut along the real axis between d and d^*. Applying Cauchy's integral theorem and working out the details, it is found that

$$\tau_{zy}^N = \frac{\tau_o}{\pi b^{1/2}}\int_{t_s}^{t} Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{\lambda_s}\right]ds - \frac{\tau_o}{\pi d^{1/2}}\int_{t_s}^{t-t_f} Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{\lambda_s-d}\right]ds - \frac{\tau_o}{\pi b^{1/2}}\int_{br}^{t} Im\left[\frac{\beta_+(\lambda_o)\partial\lambda_o/\partial s}{\lambda_o}\right]ds, \qquad (1.15a)$$

where

$$\lambda_o = \lambda_s(v_o\to 0) = -\frac{s}{r}\cos\theta + i\left(\frac{s^2}{r^2}-b^2\right)^{1/2}\sin\theta,$$

$$r=(x^2+z^2)^{1/2}.$$

By a similar procedure, we also get the result for τ_{xy}^N.

$$\tau_{xy}^N = \frac{\tau_o}{\pi b^{1/2}}\int_{t_s}^{t} Im\left[\frac{\partial\lambda_s/\partial s}{\beta_-(\lambda_s)}\right]ds + \frac{\tau_o}{\pi d^{1/2}}\int_{t_s}^{t-t_f} Im\left[\frac{\lambda_s\partial\lambda_s/\partial s}{(\lambda_s-d)\beta_-(\lambda_s)}\right]ds + \frac{\tau_o}{\pi b^{1/2}}\int_{br}^{t} Im\left[\frac{\partial\lambda_o/\partial s}{\beta_-(\lambda_o)}\right]ds, \qquad (1.15b)$$

It is very interesting to see that the full field solution of a crack propagating with constant speed subjected to loads on the new crack faces equal to the stresses along the crack line in the original stationary crack problem can be expressed by three terms, each term having its own physical meaning. The first term represents the solution due to the uniform loading on the original and new crack faces for a crack which begins to grow at constant speed at a time t_f after the loading is applied. The functional form of this first term is the same as the solution to the problem with uniform loading on the original and new crack faces with no delay time. The only dependence on delay time is through the definition of ξ. This problem corresponds to Baker's problem (Baker, 1962) in Mode III. The second term represents the solution for a crack which starts to propagate at time t_f with constant speed with loading applied uniformly on the new crack faces only. The third term represents the solution for a stationary crack subject to uniform loading on original crack faces with the crack tip located at $x=0,z=0$.

The full field solution of a crack propagating at constant speed with uniform loading on the original crack faces only can be represented by the solution shown in equations (1.15) plus the stationary crack solution, which is exactly the same as the third term in equations (1.15). Hence, the solution has only two terms as follows

$$\tau_{zy}^O = \frac{\tau_o}{\pi b^{1/2}}\int_{t_s}^{t} Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{\lambda_s}\right]ds - \frac{\tau_o}{\pi d^{1/2}}\int_{t_s}^{t-t_f} Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{\lambda_s-d}\right]ds, \qquad (1.16a)$$

$$\tau_{xy}^O = \frac{\tau_o}{\pi b^{1/2}}\int_{t_s}^{t} Im\left[\frac{\partial\lambda_s/\partial s}{\beta_-(\lambda_s)}\right]ds + \frac{\tau_o}{\pi d^{1/2}}\int_{t_s}^{t-t_f} Im\left[\frac{\lambda_s\partial\lambda_s/\partial s}{\beta_-(\lambda_s)(\lambda_s-d)}\right]ds, \qquad (1.16b)$$

If the loading condition is an incident stress wave parallel to the original crack faces, the full field solution of a propagating crack has only the first term in equations (1.16) plus the stress wave loading. The second term is exactly canceled out by the uniform loading on the new crack faces due to the incident stress wave loading, so that

$$\tau_{zy}^W = \frac{\tau_o}{\pi b^{1/2}}\int_{t_s}^{t} Im\left[\frac{\beta_+(\lambda_s)\partial\lambda_s/\partial s}{\lambda_s}\right]ds + \tau_o. \qquad (1.17a)$$

$$\tau_{xy}^W = \frac{\tau_o}{\pi b^{1/2}}\int_{t_s}^{t} Im\left[\frac{\partial\lambda_s/\partial s}{\beta_-(\lambda_s)}\right]ds. \qquad (1.17b)$$

The solution shown in equations (1.17) represents a very interesting result that the stress field of the propagating crack due to stress wave loading is independent of delay time t_f and has exactly the same functional form as the stress field of the Baker's problem in Mode III plus the stress wave loading part. This result was also observed by Achenbach and Nuismer (1971) for the Mode I case with no delay time effect. The special properties mentioned above will make it possible to determine the full field solution of the suddenly stopped crack in the next section.

In order to investigate the effect of the crack propagation speed on the stress intensity factor dominated part of the stress field, we have also calculated the ratio of the exact stresses (1.17a,b) around the crack tip and the singular part of the stress field. The singular part is given by

$$\tau_{zy}^W \approx \frac{2\tau_o(1-b/d)^{1/2}}{\pi}\left(\frac{t}{r_s b}\right)^{1/2}\cos(\theta_s/2), \qquad (1.18a)$$

$$\tau_{xy}^W \approx \frac{2\tau_o}{\pi(1+b/d)^{1/2}}\left(\frac{t}{r_s b}\right)^{1/2}\sin(\theta_s/2), \qquad (1.18b)$$

where

$$r_s = [\xi^2 + (1-b^2/d^2)z^2]^{1/2},$$

$$\theta_s = \tan^{-1}[(1-b^2/d^2)^{1/2}z/\xi].$$

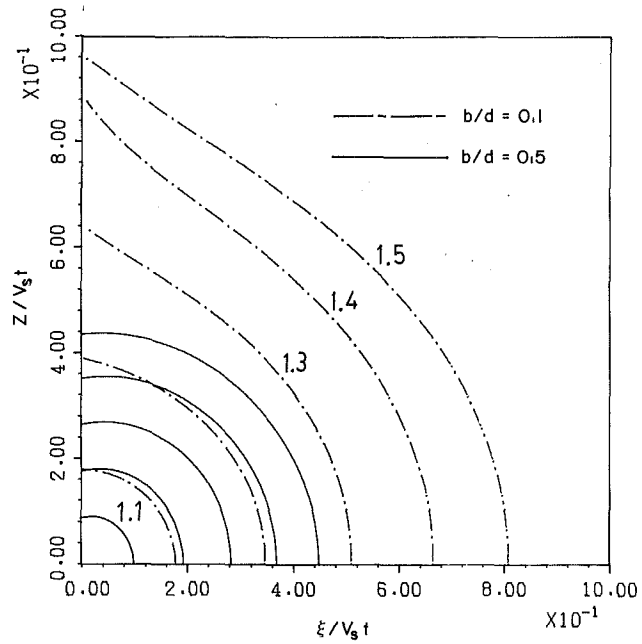


Fig. 5 The ratio of τ_{zy}^{W} for the exact solution and the stress intensity factor controlled field due to stress wave loading

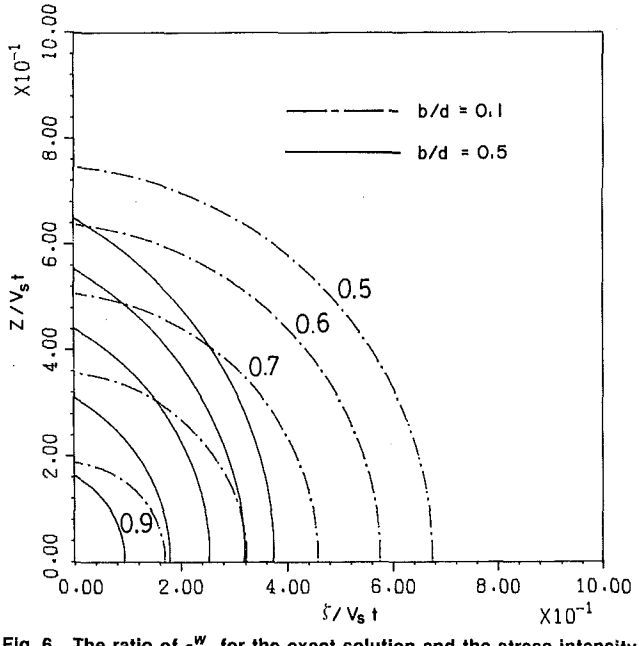


Fig. 6 The ratio of τ_{xy}^{W} for the exact solution and the stress intensity factor controlled field due to stress wave loading on the crack faces

In Fig. 5 the position of the fixed ratios 1.1, 1.2, 1.3, 1.4, and 1.5 of τ_{zy}^{W} for the full field and the stress intensity factor dominated field is plotted for the propagating crack due to stress wave loading. The dashed line is for the crack tip speed $v_o = 0.1v_s$ and the solid line is for crack tip speed $v_o = 0.5v_s$. A similar result is plotted for τ_{xy}^{W} in Fig. 6. If we compare the results for the propagating crack case (Fig. 5) with the results of the stationary crack case (Fig. 4), we see that there is little difference in the area around the crack tip dominated by the stress intensity factor for the slow crack tip speed and the stationary crack case, but that for the high crack speed $v_o = 0.5v_s$ there is a very significant difference. If we accept that a ratio of 1.1 defines the region in which the singular term adequately characterizes the actual stress field, then the region is reduced from approximately 20 percent of the distance to the shear wave front for slow crack speeds to approximately 10 percent of the distance for the crack speed $v_o = 0.5v_s$.

We have shown above that the stress field of a propagating crack subjected to stress wave loading is independent of delay time t_f. Is this also true for the displacement $w(\xi,0,t)$ on the fracture plane?

We start with the fundamental solution for the boundary condition in equations (1.8) and (1.9). The displacement on the crack faces behind the moving crack tip is

$$w^F(\zeta,0,t) = -\frac{nh}{\mu\pi\beta_+(h)}\int_{b_1}^{t/\zeta} Im\left[\frac{1}{(h-\eta)\beta_-(\eta)}\right]d\eta$$

$$-\frac{mh^2}{\mu\pi}\int_{b_1}^{t/\zeta}(t-\eta\zeta)Im\left[\frac{1}{(\eta-h)\beta_-(\eta)\beta_+(h)}\right]_h d\eta \qquad (1.19)$$

where ζ is the negative of ξ. Following the same method as for the construction of the stress component above, the displacement $w^N(\zeta,0,t)$ on the crack faces behind the moving crack tip due to the stationary crack stress distribution applied on the new crack faces yields

$$w^N(\zeta,0,t) = \int_0^{u_o^*} w^F(\zeta,0,t-t^*,m=-1,n=-t^*)(\tau_{zy}^s)_o du$$

$$= \frac{\tau_o}{\mu\pi}\int_d^{d_o^*}\left\{\int_{b_1}^{(t-t^*)/\zeta}(t-t^*-\zeta\eta)\right.$$

$$Im\left[\frac{1}{(\eta-h)\beta_-(\eta)\beta_+(h)}\right]d\eta\Big\}_h (\tau_{zy}^s)_o dh =$$

$$-\frac{\tau_o}{\mu\pi}\int_{b_1}^{t/\zeta}\frac{t-\zeta\eta}{\eta b^{1/2}}Im\left[\frac{1}{\beta_-(\eta)}\right]d\eta$$

$$+\frac{\tau_o}{\mu\pi}\int_{b_1}^{(t-t_f)/\zeta}\frac{t-t_f-\zeta\eta}{(\eta-d)\beta_+(d)}Im\left[\frac{1}{\beta_-(\eta)}\right]d\eta, \qquad (1.20)$$

where

$$u_o^* = \frac{t-t_f-\zeta b_1}{d(t-\zeta b_1)},\quad d_o^* = \frac{(t-\zeta b_1)d}{t_f},\quad t>t_f,$$

and

$$b_1 = \frac{b}{1+b/d}.$$

The first term in (1.20) is the same as for the uniform loading applied on the original and new crack faces. The second term in (1.20) is the solution due to applied uniform loading on the new crack faces. Hence, for the incident stress wave loading condition, the second term will be canceled out and only one term with remains

$$w^W(\zeta,0,t) = -\frac{\tau_o}{\mu\pi}\int_{b_1}^{t/\zeta}\frac{t-\zeta\eta}{\eta b^{1/2}}Im\left[\frac{1}{\beta_-(\eta)}\right]d\eta,\quad t>t_f. \qquad (1.21)$$

Again, the displacement on the crack faces due to stress wave loading is independent of the delay time t_f and has exactly the same functional form as for the solution of displacement for Baker's problem in Mode III.

The Stress Field Solution for the Stopped Crack

The special properties of the elastic field which have been demonstrated in last section are essentially all that is needed to obtain the stress field of the stopped crack. In order to stop the crack at the position $x = l_s < v_o(t-t_f)$ due to the stress wave loading, we must close the displacement of the crack faces for $l_s < x < v_o(t-t_f)$. The procedure to construct the stress field of stopped crack follows Freund (1973).

Consider a modified problem of a semi-infinite crack with the crack tip at $x = l_s = v_o(t_b - t_f)$ and no loads acting on the body. At time $t = t_f + l_s/v_o$, the traction distribution described

by $(\tau^s_{zy})_o[t/(\xi+v_o(t-t_f-l_s/v_o))]+\tau_o$ acts on the new crack faces $l_s<x<v_o(t-t_f)$. If we follow the same method used in the last section to construct the displacement field on the new crack faces $l_s<x<v_o(t-t_f)$ for this modified problem, we will get the same result as shown in equation (1.21) which is the solution for the stress wave loading. Subtracting the displacement solution for the modified problem from the original stress wave solution will make the crack suddenly stop at $x=l_s$. The new problem will then be an incident stress wave striking the crack at time $t=0$, the crack-tip beginning to move at constant speed v_o at time $t=t_f$, and after some time later $t=t_f+l_s/v_o$, the crack tip suddenly stopping at $x=l_s$.

The stress field of this stopped crack can be obtained by subtracting the solution for the modified problem from the stress wave loading solution. Finally, it is found that the stopped crack solution is given by

$$\tau_{zy}(x,z,t)=\tau^s_{zy}(t/[(x-l_s)^2+z^2]^{1/2})+\tau_o, \tag{2.1a}$$

$$\tau_{xy}(x,z,t)=\tau^s_{xy}(t/[(x-l_s)^2+z^2]^{1/2}), \tag{2.1b}$$

where

$$0<[(x-l_s)^2+z^2]^{1/2}<v_s(t-t_f-l_s/v_o)$$

Hence, the main result is that, upon stopping of the crack propagating at constant speed subjected to stress wave loading, a stress distribution is radiated out from the stopped crack tip, traveling with the shear wave speed which is the exact stress distribution that a stationary crack tip would have if it had been at $x=l_s$ for all time. The stress distribution is completely independent of the fact that the crack-tip had been in motion. Knowing the stress field of the stopped crack, we can construct the solution of the kinked crack subjected to this stress field.

The Stress Intensity Factor for the Kinked Crack

A semi-infinite crack is at rest for $t<0$, the incident stress wave loading strikes the crack tip at time $t=0$. At time $t=t_f$, the crack starts in motion straight ahead of the fracture plane with constant velocity v_o and stops at the position $x=l_s$. Instantaneously, the crack kinks from the stopped crack tip and propagates with constant speed v_c, making an angle δ with respect to the crack plane. The perturbation method shown in Achenbach et al. (1984) and Ma and Burgers (1986a,1986b) is applied to calculate the dynamic stress intensity factor for the kinked crack. The first order approximation of the dynamic stress intensity factor for the 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 tractions given in equations (2.1) on the kinked crack faces. Since the stress field in this problem is very similar to that in Ma and Burgers (1986a), the solution of the dynamic stress intensity factor obtained there can be used to give the solution to this problem after a slight modification.

In Ma and Burgers (1986a), the dynamic crack kinks out of the original semi-infinite crack, propagating with constant velocity v_c at an angle δ to the original crack, at some time t_f after the stress wave loading with the incident angle α initially interacts with the crack tip. The first order approximation solution for the stress intensity factor is written in closed form (Ma and Burgers,1986a) as

$$K=2\tau_o\left[\frac{2(1-b/d_c)}{\pi(1+\sin\alpha)}\right]^{1/2}\left\{\frac{1}{2b^{1/2}}\left(A_1\left[t-\frac{b}{d_c}(t-t_f)\sin\Theta_1\right]^{1/2}\right.\right.$$
$$+A_2\left[t+\frac{b}{d_c}(t-t_f)\sin\Theta_2\right]^{1/2}\Big)-\frac{(t-t_f)^{1/2}}{2d_c^{1/2}}\Big[A_1\sqrt{1-\sin\Theta_1}$$
$$\left.\left.+A_2\sqrt{1+\sin\Theta_2}\right]+\left[\frac{(1+\sin\alpha)(t-t_f)}{d_c}\right]^{1/2}\cos\Theta_1\right\}, \tag{3.1}$$

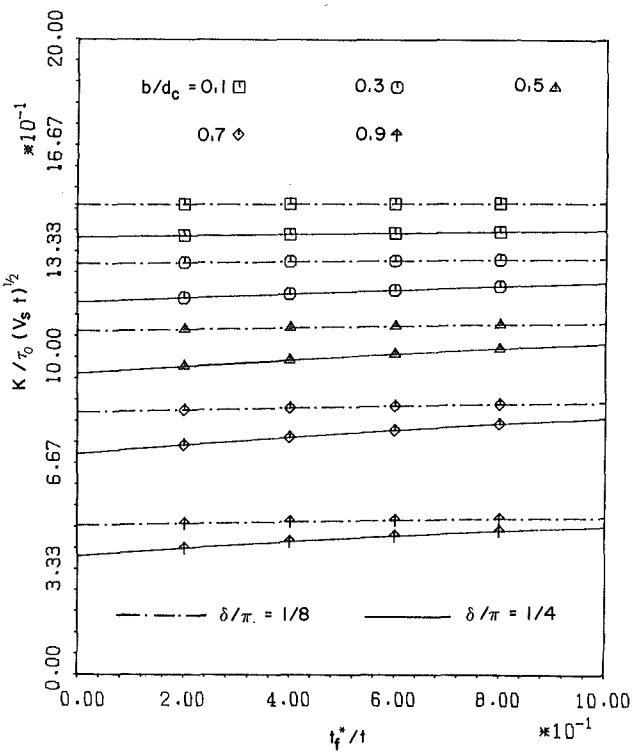


Fig. 7 The dynamic stress intensity factor of the kinked crack for kink angles $\delta=\pi/8,\pi/4$

where

$$A_1=\left[\cos\left(\alpha-\frac{\delta}{2}\right)+\sin\frac{\delta}{2}\right],$$

$$A_2=\left[\cos\left(\alpha+\frac{\delta}{2}\right)-\sin\frac{\delta}{2}\right],$$

$$\Theta_1=\delta-\alpha,\ \Theta_2=\delta+\alpha,$$

and

$$d_c=\frac{1}{v_c}.$$

This closed form solution (3.1) gives very accurate results, the error for the first order approximation with delay time equal zero when compared to the exact values in Dempsey et al. (1982) is less than 10 percent for any combination of angle of incident wave, kinked angle and propagation speed of the crack. For the kinking angle less than $\pi/4$, the error is even less than 2 percent. For the incident stress angle $\alpha=0$, equation (3.1) becomes

$$K=2\tau_o\left[\frac{2(1-b/d_c)}{\pi}\right]^{1/2}\left\{\frac{1}{2b^{1/2}}\left(B_1\left[t-\frac{b}{d_c}(t-t_f)\sin\delta\right]^{1/2}\right.\right.$$
$$\left.+B_2\left[t+\frac{b}{d_c}(t-t_f)\sin\delta\right]^{1/2}\right)-\frac{(t-t_f)^{1/2}}{2d_c^{1/2}}[B_1\sqrt{1-\sin\delta}$$
$$\left.+B_2\sqrt{1+\sin\delta}]+\left[\frac{(t-t_f)}{d_c}\right]^{1/2}\cos\delta\right\}, \tag{3.2}$$

where

$$B_1=[\cos\frac{\delta}{2}+\sin\frac{\delta}{2}],\ B_2=[\cos\frac{\delta}{2}-\sin\frac{\delta}{2}].$$

The stress intensity factor as shown in equation (3.2) is for the stationary crack kinking at some delay time t_f after the parallel stress wave loading strikes the crack tip. For the crack which stops at $x=l_s$ and kinks instantaneously, from the results of the last section, we can consider the crack to be at

$x=l_s$ for all the time and to kink after the delay time $t_f^*=t_f+l_s/v_o$ due to the same stress wave loading. Hence, just replacing t_f by t_f^* in equation (3.2) will give the dynamic stress intensity factor for the crack which kinks instantaneously after stopping. It is noted that for the kinking angle $\delta=0$, the solutions are exact. The dynamic stress intensity factor for the kinked crack is shown in Fig. 7. We choose two kinking angles $\delta=\pi/8,\pi/4$ and different kinking crack velocity. The normalized time t_f^*/t is chosen as the parameter so that the instant of initiation of the kink is at $t_f^*/t=1$, while $t_f^*/t=0$ corresponds to when the kinked crack has propagated for an infinite time compared to the delay time t_f^*. The stress intensity factor is almost constant for small kinked angles which indicates that it only weakly depends on the delay time t_f^* for these kinked angles.

The stress intensity factor just after the kinking occurs has the form:

$$K=2\tau_o\left[\frac{2t}{b\pi}\right]^{1/2}(1-b/d_c)^{1/2}\cos\frac{\delta}{2}$$

$$=K^s(1-b/d_c)^{1/2}\cos\frac{\delta}{2}\ \text{as}\ t\to t_f^*. \tag{3.3}$$

In equation (3.3), we can see that the stress intensity factor immediately after kinking is expressed in the form of an universal function of the crack tip speed $(1-b/d_c)^{1/2}$ times the stress intensity factor appropriate for the given diffraction process with no crack-tip motion K^s times a spatial angular dependence $\cos\delta/2$.

Just before kinking, the crack propagates straight in the fracture plane with velocity v_o. The stress intensity factor is given by

$$K=K^s(1-b/d)^{1/2}. \tag{3.4}$$

Hence, the stress intensity factor just before and after the kinked crack has a jump as shown in equations (3.3) and (3.4). The magnitude of this jump depends on the crack speed v_o and v_c, and the kinked angle δ.

If the crack does not kink instantaneously as the crack stops, that is, it kinks after some delay time $\hat{t}_f$, the stress intensity factor of the kinked crack for this case can be obtained by replacing t_f by $t_f+l_s/v_o+\hat{t}_f$ in equation (3.2).

The Criterion for Kinking

For Mode III fracture, the energy flux into the propagating crack tip can be written in terms of the corresponding dynamic stress intensity factor as

$$E=\frac{K^2}{2\mu d(1-b^2/d^2)^{1/2}}. \tag{4.1}$$

First, we consider a maximum energy release rate criterion for the kinked crack. This criterion suggests that the kinked crack will choose to propagate in the direction and at a velocity for which the energy flux into the crack tip is a maximum. Some results based on this criterion are discussed by Achenbach et al. (1984) and Ma and Burgers (1986a). For the case discussed in the previous section of crack subjected to a planar antiplane stress pulse parallel to the crack tip which starts to propagate and then suddenly stops, from the discussion in Ma and Burgers (1986a) using the maximum energy release rate criterion, the crack will tend to propagate straight ahead and no kink will occur.

An alternative fracture criterion was discussed by Freund (1975). Consider a propagating crack that may kink. At each instant, including the instant just prior to bifurication, the rate of energy supplied to the crack tip by the loads is equal to the rate of energy required to drive the crack at its current speed. Since the time for bifurication to occur is very short, the local stress and strain fields cannot readjust by any significant amount. Therefore, it seems reasonable to assume that the energy release rate just after kinking has the same value as immediately before kinking. The stress intensity factor just before and just after kinking are given in equations (3.4) and (3.3). The energy balance criterion is then given by

$$\frac{(1-b/d)^{1/2}}{d(1+b/d)^{1/2}}=\frac{(1-b/d_c)^{1/2}}{d_c(1+b/d_c)^{1/2}}\cos^2\delta/2. \tag{4.2}$$

Substituting some typical values for b/d give the following possiblities:

If $b/d=0.1$

$\delta=30^\circ\ b/d_c=0.108$
$\delta=45^\circ\ b/d_c=0.109$
$\delta=60^\circ\ b/d_c=0.139$

If $b/d=0.2$

$\delta=30^\circ\ b/d_c=0.219$
$\delta=45^\circ\ b/d_c=0.246$
$\delta=60^\circ\ b/d_c=0.295$

Discussion

During the initial stage of crack growth while the speed of the crack tip is relatively low, the stresses are largest directly ahead of the crack tip. The fracture surface remains planar and relatively smooth. As the crack speed increases, the maximum stresses in the crack tip region move away from a line straight ahead of the crack. This leads to microcracking at points adjacent to the main fracture path and leads to increased roughness of the fracture surface. This microcracking might be viewed as the mechanism by which the main crack probes for alternative crack paths. If no such paths are found, then the planar crack growth continues. However, if there are alternative paths, which are more favorable, then such a path will be chosen and branching (or kinking) will occur.

The analytical investigations on the kinked crack due to dynamic loading published recently, are restricted to geometries where the kinked crack initiates out of the original stationary crack tip. From experimental observations, it is clear that the crack almost always propagates straight ahead for a short period of time and then kinks. In this paper, we have tried to make the analytical model be a closer representation of the experiments than has been previously achieved in that the initial propagation straight ahead and stopping of the crack is modeled, before kinking occurs. The significant result is that upon stopping, for all angles the stresses behind the shear wave front emitted from the crack tip are the same as for a stationary crack, with its crack tip located at the position the propagating crack stopped at for all time, subjected to the same loading as the propagating crack. The stress intensity factor for the kinked crack is constructed by negating the stress distribution of the stopped crack along the kinked crack line. It is also shown that the first order approximation of the stress intensity factor immediately after kinking can be expressed in the form of a universal function of the crack tip speed times the stress intensity factor appropriate for the given diffraction process with no crack tip motion times a spatial angular dependence.

The accuracy of the first order approximation for the stress intensity factor for the kinked crack should still be checked. After the kinked crack has propagated for a long time (i.e., $t_f^*/t\to 0$), the error in the approximation is quite small, as shown in Ma and Burgers (1986a). This good agreement suggests that the wedge geometry of the kinked crack has only a minor effect on the dynamic stress intensity factor. In studying the criterion for crack kinking, it is obvious that the times of importance are immediately before and after kinking, i.e., for t_f^*/t near 1. Unfortunately, there are no other results

available which are accurate in this time range, so the accuracy of the first order approximation is still unknown.

The size and shape of the region of the stress field which is dominated by the stress intensity factor is found to be a function of the crack tip speed, becoming smaller as the crack speed increases. The implication for experiments which use optical methods and essentially measure the deformation a small distance away from the crack tip, is that the region in which the measurements are taken is smaller for higher crack speeds. This may mean that some of the available data for stress intensity factors at higher crack speeds was measured in regions where the stress intensity factor was not quite as dominant as it should be for accurate data interpretation. Therefore, older data for dynamic stress intensity factors at the higher crack speeds should be checked.

Acknowledgments

The authors would like to thank Professor L. B. Freund for helpful discussions. The support of the authors by the National Science Foundation, Solid Mechanics Program, through Grant MEA-8306644 is gratefully acknowledged. The calculations were performed in the VAX-11/780 Computational Mechanics Facility at Brown University. This faculty was made possible by grants from the National Science Foundation (Solid Mechanics Program), the General Electric Foundation, and the Digital Equipment Corporation.

References

Achenbach, J. D., Kuo, M. K., and Dempsey, J. P., 1984, "Mode III and Mixed Mode I-II Crack Kinking under Stress-Wave Loading," *Int. J. Solids Structures*, Vol. 20, pp. 395–410.

Achenbach, J. D., and Nuismer, R., 1971, "Fracture Generated by a Dilatational Wave," *International Journal of Fracture Mechanics*, 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.

Burgers, P., 1982, "Dynamic Propagation of a Kinked or Bifurcated Crack in Antiplane Strain, ASME Journal of Applied Mechanics, Vol. 49, pp. 371–376.

Burgers, P., 1983, "Dynamic Kinking of a Crack in Plane Strain," *Int. J. Solids Structures*, Vol. 19, pp. 735–752.

Burgers, P., and Dempsey, J. P., 1982, "Two Analytical Solutions for Dynamic Crack Bifurcation in Antiplane Strain," ASME Journal of Applied Mechanics, Vol. 49, pp. 366–370.

Burgers, P., and Dempsey, J. P., 1984, "Plane Strain Dynamic Crack Bifurcation," *Int. J. Solids Structures*, Vol. 20, pp. 609–618.

Dempsey, J. P., Kuo, M. K., and Achenbach, J. D., 1982, "Mode III Crack Kinking under Stress Wave Loading," *Wave Motion*, Vol. 4, pp. 181–190.

Dempsey, J. P., Kuo, M. K., and Bentley, D. L., 1986, "Dynamic Effects in Mode III Crack Bifurcation," *Int. J. Solids Structures*, Vol. 22, pp. 333–353.

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., 1974, "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., 1975, "Bifurcation of a Propagating Crack in a Brittle Solid," *12th Annual Meeting, Society of Engineering Science*, University of Texas Press, Austin, pp. 1015–1022.

Ma, C. C., and Burgers, P., 1985, Recent Research Results.

Ma, C. C., and Burgers, P., 1986, "Mode III Crack Kinking with Delay Time: An Analytical Approximation," *Int. J. Solids Structures*, Vol. 22, pp. 883–899.

Ma, C. C., and Burgers, P., 1986, "Dynamic Mode I and Mode II Crack Kinking Including Delay Time Effects," *Int. J. Solids Structures*, in press.

Ravi-Chandar, K., and Knauss, W. G., 1984a, "An Experimental Investigation into Dynamic Fracture: I. Crack Initiation and Arrest," *International Journal of Fracture*, Vol. 25, pp. 247–262.

Ravi-Chandar, K., and Knauss, W. G., 1984b, "An Experimental Investigation into Dynamic Fracture: II. Microstructural Aspects," *International Journal of Fracture*, Vol. 26, pp. 65–80.

Ravi-Chandar, K., and Knauss, W. G., 1984c, "An Experimental Investigation into Dynamic Fracture: III. On Steady-State Crack Propagation and Crack Branching," *Interational Journal of Fracture*, Vol. 26, 141–154.

Ravi-Chandar, K., and Knauss, W. G., 1984d, "An Experimental Investigation into Dynamic Fracture: IV. On the Interation of Stress Waves with Propagating Cracks," *International Journal of Fracture*, Vol. 26, 189–200.

Ravi-Chandran, G., 1986, "Dynamic Fracture under Plane Wave Loading," Ph.D. Thesis, Brown University, Providence, RI.

A Crack in an Anisotropic Layered Material Under the Action of Impact Loading

W. T. Ang*

Department of Applied Mathematics,
University of Adelaide,
Adelaide, South Australia 5001

The problem of a plane crack in an anisotropic layered material under the action of impact loading is considered in this paper. The problem is reduced in the Laplace transform domain to a set of simultaneous Fredholm integral equations of the second kind. Once these integral equations are solved, the crack tip stress intensity factors in the Laplace transform domain may be readily calculated. The dynamic stress intensity factors can then be obtained through the use of a numerical technique for inverting Laplace transforms. Numerical results are given for specific examples involving particular transversely isotropic materials.

1 Introduction

In recent years, the problem of a crack in an elastic material under the action of impact or shock loading has been a subject of considerable interest. Such a problem is of practical importance since the impact loading can give rise to displacements and stresses which are higher in magnitudes than those induced by the corresponding static loading. However, very few solutions for this type of problem can be found in the literature, compared with those already given for problems involving static loading. Sih et al. (1972) and Kassir and Bandyopadhyay (1983) have studied the problem for infinite isotropic and orthotropic materials, respectively. Chen and Sih (1977) and Chen (1978) have examined how the boundary of a material can affect the transient response of a crack to the impact loading by placing the crack in an infinitely long isotropic elastic strip.

Clements (1979) has solved a two-dimensional elastostatic problem involving a crack in an anisotropic layer sandwiched between two anisotropic half-spaces. In the present paper, we consider the same problem for the case where the anisotropic layered material is subject to an impact loading. Firstly, we remove the time dependence of the displacements and stresses by using the Laplace transform and develop the necessary basic equations for the problem. With the basic equations in hand, we then reduce the problem to a set of Fredholm integral equations of the second kind by using an analysis similar to that given in Clements. Once these integral equations are solved, the crack tip stress intensity factors in the Laplace transform domain can be readily evaluated. These factors can then be obtatined in the physical domain by using a numerical technique for inverting Laplace transforms. Numerical results for specific examples concerning particular transversely isotropic materials are obtained and the numerical procedure involved in obtaining the results is described in detail.

2 Basic Equations

The behavior of an anisotropic elastic material is governed by a system of linear partial differential equations given by

$$c_{ijkl}\frac{\partial^2 u_k}{\partial x_l \partial x_j}=\rho\frac{\partial^2 u_i}{\partial t^2}, \tag{2.1}$$

where $i, j, k, l = 1, 2, 3$, x_i are the spatial coordinates with respect to a Cartesian frame, t is the time coordinate, c_{ijkl} are the elastic moduli of the material, ρ is the density of the material (assumed to be constant here), and $u_i = u_i(x_1, x_2, x_3, t)$ are the Cartesian displacements. The usual convention of summing over a repeated Latin subscript holds here.

If we impose the initial conditions

$$u_i=0 \quad \text{and} \quad \frac{\partial u_i}{\partial t}=0 \quad \text{at} \quad t=0, \tag{2.2}$$

then, by taking Laplace transform with respect to t of both sides of equations (2.1), we obtain

$$c_{ijkl}\frac{\partial^2 U_k}{\partial x_l \partial x_j}-\rho s^2 U_i=0, \tag{2.3}$$

where

$$U_i(x_1,x_2,x_3,s)=\int_0^\infty u_i(x_1,x_2,x_3,t)\exp(-st)\,dt, \tag{2.4}$$

where s is the Laplace transform parameter.

Provided that we take the parameter s as a real number and assume that U_k are independent of x_3, a general expression for U_k which satisfy the system (2.3) may be given by

*Presently Assistant Professor, Department of Mathematics, Iowa State University, Ames, IA 50011.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, November 20, 1986; final revision July 1, 1987.

$$U_k(x_1,x_2,s) = \Re\left\{\int_0^\infty \sum_\alpha A_{k\alpha}(\xi,s)[\phi_\alpha(\xi,s)\right.$$

$$\left.\exp(i\xi z_\alpha) + \psi_\alpha(\xi,s)\exp(-i\xi z_\alpha)]d\xi\right\}, \quad (2.5)$$

where Σ_α denotes summation over $\alpha = 1, 2, 3$; $i = \sqrt{-1}$, ϕ_α and ψ_α are arbitrary functions, $z_\alpha = x_1 + \tau_\alpha(\xi, s)x_2$, and $\tau_\alpha = \tau_\alpha(\xi, s)$ are the roots with positive imaginary parts of the sextic equation

$$\det[(c_{i1k1} + \tau c_{i1k2} + \tau c_{i2k1} + \tau^2 c_{i2k2})\xi^2 + \rho s^2 \delta_{ik}] = 0, \quad (2.6)$$

where δ_{ik} is the kronecker-delta, and $A_{k\alpha}$ is a given solution of the system of linear algebraic equations

$$[(c_{i1k1} + \tau c_{i1k2} + \tau c_{i2k1} + \tau^2 c_{i2k2})\xi^2 + \rho s^2 \delta_{ik}]A_k = 0, \quad (2.7)$$

when $\tau = \tau_\alpha$. By following a line of reasoning similar to the one given in Eshelby et al. (1953), it can be shown that, since s is taken as a real number, the sextic equation in (2.6) does not admit real roots. Hence the roots of equation (2.6) occur in complex conjugate pairs. For our purpose here, we assume that these roots are distinct. Note that, in the limit as $s/\xi \to 0$, $\tau_\alpha \to \tau_\alpha^0$ and $A_{k\alpha} \to A_{k\alpha}^0$, where τ_α^0 and $A_{k\alpha}^0$ are the corresponding constants defined in Clements (1979). In fact, it is easy to show that $\tau_\alpha - \tau_\alpha^0$ and $A_{k\alpha} - A_{k\alpha}^0$ behave as $O(s^2/\xi^2)$ for small s/ξ.

Now, from equations (2.4)–(2.5) and the stress-displacement relation

$$\sigma_{ij} = c_{ijkl}\frac{\partial u_k}{\partial x_l}, \quad (2.8)$$

where $\sigma_{ij} = \sigma_{ij}(x_1, x_2, t)$ are the Cartesian stresses, we obtain

$$S_{ij}(x_1,x_2,s) = \Re\left\{\int_0^\infty i\xi \sum_\alpha L_{ij\alpha}(\xi,s)[\phi_\alpha(\xi,s)\exp(i\xi z_\alpha)\right.$$

$$\left.-\psi_\alpha(\xi,s)\exp(-i\xi z_\alpha)]d\xi\right\}, \quad (2.9)$$

where

$$S_{ij}(x_1,x_2,s) = \int_0^\infty \sigma_{ij}(x_1,x_2,t)\exp(-st)dt, \quad (2.10)$$

and

$$L_{ij\alpha}(\xi,s) = [c_{ijk1} + \tau_\alpha(\xi,s)c_{ijk2}]A_{k\alpha}(\xi,s). \quad (2.11)$$

3 Mathematical Statement of the Problem

Consider an anisotropic layered material consisting of the regions $x_2 > h$ (region 1), $-h < x_2 < h$ (region 2), and $x_2 < -h$ (region 3). For simplicity, we assume that regions 1 and 3 contain the same material (that is, material whose elastic behavior is characterized by the same elastic moduli). The material in region 2 may be different from that in regions 1 and 3. A perfect bond is assumed to exist between the materials in the three regions. Region 2 contains a plane crack in the region $|x_1| < a$, $x_2 = 0$, $-\infty < x_3 < \infty$ (Fig. 1). The crack is opened by equal and opposite tractions suddenly applied on each of its faces. We assume that the geometry and the boundary conditions of the problem here do not change in the Ox_3 direction. We are interested in determining the displacements and the stresses throughout the material.

Mathematically, the problem is one of solving the system of equations (2.1) subject to the initial conditions (2.2) and

$$\sigma_{j2}(x_1,0,t) = P_j(x_1)q(t) \quad \text{for} \quad |x_1| < a, \quad (3.1)$$

where $P_j(x_1)$ and $q(t)$ are suitably prescribed functions of x_1 and t, respectively. We shall assume here that $P_j(x_1)$ are even functions of x_1.

In the Laplace transform domain (henceforth referred to as the s domain), the problem is then to find the arbitrary functions ϕ_α and ψ_α in (2.5) and (2.9) such that

$$S_{j2}(x_1,0,s) = P_j(x_1)Q(s) \quad \text{for} \quad |x_1| < a, \quad (3.2)$$

where

$$Q(s) = \int_0^\infty q(t)\exp(-st)dt. \quad (3.3)$$

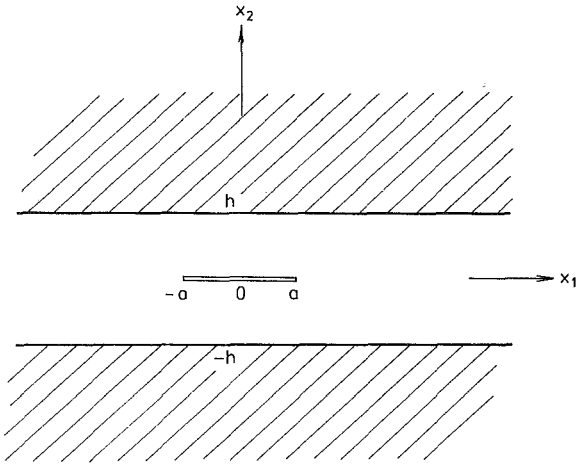


Fig. 1 A crack in an anisotropic layered material

4 Solutions in the s Domain

Guided by the analysis in Clements (1979), we choose the transformed displacements U_k in regions 1 and 3 to be given by

$$U_k(x_1,x_2,s) = \text{sgn}(x_2)\Re\left\{\int_0^\infty \sum_\alpha\right.$$

$$\left.A_{k\alpha}^*(\xi,s)\chi_\alpha(\xi,s)\exp(\text{sgn}(x_2)i\xi z_\alpha^*)d\xi\right\}, \quad (4.1)$$

where $z_\alpha^* = x_1 + \tau^*(\xi, s)x_2$, τ_α^* and $A_{k\alpha}^*$ are obtained from equations (2.6)–(2.7) using the elastic moduli of the material in regions 1 and 3, sgn (x_2) denotes the sign of x_2, and $\chi_\alpha(\xi, s)$ are arbitrary functions yet to be determined. In region 2, U_k are taken to be

$$U_k(x_1,x_2,s) = \Re\left\{\int_0^\infty \sum_\alpha A_{k\alpha}[(E_\alpha(\xi,s)\right.$$

$$+H(x_2)M_{\alpha p}\Psi_p(\xi,s))\exp(i\xi z_\alpha) - (E_\alpha(\xi,s)$$

$$\left.+H(-x_2)M_{\alpha p}\Psi_p(\xi,s))\exp(-i\xi z_\alpha)]d\xi\right\}, \quad (4.2)$$

where $z_\alpha = x_1 + \tau(\xi, s)x_2$, τ_α and $A_{k\alpha}$ are obtained from equations (2.6)–(2.7) using the elastic moduli of the material in region 2, $H(x_2)$ is the Heaviside unit step function, $[M_{\alpha p}]$ is the inverse of $[L_{j2\alpha}]$ (as defined in equation (2.11)), $E_\alpha(\xi, s)$ are arbitrary functions to be determined, and $\Psi_p(\xi, s)$ are given by

$$\Psi_p(\xi,s) = iC_{pk}(\xi,s)\int_0^a r_k(t,s)J_0(\xi t)dt, \quad (4.3)$$

where $J_0(x)$ is a Bessel function of order zero, $r_k(t, s)$ are arbitrary functions to be determined, and $[C_{pk}]$ is the inverse of $[B_{pk}]$ defined by

$$B_{kp}(\xi,s) = i\sum_\alpha [A_{k\alpha}(\xi,s)M_{\alpha p}(\xi,s)$$

$$-\overline{A}_{k\alpha}(\xi,s)\overline{M}_{\alpha p}(\xi,s)], \quad (4.4)$$

where the bar denotes the conjugate of a complex number. It can be easily verified that U_k as given by equation (4.2) together with equation (4.3) are continuous throughout region 2 except on the faces of the crack.

From equations (2.9) and (4.1)–(4.2), the transformed stresses S_{ij} in regions 1 and 3 are

$$S_{ij}(x_1,x_2,s) = \Re\left\{\int_0^\infty i\xi \sum_\alpha L^*_{ij\alpha}(\xi,s)\chi_\alpha(\xi,s)\exp(\mathrm{sgn}(x_2)i\xi z^*_\alpha)d\xi\right\}, \tag{4.5}$$

where $L^*_{ij\alpha}$ are as defined in equation (2.11) using the elastic moduli of the material in regions 1 and 3 and with $A_{k\alpha}$ and τ_α being replaced by $A^*_{k\alpha}$ and τ^*_α, respectively, while in region 2 S_{ij} are given by

$$S_{ij}(x_1,x_2,s) = \Re\left\{\int_0^\infty i\xi \sum_\alpha L_{ij\alpha}[(E_\alpha(\xi,s) + H(x_2)M_{\alpha p}\Psi_p(\xi,s))\exp(i\xi z_\alpha) + (E_\alpha(\xi,s) + H(-x_2)M_{\alpha p}\Psi_p(\xi,s))\exp(-i\xi z_\alpha)]d\xi\right\}. \tag{4.6}$$

Note that the transformed stresses S_{j2} as given in equation (4.6) are continuous in the plane containing the crack, that is, on the plane $x_2 = 0$.

Since we assume that the different materials making up the anisotropic layered material are perfectly bonded, it is required that the arbitrary functions in equations (4.1)–(4.2) are chosen in such a way that the displacements and the tractions are continuous at the interfaces separating two different materials. Setting

$$\chi_\alpha(\xi,s) = M^*_{\alpha p}(\xi,s)\exp(-i\xi\tau^*_\alpha h)\phi_p(\xi,s), \tag{4.7}$$

where

$$\phi_p(\xi,s) = \sum_\alpha [L_{p2\alpha}(\xi,s)(E_\alpha(\xi,s) + M_{\alpha k}(\xi,s)\Psi_k(\xi,s))\exp(i\xi\tau_\alpha h) - \bar{L}_{p2\alpha}(\xi,s)\bar{E}_\alpha(\xi,s)\exp(i\xi\bar{\tau}_\alpha h)], \tag{4.8}$$

we find that the requirement that the displacements and the tractions are continuous at the interfaces is satisfied if

$$\sum_\alpha [A_{k\alpha}(\xi,s)(E_\alpha(\xi,s) + M_{\alpha p}(\xi,s)\Psi_p(\xi,s))\exp(i\xi\tau_\alpha h) - \overline{A_{k\alpha}}(\xi,s)\overline{E_\alpha}(\xi,s)\exp(i\xi\bar{\tau}_\alpha h)] = \sum_\alpha A^*_{k\alpha}(\xi,s)M^*_{\alpha p}(\xi,s)\phi_p(\xi,s), \tag{4.9}$$

where $[M^*_{\alpha p}]$ is the inverse of $[L^*_{i2\alpha}]$.

If we set

$$E_\alpha(\xi,s) = W_{\alpha p}\Psi_p(\xi,s), \tag{4.10}$$

where $W_{\alpha p}$ may be determined numerically using equations (4.7)–(4.8), then from equation (4.6) S_{j2} at $x_2 = 0$ are found to be

$$S_{j2}(x_1,0,s) = -\int_0^\infty \xi N_{jk}(\xi,s)\cos(\xi x_1)\int_0^a r_k(t,s)J_0(\xi t)dtd\xi - \frac{d}{dx_1}\int_0^{\min(x_1,a)} \frac{\Omega_{jk}r_k(t,s)dt}{(x_1^2-t^2)^{1/2}}, \tag{4.11}$$

where

$$N_{jk}(\xi,s) = T_{jp}(\xi,s)C_{pk}(\xi,s) + D_{jk}(\xi,s), \tag{4.12}$$

$$T_{jp}(\xi,s) = \sum_\alpha [L_{j2\alpha}(\xi,s)W_{\alpha p}(\xi,s) + \bar{L}_{j2\alpha}(\xi,s)\overline{W}_{\alpha p}(\xi,s)], \tag{4.13}$$

$$D_{jk}(\xi,s) = C_{pk}(\xi,s) - \Omega_{jk}, \tag{4.14}$$

$$\Omega_{jk} = \lim_{\xi/s\to\infty} C_{jk}(\xi,s). \tag{4.15}$$

From equations (2.6)–(2.7), it is clear that the limits in equation (4.15) are independent of s.

From equation (4.11), conditions (3.2) give rise to a set of Abel's integral equations which may be readily inverted to obtain (for $0 < x \le a$)

$$\Omega_{jk}r_k(x,s) + x\int_0^a K_{jk}(u,x,s)r_k(u,s)du = -\frac{2x}{\pi}Q(s)\int_0^x \frac{P_j(u)du}{(x^2-u^2)^{1/2}}, \tag{4.16}$$

where

$$K_{jk}(u,x,s) = \int_0^\infty \xi N_{jk}(\xi,s)J_0(\xi x)J_0(\xi u)d\xi. \tag{4.17}$$

Numerical values of the unknown functions $r_j(x, s)$ may now be determined by solving the integral equations in (4.16) numerically.

For the special case where $h \to \infty$, we must set the functions $E_\alpha(\xi, s)$ to zero if we require the displacements and the stresses to vanish at infinity. We may then take $T_{jp}(\xi, s) = 0$. Hence for this particular case the solution given in this section reduces to that for the corresponding problem of a crack in an infinite homogeneous anisotropic elastic material.

5 Stress Intensity Factors

At the crack tip $x_1 = a$, we define the stress intensity factors $k_I(t)$ and $k_{II}(t)$ as

$$k_I = \lim_{x_1\to a^+}(x_1-a)^{1/2}\sigma_{22}(x_1,0,t), \qquad k_{II} = \lim_{x_1\to a^-}(x_1-a)^{1/2}\sigma_{12}(x_1,0,t). \tag{5.1}$$

These factors are suitable for determining criteria for the stability of the crack if the anisotropic materials possess certain symmetries which are aligned with the coordinate axes (as in the numerical examples considered in Section 6).

Taking the Laplace transform with respect to t of equations (5.1), we obtain

$$K_I = \lim_{x_1\to a^+}(x_1-a)^{1/2}S_{22}(x_1,0,s), \qquad K_{II} = \lim_{x_1\to a^-}(x_1-a)^{1/2}S_{12}(x_1,0,s), \tag{5.2}$$

where

$$K_I(s) = \int_0^\infty k_I(t)\exp(-st)dt, \qquad K_{II}(s) = \int_0^\infty k_{II}(t)\exp(-st)dt. \tag{5.3}$$

From equations (4.11) and (5.3) and using integration by parts, we have

$$K_I(s) = \frac{\Omega_{2k}r_k(a,s)}{\sqrt{2a}} \quad \text{and} \quad K_{II}(s) = \frac{\Omega_{1k}r_k(a,s)}{\sqrt{2a}}. \tag{5.4}$$

Once $r_k(a, s)$ are determined from equation (4.16), the stress intensity factors $K_I(s)$ and $K_{II}(s)$ in the s-domain may be readily evaluated from equations (5.4). Through the use of a

numerical technique for inverting Laplace transforms, the stress intensity factors $k_I(t)$ and $k_{II}(t)$ may then be obtained.

6 Numerical Examples

In this section, special cases of the problem described in Section 3 are considered. More specifically, the materials occupying regions 1, 2, and 3 are taken to be transversely isotropic in nature. The crack is assumed to be opened by a uniform pressure suddenly applied on its faces. That is, we take the functions $P_j(x_1)$ and $q(t)$ in equation (3.1) to be, respectively,

$$P_j(x_1) = -\delta_{j2} p_0 \quad \text{and} \quad q(t) = H(t), \tag{6.1}$$

where p_0 is a given positive constant and $H(t)$ is the unit step Heaviside function.

The elastic behavior of a transversely isotropic material can be characterized using five elastic constants A, N, F, C, and L. Taking the transverse plane of the material to be perpendicular to the $0x_1$ axis, we find that the only nonzero c_{ijkl} which are of interest here are

$$A = c_{2222}, \quad C = c_{1111}, \quad F = c_{1122} = c_{2211},$$

$$N = c_{2233} = c_{3322}, \quad \frac{1}{2}(A-N) = c_{2323} = c_{3232} = c_{2332} = c_{3223},$$

$$L = c_{1212} = c_{2121} = c_{1221} = c_{2112} = c_{3113} = c_{1313} = c_{3131}. \tag{6.2}$$

The system (2.1) thus reduces to

$$\begin{aligned} C\frac{\partial^2 U_1}{\partial x_1^2} + L\frac{\partial^2 U_1}{\partial x_2^2} + (F+L)\frac{\partial^2 U_2}{\partial x_1 \partial x_2} - \rho s^2 U_1 = 0, \\ L\frac{\partial^2 U_2}{\partial x_1^2} + A\frac{\partial^2 U_2}{\partial x_2^2} + (F+L)\frac{\partial^2 U_1}{\partial x_1 \partial x_2} - \rho s^2 U_2 = 0, \end{aligned} \tag{6.3}$$

and the sextic equation (2.6) becomes

$$\left[\frac{1}{2}(A-N)\xi^2\tau^2 + L\xi^2 + \rho s^2\right][AL\xi^4\tau^4 + ((AC - F^2 - 2FL)\xi^4 + (A+L)\rho s^2\xi^2)\tau^2 + (C\xi^2 + \rho s^2)(L\xi^2 + \rho s^2)] = 0. \tag{6.4}$$

If τ_1 is given by

$$(\tau_1(\xi,s))^2 = \frac{-2(L\xi^2 + \rho s^2)}{(A-N)\xi^2}, \tag{6.5}$$

then τ_2^2 and τ_3^2 are the roots of the quartic factor in equation (6.4). We may then choose the functions $A_{k\alpha}$ to be

$$[A_{k\alpha}(\xi,s)] = \begin{bmatrix} 0 & \dfrac{-i\tau_2(F+L)\xi^2}{Z_2(\xi,s)} & \dfrac{-i\tau_3(F+L)\xi^2}{Z_3(\xi,s)} \\ 0 & i & i \\ 1 & 0 & 0 \end{bmatrix}, \tag{6.6}$$

and from equation (2.11) the corresponding functions $L_{j2\alpha}$ are found to be

$$[L_{j2\alpha}(\xi,s)] = \begin{bmatrix} 0 & \dfrac{iL[(C-F\tau_2^2)\xi^2 + \rho s^2]}{Z_2(\xi,s)} & \dfrac{iL[(C-F\tau_3^2)\xi^2 + \rho s^2]}{Z_3(\xi,s)} \\ 0 & i\tau_2\left[A - \dfrac{F(F+L)\xi^2}{Z_2(\xi,s)}\right] & i\tau_3\left[A - \dfrac{F(F+L)\xi^2}{Z_3(\xi,s)}\right] \\ \dfrac{1}{2}\tau_1(A-N) & 0 & 0 \end{bmatrix}, \tag{6.7}$$

where (for $i = 2, 3$)

$$Z_i(\xi,s) = (C + L\tau_i^2)\xi^2 + \rho s^2. \tag{6.8}$$

The other functions like $M_{\alpha p}$, B_{kp}, C_{pj}, W_{jk} and T_{jp} may be computed numerically for required values of ξ and s directly using equations (6.6)–(6.7).

From equations (6.1), (6.4), and (6.6)–(6.7), assuming that the roots τ_α are purely imaginary (as they will be in the numerical examples considered here), we find that the integral equation (4.16) may be reduced to

$$\Omega r(x,s) + x\int_0^1 K(u,x,s) r(u,s)\, du = \frac{xp_0}{s} \quad \text{for } 0 < x \le 1, \tag{6.9}$$

where $\Omega = \Omega_{22}$, $r(x,s) = r_2(ax,s)/a$ and

$$K(u,x,s) = \int_0^\infty \xi N_{22}(\xi/a, s) J_0(\xi x) J_0(\xi u)\, d\xi. \tag{6.10}$$

For this particular case, r_1 is zero.

We replace the Fredholm integral equation in (6.9) by

$$\Omega r^n(x,s) + x\int_0^1 K(u,x,s) r^{n-1}(u,s)\, du = \frac{xp_0}{s} \quad \text{for } n = 1, 2, \ldots, \tag{6.11}$$

together with

$$r^0(x,s) = \frac{xp_0}{s}, \tag{6.12}$$

to yield a sequence of approximations $\{r^n(x,s)\}$ to the unknown function $r(ax,s)$. However, it is found that this sequence converges smoothly and rapidly only when the parameter s is small. In general, the convergence is oscillatory (presumably about the true solution) and extremely slow. The range of s in which the convergence is rapid is sufficiently widened when we supercede $r^{n-1}(x,s)$ in equation (6.11) by $[r^{n-1}(x,s) + r^{n-2}(x,s)]/2$ fpr $n \ge 2$.

In solving equation (6.11) together with (6.12), we have to evaluate the infinite integrals

$$\int_0^\infty \xi T_{22}(\xi,s) C_{22}(\xi,s) J_0(\xi x) J_0(\xi u)\, d\xi$$

$$\text{and} \quad \int_0^\infty \xi D_{22}(\xi,s) J_0(\xi x) J_0(\xi u)\, d\xi. \tag{6.13}$$

From equations (4.8)–(4.9), the integrand of the first integral in (6.13) is expected to decay exponentially and rapidly to zero as ξ increases. Hence the numerical evaluation of this integral poses no considerable difficulty. The integrand of the other integral behaves as $O(s^2/\xi^2)$ for large ξ and hence the convergence of this integral is extremely slow. Using the results (Watson, 1922)

$$\int_0^\infty \frac{\xi}{\xi^2 + \eta^2} J_0(\xi u) J_0(\xi x)\, d\xi = I_0(\eta u) K_0(\eta x) \quad \text{for } 0 < u \le x, \tag{6.14}$$

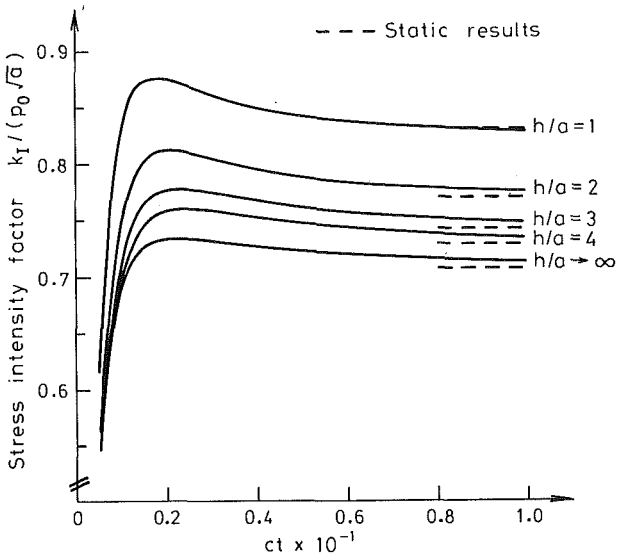


Fig. 2 Dynamic stress intensity factor for case (a)

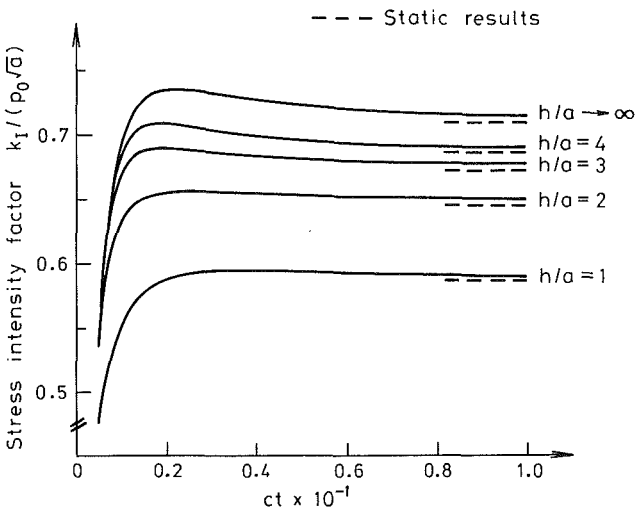


Fig. 3 Dynamic stress intensity factor for case (b)

the second integral in (6.13) may be written as

$$\int_0^\infty \xi D_{22}(\xi,s)J_0(\xi u)J_0(\xi x)d\xi$$

$$= \int_0^\infty \xi\left[D_{22}(\xi,s) - \frac{s^2\gamma_{22}}{\xi^2+\eta^2}\right]J_0(\xi u)J_0(\xi x)d\xi + s^2\gamma_{22}I_0(\eta u)K_0(\eta x) \quad \text{for} \quad 0<u\le x, \tag{6.15}$$

where $I_0(x)$ and $K_0(x)$ are the modified zero order Bessel functions of the first and second kind, respectively, γ_{22} is the limit of $\xi^2 D_{22}(\xi, s)/s^2$ as $\xi \to \infty$ and η is any real positive constant. The integrand of the integral on the right-hand side of (6.15) is $O(s^4/\xi^4)$ for large ξ. Due to the rapid diminution of this integrand, the second integral in (6.13) can now be evaluated easily by using (6.15).

To obtain some numerical results, we use the elastic constants for crystals of titanium and magnesium. For titanium, these constants (in gram per centimeter per microsecond square) are $A = 1.62$, $N = 0.92$, $F = 0.69$, $C = 1.81$, and $L = 0.47$, while, for magnesium, they are $A = 0.60$, $N = 0.26$, $F = 0.21$, $C = 0.61$, and $L = 0.16$. The densities ρ of titanium and magnesium in gram per centimeter cube are, respectively, 4.51 and 1.74. We now consider the following cases: (*a*) region 2 is occupied by the element titanium while regions 1 and 3 are occupied by magnesium, and (*b*) region 2 is occupied by the element magnesium while regions 1 and 3 are occupied by titanium. For both these cases, the unknown function $r(1, s)$ is obtained for required values of s by solving the Fredholm integral equation in equation (6.9) using equations (6.11)–(6.12). The stress intensity factor $K_I(s)/\sqrt{a} = \Omega r(1, s)/\sqrt{2}$ in the Laplace transform domain is then evaluated by using equations (5.4).

Table 1 Values of $k_I(t)/(p_0\sqrt{a})$ at $ct = 40.0$

h/a	Case (a) $k_I(t)/(p_0\sqrt{a})$	Case (a) $k_I^s/(p_0\sqrt{a})$	Case (b) $k_I(t)/(p_0\sqrt{a})$	Case (b) $k_I^s/(p_0\sqrt{a})$
1.000	0.822	0.831	0.589	0.585
2.000	0.769	0.770	0.644	0.643
3.000	0.742	0.741	0.671	0.670
4.000	0.728	0.728	0.685	0.684
∞	0.708	0.707	0.708	0.707

Table 2 Ratios of k_I^m/k_I^s

h/a	1.000	2.000	3.000	4.000	∞
k_I^m/k_I^s Case (a)	1.056	1.054	1.050	1.045	1.039
k_I^m/k_I^s Case (b)	1.015	1.019	1.028	1.035	1.038

A numerical technique by Stehfest (1970) for inverting Laplace transforms is used to obtain $k_I(t)/\sqrt{a}$. This technique is chosen because it can invert a fairly wide range of Laplace transforms with reasonable accuracy (Davies and Martin, 1979) and it is suitable for use here in that it does not require the evaluation of $K_I(s)/\sqrt{a}$ for complex values of s. According to this method, an approximation to the stress intensity factor $k_I(t)$ is given by

$$k_I(t) \approx \frac{\log 2}{t}\sum_{j=1}^{N} V_j K_I\left(\frac{j \log 2}{t}\right), \tag{6.16}$$

where N is a positive even integer and V_j is defined by

$$V_j = (-1)^{j+N/2} \sum_{p=\left[\frac{j+1}{2}\right]}^{\min(j,N/2)} \frac{p^{N/2}(2p)!}{(N/2-p)!p!(p-1)!(j-p)!(2p-j)!}. \tag{6.17}$$

The accuracy of most (if not all) existing numerical techniques for inverting Laplace transforms is highly susceptible to round-off errors. Hence in (6.16) a larger value of N does not necessarily imply greater accuracy in the approximation. In effect, if N is larger than a certain value (which depends on the accuracy with which we can evaluate $K_I(s)/\sqrt{a}$), the accuracy of the approximation starts to deteriorate. We test the approximation in (6.16) for selected values of h/a and large ct (where $c = \sqrt{L/\rho}$ using the constants of the material in region 2) using $N = 2$, 4, and 6. For large ct, $k_I(t)/\sqrt{a}$ is expected to be close to the value of the corresponding stress intensity factor $k_I^s/\sqrt{a}$ for the static problem considered in Clements (1979). We use this as a criterion for choosing the value of N which provides us with the best approximation for each of the values of h/a considered here. The best results of $k_I(t)/\sqrt{a}$ obtained at $ct = 40.0$ for cases (*a*) and (*b*) are summarized in Table 1. They are achieved using $N = 2$. It is obvious that the values of $k_I(t)/\sqrt{a}$ as given in Table 1 are reasonably close to the corresponding values of $k_I^s/\sqrt{a}$.

We plot $k_I(t)/(p_0)\sqrt{a}$ against ct for $h/a = 1.0$, 2.0, 3.0,

and 4.0 as well as for $h/a \to \infty$ in Figs. 2 and 3 for cases (*a*) and (*b*), respectively. Due to the discontinuity in the load function at $t = 0$, we find that it is difficult to obtain sensible results for $k_I(t)/\sqrt{a}$ when ct is extremely small. However, the evaluation of $k_I(t)/\sqrt{a}$ for the range of ct which is of practical interest presents no considerable difficulty. The graphs obtained exhibit various features which are similar to the ones given in Chen and Sih (1977). More specifically, $k_I(t)/\sqrt{a}$ rises rapidly, overshoots the static value $k_I^s/\sqrt{a}$, attains a peak value $k_I^m/\sqrt{a}$, and eventually settles down to $k_I^s/\sqrt{a}$. In Table 2, the effect of the parameter h/a on the ratio k_I^m/k_I^s is examined for both cases (*a*) and (*b*). For case (*a*), it appears that the ratio decreases as h/a increases, while for case (*b*) the reverse happens. At any rate, the percentage by which the peak value of $k_I(t)/\sqrt{a}$ is larger than the corresponding static value is not very wide in range for the particular examples considered here. As is clear from Table 2, for case (*a*) it ranges from roughly 4.0 percent to 6.0 percent, while for case (*b*) from about 1.5 percent to 4.0 percent. It is clear from the graphs that, as h/a increases, the results obtained tend to those for the corresponding problem involving an infinite homogeneous material. In fact, we find that the graphs for $h/a = 15.0$ and $h/a \to \infty$ are almost indistinguishable. As in Clements (1979), the crack is more stable in case (*b*) than in case (*a*), that is, the crack is less likely to extend when the layer is surrounded by a stronger material.

Acknowledgments

I wish to thank Dr. D. L. Clements for discussions related to this project. This project is supported by a research grant from The University of Adelaide.

References

Chen, E. P., 1978, "Sudden Appearance of a Crack in a Stretched Finite Strip," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 277–280.

Chen, E. P., and Sih, G. S., 1977, *Elastodynamic Crack Problems, Mechanics of Fractures*, Sih, G. S., ed., Noordhoff International Publishing, Leyden, The Netherlands, pp. 1–58.

Clements, D. L., 1979, "A Crack in an Anisotropic Layered Material," *Rozprawy Inzynierskie*, Vol. 27, pp. 171–180.

Davies, B., and Martin, B., 1979, "Numerical Inversion of the Laplace Transform: A Survey and Comparison of Methods," *Journal of Computational Physics*, Vol. 33, pp. 1–32.

Eshelby, J. D., Read, W. T., and Shockley, W., 1953, "Anisotropic Elasticity with Applications to Dislocation Theory," *Acta Metallurgica*, Vol. 1, pp. 251–259.

Kassir, M. K., and Bandyopadhyay, K. K., 1983, "Impact Response of a Cracked Orthotropic Medium," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 630–636.

Sih, G. C., Embley, G. T., and Ravera, R. S., 1972, "Impact Response of a Finite Crack in Plane Extension," *International Journal of Solids and Structures*, Vol. 8, pp. 977–993.

Stehfest, H., 1970, "Numerical Inversion of Laplace Transform," *Communications of ACM*, Vol. 13, pp. 47–49, p. 624.

Watson, G. N., 1922, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press.

A Theory of Particle-Reinforced Plasticity

G. P. Tandon*
Graduate Student.

G. J. Weng
Professor,
Mem. ASME

Department of Mechanics
and Materials Science,
Rutgers University,
New Brunswick, NJ 08903

A simple, albeit approximate, theory is developed to determine the elastoplastic behavior of particle-reinforced materials. The elastic, spherical particles are uniformly dispersed in the ductile, work-hardening matrix. The method proposed combines Mori-Tanaka's concept of average stress in elasticity and Hill's discovery of a decreasing constraint power of the matrix in polycrystal plasticity. Under a monotonic, proportional loading the latter was characterized, approximately, by the secant moduli of the matrix. The theory is established for both traction and displacement-prescribed boundary conditions, under which, the average stress and strain of the constituents and the effective secant moduli of the composite are explicitly given in terms of the secant moduli of the matrix and the volume fraction of particles. In particular, the yield stress and work-hardening modulus of the composite are shown to be inversely proportional to the deviatoric part of average stress concentration factors of the matrix, and therefore will increase (or decrease) with increasing hard (or soft) particle concentration. It is also found that, even if the matrix is plastically incompressible, the composite as a whole is not. Comparison between the theory and the experiment for a silica/epoxy system shows a reasonable agreement. The theory is also compared with a recently developed one by Arsenault and Taya; while both give the same initial yield stress for the composite, the work-hardening modulus predicted by their theory is found to be higher.

1 Introduction

The yield strength of a ductile material can be greatly enhanced by homogeneous dispersion of hard particles. These particles, such as oxides and carbides, often can deform only elastically. The ductile matrix may be metallic or polymeric, and, with the addition of particles, the overall property will change accordingly. The hard particles are important load-carrying members of the system and, by sustaining a higher stress, they can reduce the magnitude of stress experienced by the matrix. This leads to a lower plastic strain in the matrix phase—as compared to that in the absence of inclusions—under a given external load. Moreover, since the plastic strain of the composite is only contributed by the matrix alone, its magnitude is further reduced according to the volume fraction of this phase. These double reductions may significantly increase the yield stress and work-hardening modulus of the composite system.

While considerable work apparently has been done in the area of composite elasticity, the study of its elastoplastic behavior—due to the material nonlinearity—is still very limited. Among these limited studies attention has been given primarily to the fiber-reinforced composites (see, for instance, Hill, 1964; Dvorak and Bahei-El-Din, 1982; Accorsi and Nemat-Nasser, 1986). Here, instead, we shall focus our attention on the other class of materials—the particle-reinforced composite, an area which is also of vital technological importance.

The method to be introduced was prompted from a recent success in the study of particle-reinforced elasticity (Weng, 1984; Tandon and Weng, 1986). Based on Eshelby's (1957) solution of an ellipsoidal inclusion and Mori-Tanaka's (1973) concept of average stress in the matrix, the average stress and strain in the particles and at the interface were determined by means of Eshelby's equivalence principle. Originally proposed for particles with a given plastic strain, but nonetheless with an identical elastic property as the matrix—which deforms only elastically—the earlier Mori-Tanaka method only deals with the transformation problem. It was subsequently modified to an inhomogeneity problem (still within the context of elasticity) by Chow (1977), Taya and Chou (1981), and Taya and Mura (1981). By means of Eshelby's equivalent "stress-free" strain (transformation strain, or eigenstrain, Mura, 1982) the inclusion-matrix interaction at a finite concentration was accounted for by means of the average-stress concept under a prescribed traction. The analysis was extended to a prescribed displacement by Weng (1984) and, in addition to the stress-free strain, the "strain-free" stress (Hill, 1965; or polarization stress, Kröner, 1958) was also included for both types of boundary conditions. With spherical inclusions, the effective bulk and shear moduli of the composite were shown to coincide with Hashin and Shtrikman's (1963) lower bounds if the

*Current address: Adtech System Research, Fairborne, OH 45324.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 27, 1987; final revision October 26, 1987.

matrix is the softer phase, otherwise the values coincide with their upper bounds. When both phases have identical shear rigidity the effective bulk modulus is identical to Hill's (1963) exact solution. The predicted Young's modulus was shown to agree favorably with the experimental data on tungsten carbide/cobalt, glass/polyester, and some two and three-phase porous media (see Weng, 1984, and Norris, 1985, for other implications).

In dealing with plasticity, the problem is simpler if the particles are the phase which deform plastically and the matrix remains elastic. As in this case it can be solved straightforwardly by modifying Mori-Tanaka's original formulation to an equivalent transformation problem. A more difficult, and also technologically more important, problem is the one declared in the opening paragraph, in which the inclusions remain elastic and the matrix deforms plastically. Then, the constraint power of the matrix will continue to weaken in the course of plastic deformation. Such a weakening effect was first discovered by Hill (1965) in his study of polycrystal plasticity. The resulting self-consistent relation differs from the earlier, simpler version of Kröner (1961) and Budiansky and Wu (1962), in which the constraint power of the matrix remains elastic. Hutchinson's (1970) calculations indicated that the predicted stress-strain curve of the polycrystal based on the K-B-W model is higher than that based on Hill's model. Hill's formulation appears to be highly rigorous, and can be used for any incremental calculations. Under a monotonic, proportional loading, an approximate, but numerically simpler, method was later suggested by Berveiller and Zaoui (1979); this method employs the "secant" moduli of the polycrystal to characterize the weakening constraint power of the matrix. Weng's (1982) calculations have demonstrated that this method actually gives a very close approximation to Hill's model.

Encouraged by its accuracy and certainly attracted by its great simplicity, we shall also introduce the concept of *secant moduli* here to the Mori-Tanaka method, so that it will become suitable for the calculation of particle-reinforced plasticity. It should be recognized that the problem of particle-reinforced plasticity involves a clear matrix phase, and, therefore, differs fundamentally from that where both (or multi) phases exist on equal geometrical footing (such as some dual phase steels studied by Weng, 1985, or polycrystal plasticity). While the latter problem is probably best analyzed by the self-consistent scheme—with the inclusion embedded in the homogeneous, effective medium (e.g., the K-B-W method or the Hill-Hutchinson method)—such an analysis for the hard (or soft) particle problem may lead to a significant overestimate (or underestimate) on the yield strength of the composite.

2 The Properties of the Constituents

Since the theory is intended only for the monotonic, proportional loading and the secant modulus of the matrix will be used in the subsequent development, the deformation theory (instead of the incremental one) will be chosen to describe the elastoplastic properties of the matrix.

In the composite system the spherical particles are taken to be phase 1, and the ductile matrix as phase 0. The bulk and shear moduli of the rth phase are denoted by κ_r and μ_r, respectively, and its volume fraction by c_r. The uniaxial stress and plastic strain of the matrix is assumed to be representable by the modified Ludwik equation

$$\sigma = \sigma_y + h \cdot (\epsilon^p)^n, \tag{2.1}$$

where σ is the flow stress at the plastic strain ϵ^p, and σ_y, h and n are the initial yield stress, the strength coefficient, and the work-hardening exponent, in turn. These three parameters can be readily determined from a simple tensile test.

The total strain of the matrix is the sum of the elastic and plastic components, so that, under a monotonic loading, its "secant" Young's modulus decreases with increasing plastic deformation. That is

$$E_0^s = \frac{\sigma}{\epsilon^e + \epsilon^p} = \frac{1}{\dfrac{1}{E_0} + \dfrac{\epsilon^p}{\sigma_y + h \cdot (\epsilon^p)^n}}, \tag{2.2}$$

where ϵ^e is the elastic strain and E_0 the ordinary Young's modulus.

With respect to the undeformed state the "secant" bulk and shear moduli of the matrix are taken to satisfy the usual isotropic relations, as

$$\kappa_0^s = \frac{E_0^s}{3(1-2\nu_0^s)} = \kappa_0, \quad \mu_0^s = \frac{E_0^s}{2(1+\nu_0^s)}, \tag{2.3}$$

in terms of the "secant" Poisson ratio ν_0^s. The secant bulk modulus is equal to its elastic counterpart due to the plastic incompressibility, which further provides

$$\nu_0^s = \frac{1}{2} - \left(\frac{1}{2} - \nu_0\right)\frac{E_0^s}{E_0}, \tag{2.4}$$

in connection with the ordinary Poisson ratio ν_0. Under a monotonic, proportional loading, the plastic state therefore can be described by a single secant modulus, say E_0^s, and the other two elastic constants.

In a composite system the stress and strain state of the matrix is usually triaxial. Equation (2.1) can be generalized to such a state by means of von Mises' effective stress σ^* and strain ϵ^{p*} as

$$\sigma^* = \sigma_y + h \cdot (\epsilon^{p*})^n, \tag{2.5}$$

where as usual

$$\sigma^* = \left(\frac{3}{2}\sigma'_{ij}\sigma'_{ij}\right)^{1/2}, \quad \epsilon^{*p} = \left(\frac{2}{3}\epsilon^p_{ij}\epsilon^p_{ij}\right)^{1/2}, \tag{2.6}$$

σ'_{ij} being the deviatoric stress. In addition, the flow rule

$$\epsilon^p_{ij} = \frac{3}{2} \cdot \frac{\epsilon^{p*}}{\sigma^*} \cdot \sigma'_{ij}, \tag{2.7}$$

is also adopted. Now in terms of the generalized stress and strain, the secant Young's modulus can be recast as

$$E_0^s = \frac{1}{\dfrac{1}{E_0} + \dfrac{\epsilon^{p*}}{\sigma_y + h \cdot (\epsilon^{p*})^n}}, \tag{2.8}$$

for the multiaxial state.

In the following analysis the particles are taken to be uniformly dispersed, perfectly bonded with the matrix, which can continue to work harden. The average stress of the matrix—instead of the in-situ local stress—will be used to determine the yield condition. While it is appreciated that the actual yielding process will start from some local area, its contribution to the plastic strain of the composite, in view of its negligible volume fraction, is practically insignificant. For the determination of yield point of the composite, the average, or mean-field theory, should be a reasonable approximation.

Both indicial and symbolic notations will be used in the text. In the latter case a second-rank tensor will be denoted by the bold-face, lower case Greek letter and a fourth-rank tensor by the ordinary capital letter. As will become evident the latter tensors are all isotropic; their hydrostatic and deviatoric components will be expressed as, for example,

$$L_0^s = (3\kappa_0^s, 2\mu_0^s), \tag{2.9}$$

in the case of secant moduli tensor of the matrix. Likewise, these components can be written for the stress and strain tensors as

$$\sigma=(\sigma_{kk},\ \sigma'_{ij}),\ \epsilon=(\epsilon_{kk},\ \epsilon'_{ij}). \tag{2.10}$$

3 Foundation of the Theory

The present theory—which evolved from the writers' study on polycrystal plasticity and the Mori-Tanaka (1973) method—is an approximate, but rather simple one. To place it in proper perspective with regard to other mean-field theories, it is fitting to give a brief account on its origin.

The *auxiliary* problem of inclusion-matrix interaction during plastic deformation has been addressed by two major approaches: following the inquiry of Kröner (1961) and Budiansky-Wu (1962) on the one hand, and Hill (1965) on the other. The K-B-W approach—with an earlier, similar version by Hershey (1954)—identifies the difference between the plastic strains of the inclusion and the matrix as Eshelby's (1957) "stress-free" transformation strain, and takes the matrix response during Eshelby's releasing process of the restored inclusion to be purely elastic everywhere. When the inclusion and matrix have the same isotropic, elastic moduli, the local stress in the inclusion under an external $\bar{\sigma}$ can be written as

$$\sigma^{(1)}=\bar{\sigma}-2\mu_0(1-\beta_0)(\epsilon^{p(1)}-\bar{\epsilon}^p), \tag{3.1}$$

where μ_0 is the elastic shear modulus of the matrix (also of the inclusion here), $\beta_0=(2/15)\cdot(4-5\nu_0)/(1-\nu_0)$, and $\epsilon^{p(1)}$ and $\bar{\epsilon}^p$ are the average plastic strain of the inclusion and of the matrix (or the inclusion/matrix composite, as the matrix is infinitely extended). The constraint power of this model is characterized by $2\mu_0(1-\beta_0)$ and, being dependent solely on the elastic properties of the matrix, it remains constant in the course of deformation.

When the inclusion and matrix have different elastic moduli, say L_1 and L_0, respectively, Tanaka and Mori (1970) appear to be the first to also adopt this line of thought (perhaps independent of the K-B-W works) to address the problem of composite plasticity. With very dilute concentration of particles, Eshelby's equivalence principle then takes the form

$$\begin{aligned}\sigma^{(1)}=\bar{\sigma}+\sigma^{pt} &= L_1[\epsilon^0+\epsilon^{pt}-(\epsilon^{p(1)}-\epsilon^{p(0)})]\\ &= L_0[\epsilon^0+\epsilon^{pt}-(\epsilon^{p(1)}-\epsilon^{p(0)})-\epsilon^*],\end{aligned} \tag{3.2}$$

where $\bar{\sigma}=L_0\epsilon^0$, σ^{pt} and ϵ^{pt} represents the "perturbed" parts—or the differences—of average stress and strain, respectively, between the inclusion and the matrix, and ϵ^* is Eshelby's equivalent transformation strain. The perturbed strain is further connected to the total transformation strain as

$$\epsilon^{pt}=S_0[(\epsilon^{p(1)}-\epsilon^{p(0)})+\epsilon^*], \tag{3.3}$$

where, for a spherical inclusion, Eshelby's S_0 tensor, has the hydrostatic and deviatoric components

$$S_0=(\alpha_0,\ \beta_0), \tag{3.4}$$

with $\alpha_0=(1/3)\cdot(1+\nu_0)/(1-\nu_0)$ in an isotropic matrix. The K-B-W relation is readily recovered from equations (3.2) to (3.4) by the conditions of $L_0=L_1$ and plastic incompressibility. This method was subsequently applied by Tanaka et al. (1973) to study the effect of aspect ratio of inclusions on the plastic behavior of composites.

At finite concentration of particles, say with a volume fraction c_1, the average stress and strain in the matrix can no longer be represented by $\bar{\sigma}$ and ϵ^0, and some additional values $\tilde{\sigma}$ and $\tilde{\epsilon}$, ought to be included (Mori and Tanaka, 1973). This leads to an average stress and strain of $\bar{\sigma}+\tilde{\sigma}$ and $\epsilon^0+\tilde{\epsilon}$, respectively, in the matrix. Consequently, Eshelby's equivalence principle is written as

$$\begin{aligned}\sigma^{(1)}=\bar{\sigma}+\tilde{\sigma}+\sigma^{pt} &= L_1[\epsilon^0+\tilde{\epsilon}+\epsilon^{pt}-(\epsilon^{p(1)}-\epsilon^{p(0)})]\\ &= L_0[\epsilon^0+\tilde{\epsilon}+\epsilon^{pt}-(\epsilon^{p(1)}-\epsilon^{p(0)})-\epsilon^*],\end{aligned} \tag{3.5}$$

with the further constraint

$$\tilde{\sigma}=c_1\sigma^{(1)}+(1-c_1)(\bar{\sigma}+\tilde{\sigma}),\ \text{or}\ \tilde{\sigma}=-c_1\sigma^{pt}, \tag{3.6}$$

to satisfy the boundary condition. This approach has been taken by Wakashima et al. (1979), Pedersen (1983), and most recently by Arsenault and Taya (1987), to explore several features of two-phase plasticity.

Although having been improved over the years, this line of approach basically treats the difference between the plastic strains of the inclusion and the matrix as Eshelby's transformation strain, and assumes the constraint power of the matrix to remain elastic. The possible additional plastic flow in the inclusion and the matrix, especially near the interface, during Eshelby's releasing process of the restored inclusion apparently has not been accounted for.

The second approach was originated by Hill (1965). The auxiliary problem then involved a single inclusion embedded in a homogeneous matrix, both undergoing an incremental flow. The problem was treated within the context of anisotropic elasticity. The incremental stress and strain of the inclusion, denoted, respectively, by $d\sigma^{(1)}$ and $d\epsilon^{(1)}$, were found to be connected to those of the surrounding matrix, denoted by $d\bar{\sigma}$ and $d\bar{\epsilon}$, as

$$d\sigma^{(1)}=d\bar{\sigma}-L_0^{*t}(d\epsilon^{(1)}-d\bar{\epsilon}), \tag{3.7}$$

where L_0^{*t} is the (tangent) constraint tensor of the matrix, related to the tangent moduli tensor of the matrix, L_0^t, through

$$L_0^{*t}S_0^t=L_0^t(I-S_0^t), \tag{3.8}$$

where S_0^t is the corresponding Eshelby's transformation tensor in the said context of anisotropic elasticity, and I the usual fourth rank identity tensor (L_0^{*t}, L_0^t, and S_0^t correspond to Hill's L^*, L, and S, in turn). Since L_0^t continues to decrease in the course of plastic flow, the constraint power of the matrix—represented by L_0^{*t}—is seen to weaken under a monotonic, proportional loading.

While derived with his own perspective, Hill's relation can also be established by means of Eshelby's (1957) equivalence principle,

$$\begin{aligned}d\sigma^{(1)}=d\bar{\sigma}+d\sigma^{pt}&=L_1^t(d\bar{\epsilon}+d\epsilon^{pt})\\ &=L_0^t(d\bar{\epsilon}+d\epsilon^{pt}-d\epsilon^*),\end{aligned} \tag{3.9}$$

where L_1^t is the tangent moduli tensor of the inclusion, and $d\epsilon^*$ Eshelby's equivalent transformation strain. Now since

$$d\bar{\sigma}=L_0^t d\bar{\epsilon},\ \text{and}\ d\epsilon^{pt}=S_0^t d\epsilon^*, \tag{3.10}$$

one has

$$d\sigma^{pt}=-L_0^t(I-S_0^t)d\epsilon^*. \tag{3.11}$$

On the other hand, from equation (3.7), $d\sigma^{pt}=-L_0^{*t}d\epsilon^{pt}=-L_0^{*t}S_0^t d\epsilon^*$, and, comparing it to equation (3.11), one arrives at Hill's connection (3.8). The constraint power of the matrix is, therefore, implicitly reflected by the coefficient L_0^t in the equivalence principle (3.9).

Tensors L_0^{*t}, L_0^t and S_0^t are generally anisotropic; however, under a monotonic, proportional loading they would reduce to the isotropic quantities in reference to the undeformed state (Berveiller and Zaoui, 1979). In terms of the total stress and strain, this modified, approximate version can be cast into

$$\sigma^{(1)}=\bar{\sigma}-L_0^{*s}(\epsilon^{(1)}-\bar{\epsilon}), \tag{3.12}$$

$$L_0^{*s}S_0^s=L_0^s(I-S_0^s), \tag{3.13}$$

in parallel to equations (3.7) and (3.8), where the superscript s refers to the corresponding "secant" moduli, such as the secant moduli tensor of the matrix L_0^s. Tensors L_0^{*s}, L_0^s, and S_0^s are now all isotropic, having the hydrostatic and deviatoric components

$$L_0^s=(3\kappa_0,\ 2\mu_0^s),\quad S_0^s=(\alpha_0^s,\ \beta_0^s), \tag{3.14}$$

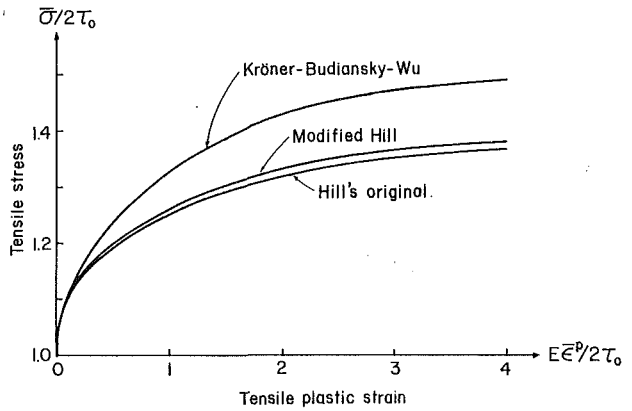


Fig. 1 The stress-strain curves of a polycrystal calculated from the Kröner-Budiansky-Wu model, Hill's original incremental model (both by Hutchinson, 1970), and the Berveiller-Zaoui modification of Hill's model (marked here as "Modified Hill," by Weng, 1982)

with

$$\alpha_0^s = \frac{1}{3}\cdot\frac{1+\nu_0^s}{1-\nu_0^s},\ \beta_0^s = \frac{2}{15}\cdot\frac{4-5\nu_0^s}{1-\nu_0^s}. \tag{3.15}$$

Like Hill's incremental relations, the B-Z approximations (3.12) and (3.13) can also be derived by the approximate equivalence

$$\sigma^{(1)} = \bar{\sigma} + \sigma^{pt} = L_1^s(\bar{\epsilon} + \epsilon^{pt})$$

$$= L_0^s(\bar{\epsilon} + \epsilon^{pt} - \epsilon^*), \tag{3.16}$$

with

$$\tilde{\sigma} = L_0^s\tilde{\epsilon}, \text{ and } \epsilon^{pt} = S_0^s\epsilon^*, \tag{3.17}$$

as in equations (3.10). The coefficient L_0^s in equation (3.16) decreases with increasing plastic flow, thereby reflecting the decreasing constraint power of the matrix.

The accuracy of the Berveiller-Zaoui (1979) approximation in the calculation of polycrystal plasticity is depicted in Fig. 1 (taken from Weng, 1982); the resulting stress-strain response, marked by "Modified Hill," is seen to be softer than that of the K-B-W model and quite close to Hill's original one (taken from Hutchinson, 1970).

The B-Z approximation—or equations (3.16) and (3.17)—applies to the single inclusion-matrix interaction. In what follows we propose to introduce Mori-Tanaka's concept of average stress in the matrix into such an approximation to account for the inclusion-matrix interaction at finite concentration.

4 Average Stress of the Constituents, and Total Strain and Effective Secant Moduli of the Composite Under a Prescribed Traction

Now consider a representative volume of the composite, which is statistically homogeneous and macroscopically isotropic. To facilitate the analysis we introduce an identically shaped comparison material, with the property of the matrix. Let the composite and the comparison material both subject to the same boundary traction which would give rise to a uniform stress $\bar{\sigma}$. As shown in Fig. 2(*a*) the strain in the comparison material ϵ^0, is given by

$$\bar{\sigma} = L_0^{(s)}\epsilon^0, \tag{4.1}$$

where $L_0^{(s)}$ (noting the presence of parenthesis in the superscript *s*) is the secant moduli tensor of the comparison material at the applied stress state.

Due to the presence of particles, the average strain of the matrix in the composite system differs from ϵ^0, say by an

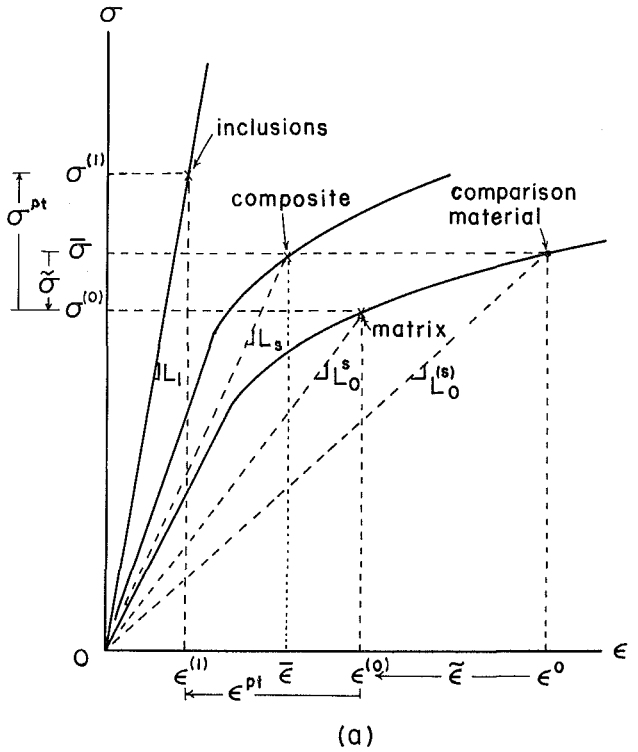


Fig. 2 Schematic representation of the average stress and strain state of the constituents under (*a*) a prescribed traction, and (*b*) a prescribed displacement

amount $\tilde{\epsilon}$. As marked in Fig. 2(*a*), the average strain of the matrix is

$$\epsilon^{(0)} = \epsilon^0 + \tilde{\epsilon}. \tag{4.2}$$

Its average stress also differs from $\bar{\sigma}$, say by $\tilde{\sigma}$, such that its stress-strain relation is now connected by

$$\sigma^{(0)} = \bar{\sigma} + \tilde{\sigma} = L_0^s(\epsilon^0 + \tilde{\epsilon}), \tag{4.3}$$

where L_0^s (noting the absence of parenthesis in the superscript s) is the secant moduli tensor of the matrix in the composite system at the current stress state. Tensor L_0^s generally differs from $L_0^{(s)}$ of the comparison material; such a difference disappears in the case of linear elasticity or of single inclusion-matrix interaction (as the B-Z formulation).

The average stress and strain of inclusions further differ from those of the surrounding matrix, say with the additional perturbations σ^{pt} and ϵ^{pt}, respectively. Now recognizing that the constraint power of the matrix in the composite system is to be represented by its secant moduli L_0^s (not $L_0^{(s)}$ of the comparison material), the approximate equivalence principle (3.16) may be combined with Mori-Tanaka's concept of average stress to take the form

$$\sigma^{(1)} = \bar{\sigma} + \tilde{\sigma} + \sigma^{pt} = L_1(\epsilon^0 + \tilde{\epsilon} + \epsilon^{pt})$$

$$= L_0^s(\epsilon^0 + \tilde{\epsilon} + \epsilon^{pt} - \epsilon^*), \tag{4.4}$$

for the elastically deforming particles. The constraint power of the matrix (now reflected by L_0^s) is seen to weaken upon continuous loading.

Since the surrounding matrix exists at the state L_0^s (not $L_0^{(s)}$), the connection between ϵ^{pt} and $\epsilon^*$is still given by the second of equations (3.17) (not $S_0^{(s)}$ associated with $L_0^{(s)}$). Also, in balance with the external stress, one still has (3.6). Namely,

$$\epsilon^{pt} = S_0^s \epsilon^*, \text{ and } \tilde{\sigma} = -c_1 \sigma^{pt}. \tag{4.5}$$

Moreover, from equations (4.1) and (4.3), $\tilde{\sigma} = (L_0^s - L_0^{(s)})\epsilon^0 + L_0^s \tilde{\epsilon}$, and from equations (4.3) and (4.4) $\sigma^{pt} = L_0^s(\epsilon^{pt} - \epsilon^*) = L_0^s(S_0^s - I)\epsilon^*$. Substituting these into the second of equations (4.5) provides

$$\tilde{\epsilon} = -c_1(S_0^s - I)\epsilon^* - (I - L_0^{s\,-1} L_0^{(s)})\epsilon^0, \text{ or}$$

$$= -c_1 \epsilon^{pt} + c_1 \epsilon^* - \epsilon^0 + L_0^{s\,-1} L_0^{(s)} \epsilon^0, \tag{4.6}$$

where $L_0^{s\,-1}$ is the inverse of L_0^s.

It is now more convenient to decompose the stress and strain into their hydrostatic and deviatoric parts. From the first of equations (4.6) one writes

$$\tilde{\epsilon}_{kk} = -c_1(\alpha_0^s - 1)\epsilon_{kk}^*, \tag{4.7}$$

$$\tilde{\epsilon}'_{ij} = -c_1(\beta_0^s - 1)\epsilon_{ij}^{*\prime} - \left(1 - \frac{\mu_0^{(s)}}{\mu_0^s}\right)\epsilon_{ij}^{0\prime}, \tag{4.8}$$

noting that $\kappa_0^{(s)} = \kappa_0^s = \kappa_0$ for the matrix material. Then from the last of equation (4.4) and also with the help of the first of equations (4.5), we have the equivalent transformation strain.

$$\epsilon_{kk}^* = -\frac{(\kappa_1 - \kappa_0)}{[c_1 + (1-c_1)\alpha_0^s](\kappa_1 - \kappa_0) + \kappa_0}\epsilon_{kk}^0, \tag{4.9}$$

$$\epsilon_{ij}^{*\prime} = -\frac{(\mu_1 - \mu_0^s)}{[c_1 + (1-c_1)\beta_0^s](\mu_1 - \mu_0^s) + \mu_0^s} \cdot \frac{\mu_0^{(s)}}{\mu_0^s}\epsilon_{ij}^{0\prime}, \tag{4.10}$$

where the strain of the comparison material is further related to the applied stress by $3\kappa_0 \epsilon_{kk}^0 = \bar{\sigma}_{kk}$ and $2\mu_0^{(s)} \epsilon_{ij}^{0\prime} = \bar{\sigma}'_{ij}$.

Central to the determination of yield criterion is the average stress in the matrix. Now that $\tilde{\epsilon}$ and ϵ^* are both established it is straightforward to find

$$\sigma_{kk}^{(0)} = a_0 \bar{\sigma}_{kk},\ a_0 = \frac{\alpha_0^s(\kappa_1 - \kappa_0) + \kappa_0}{[c_1 + (1-c_1)\alpha_0^s](\kappa_1 - \kappa_0) + \kappa_0}, \tag{4.11}$$

$$\sigma_{ij}^{\prime(0)} = b_0 \bar{\sigma}'_{ij},\ b_0 = \frac{\beta_0^s(\mu_1 - \mu_0^s) + \mu_0^s}{[c_1 + (1-c_1)\beta_0^s](\mu_1 - \mu_0^s) + \mu_0^s}, \tag{4.12}$$

from equation (4.3), where a_0 and b_0 are, respectively, the hydrostatic and deviatoric stress concentration factors of the matrix. Likewise for the inclusions one has

$$\sigma_{kk}^{(1)} = a_1 \bar{\sigma}_{kk},\ a_1 = \frac{\kappa_1}{[c_1 + (1-c_1)\alpha_0^s](\kappa_1 - \kappa_0) + \kappa_0}, \tag{4.13}$$

$$\sigma_{ij}^{\prime(1)} = b_1 \bar{\sigma}'_{ij},\ b_1 = \frac{\mu_1}{[c_1 + (1-c_1)\beta_0^s](\mu_1 - \mu_0^s) + \mu_0^s}, \tag{4.14}$$

from equation (4.4). The stress concentration factors satisfy $c_1 a_1 + (1-c_1)a_0 = 1$, $c_1 b_1 + (1-c_1)b_0 = 1$, as required. As will become evident the factor b_0 plays an especially vital role in the determination of composite plasticity.

The total strain of the composite is given by the weighted average of those of its constituents; this leads to

$$\bar{\epsilon} = \epsilon^0 + \tilde{\epsilon} + c_1 \epsilon^{pt} = L_0^{s\,-1} L_0^{(s)} \epsilon^0 + c_1 \epsilon^*, \tag{4.15}$$

where the last was arrived with the help of the second of equations (4.6). In terms of the hydrostatic and deviatoric components, it becomes

$$\bar{\epsilon}_{kk} = \epsilon_{kk}^0 + c_1 \epsilon_{kk}^* = \frac{\bar{\sigma}_{kk}}{3\kappa_0} + c_1 \epsilon_{kk}^*, \tag{4.16}$$

$$\bar{\epsilon}'_{ij} = \frac{\mu_0^{(s)}}{\mu_0^s}\epsilon_{ij}^{0\prime} + c_1 \epsilon_{ij}^{*\prime} = \frac{\bar{\sigma}'_{ij}}{2\mu_0^s} + c_1 \epsilon_{ij}^{*\prime}. \tag{4.17}$$

With ϵ^* given by equations (4.9) and (4.10), the effective "secant" bulk and shear moduli of the composite, defined by $\bar{\sigma}_{kk} = 3\kappa_s \bar{\epsilon}_{kk}$ and $\bar{\sigma}'_{ij} = 2\mu_s \bar{\epsilon}'_{ij}$, readily follow as

$$\frac{\kappa_s}{\kappa_0} = 1 + \frac{c_1(\kappa_1 - \kappa_0)}{(1-c_1)\alpha_0^s(\kappa_1 - \kappa_0) + \kappa_0}, \tag{4.18}$$

$$\frac{\mu_s}{\mu_0^s} = 1 + \frac{c_1(\mu_1 - \mu_0^s)}{(1-c_1)\beta_0^s(\mu_1 - \mu_0^s) + \mu_0^s}, \tag{4.19}$$

and then so do the "secant" Young's modulus and "secant" Poisson ratio from the isotropic relations

$$E_s = \frac{9\kappa_s \mu_s}{3\kappa_s + \mu_s},\ \nu_s = \frac{3\kappa_s - 2\mu_s}{2(3\kappa_s + \mu_s)}. \tag{4.20}$$

Since α_0^s is not constant, the effective secant bulk modulus κ_s also changes with increasing plastic deformation; thus while the matrix phase is plastically incompressible the composite is not! Within the elastic limit these results reduce to those derived by Weng (1984) (with the super and subscript s deleted in all equations).

Therefore, at a given applied stress $\bar{\sigma}$ the total strain of the composite $\bar{\epsilon}$ and its secant moduli, κ_s and μ_s, can be found from equations (4.16) and (4.17), and (4.18) and (4.19), respectively, and so can the average stress and strain state of its constituent phases. However, as μ_0^s of the matrix is not known at the outset, an iterative procedure is usually required for the solution. Such a process is also necessary when the boundary condition is a prescribed displacement, so we shall defer its discussion until after the next section.

5 Average Strain of the Constituents, and the Stress and Effective Secant Moduli of the Composite Under a Prescribed Displacement

Alternatively the composite and the comparison material are now subjected to a prescribed displacement corresponding to a uniform strain $\bar{\epsilon}$. The stress required for the comparison material, referrring to Fig. 2(b), is

$$\sigma^0 = L_0^{(s)} \bar{\epsilon}, \tag{5.1}$$

noting again the presence of parentheses in the superscript s for the moduli tensor of the comparison material. The average stress and strain of the matrix in the composite system will differ from these by, say $\tilde{\sigma}$ and $\tilde{\epsilon}$, respectively. As marked in the same figure these states are connected to each other by this secant moduli L_0^s

$$\sigma^{(0)} = \sigma^0 + \tilde{\sigma} = L_0^s(\bar{\epsilon} + \tilde{\epsilon}), \tag{5.2}$$

noting also the absence of parentheses in the superscript s for the moduli tensor of the matrix of the composite system.

The average stress and strain of the inclusions further differ from these by σ^{pt} and ϵ^{pt}, respectively. As in equation (4.4), the approximate equivalence can be written as

$$\sigma^{(1)} = \sigma^0 + \tilde{\sigma} + \sigma^{pt} = L_1(\bar{\epsilon} + \tilde{\epsilon} + \epsilon^{pt})$$

$$= L_0^s(\bar{\epsilon} + \tilde{\epsilon} + \epsilon^{pt} - \epsilon^*), \tag{5.3}$$

where ϵ^* is the equivalent transformation strain.

In balance the weighted average of $\epsilon^{(r)}$ must be equal to $\bar{\epsilon}$; this gives

$$\tilde{\epsilon} = -c_1\epsilon^{pt}. \tag{5.4}$$

Substituting equation (5.4) and the first of equations (4.5) into the last equality of (5.3), we have

$$\epsilon_{kk}^* = -\frac{\kappa_1 - \kappa_0}{(1-c_1)\alpha_0^s(\kappa_1 - \kappa_0) + \kappa_0}\bar{\epsilon}_{kk}, \tag{5.5}$$

$$\epsilon_{ij}^{*\prime} = -\frac{\mu_1 - \mu_0^s}{(1-c_1)\beta_0^s(\mu_1 - \mu_0^s) + \mu_0^s}\bar{\epsilon}_{ij}'. \tag{5.6}$$

It is then straightforward to show that the average stress and strain of the constituents are given by

$$\sigma_{kk}^{(0)} = 3\kappa_0\epsilon_{kk}^{(0)},\ \epsilon_{kk}^{(0)} = p_0\bar{\epsilon}_{kk},\ p_0 = \frac{\alpha_0^s(\kappa_1 - \kappa_0) + \kappa_0}{(1-c_1)\alpha_0^s(\kappa_1 - \kappa_0) + \kappa_0}, \tag{5.7}$$

$$\sigma_{ij}'^{(0)} = 2\mu_0^s\epsilon_{ij}'^{(0)},\ \epsilon_{ij}'^{(0)} = q_0\bar{\epsilon}_{ij}',\ q_0 = \frac{\beta_0^s(\mu_1 - \mu_0^s) + \mu_0^s}{(1-c_1)\beta_0^s(\mu_1 - \mu_0^s) + \mu_0^s}, \tag{5.8}$$

$$\sigma_{kk}^{(1)} = 3\kappa_1\epsilon_{kk}^{(1)},\ \epsilon_{kk}^{(1)} = p_1\bar{\epsilon}_{kk},\ p_1 = \frac{\kappa_0}{(1-c_1)\alpha_0^s(\kappa_1 - \kappa_0) + \kappa_0}, \tag{5.9}$$

$$\sigma_{ij}'^{(1)} = 2\mu_1\epsilon_{ij}'^{(1)},\ \epsilon_{ij}'^{(1)} = q_1\bar{\epsilon}_{ij}',\ q_1 = \frac{\mu_0^s}{(1-c_1)\beta_0^s(\mu_1 - \mu_0^s) + \mu_0^s}, \tag{5.10}$$

where p_r and q_r are the hydrostatic and deviatoric strain concentration factors, respectively, of the rth phase.

The average stress of the composite is given by the weighted average of $\sigma^{(r)}$, and this provides $\bar{\sigma} = \sigma^{(0)} + c_1\sigma^{pt} = \sigma^{(0)} + c_1L_0^s(S_0^s - I)\epsilon^*$. Recasting $\sigma^{(0)}$ and ϵ^* in terms of $\bar{\epsilon}$ and identifying the coefficients as the effective secant moduli of the composite, one arrives at

$$\bar{\sigma}_{kk} = 3\kappa_s\bar{\epsilon}_{kk},\ \bar{\sigma}_{ij}' = 2\mu_s\bar{\epsilon}_{ij}', \tag{5.11}$$

$$\frac{\kappa_s}{\kappa_0} = 1 + \frac{c_1(\kappa_1 - \kappa_0)}{(1-c_1)\alpha_0^s(\kappa_1 - \kappa_0) + \kappa_0}, \tag{5.12}$$

$$\frac{\mu_s}{\mu_0^s} = 1 + \frac{c_1(\mu_1 - \mu_0^s)}{(1-c_1)\beta_0^s(\mu_1 - \mu_0^s) + \mu_0^s}. \tag{5.13}$$

Comparison between equations (5.12) and (5.13) and (4.18) and (4.19) indicates that the analyses under these two different boundary conditions lead to the identical effective properties.

6 Onset of Yielding, Iterative Scheme, and Plastic Strain

The composite begins its noticeable plastic deformation when the average stress in the matrix reaches its initial yield point. As the stress state of the matrix is usually triaxial, the yield condition needs to be established with its generalized effective stress σ^*. For convenience we also introduce the effective stress of the composite as

$$\bar{\sigma}^* = \left(\frac{3}{2}\bar{\sigma}_{ij}'\bar{\sigma}_{ij}'\right)^{1/2}, \tag{6.1}$$

in parallel to the first of equations (2.6). For a traction-prescribed condition $\sigma_{ij}'^{(0)} = b_0\bar{\sigma}_{ij}'$ and thus $\sigma^* = b_0\bar{\sigma}^*$. The initial yield condition of the composite is, therefore, satisfied when

$$\bar{\sigma}^* = \sigma_y/b_0, \tag{6.2}$$

where b_0—given by (4.12)—is to be evaluated for the elastic case (the superscript s deleted).

For a displacement-prescribed condition $\sigma_{ij}'^{(0)} = 2\mu_0^s\epsilon_{ij}'^{(0)} = 2\mu_0^s q_0\bar{\epsilon}_{ij}'$; the initial yielding occurs at

$$\left(\frac{3}{2}\bar{\epsilon}_{ij}'\bar{\epsilon}_{ij}'\right)^{1/2} = \sigma_y/2\mu_0 q_0, \tag{6.3}$$

with q_0 also calculated for the elastic state.

When the applied $\bar{\sigma}$ or $\bar{\epsilon}$ exceeds the value for the initial yielding, the matrix will be in the plastic state. Since the secant moduli L_0^s is not known *a priori*, an iterative procedure is usually required.

We may start by assuming an ϵ^{p*} for the matrix, and calculate the corresponding E_0^s by equation (2.8). The other secant moduli then follow from equations (2.3) and (2.4), and so do the values of α_0^s and β_0^s from equations (3.15).

For the traction-prescribed condition the stress concentration factor b_0 and consequently $\sigma_{ij}'^{(0)}$ then can be evaluated from equations (4.12), and this allows one to determine σ^* from the first of equations (2.6). If this σ^* is identical, or very close to the one calculated from the constitutive equation (2.5) with the assumed ϵ^{p*}, the solution is found. Otherwise a new ϵ^{p*}, reflecting the calculated σ^*, should be assumed to continue the search until the solution is found.

For the displacement-prescribed condition the strain concentration factor q_0 can be determined from equations (5.8), and so can the deviatoric stress $\sigma_{ij}'^{(0)}$ and σ^*. The rest of the procedure is identical to that for a prescribed traction.

Once the solution is found, the plastic strain of the composite can be determined also. While the total strain of the composite is given by the weighted average of those of its constituents, its plastic strain—due to the elastic heterogeneity—cannot be so calculated from its microscopic counterparts (Hill, 1967). Instead it should be determined from the unloading (elastic) process; that is

$$\bar{\epsilon}_{ij}^p = \left(\frac{1}{2\mu_s} - \frac{1}{2\mu}\right)\bar{\sigma}_{ij}' + \frac{1}{3}\delta_{ij}\left(\frac{1}{3\kappa_s} - \frac{1}{3\kappa}\right)\bar{\sigma}_{kk}, \tag{6.4}$$

where μ and κ are its effective elastic moduli, calculated from equations (4.18) and (4.19) with the sub- and the superscript s deleted, and δ_{ij} the Kronecker delta. Since

$$\epsilon_{ij}^p = \left(\frac{1}{2\mu_0^s} - \frac{1}{2\mu_0}\right)\sigma_{ij}'^{(0)},\quad \text{with } \sigma_{ij}'^{(0)} = b_0\bar{\sigma}_{ij}' \tag{6.5}$$

it is straightforward to show that, in terms of the weighted mean, $(1-c_1)\epsilon_{ij}^p$, the plastic strain is given by

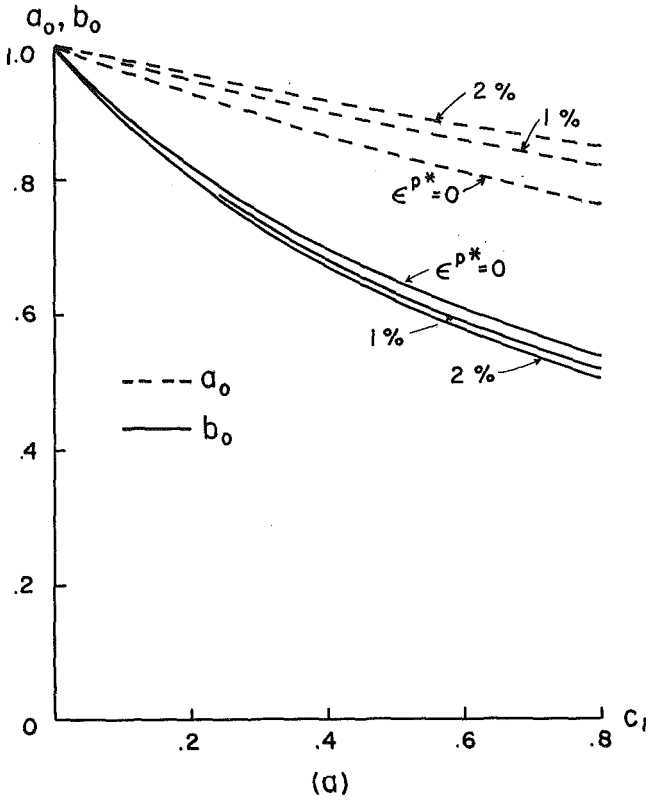
$$\bar{\epsilon}_{ij}^p = (1-c_1)\epsilon_{ij}^p + \left[1 - \frac{\mu_0}{\mu} - \frac{c_1(\mu_1 - \mu_0)}{[c_1 + (1-c_1)\beta_0^s](\mu_1 - \mu_0^s) + \mu_0^s}\right]\frac{\bar{\sigma}_{ij}'}{2\mu_0}$$

$$+ \frac{1}{3}\delta_{ij}\left[1 - \frac{\kappa_0}{\kappa} - \frac{c_1(\kappa_1 - \kappa_0)}{[c_1 + (1-c_1)\alpha_0^s](\kappa_1 - \kappa_0) + \kappa_0}\right]\frac{\bar{\sigma}_{kk}}{3\kappa_0}. \tag{6.6}$$

The plastic strain of the composite is seen to be equal to the weighted average only when both phases share the same elastic moduli.

7 Numerical Results and Comparison With Experiments

It is now of interest to apply the developed theory to a practical system. Let us consider the composite of an epoxy matrix reinforced with silica particles. The composite and the resin



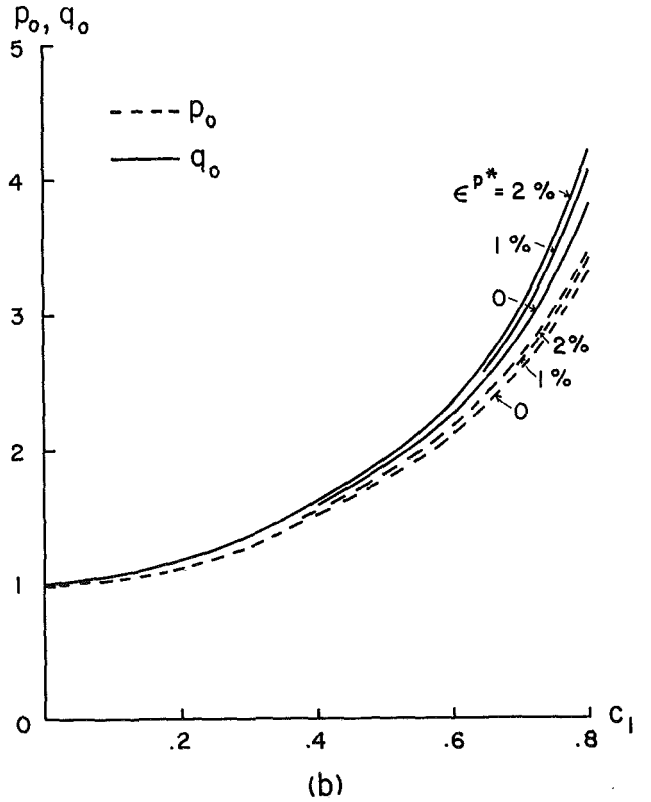


Fig. 3 (a) The average hydrostatic and deviatoric stress concentration factors, and (b) strain concentration factors, of the matrix phase under a prescribed traction and a prescribed displacement, respectively

were tested by Young and Beaumont (1977), who also gave the Young's moduli of the epoxy and silica as $E_0 = 3.16$ GPa and $E_1 = 73.1$ GPa, respectively. The Poisson ratios of the two are $\nu_0 = 0.35$ and $\nu_1 = 0.18$ (Kaye and Laby, 1966). Within the work-hardening region the experimental curve of the epoxy under compression gives $\sigma_y = 75.86$ MPa, $h = 32.18$ MPa, and $n = 0.26$. This system—with its hard particles and ductile matrix—is typical of the wide range of other practical ones and therefore the numerical results calculated with these constants should represent a general trend for this class of composites.

First the dependence of the hydrostatic and deviatoric stress concentration factors, a_0 and b_0, on the volume fraction of particles, c_1, are shown in Fig. 3(a) for three selected effective plastic strains of the matrix: $\epsilon^{p*} = 0$ (proportional limit), 1 percent and 2 percent. It is evident that these factors are strongly dependent on c_1. Especially the deviatoric component b_0, which determines the yield condition of the composite, is seen to decrease drastically with increasing c_1, thereby effectively improving its yield strength. At a given c_1 both a_0 and b_0 are also dependent on ϵ^{p*}, but plastic strain seems to have opposite effects to these two parameters. The strain concentration factors p_0 and q_0 under a prescribed displacement are displayed in Fig. 3(b). While the deviatoric component appears to be more strongly affected by c_1, the separation of the two is not as wide as in Fig. 3(a). Due to the ductility of the matrix, its strain concentration factors are much higher than the stress concentration factor of the inclusions (which typically are less than 2).

With these constituent properties we now apply the present theory to predict the stress-strain curves of the composite at five selected c_1. While under a prescribed $\bar{\sigma}$ or $\bar{\epsilon}$ an iterative procedure is usually required to find the average stress and strain of the constituents, such a process is *not* necessary at all for the derivation of stress-strain curves of the composite. Since $\sigma^* = b_0 \bar{\sigma}^*$, the constitutive equation (2.5) can be rewritten to provide the flow stress of the composite as

$$\bar{\sigma}^* = \frac{1}{b_0} \cdot [\sigma_y + h \cdot (\epsilon^{p*})^n], \tag{7.1}$$

which, under uniaxial compression, $\bar{\sigma}^* = \bar{\sigma}_{11}$.

Thus as ϵ^{p*} of the matrix gradually increases from 0 to certain value, its secant Young's modulus E_0^s also changes according to equation (2.8), and so do its ν_0^s, μ_0^s, α_0^s, and β_0^s according to equations (2.3), (2.4), and (3.15). The stress concentration factor b_0 at a given c_1 then follows from equations (4.12) (some have already been plotted in Fig. 3(a)), and this allows us to determine $\bar{\sigma}^*$ (or $\bar{\sigma}_{11}$ for this matter) from equation (7.1). The strain of the composite on the other hand can be calculated from the effective secant moduli κ_s and μ_s in equations (4.18) and (4.19). The plastic strain of the composite also follows from equation (6.4), which, for the uniaxial case, reduces to

$$\bar{\epsilon}_{11}^p = \left(\frac{1}{E_s} - \frac{1}{E}\right) \bar{\sigma}_{11}. \tag{7.2}$$

The experimental stress-plastic strain curves of the five composites, and that of the epoxy matrix, are reproduced in Fig. 4(a), and the theoretical predictions are given in Fig. 4(b). A more direct comparison between the two is depicted in Fig. 4(c) for the yield stress at the offset plastic strain $\bar{\epsilon}_{11}^p = 1$ percent. The experimental data in the latest figure are taken directly from Table III of Young and Beaumont (1977) and the theoretical curve is simply a smooth connection of the values of $\bar{\sigma}_{11}$ in Fig. 4(b) at this plastic strain. With the exception of $c_1 = 0.52$, the theoretical predictions are seen to be within the reasonable range of agreement. (In retrospect it appears from Fig. 4(a) that the increase of flow stress from $c_1 = 0.47$ to $c_1 = 0.52$—a 5 percent difference—is substantially higher than the increase from $c_1 = 0.42$ to $c_1 = 0.47$).

Some qualifications for the experimental comparison are needed at this point. First, unlike common metals, most

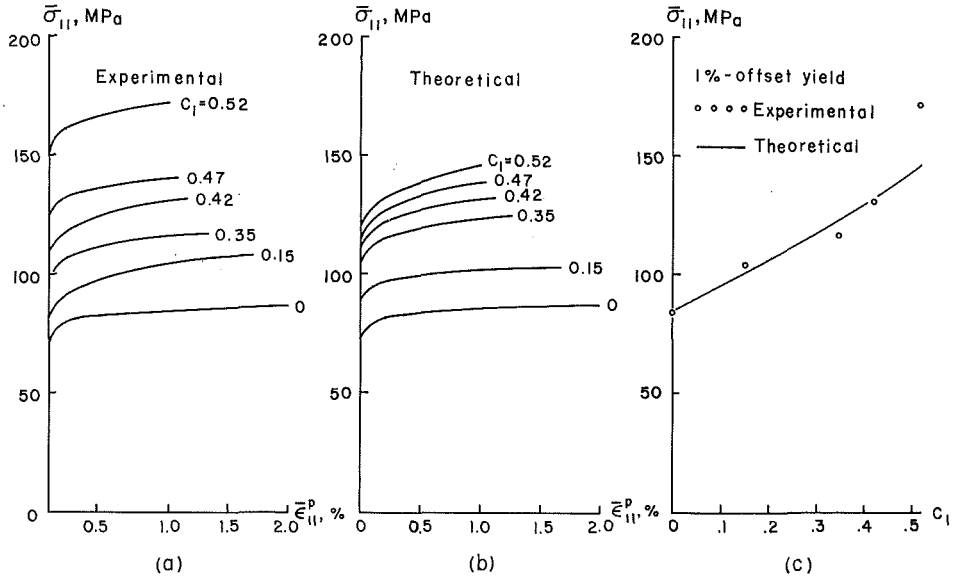


Fig. 4 Stress-plastic strain curves of the silica-particles/epoxy-resin composites under the uniaxial compression: (a) experimental data, (b) theoretical predictions, and (c) direct comparison for the 1 percent offset yield strength between the experiment and the theory

polymers experience a certain degree of pressure dependence in their yield behavior and this leads to the well-known strength-differential (S-D) effect between tension and compression. The S-D effect is usually more pronounced for thermoplastics and less so for thermosets, such as epoxy. The authors have not been able to find the precise value of S-D factor for the epoxy resin and its possible existence is neglected in the present calculation. Fortunately both the matrix and the composite systems were tested under compression; therefore, even if the S-D effect does exist, such a neglect should not introduce a gross error. (This would not do if the matrix was tested under compression whereas the composite were tested under tension, for instance.) Secondly, also unlike the behavior of common metals, the plastic deformation of polymers may introduce some volume expansion. However, Spitzig and Richmond's (1979) measurements have shown that the volume expansion for polyethylene is very small—only about one-sixth of that predicted from the normality structure (assuming a pressure-dependence), and that, for polycarbonate, it actually decreases—in complete violation of the associated flow rule. Indeed, as reported by Drucker (1973), "the major portion of the plastic volume expansion will have taken place well before the large offset yield strength is reached." No specific data on the plastic volume expansion for epoxy appears to have been reported and, in view of its possible small amount and uncertainty, such a term is also neglected here. But again since we are only concerned about the uniaxial stress-strain curve in the comparison, such an approximation should not result in a serious error. Thirdly, the silica particles in the specimens are of course not perfectly spherical, as assumed in the model. Despite these shortcomings, Young and Beaumont's data are preciously the only one we could find in the open literature.

8 Constitutive Equation of the Composite, and Comparison With Arsenault and Taya's Theory

For engineering applications, it is highly desirable to have a more explicit (perhaps even though approximate) constitutive equation of the composite, so that its stress-strain relation at a given c_1 can be readily deduced from that of the matrix and other pertinent properties of its constituents. To this end let us first examine the degree of accuracy if the plastic strain of the composite is calculated directly from the weighted average $(1-c_1)\epsilon^p_{ij}$, as an approximation. For this, the error—in percent—is plotted in Fig. 5 for four selected volume fractions as a function of $\bar{\epsilon}^p_{11}$. While one might tend to anticipate that the error would increase with increasing c_1, it is so only when the plastic strain is very small. Moreover, as $\bar{\epsilon}^p_{11}$ increases, the error tends to decrease. Within $c_1 \leq 0.4$, the error is below 6 percent. Since the tangent modulus decreases monotonically (in some cases the stress-strain curves are very flat; see Fig. 4) a 6 percent error in the plastic strain actually introduces a very small error in the flow stress. For the purpose of calculating the stress-strain curves of the composite the simple weighted average, therefore, should serve as a good approximation.

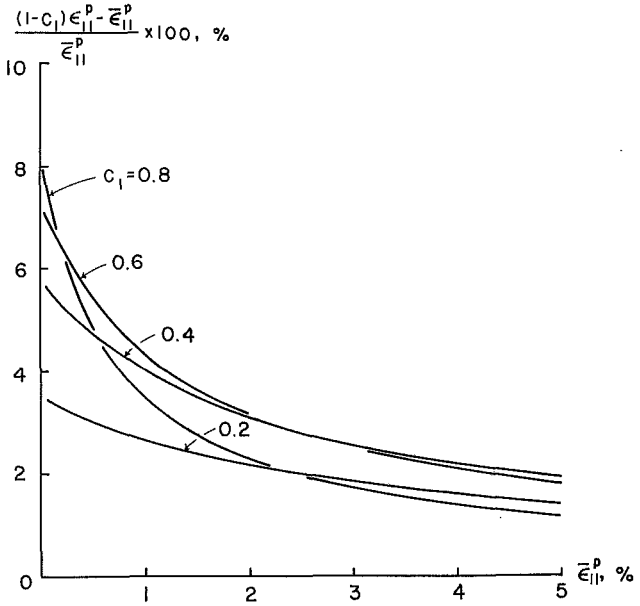


Fig. 5 The error percentage of the plastic strain of the composite as calculated from the direct weighted mean, $(1-c_1)\epsilon^p_{11}$

Then, using such an approximation, the constitutive equation of the composite can be derived from equation (7.1) as

$$\bar{\sigma}^* = \bar{\sigma}_y + H\bullet(\bar{\epsilon}^{p*})^n, \tag{8.1}$$

where

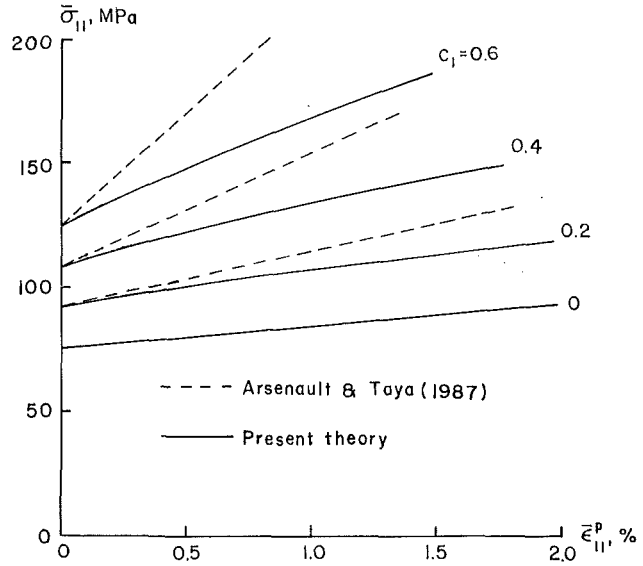


Fig. 6 The stress-strain curves of a silica/epoxy composite system calculated from the Arsenault and Taya (1987) theory, and the present one (assuming a bilinear stress-strain curve for the matrix)

$$\bar{\sigma}_y = \frac{\sigma_y}{b_0},\quad H = \frac{h}{b_0(1-c_1)^n},\quad \text{and}\quad \bar{\epsilon}^{p*} = (1-c_1)\epsilon^{p*}. \tag{8.2}$$

Since the values of b_0 and $(1-c_1)$ decrease with increasing c_1, the initial yield stress and the work-hardening modulus of the composite are seen to enhance with increasing hard particle concentration. The result is reversed with soft particles.

In light of the theoretical evolution outlined in Section 3, it is useful to give some quantitative comparison between the present theory and that of Arsenault and Taya (1987), which represents the latest development of the K-B-W line of approach. Specifically addressing the tensile stress-strain relation of a composite with a linearly hardening matrix (bilinear curve, $n=1$), they gave the initial yield stress $\bar{\sigma}_y$, and work-hardening modulus H, of the composite as

$$\bar{\sigma}_y/\sigma_y = C_0,\ H/h = C_2 + C_3\cdot(E_0/h), \tag{8.3}$$

where, after some algebra, we find their C_0, C_2, and C_3 to be

$$C_0 = \frac{1}{b_0},\qquad C_2 = \frac{1}{b_0(1-c_1)},$$

$$C_3 = \frac{3c_1(1-\beta_0)\mu_1}{2(1+\nu_0)[\beta_0(\mu_1-\mu_0)+\mu_0](1-c_1)}, \tag{8.4}$$

with b_0 evaluated from equation (4.12) at the *elastic* state (the superscript s deleted). (Note: A typo was found in their equation (22), or the last term of (8.3) here, in which E_0/h was given as h/E_0—or E/E_T versus E_T/E in their notations.) The hardening modulus H was also evaluated by assuming the weighted mean, $\bar{\epsilon}^p = (1-c_1)\epsilon^p$, for the composite.

Comparing equations (8.3) and (8.4) to (8.1) and (8.2) immediately suggests that both theories predict the same initial yield point for the composite but different hardening modulus. Since their theory was written for a linearly hardening matrix, we set $n=1$ and, using the flow stress at $\epsilon^p = 1$ percent as the basis, subsequently found $h = 814$ MPa for the above epoxy. The predicted stress-strain curves of the composite by these two theories are displayed in Fig. 6. As the K-B-W theory versus the B-Z theory in polycrystal plasticity (see again Fig. 1), the Arsenault-Taya theory generally leads to a stiffer response for the composite as compared to the present one. The difference between the two appears to be less significant at low plastic strain and particle content, but increases with increasing deformation and concentration.

Finally it should be reminded that the present theory—like many others mentioned in Section 3—is a mean-field one. The mean field approach is known to be highly accurate for the determination of linear effective properties, but generally less so for the nonlinear ones; for the latter class of problems, such an approach is also more reliable at low concentration of particles. Thus, while the present theory can in principle account for the inclusion-matrix interaction at finite concentration, its practical applications perhaps should be restricted to the lower range. More sophisticated theories ought to incorporate the information of point-to-point stress and strain distribution in the nonlinear analysis, and any simple results thus arrived at would be most attractive.

Acknowledgment

This work was supported by the National Science Foundation, Solid and Geo-Mechanics Program, under Grant MSM 86-14151.

References

Accorsi, M. L., and Nemat-Nasser, S., 1986, "Bounds on the Overall Elastic and Instantaneous Elastoplastic Moduli of Periodic Composites," *Mechanics of Materials*, Vol. 5, pp. 209–220.

Arsenault, R. J., and Taya, M., 1987, "Thermal Residual Stress in Metal Matrix Composite," *Acta Metallurgica*, Vol. 35, pp. 651–659.

Berveiller, M., and Zaoui, A., 1979, "An Extension of the Self-Consistent Scheme to Plastically-Flowing Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 26, pp. 325–344.

Budiansky, B., and Wu, T. T., 1962, "Theoretical Prediction of Plastic Strains of Polycrystals," *Proceedings, 4th U. S. National Congress of Applied Mechanics*, ASME, New York, pp. 1175–1185.

Chow, T. S., 1978, "Effect of Particle Shape at Finite Concentration on the Elastic Moduli of Filled Polymers," *Journal of Polymer Science: Polymer Physics Edition*, Vol. 16, pp. 959–965.

Drucker, D. C., 1973, "Plasticity Theory, Strength-Differential (SD) Phenomenon, and Volume Expansion in Metals and Plastics," *Metallurgical Transactions*, Vol. 4, pp. 667–673.

Dvorak, G. J., and Bahei-El-Din, Y. A., 1982, "Plasticity Analysis of Fibrous Composites," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 327–335.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems, *Proceedings of the Royal Society, London*, Vol. A241, pp. 376–396.

Hashin, Z., and Shtrikman, S., 1963, "A Variatiónal Approach to the Theory of the Elastic Behavior of Multiphase Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 11, pp. 127–140.

Hershey, A. V., 1954, "The Plasticity of an Isotropic Aggregate of Anisotropic Face-Centered Cubic Crystals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 21, pp. 241–249.

Hill, R., 1963, "Elastic Properties of Reinforced Solids: Some Theoretical Principles," *Journal of the Mechanics and Physics of Solids*, Vol. 11, pp. 357–372.

Hill, R., 1964, "Theory of Mechanical Properties of Fibre-Strengthened Materials: II. Inelastic Behaviour," *Journal of the Mechanics and Physics of Solids*, Vol. 12, pp. 213–218.

Hill, R., 1965, "Continuum Micro-Mechanics of Elastoplastic Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 13, pp. 89–101.

Hill, R., 1967, "The Essential Structure of Constitutive Laws for Metal Composites and Polycrystals," *Journal of the Mechanics and Physics of Solids*, Vol. 15, pp. 79–95.

Hutchinson, J. W., 1970, "Elastic-Plastic Behavior of Polycrystalline Metals and Composites," *Proceedings of the Royal Society, London*, Vol. A319, pp. 247–272.

Kaye, G. W. C., and Laby, T. H., 1966, *Tables of Physical and Chemical Constants*, Longmans, London.

Kröner, E., 1958, *Kontinuumstheorie der Versetzungen und Eigenspannungen*, Springer, Berlin.

Kröner, E., 1961, "Zur Plastischen Verformung des Vielkristalls," Acta Metallurgica, Vol. 9, pp. 155–161.

Mori, T., and Tanaka, K., 1973, "Average Stress in the Matrix and Average Elastic Energy of Materials with Misfitting Inclusions," *Acta Metallurgica*, Vol. 21, pp. 571–574.

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

Norris, A. N., 1985, "A Differential Scheme for the Effective Moduli of Composites," Mechanics of Materials, Vol. 4, pp. 1–16.

Pedersen, O. B., 1983, "Thermoelasticity and Plasticity of Composites-I. Mean Field Theory," *Acta Metallurgica*, Vol. 31, pp. 1795–1808.

Spitzig, W. A., and Richmond, O., 1979, "Effect of Hydrostatic Pressure on the Deformation Behavior of Polyethylene and Polycarbonate in Tension and in Compression," *Polymer Engineering and Science*, Vol. 19, pp. 1129–1139.

Tanaka, K., and Mori, T., 1970, "The Hardening of Crystals by Non-Deforming Particles and Fibers," *Acta Metallurgica*, Vol. 18, pp. 931–941.

Tanaka, K., Wakashima, K., and Mori, T, 1973, "Plastic Deformation Anisotropy and Work-Hardening of Composite Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 21, pp. 207–214.

Tandon, G. P., and Weng, G. J., 1986, "Stress Distribution in and Around Spheroidal Inclusions and Voids at Finite Concentration," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 511–518.

Taya, M., and Chou, T.-W., 1981, "On Two Kinds of Ellipsoidal Inhomogeneities in an Infinite Elastic Body: An Application to a Hybrid Composite," *International Journal of Solids and Structures*, Vol. 17, pp. 553–563.

Taya, M., and Mura, T., 1981, "The Stiffness and Strength of an Aligned Short-Fiber Reinforced Composite Containing Fiber-End Cracks Under Uniaxial Applied Stress," ASME JOURNAL OF APPLIED MECHANICS, Vol. 43, pp. 361–367.

Wakashima, K., Suzuki, Y., and Umekawa, S., 1979, "A Micromechanical Prediction of Initial Yield Surfaces of Unidirectional Composites," *Journal of Composite Materials*, Vol. 13, pp. 288–302.

Weng, G. J., 1982, "A Unified, Self-Consistent Theory for the Plastic-Creep Deformation of Metals," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 728–734.

Weng, G. J., 1984, "Some Elastic Properties of Reinforced Solids, With Special Reference to Isotropic Ones Containing Spherical Inclusions," *International Journal of Engineering Science*, Vol. 22, pp. 845–856.

Weng, G. J., 1985, "Micromechanical Determination of Two-Phase Plasticity," *International Journal of Plasticity*, Vol. 1, pp. 275–287.

Young, R. J., and Beaumont, P. W. R., 1977, "Failure of Brittle Polymers by Slow Crack Growth: Part 3. Effect of Composition Upon the Fracture of Silica Particle-Filled Epoxy Resin Composites," *Journal of Materials Science*, Vol. 12, pp. 684–692.

Optimal Design of Laminated-Composite Circular-Cylindrical Shells Subjected to Combined Loads

G. Sun
Shanghai Jiao Tong University,
Shanghai, People's Republic of China

J. S. Hansen
Institute for Aerospace Studies,
University of Toronto,
Downsview, Ontario, Canada M3H 5T6

Optimization of the buckling load of a laminated-composite, circular-cylindrical shell subjected to axial compression, external pressure, torsion, or a combination thereof is undertaken. In the optimization procedure it is assumed that the shell has a fixed weight (length, radius and thickness); the buckling load is taken as the objective function which is maximized by adopting the lamina fiber orientations as the optimizing parameters. For the shell analysis a perturbation approach is used and the boundary conditions and nonlinear prebuckling effects are included; the analysis yields both the buckling load and the post-buckling character of the shell. The procedure developed is demonstrated for eight loading configurations. In addition, selected laminates were chosen for an experimental programme involving a series of graphite/epoxy shells. The predicted analytical and the measured experimental buckling loads are in very good agreement.

Introduction

The buckling and post-buckling character of shell structures encompasses a class of problems which have received extensive attention in the literature. Of this work the majority of the emphasis has been placed on calculations for isotropic, homogeneous material constitutive laws. Less work, although still significant, has addressed buckling problems of composite cylinders. The present work deals with the latter case where the intention is to maximize the buckling load of a laminated-composite, circular-cylindrical shell as a function of ply orientation.

Perhaps the dominant contribution to the theory of elastic stability was that due to Koiter (1945), wherein it was demonstrated conclusively that initial geometric imperfections play a dominant role in reducing the buckling load of shell structures. An equivalent variation of Koiter's approach, written in a version of the principle of virtual work, was developed by Budiansky and Hutchinson (1964). Koiter (1963) also developed a special theory in which the effect of a small but finite imperfection in the shape of the classical axisymmetric buckling mode of a perfect cylindrical shell was specified. Further refinement in this theory allows the inclusion of boundary conditions and nonlinear prebuckling deformations and demonstrates the importance of including these effects in a stability analysis.

The optimal design of laminated cylinders for buckling has been addressed by relatively few authors. Papers which have contributed to this area are by Anderson and Stroud (1979), Dickson et al. (1980), Bauld and Khot (1982), Bushnell (1983), Knot and Bauld (1983), Nshanian and Pappas (1983), as well as Zimmerman (1985). The present study was prompted by the work of Tennyson and Hansen (1983) wherein it was demonstrated that the buckling strength of an axially loaded cylindrical shell could be increased significantly through a judicious choice of laminate configuration. The present work goes much further by combining the shell buckling calculation with an optimization algorithm. The optimization procedure adopted involves a two step approach. That is, a random search was employed to determine an initial guess for a systematic search based on the method due to Powell (1964). Such a strategy was dictated by the fact that the functional relationship between buckling load and ply orientation exhibits many local maxima. The random search was therefore intended to isolate the greatest of these local maxima. The objective function (buckling load) was calculated as an accurate numerical eigen-solution of the von Kármán-Donnell thin shell equations and incorporated combined loading (axial compression, external pressure, torsion), nonlinear prebuckling deformations, boundary conditions, and initial imperfections. In addition, the Koiter postbuckling coefficient b was evaluated to indicate the sensitivity to asymmetric imperfections. The procedure developed is demonstrated for eight example load cases; for each of these cases an optimal configuration was determined. In adddition, experiments for selected optimal, four-ply, graphite/epoxy cylinders were performed and comparisons were made to cylinders manufactured using

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 18, 1987; final revision October 7, 1987.

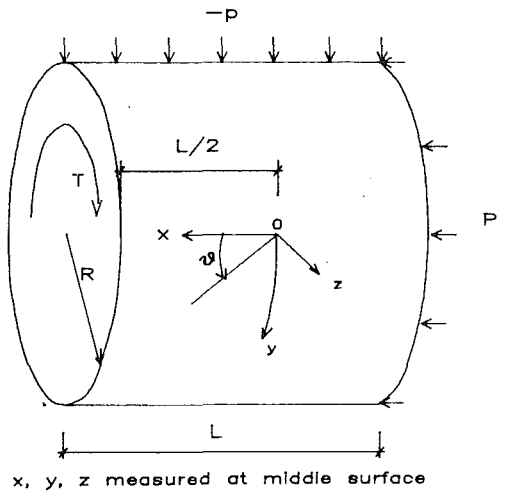


Fig. 1 Geometry and loading of the circular cylindrical shell

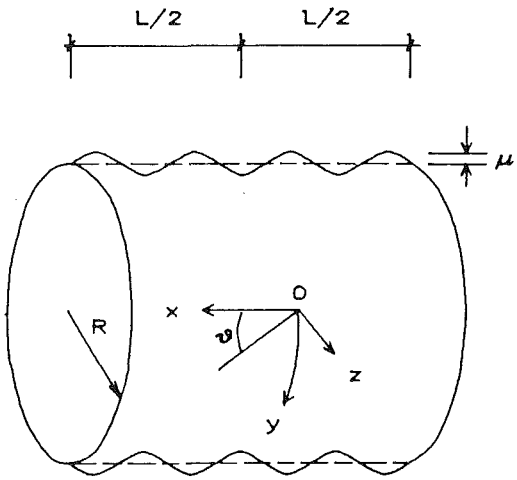


Fig. 2 Circular cylinder containing an axisymmetric geometric imperfection

reference laminates. The correlation between theory and experiment was found to be good and the improvements in buckling strength due to the optimization were significant.

Theoretical Analysis

Let the midsurface of a perfect circular cylindrical shell be the reference surface, and the origin of the coordinates be located at the midlength of the cylinder (Fig. 1). The coordinates x, y, z are measured in the axial, circumferential, and normal directions, respectively. The cylinder has length L, radius R and wall thicknss t. The parameter Z is defined as $Z=L^2/(Rt)$. Nondimensional parameters are defined as follows:

$$\bar{x}=\sqrt{Rt}\,x, \quad \bar{y}=\sqrt{Rt}\,y, \quad \bar{W}=tW$$

$$\bar{A}^*_{ij}=\frac{1}{E_{11}t}A^*_{ij}, \quad \bar{B}^*_{ij}=tB^*_{ij}, \quad \bar{D}^*_{ij}=E_{11}t^3D^*_{ij} \qquad (1)$$

$$(i,j=1,2,6)$$

$$\left[\begin{array}{c|c} \nabla^4_A{}^* & -\nabla^4_B{}^* \\ \hline \nabla^4_B{}^* & \nabla^4_D{}^* \end{array}\right]\left[\begin{array}{c} F \\ \hline W \end{array}\right]=\left[\begin{array}{c} (W_{,xy})^2-W_{,yy}(\tilde{W}_{,xx}+W_{,xx})+W_{,xx} \\ \hline F_{,yy}(\tilde{W}_{,xx}+W_{,xx})-2F_{,xy}W_{,xy}+F_{,xx}W_{,yy}-F_{,xx}+p \end{array}\right] \qquad (4)$$

$$\bar{p}=\frac{E_{11}t^2}{R^2}p, \quad \bar{F}=E_{11}t^3F, \quad (\bar{N}_x,\bar{N}_y,\bar{N}_{xy})=\frac{E_{11}t^2}{R}(N_x,N_y,N_{xy})$$

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} the compliance coefficients of the laminate (Ashton et al., 1969) are constant, F is the Airy stress function, p is the external pressure, E_{11} is the elastic modulus of a lamina in the fiber direction, while N_x, N_y, N_{xy} are the axial, circumferential, and shear stress resultants. Following the method of Tennyson et al. (1971) and Booton (1976), a small but finite initial axisymmetric midsurface deviation is assumed to be (Fig. 2),

$$\tilde{W}(x)=-\mu\cos(\omega x) \qquad (2)$$

where

$$\mu=\sqrt{2}\delta_{rms}/t, \quad \omega=(A^*_{22}D^*_{11})^{-1/4}$$

In the above, δ_{rms} is the measured root-mean-square value of the actual wall midsurface profiles and μ is the equivalent amplitude of the midsurface imperfection in the shape of the classical axisymmetric buckling mode. Also, ω is the normalized wave length parameter which corresponds to the classical axisymmetric buckling mode. As shown theoretically and experimentally, this imperfection yields a worst case buckling load under axial compression (Tennyson et al., 1971). It should be mentioned that it is known that the axisymmetric imperfection shape does not necessarily yield the worst form of imperfection for either pressure or torsional loads. However, in the work by Booton (1976), excellent correlation between theoretical predictions and experimental results were obtained for pressure and torsional loading cases based on the equivalent axisymmetric representation of the initial imperfection. Thus the same procedure was adopted in the present study. Introducing the Airy stress function,

$$F_{,yy}=N_x, \quad F_{,xx}=N_y, \quad F_{,xy}=-N_{xy} \qquad (3)$$

and using the nonlinear von Kármán-Donnell strain displacement relations, the compatibility and equilibrium equations in terms of W and F can be derived for the laminated circular cylinder containing an initial axisymmetric imperfection $\tilde{W}$ as follows,

where the operators

$$\nabla^4_A{}^*=A^*_{22}\frac{\partial^4}{\partial x^4}-2A^*_{26}\frac{\partial^4}{\partial x^3\partial y}+(2A^*_{12}+A^*_{66})\frac{\partial^4}{\partial x^2\partial y^2}-2A^*_{16}\frac{\partial^4}{\partial x\partial y^3}+A^*_{11}\frac{\partial^4}{\partial y^4}$$

$$\nabla^4_B{}^*=B^*_{21}\frac{\partial^4}{\partial x^4}+(2B^*_{26}-B^*_{61})\frac{\partial^4}{\partial x^3\partial y}+(B^*_{11}+B^*_{22}-2B^*_{66})\frac{\partial^4}{\partial x^2\partial y^2}+(2B^*_{16}-B^*_{62})\frac{\partial^4}{\partial x\partial y^3}+B^*_{12}\frac{\partial^4}{\partial y^4} \qquad (5)$$

$$\nabla^4_D{}^*=D^*_{11}\frac{\partial^4}{\partial x^4}+4D^*_{16}\frac{\partial^4}{\partial x^3\partial y}+2(D^*_{12}+2D^*_{66})\frac{\partial^4}{\partial x^2\partial y^2}+4D^*_{26}\frac{\partial^4}{\partial x\partial y^3}+D^*_{22}\frac{\partial^4}{\partial y^4}$$

The axisymmetric prebuckling deformation of the shell can be written as,

$$W_0 = W_0(x,\lambda), \quad F_0 = \frac{1}{2}N_x^0 y^2 - N_{xy}^0 xy + f_0(x,\lambda) \tag{6}$$

where the subscript 0 on W, F, and f, and the superscript 0 on N_x, N_{xy} indicates that they are prebuckling quantities. Substituting equation (6) into (4) leads to a fourth order ordinary differential equation for W_0,

$$(A_{22}^* D_{11}^* + B_{21}^{*2})\, W_{0,xxxx} + (2B_{21}^* - A_{22}^* N_x^0) W_{0,xx} + W_0$$

$$= N_x^0 (A_{12}^* + A_{22}^* \tilde{W}_{,xx}) + A_{22}^* p + A_{26}^* N_{xy}^0 \tag{7}$$

and an auxiliary equation for f_0,

$$f_{0,xx}(x) = N_y^0(x) = \frac{1}{A_{22}^*}(B_{21}^* W_{0,xx} + W_0 - A_{12}^* N_x^0 - A_{26}^* N_{xy}^0) \tag{8}$$

In order that the external loading be characterized by a single parameter, N_x^0, p, and N_{xy}^0 are expressed as

$$-N_x^0 = R_x\lambda; \quad -p = R_y\lambda; \quad N_{xy}^0 = R_{xy}\lambda \tag{9}$$

Here, the load parameter λ is assumed to be increasing from zero while R_x, R_y, R_{xy} are factors which are chosen to yield the desired ratio between the axial, pressure, and torsional loads. Thus the external loads (axial compression force $\bar{P}$, external pressure $\bar{p}$, and torque $\bar{T}$) are defined in terms of λ as

$$\bar{P} = 2\pi E_{11} t^2 R_x \lambda, \; \bar{p} = \frac{t^2}{R^2} E_{11} R_y \lambda, \; \bar{T} = 2\pi R E_{11} t^2 R_{xy} \lambda \tag{10}$$

The derivatives of prebuckling deformations W_0, f_0 with respect to the load parameter λ are defined as,

$$\dot{W}_0(x) = \frac{\partial W_0(x)}{\partial \lambda}, \quad \dot{f}_0 = \frac{\partial f_0(x)}{\partial \lambda} \tag{11}$$

These quantities are required at various prebuckling load levels in the solution of the buckling problem and the evaluation of postbuckling coefficient b. The equations for $\dot{W}_0$ and $\dot{f}_0$ can be directly derived by differentiating equations (7) and (8) with respect to λ. Considering (9), the fourth-order ordinary differential equation for $\dot{W}_0$ is,

$$(A_{22}^* D_{11}^* + B_{21}^{*2})\dot{W}_{0,xxxx} + (2B_{21}^* + A_{22}^* R_x \lambda)\dot{W}_{0,xx} + \dot{W}_0$$

$$= -R_x(A_{12}^* + A_{22}^* \tilde{W}_{,xx} + A_{22}^* W_{0,xx}) - A_{22}^* R_y + A_{26}^* R_{xy} \tag{12}$$

and from equation (8) the expression for $\dot{f}_0$ becomes

$$\dot{f}_{0,xx} = \frac{1}{A_{22}^*}(B_{21}^* \dot{W}_{0,xx} + \dot{W}_0 + A_{12}^* R_x - A_{26}^* R_{xy}) \tag{13}$$

The classical buckling load of the cylinder is designated as λ_c. In general, λ_c is a function of the initial axisymmetric imperfection $\tilde{W}$ of the cylinder; to avoid confusion in the presentation of results when $\tilde{W}$ is nonzero, λ_c will be written as λ_c^i, while if $\tilde{W}$ is zero it will remain as λ_c. Following the method outlined in Hutchinson and Frauenthal (1969) an asymptotic perturbation expansion of the solution valid in the neighborhood of the bifurcation point is taken in the form,

$$\begin{bmatrix} W \\ -- \\ F \end{bmatrix} = \begin{bmatrix} W_0 \\ -- \\ F_0 \end{bmatrix} + \epsilon \begin{bmatrix} W_{\mathrm{I}} \\ -- \\ F_{\mathrm{I}} \end{bmatrix} + \epsilon^2 \begin{bmatrix} W_{\mathrm{II}} \\ -- \\ F_{\mathrm{II}} \end{bmatrix} + \ldots\ldots \tag{14}$$

where W_0, F_0 are the prebuckling field, W_{I}, F_{I} are the buckling field and W_{II}, F_{II} are the second order field, and ϵ is the normalized amplitude of the buckling mode W_{I} and serves as the perturbation parameter. The load λ in the vicinity of the buckling load λ_c is expressed in the form,

$$\frac{\lambda}{\lambda_c} = 1 + a\epsilon + b\epsilon^2 + \ldots\ldots \tag{15}$$

where a and b are postbuckling coefficients. In every case considered in this study the first coefficient a is identically zero due to the sinusoidal nature assumed for the buckling mode. Thus the initial postbuckling behavior of the structure hinges on the sign and amplitude of b. A negative b means that the load carrying ability diminishes immediately after buckling and the shell is imperfection sensitive, while a positive b means that the shell retains the ability to withstand increased loads after buckling.

Substituting (14) into the governing equations (4) and setting the coefficients of ϵ to zero yields the following buckling equations,

$$\begin{bmatrix} \nabla_A^{4*} & -\nabla_B^{4*} \\ \nabla_B^{4*} & \nabla_D^{4*} \end{bmatrix} \begin{bmatrix} F_{\mathrm{I}} \\ W_{\mathrm{I}} \end{bmatrix} = \begin{bmatrix} -(\tilde{W}_{,xx} + W_{0,xx}^c) W_{\mathrm{I},yy} + W_{\mathrm{I},xx} \\ (W_{0,xx}^c + \tilde{W}_{,xx}) F_{\mathrm{I},yy} + R_x \lambda_c W_{\mathrm{I},xx} + 2R_{xy}\lambda_c W_{\mathrm{I},xy} \\ + F_{0,xx}^c W_{\mathrm{I},yy} - F_{\mathrm{I},xx} \end{bmatrix} \tag{16}$$

where the superscript c on W_0 and F_0 indicates that they are evaluated at the critical load λ_c. The solution of these equations is assumed in the separable form,

$$W_{\mathrm{I}} = w_1(x)\cos(Ny) + w_2(x)\sin(Ny)$$

$$F_{\mathrm{I}} = f_1(x)\cos(Ny) + f_2(x)\sin(Ny) \tag{17}$$

where the normalized wave number N is related to the integer circumferential wave number n by $N = \sqrt{t/R}\, n$.

Substituting equations (17) into (16) leads to four coupled fourth order ordinary differential equations which are discretized using the finite difference technique. The solution for the buckling load is found by using an extension of the standard inverse power method. This extension is required because the present eigenproblem does not have the simple form $[A]x = \lambda[B]x$; the prebuckling displacement W_0 is a nonlinear function of λ. Therefore, W_0 is written as $W_0 = W_{0_g} + \lambda \dot{W}_{0_g}$ where W_{0_g} and $\dot{W}_{0_g}$ are values of W_0 and $\dot{W}_0$ evaluated at λ_g (an initial guess of the least eigenvalue). After this substitution the inverse power method is applied to the linearized eigenproblem and the least eigenvalue is obtained using an iterative procedure.

Substituting equation (14) into the governing equations (4) and equating the coefficients of ϵ^2 to zero yields the equations for the second order field,

$$\begin{bmatrix} \nabla_A^{4*} & -\nabla_B^{4*} \\ \nabla_B^{4*} & \nabla_D^{4*} \end{bmatrix} \begin{bmatrix} F_{\mathrm{II}} \\ W_{\mathrm{II}} \end{bmatrix} = \begin{bmatrix} -(\tilde{W}_{,xx} + W_{0,xx}^c) W_{\mathrm{II},yy} + W_{\mathrm{II},xx} + W_{\mathrm{I},xy}{}^2 - W_{\mathrm{I},xx} W_{\mathrm{I},yy} \\ (W_{0,xx}^c + \tilde{W}_{,xx}) F_{\mathrm{II},yy} + R_x \lambda_c W_{\mathrm{II},xx} + 2R_{xy}\lambda_c W_{\mathrm{II},xy} - F_{\mathrm{II},xx} \\ + F_{0,xx}^c W_{\mathrm{II},yy} + F_{\mathrm{I},yy} W_{\mathrm{I},xx} - 2F_{\mathrm{I},xy} W_{\mathrm{I},xy} + F_{\mathrm{I},xx} W_{\mathrm{I},yy} \end{bmatrix}$$

The solution to the postbuckling equations is taken in the form

$$W_{II} = w_3(x) + w_4(x)\cos(2Ny) + w_5(x)\sin(2Ny)$$

$$F_{II} = f_3(x) + f_4(x)\cos(2Ny) + f_5(x)\sin(2Ny) \tag{19}$$

With the solution of the above sets of equations, it is possible to evaluate the post-buckling coefficient b (Hutchinson and Frauenthal, 1969),

$$b = -\frac{b_1}{b_2}\frac{b_3}{b_4} \tag{20}$$

where

$$\begin{aligned} b_1 &= F_{II}^*(W_I, W_I) + 2F_I^*(W_I, W_{II}) \\ b_2 &= \lambda_c[\dot{F}_0^c{}^*(W_I, W_I) + 2F_I^*(\dot{W}_0^c, W_I)] \\ b_3 &= \{F_0^c{}^*(W_I, W_I) + F_I^*(W_0^c, W_I)\}^2 \\ b_4 &= \{\lambda_c[\dot{F}_0^c{}^*(W_I, W_I) + 2F_I^*(\dot{W}_0^c, W_I)]\}^2 \end{aligned} \tag{21}$$

and the following shorthand notation has been used,

$$A^*(B,C) = \iint [A_{,xx}B_{,y}C_{,y} + A_{,yy}B_{,x}C_{,x} - A_{,xy}(B_{,x}C_{,y} + B_{,y}C_{,x})]dxdy \tag{22}$$

It should be noted that in the present work the b coefficient is evaluated based on a cylinder containing an initial axisymmetric imperfection; the usual approach evaluates b based on a perfect cylinder.

In order to provide a valid comparison with experiments on cylinders of finite length, fully clamped boundary conditions at the end $x = \sqrt{Z}/2$ are imposed. That is,

$$W = 0, \quad W_{,x} = 0, \quad U_{,y} = 0, \quad V_{,y} = 0 \tag{23}$$

where U and V are axial and circumferential displacements, respectively. The boundary conditions at the midlength of the cylinder are obtained by imposing symmetry and continuity requirements on W and F. A precise specification of these boundary conditions in terms of W and F is given in Sun (1986).

Optimization Procedure

Cylinders manufactured from four-ply Hercules AS4/3501-6 graphite/epoxy with lamina properties $E_{11} = 14.6 \times 10^{10}$ N/m^2, $E_{22} = 1.08 \times 10^{10}$ N/m^2, $G_{12} = 0.58 \times 10^{10}$ N/m^2, $\nu_{12} = 0.29$ were selected for investigation. The buckling load of the perfect cylinder ($\mu = 0$) with the nominal geometry $R/t = 165$, $Z = L^2/(Rt) = 500$ was taken as the objective function and the orientation of each layer was taken as the design parameter. It may seem strange to conduct the optimization for the perfect cylinder ($\mu = 0$) rather than the imperfect cylinder ($\mu \neq 0$). The problem is that the initial imperfections are random variables and thus are not known prior to manufacture of the cylinder. On the other hand, once a cylinder has been manufactured it is too late to perform an optimization. If the statistics of the random imperfections were known it would be possible to perform the optimization based on these data; however, in the present case the imperfection statics are not available and thus the optimization was carried out on the perfect ($\mu = 0$) cylinders and it was assumed that this yielded an optimum (or near optimum) for the imperfect case.

A very powerful unconstrained optimization technique (not requiring the use of derivatives), due to Powell (1964), was employed. The computer code for calculating buckling loads was merged with an alogrithm developed by Powell (Kuester and Mize, 1973) to conduct a systematic search. Powell's method requires the objective function to possess a unique maximum; however, in most cases the objective function for the buckling load has many local maxima. Thus, the initial point for the systematic search was obtained using a random search technique. That is, a set of uniformly distributed random laminate configurations was produced and the buckling loads corresponding to this set of laminates were calculated; the particular laminate producing the highest buckling load was then chosen as the starting laminate for Powell's method. As listed in Table 1, optimal laminate configurations were evaluated for eight example loading cases. They were: (1) pure axial compression; (2) pure external pressure; (3) pure torsion; (4) axial compression and torsion with a load ratio $R_x:R_{xy} = 3:1$; (5) and (6) axial compression and external pressure with load ratios $R_x:R_y = 3:1$ and $R_x:R_y = 8:1$; (7) and (8) torsion and external pressure with load ratios $R_{xy}:R_y = 1:1$ and $R_{xy}:R_y = 6:1$. It is noted that for torsional loading a positive torque was assumed; the laminate achieved is not optimal for a negative torque and if this was desired a different strategy should be adopted. Several runs of the optimization procedure were conducted for each loading case, the results of which are presented in Table 1. In that table the starting laminates for Powell's method, obtained from the random search, are listed in the third column while the optimal configurations obtained from Powell's method are listed in the fourth column. It was found that the starting laminates resulting from the random search exhibited a large scatter, but the optimal configurations are essentially reproducible (within the convergence criteria adopted). The exception to this general result was the one run for loading case (4), $R_x:R_{xy} = 3:1$ where it may be seen that Powell's method converged to a local optimum. This was the only such case encountered, but it does indicate the need for a series of runs in order to assure that a global optimum has been achieved.

Table 1 Optimal configurations for example loading cases

Loading Case	No. of Random Runs	Starting Laminate	Optimal Configuration	λ_c	b
(1) R_x=1	400	21,-44,90,19 40,-40,86,16 27,-44,70,-06 26,-51,71,-10	26,-42,76,-03 28,-38,81,-10 27,-41,77,-07 27,-38,77,-10	0.199 0.197 0.199 0.197	-0.048 -0.050 -0.053 -0.051
(2)*** R_y=1	50	89,-34,-14,-84 65,89,53,84 80,-52,12,70 81,11,50,-59 89,37,-13,-79	84,-40,16,90 85,-39,15,90 85,-39,16,90 86,-35,21,90 85,-40,17,90	0.0365 0.0365 0.0365 0.0366 0.0366	-0.012 -0.009 -0.012 -0.012 -0.009
(3) R_{xy}=1	50	-60,59,12,-76 -69,-22,64,-88 -68,67,-45,62	-67,30,26,-70 -67,29,25,-71 -66,33,19,-71	0.111 0.111 0.111	-0.023 -0.023 -0.018
(4) R_x:R_{xy}=3:1	400	-62,-04,42,-43 -56,23,90,-52 -53,27,57,-49 -52,83,28,-56	-59,05,51,-59 -55,16,62,-48 -55,14,56,-52 -53,-84,38,-56*	0.0547 0.0547 0.0546 0.0486	-0.072 -0.045 -0.058 -0.038
(5)** R_x:R_y=3:1	400	84,19,-11,77 88,31,-13,-87 75,-41,34,84	-87,35,-22,89 -86,36,-22,89 -87,33,-23,88	0.0299 0.0299 0.0299	-0.029 -0.029 -0.029
(6)** R_x:R_y=8:1	400	78,-69,04,-78 82,-62,-11,85	79,-68,10,90 82,-64,12,-87	0.0166 0.0165	-0.066 -0.080
(7)** R_{xy}:R_y=1:1	400	-82,-17,45,-74 -76,50,-03,-83 -89,42,31,87	-83,32,17,-87 -82,33,14,-86 -82,32,16,-86	0.0337 0.0337 0.0337	-0.017 -0.017 -0.017
(8)** R_{xy}:R_y=6:1	400	-82,40,40,-56 -68,34,-05,-71 -72,48,02,-80	-72,26,27,-74 -71,29,25,-74 -70,29,24,-73	0.0145 0.0145 0.0145	-0.023 -0.022 -0.023

*This run obviously converged to a local maximum
**R_y is hydrostatic pressure, i.e., an additional $R_y/2$ axial compression load is included.
***R_y is lateral pressure.

Table 2 Comparison cylindrical shell data for experiments radius $R = 3.28$ in., length $L = 6$ in.

Shell No.	Optimized for	Laminate Configuration	t	μ	R/t	Z
1	R_x=1	26,-42,76,-03	0.0198	0.038	166	554
2	R_y=1	-85,38,-18,90	0.0198	0.034	166	554
3	R_{xy}=1	-67,31,23,-71	0.0198	0.057	166	554
4	R_x:R_{xy}=3:1	-59,05,51,-59	0.0196	0.055	168	559
5	R_x:R_y=3:1	-87,34,-22,89	0.0197	0.073	167	557
6	R_x:R_y=8:1	79,-68,10,90	0.0194	0.055	169	565
7	R_{xy}:R_y=1:1	-83,37,15,-86	0.0197	0.069	167	557
8	R_{xy}:R_y=6:1	-70,27,24,-73	0.0198	0.048	166	554
9	Reference	90,0,0,90	0.0194	0.049	169	565
10	Reference	0,45,-45,90	0.0197	0.045	167	557
11	Reference	0,90,90,0	0.0199	0.029	165	551

t, Wall thickness (in.).
μ, Amplitude of axisymmetric imperfection (normalized to wall thickness).

Table 3 Comparison of experiments with theory

Shell No.	Loading Ratio	Laminate Configuration	Theory λ_c^1 b	Experiment λ_c^1	% of theory
1.	R_x=1	26,-42,76,-03*	0.153 -0.28	0.131	85.5 %
11.		0, 90,90,0	0.11 -0.2	0.095	86.2 %
4.		-59,05,51,-59	0.099 -0.069	0.082	83.0 %
9.		90,0,0,90	0.090 0.002	0.0828	91.8 %
7.		-83,37,15,-86	0.084 -0.087	0.081	96.7 %
2.	R_y=1	-85,38,-18,90*	0.0332 -0.017	0.0294	88.5 %
5.		-87,34,-22,89	0.0326 -0.011	0.0284	87.1 %
10.		0,45,-45,90	0.0136 -0.018	0.0119	87.1 %
1.		26,-42,76,-03	0.0108 -0.028	0.0094	87.3 %
11.		0,90,90,0	0.0105 -0.019	0.0088	84.2 %
3.	R_{xy}=1	-67,31,23,-71*	0.107 -0.022	0.0978	91.4 %
7.		-83,37,15,-86	0.096 -0.020	0.0814	84.8 %
9.		90,0,0,90	0.056 -0.009	0.051	90.6 %
4.	R_x:R_{xy}=3:1	-59,05,51,-59*	0.0408 -0.061	0.0397	97.3 %
7.		-83,37,15,-86	0.0275 -0.089	0.027	98.2 %
9.		90,0,0,90	0.0238 -0.021	0.0222	93.4 %
6.	R_x:R_y=8:1	79,-68,10,90*	0.0112 -0.084	0.0108	97.2 %
10.		0,45,-45,90	0.0096 -0.066	0.0080	82.5 %
1.		26,-42,76,-03	0.0085 -0.076	0.0068	79.8 %
8.	R_{xy}:R_y=6:1	-70,27,24,-73*	0.0140 -0.021	0.0134	95.4 %
6.		79,-68,10,90	0.0092 -0.012	0.0083	90.6 %
10.		0,45,-45,90	0.0043 -0.017	0.0040	93.7 %
11.		0,90,90,0	0.0038 -0.021	0.0037	97.4 %

*The optimized configuration for the corresponding loading ratio.

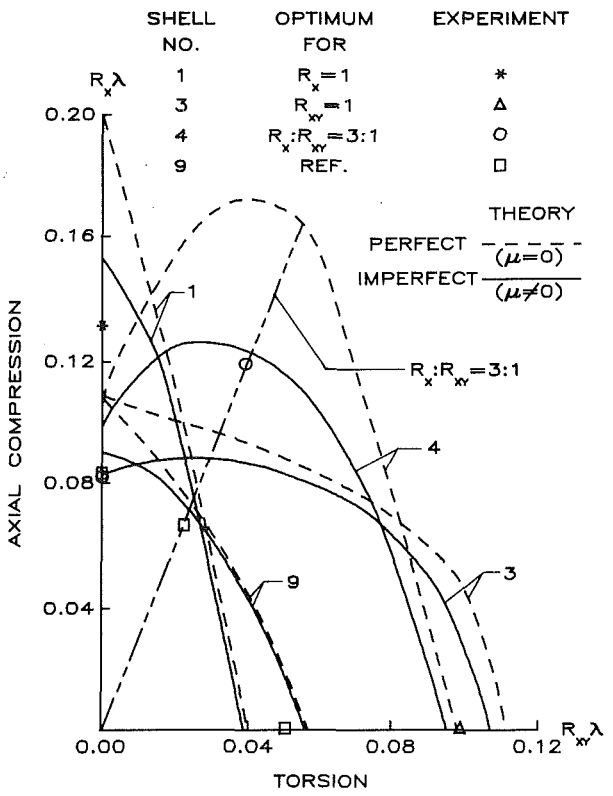


Fig. 3 Critical combinations of axial-compressive and torsional loads

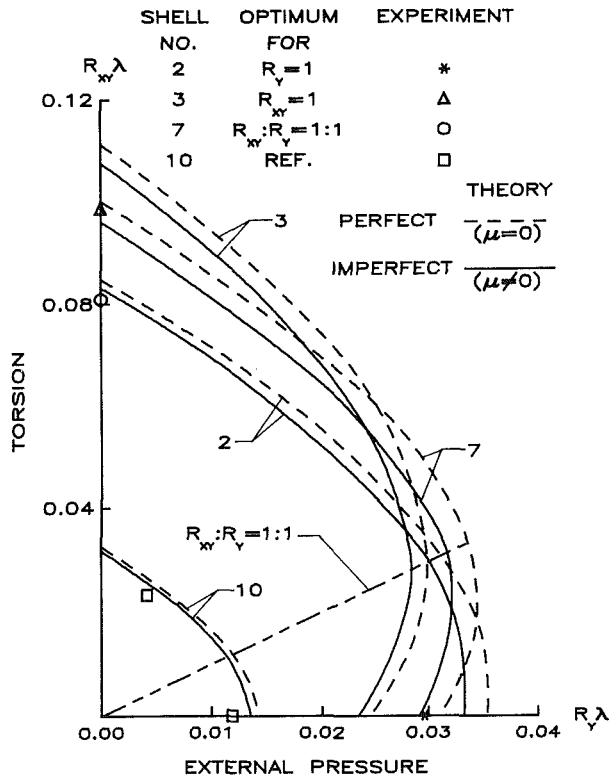


Fig. 4 Critical combinations of torsional and external-pressure loads

Experiments and Discussions

A total of 11 four-ply cylinders were manufactured having dimensions of 15.24 cm in length, 16.66 cm in diameter, and slightly less than a 0.05 cm wall thickness (Table 2). Shells No. 1–8 were laminated according to the eight optimized configurations. Shells No. 9–11 have wall configurations (90,0,0,90), (0,45,−45,90), and (0,90,90,0) and were manufactured to serve as reference data. (Numbers in parenthesis are orientations of each ply from the inner wall of the shell out; the degree sign notation has been omitted.) In order to expand the data set, each shell was tested under a variety of loading ratios if no detectable damage was found from previous experiments. The amplitude μ of an equivalent axisymmetric imperfection was determined by measuring the surface profiles of the test cylinders as described in Tennyson et al. (1971). These equivalent axisymmetric imperfections were then included in the calculated buckling loads which were determined in order to provide comparisons for the experimentally measured buckling loads. As may be seen in Table 3, the experimental buckling loads range from 78.9 percent to 98.2 percent of the numerically predicted values. This

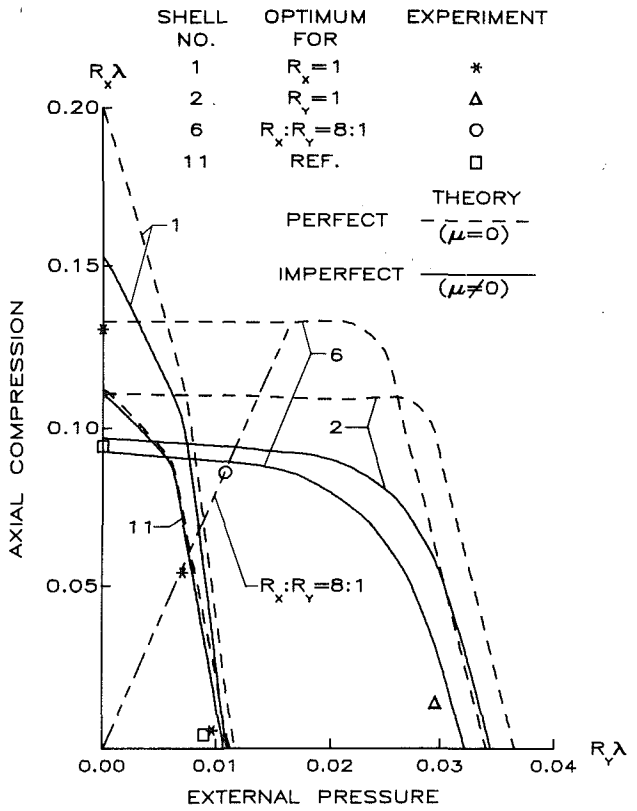


Fig. 5 Critical combinations of axial-compressive and external-pressure loads

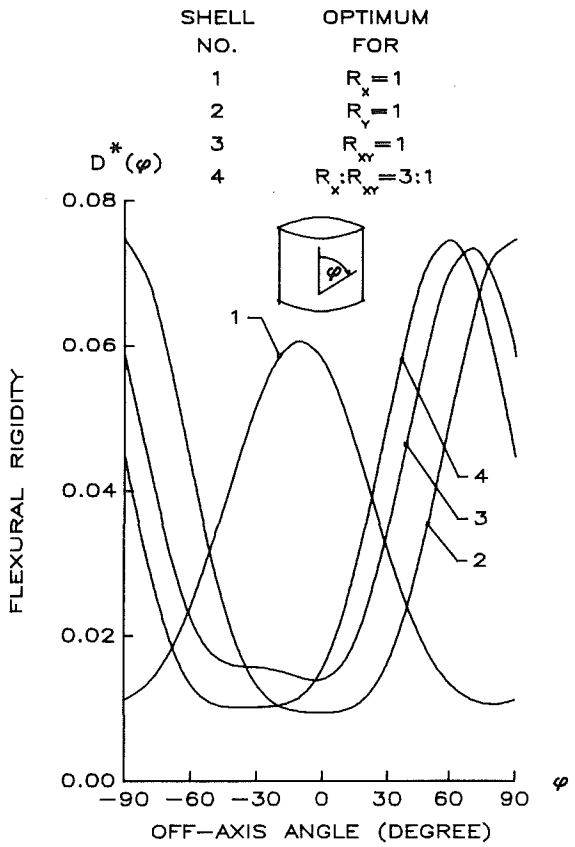


Fig. 6 Distribution of flexural rigidity $D^*(\phi)$

is felt to be very good agreement; the discrepancies may be attributed to factors such as the equivalent axisymmetric imperfection hypothesis, eccentricities in the loading, other manufacturing defects as well as theoretical shortcomings.

It is furthermore to be noted that the experimentally determined buckling loads for the optimal configurations are found to be 1.62, 3.19, 1.92, 1.79, 1.59, and 3.62 times as high as the least buckling loads of the reference cylinders, respectively, for the six loading cases. The other two loading cases omitted from Table 3 (as compared to Table 1) were not tested because of limitations of the experimental equipment.

In Figs. 3–5, theoretical interaction curves for the buckling load under the combination of two of the three load components are plotted. In each figure, theoretical curves for three optimal configurations (two of which were optimized for single loads and the third for combined loads) and one reference configuration are plotted against experimental results. It is obvious that the buckling strength is enhanced at the particular loading ratio for which the cylinder is optimized.

The optimal configurations determined do not appear to be easily rationalized. However, it is interesting to note the character of the flexural rigidity of the shell as a function of orientation since the flexural rigidity may be expected to play an important role in the buckling process. For the laminated cylinders considered, the flexural rigidity varies as a function of orientation; here it is designated as $D^*(\phi)$ where ϕ is the angle measured from the x axis. In Fig. 6, the $D^*(\phi)$ is plotted as a function of the angle ϕ for shells No. 1–4. Even though none of the optimized laminates exhibit any apparent symmetry, it is found that the $D^*(\phi)$ curve of Shell 1, which is optimized for axial compression, has a maximum nearly symmetric about axial direction; furthermore the curve for Shell 2, which is optimized for lateral pressure, has a maximum nearly symmetric about the circumferential direction. Between the two extremes are the other two maxima of $D^*(\phi)$ curves for Shells No. 3 and 4. It seems that the laminates are arranged to have a high flexural rigidity approximately along the principal compressive axis for the particular loading state in question in order to resist buckling.

In the present study only four-ply cylinders were considered, which simplified both the numerical and experimental work. It is anticipated that if more plies are considered the buckling load may be increased. To demonstrate this point, an additional optimization run was carried out for a fictitious six-ply cylinder with the same geometry as the four-ply cylinders. The resulting optimal configuration is $(31, -36, 85, 16, 83, -06)$. The buckling load is $\lambda_c = 0.226$ for axial compression, 1.14 times as high as Shell 1 with $\lambda_c = 0.198$.

Conclusion

Based on a clamped, finite-length four-ply graphite/epoxy cylinder and the finite difference solution to von Kármán-Donnell shell equations, it was demonstrated theoretically and experimentally that an optimal laminate configuration for buckling exists for each of the eight example loading cases. In addition it was demonstrated that the improvement in the buckling strength resulting from optimization is significant. Therefore, the present work provides a sound analysis as well as a feasible algorithm for the optimal design of laminated cylinders for buckling.

Acknowledgment

The Authors wish to express their appreciation to Prof. R. C. Tennyson for valuable discussions. This work was supported by the Natural Sciences and Engineering Research Council of Canada under Grant A3663.

References

Anderson, M. S., and Stroud, W. J., 1979, "General Panel Sizing Computer Code and Its Application to Composite Structural Panels," *AIAA Journal*, Vol. 17, No. 8, pp. 892–897.

Ashton, J. E., Halpin, J. C., and Petit, P. H., 1969, *Primer on Composite Materials: Analysis*, Technomic Publication, Westport, CT.

Bauld, N. R., Jr., and Khot, N. S., 1982, "A Numerical and Experimental Investigation of the Buckling Behavior of Composite Panels," *Computers and Structures*, Vol. 15, pp. 393–403.

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, Julius Springer-Verlag, Berlin, pp. 636–651.

Bushnell, D., 1983, "PANDA—Interactive Program for Minimum Weight Design of Stiffened Cylindrical Panels and Shells," *Computers and Structures*, Vol. 16, pp. 167–185.

Dickson, J. N., Biggers, S. B., and Wang, J. T. S., 1980, "Design of Stiffened Composite Panels in the Post-Buckling Range," *Fibrous Composites in Structural Design*, E. M. Lenoe, D. W. Oplinger, J. J. Burke, eds., Plenum Press, New York, pp. 313–327.

Hutchinson, J. W., and Frauenthal, J. C., 1969, "Elastic Postbuckling Behavior of Stiffened and Barreled Cylindrical Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 784–790.

Khot, N. S., and Bauld, N. R. Jr., 1983, "Further Comparison of the Numerical and Experimental Buckling Behaviors of Composite Panels," *Computers and Structures*, Vol. 17, pp. 61–68.

Koiter, W. T., 1945, "On the Stability of Elastic Equilibrium," Thesis Delft, H. J. Paris, Amsterdam; English Translation AFFDL-TR-70-25, 1970.

Koiter, W. T., 1963, "The Effect of Axisymmetric Imperfections on the Buckling of Cylindrical Shells Under Axial Compression," TR 6-90-63-86, Lockheed Missiles and Space Co.

Kuester, J. L., and Mize, J. H., 1973, *Optimization Techniques with Fortran*, McGraw-Hill, New York.

Nshanian, Y. S., and Pappas, M., 1983, "Optimal Laminated Composite Shells for Buckling and Vibration," *AIAA Journal*, Vol. 21, No. 3, pp. 430–437.

Powell, M. J. D., 1964, "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," *Computer J.*, Vol. 7, pp. 155–162.

Sun, G., 1986, "Optimization of Laminated Cylinders for Buckling," Doctoral Thesis, Institute for Aerospace Studies, University of Toronto, Toronto, Ontario, Canada.

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 Practise*, Thompson, J. M. T., and Hunt, G. W., eds., Cambridge Univ. Press, Cambridge.

Zimmerman, R., 1986, "Optimization of Axially Compressed CFRP Cylinders," *Spacecraft Structures. Int. Conference*, CNES, Toulouse, Dec. 3–6, 1985, *Proceedings*, ESA SP-238, Apr. 1986, pp. 407–412.

An Approximate Analysis of Stresses in Multilayered Elastic Thin Films

E. Suhir
AT&T Bell Laboratories,
Murray Hill, NJ 07974
Mem. ASME

The analysis contains an engineering method for the approximate evaluation of thermally induced stresses in single and multilayered heteroepitaxial structures fabricated on thick substrates, with consideration of the finite size of the structure. The examined stresses include normal stresses, acting in the film layers themselves and responsible for their ultimate and fatigue strength, as well as interfacial stresses, responsible for film blistering and peeling. The developed formulas are simple, easy-to-use, and clearly indicate how material and structural characteristics affect the magnitude and the distribution of stresses and deflections. Some recommendations for smaller stresses in film structures are presented. The obtained results can be utilized as a guidance for physical design of multilayered heteroepitaxial structures in microelectronics.

Introduction

The first theoretical formula for the evaluation of stresses, arising in a thin film prepared on a thick substrate, was suggested by G. G. Stoney (1909), and is still widely utilized for stress calculation from the measured deformation of the substrate. This formula can be written down as follows:

$$\sigma_f = E_s^0 \frac{h_s^2}{6\rho h_f}, \tag{1}$$

where σ_f is the stress in the film, $E_s^0 = E_s/1-\nu_s$ is the generalized Young's modulus for the substrate material, E_s and ν_s are elastic constants of this material, h_s and h_f are thicknesses of the substrate and the film, respectively, and ρ is the radius of curvature.

The formula (1) can be obtained on the basis of the following elementary considerations. The bending moment, which a film, experiencing stress σ_f, applies to the substrate, is $M_f = \sigma_f h_f (h_s/2)$. On the other hand, this moment is related to the moment of inertia $I_s = h_s^3/12$ of the substrate cross-sectional area of unit width by the formula $M_f = E_s^0 (I/\rho) = (E_s^0 h_s^3)/(12\rho)$, where the generalized Young's modulus E_s^0 is used to account for the two-dimensional stress condition. The above two formulas for the bending moment result in the equation (1).

Stoney's formula has to be used, as long as the elastic constants of the film material and its thermal expansion mismatch with the substrate are unavailable. Otherwise the formula

$$\sigma_f = E_f^0 \Delta\alpha \Delta t, \quad \Delta\alpha = \alpha_f - \alpha_s, \tag{2}$$

should be utilized. This formula follows from the fact, that the strain $\epsilon_s = \alpha_s \Delta t + F/E_s^0 h_s$ in the substrate must be equal to the strain $\epsilon_f = \alpha_f \Delta t - F/E_f^0 h_f$ in the film. The force F, arising between the film and the substrate, can be found on the basis of the strain compatibility condition and is as follows: $F \cong E_f^0 h_f \Delta\alpha \Delta t$. Then the stress in the film is expressed by the formula (2), which, unlike Stoney's formula (1), reflects the role of the actual factors, affecting the stress in the film, while formula (1) can be even misleading, if used for the purpose of optimal structural design.

The formulas for the curvature and the maximum bow can be easily obtained from (1) and (2):

$$\frac{1}{\rho} = 6 \frac{E_f^0}{E_s^0} \frac{h_f}{h_s^2} \Delta\alpha \Delta t, \quad w_0 = 3 \frac{E_f^0}{E_s^0} \left(\frac{\ell}{h_s}\right)^2 h_f \Delta\alpha \Delta t, \tag{3}$$

where ℓ is half the film length. The existing methods of stress calculations for multilayered heteroepitaxial structures (Reinhart and Logan, 1973; Röll, 1976; Olsen and Ettenberg, 1977; Vilms and Kerps, 1982) are based on the formulas (1)–(3). Since these methods deal with the stresses in the film layered themselves, they can be used only for an indirect judgement of the level of the interfacial stresses, responsible for film blistering and peeling.

Therefore the major objective of the analysis below is to develop an engineering method for an approximate evaluation of the interfacial stresses in single and multilayered heteroepitaxial structures, fabricated on thick substrates, and to find out how the material and structural characteristics affect these stresses. This would enable one to decide what could be done in order to reduce the interfacial stresses if necessary. The suggested approach, based on the concept of the interface

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February 23, 1987; final revision, September 24, 1987.

Paper No. 88-WA/APM-14.

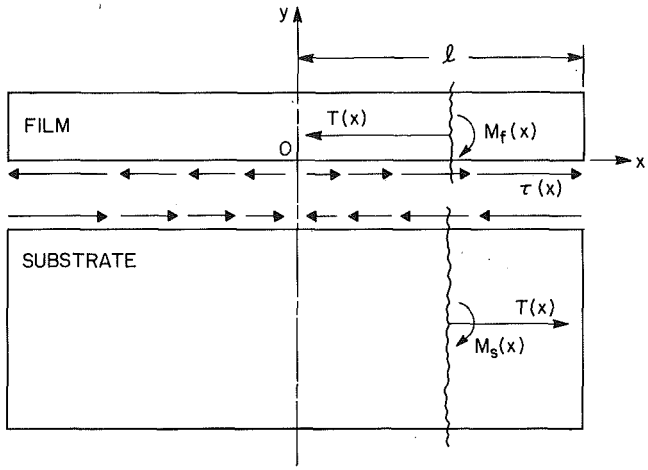


Fig. 1 Stress analysis model

compliance (Suhir, 1986), makes it possible to determine the magnitude and the distribution of the shearing and peeling stresses in the interfaces, as well as the normal stresses in the layers, with consideration of the finite size of the structure. It should be noted that the developed approach is equally applicable to other areas of physics and material science where thermal or lattice-mismatched stresses occur (Luryi and Suhir, 1986). It can also be used in those areas of engineering where lap shear joints subjected to external or thermally induced loading are utilized (Suhir, 1986).

Theory

1 Single-Layered Structure. Examine first the simplest case of a single layer structure fabricated at an elevated temperature and subsequently cooled (Fig. 1). The longitudinal displacements $u_f(x)$ and $u_s(x)$ of the lower extreme fiber of the film and the upper extreme fiber of the substrate, respectively, can be expressed by the formulas:

$$\left.\begin{aligned} u_f(x) &= \alpha_f \Delta t x - \lambda_f \int_0^x T(\xi)d\xi + \kappa_f \tau(x) + \frac{1}{2} h_f \int_0^x \frac{d\xi}{\rho(\xi)} \\ u_s(x) &= \alpha_s \Delta t x + \lambda_s \int_0^x T(\xi)d\xi - \kappa_s \tau(x) - \frac{1}{2} h_s \int_0^x \frac{d\xi}{\rho(\xi)} \end{aligned}\right\}, \quad (4)$$

where $\tau(x)$ is the shearing stress in the interface,

$$T(x) = \int_{-\ell}^{x} \tau(\xi)d\xi \quad (5)$$

is the shearing force per unit film width for the given cross section x, ℓ is half the film length, $\rho(x)$ is the radius of curvature, α_f and α_s are thermal expansion coefficients for the film and the substrate materials, $\lambda_f = (E_f^0 h_f)^{-1}$ and $\lambda_s = (E_s^0 h_s)^{-1}$ are coefficients of axial compliance for the film and the substrate, $E_f^0 = E_f/(1-\nu_f)$ and $E_s^0 = E_s/(1-\nu_s)$ are generalized Young moduli of the materials, h_f and h_s are the thicknesses of the film and the substrate,

$$\kappa_f = 2/3\,(1+\nu_f)/(1-\nu_f)\; h_f/E_f^0$$

and

$$\kappa_s = 2/3\,(1+\nu_s)/(1-\nu_s)\; h_s/E_s^0$$

are coefficients of interfacial compliance (Suhir, 1986), E_f and E_s are Young moduli, ν_f and ν_s are Poisson ratios for the film and the substrate materials, and Δt is the temperature differential. The origin O of the rectangular coordinates x, y is in the middle of the structure on the interface.

The first terms in equations (4) are unrestricted thermal contractions. The second terms are due to the forces (5) and are calculated under an assumption that these forces are uniformly distributed over the film and the substrate thicknesses. The third terms account for the additional displacements due to the actual (nonuniform) distribution of the above forces, and are calculated under an assumption that the corresponding corrections are directly proportional to the shearing stress in the given cross section and are not affected by the stresses in other cross sections. The last terms are due to bending.

The equation of equilibrium for a portion of the film-substrate structure is as follows:

$$\frac{h_f + h_s}{2} T(x) = M_f(x) + M_s(x), \quad (6)$$

where

$$M_f(x) = \frac{E_f^0 h_f^3}{12\rho(x)}, \quad M_s(x) = \frac{E_s^0 h_s^3}{12\rho(x)} \quad (7)$$

are bending moments acting over the cross sections of the film and the substrate. From the equations (6) and (7) we have:

$$\frac{1}{\rho(x)} = 6\,\frac{h_f + h_s}{E_f^0 h_f^3 + E_s^0 h_s^3}\, T(x) \cong 6\,\frac{T(x)}{E_s^0 h_s^2}. \quad (8)$$

After substituting (8) in (4) and using the condition $u_f(x) = u_s(x)$ of the displacement compatibility, we obtain the following basic equation for the unknown shearing stress function $\tau(x)$:

$$\tau(x) - k^2 \int_0^x T(\xi)d\xi = -\frac{\Delta\alpha\Delta t}{\kappa} x, \quad (9)$$

where

$$\left.\begin{aligned} k &= \sqrt{\frac{\lambda}{\kappa}}, \quad \lambda = \lambda_f + \lambda_s + \frac{(h_f+h_s)^2}{E_f^0 h_f^3 + E_s^0 h_s^3} \cong \lambda_f = \frac{1}{E_f^0 h_f}, \\ \kappa &= \kappa_f + \kappa_s = \frac{2}{3}\left(\frac{1+\nu_f}{1-\nu_f}\,\frac{h_f}{E_f^0} + \frac{1+\nu_s}{1-\nu_s}\,\frac{h_s}{E_s^0}\right), \\ \Delta\alpha &= \alpha_f - \alpha_s \end{aligned}\right\}. \quad (10)$$

The equation (9) has the following solution:

$$\tau(x) = -E_f^0 h_f \chi_1(x) \Delta\alpha\Delta t, \quad (11)$$

where the function $\chi_1(x) = k\,(\sinh kx)/(\cosh k\ell)$ characterizes the distribution of the shearing stress along the film. The maximum shearing stress is at the end of the film:

$$\tau_{\max} = \tau(\ell) = -kE_f^0 h_f \Delta\alpha\Delta t \tanh k\ell. \quad (12)$$

The solution (11) satisfies the condition $\tau(0) = 0$ of symmetry and the boundary condition $T(\ell) = 0$.

Calculations show that for the structures in question the k values are very great. Therefore, formulas (11) and (12) can be simplified as follows:

$$\tau(x) = \tau_{\max} e^{-k(\ell - x)}, \quad \tau_{\max} = -kE_f^0 h_f \Delta\alpha\Delta t. \quad (13)$$

As is evident from these formulas, the maximum shearing stress is independent from the film size, and the stresses drop exponentially with the decrease in x, i.e., they concentrate near the film ends. The length of the zone of high stresses can be defined, for instance, as such a length, for which the stresses decrease by, say, 95 percent of their maximum values at the edges. This results in the following formula for the length of the zone of appreciable shearing stresses: $\ell_i = -(\ln 0.05/k) \cong 3/k$.

After substituting equation (11) in (5) we find:

$$T(x) = E_f^0 h_f \chi_0(x) \Delta\alpha\Delta t, \quad (14)$$

where the function $\chi_0(x) = 1 - (\cosh kx)/(\cosh k\ell)$ characterizes the distribution of the forces $T(x)$ and the

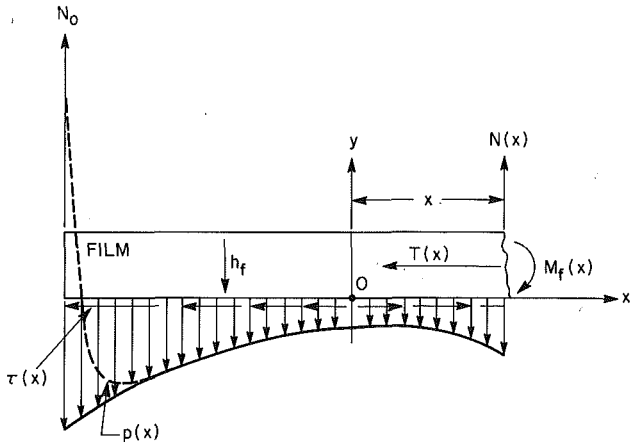


Fig. 2 Peeling stresses

resulting normal stresses along the film. Note, that $\chi_1(x) = -(d\chi_0(x))/(dx)$.

By substituting the equation (14) in (8) we obtain the following formula for the curvature:

$$\frac{1}{\rho(x)} = 6\frac{E_f^0 h_f}{E_s^0 h_s^2}\chi_0(x)\Delta\alpha\Delta t. \tag{15}$$

Then the formulas (7) yield:

$$\left.\begin{aligned} M_f(x) &= \frac{(E_f^0)^2 h_f^4}{2E_s^0 h_s^2}\chi_0(x)\Delta\alpha\Delta t, \\ M_s(x) &= \frac{E_f^0 h_f h_s}{2}\chi_0(x)\Delta\alpha\Delta t. \end{aligned}\right\} \tag{16}$$

Since the thickness of the film is small, the bending moment $M_f(x)$ is also small, and therefore the stresses in the film are due to the forces $T(x)$ only:

$$\sigma_f(x) = \frac{T(x)}{h_f} = E_f^0\chi_0(x)\Delta\alpha\Delta t. \tag{17}$$

For great k values, the function $\chi_0(x)$ can be calculated by the formula $\chi_0(x) = 1 - e^{-k(\ell-x)}$. As evident from this formula, for small enough x values, i.e., for the cross-sections sufficiently remote from the film ends, the factor $\chi_0(x)$ is close to unity, and the stresses in the film are independent from the location of the given cross-section along the film length. Near the ends, where the x coordinate is on the same order as half the film length ℓ, the stresses rapidly drop and turn to zero at the edges.

Stresses in the substrate, unlike the stresses in the film, are due both to the force $T(x)$ and the bending moment $M_s(x)$. The total stresses are maximum at the interface and can be evaluated as follows:

$$\sigma_s(x) = -\frac{T(x)}{h_s} - 6\frac{M_s(x)}{h_s^2} = -4E_f^0\frac{h_f}{h_s}\chi_0(x)\Delta\alpha\Delta t. \tag{18}$$

The deflection function can be found from the equation (15), since for small curvatures $1/(\rho(x)) \cong (d^2w(x))/(dx^2)$. For sufficiently great k values this formula results in the following equation for the maximum bow:

$$w_0 = 3\frac{E_f^0}{E_s^0}\left(\frac{\ell}{h_s}\right)^2 h_f\Delta\alpha\Delta t. \tag{19}$$

The expected distribution of the transverse normal (peeling) stresses is schematically shown in Fig. 2 by the dotted line. To simplify the analysis, however, we replace the stresses, directed upward and distributed over small areas near the edges, by concentrated forces N_0 applied at the edges. These forces, as well as the distributed peeling stress $p(x)$, can be found on the basis of the following equation of equilibrium for the portion of the film:

$$(x+\ell)N_0 - \int_{-\ell}^{x}\int_{-\ell}^{\xi} p(\xi')d\xi' d\xi' = M_f(x) - \frac{h_f}{2}T(x) \cong -\frac{h_f}{2}T(x).$$

Using the equation (14) we obtain:

$$N(x) = \int_{-\ell}^{x} p(\xi)d\xi = N_0 - \frac{1}{2}E_f^0 h_f^2\chi_1(x)\Delta\alpha\Delta t = N_0 + \frac{1}{2}h_f\tau(x), \tag{20}$$

$$p(x) = \frac{dN(x)}{dx} = -\frac{1}{2}E_f^0 h_f^2\Delta\alpha\Delta t\frac{d\chi_1(x)}{dx} = \frac{1}{2}E_f^0 h_f^2\chi_2(x)\Delta\alpha\Delta t, \tag{21}$$

where the function $\chi_2(x) = -(d\chi_1(x))/(dx) = -k^2(\cosh kx)/(\cosh k\ell)$ characterizes the longitudinal distribution of the peeling stress. Since the equilibrium condition requires that $N(\ell) = 0$, then

$$N_0 = -\frac{1}{2}h_f\tau_{\max} = \frac{1}{2}kE_f^0 h_f^2\Delta\alpha\Delta t. \tag{22}$$

The distributed peeling stress $p(x)$ is maximum at the end cross sections:

$$p_{\max} = p(\ell) = -\frac{1}{2}E_f^0(kh_f)^2\Delta\alpha\Delta t = \frac{1}{2}kh_f\tau_{\max}. \tag{23}$$

For great k values the formula (21) can be simplified as follows:

$$p(x) = p_{\max}e^{-k(\ell-x)}. \tag{24}$$

Hence, the distribution of the peeling stresses in this case is similar to the distribution of the shearing stress.

It is noteworthy, that while the normal stress in a thin film is independent from the film thickness, the interfacial stresses increase with an increase in the thickness of the film. Therefore, there is an incentive to reduce, for smaller interfacial stresses, the film thickness in the zone of high stresses by slanting the film edges. In addition, the interfacial stresses, unlike the stresses in the film, depend on the parameter k of the interfacial compliance, which, in its turn, depends on the thickness and the Young's modulus of the substrate. The stress in the film itself, however, is independent from the thickness and the Young modulus of the substrate, as long as these parameters are substantially greater than the thickness and the Young modulus of the film.

2 Multilayered Structure. Examine now a multilayered heteroepitaxial structure (Fig. 3). We assume, for instance, that the coefficients α_i ($i = 0, 1, 2, \ldots, m$ with $i = 0$ referring to substrate) of linear thermal expansion of the layers increase with an increase in the number i, and that the entire structure is subjected to uniform cooling.

The analysis, carried out for a single-layered structure, has indicated that the formula for the shearing force $T(x)$ in the film can be obtained by multiplying the force $T = E_f^0 h_f\Delta\alpha\Delta t$, calculated under an assumption that it is constant along the film, by the factor $\chi_0(x)$, which considers the longitudinal distribution of this force. The interfacial stresses can be then evaluated by the formulas

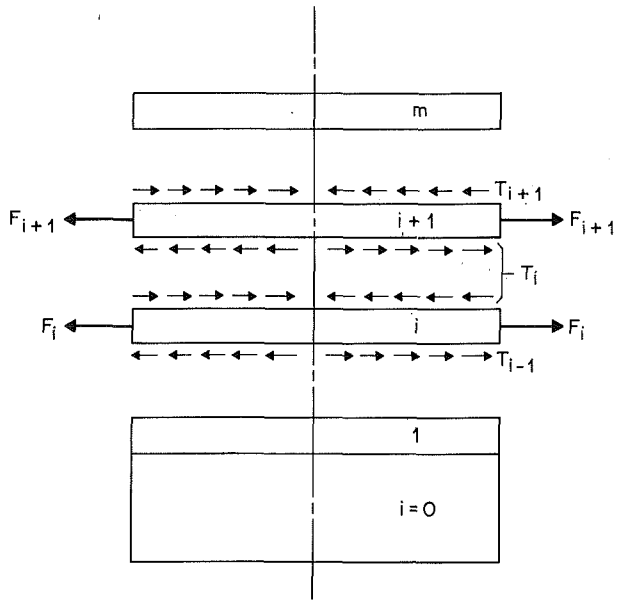


Fig. 3 Multilayered structure

$$\tau(x) = T\frac{d\chi_0(x)}{dx} = -T\chi_1(x), \quad N_0 = -\frac{1}{2}h_f\tau_{\max},$$

$$p(x) = \frac{1}{2}h_f T\chi_2(x). \tag{25}$$

This enables one to simplify the analysis for a multilayered structure and to avoid forming and solving a system of integral equations of the type (9). Therefore, the analysis below is based on the strain compatibility conditions rather than on the conditions of the compatibility of displacements. The longitudinal strains e_{i-1}^{+} and e_i^{-} of the upper extreme fiber of the $(i-1)$-*st* layer and the lower extreme fiber of the ith layer, respectively, can be expressed as follows:

$$e_{i-1}^{+} = \alpha_{i-1}\Delta t - \lambda_{i-1}F_{i-1} + \frac{h_{i-1}}{2\rho}, \quad e_i^{-} = \alpha_i\Delta t - \lambda_i F_i - \frac{h_i}{2\rho},$$
$$i = 0, 1, 2, \ldots, m. \tag{26}$$

Here α_i is the coefficient of thermal expansion for the ith layer material, $\lambda_i = (E_i^0 h_i)^{-1}$ is the axial compliance of this layer, F_i is the force acting in the ith layer, ρ is the radius of curvature, and Δt is the temperature differential. Using the condition $e_{i-1}^{+} = e_i^{-}$ of strain compatibility, we obtain:

$$\lambda_i F_i - \lambda_{i-1}F_{i-1} + \frac{h_i + h_{i-1}}{2\rho} = (\alpha_i - \alpha_{i-1})\Delta t,$$
$$i = 0, 1, 2, \ldots, m. \tag{27}$$

After summarizing these equations from $i = 0$ to i we have:

$$F_i = \frac{1}{\lambda_i}\left(\lambda_0 F_0 + \Delta\alpha_i\Delta t - \frac{a_i}{\rho}\right), \quad i = 0, 1, 2, \ldots, m, \tag{28}$$

where $\Delta\alpha_i = \alpha_i - \alpha_0$ is the thermal expansion mismatch between the ith layer and the substrate materials, and

$$a_i = \sum_{j=0}^{i} h_j - \frac{h_0 + h_i}{2}. \tag{29}$$

By summarizing all the equations (28) and considering the obvious equilibrium condition $\sum_{i=0}^{m} F_i = 0$, we find:

$$\lambda_0 F_0 = -s\Delta t + \frac{s_1}{\rho}, \tag{30}$$

where

$$s = \frac{\sum_{i=0}^{m}\Delta\alpha_i/\lambda_i}{\sum_{i=0}^{m}1/\lambda_i}, \quad s_1 = \frac{\sum_{i=0}^{m}a_i/\lambda_i}{\sum_{i=0}^{m}1/\lambda_i}. \tag{31}$$

For a multilayered structure on a thick substrate ($h_0 \gg \sum_{i=0}^{m} h_i$) the formula (29) yields: $a_i \cong h_0/2$. In this case $s_1 = h_0/2$ and $s \cong 0$. Then the equation (30) results in the formula: $\lambda_0 F_0 \cong h_0/2\rho$, and the equation (28) yields:

$$F_i = \frac{\Delta\alpha_i\Delta t}{\lambda_i} = E_i^0 h_i\Delta\alpha_i\Delta t, \quad i = 1, 2, \ldots, m. \tag{32}$$

The corresponding stresses in the film layers are as follows:

$$\sigma_i = \frac{F_i}{h_i} = E_i^0\Delta\alpha_i\Delta t, \quad i = 1, 2, \ldots, m. \tag{33}$$

The shearing force acting in the ith interface, i.e., in the interface between the ith and the $(i+1)$-*st* layers, is

$$T_i = \sum_{j=i}^{m} F_j = \Delta t\sum_{j=i}^{m} E_j^0 h_j\Delta\alpha_j, \quad i = 0, 1, 2, \ldots, m-1. \tag{34}$$

We present the k^2 value in the formulas (10) for the case of a multilayered structure in the following approximate way:

$$k^2 = \frac{\sum_{i=0}^{m}\lambda_i}{\sum_{i=0}^{m}\kappa_i} \cong \frac{3}{4}\frac{\sum_{i=0}^{m}\lambda_i}{\sum_{i=0}^{m}h_i^2\lambda_i}, \tag{35}$$

where the interfacial compliances κ_i of the layers can be assessed, assuming $\nu_1 = 1/3$, by the formula:

$$\kappa_i = \frac{2}{3}\frac{1+\nu_i}{1-\nu_i}\frac{h_i}{E_i^0} \cong \frac{4}{3}\frac{h_i}{E_i^0} = \frac{4}{3}h_i^2\lambda_i. \tag{36}$$

The distributed stresses in the ith interface can be calculated on the basis of the formulas (25) as follows:

$$\left.\begin{aligned} \tau_i(x) &= -T_i\chi_1(x) = -kT_i\frac{\sinh kx}{\cosh k\ell}, \\ p_i(x) &= \frac{1}{2}T_i\chi_2(x)\sum_{j=i+1}^{m}h_j \\ &= -\frac{1}{2}k^2T_i\frac{\cosh kx}{\cosh k\ell}\sum_{j=i+1}^{m}h_i. \end{aligned}\right\} \tag{37}$$

The maximum values of the interfacial forces occur at the end cross sections:

$$\left.\begin{aligned} \tau_{i,\max} &\cong -kT_i = -k\Delta t\sum_{j=i}^{m}E_j^0 h_j\Delta\alpha_j, \\ N_i &= -\frac{1}{2}\tau_{i,\max}\sum_{j=i+1}^{m}h_j \\ p_{i,\max} &= -\frac{1}{2}k^2T_i\sum_{j=i+1}^{m}h_j \\ &= -\frac{1}{2}k^2\Delta t\sum_{j=i+1}^{m}h_j\sum_{j=i}^{m}E_j^0 h_j\Delta\alpha_j, \\ i &= 0, 1, 2, \ldots, m-1 \end{aligned}\right\} \tag{38}$$

If the layers are applied at different temperatures, the above formulas for the $\tau_{i,\max}$ and $p_{i,\max}$ could be written as follows:

$$\left.\begin{aligned}
\tau_{i,\max} &= -k\sum_{j=i}^{m} E_j^0 h_j \Delta\alpha_j \Delta t_j = -k\sum_{j=i}^{m}\sigma_j h_j,\\
p_{i,\max} &= -\frac{1}{2}k^2\sum_{j=i+1}^{m} h_j \sum_{j=i}^{m} E_j^0 \mathrm{k}_j \Delta\alpha_j \Delta t_j\\
&= \frac{1}{2}k\tau_{i,\max}\sum_{j=i+1}^{m} h_j,\\
i &= 0, 1, 2, \ldots, m-1,
\end{aligned}\right\} \quad (39)$$

where Δt_i is the temperature change for the ith layer, i.e., the difference between the application temperature of the material and the given temperature. On the interface with the substrate ($i=0$) we have:

$$\left.\begin{aligned}
\tau_{0,\max} &= -k\sum_{i=0}^{m} E_i^0 h_i \Delta\alpha_i \Delta t_i,\\
N_0 &= -\frac{1}{2}\tau_{0,\max}\sum_{i=1}^{m} h_i,\\
p_{0,\max} &= \frac{1}{2}k\tau_{0,\max}\sum_{i=1}^{m} h_i.
\end{aligned}\right\} \quad (40)$$

The maximum bow can be calculated by the formula:

$$w_0 = \frac{3}{E_s^0}\left(\frac{\ell}{h_s}\right)^2 \sum_{i=1}^{m}\sigma_i h_i. \quad (41)$$

Numerical Example

The numerical example is executed for a hypothetical multilayered structure, where a 0.25 μm thick SiO_2 layer, 1.5 μm thick Al layer, and 5 μm thick polymide layer are applied on a 0.51 mm thick Si substrate. The stresses and maximum bow are evaluated for the room temperature conditions (t = 20°C). Calculations are performed in Table 1. The half length of the structure, affecting the maximum bow, is ℓ = 25.4 mm. The obtained results indicate that the thermally induced stresses are rather great and could possibly result in insufficient ultimate and fatigue strength.

Obviously, in order to be able to make a final conclusion regarding the cohesive and the adhesive strength of the materials, one has to know not only the actual stresses, but also the allowable (design) stress values. The latters are, unfortunately, not available for many materials utilized in multilayered structures. For this reason, until an appropriate strength data is available, our analysis, dealing only with the left part of the strength condition "actual stress ≤ design stress," can be used for the prediction of the stress level and could, of course, be utilized as a guidance for structural design of low stress multilayered structures.

Discussion

We now consider how the results of the above analysis might help our understanding of the thermally induced stresses in multilayered structures. In addition to normal stresses acting in the film layers themselves, there are shearing and transverse normal (peeling) stresses, acting in the interfaces. While the stresses in the layers themselves are responsible for the strength of the film materials, the interfacial stresses are responsible for blistering and peeling.

The normal stresses in the film layers are practically uniformly distributed along the film. Only at the very ends of the structure these stresses rapidly drop to zero. The interfacial stresses, on the contrary, concentrate near the edges at the distances on the same order of magnitude as the thickness of the multilayered structure. Since thin films do not experience bending stresses, the normal stresses in thin films are independent from the layer thicknesses. If, in addition, the substrate is thick, compared to the total thickness of the heteroepitaxial structure, then the normal stresses in the given film layer become independent from the mechanical properties of the other film layers and are characterized by the generalized Young modulus of the given layer, temperature differential, and the thermal expansion mismatch between this layer and the substrate materials. The corresponding normal stresses in the layers can be calculated by formula (33).

The maximum interfacial stresses arising in the ith layer of a multilayered heteroepitaxial structure, fabricated on a thick substrate, can be evaluated by formulas (37). These stresses increase with an increase in the normal stress level and in the thicknesses of the layers, located on the "free surface" of the structure. The interfacial stresses increase with an increase in the interfacial stiffness. This effect is accounted for by the factor k, which is affected by the stiffness of all the heteroepitaxial layers.

Thus, if there is a need to reduce the stress level, one should use materials with small Young moduli and small thermal expansion mismatch with the substrate. In the case of interfacial stresses, in addition to the above measure, smaller stresses

Table 1 Stresses in Multilayered Structures (Calculation Sheet)

COMPONENT NUMBER, i	MATERIAL	THICKNESS, h_i, m	GENERALIZED YOUNG MODULUS E_i^0, PA	COEFFICIENT OF THERMAL EXPANSION, α_i, 1/°C	TEMPERATURE DIFFERENTIAL, Δt_i, °C	STRESSES IN FILMS, σ_i, PA: FILM STRESS FACTOR $f_i = E_i^0\Delta\alpha_i$, PA/°C	STRESSES, $\sigma_i = f_i\Delta t_i$, PA	MAXIMUM SHEARING: STIFFNESS FACTOR k, 1/m: AXIAL COMPLIANCE $\lambda_i = (E_i^0 h_i)^{-1}$, m/N	INTERFACIAL COMPLIANCE $\aleph_i = 4/3\, h_i^2\lambda_i$, m^3/N	MAXIMUM SHEARING STRESSES, τ_i, PA: SHEARING STRESS FACTOR, $\sigma_i h_i$, N/m	$S_i^I = \sum_{i=1}^{i}\sigma_i h_i$, N/m	STRESSES, $\tau_i = kS_i^I$, PA	MAXIMUM PEELING STRESSES, p_i, PA: $S_i'' = \sum_{i=i+1}^{m}\sigma_i h_i$, m	STRESSES, $p_i = 1/2\, k\tau_i S_i''$, PA	PELLING FORCE N_0, N/m: $N_0 = -1/2\,\tau_i S_i''$, N/m
0	Si	0.508×10^{-3}	1.80×10^{7}	3.2×10^{-6}	0	0	0	1.094×10^{-4}	3.764×10^{-11}	0					
1	SiO_2	2.50×10^{-7}	1.60×10^{7}	0.9×10^{-6}	980	-36.81	-36074	0.250	0.208×10^{-13}	-0.0090	0.044	569	6.75×10^{-6}	248	0.0019
2	Al	1.50×10^{-6}	1.08×10^{7}	23.6×10^{-6}	0	219.54	0	0.0617	1.850×10^{-13}	0	0.0134	1733	6.50×10^{-6}	728	0.0056
3	POLY	5.00×10^{-6}	0.276×10^{6}	38.0×10^{-6}	280	9.60	2688	0.725	241.6×10^{-13}	0.0134	0.0134	1733	5.00×10^{-6}	560	0.0043
							Σ	1.0368	6.201×10^{-11}	0.0044	MAXIMUM BOW FACTOR $f_b = \frac{3}{E_0^0}\left(\frac{\ell}{h_0}\right)^2 = 4.167 \times 10^{-4}\ \frac{m^2}{N}$				
								$k = \sqrt{\frac{1.0368}{6.201 \times 10^{-11}}} = 1.293 \times 10^{5}$ 1/m			MAXIMUM BOW $w_0 = f_b\sum_{i=1}^{m}\sigma_i h_i = 0.183 \times 10^{-5}$ m			

Table 2 Summary of stress characteristics

STRESSES	NORMAL	SHEAR	PEELING
	IN THE FILMS	IN THE INTERFACES	
RESPONSIBLE FOR:	STRENGTH OF THE FILMS	BLISTERING AND PEELING (INTERFACIAL STRENGTH)	
DISTRIBUTION	UNIFORMLY DISTRIBUTED OVER THE FILM LENGTH	CONCENTRATE NEAR THE FILM ENDS AT DISTANCES ON THE ORDER OF THE FILM THICKNESS	
CALCULATION FORMULAS	$\sigma_j = E_j^0 \Delta\alpha_j \Delta t$ $E_i^0 = \frac{E_i}{1-\nu_i}$ $\Delta\alpha_i = \alpha_i - \alpha_s$	$\tau_i = k_i \sum_{j=i+1}^{m} h_j \sigma_j$ $k_i^2 = \frac{3}{4} \frac{\sum_{j=0}^{i+1} \lambda_j}{\sum_{j=0}^{i+1} h_j^2 \lambda_j}$	$P_j = \frac{1}{2} k_i \tau_j \sum_{j=i+1}^{m} h_j$ $N_i = -\frac{1}{2} \tau_i \sum_{j=i+1} h_j$ $\lambda_i = \frac{1}{E_i^0 h_i}$
DEPEND ON:	YOUNG MODULUS; THERMAL EXPANSION MISM. w/SUBSTRATE; TEMP. DIFFERENTIAL	NORMAL STRESSES AND THICKNESS OF THE GIVEN LAYER STRUCTURAL STIFFNESS (FACTOR k)	SHEAR STRESSES AND THICKNESSES OF THE GIVEN LAYER
CAN BE REDUCED BY:	USING SOFTER MATERIALS, HAVING SMALLER THERMAL MISMATCH WITH THE SUBSTRATE	REDUCING THE NORMAL STRESSES IN THE LAYERS ABOVE THE GIVEN LAYER, SLOPING THE ENDS OF THE LAYERS, REDUCING STRUCTURAL STIFFNESS (FACTOR k)	

could be obtained by a proper slanting of the layer ends and by increasing the structural compliance (resulting in smaller values of the stiffness factor k). The summary of the results is given in Table 2. Note, that the developed theory is equally applicable to heteroepitaxial lattice mismatched structures. In this case the thermal mismatch strain $\Delta\alpha\Delta t$ should be substituted by the lattice mismatch strain f (Luryi and Suhir, 1986).

Conclusion

An approximate engineering theory of stresses in multilayered thin film structure fabricated on a thick substrate is developed. The obtained formulas are simple, easy-to-use, and clearly indicate the role of the major factors affecting the stresses. The results of analysis can be utilized for guidance in the physical design of multilayered heteroepitaxial structures in microelectronics.

Acknowledgment

The author thanks H. J. Levinstein, W. J. Bertram, C. J. Bartlett, and J. M. Segelken for useful discussions, and L. B. Freund for indicating his unpublished work, where an approach, similar to ours, was taken.

References

Freund, L. B., 1981, "Shear Stress at a Film-Substrate Interface Due to a Differential Thermal Strain," Division of Engineering, Brown University, Providence, RI, unpublished.

Luryi, S., and Suhir, E., 1986, "New Approach to the High Quality Epitaxial Growth of Lattice-Mismatched Materials," *Applied Physics Letters*, Vol. 49, No. 3, pp. 140–142.

Olsen, G. H., and Ettenberg, M., 1977, "Calculated Stresses in Multilayered Heteroepitaxial Structures," *Journal of Applied Physics*, Vol. 48, No. 6, pp. 2543–2547.

Reinhart, F. K., and Logan, R. A., 1973, "Interface Stress of $Al_xGa_{1-x}As$-GaAs Layer Structures," *Journal of Applied Physics*, Vol. 44, No. 7, pp. 3171–3175.

Röll, K., 1976, "Analysis of Stress and Strain Distribution in Thin Films and Substrates," *Journal of Applied Physics*, Vol. 47, No. 7, pp. 3224–3229.

Stoney, G. G., 1909, "The Tension of Metallic Films Deposited by Electrolysis," *Royal Soc. Proc.*, A82, London, pp. 172–175.

Suhir, E., 1986, "Stresses in Bi-Metal Thermostats," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 657–660.

Suhir, E., 1986, "Stresses in Adhesively Bonded Assemblies Used in Electronic Packaging," *Mat. Res. Soc. Symp. Proc.*, Vol. 72, "Electronic Packaging Materials Science II," pp. 133–138.

Suhir, E., 1986, "Calculated Thermally Induced Stresses in Adhesively Bonded and Soldered Assemblies," *Int. Symp. on Microelectronics Proc.*, Atlanta, GA, Oct. 1986, pp. 383–392.

Suhir, E., 1986, "Calculated Stresses in Shear Joints," Presentation at ASME 1986 WAM, Boston, MA, Dec. 1986, unpublished.

Vilms, L., and Kerps, D., 1982, "Simple Stress Formula for Multilayered Thin Films on a Thick Substrate," *Journal of Applied Physics*, Vol. 53, No. 3, pp. 1536–1537.

Finite Deformation and Nonlinear Elastic Behavior of Flexible Composites

Shen-Yi Luo
Research Assistant.

Tsu-Wei Chou
Professor.

Center for Composite Materials and
Department of Mechanical Engineering,
University of Delaware,
Newark, DE 19716

The flexible composites discussed in this paper are composed of continuous fibers in an elastomeric matrix. The usable range of deformation of these composites is much larger than that of conventional rigid composites. Due to the material as well as geometric factors, the stress-strain relations for these composites are generally nonlinear under finite deformations. A constitutive model has been developed based upon the Eulerian description. The material nonlinear stress-strain relation is derived by using the stress energy density referring to the deformed volume. The stretching-shear coupling and the effects of the in-plane reorientation of fibers are also considered in the theoretical analysis. Comparisons are made between predictions of the present theory and experimental data for tirecord/rubber and Kevlar/silicone-elastomer flexible composite laminae; very good correlations have been found.

1 Introduction

The flexible composites discussed in this paper are composed of continuous fibers in an elastomeric matrix. The usable range of deformation of these composites is much larger than that of conventional rigid composites (Chou and Takahashi, 1986). In addition, elastomeric flexible composites are highly anisotropic (i.e., $E_{11}/E_{22} >> 1$) and hence, their effective elastic properties are very sensitive to the fiber orientation. These composites also show low shear modulus and hence large shear deformation, which allows fibers to change their orientation within the lamina. The geometric nonlinearity of the flexible composite is mainly caused by the reorientation of these fibers. The material nonlinearity is also pronounced in elastomeric composites under large deformation. The conventional theory based on the infinitesimal strain assumption for rigid matrix composites is no longer applicable, and a better understanding of the nonlinear behavior of these composites needs to be developed.

The general nonlinear theory of anisotropic material has been studied by Rivlin (1964) and Green and Adkins (1970). However, in treatment of the nonlinear problems, difficulties often rise in the analytical determination of the response function. In some of the existing composite analyses, the nonlinear stress-strain curves in the principal material directions are determined experimentally. A number of methods and various response functions have been adopted to represent these curves (Petit and Waddoups, 1969; Hahn and Tsai, 1973; Jones and Morgan, 1977; and Ishikawa and Chou, 1983). However, the geometric nonlinearity, which is important in the analysis of flexible composites is not emphasized in these studies. The influence of the geometric change of fibers has been summarized and studied by the theory of ideal fiber-reinforced material (Pipkin, 1974). There, the kinematic constraints are substituted for the material stress-strain relations by assuming that the fibers are inextensible. The theory of flexible composites, consisting of sinusoidally-shaped extensible fibers and ductile matrices under finite deformation, was first developed by Chou and Takahashi (1986). Their theoretical predictions of the overall stress-strain relation, based upon a step-wise incremental analysis and laminated plate theory, have considered the fiber geometric nonlinearity as well as the matrix material nonlinearity.

In this paper, the nonlinear elastic behavior of flexible composites is further examined. A different mathematical model of the constitutive relation is established based upon the Eulerian description to account for the material nonlinearity including stretching-shear coupling. The geometric nonlinear behavior, due to the reorientation of fibers, is analyzed through an iterative calculation. Comparisons are made between the present theory and experimental data for tirecord/rubber and Kevlar/silicone-elastomer flexible composite laminae.

2 Analysis

2.1 Eulerian Strains. Both the Lagrangian and Eulerian descriptions have been used in the theory of finite elasticity (for example, Fung, 1965). The strain tensor associated with the Lagrangian system (Lagrangian strain, E_{ij}) is known as Green's strain tensor. The strain tensor associated with the Eulerian system (Eulerian strain, e_{ij}) is known as Almansi's

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 23, 1986; final revision, July 8, 1987.

Paper No. 88-WA/APM-8.

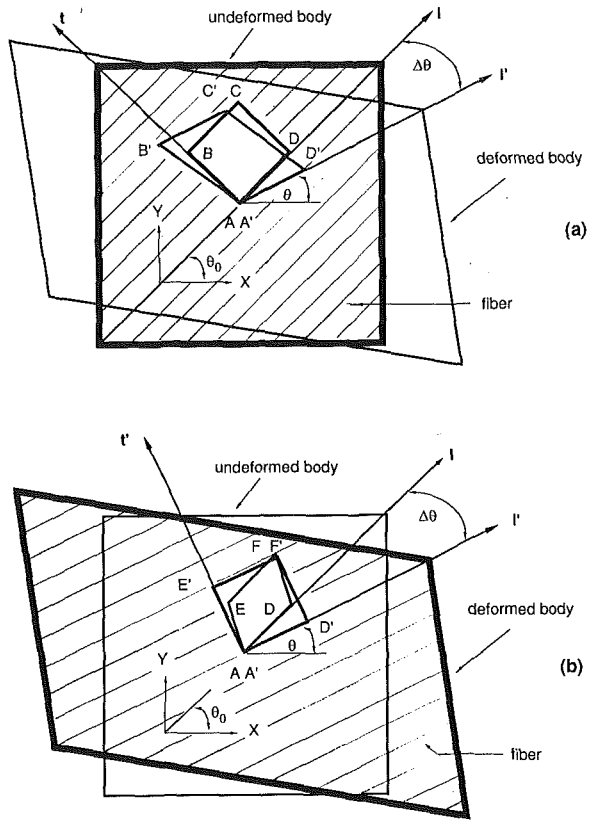


Fig. 1 A rectangular element of composite lamina before and after loading (a) in the Lagrangian system, (b) in the Eulerian system

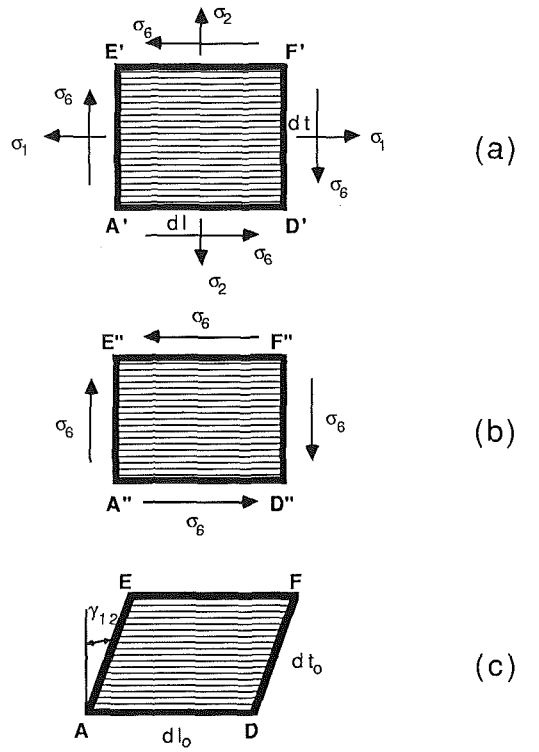


Fig. 2 Illustration of the deformation of a unidirectional rectangular element in the Eulerian system

strain tensor for large deformation and Cauchy's strain tensor for infinitesimal deformation. In spite of the popularity of Lagrangian description, in elasticity theory, the Eulerian strain tensor which refers to the deformed configuration is convenient for the present analysis.

In a rectangular Cartesian coordinate system x_i, the Eulerian strains are expressed as

$$e_{ij} = 1/2(u_{i,j} + u_{j,i} - u_{k,i}u_{k,j}) \tag{1}$$

where u_i denotes the displacement in the x_i direction. In the two-dimensional case,

$$\begin{aligned} e_{11} &= \partial u_1/\partial x_1 - 1/2[(\partial u_1/\partial x_1)^2 + (\partial u_2/\partial x_1)^2] \\ e_{22} &= \partial u_2/\partial x_2 - 1/2[(\partial u_1/\partial x_2)^2 + (\partial u_2/\partial x_2)^2] \\ e_{12} &= 1/2[(\partial u_1/\partial x_2) + (\partial u_2/\partial x_1) - (\partial u_1/\partial x_1)(\partial u_1/\partial x_2) \\ &\quad - (\partial u_2/\partial x_1)(\partial u_2/\partial x_2)] \end{aligned} \tag{2}$$

2.2 Finite Deformation of a Composite Lamina. For a composite lamina under finite deformation, the fiber orientation generally deviates significantly from its initial position. The Lagrangian description of such deformation is considered first. In Fig. 1(*a*), the initial fiber orientation is at an angle θ_0 with respect to the x axis. The Cartesian coordinates l–t coincide with the initial principal material directions; namely, l and t are along the fiber and transverse directions, respectively. Under loading, the rectangular element $ABCD$ is changed into a quadrilateral element $A'B'C'D'$. There is an angle $\Delta\theta$ between AD and $A'D'$. Corresponding to this change, the current fiber orientation l' is at an angle θ with respect to the x axis, and

$$\theta = \theta_0 + \Delta\theta. \tag{3}$$

The change of fiber orientation not only alters the geometric configuration of the lamina, but also the elastic properties with respect to the initial coordinate system. All these effects need to be considered for the prediction of the stress-strain relation. Thus, it is very cumbersome to investigate such changes based on the Lagrangian system.

In the following analysis, the Eulerian description is adopted, and the rectangular element $A'E'F'D'$ in the deformed body of Fig. 1(*b*) is considered. The sides $A'D'$ and $A'E'$ coincide with the Cartesian coordinates l'–t', with l' referring to the current fiber direction. Thus, the quadrilateral element $AEFD$ corresponds to the rectangular element $A'E'F'D'$ in the undeformed state. One may assume that the rectangle $A'E'F'D'$ undergoes two stages of deformation in restoring to its initial shape $AEFD$. These stages are illustrated in Fig. 2. First, $A'E'F'D'$ becomes a smaller rectangle $A''E''F''D''$ by removing the normal stresses; then it reverses to $AEFD$ by removing the shear stress. Obviously, γ_{12} is the angular deviation from a right-angle in the undeformed lamina between two intersecting line elements which coincide with the coordinates l'–t' in the deformed lamina.

The deformations depicted in Fig. 2 can be related to the Eulerian strain components. Let the line elements $AD = dl_0$ and $AE = dt_0$ in the undeformed lamina (Fig. 2(*c*)); also define $A'D' = dl$ and $A'E' = dt$ in the deformed lamina (Fig. 2(*a*)). Then, the physical significance of the Eulerian strains can be explained as

$$\begin{aligned} [(dl)^2 - (dl_0)^2]/(dl)^2 &= 2e_{11} \\ [(dt)^2 - (dt_0)^2]/(dt)^2 &= 2e_{22} \\ \sin\gamma_{12} &= 2e_{12}/(1-2e_{11})^{1/2}/(1-2e_{22})^{1/2} \end{aligned} \tag{4}$$

Let the axial "engineering strain" be defined as

$$\epsilon_1 = (dl - dl_0)/dl_0 \tag{5}$$

Then, it follows

$$\epsilon_1 = 1/(1-2e_{11})^{-1/2} - 1$$

and

$$e_{11} = [1 - (1+\epsilon_1)^{-2}]/2 \tag{6}$$

2.3 Constitutive Relations. We consider the stress-energy per unit area of the deformed lamina, and adopt the following polynomial expression for the plane-stress complementary energy function in the Eulerian system, in terms of σ_1, σ_2, and σ_6^2.

$$\begin{aligned} W^* = {} & (1/2)S_{11}\sigma_1^2 + (1/3)S_{111}\sigma_1^3 + (1/4)S_{1111}\sigma_1^4 \\ & + S_{12}\sigma_1\sigma_2 + (1/2)S_{22}\sigma_2^2 + (1/3)S_{222}\sigma_2^3 \\ & + (1/4)S_{2222}\sigma_2^4 + (1/2)S_{66}\sigma_6^2 + (1/4)S_{6666}\sigma_6^4 \\ & + S_{166}\sigma_1\sigma_6^2 + S_{2266}\sigma_2^2\sigma_6^2 \end{aligned} \tag{7}$$

where the subscripts 1 and 2 refer, respectively, to the l' and t' directions in Fig. 1(*b*). Also, the short-handed notations are used, namely, $\sigma_1 = \sigma_{11}$, $\sigma_2 = \sigma_{22}$, and $\sigma_6 = \sigma_{12}$. S_{ij}, S_{ijk} and S_{ijkl} are the compliance constants.

By using the relation

$$e_i = \partial W^*/\partial\sigma_i \tag{8}$$

the following constitutive equation for the rectangular element $A'E'F'D'$ (Fig. 2) in the Eulerian system are obtained.

$$\begin{aligned} e_1 &= S_{11}\sigma_1 + S_{111}\sigma_1^2 + S_{1111}\sigma_1^3 + S_{12}\sigma_2 + S_{166}\sigma_6^2 \\ e_2 &= S_{22}\sigma_2 + S_{222}\sigma_2^2 + S_{2222}\sigma_2^3 + S_{12}\sigma_1 + 2S_{2266}\sigma_2\sigma_6^2 \\ e_6 &= S_{66}\sigma_6 + S_{6666}\sigma_6^3 + 2S_{166}\sigma_1\sigma_6 + 2S_{2266}\sigma_2^2\sigma_6 \end{aligned} \tag{9}$$

Here $e_1 = e_{11}$, $e_2 = e_{22}$, and $e_6 = 2e_{12}$. The choice of compliance constants in equation (7) is made on the following basis. First, the terms S_{11}, S_{22}, S_{12}, and S_{66} are needed for the linear deformation. Second, the terms S_{111} and S_{222} are adopted for representing the bimodulus behavior in the axial and transverse directions, respectively. Third, the nonlinear terms are given by S_{1111}, S_{2222}, and S_{6666}. Lastly, the greatest uncertainty involves the coupling terms between the normal and the shear deformations. Unlike in rigid composites, the coupling effects may not be negligible in flexible composites. Two terms, S_{166} and S_{2266}, are retained to represent the interactions between axial and shear deformations in the present analysis.

Equation (9) is similar to Hahn and Tsai's (1973) expressions in its mathematical forms. However, there are some basic differences: (*a*) the strains on the left-hand side of equation (9) are Eulerian strains; (*b*) the stresses on the right-hand side of the equation act on a rectangular element in the deformed body; and (*c*) the coordinates l'-t' depend on the deformation, and the configuration of this element in the undeformed body is initially "unknown."

The compliance constants in equation (9) can be determined experimentally. The second-order constants (S_{11}, S_{22}, S_{12}, and S_{66}) are based on linear theory. The other constants are obtained by fitting the theoretical curves with experimental data. For example, when $\sigma_1 \neq 0$ and $\sigma_2 = \sigma_6 = 0$, equation (9) becomes

$$e_1 = S_{11}\sigma_1 + S_{111}\sigma_1^2 + S_{1111}\sigma_1^3 \tag{10}$$

The nonlinear $\sigma_1 - e_1$ curve can be obtained from simple experiments. Then, S_{11} is obtained from the initial slope of the experimental curve (i.e., $S_{11} = 1/\text{Young's modulus}$). S_{111} (which reflects bimodulus behavior) and S_{1111} are determined by fitting the theoretical curves to both tension and compression experimental data. The shear and stretching-shear coupling terms can be obtained from off-axis tensile tests with various initial fiber orientations, and equation (12) as derived below.

First, equation (9) can be written in the matrix form

$$\begin{Bmatrix} e_1 \\ e_2 \\ e_6 \end{Bmatrix} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{16} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{26} \\ \mathbf{S}_{61} & \mathbf{S}_{62} & \mathbf{S}_{66} \end{bmatrix} \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_6 \end{Bmatrix} \tag{11}$$

$$\begin{aligned} \mathbf{S}_{11} &= S_{11} + S_{111}\sigma_1 + S_{1111}\sigma_1^2 \\ \mathbf{S}_{22} &= S_{22} + S_{222}\sigma_2 + S_{2222}\sigma_2^2 \\ \mathbf{S}_{66} &= S_{66} + S_{6666}\sigma_6^2 \\ \mathbf{S}_{12} &= \mathbf{S}_{21} = S_{12} \\ \mathbf{S}_{16} &= S_{166}\sigma_6 \\ \mathbf{S}_{61} &= 2S_{166}\sigma_6 \\ \mathbf{S}_{62} &= \mathbf{S}_{26} = 2S_{2266}\sigma_2\sigma_6 \end{aligned}$$

The nonlinear Eulerian stress-strain relations in an arbitrary coordinate x–y can be obtained by using the transformation equation, where the fiber orientation angle, θ, is defined with respect to the x axis.

$$\{e\} = [T]^{-1}[\mathbf{S}][T]\{\sigma\} = [\mathbf{S}^*]\{\sigma\} \tag{12}$$

where

$$\{\mathbf{e}\} = \begin{Bmatrix} \mathbf{e}_{xx} \\ \mathbf{e}_{yy} \\ \mathbf{e}_{xy} \end{Bmatrix}, \quad \{\sigma\} = \begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{xy} \end{Bmatrix}$$

$$[T] = \begin{bmatrix} m^2 & n^2 & 2mn \\ n^2 & m^2 & -2mn \\ -mn & mn & m^2 - n^2 \end{bmatrix}$$

$m = \cos\theta$, $n = \sin\theta$

$$[\mathbf{s}^*] = \begin{bmatrix} \mathbf{S}_{11}^* & \mathbf{S}_{12}^* & \mathbf{S}_{16}^* \\ \mathbf{S}_{21}^* & \mathbf{S}_{22}^* & \mathbf{S}_{26}^* \\ \mathbf{S}_{61}^* & \mathbf{S}_{62}^* & \mathbf{S}_{66}^* \end{bmatrix} =$$

$$= \begin{bmatrix} m^4 S_{11} + 2m^2n^2S_{12} + n^4S_{22} + m^2n^2S_{66} & m^2n^2S_{11} + (m^4+n^4)S_{12} + m^2n^2S_{22} - m^2n^2S_{66} & 2m^3nS_{11} - 2mn(m^2-n^2)S_{12} - 2mn^2S_{22} - mn(m^2-n^2)S_{66} \\ m^2n^2S_{11} + (m^4+n^4)S_{12} + m^2n^2S_{22} - m^2n^2S_{66} & n^4S_{11} + 2m^2n^2S_{12} + m^4S_{22} + m^2n^2S_{66} & 2mn^3S_{11} + 2mn(m^2-n^2)S_{12} - 2m^3nS_{22} + mn(m^2-n^2)S_{66} \\ m^3nS_{11} - mn(m^2-n^2)S_{12} - mn^3S_{22} - mn(m^2-n^2)S_{66}/2 & mn^3S_{11} + mn(m^2-n^2)S_{12} - m^3nS_{22} + mn(m^2-n^2)S_{66}/2 & 2m^2n^2S_{11} - 4m^2n^2S_{12} + 2m^2n^2S_{22} + (m^2-n^2)S_{66}/2 \end{bmatrix}$$

$$+ (S_{111}\sigma_1 + S_{1111}\sigma_1^2) \begin{bmatrix} m^4 & m^2n^2 & 2m^3n \\ m^2n^2 & m^2n^2 & 2mn^3 \\ m^3n & mn^3 & 2m^2n^2 \end{bmatrix}$$

$$+(S_{222}\sigma_2+S_{2222}\sigma_2^2)\begin{bmatrix} n^4 & m^2n^2 & -2mn^3 \\ m^2n^2 & m^2n^2 & -2m^3n \\ -mn^3 & -m^3n & 2m^2n^2 \end{bmatrix}$$

$$+S_{6666}\sigma_6^2\begin{bmatrix} m^2n^2 & -m^2n^2 & -mn(m^2-n^2) \\ -m^2n^2 & -m^2n^2 & mn(m^2-n^2) \\ -mn(m^2-n^2)/2 & mn(m^2-n^2)/2 & (m^2-n^2)/2 \end{bmatrix}$$

$$+S_{166}\sigma_6\begin{bmatrix} -3m^3n & m^3n-2mn^3 & m^4-5m^2n^2 \\ 2m^3n-2mn^3 & 3mn^3 & -n^4+5m^2n^2 \\ m^4-2m^2n^2 & -n^4+2m^2n^2 & 3m^3n-3mn^3 \end{bmatrix}$$

$$+S_{2266}\sigma_2\sigma_6\begin{bmatrix} -4mn^3 & 2mn^3-2m^3n & -2n^4+6m^2n^2 \\ 2mn^3-2m^3n & 4m^3n & 2m^4-6m^2n^2 \\ -n^4+3m^2n^2 & m^4-3m^2n^2 & 4mn^3-4m^3n \end{bmatrix}$$

2.4 Reorientation of Fibers. The fiber reorientation due to finite deformation can be predicted as follows. First, the angle DAD' and EAE' are defined as α and β, respectively, in Fig. 1(*b*). Then,

$$\Delta\theta=(\alpha+\beta)/2+(\alpha-\beta)/2 \tag{13}$$

$(\alpha+\beta)/2$, the symmetric part of $\Delta\theta$, equals $\gamma_{12}/2$. From equation (4),

$$\gamma_{12}=\sin^{-1}[2e_{12}/(1-2e_{11})^{1/2}/(1-2e_{22})^{1/2}] \tag{14}$$

where, e_{11}, e_{22}, and e_{12} can be determined from equation (9). $(\alpha-\beta)/2$, the antisymmetric part of $\Delta\theta$, is defined as ω. It is understood that ω is the rigid body rotation, which is independent of the coordinate system but dependent on the boundary conditions. Then equation (13) can be rewritten as

$$\Delta\theta=\gamma_{12}/2+\omega \tag{15}$$

From equations (3), (11), (12), and (15), the current strains and fiber orientations can be determined by an iterative calculation technique. The procedure is shown in the flow chart of Fig. 3. The determination of ω will be demonstrated in Section 4 by two examples.

3 Experiments

3.1 Material and Fabrication. Experiments have been conducted to study the nonlinear elastic behavior of composites under finite strain. Two types of flexible composites are selected for the experimental work. One is the tirecord/rubber sheet widely used in the tire industry, which is supplied by the BFGoodrich Company. The other is Kevlar-49/silicone elastomer lamina, which is fabricated by the authors.

The fiber material used in the fabrication is Kevlar-49, 1420 denier (*1 denier = 1 gram/9000 meters*) with 1000 filaments in a yarn, manufactured by Du Pont Co. The resin is Sylgard 184 silicone elastomer, supplied by Dow Corning Co. The resin is transparent and has an elongation of up to 100 percent.

The unidirectional composite laminae are made by a frame-winding method (Parvizi, 1978). The yarns are carefully wound onto a metal frame (40 cm × 33 cm) with scales on the two opposite ends. The yarn alignment and uniformity are controlled by the number of yarns per unit length of the frame. After winding, the frame is first soaked in the Sylgard prime coat liquid (supplied by Dow Corning Co.) to achieve a better adhesion between fibers and matrix. Then, the frame is placed in a mold filled with resin matrix. Also, procedures are taken to remove the air bubbles and control the thickness of the specimens. The curing processes is performed in an autoclave.

The procedure of making flexible composite laminae with wavy fibers is basically the same as that for making unidirectional fiber laminae, except that the precoated fibers are pressed by a mold to form a predesigned shape. Then, these wavy fibers are aligned and embedded in the resin.

3.2 Tests and Results. All tensile tests are performed on an Instron tester at the cross-head speed of 5 mm/min. Longitudinal, transverse, and off-axis tensile tests are performed for unidirectional fiber specimens. There are at least

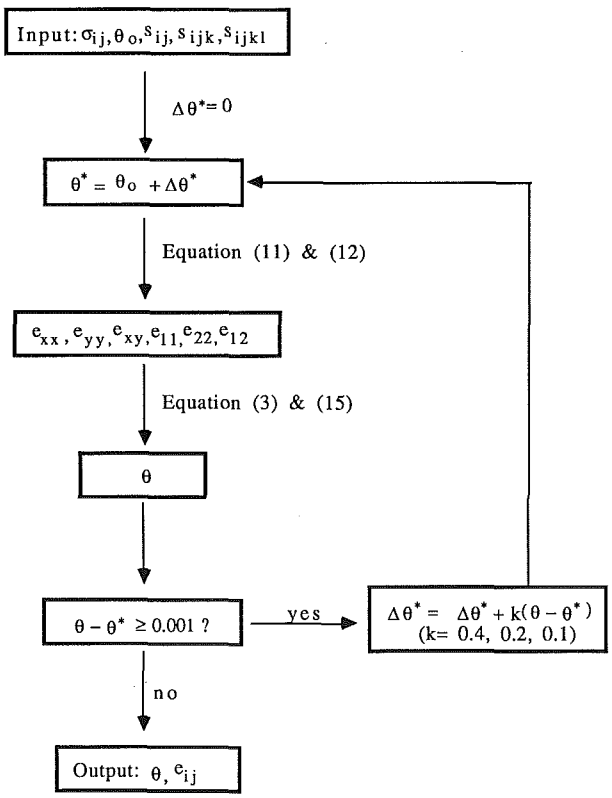


Fig. 3 Flow diagram of the calculation procedure

three specimens used in each test. The dimensions of the specimens are 2 mm × 20 mm × 280 mm. For wavy fiber specimens, only longitudinal properties are measured. The dimensions of the specimens are 2 mm × 30 mm × 250 mm.

Conventional strain-gauging technique is not suitable for measuring the large strains of soft materials in the present work. Consequently, extensometers with a single point edge knife are used for measuring the displacement of the unidirectional fiber specimens. For wavy fiber specimens, a photographic method is used. At each predetermined strain level, a photograph is taken to record the deformed configuration of the grid premarked on the specimen. Measurements of the locations of the marks on the specimen determine the displacement. In order to eliminate the influence of stress relaxation of the material, the specimen is kept at constant strain for 5 seconds before a photograph is taken (Rivlin, 1986). The measurement error of this method is believed to be small compared to the scale of deformation.

The experimentally determined elastic properties of tirecord/rubber and Kevlar/silicone-elastomer are listed in Table 1. Since only the tensile properties are discussed in this work, the compliance constants S_{111} and S_{222} are neglected in the curve-fitting function.

4 Numerical Examples

Two examples of flexible composites are considered: (1) unidirectional off-axis laminae and (2) a lamina with sinusoidally-shaped fibers. Uniaxial loading is applied (i.e., $\sigma_{xx} \neq 0$, $\sigma_{yy} = \sigma_{xy} = 0$). The theoretical calculations are based on the analysis of Sections 2.3 and 2.4. The total Eulerian strains and the current fiber orientation are calculated from equations (3), (11), (12), and (15) by an iterative trial-and-error technique. This method is different from the incremental method, with two advantages: (*a*) the local strains can be obtained with physical nonlinearities in tension, shear, and stretching-shear coupling, and (*b*) the question of validity of superposition, encountered in the incremental method for finite deformation problems, need not be considered here.

Table 1 Elastic properties of elastomeric composite laminae

	Tirecord/rubber	Kevlar-49/elastomer ($V_f = 9\%$)
$S_{11}(\text{MPa})^{-1}$	0.165*E*-3	0.114*E*-3
$S_{1111}(\text{MPa})^{-3}$	0	0
$S_{12}(\text{MPa})^{-1}$	−65.9*E*-6	−69.9*E*-6
$S_{22}(\text{MPa})^{-1}$	0.121	0.306
$S_{2222}(\text{MPa})^{-3}$	51.4*E*-3	0.563
$S_{66}(\text{MPa})^{-1}$	0.408	0.387
$S_{6666}(\text{MPa})^{-3}$	0.183	77.5*E*-3
$S_{166}(\text{MPa})^{-2}$	0.131*E*-3	3.43*E*-6
$S_{2266}(\text{MPa})^{-3}$	0.469	56.3*E*-3

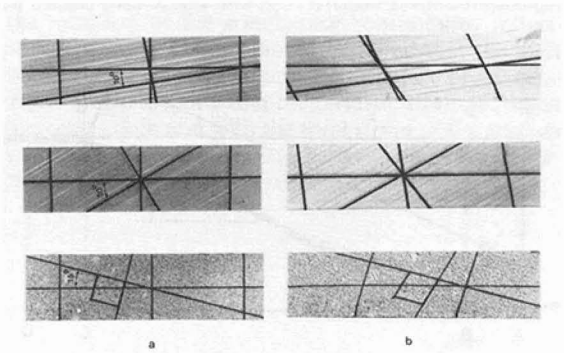


Fig. 4 Off-axis specimens of elastomeric composite laminae (*a*) without loading, (*b*) with loading. The 15 deg one is tirecord/rubber specimen, and the 10 deg and 30 deg ones are Kevlar/silicone elastomer specimens.

4.1 Unidirectional Off-axis Laminae. The elastic properties of the experimental materials are given in Table 1. Figures 4(*a*) and 4(*b*) are the photographs of the off-axis specimens without and with loading, respectively. Figure 4(*b*) shows that the central lines of specimens remain straight under loading. Thus, the rigid rotation in equation (15) can be written as

$$\omega = -\gamma_{xy}/2 \tag{16}$$

where γ_{xy} is the angular deviation from a right-angle in the undeformed lamina between two line elements which coincide with the coordinates x–y in the deformed lamina. From equation (4),

$$\gamma_{xy} = \sin^{-1}[2e_{xy}/(1-2e_{xx})^{1/2}/(1-2e_{yy})^{1/2}] \tag{17}$$

Figure 5 shows the comparisons between analytical predictions and experimental results on off-axis responses of tirecord/rubber composites. The fiber initial orientations are 15 deg, 30 deg, and 60 deg. The maximum strains are over 6 percent. The same comparisons for Kevlar-49/silicone elastomer specimens are shown in Fig. 6. The fiber initial orientations are 10 deg, 30 deg, and 60 deg in this case. Very good correlations have been found.

4.2 Flexible Composite Specimens With Wavy Fibers. Figure 7 illustrates the deformations of a flexible composite sample with wavy fibers under different levels of loading. Figure 8 is a photoelastic view of a flexible composite sample under longitudinal loading, which shows that relatively uniform strain is maintained in distinct regions along the x axis. It should be noted that although all the experimental data collected are based upon Kevlar-49/silicone elastomer system, the photographs shown in Figs. 7 and 8 are based upon graphite fiber as reinforcement materials, so better contrast between the fiber and matrix in the photographs can be achieved.

The deformation of this flexible composite is best understood by examining a representative element of the wavy fiber which contains a half-wave of the sinusoidal curve (Fig.

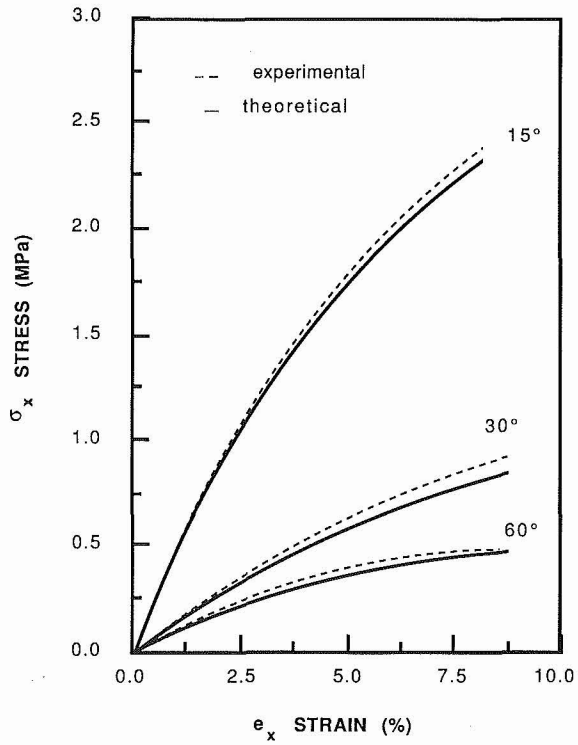


Fig. 5 Comparison between theoretical curves and experimental results on 15 deg, 30 deg, and 60 deg off-axis stress-strain responses of tirecord/rubber composite laminae

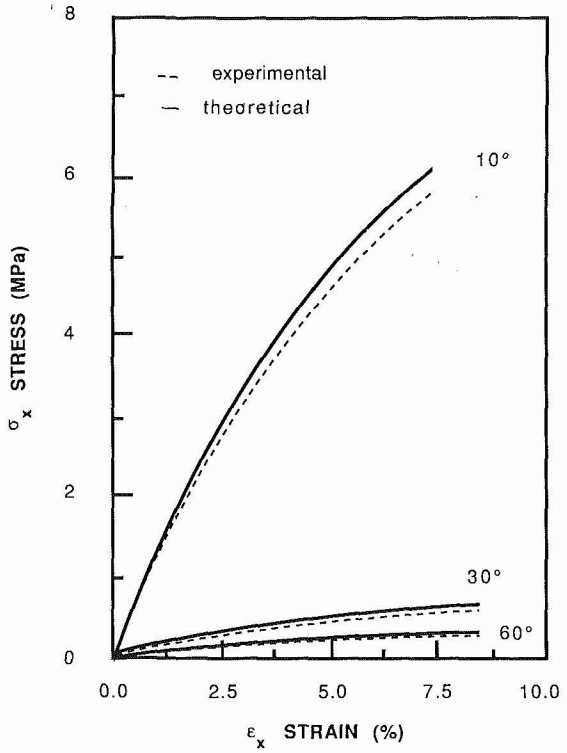


Fig. 6 Comparison between theoretical predictions and experimental results on 10 deg, 30 deg, and 60 deg off-axis stress-strain responses of Kevlar-49/silicone elastomer composite laminae

9). This element can be divided into subelements along the x axis. Then, its longitudinal deformation is the summation of the deformations of all of these subelements. For the purpose of the present calculation, the fiber element is divided into three regions along the x axis: (1) AB, (2) CD, and (3) BC. In the region AB and CD, the fiber segments are considered to be parallel to the x axis. Let $AB+CD=a$. In the region BC, the fiber is assumed to have an average inclination angle of θ_0 with respect to the x axis in the initial state; also $BC=b$. Then, the overall "engineering strain" of the specimen in the x direction is

$$\epsilon_x = \epsilon_a[a/(a+b)] + \epsilon_b[b/(a+b)], \tag{18}$$

Here, ϵ_a is the strain in the AB and CD parts, which is obtained from equations (12) and (6) by setting $\theta=0$. ϵ_b is the average strain of the region BC, which is treated as an off-axis lamina. Referring to equation (5), ϵ_b is defined as

$$\epsilon_b = (L\cos\theta - L_0\cos\theta_0)/(L_0\cos\theta_0)$$
$$= (1-2e_{11})^{-1/2}\cos\theta/\cos\theta_0 - 1 \tag{19}$$

where L and θ, respectively, are the instantaneous values of L_0 and θ_0 under loading.

In this wavy fiber composite, the central line of the specimen does not remain straight under loading, but the lines BB' and CC' remain perpendicular to the x axis as shown in Figs. 7 and 8. Thus,

$$\omega = +\gamma_{xy}/2 \tag{20}$$

Figure 10 shows the theoretical and experimental relation of longitudinal stress, σ_x, versus average strain, ϵ_x, of the flexible composite specimens. The initial fiber orientation angle θ_0 in the assumed off-axis part BC is in the range of 19 deg to 28 deg, as can be seen in Fig. 9. Thus, the average value $\theta_0=23.5$ deg has been assumed. The theoretical curves are obtained for $\theta_0=19$ deg, 23.5 deg, and 28 deg. The experimental strain values of ϵ_x are obtained by the photographic method.

Furthermore, the theoretical local Eulerian strain in the specimen can be obtained directly from equation (12). Figure 11 shows the relations between the applied stress σ_x and local Eulerian strains e_{xx}, e_{yy}, and e_{xy} in region BC of Fig. 9. The initial angle $\theta_0=23.5$ deg is used.

The experimental results of the variation of fiber orientation θ in region BC with applied stress σ_x also match very well with theoretical predictions. The fiber orientation angle decreases rapidly at low stress level and the specimen is broken before the fibers are totally straightened out. The maximum local tensile strain of the fibers occurs in the initially straight segments (near point A or D' in Fig. 9). Hence, the strength of the flexible composite may be determined by the maximum tensile strain in region AB or CD and the maximum shear strain in region BC of Fig. 9.

5 Concluding Remarks

A constitutive model for flexible composites has been

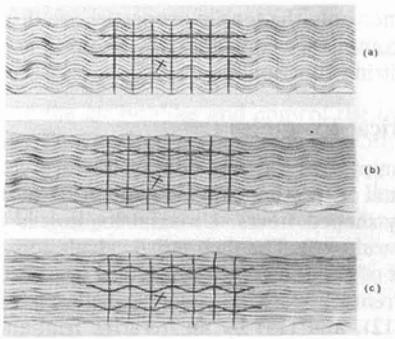


Fig. 7 Photographs of a flexible composite lamina sample (Graphite/silicone elastomer) with sinusoidally shaped fibers: (a) load = 0 kg; (b) load = 6 kg; (c) = 20 kg

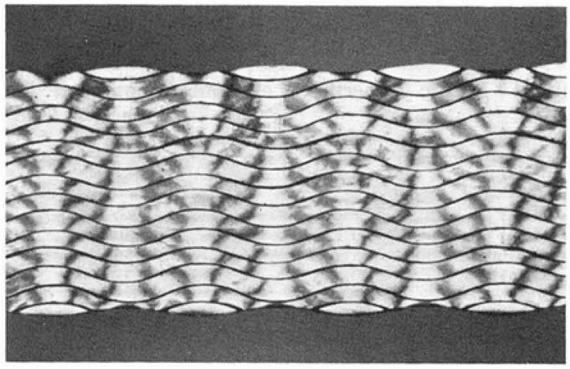


Fig. 8 A photoelastic view of a flexible composite lamina (Graphite/photostress material PL-2) with sinusoidally-shaped fibers

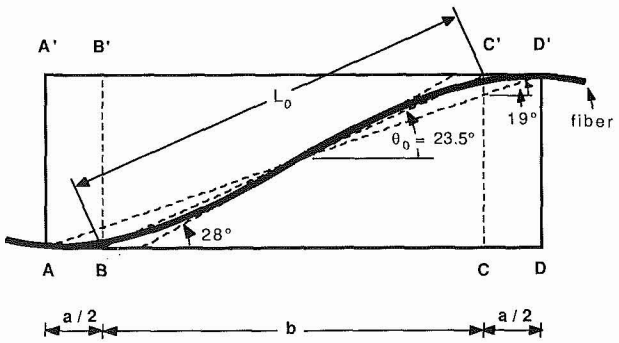


Fig. 9 The original shape of a wavy fiber in a flexible composite specimen

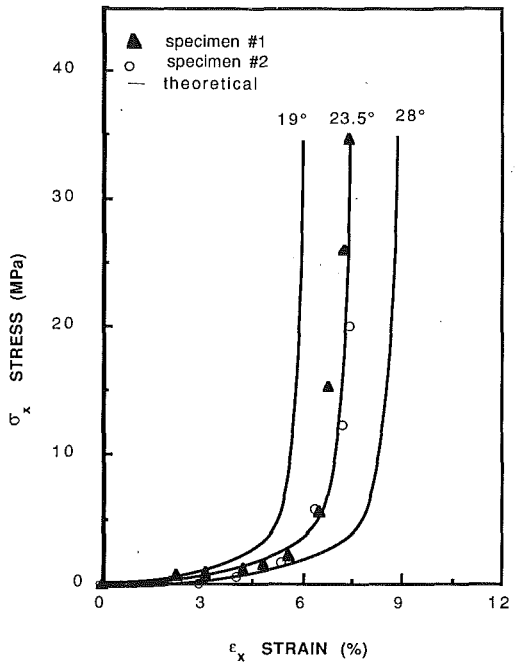


Fig. 10 Stress σ_x versus overall average strain, ϵ_x for a flexible composite specimen with wavy fibers

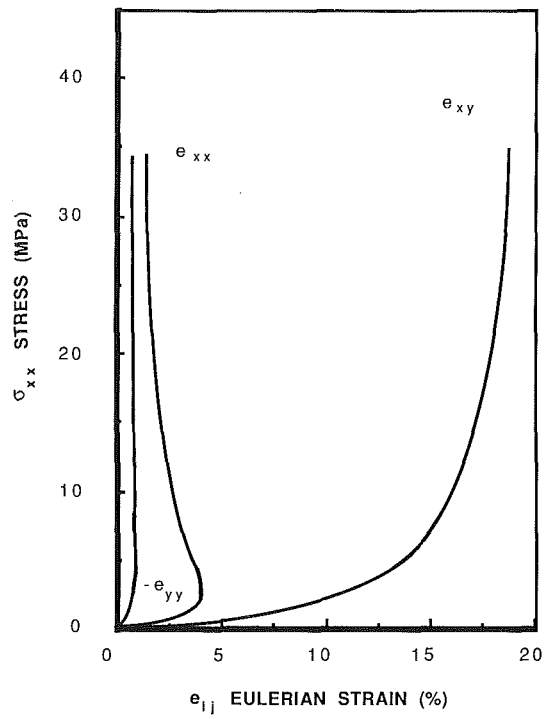


Fig. 11 Stress σ_x versus local Eulerian strain, ϵ_{ij} in x-y coordinates for a flexible composite specimen with wavy fibers

developed based upon the Eulerian description of deformation; consequently, it is not restricted to infinitesimal strains. The constitutive relation takes into account the material nonlinearities in tension, shear, and stretching-shear coupling. An iterative calculation procedure has been adopted to track the fiber reorientation. Good agreement has been found between the theoretical predictions and experimental results for unidirectional off-axis laminae and a lamina with sinusoidally-shaped fibers. The theoretical local Eulerian strains of the flexible composite are obtained directly from the constitutive equation. This information can be used for predicting the failure of flexible composites under finite deformation.

The mathematical form of the constitutive equation used in this work is based on the stress energy density function referring to the deformed volume. The use of strain energy density function referring the undeformed volume for problems of anisotropic material under large deformation is being investigated (Rivlin, 1986); the results will be reported later.

It should be noted that various combinations of a large number of coupling terms were compared in this work in curve fitting of experimental results. It appears that the use of S_{166} and S_{2266} can most adequately represent the experimental data. However, there is still a lack of sound analytical basis for the selection of the compliance constants to reflect the elastic interactions. It also should be noted that the inelastic behavior generally occurs in large deformation problems. Our experiments also indicate that the inelastic strain is very small (<5 percent) compared with the total strain of the specimens. Thus, for simplification, elastic behavior is assumed for the material systems used in this work.

Acknowledgment

We would like to thank Drs. R. Moghe and S. Parhizgar of the Research and Development Center, BFGoodrich Co., for supplying the tirecord/rubber sheets used in this work. This research was partially supported by the National Science Foundation through the Center for Composites Manufacturing Science and Engineering of the University of Delaware.

References

Chou, T. W., and Takahashi, K., 1986, "Nonlinear Elastic Behavior of Flexible Composites," *Composites*, Vol. 18, No. 1, pp. 25–34.

Fung, Y. C., 1965, *Foundations of Solid Mechanics*, Prentice-Hall, Englewood.

Green, A. E., and Adkins, J. E., 1970, *Large Elastic Deformations*, Oxford University Press, London.

Hahn, H. T., and Tsai, S. W., 1973, "Nonlinear Elastic Behavior of Unidirectional Composite Laminae," *J. Comp. Mat.*, Vol. 7, pp. 102–118.

Ishikawa, T., and Chou, T. W., 1983, "Nonlinear Behavior of Woven Fabric Composites," *J. Comp. Mat.*, Vol. 17, pp. 399–413.

Jones, R. S., and Morgan, H. S., 1977, "Analysis of Nonlinear Stress-Strain Behavior of Fiber-Reinforced Composite Materials," *AIAA J.*, Vol. 15, pp. 1669–1676.

Parvizi, Azar-Dokht, 1978, "The Transverse Cracking in Glass Fiber Reinforced Plastic Laminates," PhD Dissertation, University of Surrey.

Petit, P. H., and Waddoups, M. E., 1969, "A Method of Predicting the Nonlinear Behavior of Laminated Composites," *J. Comp. Mat.*, Vol. 3, pp. 2–19.

Pipkin, A. C., 1977, "Finite Deformation in Material Reinforced With Inextensible Cords," *Finite Elasticity*, R. S. Rivlin, ed., ASME, New York, pp. 91–102.

Rivlin, R. S., 1964, "Networks of Inextensible Cords," *Nonlinear Problems of Engineering*, Academic Press, New York.

Rivlin, R. S., 1986, Private Communications.

Nonlinear Composite Beam Theory

O. A. Bauchau
Assistant Professor.

C. H. Hong
Graduate Assistant.

Department of Mechanical Engineering,
Aeronautical Engineering and Mechanics,
Rensselaer Polytechnic Institute,
Troy, New York, 12180-3590

The modeling of naturally curved and twisted beams undergoing arbitrarily large displacements and rotations, but small strains, is a common problem in numerous engineering applications. This paper has three goals: (1) present a new formulation of this problem which includes transverse shearing deformations, torsional warping effects, and elastic couplings resulting from the use of composite materials, (2) show that the small strain assumption must be applied in a consistent fashion for composite beams, and (3) present some numerical results based on this new formulation to assess its accuracy, and to point out some distinguishing feature of anisotropic beam behavior. First, the predictions of the formulation will be compared with experimental results for the large deflections and rotations of an aluminum beam. Then, the distinguishing features of composite beams that are likely to impact the design of rotating blades (such as helicopter blades) will be discussed. A first type of extension-twisting coupling introduced by the warping behavior of a pretwisted beam is discussed, then, a shearing strain squared term, usually neglected in small strain analyses, is shown to introduce a coupling between axial extension and twisting behavior, that can be significant when the ratio E/G *is large* (E *and* G *are Young's and shearing moduli of the beam, respectively*). *Finally, the impact of inplane shearing modulus changes and torsional warping constraints on the behavior of beams exhibiting elastic couplings is investigated.*

1 Introduction

The nonlinear static and dynamic analysis of naturally curved and twisted beams has received considerable attention in recent years (Friedmann, 1977; Rosen, 1979; Rosen et al., 1987; Bauchau et al., 1987c), and has many important applications in aeronautical engineering such as helicopter blades and flexible space structures. The nonlinear kinematics involved in this problem are complex since both deformed and undeformed configurations of the beam are three-dimensional. Moreover, laminated composite materials are increasingly used for the construction of such structures, causing several nonclassical effects of beam theory to become more pronounced, such as transverse shearing deformations, torsion related warpings (Bauchau et al., 1985, 1987a), and elastic couplings.

The assumption of indeformability of the cross-section in its own plane is fundamental in beam theory as it reduces a three-dimensional problem to a one dimensional model along the axis of the beam. In the Euler-Bernoulli theory, cross sections are assumed to remain plane during deformation, however, with anisotropic materials, out-of-plane warping can become important, and torsion related warpings should be included in the analysis. In many applications, large deflections and rotations of the beam will occur, however the strain level remains low. For instance, in a helicopter blade fatigue life is a major concern, hence the operating strain level must remain well within the linear-elastic range of the material. As a result, most analyses (Rosen et al., 1979, 1987; Bauchau et al., 1987b, 1987c) are based on a small strain assumption that considerably simplifies the formulation and the resulting equations.

The small strain assumption has important implications. First, the Green-Lagrange strain components often used in the derivation of the nonlinear kinematics (Rosen et al., 1979, 1987) can be equated to the engineering strain components, and hence the usual stress-strain relationships of the material can be used. Second, the changes in surface area of a differential volume element due to deformation are negligible, and finally the strain-displacement equations can be considerably simplified since all second order terms (i.e, strain square terms) can be neglected.

The objective of this paper is to present a consistent, nonlinear analysis of naturally curved and twisted beams undergoing arbitrarily large deflections and rotations. Transverse shearing deformations, torsional warping effects, and elastic couplings resulting from the use of composite materials are all included in the formulation. Numerical results based on a finite element approximation of the resulting equations are presented. Special attention is devoted to the extension twisting behavior of composite beams which is important to the design of aeroelastically tailored composite helicopter blades.

2 Geometry of the Beam and Kinematics of the Deformation

Consider the naturally curved and twisted beam depicted in

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, April 1, 1987; final revision October 9, 1987.

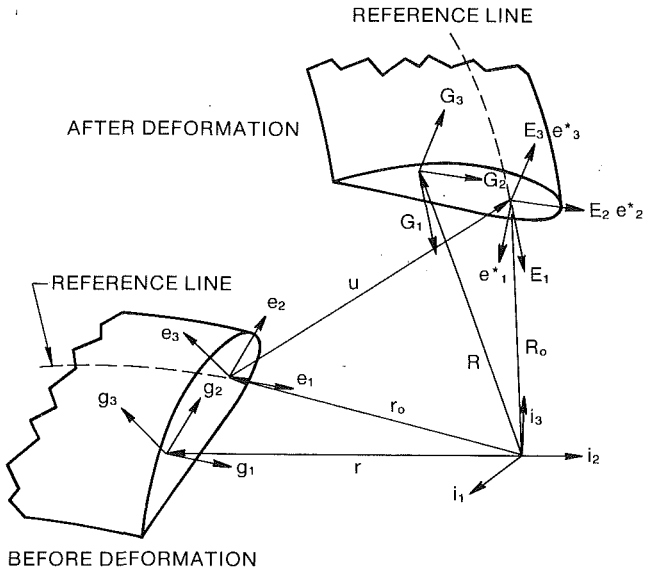


Fig. 1 Geometry of the beam before and after deformation

Fig. 1. The triad $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ is fixed in space and the triad $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is attached to a reference line along the axis of the beam. $\mathbf{e}_1$ is chosen tangent to the reference line and $\mathbf{e}_2$, $\mathbf{e}_3$ define the plane of the cross section. The curvilinear coordinates along this triad are x_1, x_2, and x_3, respectively. The position vector of a particle of the beam in the undeformed configuration is

$$\mathbf{r} = \mathbf{r}(x_1, x_2, x_3). \tag{1}$$

After deformation the same particle has a position vector

$$\mathbf{R} = \mathbf{R}(x_1, x_2, x_3) \tag{2}$$

The corresponding vectors at the reference line are

$$\mathbf{r}_0 = \mathbf{r}_0(x_1, 0, 0) \text{ and } \mathbf{R}_0 = \mathbf{R}_0(x_1, 0, 0), \tag{3}$$

and displacement vector of the reference line is given by

$$\mathbf{u} = \mathbf{R}_0 - \mathbf{r}_0. \tag{4}$$

The base vectors (Washizu, 1975; Wempner 1981) in the undeformed and deformed positions, respectively, are defined as

$$\mathbf{g}_i = \mathbf{r}_{,i} \text{ and } \mathbf{G}_i = \mathbf{R}_{,i}, \tag{5}$$

where the notation $(.)_{,i}$ means derivative with respect to x_i. At the reference line the base vectors are

$$\mathbf{e}_i = \mathbf{r}_{0,i} \text{ and } \mathbf{E}_i = \mathbf{R}_{0,i}. \tag{6}$$

$\mathbf{e}_i$ forms a triad since the derivatives in equation (6) are taken with respect to the natural coordinates of the beam. The triad $\mathbf{e}_i$ can be viewed as a rotation of the basic reference triad $\mathbf{i}_i$ through a given rotation matrix $t(x_1)$ such that:

$$\mathbf{e}_i = t^T(x_1)\mathbf{i}_i. \tag{7}$$

The derivatives of this triad are readily calculated as:

$$\begin{bmatrix} \mathbf{e}_1' \\ \mathbf{e}_2' \\ \mathbf{e}_3' \end{bmatrix} = \begin{bmatrix} 0 & k_3 & -k_2 \\ -k_3 & 0 & k_1 \\ k_2 & -k_1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix} \tag{8}$$

where $(.)'$ means derivative with respect to x_1; k_1 is the natural twist (or pretwist), k_2 and k_3 the natural curvature (or prebend) of the beam. The position vector of an arbitrary point of the cross section can be expressed as:

$$\mathbf{r} = \mathbf{r}_0 + x_2\,\mathbf{e}_2 + x_3\,\mathbf{e}_3, \tag{9}$$

hence the base vectors become:

$$\begin{aligned} \mathbf{g}_1 &= \sqrt{g}\mathbf{e}_1 - x_3 k_1\,\mathbf{e}_2 + x_2 k_1\,\mathbf{e}_3\,, \\ \mathbf{g}_2 &= \mathbf{e}_2\,, \\ \mathbf{g}_3 &= \mathbf{e}_3\,. \end{aligned} \tag{10}$$

The metric tensor is obtained as $\mathbf{g}_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$ and its determinant is g. The fundamental assumption in beam theory is that the cross section does not deform in its own plane. This means that the base vector $\mathbf{E}_2$ and $\mathbf{E}_3$ which are in the plane of the cross section after deformation simply correspond to a translation and rotation of the base vectors $\mathbf{e}_2$ and $\mathbf{e}_3$ of the original configuration. Note that arbitrarily large displacements and rotations can occur, but no deformation of the cross section is allowed, i.e., $\mathbf{E}_2$ and $\mathbf{E}_3$ are mutually orthogonal unit vectors. In contrast $\mathbf{E}_1$ is no longer unit nor orthogonal to $\mathbf{E}_2$ or $\mathbf{E}_3$, as axial and shearing strains are allowed. The triad $\mathbf{e}_i^*$ is defined as follows:

$$\mathbf{e}_2^* = \mathbf{E}_2, \mathbf{e}_3^* = \mathbf{E}_3, \mathbf{e}_1^* = \mathbf{e}_2^* \times \mathbf{e}_3^* \tag{11}$$

The vector $\mathbf{E}_1$ can be resolved in this triad as:

$$\mathbf{E}_1 = (1 + \bar{e}_{11})\mathbf{e}_1^* + 2\bar{e}_{12}\mathbf{e}_2^* + 2\bar{e}_{13}\mathbf{e}_3^* \tag{12}$$

At this point, $\bar{e}_{11}$, $\bar{e}_{12}$, and $\bar{e}_{13}$ are the unknown components of $\mathbf{E}_1$ along the triad $\mathbf{e}_i^*$, however they will be identified later as strain quantities. Here again the triad $\mathbf{e}_i^*$ can be related to the basic reference triad $\mathbf{i}_i$ through an unknown rotation matrix $T(x_1)$ such that:

$$\mathbf{e}_i^* = T^T(x_1)\mathbf{i}_i. \tag{13}$$

The derivatives of this triad are:

$$\begin{bmatrix} \mathbf{e}_1'^* \\ \mathbf{e}_2'^* \\ \mathbf{e}_3'^* \end{bmatrix} = \begin{bmatrix} 0 & K_3 & -K_2 \\ -K_3 & 0 & K_1 \\ K_2 & -K_1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{e}_1^* \\ \mathbf{e}_2^* \\ \mathbf{e}_3^* \end{bmatrix} \tag{14}$$

where K_1 is the twist, and K_2 and K_3 the curvatures of the deformed beam. Since the cross section does not deform in its own plane, the position vector in the deformed configuration can be written as:

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_0 + x_2\mathbf{E}_2 + x_3\mathbf{E}_3 \\ &\quad + [\delta(x_1)W_1(x_2, x_3) + 2\hat{e}_{12}W_2(x_2, x_3) + 2\hat{e}_{13}W_3(x_2, x_3)]\mathbf{e}_1^* \end{aligned} \tag{15}$$

The first three terms represent large translations and rotations of the cross section and can be interpreted geometrically as plane sections remaining plane, but not necessarily normal to the deformed axis of the beam. The last term represents a small displacement in the direction of $\mathbf{e}_1^*$, that is out-of-plane warping of the cross section. This warping has three components: one torsion related warping displacement W_1, and two bending related warpings W_2, and W_3. These warping displacements are selected as the St. Venant warping functions (5). $\delta(x_1)$ is an unknown function characterizing the magnitude of the torsional warping, whereas the magnitude of the bending related warpings are chosen proportional to the average transverse shearing strains $2\hat{e}_{12}$ and $2\hat{e}_{13}$. Combining equations (5), (6), (12), and (14) the base vectors of the deformed configuration become:

$$\begin{aligned} \mathbf{G}_1 &= [(1 + \bar{e}_{11}) - x_2K_3 + x_3K_2 + \delta' W_1]\mathbf{e}_1^* \\ &\quad + [2\bar{e}_{12} - x_3K_1]\mathbf{e}_2^* + [2\bar{e}_{13} + x_2K_1]\mathbf{e}_3^* \\ \mathbf{G}_2 &= [\delta W_{1,2} + 2\hat{e}_{12}W_{2,2} + 2\hat{e}_{13}W_{3,2}]\mathbf{e}_1^* + \mathbf{e}_2^* \\ \mathbf{G}_3 &= [\delta W_{1,3} + 2\hat{e}_{12}W_{2,3} + 2\hat{e}_{13}W_{3,3}]\mathbf{e}_1^* + \mathbf{e}_3^* \end{aligned} \tag{16}$$

In equations (16) all higher order terms containing warping quantities have been neglected, and differential bending related warpings are neglected as well (Bauchau et al., 1985, 1987). The Green-Lagrange strain components in a curvilinear coordinate system (Washizu, 1975) are given as

$f_{ij} = 1/2(G_{ij} - g_{ij})$ where G_{ij} is the metric tensor in the deformed configuration defined as $\mathbf{G}_{ij} = \mathbf{G}_i \cdot \mathbf{G}_j$. It is straightforward to verify that $f_{22} = f_{33} = f_{23} = 0$ as a direct implication of the indeformability of the cross section in its own plane. The other strain components are the transverse shearing strains $\sqrt{g}\, e_{12} = f_{12}$ and $\sqrt{g}\, e_{13} = f_{13}$, and the axial strain component $g e_{11} = f_{11} + 2x_3 k_1 f_{12} - 2x_2 k_1 f_{13}$, where e_{11}, e_{12}, and e_{13} are the components of the Green-Lagrange strain tensor in the local orthogonal coordinate system $\mathbf{e}_i$. The initial curvature of the beam k_2 and k_3 are now assumed to be small, hence $\sqrt{g} = 1 - x_2 k_3 + x_3 k_2 \cong 1$. This assumption is realistic in many practical applications and is satisfied for helicopter blades. The strain components now become:

$$2e_{12} = 2\bar{e}_{12} - x_3\kappa_1 + \delta W_{1,2} + 2\hat{e}_{12} W_{2,2} + 2\hat{e}_{13} W_{3,2} \tag{17}$$

$$2e_{13} = 2\bar{e}_{13} + x_2\kappa_1 + \delta W_{1,3} + 2\hat{e}_{12} W_{2,3} + 2\hat{e}_{13} W_{3,3} \tag{18}$$

$$\begin{aligned} e_{11} = {} & \bar{e}_{11} + 1/2\bar{e}_{11}^2 - x_2(1+\bar{e}_{11})\kappa_3 + x_3(1+\bar{e}_{11})\kappa_2 + \delta' W_1 \\ & + \delta k_1 (x_3 W_{1,2} - x_2 W_{1,3}) \\ & + 1/2(2\bar{e}_{12} - x_3\kappa_1)^2 + 1/2(2\bar{e}_{13} + x_2\kappa_1)^2 \\ & + 1/2(x_2\kappa_3 - x_3\kappa_2)^2 \end{aligned} \tag{19}$$

where $\kappa_i = K_i - k_i$. To complete the formulation, the coefficients $\bar{e}_{11}$, $\bar{e}_{12}$, and $\bar{e}_{13}$ first introduced in equation (12) must now be related to the displacements and rotations; differentiating equation (4) with respect to x_1 and using equations (6) we obtain:

$$\mathbf{E}_1 = \mathbf{e}_1 + \mathbf{u}' \tag{20}$$

or

$$\mathbf{E}_1 = (u_1' + t_{11})\mathbf{i}_1 + (u_2' + t_{21})\mathbf{i}_2 + (u_3' + t_{31})\mathbf{i}_3 \tag{21}$$

where u_i are the components of the displacement vector in the basic reference triad $\mathbf{i}_i$, and t_{ij} the components of the rotation matrix t. On the other hand, combining equations (12) and (13) yields another relation for $\mathbf{E}_1$ that can be identified with equations (21) to obtain:

$$\begin{bmatrix} 1+\bar{e}_{11} \\ 2\bar{e}_{12} \\ 2\bar{e}_{13} \end{bmatrix} = T^T \begin{bmatrix} u_1' + t_{11} \\ u_2' + t_{21} \\ u_3' + t_{31} \end{bmatrix} \tag{22}$$

This completes the strain analysis. It is important to note that this development is valid for arbitrarily large displacements, rotations, and strains. The unknowns of the problem are the three displacement components u_i, the three rotation parameters implicitly defined in the rotation matrix T, and the torsional warping amplitude.

3 Strain Energy in the Beam

The total potential energy for the present problem is:

$$\Pi = \frac{1}{2}\int_0^L \int_\Omega \sigma_{ij} e_{ij} \sqrt{g}\, d\Omega\, dx_1 - W \tag{23}$$

where σ_{ij} are the components of the stress tensor measured per unit surface of the undeformed configuration and e_{ij} are the components of the Green-Lagrange strain tensor, both measured in the local coordinate system $\mathbf{e}_i$ (Washizu, 1975); W the work done by the applied loads, Ω the cross-sectional area of the beam, and L its length. We now assume that the strain components remain small enough to render negligible the changes in surface area of a differential volume element due to deformation. This situation occurs in many practical applications, such as helicopter blades. Hence, the stress components σ_{ij} can be equated to the components of the Cauchy stress tensor in the local axis system $\mathbf{e}_i$, and the Green-Lagrange strain components e_{11}, $2e_{12}$, and $2e_{13}$ become equal to the engineering strain components ϵ_{11}, γ_{12}, and γ_{13}, respectively.

At this point, we will focus on thin-walled beams which are a realistic model for helicopter blades. This is not a restrictive assumption as the present theory could be applied to a solid cross section as well. For thin-walled sections the only nonvanishing stress components are the axial and shear stress flows, n and q, respectively; hence equation (23) now becomes:

$$\Pi = \frac{1}{2}\int_0^L \int_\Gamma (n\, e + q\, \gamma) ds\, dx_1 - W \tag{24}$$

where s is the curvilinear coordinate describing the thin-walled contour Γ, e is the axial strain, and γ the engineering shearing strain between a line element along the contour and one perpendicular to the cross section, given by:

$$\gamma = 2e_{12} x_2^+ + 2e_{13} x_3^+, \tag{25}$$

where $(.)^+$ denotes derivative with respect to s. The axial strain is given by equation (19), and the shearing strain is found by combining equations (17), (18), and (25) to obtain:

$$\gamma = 2\bar{e}_{12} G_2 + 2\bar{e}_{13} G_3 + \delta W_1^+ + \rho\, \kappa_1, \tag{26}$$

where $G_2 = x_2^+ + W_2^+$ and $G_3 = x_3^+ + W_3^+$ are the shearing strain distributions corresponding to unit transverse strains, and $\rho = (x_2 - x_{2k})x_3^+ - (x_3 - x_{3k})x_2^+ + x_{2k}G_3 + x_{3k}G_2$ where x_{2k} and x_{3k} are the coordinates of the shear center. Finally, the constitutive relationships for a thin-walled beam made of laminated composite are:

$$\begin{bmatrix} n \\ q \end{bmatrix} = \begin{bmatrix} A_{nn} & A_{nq} \\ A_{nq} & A_{qq} \end{bmatrix} \begin{bmatrix} e \\ \gamma \end{bmatrix} \tag{27}$$

where the stiffness coefficients A_{ij} are related to the inplane stiffness matrix of the laminate (Tsai et al., 1983). A nonvanishing A_{nq} corresponds to inplane extension-shearing coupling for the laminate and results in elastic couplings for the beam. The total potential energy finally becomes:

$$\Pi = \frac{1}{2}\int_0^L \int_\Gamma (A_{nn} e + A_{qq}\gamma + 2A_{nq} e\gamma) ds\, dz - W \tag{28}$$

This expression can be used as a basis for a finite element approximation of the problem.

4 The Strain-Displacement Expressions

In the derivation of the strain expressions (17)–(19), no assumptions were made about the magnitude of the displacements, rotations, or strains, hence, these expressions are valid for beams with small initial curvatures undergoing arbitrarily large displacements, rotations, and strains. On the other hand, in the derivation of the total potential energy expression (28), strain components were assumed small enough to render negligible changes in area due to deformation, and to equate Green-Lagrange to engineering strains. This requires both axial and shearing strains to be much smaller than unity, i.e., $e \ll 1$, and $\gamma \ll 1$; however, nothing is assumed about the relative magnitude of e versus γ. For consistency, the same assumptions must now be introduced into the strain-displacements equations (17)–(19), to obtain:

$$\begin{aligned} \epsilon_{11} = {} & \bar{e}_{11} - x_2\kappa_3 + x_3\kappa_2 + \delta' W_1 + \delta k_1 (x_3 W_{1,2} - x_2 W_{1,3}) \\ & + 1/2(2\bar{e}_{12} - x_3\kappa_1)^2 + 1/2(2\bar{e}_{13} + x_2\kappa_1)^2; \end{aligned} \tag{29}$$

$$\gamma_{12} = 2\bar{e}_{12} - x_3\kappa_1 + \delta W_{1,2} + 2\hat{e}_{12} W_{2,2} + 2\hat{e}_{13} W_{3,2}; \tag{30}$$

$$\gamma_{13} = 2\bar{e}_{13} + x_2\kappa_1 + \delta W_{1,3} + 2\hat{e}_{12} W_{2,3} + 2\hat{e}_{13} W_{3,3}. \tag{31}$$

The last term appearing in equation (19) is negligible since it represents the square of the axial strain components due to bending. If we now introduce the additional assumption that axial and shearing strains are of the same order of magnitude,

then $\gamma^2 \ll e$, and the two last terms of equation (29) can be neglected, since they are square of the shearing strain component (30) and (31), respectively; this yields:

$$\begin{aligned} \epsilon_{11} &= \bar{e}_{11} - x_2\kappa_3 + x_3\kappa_2 \quad + \delta' W_1 + \delta k_1 (x_3 W_{1,2} - x_2 W_{1,3}) \\ \gamma_{12} &= 2\bar{e}_{12} - x_3\kappa_1 \quad + \delta W_{1,2} + 2\hat{e}_{12} W_{2,2} + 2\hat{e}_{13} W_{3,2}; \\ \gamma_{13} &= 2\bar{e}_{13} + x_2\kappa_1 \quad + \delta W_{1,3} + 2\hat{e}_{12} W_{2,3} + 2\hat{e}_{13} W_{3,3} \end{aligned} \tag{32}$$

These expressions are often successfully used as the basis for beam models involving large displacements and rotations, but small strains (Rosen et al., 1979, 1987; Bauchau et al., 1987b, 1987c; Washizu, 1964). However, it is interesting to note that one additional assumption was required ($\gamma^2 \ll e$), that might not be adequate when dealing with highly anisotropic composite materials.

The seven unknown functions of the problem are the three displacements, the three rotation parameters implicitly defined in the rotation matrix T (in this effort the Euler angles were selected), and the magnitude of the torsional warping δ. These functions can be discretized using standard finite element techniques, and in this study, four noded elements (cubic interpolation functions) were used. The resulting expression for the total potential energy is strongly nonlinear since it depends on powers of the displacement functions, and products of trigonometric functions of the Euler angles. A Newton-Ralphson technique was used to iteratively solve the resulting finite element equations.

5 Numerical Examples

The first example was selected to demonstrate the ability of the program to deal with arbitrarily large displacements and rotations; the finite element results will be compared with experimental measurements obtained by Dowell and Traybar (1975) for a slender, cantilevered beam with a solid aluminum rectangular cross section subjected to a tip load P. The angle Φ_0 between the loading direction and the major axis of the cross section (the "loading angle") can be varied from 0 to 90 deg in the experimental set-up, yielding a wide range of nonlinear problems where torsion and bending are coupled. Figures 2(*a*), (*b*), and (*c*) show the tip twist angle, the flatwise tip displacement (i.e., the displacement perpendicular to the major axis of the cross section), and the edgewise tip displacement (parallel to the major axis), respectively, versus the loading angle, for three values of the tip load ($P = 1$, 2, and 3 lbs). In all cases excellent correlation is found between the finite element predictions, a theoretical prediction by Rosen et al. (1987), and the experimental results. The strain fields (29)–(31) or (32) gave identical predictions.

The following examples all deal with the extension-twisting behavior of composite beams. This problem is important for helicopter blade analysis since large axial centrifugal loads are present in the blade, and small changes in torsional stiffness result in angle of attack variations that could in turn result in large changes in the blade's aerodynamic loading. Moreover, any coupling between extension and twisting could have a stabilizing or destabilizing effect on the blade's behavior. This is especially important in the design of a flexbeam, that structure connecting the blade's root to the rotor hub. The flex-

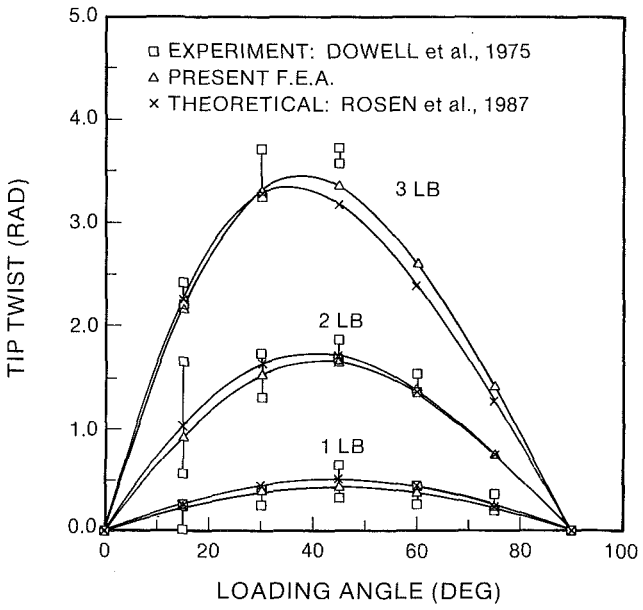


Fig. 2(*a*) Tip twist angle

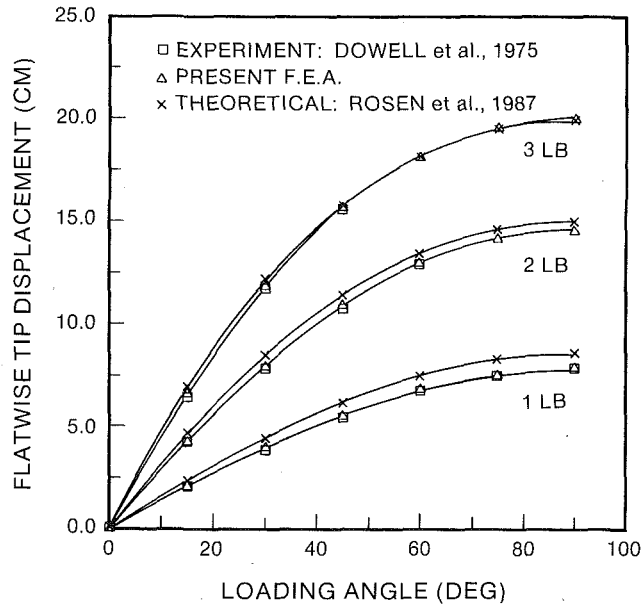


Fig. 2(*b*) Flatwise tip displacement

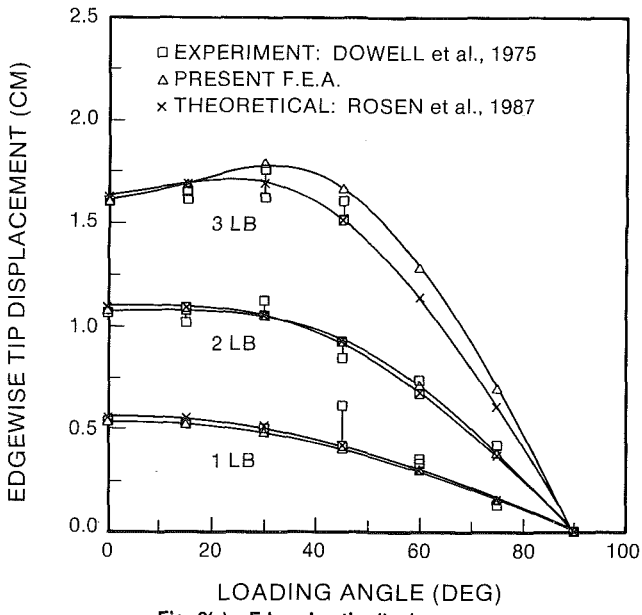


Fig. 2(*c*) Edgewise tip displacement

Fig. 2 Tip displacements of the beam versus loading angle

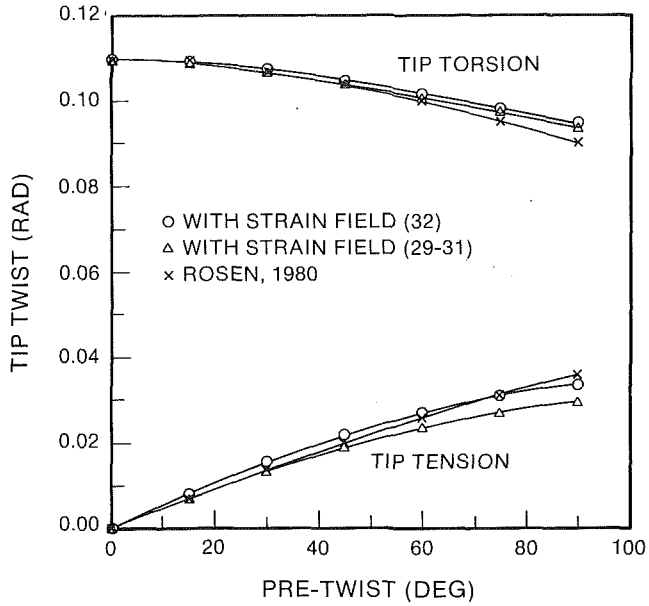


Fig. 3 Tip twist of a thin-walled box beam versus pretwisting angle

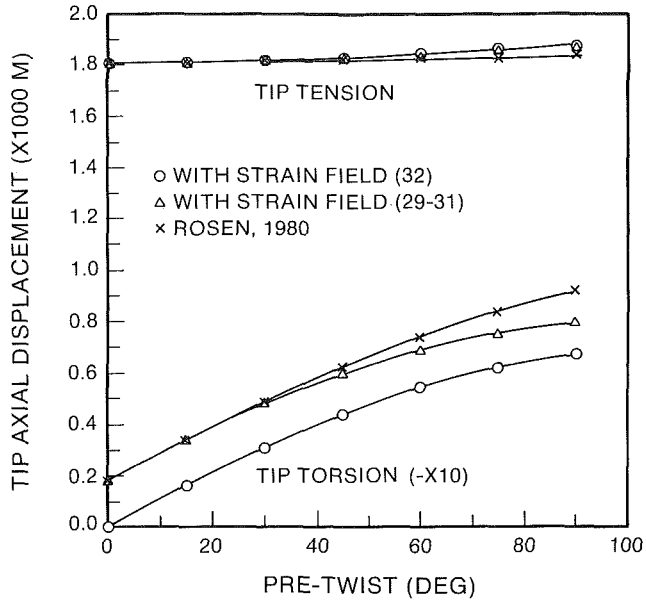


Fig. 4 Tip displacement of a thin-walled box beam versus pretwisting angle

beam must have a very high axial strength (to carry the centrifugal load) together with very low torsional stiffness (to allow easy control of the blade's pitch). This results in a highly anisotropic structure prone to extension twisting coupling, and extremely sensitive to such coupling.

Extension-twisting coupling can originate from three sources. First, naturally twisted beams will tend to untwist under an axial load; this effect is purely linear and results from differential torsional warping (Rosen, 1980), hence the importance of including this warping effect in the analysis. Second, a straight beam will tend to stiffen under an applied axial load. This effect is independant of warping and will occur even for a circular cross section which exhibits no torsional warping. It is a nonlinear effect due to the square shearing strain term discussed in equations (32); this term is generally small, however it can become significant for highly anisotropic beams. Finally, extension-twisting coupling can result from elastically tailoring a composite lay-up to that effect. This is a linear effect originating from a nonvanishing A_{nq} (27).

To demonstrate the first effect, the torsional behavior of a naturally twisted rectangular box beam under tip torque and axial load is analyzed. The cross-sectional width and height are 127.0 and 25.4 mm, respectively, the beam's length is 0.635 m. The thin walls consist of a laminated Graphite/Epoxy composite with a $[0_2^\circ, \pm 15^\circ]_s$ lay-up corresponding to stiffness coefficients $A_{nn} = 115.0$ MN/m, $A_{qq} = 8.5$ MN/m, and $A_{nq} = 0$, and a wall thickness of 1 mm. The upper curve of Fig. 3 shows the tip twist under a fixed tip torque as a function of the initial pretwist angle, and the lower curve shows the tip twist under a fixed axial load. Figure 4 shows the tip displacement under the same conditions. Note the torsional stiffening under tip torque, and the "untwisting" under tip tension, which are linear effects. The choice of the strain field (29)–(31) versus equations (32) has a minimum effect on the major displacement component (i.e., displacement under tip load or twist under tip torque), but a more significant effect on the coupled displacement components (twist under tip load and displacement under tip torque). The results are in good agreement with theoretical predictions by Rosen (1980). Note that Rosen also included the shearing strain square term in his analysis (i.e., strain field (29)–(31)). The discrepancy between the present analysis and Rosen's results at large pretwist angles comes from the fact that in the latter analysis the amplitude of the torsional warping displacement is chosen proportional to the twist rate, but is an independant

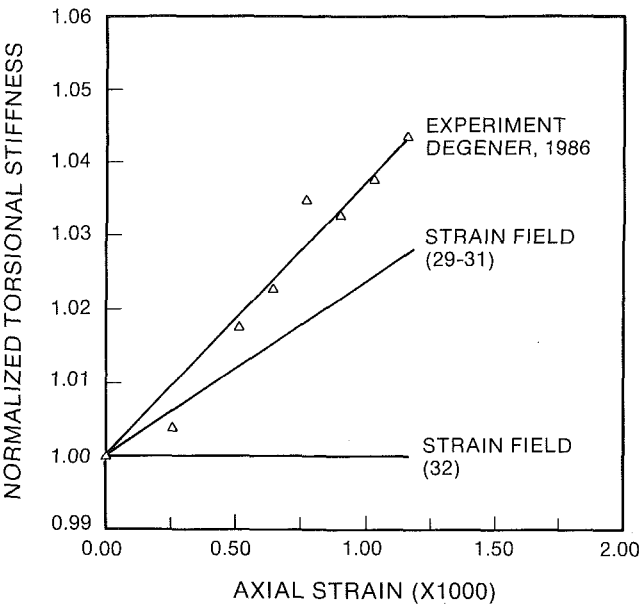


Fig. 5 Normalized torsional stiffness of a Kevlar Beam under axial tension

quantity in the present analysis resulting in a greater flexibility of the beam.

Consider next a thin-walled Kevlar beam ($L = 0.6$ m) with a circular cross section (outer diameter is 6 mm, and the wall thickness is 0.125 mm); the lay-up consists of unidirectional fibers aligned with the axis of the beam, and the stiffness coefficients are $A_{nn} = 8.25$ MN/m, $A_{qq} = 0.35$ MN/m, and $A_{nq} = 0$. In an experimental study (Degener, 1986) the torsional stiffness of this beam was measured for various levels of applied axial strain. Figure 5 shows the measured torsional stiffness (normalized by the initial torsional stiffness) versus the axial strain. A model of this problem based on the strain field (32) predicts a constant torsional stiffness, in contrast with the observed experimental results. For this simple case, the strain field (32) reduces to:

$$\epsilon_{11} = \bar{e}_{11} \text{ and } \gamma = \rho\kappa_1, \tag{29}$$

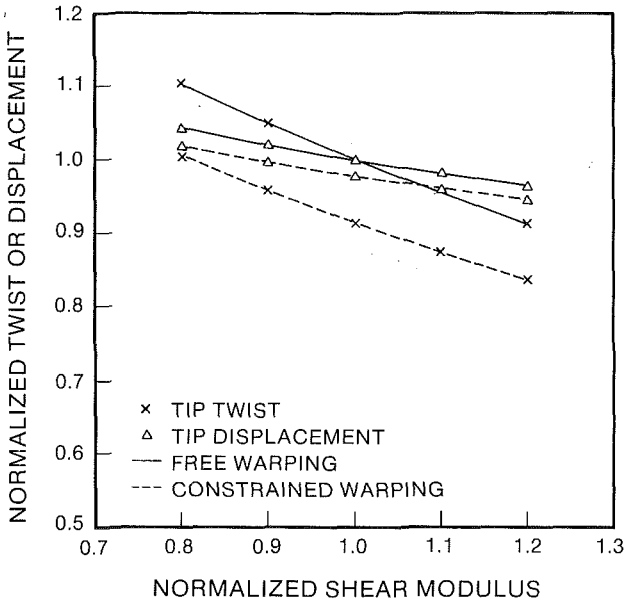


Fig. 6 Normalized tip twist and displacement of a rectangular box beam with extension-twisting couplings, under tip tension

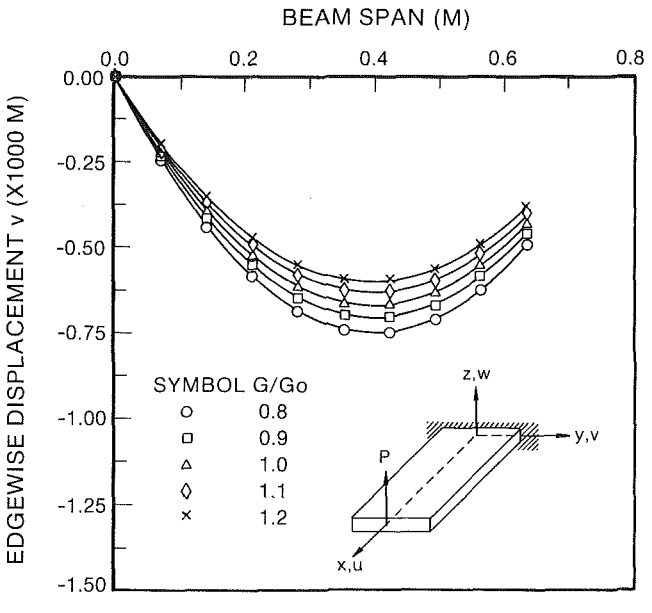


Fig. 7 Edgewise displacement of a rectangular box beam with extension-twisting couplings, under tip transverse load

which clearly show that axial and shearing strains are uncoupled. On the other hand, the strain field (29)–(31) gives:

$$\epsilon_{11} = \bar{e}_{11} + \frac{1}{2}\rho^2\kappa_1^2 \text{ and } \gamma = \rho\kappa_1, \tag{30}$$

which now exhibits a kinematic coupling term between tension and torsion. The strain energy then becomes:

$$U = \frac{1}{2}\int_0^L \int_\Gamma [A_{nn}(\bar{e}_{11} + \frac{1}{2}\rho^2\kappa_1^2)^2 + A_{qq}\rho^2\kappa_1^2]ds\,dx \tag{31}$$

and neglecting the term in κ_1^4 the torsional stiffness GJ becomes:

$$GJ = A_{qq}I_p(1 + e_{11}A_{nn}/A_{qq}). \tag{32}$$

Where I_p is the polar moment of inertia of the cross section. In this expression the dependency of the torsional stiffness on the axial strain level clearly appears. For most isotropic and homogeneous engineering materials Poisson's ratio is about 0.3, hence the ratio $A_{nn}/A_{qq} = 2.6$, and GJ remains virtually constant for small axial strain levels. However, for anisotropic materials this ratio could be much larger (about 24 for unidirectional Kevlar) resulting in a significant increase in torsional stiffness.

Figure 5 shows that the predictions based on the strain field (29)–(31) are in closer agreement with the experimentally measured torsional stiffnesses which increased of about 4 percent for an axial strain of 1000 μm/m. In a typical flexbeam, the axial strain level would be significantly higher (about 5000 μm/m) resulting in drastic torsional stiffness changes. It is also important to note that the additional strain term introduces a kinematic coupling between extension and twisting that is inexistent when this term is neglected. Such a coupling could potentially affect the stability of the blade. The observed discrepancy between the experimental results and the present calculations might be due to simplified material constitutive relationship used in this model (however, Kevlar exhibits a linear stress-strain curve in the fiber direction for $e \leq 1000$ μm/m, and the shearing strain was kept below 50 μm/m during the experiment).

The last example involves a composite beam exhibiting elastic couplings. The geometry is once again a thin-walled rectangular box section of dimensions given above. The web has the following stiffness parameters: $A_{nn} = 90.3$ MN/m, $A_{qq} = 23.3$ MN/m, and $A_{nq} = 0$, corresponding to a $[\pm 45\ \text{deg}]_s$ lay-up of a T300/5208 Graphite/Epoxy composite (0 deg corresponds to fibers aligned with the axis of the beam). Two different flange configurations were considered: in the first case both upper and lower flanges consist of a $[15_2, 0, 30]_s$ lay-up, resulting in extension-twisting coupling (stiffness parameters are $A_{nn} = 119.4$ MN/m, $A_{qq} = 14.1$ MN/m, and $A_{nq} = 22.5$ MN/m). In the second case the lay-ups are $[15_2, 0, 30]_s$ and $[-15_2, 0, -30]_s$, for the upper and lower flanges, respectively, resulting in bending-twisting coupling.

This beam will be used as a rough model of an elastically tailored helicopter blade; two important features of such structures will be illustrated. The first feature is the marked dependency of their bending and torsional behavior on the ply shearing modulus G_{LT} (Shyprykevitch, 1982). Whenever possible, a designer will select the composite lay-up that optimizes strength or stiffness by taking advantage of the high strength and stiffness in the fiber direction, hence minimizing the effect of ply shearing and transverse properties on overall laminate behavior. However, elastic coupling in a thin-walled beam is obtained through extension-shearing coupling of the laminate which is itself strongly dependent on G_{LT}, hence, it is not possible to obtain elastic couplings without a strong dependency on this modulus. To illustrate this point the behavior of coupled beams will be presented as a function of normalized ply shearing modulus, where the normalization factor is chosen as the nominal G_{LT}. The second feature is the more pronounced dependency of the overall behavior on nonclassical effects of beam theory such as torsional warping constraints.

Figure 6 shows the normalized tip twist the displacement under tip axial load of 100 KN for the beam with extension-twisting couplings (this would model the effect of centrifugal loads on a helicopter blade). Two types of boundary conditions are considered at the root: the torsional warping is either free or constrained. As expected, both normalized G_{LT} and torsional warping significantly affected the torsional behavior of the beam, however, the same factors also affect the extensional behavior, though less significantly.

Another important point is that this lay-up, designed to exhibit extension-twisting coupling, will also present several other types of couplings; consider the same beam under a tip-

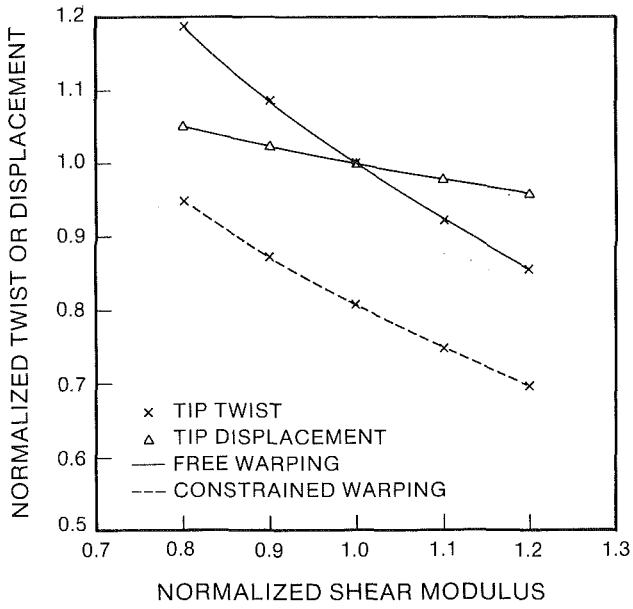


Fig. 8 Normalized tip twist and displacement of a rectangular box beam with extension-twisting couplings, under tip transverse load

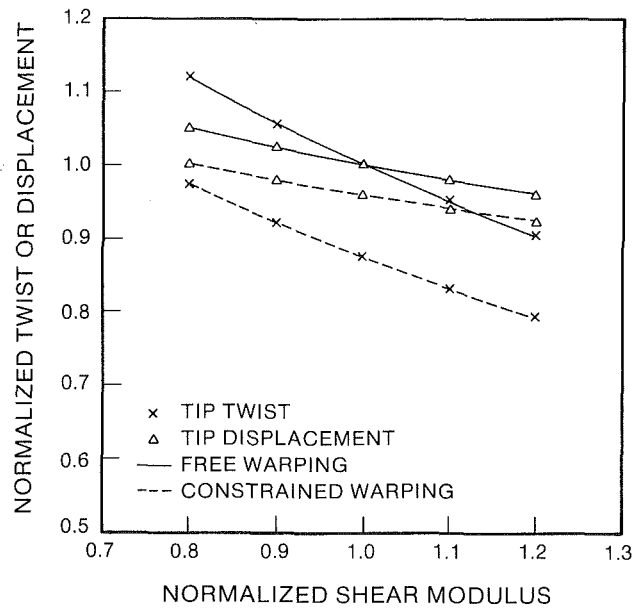


Fig. 9 Normalized tip twist and displacement of a rectangular box beam with bending-twisting coupling, under tip transverse load

flatwise load $P = 1$ KN (this would model the lift force on the blade). Figure 7 shows the resulting edgewise displacement v. This load generates a uniform shear force over the span of the beam, and a linearly varying bending moment. This bending moment couples with edgewise shearing to generate negative v displacements, and is dominant near the root of the beam. On the other hand, the shear force couples with bending about the y axis to generate positive v displacements, and is dominant near the tip of the beam where the bending moment vanishes. This explains the deformed shape of the beam shown in Fig. 7. This coupling between flatwise and edgewise motions is of importance to helicopter blade stability because of the very different aerodynamic damping characteristics for these two types of motion. In these examples, the strain fields (29)–(31) or (32) gave identical predictions.

When nonlinear effects are taken into account, the edgewise motion v displaces the point of application of the load P, causing twisting of the beam. Figure 8 shows this tip twist and the w displacement versus the normalized shear modulus. Torsional warping significantly affect the tip twist (about 20 percent) but has no effect on the axial displacements (the two curves superpose on Fig. 8), and both tip twist and displacement vary with G_{LT}.

Finally the behavior of the beam with bending-twisting couplings is shown in Fig. 9. Once again, both tip twist and displacement are affected by torsional warping and ply shearing modulus.

7 Conclusions

This paper has presented a consistent formulation for the modeling of beams with small initial curvature undergoing arbitrarily large displacements and rotations but small strains. Transverse shearing deformations, torsional warping effects, and elastic couplings resulting from elastically tailored composite beams were included in the analysis. Numerical results based on a finite element approximation of the resulting equations were presented. These examples focused on the extension-twisting behavior of beams as it impacts the design of helicopter composite blades.

The extension-twisting coupling of a pretwisted beam due to torsional warping is included in the model and is in agreement with previous modeling of this effect. Next, a shearing strain square term was shown to affect the behavior of beams with pronounced anisotropy. This nonlinear effect is usually neglected in small strain analyses, however it introduces a kinematic coupling between extension and twisting that should be important for the design and stability of helicopter flexbeam attachments. Finally elastic tailoring of helicopter blades is an increasingly important issue for performance optimization, however some important facts should be noted. First, multiple coupling terms are likely to be present, for instance beam exhibiting extension-twisting or bending twisting couplings will often present bending-transverse shearing or transverse shearing-warping couplings that affect the overall behavior of the structure. Second, several nonclassical effects of beam theory become more prevalent on the overall behavior; for instance torsional warping constraints are known to affect the torsional behavior of a beam, however, when the beam exhibits elastic couplings, torsional warping constraints will also significantly affect bending and axial behavior. Finally, the material shearing modulus also becomes an important factor affecting not only torsional, but axial and bending behavior as well.

Acknowledgments

This research was sponsored by the Army Research Office under Grant DAAG 29-82-K-0093, and Dr. Robert Singleton was contract monitor.

References

Bathe, K. J., 1982, *Finite Element Procedures in Engineering Analysis*, Perentice-Hall, Englewood Cliffs, NJ.

Bauchau, O. A., 1985, "A Beam Theory for Anisotropic Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 107, pp. 416-422.

Bauchau, O. A., Coffenberry, B. S., and Rehfield, L. W., 1987a, "Composite Box Beam Analysis: Theory and Experiments," *J. Reinforced Plastics and Composites*, Vol. 6, No. 1.

Bauchau, O. A., and Hong, C. H., 1987b, "Finite Element Approach to Rotor Blade Modeling," *J. Amer. Hel. Soc.*, Vol. 32, No. 1, pp. 60-67.

Bauchau, O. A., and Hong, C. H., 1987c, "Large Displacement Analysis of Naturally Curved and Twisted Composite Beams," *J. AIAA*, Vol. 25, No. 10.

Degener, M., 1986, "Experimental Investigation of the Torsional Stiffness of a Composite Beam Under Axial Loading," Report of DFVLR Institute of Aeroelasticity.

Dowell, E. H., and Traybar, J. J., 1975, "An Experimental Study of the

Nonlinear Stiffness of a Rotor Blade Undergoing Flap, Lag, and Twist Deformations," Princeton University, Aerospace and Mechanical Science Report No. 1257.

Friedmann, P. P., 1977, "Recent Developments in Rotary-wing Aeroelasticity," *J. Aircraft*, Vol. 14, No. 11, pp. 1027–1041.

Rosen, A., and Friedmann, P. P., 1979, "The Nonlinear Behaivor of Elastic Slender Straight Beams Undergoing Small Strain and Rotations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 161–168.

Rosen, A., 1980, "The Effect of Initial Twist on the Torsional Rigidity of Beams: Another Point of View," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 389–392.

Rosen, A., Loewy, R. G., and Mathew, M. B., 1987, "Nonlinear Analysis of Pretwisted Rods Using 'Principal Curvature Transformation'—Part I—Theoretical Derivation, Part II—Numerical Results," *AIAA J.*, Vol. 25, No. 3, pp. 470–478.

Shyprykevich, P., 1982, "Characterization of Advanced Composite Laminates for Aeroelastic Tailoring," Grumman Aerospace Corporation, Report NAS 1-14759.

Tsai, S. W., and Hahn, H. T., 1983, *Introduction to Composite Materials*, Technomic Publishing Co.

Washizu, K., 1975, *Variational Methods in Elasticity and Plasticity*, Pergamon Press, Elmsford, NY.

Washizu, K., 1964, "Some Considerations on a Naturally Curved and Twisted Slender Beam," *J. Math and Phys.*, Vol. 48, No. 2, pp. 111–116.

Wempner, G., 1981, *Mechanics of Solids with Application to Thin Bodies*, Sijthoff and Noordhoff.

Optimal Design of a Rotating Disk for Kinetic Energy Storage

M. Berger

I. Porat

Faculty of Mechanical Engineering,
Technion–Israel Institute of Technology,
Haifa, ISRAEL

A thin homogeneous rotating disk of variable thickness is considered for the purpose of storing kinetic energy. The objective of the design is to find the optimal shape of the disk for which, in the presence of constraints on the geometry and strength of the disk, the Specific Kinetic Energy (SKE) is maximal. An upper bound for the SKE of a finite diameter disk is derived and a discrete formulation is presented by which an approximate optimal profile for arbitrary design parameters and rotational speeds can be obtained numerically. Applying a parametric study in which optimal designs for a sequence of rotational speeds are observed, a general configuration of the exact optimal profile is presented. The parametric study reveals the existence of three speed intervals, each characterized by a common type of optimal design. The optimal SKE corresponding to the ultimate rotational speed reaches a value very close to the theoretical upper bound, namely twice that of a thin ring. The model gives insight into the nature of optimal designs and serves as a simple and rapid computational tool for finding the optimal profile for arbitrary disk parameters and rotational speeds.

1 Introduction

In the last twenty years there has been a growing interest in the ancient technique of storing energy in a flywheel (Dann, 1973; Lawson, 1973; Rabenhorst, 1974; Hodson, 1978; Davis and Csomor, 1980; Sapowith et al., 1981; Coppa, 1983). This technique was particularly encouraged by high strength-low density filament materials, and a substantial research effort was directed towards the development of a satisfactory flywheel structure using these materials. However, due to certain superior mechanical properties of isotropic metals and since composites are still being improved, it is worthwhile to investigate how such materials could be used most efficiently.

Optimal shapes of rotating disks concerning various design objectives already have been studied by Mahmoodi (1969), Seireg and Surana (1970), Chern and Prager (1970), Distefano (1972), Curtis and Berger (1977), Oda and Yamazaki (1979), Bhavikatti and Ramakrishnan (1980), and Sandgren and Ragsdell (1983). Among these objectives are Minimum Weight, Maximum Plane Stress, Edge Displacement, and Maximum Specific Kinetic Energy (SKE). In those studies, however, only *smooth* shapes have been considered as candidates for optimal solutions and merely *approximate* solutions have been derived. The contribution of this study is a derivation of a general shape configuration from which, for arbitrary design parameters and rotational speeds, an *exact* optimal shape which maximizes the SKE can be determined. It turned out that, among the family of continuous functions, the exact optimal shape is one with *piecewise smooth derivatives*. Applying a parametric study in which optimal designs for a sequence of rotational speeds are observed reveals the existence of three speed intervals, each characterized by a common type of optimal design.

2 The Thin Disk

A solid disk composed of a homogeneous isotropic material of mass density Γ (kg/m^3) is considered. The disk rotates about the Z axis at a maximum angular speed Ω (Rad/s). The disk of diameter $2R_0$ may attain a continuously varying thickness $W(R)$ which is axisymmetrical and symmetrical with respect to the center plane $Z=0$. The thickness is bounded in its relative size as well as in profile slope so that the disk can be treated as a two-dimensional axisymmetrical problem, the strength of which is represented by the radial and tangential stresses $S_R(R)$, $S_T(R)$. The only body-force applied to the disk is the centrifugal force (its weight being neglected). External forces, if such exist at all, are applied radially at the boundary $R=R_0$.

First, define the dimensionless variables

$$r=R/R_0 \quad ; \quad w(r)=W(R)/R_0 \tag{1}$$

$$s_R(r)=S_R(R)/S_{\mathrm{MAX}} \quad ; \quad s_T(r)=S_T(R)/S_{\mathrm{MAX}} \tag{2}$$

$$f_0=F_0/(S_{\mathrm{MAX}}R_0)\geq 0 \tag{3}$$

where S_{MAX} is the yield strength of the material and F_0 the total load per unit length at the circumference of the disk (positive in tension). $w(r)$ will be called the "shape" or "profile" of the disk.

The ultimate performance of a rotating thin ring of the same diameter and material is

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, December 30, 1986; final revision June 27, 1987.

$$V_R^2 = S_{MAX}/\Gamma; \quad E_R = \frac{1}{2}V_R^2 \tag{4}$$

where V_R is the maximum tangential velocity and E_R the maximum SKE of the ring. Thus, choosing the ultimate performance of the ring as a reference level, the edge tangential velocity V_D and the SKE E_D of the disk can be expressed in normalized units by the variables

$$k = V_D/V_R; \quad e = E_D/E_R. \tag{5}$$

Hence,

$$e(k, w) = k^2\left[\int_0^1 r^3 w(r)dr \Big/ \int_0^1 rw(r)dr\right] \tag{6}$$

$$\frac{ds_R}{dr} = -\frac{s_R - s_T}{r} - \left[\frac{w'(r)}{w(r)}s_R + k^2 r\right] \qquad 0 \le r \le 1 \tag{7}$$

$$\frac{ds_T}{dr} = \frac{s_R - s_T}{r} - \mu\left[\frac{w'(r)}{w(r)}s_R + k^2 r\right]$$

$$s_R(0) = s_T(0); \; s_R(1) = f_0/w(1) \tag{8}$$

where μ is Poisson's ratio. Note that the stress derivatives at $r=0$ are singular. However, this singularity can be removed by defining

$$\left.\frac{ds_R}{dr}\right|_{r=0} = -\frac{2+\mu}{3}\frac{w'(0)}{w(0)}s_R(0)$$

$$\left.\frac{ds_T}{dr}\right|_{r=0} = -\frac{1+2\mu}{3}\frac{w'(0)}{w(0)}s_R(0)$$

The calculation of the stresses for a given shape $w(r)$ generally involves the numerical solution of a two-point boundary value problem, which can be solved by a shooting method.

3 The Optimum Design Problem

The objective of the design is to maximize the SKE subject to certain geometrical restrictions and strength requirements, i.e., given k and δ_L, δ_U, δ', μ $(0 < \delta_L < \delta_U;\ \delta' > 0)$, find a continuous function $w(r) \epsilon C^0$ for which

$$e(k, w) \to \text{MAX} \tag{9}$$

and such that for any $0 \le r \le 1$ the following constraints hold:

$$\delta_L \le w(r) \le \delta_U \tag{10}$$

$$|w'(r)| \le 2\delta' \tag{11}$$

$$s_R(r) \le 1; \; s_T(r) \le 1 \tag{12}$$

Note that the optimal solution entirely depends on four design parameters: three "shape" parameters (δ_L, δ_U, δ') and μ. By choosing small enough values for δ_U and δ', constraints (10) and (11) ensure a plane stress state in the disk. A typical set of parameters which will be used throughout this study for numerical evaluations is

$$\delta_L = 0.01; \quad \delta_U = 0.1; \quad 2\delta' = 1.8; \quad \mu = 0.3. \tag{13}$$

Also note that the SKE, as it appears in equation (6), depends on two virtually independent factors: a kinematic and a shape factor. However, the range of the shape factor implicitly depends on the kinematic factor by virtue of (12) and (7). The following theorem gives an estimate of the ultimate SKE of a finite diameter disk.

Theorem. The ultimate SKE of a rotating disk loaded by $f_0 \ge 0$ and fulfilling constraints (10)–(12) is bounded by twice that of a thin ring (i.e., $e(k, w) < 2$).

Proof. Multiplying the first equation of (7) by $r^2 w(r)$ and integrating gives

$$k^2 \int r^3 w(r)dr = -\int rw(r)(s_R - s_T)dr - \int r^2 \frac{d}{dr}[w(r)s_R]dr$$

and integration by parts of the second term on the right-hand side yields

$$k^2 \int r^3 w(r)dr = \int rw(r)(s_R + s_T)dr - r^2 w(r)s_R. \tag{14}$$

Hence, by virtue of the boundary conditions (8), one gets in the region [0, 1]

$$k^2 \int_0^1 r^3 w(r)dr = \int_0^1 rw(r)(s_R + s_T)dr - f_0. \tag{15}$$

However, due to the strength constraints (12), there exists a nonnegative function s_C^2 defined by

$$s_C^2(r) = 2 - [s_R(r) + s_T(r)] \quad 0 \le r \le 1. \tag{16}$$

This function describes the complement of the sum of the actual stresses to the maximum permitted level. Thus, substituting for $s_R + s_T$ into equation (15), one gets, by equation (6),

$$e(k, w) = 2 - \left[\int_0^1 rws_C^2\, dr + f_0\right] \Big/ \int_0^1 rwdr \le 2. \tag{17}$$

The value $e = 2$ can never be reached since at least one of the terms in the numerator must be positive. If $f_0 = 0$, then $s_C(r)$ cannot vanish at the outer section of the disk and the integral in the numerator will be strictly positive. Conversely, if $s_C(r) \equiv 0$, then by (8) $f_0 = w(1) > 0$. Hence,

$$e(k, w) < 2. \qquad \text{Q.E.D.} \tag{18}$$

Before we proceed to derive the optimal shapes, two classical shapes are evaluated.

A Constant-Thickness Disk. Denoting the thickness by w_0, we get from equations (6)–(8)

$$e = (1/2)k^2 \tag{19}$$

$$s_R(r) = \frac{f_0}{w_0} + \frac{3+\mu}{8}k^2[1 - r^2];$$
$$s_T(r) = \frac{f_0}{w_0} + \frac{3+\mu}{8}k^2\left[1 - \frac{1+3\mu}{3+\mu}r^2\right]. \tag{20}$$

Hence, by (12),

$$\text{MAX } e = 4/(3+\mu); \quad \text{MAX } k^2 = 8/(3+\mu) \tag{21}$$

which for $\mu = 0.3$ yields

$$\text{MAX } e = 1.21; \quad \text{MAX } k^2 = 2.42. \tag{22}$$

An Exponential Profile. Consider a shape adapted to the maximal running speed:

$$w(r) = w(1)e^{-k^2(r^2-1)/2}. \tag{23}$$

For a disk loaded maximally at its circumference, i.e., $f_0 = w(1)$, the exponential shape adapted to speed k serves as a constant (maximum) stress shape both in the radial and tangential directions:

$$s_R(r) \equiv s_T(r) \equiv 1 \qquad 0 \le r \le 1. \tag{24}$$

One can easily show that the ultimate performance of the exponential profile for the design parameters (13) and $w(1) = \delta_L$ is

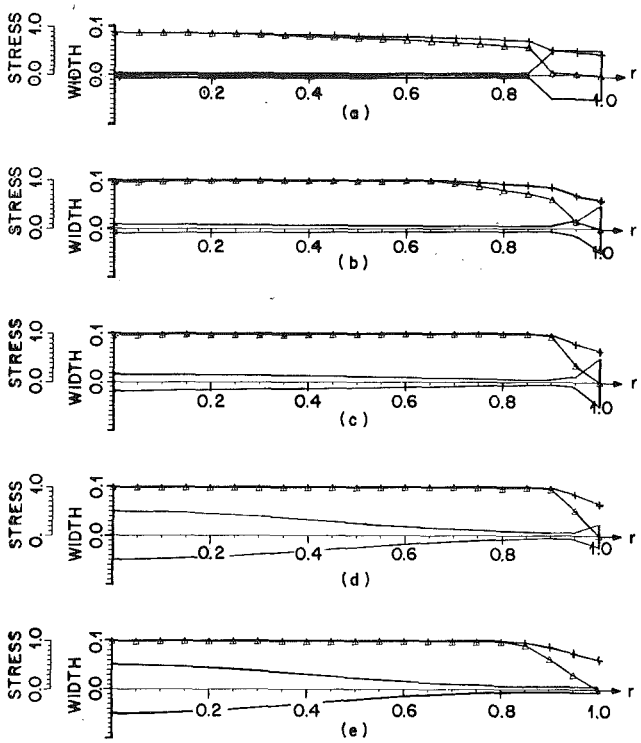


Fig. 1 Optimal discrete shapes and stresses versus r for the Design Parameters (13) ($\delta_L = 0.01$, $\delta_U = 0.1$, $2\delta' = 1.8$, $\mu = 0.3$). The Δ and $+$ symbols trace the s_R and s_T curves correspondingly: (a) $k^2 = 1.0$; (b) $k^2 = 2.1$; (c) $k^2 = 3.1$; (d) $k^2 = 5.5$; (e) $k^2 = 6.5$.

$$\text{MAX } e = 1.49; \quad \text{MAX } k^2 = 4.61. \tag{25}$$

The gap between the ultimate SKE in equations (25) and the upper bound ($=2$) surely justifies the search for an optimal shape.

4 Discrete Solution

Problem (9)–(12) is now approximated by a discrete formulation in which the disk is divided into N concentric rings of conic shape separated by $N+1$ equally-spaced division points r_i, $i = 1, \ldots, N+1$, ($r_1 = 0$; $r_{N+1} = 1$). Denoting the thickness at r_i by w_i and approximating the shape of each element by a profile $\hat{w}$ varying linearly, we have for $i = 1, \ldots, N$

$$\begin{aligned} &r_{i+1} - r_i = 1/N \\ &\hat{w}(r) = w_i + N(w_{i+1} - w_i)(r - r_i) \qquad r_i \le r \le r_{i+1}. \end{aligned} \tag{26}$$

Denoting the stress functions in the ith ring by $\hat{s}_R^{(i)}(r)$, $\hat{s}_T^{(i)}(r)$, these stresses can be calculated by solving an $N+1$ point boundary value problem related to a system of N rings, i.e., one must solve

$$\frac{d\hat{s}_R^{(i)}}{dr} = -\frac{\hat{s}_R^{(i)} - \hat{s}_T^{(i)}}{r} - \left[\frac{N[w_{i+1} - w_i]}{\hat{w}(r)}\hat{s}_R^{(i)} + k^2 r\right]$$

$$r_i \le r < r_{i+1}, \quad i = 1, \ldots, N$$

$$\frac{d\hat{s}_T^{(i)}}{dr} = -\frac{\hat{s}_R^{(i)} - \hat{s}_T^{(i)}}{r} - \mu\left[\frac{N[w_{i+1} - w_i]}{\hat{w}(r)}\hat{s}_R^{(i)} + k^2 r\right]$$

subject to the boundary conditions

$$\begin{aligned} &\hat{s}_R(0) = \hat{s}_T(0); \quad \hat{s}_R(1) = f_0/w_{N+1} \\ &\hat{s}_R^{(i)}(r_i) = \hat{s}_R^{(i-1)}(r_i); \quad \hat{s}_T^{(i)}(r_i) = \hat{s}_T^{(i-1)}(r_i) \qquad i = 2, \ldots, N. \end{aligned}$$

Numerical solutions of this system were calculated in this study by a method based on Biezeno and Grammel (1954) and Manna (1968).

Thus, the original optimization problem (9)–(12) is replaced by the following discrete scheme: Given k and δ_L, δ_U, δ', μ, find $\{w_i\}$ which maximizes

$$\hat{e}(k, \hat{w}) = k^2 \left[\sum_{i=1}^{N} \int_{r_i}^{r_{i+1}} r^3 \hat{w}(r)\, dr\right] \Big/ \left[\sum_{i=1}^{N} \int_{r_i}^{r_{i+1}} r\hat{w}(r)\, dr\right] \tag{27}$$

subject to

$$\begin{aligned} &\delta_L \le w_i < \delta_U && i = 1, \ldots, N+1 \\ &-2\delta'/N \le (w_{i+1} - w_i) \le 2\delta'/N && i = 1, \ldots, N \\ &\hat{s}_R(r_i) \le 1; \ \hat{s}_T(r_i) \le 1 && i = 1, \ldots, N \\ &w_{N+1} \ge f_0 \end{aligned} \tag{28}$$

Problem (27)–(28) retains the stress constraints as nonlinear functions of the variables w_i. Thus, a Nonlinear Programming algorithm which can treat nonlinear inequality constraints must be applied (Gill and Murray, 1974). The question of choosing the most efficient numerical scheme for stress calculation and shape optimization was discussed in earlier publications (Curtis and Berger, 1977; Sandgren and Ragsdell, 1983). However, this issue is not of prime interest in this study, since the discrete solution serves merely as an indication and a guide leading to the analytic optimal shape.

At this point, we proceed to a parametric study in which problem (27)–(28) is solved parametrically for a sequence of k values. Figures 1(a)–1(e) demonstrate some of the optimal discrete shapes of a free disk, each accompanied with its stress distribution curves. These designs are derived for the sequence $k^2 = 1.0, 2.1, 3.1, 5.5, 6.5$ which correspond to the rim tangential velocities $V_D = 436, 632, 767, 1022, 1111$ m/s (rounded values) of a Marage Steel disk ($S_{\text{MAX}}/\Gamma = 1.9 \times 10^5$ m^2/s^2) and use the design parameters in equations (13) with $N = 20$.

Observing the evolvement of the optimal designs in Fig. 1 as a function of k, one realizes the following trend: Starting at low speeds, the optimal design appears to accumulate a bundle of mass at the rim which gradually shrinks as k climbs to higher values, until almost vanishing. Regarding the inner part of the disk, the optimal design is of minimal thickness for low speeds but, as k increases, a constant (maximum) stress shape seems to evolve in this region, the size of which extends both in thickness and in length. Then, for very high speeds (see Fig. 1(e)), the constant stress section surprisingly *retreats* back towards the center and covers a shorter inner section of the disk.

Note: The stress curves in Fig. 1 show some deviations around the maximum stress level at regions where a constant maximum stress $s_R = s_T = 1$ is expected. This is due to a termination of the computer runs at a stage seeming satisfactory to point out the general trend of the optimal shapes.

The discrete scheme was applied to a variety of design parameters' values δ_L, δ_U, δ' used in common practice, namely in the range

$$\delta_U \le 0.2; \ 0.1 \le \delta_L/\delta_U \le 1; \ \delta' \ge 0.2$$

and the character of the optimal designs as presented above was found typical of all. As a consequence, relying on the discrete results and supported by physical arguments, an analytical exact optimal solution is proposed in the next section.

5 A General Optimal Analytical Shape

Let us define a general configuration of piecewise differentiable disk profiles constructed from four basic shapes as shown in Fig. 2 and defined analytically by

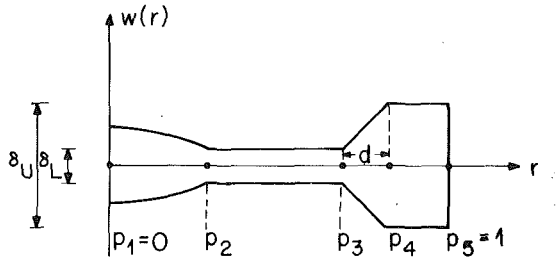


Fig. 2 The general configuration of the exact optimal shapes

$$w(r)=\begin{cases} w_1(r)=\delta_L\exp[(k^2/2)(p_2^2-r^2)] & p_1\le r\le p_2 \\ w_2(r)=\delta_L & p_2\le r\le p_3 \\ w_3(r)=\delta_L+2\delta'(r-p_3) & p_3\le r\le p_4 \\ w_4(r)=\delta_U & p_4\le r\le p_5 \end{cases} \tag{29}$$

The four basic shapes are the exponential constant (maximum) stress shape, the minimum thickness shape, the maxium slope shape, and the maximum thickness shape.

p_1 and p_5 are fixed at $p_1=0$, $p_5=1$. The location of p_3 is a free parameter. For a certain choice of p_3, the location of p_4 is chosen so as to set up maximum mass at the rim, i.e., at a point where the maximum slope ray originating from p_3 reaches either the maximum thickness line or the edge of the disk (the one which occurs first). Formally, if $d=(\delta_U-\delta_L)/2\delta'$, then

$$p_4=\begin{cases} p_3+d & p_3+d<1 \\ 1 & p_3+d\ge 1 \end{cases} \tag{30}$$

Getting p_4 allocated, p_2 is located so as to satisfy the strength requirement at the inner section of the disk while applying a minimum amount of mass in that section. Suppose the minimum thickness shape w_2 in Fig. 2 is extended to the whole section $[0, p_3]$. As a result of the rotation, stresses s_R, s_T appear in section $[0, p_3]$ which monotonously increase towards the center, coinciding to a common maximum level (see equations (8)). For low enough k, the stresses at the center may happen to lie under the maximum permitted level, in which case $p_2=0$ is chosen and w_2 prevails in $[0, p_3]$. But for high enough k, such that the stresses at the center exceed the maximum permitted level, we locate a point q_2 where a strength constraint is first violated, as shown in Fig. 3(*a*). The shape w_2 is now replaced in section $[0, q_2]$ by an exponential w_1 drawn from q_2 backwards to the center. As a result, the stresses drop in the whole disk below the maximum permitted level. In order to maintain the strength conditions by applying minimum mass at the center section, point q_2 together with the exponential profile is pushed to the left (see Fig. 3(*b*)) until a critical point $q_2=p_2$ is reached for which the condition $s_R(p_2)=s_T(p_2)=1$ is met. At this point, the exponential profile serves as a constant stress profile, i.e., $s_R(r)\equiv s_T(r)\equiv 1$ for $0\le r\le p_2$.

Note that for configuration (29) *at every point* along the disk radius at least one of the constraints (10)–(12) is active (i.e., the equality sign holds).

For a given triple δ_L, δ_U, δ' and given k, the "Analytic Profile" (29) defines a one-parameter family of profiles depending on p_3. It will be shown that the exact optimal solution of problem (9)–(12) is one of the family (29) with $p_3=p_3^*$ for which the SKE is maximized. Thus, for fixed δ_L, δ_U, δ', a family of optimal designs characterized by $p_3^*=p_3^*(k)$ exists.

Moreover, the family of optimal designs can be separated into three groups corresponding to three consecutive speed intervals, namely "low" range $[0, k_1]$, "medium" range $[k_1, k_2]$, and "high" range $[k_2, k_3]$ speeds. Each range is characterized by a subset of constraints (10)–(12) being active in that range and a specific rule of choosing p_2. The speed k_3 is the ultimate feasible speed of a disk.

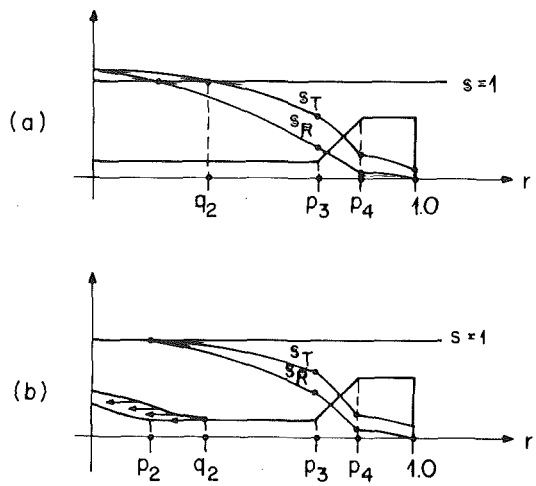


Fig. 3 Construction of the concentric constant strength section: (*a*) nonfeasible state; (*b*) feasible state and shape

6 The Optimal Shape for Low Speeds $[0, k_1]$

Denote

$$E(p_i)=\int_{p_i}^{p_{i+1}} r^3 w_i(r)\,dr \tag{31}$$

$$M(p_i)=\int_{p_i}^{p_{i+1}} r w_i(r)\,dr \tag{32}$$

and denote the stress functions related to a profile of the family (29) by

$$s_R(r)=s_1[r, k, p_2, p_3];\quad s_T(r)=s_2[r, k, p_2, p_3] \tag{33}$$

Assertion. The optimal design of a free disk rotating at "low" speeds is a profile (29) with $0\le p_3\le 1$ which maximizes

$$e_1(p_3)=\frac{E(p_2)+E(p_3)+E(p_4)}{M(p_2)+M(p_3)+M(p_4)} \tag{34}$$

with

$$p_2=p_1=0;\quad p_4 \text{ determined by equation (30).} \tag{35}$$

The idea behind these equations is that for low enough speeds the stress constraints are nonactive and the optimal mass distribution is determined only by the geometrical constraints. The optimal profile for such speeds is presented in Fig. 4(*a*) and is characterized by a minimum thickness center section, maximum slope section, and a maximum thickness outer section. The optimal location p_3^* is a transition point, to the left of which it is optimal to apply a minimum amount of mass, whereas to the right of it maximum mass accumulation is advantageous.

The optimal profile is determined by solving equations (34) numerically for p_3, which involves quite a simple task of maximizing a function of one variable. Note that the solution of problem (34)–(35) is independent of k and μ. Thus, the optimal shape is unique for a range of speeds $k\ge 0$ and valid as long as the stress constraints (12) are not violated. From equations (20), the maximum stress is at the center and monotonously increases as k does. Hence, the optimal shape (34) is feasible up to $k=k_1$ for which:

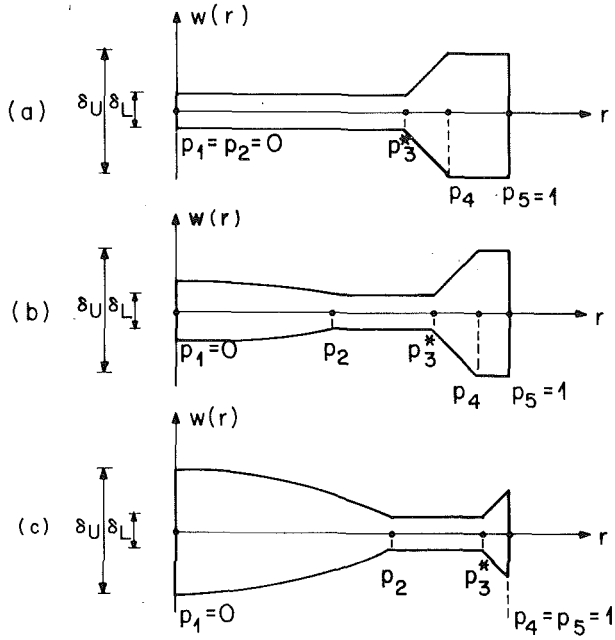


Fig. 4 Typical exact optimal shapes: (a) low speeds range; (b) medium speeds range; (c) high speeds range

$$s_1(0, k_1, 0, p_3^*) = 1. \tag{36}$$

Solving equation (36) with the design parameters (13), one obtains $k_1^2 = 1.131$ and $e = 0.859$. The shape factor of this design is $e/(k_1^2) = 0.76$, while that of a constant thickness shape is only 0.5.

7 The Optimal Shapes for Medium Speeds $[k_1, k_2]$

In the range of $[0, k_1]$, only the geometrical design parameters actively participated in shaping the optimal disk. As k increased, though initially fulfilled safely, the strength constraints steadily approached the maximum stress bound until, upon reaching $k = k_1$, they became active at the center $r = 0$. We pass now into a range of speeds for which the strength constraints actively participate in forming the optimal shape.

Assertion. The optimal shape of a free disk rotating at speed k in the "medium" speeds range is one of (29) which maximizes the function in p_3:

$$e_2(p_3) = k^2 \frac{E(p_1) + E(p_2) + E(p_3) + E(p_4)}{M(p_1) + M(p_2) + M(p_3) + M(p_4)} \tag{37}$$

with $p_2 (p_2 < p_3 \le 1)$ determined as the solution of

$$s_1(0, k, p_2, p_3) = 1. \tag{38}$$

Assertion (37)–(38) states that the optimal shape for medium speeds must follow in principle the strategy of the transition point p_3^* in the same manner as stated for low speeds, i.e., a minimum amount of mass should be applied to the left of p_3^* and a maximum amount to the right of it. However, the "minimum mass" goal in the section to the left of p_3^* must be realized for medium speeds partly by an exponential and partly by a minimum thickness profile. The objective of the exponential profile is to maintain the strength conditions in the center section $[0, p_2]$ by applying minimum mass. A typical optimal profile is demonstrated in Fig. 4(*b*).

In practice, the solution of (37)–(38) is carried out numerically by a combined single variable function minimizer and an equation solver. For a tentative p_3 examined by the function minimizer, the related p_2 location is calculated by an equation solver. The optimal designs for an increasing sequence of speeds starting at k_1 can be determined recursively. In each step, the optimal pair p_2^*, p_3^* of the last optimal design is used as an initial guess in the present step. The particular set of optimal designs corresponding to the design parameters (13) is presented in Fig. 7 by means of the optimal pairs p_2^*, p_3^* as a function of k. One can see that as k increases, the exponential section expands to the right, which causes a monotonously increasing thickness at the center $r = 0$. Hence, the solution (37)–(38) can serve as long as $w(0)$ does not violate the maximum thickness constraint. Thus, the ultimate k for the "medium" speeds range is determined from (29) by imposing $w_1(0) = \delta_U$, which yields the condition

$$k_2^2 = -2 \ln\delta / p_2^2; \quad \delta = \delta_L / \delta_U. \tag{39}$$

Note that in the medium range, as k increases, the transition point p_3^* moves to the right and the bundle of mass at the rim shrinks. This implies lower and lower shape factors, but the product of the kinematic and shape factors still yields monotonously increasing SKE values.

8 The Optimal Shapes for High Speeds $[k_2, k_3]$

The medium range $[k_1, k_2]$ was dominated by the evolvement of the exponential profile in some inner section of the disk. The thickness throughout this section was completely ruled by the location of the critical point p_2. We pass to the "high" range speeds in which our concern is shifted from the critical point p_2 to the center of the disk. From here on the center thickness must be guarded from violating both the stress and the maximum thickness constraints.

Assertion. The optimal shape of a free disk rotating at speed k in the "high" speed range is one of (29) which solves the equation in p_3:

$$s_1(0, k, p_2^*, p_3) = 1 \tag{40}$$

with

$$(p_2^*)^2 = -2\ln\delta / k^2. \tag{41}$$

Assertion (40)–(41) states that for high speeds, i.e., for $k > k_2$, the strategy of the transition point p_3 still holds. However, the "minimum mass" goal in section $[0, p_2]$ as implemented for medium speeds must now be modified. Instead of drawing an exponential from the bottom point (p_2^*, δ_L), an exponential profile originating at the top center point $(0, \delta_U)$ should be drawn forward down to the minimum thickness line. The analytic form of this profile is

$$w(r) = \delta_U e^{-k^2 r^2/2}. \tag{42}$$

The intersection of this curve with the minimum thickness line determines the point p_2^* (see equation (41)).

The "maximum mass" goal to the right of the transition point p_3 is implemented by locating p_3 so as to maintain a constant maximum strength state in section $[0, p_2^*]$ (see equation (40)). Locating p_3 to the left of this point will cause a violation of the strength conditions, while locating it to the right will diminish the SKE unnecessarily and cause a suboptimal design. A typical optimal shape is presented in Fig. 4(*c*).

By gradually increasing the speed beyond k_2 one finds out that, though the bundle of mass at the rim rapidly shrinks, the optimal SKE is still monotonously increasing. This process is eventually terminated at the instant the bundle vanishes, i.e., when p_3^* gets to the edge $r = 1$. At this instant the optimal shape consists of only two of the four basic shapes, namely the exponential and minimum thickness shapes (see Fig. 5). The speed k_3 at this instant is given by

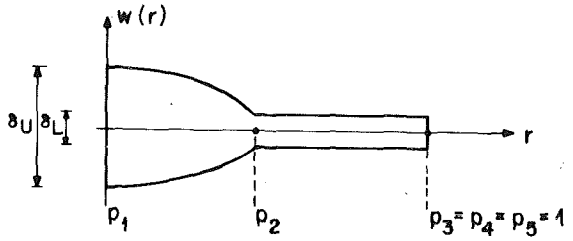


Fig. 5 The optimal shape for ultimate speed k_3

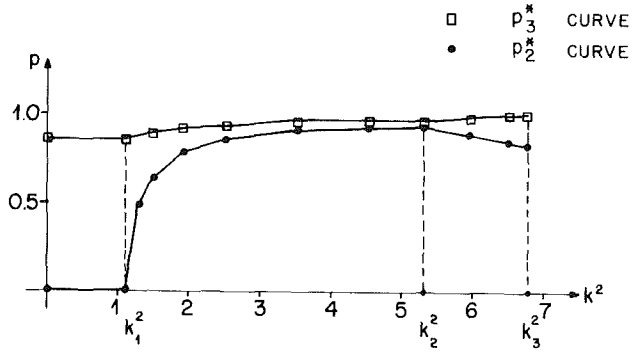


Fig. 6 Optimal location of p_2 and p_3 as a function of k^2 (for $\delta_L = 0.01$, $\delta_U = 0.1$, $2\delta' = 1.8$, $\mu = 0.3$)

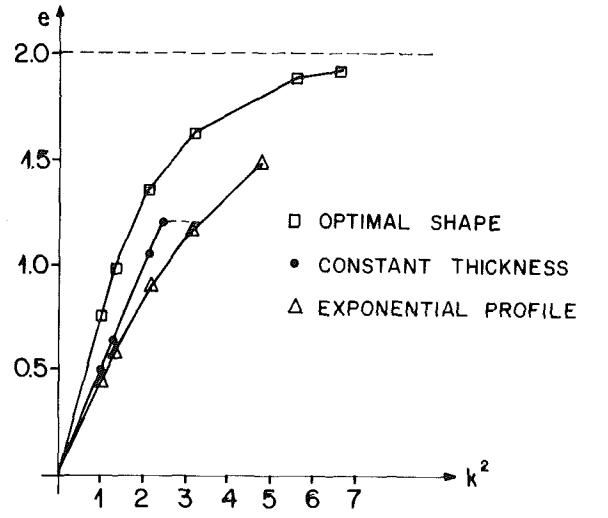


Fig. 7 The SKE as a function of k^2 for various designs. The optimal shape curve corresponds to $\delta_L = 0.01$, $\delta_U = 0.1$, $2\delta' = 1.8$, $\mu = 0.3$.

$$k_3^2 = \frac{4}{3+\mu}\left[\left(1-\frac{1+\mu}{2}\ln\delta\right)+\left[\left(1-\frac{1+\mu}{2}\ln\delta\right)^2 + \frac{1}{4}(3+\mu)(1-\mu)(\ln\delta)^2\right]^{1/2}\right] \tag{43}$$

This value was derived from a bored constant thickness disk of inner radius $r_1 = p_2^*$ by imposing the conditions $s_R(1) = 0$, $s_R(r_1) = s_T(r_1) = 1$ and condition (41). k_3 is the *ultimate* rotational speed one can apply to a solid disk. It depends on the minimum/maximum thickness ratio and μ. The ultimate speed for design parameters (13) is $k_3^2 = 6.72181$.

Note that by equation (41) p_2^* is monotonously decreasing in the "high" range, which means that the constant strength section retards backwards towards the center. This effect is implied by holding the exponential origin tied to a fixed thickness at the center and due to steeper declining slopes of the exponential as k increases. Consequently, this phenomenon which was first noticed with the numerical solutions is now theoretically explained by the analytic model (29).

Table 1 Optimal SKE values for design parameters set (13). "XXX" indicates the speeds for which the constant or exponential shapes are unrealizable while the optimal shape is still feasible and improving in performance

k^2	1.0	2.1	3.1	5.5	6.5
Optimal SKE (Discrete)	0.7587	1.340	1.613	1.879	1.907
Optimal SKE (exact)	0.7588	1.367	1.637	1.897	1.926
SKE ratio (Optimal/Constant)	1.52	1.30	XXX	XXX	XXX
SKE ratio (Optimal/Exponential)	1.65	1.57	1.41	XXX	XXX

9 A Comparison With the Discrete Designs

A complete set of exact analytic optimal designs has been calculated as function of k for the design parameters stated in equations (13). Figure 6 shows the optimal pairs p_2^*, p_3^* as a function of k and Fig. 7 demonstrates the corresponding optimal SKE. Note that the SKE curves of the constant thickness and exponential shapes relate to one and the same shape for all k, whereas the optimal SKE curve involves a different design for each $k > k_1$. Figure 4 shows three optimal designs typical to the three ranges of speed and Fig. 5 shows that for the ultimate speed. One can readily see that the discrete shape of Fig. 1(*a*) relates to the low range, shapes 1(*b*)-1(*c*) to the medium range, and shapes 1(*d*)-1(*e*) to the high range of speeds. The analytic shapes closely match the discrete shapes and attain, as expected, a slightly higher SKE than those of the corresponding discrete solutions, as shown in Table 1. Table 1 also shows the relative improvement of the optimal designs over the constant and exponential shapes.

Note that the analytic model calls for a maximum slope to the right of the transition point p_3, while the discrete solutions appear to show a change in slope less than the maximum allowable value (see Fig. 1(*b*)-1(*d*)) due to a discretization effect common to finite element approximations. In fact, since in the discrete formulation (26) the disk thickness changes slope only at fixed division points, a turn of maximum slope at the accurate optimal position is unrealizable by this model if the optimal location of p_3 happens to fall between two division points. In such a case the discrete optimizer chooses alternative optimal slopes at one or two consecutive divisions containing p_3 which partly compensate for the false position of the "corner."

10 Conclusion

An upper bound of the SKE of a *finite diameter* rotating thin disk subject to shape and stress constraints has been derived. A general model has been established from which an *exact* optimal shape for an arbitrary rotational speed can be constructed, and a rapid computational tool to calculate the optimal shape was formulated. This model is supported by approximate discrete solutions and by mathematical and mechanical arguments. For typical design parameters used in practice, optimal designs very close to the upper bound were derived.

Some interesting conclusions should be indicated:

(1) Contrary to traditional smooth shapes, thought of as optimal designs, the true optimal designs are found to involve piecewise differentiable shapes generally containing three "corners."

(2) Upon applying shape constraints as part of the design requirements, the traditional "constant strength" design is suboptimal (unless $f_0 = 1$). The true optimal design is superior for *any* rotational speed.
(3) As a rule, the optimal shapes are constrained at *all* points along the disk either by a shape or a strength constraint or both.
(4) The ultimate SKE value, as well as the ultimate rotational speed, depends on Poisson's ratio and parameter δ only.
(5) The optimal design which maximizes the SKE for the ultimate speed is also the optimal design for a Minimum Weight objective.
(6) The present study discussed the case of a free solid disk. Applying a compressive load at the disk rim ($f_0 < 0$) may elevate the maximal SKE to a value higher than 2 (see equation (17)).

References

Bhavikatti, S. S., and Ramakrishnan, C. V., 1980, "Optimum Shape Design of Rotating Disks," *Computers and Structures*, Vol. 11, pp. 397–401.

Biezeno, C. B., and Grammel, R., *Engineering Dynamics*, Vol. 3, Blackie and Son, 1954.

Chern, J. M., and Prager, W., 1970, "Optimal Design of Rotating Disk for given Radial Displacement of Edge," *JOTA*, Vol. 6, No. 2, pp. 161-170.

Coppa, A. P., 1983, "Composite Ring-Disk Flywheel, Design Fabrication, and Testing," Report UCRL-15566, General Electric Co., Philadelphia, PA.

Curtis, D. M., and Berger, B. S., 1977, "Optimum Design of the Homogeneous Plane-Stress Flywheel," ASME Paper 77-DET-112.

Dann, R. I., 1973, "The Revolution in Flywheels," *Machine Design*, May 73, pp. 130–135.

Davis, D., and Csomor, A., 1980, "The New Age of High Performance Kinetic Energy Storage Systems," *Proceedings of the 15th Intersociety Energy Conversion Engineering Conference*, pp. 1507–1512.

Distefano, N., 1972, "Dynamic Programming and the Optimum Design of Rotating Disks," *JOTA*, Vol. 10, No. 2, pp. 109–128.

Gill, P. E., and Murray, W., 1974, *Numerical Methods for Constrained Optimization*, Academic Press.

Hodson, D. R., 1978, "High Energy Storage Flywheel Module," Report RI/RD78-207, Rockwell International, Canoga Park, CA.

Lawson, L. J., 1973, "Kinetic Energy Storage: A 'New' Propulsion Alternative for Mass Transportation," ASME paper 73-ICT-10.

Mahmoodi, P., 1969, "On the Optimum Design of a Rotating Disk of Nonuniform Thickness," ASME Paper 69-DE-8.

Manna, F., 1968, "Rotating Disks of Unconventional Profile," *Meccanica*, Italian Assoc. Theor. and Appl. Mech., pp. 274–282.

Oda, J., and Yamazaki, K., 1979, "On a Technique to obtain an Optimum Strength Shape by the Finite Element Method—Application to the Problem Under Body Force," *Bull. JSME*, Vol. 22, No. 164, pp. 131–140.

Rabenhorst, D. W., 1974, "The Multirim Superflywheel," Report APL/JHU-TG 1240, Johns Hopkins University applied Physics Laboratory, Silver Spring, MD.

Sandgren, E., and Ragsdell, K. M., 1983, "Optimal Flywheel Design With General Thickness Form Representation," *J. Mechanisms. Transm. Autom. Design*, Vol. 105, No. 3, pp. 425–433.

Sapowith, A. D., et al., 1981, "Evaluation and Design Considerations of Woven Composite Flywheel Materials Constructions," Report UCRL-15415, Lawrence Livermore National Lab.

Seireg, A., and Surana, K. S., 1970, "Optimum Design of Rotating Disks," *J. Eng. Ind.*, pp. 1–10.

Analytical and Experimental Study of Beam Torsional Stiffness With Large Axial Elongation

M. Degener
Research Scientist,
Institute for Aeroelasticity,
German Aerospace Research
Establishment (DFVLR),
Göttingen, West Germany

D. H. Hodges
Professor,
School of Aerospace Engineering,
Georgia Institute of Technology,
Atlanta, Georgia

D. Petersen
Research Scientist,
Institute for Structural Mechanics,
German Aerospace Research
Establishment (DFVLR),
Braunschweig, West Germany

The axial force and effective torsional stiffness versus axial elongation are investigated analytically and experimentally for a beam of circular cross section and made of an incompressible material that can sustain large elastic deformation. An approach based on a strain energy function identical to that used in linear elasticity, except with its strain components replaced by those of some finite-deformation tensor, would be expected to provide only limited predictive capability for this large-strain problem. Indeed, such an approach based on Green strain components (commonly referred to as the geometrically nonlinear theory of elasticity) incorrectly predicts a change in volume and predicts the wrong trend regarding the experimentally determined axial force and effective torsional stiffness. On the other hand, use of the same strain energy function, only with the Hencky logarithmic strain components, correctly predicts constant volume and provides excellent agreement with experimental data for lateral contraction, tensile force, and torsional stiffness—even when the axial elongation is large. For strain measures other than Hencky, the strain energy function must be modified to consistently account for large strains. For comparison, theoretical curves derived from a modified Green strain energy function are added. This approach provides results identical to those of the Neo-Hookean formulation for incompressible materials yielding fair agreement with the experimental results for coupled tension and torsion. An alternative approach, proposed in the present paper and based on a modified Almansi strain energy function, provides very good agreement with experimental data and is somewhat easier to manage than the Hencky strain energy approach.

1 Introduction

This paper concerns the elastic behavior of beams made of materials that can suffer large strains without permanent deformation and the feasibility of modeling the large-strain behavior of such beams *without* explicit consideration of material nonlinearity (i.e., without higher-order elastic constants). We consider several possible strain energy functions and their influence on the predicted effective torsional stiffness of a beam under an axial tensile force. The efficacy of these formulations is examined through analysis of the behavior of this simple structure and comparison of the results with experimental data. Large elongations are included mainly to accentuate the differences among the various approaches, and only small torsion is considered for simplicity.

1.1 Background. Analysis of deformable elastic structures, such as beams, plates, and shells, frequently necessitates consideration of kinematical nonlinearities. The reason for this is that, even for small strains, such structures can deform significantly. Indeed, in extreme cases displacements may remarkably exceed the thin dimension(s) of the structure and rotations may be large. Nevertheless, it is possible for the strains to remain sufficiently small compared to unity so that Hooke's law remains essentially valid, although equilibrium must be established with respect to the deformed structure.

When strains are not restricted to remain small, material nonlinearities can be included in the formulation in terms of explicit, higher-order elastic constants. When there is no permanent deformation, an alternative is to compare theoretical predictions in terms of how well various strain energy functions (or constitutive laws) fare that are based only on material constants evaluated at small strains. For this approach, elasticity laws are expressed as a relationship between some measure of stress and some measure of strain, the measures being basically arbitrary. Because most strain definitions are not linear in principal elongations, however, the material law embodies a certain material nonlinearity—albeit implicitly. It is reasonable to assume that there may exist problems for which the implied material nonlinearity in one formulation is superior to that of another based on the same material con-

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, February 24, 1986; final revision April 30, 1987.

stants. For isotropic materials, the question could be posed as how well a two-parameter strain energy function can perform when the strains become large and when the two parameters are the same ones with the same numerical values as those used in a corresponding linear theory.

In most work involving small-strain mechanics of deformable structures, a strain energy approach based on Green strain components is used as presented by Wempner (1981), for example. This approach is equivalent to application of the principle of virtual work with a material law that is a linear relationship between the second Piola-Kirchhoff stress components and the Green strain components as shown in Washizu (1968). We emphasize that, for this formulation, the strains may exhibit considerable kinematical nonlinearity but still remain small. This type of formulation is the basis for most work done involving the geometrically nonlinear analysis of deformable structures. It is also the basis for all the work cited below unless otherwise specified and serves as a starting point to facilitate comparisons.

1.2 Previous Work. The interaction of normal force and torsion in a beam was investigated for the first time by Wagner (1929) in his analysis of stability phenomena. In what has come to be known as Wagner's hypothesis, he hypothesized that a beam consists of longitudinal fibers in its natural state. Wagner predicted that the effective torsional stiffness of beams with open cross section would decrease under a compressive load and increase for a tensile load. The analysis did not consider beams of closed cross section.

The torsion of a prismatic beam with initial axial tensile force was solved by Biot (1939) without making use of Wagner's hypothesis. Biot recognized, however, that his material law is not strictly valid in the presence of initial stress, and that there is an error in his solution of the order of the initial tensile stress over the shear modulus. From this Biot concluded, in agreement with Wagner's solution, that beams with low torsional rigidity in their natural state (e.g., beams with open cross sections) show an increase in effective torsional rigidity due to tensile stress. On the other hand, torsional rigidity of beams with circular cross section, where Saint-Venant's torsional moment of inertia is equal to the polar moment of inertia, are not affected significantly by the initial stress.

Most beam analyses are not concerned with the influence of axial tensile stress. One class of papers in which it is considered, however, is that dealing with the geometrically nonlinear behavior of helicopter rotor blades. In the mainstream of papers associated with helicopter rotor blade analysis prior to 1980 (e.g., see Houbolt and Brooks, 1958; Hodges and Dowell, 1974; and Friedmann, 1977), the Wagner hypothesis was used, at least implicitly, in formulation of the equations. In Hodges (1980), a special case of the problem was carefully rederived without the Wagner hypothesis. Nevertheless, the term that reflects an increase in the torsional rigidity due to axial tensile force is in the same form in these four references; it shows that the torsional stiffness will increase due to tensile force regardless of whether the beam warps. For warp-free beams, the relative magnitude of this increase is of the order of elongation compared to unity, however, and thus in this case it is of little practical significance. However, the *direction* of the change in effective torsion stiffness versus elongation may be of practical significance, possibly impacting the accuracy of predicted behavior for beams of general cross section.

Petersen (1982) presented an analysis of beam torsion in the presence of axial tensile force. His analysis was based on a strain energy function written in terms of Almansi's strain components. He interpreted this as implying a linear relationship between the Cauchy-Euler stress components and the Almansi strain components. His expression for the effective torsional stiffness contains a term proportional to axial stress that vanishes for beams with warp-free cross sections but that essentially agrees with the analogous term from Wagner (1929) and Biot (1939) for beams with cross sections prone to warp.

In a rigorous treatment of beam problems in which the strains are small compared to unity, Hodges (1987) confirmed Petersen's result that the effective torsional stiffness for nonwarping beams is independent of elongation. (Both Petersen, 1982, and Hodges, 1987, assume that Poisson contraction effects can be neglected, as do the works cited above.) He also showed that the predicted expression for the effective torsional stiffness depends on the constitutive law assumed in the analysis, even for small strains. For small strains, the differences are only noticeable for transversely isotropic beams with $G \ll E$, where G is the shear modulus and E is the Young's modulus, but for large strain such differences may not be inconsequential.

Both Petersen's (1982) and Hodges' (1987) predictions are in qualitative agreement with Biot's (1939) but are different from those of Houbolt and Brooks (1958), Hodges and Dowell (1974), Friedmann (1977), and Hodges (1980). The differences are noticeable only for large elongation or for transversely isotropic materials with $G \ll E$. While the former is not of any practical concern for helicopter problems and is beyond the limiting assumptions involved in formulation of these analyses, for a composite bearing less rotor flexbeams the difference in the predicted behavior for the latter case could prove to be important.

In all the work cited above concerning coupled extensional-torsional behavior of beams, the strains are assumed to be sufficiently small that the Poisson contraction effects can be neglected. In view of the large elongation treated in the present investigation, however, these effects are not likely to be insignificant. Concerning other aspects of this problem associated with large strains, Truesdell (1966) and Green and Zerna (1968) discuss various special cases including the combined extension and torsion of circular cylinders. Although the constitutive law is only generally formulated with the aid of a strain energy function in terms of strain invariants, the results are expressed as the "universal relation of Rivlin." This expression relates the tensile force to the torsional moment and is independent of constitutive law. When applied to explicit formulae developed from various approaches, it serves as a check for theoretical consistency.

In summary of the small-strain analyses, Houbolt and Brooks (1958), Hodges and Dowell (1974), Friedmann (1977), and Hodges (1980) show that the effective torsional stiffness of a warp-free beam increases due to tensile load. This increase is not negligible for either large elongations or for materials with $G \ll E$. Petersen (1982) and Hodges (1987), on the other hand, show that the effective torsional stiffness does not change in a warp-free beam under a tensile load. The reason behind this discrepancy is the choice of material law. Although these references all use linear stress-strain laws, different stress and strain measures are involved. Since material laws are empirical in nature, a means of establishing the accuracy of the analyses was needed. This need led to the present paper which presents a detailed rederivation of certain restricted versions of the problem with various strain energy functions and corresponding experiments.

1.3 Present Approach. The small-strain approaches referenced above cannot be simply extrapolated for application to large-strain problems. The reason for this is that these formulations contain a term involving the trace of the strain tensor components—the volume change in the case of a linear theory. Clearly, however, the trace of the strain tensor components is *not necessarily* the correct expression for the

volume change if the strains are large. In this case, the volume change must be evaluated from the determinant of the deformation gradient.

A strain energy approach based on the Hencky (logarithmic) strain tensor was formulated by Anand (1979) in which a linear relationship is posed between Hencky strain components and their conjugate stress components. This approach was shown to be accurate for moderate deformation and for a variety of materials and loading conditions. This type of formulation is advantageous in that the form of the strain energy function and the elastic constants therein are the same as those of classical linear elasticity. Moreover, this formulation is consistent regarding the volume change term. If an approach similar to this were undertaken, only with Green strain components, we should not expect the results to correlate well with experimental data because the volume change term would be incorrect for large strain. Despite our knowledge that this approach is inconsistent for large strain, we believe that this limitation of the so-called geometrically nonlinear theory of elasticity is not widely understood. Therefore, we present results from it for comparison with the Hencky strain energy and with two other consistent approaches described below.

Strain energy functions (or alternatively, constitutive laws) other than the one on which classical linear elasticity theory is based can be formulated that yield a consistent expression for the volume change term (see, for example, Ogden, 1984). An analysis based on a modified Green strain energy function, when restricted to the case of an isotropic, incompressible material, yields results that are identical to those from a Neo-Hookean formulation. It results in a constitutive law that *nonlinearly* relates the Cauchy-Euler stress components to the Green strain components but with the two material constants of isotropic linear elasticity. An alternative strain energy function is proposed in the present work, written in terms of the inverse of the deformation gradient. It results in a constitutive law that relates the Cauchy-Euler stress components to the Almansi strain components nonlinearly and with the same two material constants as the modified Green strain energy function.

In this paper, tensile force and effective torsional stiffness versus elongation are compared theoretically and experimentally. Experimental results are compared with theoretical results obtained from the small-strain version (inconsistent because we apply it to a large-strain problem) of the Green strain energy approach, from the Hencky strain energy approach, from a modified Green strain energy function, and from a modified Almansi strain energy function. The modifications for the last two formulations result in a consistent treatment of large strains which is easily checked against the universal relation of Rivlin. Unlike most of the work cited above, Poisson contraction effects are considered.

2 Theoretical Formulations

In this section, after some general discussion of the coordinate systems and strain definitions, we present a basis for the various theoretical formulations discussed above. These formulations are developed in the context of the principle of virtual work and utilizing the concept of strain energy and its relationship to the constitutive equations. Subsequent definition of the displacement field for the beam problem allows us to obtain explicit expressions for the tensile force and torsional moment versus elongation for each of the formulations.

2.1 Coordinate Systems. Consider Cartesian coordinate systems x_i, associated with the undeformed (reference) state, and y_i, associated with the deformed state. The difference between them is given by the displacement components u_i expressed in the undeformed state basis such that $y_i = x_i + u_i$. The components of the deformation gradient are obtained by differentiation of the deformed state coordinates with respect to the undeformed state coordinates. Expressed in the undeformed state bases, this yields

$$\frac{\partial y_i}{\partial x_j} = C_{ij} = \delta_{ij} + \frac{\partial u_i}{\partial x_j} \tag{1}$$

The inverse of the deformation gradient is also possible by differentiation of the undeformed state coordinates with respect to the deformed state coordinates. Thus,

$$\frac{\partial x_i}{\partial y_j} = B_{ji} = \delta_{ij} - \frac{\partial u_i}{\partial y_j} \tag{2}$$

All the formulations to be used below make use of the matrix of deformation gradient components, C, or its inverse, B^T. Note also that $BC^T = CB^T = I$ where I is the identity matrix. These matrices can be determined only after the displacement field is specified.

2.2 Strain Definitions. The typical geometrically nonlinear analysis makes use of the Green strain components (see Wempner, 1981). The matrix of the deformation gradient components times its transpose yields the matrix of metric coefficients from which the Green strain components are formed by subtraction of the metric of the undeformed state, the identity matrix. This operation can be expressed in matrix form

$$\gamma = \frac{1}{2}(C^T C - I) = \frac{1}{2}(U^2 - I) \tag{3}$$

where U is the matrix of right stretch components. From this expression it is easy to define quite a general family of strains which is, for example, given by Ogden (1984) as

$$e^{(m)} = \frac{1}{m}(U^m - I) \qquad m \neq 0 \tag{4}$$

The value of m may now be chosen freely except that $m = 0$ yields the so-called Hencky strain

$$\epsilon = \ln(U) = \frac{1}{2}\ln(C^T C) \tag{5}$$

Another strain measure that is pertinent to the following development is the Almansi strain[1] for which $m = -2$. The Almansi strain can also be expressed in terms of the inverse of the deformation gradient so that

$$\begin{aligned} \alpha &= -\frac{1}{2}(U^{-2} - I) = \frac{1}{2}(I - U^{-2}) \\ &= \frac{1}{2}(I - B^T B) \end{aligned} \tag{6}$$

In short, these three strains can be characterized as follows: Green strain is related to the reference state; Hencky strain incorporates the instantaneous rate of change of the deformation resulting in the logarithmic function; Almansi strain deals with the inverse of the deformation gradient which is gained by differentiation with respect to the deformed state coordinates (see equation (2)).

2.3 Virtual Work, Strain Energy, and Hooke's Law. The virtual work of internal forces must balance that of external forces. The virtual work of internal forces is derived from a strain energy function whose density per unit volume of undeformed structure is denoted by W. The virtual work of external forces is derived from applied tip loads T, the tensile

[1] As noted by Ogden (1984), this definition of Almansi strain is Lagrangean since it is based on U. The Eulerian version of Almansi strain is given by formulas identical to equations (6) only with U replaced by the matrix of left stretch tensor components V where $CC^T = V^2$ and with $B^T B$ replaced by BB^T.

force, and M, the twisting moment. The expression of the principle of virtual work is then given by

$$\iiint_V \delta W dx_1 dx_2 dx_3 = \int_0^\ell (T\delta u' + M\delta\theta')dx_1 \tag{7}$$

where V is the volume of the beam in its undeformed state, ℓ is the length of the beam in its natural state, u is the axial displacement, θ is the torsional rotation, and $(\ )'$ denotes derivative with respect to x_1.

If the small-strain formulation found in engineering texts (see Wempner, 1981, for example) is applied to this problem, the components of the second Piola-Kirchhoff stress tensor S_{ij} are expressed as linear functions of the components of γ_{ij}. For isotropic materials, the resulting strain energy density can be expressed as

$$W = G\left(\gamma_{ij}\gamma_{ij} + \frac{\nu\gamma_{kk}^2}{1-2\nu}\right) \tag{8}$$

where G is the shear modulus, ν is Poisson's ratio, and with the Young's modulus $E = 2(1 + \nu)G$. For this strain energy it is clear that the components of the second Piola-Kirchhoff stress tensor

$$S_{ij} = \frac{\partial W}{\partial \gamma_{ij}} = 2G\left(\gamma_{ij} + \frac{\nu\gamma_{kk}}{1-2\nu}\delta_{ij}\right) \tag{9}$$

are linear in the components of the Green strain tensor giving an expression of Hooke's law identical in form to that of linear elasticity. Equations (8) and (9) form the basis of what is commonly termed the geometrically nonlinear theory of elasticity.

Equations (8) and (9) are not valid for incompressible materials, because in this case the strain components cease to be independent quantities. Presently, however, we intend to apply these equations to a problem that involves only uniaxial stress, and in this case they are valid subject to certain restrictions. The denominator of the second term of these equations vanishes when $\nu = 1/2$ in which case the volume should also remain constant. Thus, the numerator of this term must vanish when describing deformation for which the volume remains constant in order to keep the term from becoming infinite. The difficulty is, however, *that the trace of the strain components γ_{kk} is not equal to the volume change unless the strains are small or the strain measure is logarithmic*—a fact that is not frequently pointed out in the literature.

It should be clear that there are other strain energy functions—still based on the same elastic constants—which may be useful in some contexts. In addition to one based on equation (8) that is inconsistent for large strain, we consider three other strain energy formulations: (1) a function identical in form to equation (8) except that instead of being written in terms of the Green strain components, γ, we write it in terms of the Hencky strain components ϵ as suggested by Anand (1979); (2) a modified Green strain energy function from Ogden (1984); and (3) a modified Almansi strain energy function.

These formulations deserve some explanation. First, the Hencky strain energy is based on the strain energy function

$$W = G\left(\epsilon_{ij}\epsilon_{ij} + \frac{\nu\epsilon_{kk}^2}{1-2\nu}\right) \tag{10}$$

Since the Hencky (logarithmic) strain is used, its conjugate stress components will contain $\epsilon_{kk} = \ln J$ where $J = \det(C)$, representing the volume change exactly for large strains. Thus,

$$\frac{\partial W}{\partial \epsilon_{ij}} = 2G\left(\epsilon_{ij} + \frac{\nu \ln J}{1-2\nu}\delta_{ij}\right) \tag{11}$$

Now when the material is incompressible, the denominator vanishes, and the numerator reflects the unchanging volume in the simultaneous vanishing of $\ln J$.

Ogden (1984) has given a constitutive law in terms of Green strain for large strains in which the volume change term behaves appropriately for incompressible materials; viz,

$$J\tau_{ij} = 2G\left[\gamma_{ij} + \frac{\nu}{1-2\nu}J(J-1)\delta_{ij}\right] \tag{12}$$

where τ is the component matrix of the Cauchy-Euler stress tensor and $J\tau$ is the weighted Cauchy-Euler stress tensor sometimes called the Kirchhoff stress tensor (but not to be confused with the second Piola-Kirchhoff stress tensor (see Ogden, 1984). This law can be derived from a strain energy function, for simplicity written in terms of principal elongations β_i, as

$$W = \frac{G}{2}(\beta_i\beta_i - 3 - 2\ln J) + \frac{G\nu}{1-2\nu}(J-1)^2 \tag{13}$$

Note that when specialized for incompressible materials, equation (13) reduces to the strain energy function associated with Neo-Hookean materials. Also, when specialized for small strains, equation (13) reduces to the form of equation (8), only with the strain measure being that associated with $m = 1$ from equation (4).

Motivated by the relationship of the constitutive law used in Petersen (1982) to the usual one in small-strain formulations, we propose a law in terms of Almansi strain components

$$J\tau_{ij} = 2G\left[\alpha_{ij} + \frac{\nu}{1-2\nu}K(1-K)\delta_{ij}\right] \tag{14}$$

where $K = 1/J$. This law can also be derived from a strain energy function. For simplicity, we write this strain energy function in terms of inverse principal elongations $\kappa_i = 1/\beta_i$ so that

$$W = \frac{G}{2}(\kappa_i\kappa_i - 3 - 2\ln K) + \frac{G\nu}{1-2\nu}(K-1)^2 \tag{15}$$

When specialized for small strains, equation (15) reduces to the form of equation (8), only with the strain measure being that associated with $m = -1$ from equation (4).

After the beam displacement field is obtained, results from these four formulations (i.e., equations (8), (10), (13), and (15)) are determined in Sections 2.5. 2.6, 2.7, and 2.8, respectively.

2.4 Beam Displacement Field. In this paper we propose to answer the question of how the effective torsional stiffness of a beam will be changed by a large elongation. Therefore, in all four of the formulations considered, a small torsion is superimposed on the elongated beam which will yield the influence of the elongation on the effective torsional stiffness. The position of a given point in the beam's cross section is represented by a vector:

$$\mathbf{R} = \left(x_1 + u + \frac{\lambda\theta'}{1+u'}\right)\hat{\mathbf{e}}_1 + x_2(1+f)\hat{\mathbf{e}}_2 + x_3(1+f)\hat{\mathbf{e}}_3 \tag{16}$$

The basis $\hat{\mathbf{e}}_i$ is related to the reference basis $\mathbf{e}_i$ by a rotation due to torsion.

$$\begin{Bmatrix}\hat{\mathbf{e}}_1\\ \hat{\mathbf{e}}_2\\ \hat{\mathbf{e}}_3\end{Bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{bmatrix}\begin{Bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\\ \mathbf{e}_3\end{Bmatrix} \tag{17}$$

The axial displacement of a given section point is composed of the axial displacement u and the warping term of Vlasov (1961) $\theta'\lambda/(1 + u')$ where λ is the warping function, a section property. An auxiliary function f is introduced to determine the lateral contraction, which cannot be regarded as negligible for large elongations. The deformation gradient components defined in equation (1) in fixed bases $(\mathbf{e}_i)$ can be expressed in mixed bases to obtain a simpler expression without any loss of generality. Making use of equations (16) and (17) and

calculating the component matrix C such that the tensor is given by $\hat{\mathbf{e}}_i C_{ij} \mathbf{e}_j$, we obtain

$$C = \begin{bmatrix} 1+u' & \dfrac{\theta'\lambda_2}{1+u'} & \dfrac{\theta'\lambda_3}{1+u'} \\ -\theta' x_3(1+f) & 1+f & 0 \\ \theta' x_2(1+f) & 0 & 1+f \end{bmatrix} \tag{18}$$

where $\lambda_\alpha = \partial\lambda/\partial x_\alpha$. It is assumed that the applied tensile force and torsional moment are constant over the length of the beam and that f is independent of the section coordinates. Thus, $u'' = \theta'' = f' = \partial f/\partial x_2 = \partial f/\partial x_3 = 0$. Furthermore, only the axial displacement u is considered large. The torsion is supposed to be infinitesimally small, and therefore squares of θ' are neglected below except where noted.

2.5 Green Strain Energy—Small-Strain Analysis. Geometrically nonlinear analysis of structures is usually formulated in terms of the Green strain energy based on equation (8). This formulation is restricted to small strain problems, but we include it for comparison with consistent large-strain analyses. The Green strain components are

$$\begin{aligned} \gamma_{11} &= u' + \frac{1}{2}u'^2 + \frac{1}{2}(1+f)^2(x_2^2+x_3^2)\theta'^2 \\ \gamma_{12} &= \gamma_{21} = \frac{1}{2}\theta'[\lambda_2 - x_3(1+f)^2] \\ \gamma_{13} &= \gamma_{31} = \frac{1}{2}\theta'[\lambda_3 + x_2(1+f)^2] \\ \gamma_{22} &= \gamma_{33} = f + \frac{1}{2}f^2 \\ \gamma_{23} &= \gamma_{32} = 0 \end{aligned} \tag{19}$$

The variation of the longitudinal strain component must include all terms that are linear in θ' in order to pick up all linear terms in the final expression for the torsional moment. These terms can only be obtained from expressions containing quadratic terms in θ', which is the reason for the θ'^2 term being retained in γ_{11}. This term can be ignored in determining the auxiliary function.

The auxiliary function f can be determined from the condition that no lateral stress can occur such that

$$\gamma_{22} = \gamma_{33} = -\nu\gamma_{11} \tag{20}$$

The auxiliary function then turns out to be

$$1+f = \sqrt{1-2\nu\left(u' + \frac{1}{2}u'^2\right)} \tag{21}$$

which will result in a changed volume given by the determinant of C, again ignoring θ', from equation (18)

$$J = \det(C) = (1+u')(1+f)^2 = (1+u')\left[1-2\nu\left(u'+\frac{1}{2}u'^2\right)\right] \tag{22}$$

One indication of the inadequacy of this approach is now obvious since for $\nu = 1/2$ we do not obtain $J = 1$ as we should. It is of interest, however, to proceed with this formulation and to compare its predictions with the other three in this paper. It is possible to add a constraint to this formulation and determine f so that $J = 1$, but then the lateral stresses would not vanish as they should. The present form of this approach is retained so that it is parallel to the others as much as possible.

Now the variation of the strain energy can be evaluated with the Green strains. Comparison with equation (7) yields the relations for tensile force and torsional stiffness versus elongation. Here only warp-free sections, with $\lambda = 0$ and Saint-Venant torsion constant equal to polar moment of inertia I_p, are considered. The results are given by

$$\frac{T}{EA} = \left(u' + \frac{1}{2}u'^2\right)(1+u') \tag{23}$$

$$\frac{M}{GI_p\theta'} = (1+u')^2\left[1-2\nu\left(u'+\frac{1}{2}u'^2\right)\right] \tag{24}$$

For $\nu = 1/2$ and $E = 3G$, the above-mentioned universal relation of Rivlin requires that T and M satisfy

$$\frac{TI_p\theta'}{MA} = (1+u')^2 - (1+u')^{-1} \tag{25}$$

Upon substituting T from equation (23) and M from equation (24) into Rivlin's relation from equation (25), we find that this formulation does not satisfy Rivlin's relation unless $u' \ll 1$. This is another reflection of the inadequacy of the small-strain formulation to accurately account for large strains. Note that when $u' \ll 1$, we obtain $T = EAu'$ and $M = GI_p\theta'$ in agreement with Petersen (1982) and Hodges (1987).

2.6 Hencky Strain Energy Analysis. The strain definition of equation (5) is now applied in the same, simple strain energy expression used in the previous section except that ϵ_{ij}, the Hencky logarithmic strain tensor components, are used as in equation (10). Anand (1979) found this formulation to be very good in characterizing the real behavior of several materials and loading conditions.

With the aid of Lagrange interpolation polynomials (see Lancaster, 1969), the Hencky strains can be evaluated in terms of the three distinct eigenvalues of C^TC. The algebraic manipulation is rather cumbersome, and symbolic manipulation with the program MACSYMA was used. To save space, the details of the algebraic operations are not shown here. As with the small-strain analysis, a term quadratic in θ' is present in the longitudinal component of strain that can be ignored in determining the auxiliary function. For zero lateral stresses, equation (11) can be used to show that

$$\epsilon_{22} = \epsilon_{33} = \ln(1+f) = -\nu\epsilon_{11} = -\nu\ln(1+u') \tag{26}$$

Then the auxiliary function to describe the lateral contraction is found to be

$$1+f = (1+u')^{-\nu} \tag{27}$$

This equation yields the correct volume change for all values of ν. Now straightforward evaluation of the strain energy from equation (7) yields the relations for tensile force and torsional moment versus elongation. As with the previous strain energy approach above, in order to obtain the complete expression for the torsional stiffness versus elongation, it is necessary and sufficient to retain terms through the first degree in θ' from the shear strains and terms through the second degree in θ' from the longitudinal strain. The resulting tensile force versus elongation is independent of ν and is given by

$$\frac{T}{EA} = \frac{\ln(1+u')}{1+u'} \tag{28}$$

Consider the axial component of the physical stress, the tensile force T divided by the current area $(1+f)^2 A$. For the case of an incompressible material only, equation (28) shows that this stress measure is linear in the axial component of the logarithmic strain measure.[2]

[2]This observation was pointed out by Professor C. V. Smith, Jr., in private communication with the second author on April 15, 1987.

Table 1 Dimensions and material properties of test specimens

Specimen (Symbol)	# 1 (Δ)	# 2 (×)	# 3 (◇)
Outer Diameter d_o	1.52 cm	1.55 cm	1.77 cm
Inner Diameter d_i	0.56 cm	0.50 cm	0.80 cm
Cross-sectional Area A	1.58 cm^2	1.69 cm^2	1.95 cm^2
Length ℓ	61.5 cm	71 cm	66 cm
Density ρ	1.22 gm/cm^3	1.12 gm/cm^3	1.09 gm/cm^3
Modulus of Elasticity E	180 N/cm^2	390 N/cm^2	300 N/cm^2
Poisson's Ratio ν	0.5	0.5	0.5

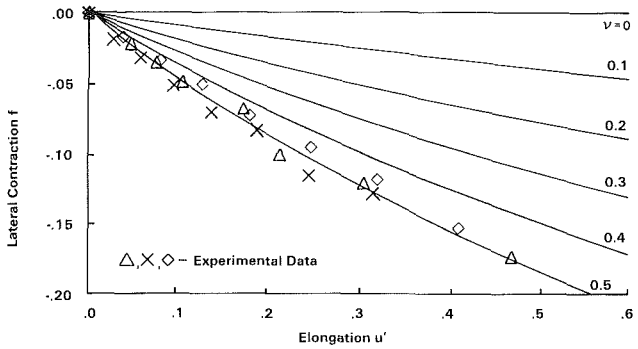


Fig. 1 Lateral contraction f versus elongation u' for various values of Poisson's ratio ν

The torsional stiffness versus elongation is

$$\frac{M}{GI_p\theta'}=\frac{2(1+\nu)\ln(1+u')}{(1+u')^{2(1+\nu)}-1} \tag{29}$$

These results obey the universal relation of Rivlin from equation (25) for $\nu = 1/2$.

2.7 Green Strain Energy—Modified Large-Strain Analysis. The material behavior of rubber is conveniently represented by an elastic potential of strain invariants. Several developments based on this concept are discussed by Treloar (1976). Here we present these results in the context of the modified Green strain energy function given in equation (13). Straightforward application of the principle of virtual work from equation (7), using the strain energy function in equation (13) and setting the lateral stresses equal to zero, leads to an expression for the auxiliary function

$$(1+f)^2=\frac{1}{2(1+u')}\left[1-\frac{1-2\nu}{2\nu(1+u')}\right]$$

$$+\frac{1}{2(1+u')}\sqrt{\left[1-\frac{1-2\nu}{2\nu(1+u')}\right]^2+\frac{2(1-2\nu)}{\nu}} \tag{30}$$

This equation yields constant volume for $\nu = 1/2$. The tensile force obtained is

$$\frac{T}{EA}=\frac{1}{2(1+\nu)}\left[1+u'-\frac{(1+f)^2}{(1+u')}\right] \tag{31}$$

and the effective torsional stiffness is

$$\frac{M}{GI_p\theta'}=(1+f)^2 \tag{32}$$

For $\nu = 1/2$ these equations satisfy the universal relation of Rivlin from equation (25) and correspond to published results for Neo-Hookean materials in Treloar (1976). He commented on this theory as follows: "There remained, however, some outstanding problems, the most serious of which was probably the deviation in the observed form of the force-extension curve for rubber in simple extension from that predicted by the statistical theory. Rather surprisingly, this particular difficulty has remained substantially unresolved over the subsequent quarter of a century, despite numerous tentative claims for its explanation." He discussed several attempts of improvement and concluded that ". . . that any advances in the methods of analysis and presentation of the actual mechanical properties of rubber . . . should nevertheless provide a more helpful basis for the attack on this more difficult, and so far unsolved, problem."

2.8 Almansi Strain Energy—Modified Large-Strain Analysis. As explained earlier, motivated by Petersen's attempt to represent the behavior of beams by a constitutive law in terms of Almansi's strain components, we formulated the modified Almansi strain energy function from equation (15). Straightforward application of the principle of virtual work from equation (7), using the strain energy function in equation (15) and setting the lateral stresses equal to zero, leads to an expression for the auxiliary function

$$\frac{1}{(1+f)^2}=\frac{(1+u')}{2}\left[1-\frac{1-2\nu}{2\nu}(1+u')\right]$$

$$+\frac{(1+u')}{2}\sqrt{\left[1-\frac{1-2\nu}{2\nu}(1+u')\right]^2+\frac{2(1-2\nu)}{\nu}} \tag{33}$$

This equation yields constant volume for $\nu = 1/2$. The tensile force obtained is

$$\frac{T}{EA}=\frac{1}{2(1+\nu)}\left[1-\frac{(1+f)^2}{(1+u')^2}\right] \tag{34}$$

and the effective torsional stiffness is

$$\frac{M}{GI_p\theta'}=\frac{(1+f)^2}{1+u'} \tag{35}$$

For $\nu = 1/2$ these equations also satisfy the universal relation of Rivlin from equation (25).

3 Comparison of Theories and Experiment

In this section the experimental investigation is described along with the comparison of the results with those of the various theoretical formulations. Here we compare the tensile force and effective torsional stiffness versus elongation from theory and experiment.

3.1 Description of Experiment. Experimental investigations were performed to verify the theoretical approaches. For these tests circular (warp-free) soft rubber tubes were used in order to allow high strain levels. To obtain a general validity of the test results, different types of rubber materials were investigated. See Table 1 for a detailed description of the stiffness, mass, and geometric properties of each of the beams tested. The values of the geometric properties given in Table 1 are those of the undeformed beams. The symbols used in tabulation of the experimental results correspond to the beams according to the legend shown in the figures. To perform the tests the rubber tubes were suspended vertically, and different tensile forces and torsional moments were applied. The elongation, the cross section diameter, and the torsional angle were measured.

To correlate the above theoretical expressions for the tensile force and torsional stiffness versus elongation with the experimental results, the constants for homogeneous and

isotropic materials, the modulus of elasticity and Poisson's ratio, had to be determined. The modulus of elasticity was measured at very small strains—in the range of applicability for Hooke's law. For this purpose an electro-optical displacement transducer was used, which allows highly accurate displacement measurements. The modulus of elasticity was measured very reliably by this device at small strains of $u' < 0.01$. The materials exhibit values of the moduli between 180 and 390 N/cm^2. All measurements were performed statically to avoid viscoelastic effects.

Poisson's ratio was evaluated by plotting the exact lateral contraction from the Hencky strain energy approach, determined from equation (27), versus elongation. In Fig. 1 these data are compared with the theoretical curves for various values of Poisson's ratio. The figure shows that for all specimens Poisson's ratio is one half; that is, the rubber materials are indeed incompressible.

3.2 Tensile Force Versus Elongation. The above analyses result in nonlinear relations between the tensile force applied to the beam and the elongation. Due to the various formulations of the strain energy function, we obtain correspondingly different expressions as seen in comparison of equations (23), (28), (31), and (34).

Figure 2 shows the results for the tensile force versus elongation. The experimental results agree very well with the theoretical approaches based on the Hencky strain energy and the modified Almansi strain energy formulations. Taking into account that the experimental data may have an error margin of approximately 5 percent, both formulations seem to be adequate. The Hencky strain energy and the modified Almansi strain energy analyses predict a slope of nearly the same magnitude. On the other hand, the small-strain approach yields an excessively high tensile force for moderately large strains. The modified Green strain energy (neo-Hookean) results agree only fairly well and tend toward a straight line for $u' > 0.3$, which means a constant stiffness under large strain.

3.3 Effective Torsional Stiffness Versus Elongation. The analysis was restricted to small torsion to avoid a complicated derivation which would only serve to make the desired insight more difficult. Our original aim was to study the effect of tensile force on torsional rigidity for beams with warp-free cross sections since there the analytical differences are more noticeable. The various expressions for the change of the torsional stiffness are given in equations (24), (29), (32), and (35).

Experimental results are shown along with theoretical predictions for the effective torsional stiffness versus elongation in Fig. 3. Although there is some scatter in the experimental data for small strains, both the Hencky strain energy and modified Almansi strain energy formulations are again well confirmed by the experimental data. The small-strain analysis based on the Green strain energy formulation shows an increasing torsional stiffness for small and moderately large elongations, which contradicts the experimental results. Although the modified Green strain energy (neo-Hookean) results show a decrease of effective torsional stiffness, the descent of this curve is less steep than those of the Hencky and modified Almansi strain energy approaches and does not agree as well with the experimental data.

In the case where $\lambda = 0$ and the Saint-Venant torsional constant is equal to the polar moment of inertia, the torsional stiffness decreases with increasing elongations or axial tensile force. This contradicts the classical theories of Wagner (1929) and Biot (1939) under the assumption that they might be applied to beams with a closed as well as an open cross section. It also contradicts qualitative predictions from previous developments based on a small-strain approach when ex-

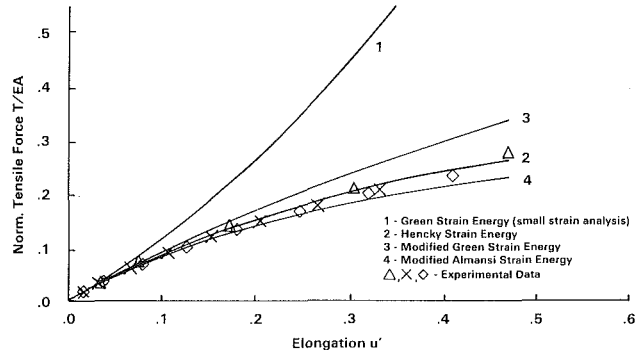


Fig. 2 Normalized tensile force *T/EA* versus elongation u' for various formulations

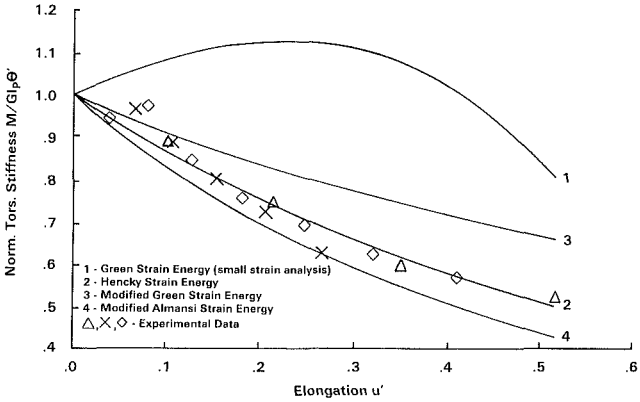


Fig. 3 Normalized effective torsional stiffness $M/GI_p\theta'$ versus elongation u' for various formulations

trapolated to this large-strain problem, such as Houbolt and Brooks (1958), Hodges and Dowell (1974), Friedmann (1977), and Hodges (1980), which predict an increase of torsional stiffness for an axial tensile force, and Petersen (1982) and Hodges (1987), which predict no change in effective torsional stiffness versus tensile force. Evidently, there are two reasons that these analyses fail to predict the correct trend in the effective torsional stiffness. One is the choice of constitutive law; all of these works except Petersen (1982) and Hodges (1987) use an approach that is equivalent to the small-strain analysis based on Green strain energy. The other, which applies more to Petersen (1982) and Hodges (1987) than to the others, is the neglect of Poisson contraction effects.

4 Conclusion

The aim of this investigation was to examine four nonlinear formulations and how well they predict the behavior of a beam that can sustain large strains. As a first step, only uniform beams of circular cross section and incompressible materials were considered. The behavior of the axial tensile force and effective torsional stiffness versus axial elongation was studied. A strain energy approach in terms of the Hencky logarithmic strain was shown to satisfy the universal relation of Rivlin, as given in Truesdell (1966) and Green and Zerna (1968), to correctly yield constant volume for an incompressible material, and to yield excellent correlation with the experimental data for lateral contraction, tensile force, and effective torsional stiffness when $u' < 1$. This formulation is advantageous in that both the form of the strain energy expression and the constants therein are identical to those of classical linear elasticity.

On the other hand, a small-strain analysis based on Green

strain energy (commonly referred to as the geometrically nonlinear theory of elasticity and equivalent to the principle of virtual work and a material law relating the second Piola-Kirchhoff stress components and Green strain components) does not predict a constant volume during deformation for the incompressible material. Furthermore, the results for tensile force and effective torsional stiffness do not agree with the experimental data and do not satisfy the universal relation of Rivlin. Known limitations in this theory that restrict its validity to small strains are highlighted.

We also considered the efficacy of strain energy formulations that are modified to consistently treat large strains. Results from the modified Green strain energy approach (for the special case we consider, this is equivalent to a Neo-Hookean strain energy function) satisfy the universal relation of Rivlin but agree only fairly well with experimental data. Results from a modified Almansi strain energy analysis also satisfy the universal relation of Rivlin but agree very well with experimental data for tensile force and effective torsional stiffness. This approach appears to be suitable for this problem.

In addition to their impact on the determination of a suitable stress-strain formulation in geometrically nonlinear analyses, the present results show that small-strain analyses of beams are not correct in predicting the trends of tensile force and effective torsional stiffness for large elongations. The comparisons above show that only advanced concepts which include nonlinearities in a more complicated manner and take lateral contraction into account are reliable for problems in which the strains are not small. The present paper confirms that the predicted effective torsional stiffness of a beam depends on the assumed constitutive law as pointed out by Hodges (1987). Even for problems in which the strains are small, some modification of the usual small-strain approach based on the geometrically nonlinear theory of elasticity may be required to accurately predict the effective torsional stiffness for transversely isotropic beams in which $G \ll E$.

For further applications that involve large strains in structural analysis, the Hencky strain energy and the modified Almansi strain energy formulations are contenders. The modified Green strain energy function is closely related to the Green strain measure. Within the scope of this investigation its predictions are less close to the experimental data. The Hencky strain energy involves transcendental functions of matrices which can be rather complicated. Indeed, the algebra was complicated enough that symbolic manipulation by computer was considered essential. Alternatively, the modified Almansi strain energy formulation involves only algebraic expressions which are easier to handle, and considerations of higher-order terms are simpler. It is recommended that a more general problem be studied to see how well the Hencky strain energy and the modified Almansi strain energy formulations perform.

Acknowledgments

The authors are grateful for the Memorandum of Understanding between the Aeroflightdynamics Directorate, U.S. Army Aviation Research and Technology Activity (AVSCOM), Ames Research Center, Moffett Field, California, and the DFVLR centers in West Germany where the first and third authors are affiliated. The present work was conceived and initiated in 1984 when the second author (who at that time was Research Scientist and Theoretical Group Leader, Rotorcraft Dynamics Division, Aeroflightdynamics Directorate) was guest scientist at the Institute for Structural Mechanics in Braunschweig, West Germany. The second author was also partially supported by grant E-16-697 from the Georgia Tech Research Corporation. Technical discussions with Professors S. Hanagud and C. V. Smith, Jr., are gratefully acknowledged. The authors express their thanks to one of the referees whose constructive comments helped to improve the paper.

References

Anand, L., 1979, "On H. Hencky's Approximate Strain-Energy Function for Moderate Deformations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 78–82.

Biot, M. A., 1939, "Increase of Torsional Stiffness of a Prismatic Bar Due to Axial Tension," *J. Appl. Phys.*, Vol. 10, pp. 860–864.

Friedmann, P., 1977, "Influence of Structural Damping, Preconing, Offsets and Large Deflections on the Flap-Lag-Torsional Stability of a Cantilevered Rotor Blade," *AIAA Journal*, Vol. 15, No. 2, pp. 149–158.

Green, A. E., and Zerna, W., 1968, *Theoretical Elasticity*, Second Ed., Clarendon Press, Oxford, equations (3.5.9) and (4.3.31).

Hodges, D. H., 1980, "Torsion of Pretwisted Beams Due to Axial Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 393–397.

Hodges, D. H., 1987, "Nonlinear Beam Kinematics for Small Strains and Finite Rotations," *Vertica*, Vol. 11, No. 3, pp. 573–589.

Hodges, D. H., and Dowell, E. H., 1974, "Nonlinear Equations of Motion for the Elastic Bending and Torsion of Twisted Nonuniform Rotor Blades," NASA TN D-7818.

Houbolt, J. C., and Brooks, G. W., 1958, "Differential Equations of Motion for Combined Flapwise Bending, Chordwise Bending, and Torsion of Twisted Nonuniform Rotor Blades," NACA Report 1346.

Lancaster, P., 1969, *The Theory of Matrices*, Academic Press, New York, Ch. 5.

Ogden, R. W., 1984, *Non-Linear Elastic Deformations*, Halsted Press, Wiley, New York, Chapters 3 and 4.

Petersen, D., 1982, "Interaction of Torsion and Tension in Beam Theory," *Vertica*, Vol. 6, pp. 311–325.

Treloar, L. R. G., 1976, "The Mechanics of Rubber Elasticity," *Proceedings of the Royal Society*, London, A 351, pp. 301–330.

Truesdell, C., 1966, *The Elements of Continuum Mechanics*, Springer-Verlag, Berlin, equation (13.23).

Vlasov, V. Z., 1961, *Thin-Walled Elastic Beams*, National Science Foundation, Washington, D.C.

Wagner, H., 1929, "Verdrehung und Knickung von Offenen Profilen," *Festschrift 25 Jahre TH Danzig*, pp. 329–343, English translation: NACA TM 807, Washington, D.C., 1936.

Washizu, K., 1968, *Variational Methods in Elasticity and Plasticity*, 2nd Ed., Pergamon Press, Oxford, Chapter 4.

Wempner, G. A., 1981, *Mechanics of Solids with Applications to Thin Bodies*, Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands, Chapters 4 and 9.

A Beam Theory for Large Global Rotation, Moderate Local Rotation, and Small Strain

D. A. Danielson
Associate Professor,
Department of Mathematics,
Naval Postgraduate School,
Monterey, CA 93943
Mem. ASME

D. H. Hodges
Professor,
School of Aerospace Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332

Kinematical relations are derived to account for the finite cross-sectional warping occurring in a beam undergoing large deflections and rotations due to deformation. The total rotation at any point in the beam is represented as a large global rotation of the reference triad (a frame which moves nominally with the reference cross section material points), a small rotation that is constant over the cross section and is due to shear, and a local rotation whose magnitude may be small to moderate and which varies over a given cross section. Appropriate variational principles, equilibrium equations, boundary conditions, and constitutive laws are obtained. Two versions are offered: an intrinsic theory without reference to displacements, and an explicit theory with global rotation characterized by a Rodrigues vector. Most of the formulas herein have been published, but we reproduce them here in a new concise notation and a more general context. As an example, the theory is shown to predict behavior that agrees with published theoretical and experimental results for extension and torsion of a pretwisted strip. The example also helps to clarify the role of local rotation in the kinematics.

I Introduction

We are concerned here with the general theory of beams or rods undergoing large deflections but small strains. In Hodges (1987b) and Danielson and Hodges (1987), a kinematical basis for beam theory is derived based on the concept of decomposition of the rotation tensor. The beam cross section is postulated to displace and rotate as a rigid body without explicit restriction on the magnitude of the motion. Further deformation of the cross section is characterized by quantities whose magnitude is small in some sense. In particular, the total rotation at any point in the beam is represented as: (1) a large global rotation of the reference triad, which is a frame whose rotation captures all rigid body motion of the reference cross section; (2) a rotation due to shear which is constant over the cross section but which does *not* represent a rigid-body rotation of the reference cross section; and (3) a local rotation whose magnitude may be small to moderate, which varies over a given cross section, and which vanishes at the reference triad.

In Section II of this paper we review the kinematical relationships derived in our preceding papers, Hodges (1987b) and Danielson and Hodges (1987). Background on the mathematics used may be obtained from Simmonds (1982) and Ogden (1984). In Section III we develop the remaining equations of the theory in terms of generalized strains. Most of the results in this section are well known and may be found in Wempner (1981). In Section IV we develop an alternate formulation in terms of generalized displacements. Our work in this section has been based on that of Simmonds and Danielson (1972) and Reissner (1973).

In Section V we apply the theory to the interesting problem of the torsion and extension of an initially twisted thin strip. The solution to the problem of the torque of an initially untwisted strip is contained in Timoshenko (1976). The effect of pretwist was included by Chu (1951). The problem of the axial tension of an initially twisted beam was solved by Hodges (1980). All of these effects were included in the work by Rosen (1983), containing both theoretical and experimental results.

The following assumptions are made at successive stages of this paper:

(*i*) Each reference cross section does not distort in its plane.

(*ii*) The deformation normal to the plane of a cross section is proportional to the warping function of the Saint-Venant torsion theory for a prismatic beam of the same cross section.

(*iii*) The maximum magnitude of the strain is of order ϵ, a quantity that is negligible compared to unity.

(*iv*) The maximum magnitude of the local rotation is of order $\epsilon^{1/2}$; thus, its square is negligible compared to unity.

(*v*) The cross section dimensions are negligible compared to initial radii of curvature.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, January 27, 1987; final revision, July 6, 1987.

Paper No. 88-WA/APM-10.

(*vi*) "Distortion" shear and transverse normal stresses are ignored in the strain energy expression.

(*vii*) The material is linearly elastic and transversely isotropic.

Our general theory requires only assumptions (*iii*) and (*iv*). The other assumptions are imposed for the particular application of the general theory considered in this paper. These assumptions are adequate for a variety of beams with either open or closed cross sections, although assumptions (*i*) and (*vi*) may need to be relaxed for treatment of certain thin-walled beams with large initial curvature and twist (see Axelrad, 1973).

II Kinematics

Let x_1 denote length along a reference line r within an undeformed beam. Call each plane perpendicular to r a reference cross section, and let x_α denote lengths along orthogonal lines in each plane. Throughout the paper, Greek indices assume values 2 and 3, and Latin indices assume values 1, 2, and 3. At each point along r define an orthogonal reference triad $\mathbf{b}_i^r(x_1)$ tangent to the coordinate curves at r with $\mathbf{b}_1^r$ tangent to r. The position vector to points in the undeformed beam is then given by

$$\mathbf{r}(x_1,x_2,x_3)=\bar{\mathbf{r}}(x_1)+x_\alpha\mathbf{b}_\alpha^r \tag{1}$$

where $\bar{\mathbf{r}}(x_1)=\mathbf{r}(x_1,0,0)$ is the position vector to points on the reference line. The covariant base vectors $\mathbf{g}_i=\partial\mathbf{r}/\partial x_i$ are obtained by differentiation of equation (1) and use of the formulas

$$\begin{aligned}\bar{\mathbf{r}}'&=\mathbf{b}_1^r\\ (\mathbf{b}_i^r)'&=\mathbf{k}\times\mathbf{b}_i^r\end{aligned} \tag{2}$$

Here $\mathbf{k}=k_i\mathbf{b}_i^r$ is the curvature vector of the undeformed beam, as in Danielson and Hodges (1987), and primes denote differentiation with respect to x_1. The contravariant base vectors $\mathbf{g}^i$ can be obtained by standard means.

After deformation, the locus of material points on the reference line is denoted by R. Making assumption (*i*), we can introduce at each point along R an orthogonal reference triad $\mathbf{b}_i^R$ such that $\mathbf{b}_\alpha^R$ is tangent to the x_α coordinate curve at R. (The geometrical meaning of these vectors, along with other possible choices for the deformed beam base vectors, is discussed further in the Appendix.) With the addition of assumption (*ii*), the position vector to a particle in the deformed beam, which had position $\mathbf{r}(x_1,x_2,x_3)$ in the undeformed beam, is given by

$$\mathbf{R}(x_1,x_2,x_3)=\bar{\mathbf{r}}+\mathbf{u}+x_\alpha\mathbf{b}_\alpha^R+\lambda\psi\mathbf{b}_1^R \tag{3}$$

Here $\mathbf{u}(x_1)=u_i\mathbf{b}_i^r$ is the displacement vector of points on the reference line, $\lambda(x_2,x_3)$ is the Saint-Venant warping function for the local cross section, and $\psi(x_1)$ is the warp amplitude. The reference axis is chosen so that $\lambda(0,0)=\lambda_\alpha(0,0)=0$, where $\lambda_\alpha=\partial\lambda/\partial x_\alpha$. The covariant base vectors $\mathbf{G}_i=\partial\mathbf{R}/\partial x_i$ of the deformed beam are obtained by differentiation of equation (3) and use of the formulas

$$\begin{aligned}\bar{\mathbf{R}}'&=\mathbf{b}_1^R+\bar{\gamma}\\ (\mathbf{b}_i^R)'&=\mathbf{K}\times\mathbf{b}_i^R\end{aligned} \tag{4}$$

Here $\bar{\mathbf{R}}(x_1)=\mathbf{R}(x_1,0,0)=\bar{\mathbf{r}}+\mathbf{u}$ is the position vector to points on R, $\bar{\gamma}=\bar{\gamma}_{11}\mathbf{b}_1^R+2\bar{\gamma}_{1\alpha}\mathbf{b}_\alpha^R$ is the strain vector at the reference line, and $\mathbf{K}=K_i\mathbf{b}_i^R$ is the curvature vector of the deformed beam. Using assumption (*iii*), one can show that the components of $\mathbf{K}$ are the twist and curvatures of the deformed beam.

Following the analysis of Danielson and Hodges (1987), we define the deformation gradient tensor $\mathbf{A}$ by

$$\mathbf{A}=\mathbf{G}_i\mathbf{g}^i \tag{5}$$

The polar decomposition theorem states that $\mathbf{A}$ can be uniquely decomposed into an orthogonal rotation tensor $\mathbf{C}$ dotted into a symmetric right stretch tensor $\mathbf{U}$:

$$\mathbf{A}=\mathbf{C}\cdot\mathbf{U} \tag{6}$$

As in Hodges (1987b) and Danielson and Hodges (1987), we use the Jaumann-Biot-Cauchy strain tensor defined by

$$\boldsymbol{\Gamma}=\mathbf{b}_i^r\gamma_{ij}\mathbf{b}_j^r=\mathbf{U}-\mathbf{I} \tag{7}$$

where $\mathbf{I}$ is the identity tensor.[1] The rotation may be decomposed into three parts: (1) a large *global* rotation from the undeformed beam reference triad to the deformed beam reference triad ($\mathbf{C}^{Rr}$) which reflects a rigid-body rotation of the reference cross section; (2) a rotation of axes due to *shear* that is small and constant over the cross section ($\mathbf{C}^{SR}$) which does not reflect a rigid-body rotation of the reference cross section; and (3) an additional *local* rotation that varies over the cross section (characterized by a finite rotation vector $\boldsymbol{\phi}=\phi_i\mathbf{b}_i^R$ chosen to have direction along the axis of local rotation and amplitude of the angle of local rotation). Thus

$$\mathbf{C}(x_1,x_2,x_3)=e^{\boldsymbol{\phi}\times\mathbf{I}}\cdot\mathbf{C}^{SR}\cdot\mathbf{C}^{Rr} \tag{8}$$

where $\mathbf{C}(x_1,0,0)=\mathbf{C}^{SR}\cdot\mathbf{C}^{Rr}$. Note that with this definition, the local rotation vanishes at the reference line, as in Hodges (1987b). However, $\mathbf{C}^{SR}$ was included with the local rotation of Danielson and Hodges (1987).

Now in order to simplify the strain expression, we make assumptions (*iii*) and (*iv*). As shown by Hodges (1987b) and Danielson and Hodges (1987), the strain components can be expressed in terms of the matrix of mixed components of the deformation gradient $A_{ij}=\mathbf{b}_i^R\cdot\mathbf{G}_k\mathbf{g}^k\cdot\mathbf{b}_j^r$. Following these earlier works, we denote $E=\frac{1}{2}(A+A^T)-I$ where I is the three-by-three identity matrix and $\phi_i=-\frac{1}{2}e_{ijk}[A_{jk}-A_{jk}(x_1,0,0)]$ where e_{ijk} are components of the permutation tensor in a Cartesian coordinate system. Imposing assumption (*v*) and noting the symmetry of E, that $E_{\alpha\beta}=0$, and $\phi_1=0$, we obtain

$$\begin{aligned}E_{11}&=\bar{\gamma}_{11}+x_3\kappa_2-x_2\kappa_3+\lambda\psi'+\psi k_1(x_3\lambda_2-x_2\lambda_3)\\ 2E_{12}&=2\bar{\gamma}_{12}-x_3\kappa_1+\lambda_2\psi+\lambda\psi(\kappa_3+k_3)\\ 2E_{13}&=2\bar{\gamma}_{13}+x_2\kappa_1+\lambda_3\psi-\lambda\psi(\kappa_2+k_2)\\ \boldsymbol{\phi}&=\{[-x_2\kappa_1+\lambda_3\psi+\lambda\psi(\kappa_2+k_2)]\mathbf{b}_2^R\\ &\quad+[-x_3\kappa_1-\lambda_2\psi+\lambda\psi(\kappa_3+k_3)]\mathbf{b}_3^R\}/2\end{aligned} \tag{9}$$

where $\kappa_i=K_i-k_i$ is the *moment strain measure*. In light of assumptions (*vi*) and (*vii*), strain components $\gamma_{\alpha\beta}$ will not appear in the formulation. From Danielson and Hodges (1987), the remaining components are

$$\begin{aligned}\gamma_{11}&=E_{11}+(E_{12}+\bar{\gamma}_{12})\phi_3-(E_{13}+\bar{\gamma}_{13})\phi_2+\frac{1}{2}(\phi_2^2+\phi_3^2)\\ 2\gamma_{12}&=2E_{12}-E_{11}\phi_3\\ 2\gamma_{13}&=2E_{13}+E_{11}\phi_2\end{aligned} \tag{10}$$

When the local rotations are *moderate* (finite, as in assumption *iv*), all of the terms in equation (10) must be kept. When the local rotations are *small* (of the order of the strain), the terms involving ϕ_α in equation (10) should be dropped.

III Intrinsic Theory

It follows from assumptions (*v*), (*vi*), and (*vii*) that the total strain energy may be taken to be

$$W_I=\int\!\!\int\!\!\int_V\left(\frac{E\gamma_{11}^2}{2}+2G\gamma_{1\alpha}\gamma_{1\alpha}\right)dx_1dx_2dx_3 \tag{11}$$

[1] Using assumption (*iii*) one can show that all strain measures based on $\mathbf{U}$ are equivalent to equation (7). We do not base our definition on the left stretch tensor $\mathbf{V}$ where $\mathbf{A}=\mathbf{V}\cdot\mathbf{C}$, however, because the strain energy would then depend on $\mathbf{C}$ for anisotropic materials (see Malvern, 1969).

where V denotes the entire volume of the undeformed beam. Substituting equations (9) and (10) into equation (11), making use of assumption (*iii*) to drop terms $0(\epsilon^3)$, and performing the integrations over x_α, we can reduce the strain energy to the form

$$W_I = \int_a^b U(\bar{\gamma}_{1i}, \kappa_i, \psi, \psi')\, dx_1 \tag{12}$$

where a and b are the values of x_1 at each end of the beam. The strain energy density U is a function of the generalized strains ($\bar{\gamma}_{1i}$, κ_i, ψ, ψ'). The virtual work of the internal forces is obtained by taking the variation of equation (12):

$$\delta W_I = \int_a^b [\mathbf{F}\cdot(\mathbf{b}_1^R \delta\bar{\gamma}_{11} + 2\mathbf{b}_\alpha^R \delta\bar{\gamma}_{1\alpha}) + M\cdot(\mathbf{b}_i^R \delta\kappa_i) + B\delta\psi + D\delta\psi']\, dx_1 \tag{13}$$

Here we have defined the generalized forces

$$\mathbf{F} = \frac{\partial U}{\partial \bar{\gamma}_{11}}\mathbf{b}_1^R + \frac{1}{2}\frac{\partial U}{\partial \bar{\gamma}_{1\alpha}}\mathbf{b}_\alpha^R;\quad \mathbf{M} = \frac{\partial U}{\partial \kappa_i}\mathbf{b}_i^R$$
$$B = \frac{\partial U}{\partial \psi};\quad D = \frac{\partial U}{\partial \psi'} \tag{14}$$

We introduce in association with the virtual strains virtual translational and rotational displacements $\delta\mathbf{u}$ and $\delta\boldsymbol{\eta}$:

$$\delta\mathbf{u} = \delta\bar{\mathbf{R}};\quad \delta\mathbf{b}_i^R = \delta\boldsymbol{\eta}\times\mathbf{b}_i^R \tag{15}$$

It may then be shown that:

$$\mathbf{b}_1^R\delta\bar{\gamma}_{11} + 2\mathbf{b}_\alpha^R\delta\bar{\gamma}_{1\alpha} = (\delta\mathbf{u})' + \bar{\mathbf{R}}'\times\delta\boldsymbol{\eta}$$
$$\mathbf{b}_i^R\delta\kappa_i = (\delta\boldsymbol{\eta})' \tag{16}$$

Insertion of equations (16) into equation (13) and integration by parts leads to

$$\delta W_I = [\mathbf{F}\cdot\delta\mathbf{u} + \mathbf{M}\cdot\delta\boldsymbol{\eta} + D\delta\psi]_a^b - \int_a^b [\mathbf{F}'\cdot\delta\mathbf{u} + (\mathbf{M}' + \bar{\mathbf{R}}'\times\mathbf{F})\cdot\delta\boldsymbol{\eta} + (D' - B)\delta\psi]\, dx_1 \tag{17}$$

The external forces acting upon the rod include the body force $\mathbf{p}$, tractions $\mathbf{t}$ upon the lateral surface, and tractions $\mathbf{t}_a$ and $\mathbf{t}_b$ upon the ends at $x_1 = a, b$. We restrict our discussion to beam statics. The virtual work of the external forces is then

$$\delta W_E = \iint_{A_b}\mathbf{t}_b\cdot\delta\mathbf{R}\, dx_2 dx_3 - \iint_{A_a}\mathbf{t}_a\cdot\delta\mathbf{R}\, dx_2 dx_3 + \int_a^b\left\{\iint_A \mathbf{p}\cdot\delta\mathbf{R}\, dx_2 dx_3 + \int_c \mathbf{t}\cdot\delta\mathbf{R}\, dc\right\} dx_1 \tag{18}$$

where A_a, A_b, A denote the areas of undeformed cross sections at a, b, x_1, and c denotes the length of the curve which circumscribes the undeformed cross section. Let us define the following averages of external forces:

$$\mathbf{P} = \iint_A \mathbf{p}\, dx_2 dx_3 + \int_c \mathbf{t}\, dc$$

$$\mathbf{N} = \iint_A (\mathbf{R} - \bar{\mathbf{R}})\times\mathbf{p}\, dx_2 dx_3 + \int_c (\mathbf{R} - \bar{\mathbf{R}})\times\mathbf{t}\, dc$$

$$f = \mathbf{b}_1^R\cdot\iint_A \mathbf{p}\lambda\, dx_2 dx_3 + \mathbf{b}_1^R\cdot\int_c \mathbf{t}\lambda\, dc$$

$$\mathbf{F}_a = \iint_{A_a}\mathbf{t}_a\, dx_2 dx_3$$

$$\mathbf{F}_b = \iint_{A_b}\mathbf{t}_b\, dx_2 dx_3 \tag{19}$$

$$\mathbf{M}_a = \iint_{A_a}(\mathbf{R} - \bar{\mathbf{R}})\times\mathbf{t}_a\, dx_2 dx_3$$

$$\mathbf{M}_b = \iint_{A_b}(\mathbf{R} - \bar{\mathbf{R}})\times\mathbf{t}_b\, dx_2 dx_3$$

$$D_a = \mathbf{b}_1^R\cdot\iint_{A_a}\mathbf{t}_a\lambda\, dx_2 dx_3$$

$$D_b = \mathbf{b}_1^R\cdot\iint_{A_b}\mathbf{t}_b\lambda\, dx_2 dx_3$$

Then the external virtual work from equation (18) becomes

$$\delta W_E = \mathbf{F}_b\cdot\delta\mathbf{u}_b + \mathbf{M}_b\cdot\delta\boldsymbol{\eta}_b + D_b\delta\psi_b - \mathbf{F}_a\cdot\delta\mathbf{u}_a - \mathbf{M}_a\cdot\delta\boldsymbol{\eta}_a - D_a\delta\psi_a + \int_a^b (\mathbf{P}\cdot\delta\mathbf{u} + \mathbf{N}\cdot\delta\boldsymbol{\eta} + f\delta\psi)\, dx_1 \tag{20}$$

Enforcing the principle of virtual work

$$\delta W_I = \delta W_E \tag{21}$$

for all admissible virtual displacements, we obtain from equations (17) and (20) the equilibrium equations

$$\mathbf{F}' + \mathbf{P} = 0$$
$$\mathbf{M}' + \bar{\mathbf{R}}'\times\mathbf{F} + \mathbf{N} = 0 \tag{22}$$
$$D' - B + f = 0$$

and the boundary conditions $\mathbf{F}_a = \mathbf{F}|_{x_1 = a}$, etc. Substitution of equations (4) and (14) into equations (22) results in a system of seven scalar differential equations in the seven unknowns ($\bar{\gamma}_{1i}$, κ_i, ψ).

IV Explicit Theory

In order to obtain an explicit set of strain-displacement relations, we must connect the triad $\mathbf{b}_i^R$ with the triad $\mathbf{b}_i^r$ by introduction of a suitable set of angular displacement parameters. The triad $\mathbf{b}_i^R$ may be obtained by rotating the triad $\mathbf{b}_i^r$ by an angle α about an axis defined by a unit vector $\mathbf{e}$. Transformation formulas then result from use of the Rodrigues vector as introduced in Simmonds and Danielson (1972) and Reissner (1973).

$$\boldsymbol{\Theta} = \theta_i\mathbf{b}_i^r = \theta_i\mathbf{b}_i^R = 2\,\mathbf{e}\tan\frac{\alpha}{2} \tag{23}$$

Introducing this vector, we find that

$$\mathbf{b}_i^R = \frac{(1 - \tfrac{1}{4}\boldsymbol{\Theta}\cdot\boldsymbol{\Theta})\mathbf{I} + \tfrac{1}{2}\boldsymbol{\Theta}\boldsymbol{\Theta} + \boldsymbol{\Theta}\times\mathbf{I}}{1 + \tfrac{1}{4}\boldsymbol{\Theta}\cdot\boldsymbol{\Theta}}\cdot\mathbf{b}_i^r \tag{24}$$

The curvature vector $\mathbf{K}$ may be expressed in terms of $\boldsymbol{\Theta}$ by use of equation (24) and the formula

$$\mathbf{K} = \frac{\boldsymbol{\Theta}' + \tfrac{1}{2}\boldsymbol{\Theta}\times\boldsymbol{\Theta}'}{1 + \tfrac{1}{4}\boldsymbol{\Theta}\cdot\boldsymbol{\Theta}} + k_i\mathbf{b}_i^R \tag{25}$$

Substitution into equations (22) results in a system of seven scalar differential equations in the seven unknowns ($\bar{\gamma}_{1i}$, θ_i, ψ). The strain vector $\bar{\boldsymbol{\gamma}}$ may be written in terms of $\mathbf{u}$ and $\boldsymbol{\Theta}$ by use of equation (24) and the formula

$$\bar{\boldsymbol{\gamma}} = \mathbf{u}' + \mathbf{b}_1^r - \mathbf{b}_1^R \tag{26}$$

Use of equation (26) results in a system of seven scalar differential equations in the seven unknowns (u_i, θ_i, ψ). A variational principle involving $\boldsymbol{\Theta}$ may be obtained from the principle of virtual work by use of the preceding equations and the relation

$$\delta\boldsymbol{\eta}=\frac{\delta\boldsymbol{\Theta}+\tfrac{1}{2}\boldsymbol{\Theta}\times\delta\boldsymbol{\Theta}}{1+\tfrac{1}{4}\boldsymbol{\Theta}\cdot\boldsymbol{\Theta}} \tag{27}$$

As shown in Hodges (1987a), equations (24)-(27) can be expressed easily and concisely in terms of matrices, and the Rodrigues formulation is simpler than other methods of describing finite rotation.

V Pretwisted Strip

We consider a strip of thin rectangular cross section having thickness t which is negligible compared to its width w, so that its Saint-Venant warping function is $\lambda = x_2x_3$. (Here $-t/2 \le x_2 \le t/2$ and $-w/2 \le x_3 \le w/2$). The beam has constant initial twist $\mathbf{k} = k\mathbf{b}_1'$. We suppose that the beam is subjected to an axial force $\mathbf{F}_b = \mathbf{F}_a = F\mathbf{b}_1'$ and an axial torque $\mathbf{M}_b = \mathbf{M}_a = M\mathbf{b}_1'$. These cause no bending of the reference line r, so that the axial base vector for the deformed beam $\mathbf{b}_1^R = \mathbf{b}_1'$, and the curvature vector of the deformed beam $\mathbf{K} = K\mathbf{b}_1^R$. The strain at the reference axis is due only to elongation and is given by $\bar{\gamma} = \bar{\gamma}\mathbf{b}_1^R$. We further assume that the ends of the bar are fee to warp, so that the elastic twist $\kappa = K - k$ is constant and the warping amplitude $\psi = \kappa$. We then obtain the local rotation vector and strain components from equations (9) and (10):

$$\begin{aligned}\boldsymbol{\phi}&=-x_3\kappa\mathbf{b}_3^R\\ \gamma_{11}&=\bar{\gamma}+\left(\frac{\kappa^2}{2}+k\kappa\right)x_3^2\\ 2\gamma_{12}&=(\bar{\gamma}+k\kappa x_3^2)\kappa x_3\\ 2\gamma_{13}&=2x_2\kappa\end{aligned} \tag{28}$$

It is interesting to note that the local rotation and the shear component of the strain are *not* of the same order of magnitude. Because the cross section is thin, the maximum value of ϕ_3 is w/t times the maximum value of γ_{13}. Substituting equations (28) into equations (11) and (12), we obtain the strain energy per unit length as

$$\begin{aligned}U=&\frac{Ewt}{2}\bar{\gamma}^2+\frac{Ew^3t}{12}\bar{\gamma}\left(k\kappa+\frac{\kappa^2}{2}\right)\\&+\frac{Ew^5t}{160}\left(k\kappa+\frac{\kappa^2}{2}\right)^2+\frac{Gwt^3}{6}\kappa^2\end{aligned} \tag{29}$$

Here we have used assumption (*iii*) to drop $2G\gamma_{12}^2$ from equation (11) since γ_{12} is $O(\epsilon^{3/2})$. The axial force F and torque M follow according to equations (14)

$$F=\frac{\partial U}{\partial\bar{\gamma}};\quad M=\frac{\partial U}{\partial\kappa} \tag{30}$$

Carrying out these differentiations and eliminating $\bar{\gamma}$ from these equations, we finally get

$$\begin{aligned}M=&\frac{Fw^2k}{12}+\frac{Gwt^3\kappa}{3}+\frac{Ew^5tk^2\kappa}{180}\\&+\underline{\frac{Fw^2\kappa}{12}+\frac{Ew^5tk\kappa^2}{120}+\frac{Ew^5t\kappa^3}{360}}\end{aligned} \tag{31}$$

Equation (31) agrees with equation (20) of Rosen (1983). The underlined terms stem from the influence of moderate local rotation in the analysis. Note that unless $F = 0$, even when κ is infinitesimal, not all of the underlined terms may be neglected. All of the terms are necessary to account for finite elastic twist κ in the presence of an axial force F. This example shows that without local rotations, there is no *nonlinear* coupling between extension and torsion. Indeed, for small strains and no local rotations (such as in a beam without warping) there can be no nonlinear coupling between extension and torsion, as discussed in Hodges (1987b).

VI Conclusion

We have obtained in Section III a complete, static, nonlinear beam theory expressed in terms of the generalized strains ($\bar{\gamma}_{1i}$, κ_i, ψ). This intrinsic theory would be appropriate when the generalized forces ($\mathbf{F}$, $\mathbf{M}$, D) are prescribed at the ends of the beam. In Section IV we obtained an equivalent theory expressed in terms of the generalized displacements (u_i, θ_i, ψ) to use when they are prescribed. We have obtained an analytical solution of the intrinsic problem in Section V, but for most other applications the equations will have to be solved numerically.

In future research, effects such as in-plane distortion of the cross section and warping due to shear could be included. Progress in this direction has been made by Axelrad (1973) for problems restricted to small local rotations. Extension of the theory to dynamics would also be useful. Our general procedure for simplifying the kinematics can be applied to any theory in which assumptions (*iii*) and (*iv*) are valid.

Acknowledgments

This paper is declared a work of the U.S. Government and is not subject to copyright protection in the United States. The first author was supported by the Naval Postgraduate School Foundation Research Program. The second author was supported by Senior Research Leadership Grant E-16-697 from the Georgia Tech Research Corporation.

References

Axelrad, E. L., 1973, "Statics of Elastic Beams," Strojizdat, Leningrad, translated from Russian in personal communication to the first author.

Chu, C., 1951, "The Effect of Initial Twist on the Torsional Rigidity of Thin Prismatical Bars and Tubular Members," *Proceedings of the 1st U.S. National Congress of Applied Mechanics*, pp. 265-269.

Danielson, D. A., and Hodges, D. H., 1987, "Nonlinear Beam Kinematics by Decomposition of the Rotation Tensor," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, No. 2, pp. 258-262.

Hodges, D. H., 1980, "Torsion of Pretwisted Beams Due to Axial Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, No. 2, pp. 393-397.

Hodges, D. H., 1987a, "Finite Rotations and Nonlinear Beam Kinematics," *Vertica*, Vol. 11, No. 1/2, pp. 297-307.

Hodges, D. H., 1987b, "Nonlinear Beam Kinematics for Small Strains and Finite Rotations," *Vertica*, Vol. 11, No. 3, pp. 573-589.

Malvern, L. E., 1969, *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, NJ, Chapter 6.

Ogden, R. W., 1984, *Non-Linear Elastic Deformations*, Ellis Horwood Limited, Chichester.

Reissner, E., 1973, "On One-Dimensional Large-Displacement Finite Strain Beam Theory," *Studies in Applied Mathematics*, Vol. 52, No. 2, pp. 87-95.

Rosen, A., 1983, "Theoretical and Experimental Investigation of the Nonlinear Torsion and Extension of Initially Twisted Bars," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, No. 2, pp. 321-326.

Simmonds, J. G., 1982, *A Brief on Tensor Analysis*, Springer-Verlag, New York.

Simmonds, J. G., and Danielson, D. A., 1972, "Nonlinear Shell Theory with Finite Rotation and Stress-Function Vectors," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, No. 4, pp. 1085-1090.

Simo, J. C., 1985, "A Finite Strain Beam Formulation. The Three-Dimensional Dynamic Problem. Part I," *Computer Methods in Applied Mechanics and Engineering*, Vol. 49, pp. 55-70.

Timoshenko, S., 1976, *Strength of Materials, Part II: Advanced Theory and Problems*, 3rd Edition, Krieger, Huntington, New York, Chapter VII.

Wempner, G. A., 1981, *Mechanics of Solids with Application to Thin Bodies*, Sijthoff and Noordhoof, Alphen aan den Rijn, The Netherlands, Chapter 8.

APPENDIX

In this paper we have used $\mathbf{C}^{Rr} = \mathbf{b}_i^R \mathbf{b}_i^r$ as the global finite rotation tensor and $\mathbf{b}_i^R = C_{ij}^{Rr} \mathbf{b}_j^r$ for the corresponding orthogonal base vectors of the deformed beam. Elements of the matrix C^{Rr} are the direction cosines of the transformation that relates the two triads. (Reversal of the superscripts on either the tensor or the matrix of components indicates the transposed tensor or matrix, respectively.) The vectors $\mathbf{b}_\alpha^R$ lie in the plane of the deformed beam reference cross section if beam deformation is constrained so that originally plane reference cross sections remain plane in the deformed beam. Note that $\mathbf{b}_1^R = \mathbf{b}_2^R \times \mathbf{b}_3^R$ is not necessarily tangent to R and that $\mathbf{b}_1^R$ is *not* equal to $\mathbf{C}(x_1, 0, 0) \cdot \mathbf{b}_i^r$ unless the Euler-Bernoulli hypothesis is adopted. These are two other possible candidates: (1) $\mathbf{C}^{Tr} = \mathbf{b}_i^T \mathbf{b}_i^r$ as the global finite rotation tensor and $\mathbf{b}_i^T = C_{ij}^{Tr} \mathbf{b}_j^r$ for the corresponding orthogonal base vectors where $\mathbf{b}_1^T$ is tangent to the deformed beam reference axis R; and (2) $\mathbf{C}^{Sr} = \mathbf{b}_i^S \mathbf{b}_i^r$ as the global finite rotation tensor and $\mathbf{b}_i^S = C_{ij}^{Sr} \mathbf{b}_j^r$ for the corresponding orthogonal base vectors where $\mathbf{C}^{Sr} = \mathbf{C}(x_1, 0, 0)$ represents the total rotation at the deformed beam reference line. In this appendix we describe these possible choices for orthogonal base vectors for the deformed beam and propose to establish that our choice ($\mathbf{b}_i^R$) is at least as "natural" as any other possibility. We also illustrate the geometrical relationships among these three candidates with an elementary example.

Consider a simple example in which the total rotation is constant over the cross section such as the first example from Hodges (1987b). The beam is initially straight and undergoes bending in a plane (see Fig. 1). The position vector to an arbitrary point in the undeformed beam is

$$\mathbf{r} = x_i \mathbf{b}_i^r \tag{A-1}$$

Assume that bending occurs in the 1-2 plane so that the 3-direction can be ignored for simplicity, and that the individual cross sections do not change shape. The position vector of the same material point in the deformed beam is

$$\mathbf{R} = (x_1 + u_1)\mathbf{b}_1^r + u_2 \mathbf{b}_2^r + x_2 \mathbf{b}_2^R \tag{A-2}$$

where

$$C^{Rr} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \tag{A-3}$$

The undeformed state base vectors are very simple

$$\mathbf{g}_i = \frac{\partial \mathbf{r}}{\partial x_i} = \mathbf{b}_i^r = \mathbf{g}^i \tag{A-4}$$

The deformed state covariant base vectors $\mathbf{G}_i$ are

$$\mathbf{G}_1 = (1 + u_1')\mathbf{b}_1^r + x_2(\mathbf{b}_2^R)' + u_2' \mathbf{b}_2^r$$
$$\mathbf{G}_2 = \mathbf{b}_2^R \tag{A-5}$$

where $(\ )'$ denotes a partial derivative with respect to x_1. From equations (A-3) and (A-5) it is determined that the matrix of deformation gradient components in mixed bases $\mathbf{b}_i^R \cdot \mathbf{A} \cdot \mathbf{b}_j^r = A_{ij}^{Rr}$ is given by

$$A^{Rr} = \begin{bmatrix} 1 + \gamma_{11} & 0 \\ 2\gamma_{12} & 1 \end{bmatrix} \tag{A-6}$$

where γ_{11} and γ_{12} are the nonzero strain components given by

$$\gamma_{11} = (1 + u_1')\cos\alpha + u_2' \sin\alpha - 1 - x_2\alpha'$$
$$2\gamma_{12} = -(1 + u_1')\sin\alpha + u_2' \cos\alpha \tag{A-7}$$

Utilizing assumption (*iii*), one can determine that the magnitude of the global rotation α is composed of a rotation due to bending plus a rotation due to shear so that

$$\alpha = \tan^{-1}\left(\frac{u_2'}{1 + u_1'}\right) - 2\gamma_{12} \tag{A-8}$$

This equation represents the magnitude of the rotation of the entire reference cross section plane as a rigid body. The first term represents a potentially large component due to bending while the second is a small component due to shear. C^{Tr} has

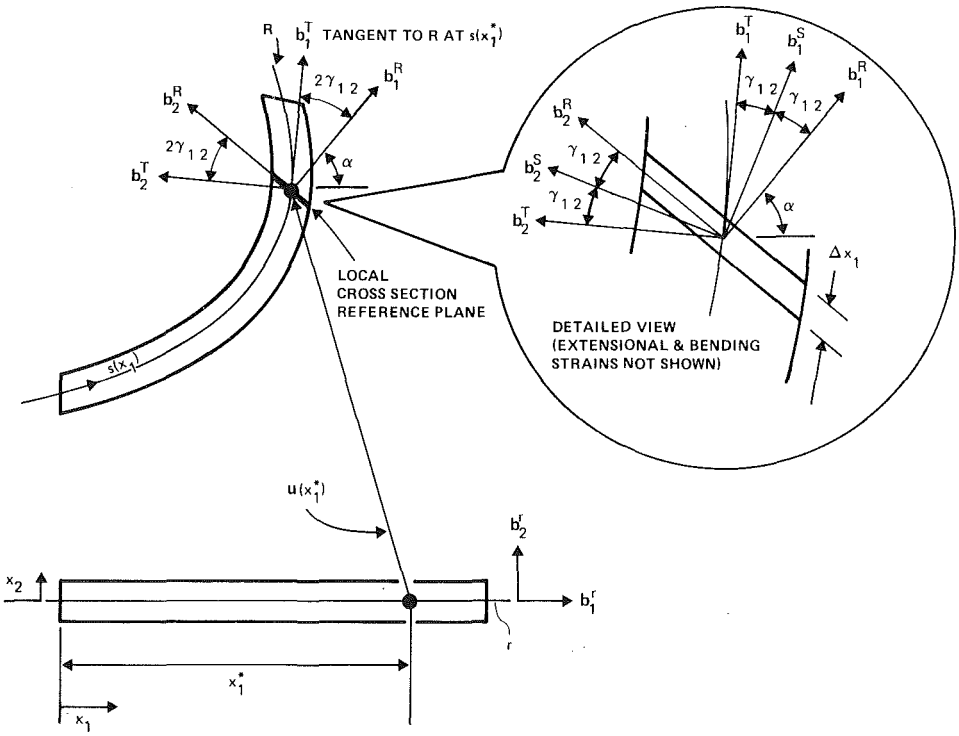


Fig. 1 Schematic of an initially straight beam undergoing planar deformation

exactly the same form as C^{Rr} except that α would be evaluated from equation (A-8) with $\gamma_{12} = 0$. When the deformation gradient components at the reference axis are expressed in mixed bases $\mathbf{b}_i^T \cdot \mathbf{A} \cdot \mathbf{b}_j^r = A_{ij}^{Tr}$, again utilizing assumption (*iii*), we obtain

$$A^{Tr} = \begin{bmatrix} 1+\gamma_{11} & 2\gamma_{12} \\ 0 & 1 \end{bmatrix} \tag{A-9}$$

It is interesting that neither A^{Rr} nor A^{Tr} is symmetric and one is the transpose of the other. The natural bases for the decomposition would appear to be those for which the matrix of deformation gradient components would be symmetric. These orthogonal base vectors turn out to be $\mathbf{b}_i^S$ and are related to the others by

$$C^{SR} = C^{TS} = \begin{bmatrix} 1 & \gamma_{12} \\ -\gamma_{12} & 1 \end{bmatrix} \tag{A-10}$$

The matrix C^{SR} represents a finite rotation of each material point due to shear identical to that determined by Hodges (1987b). Although this component of rotation is constant over the cross section, unlike the global rotation above it does *not* represent or contribute to a rigid-body rotation of the reference cross section. This point will be elaborated upon below.

With the only criterion being the rotational decomposition, it would appear that the natural finite rotation tensor would be $\mathbf{C}^{Sr}$, which is halfway between the other candidates. Indeed, the decomposition of rotation used in the present paper is

$$\begin{aligned} \mathbf{C} &= e^{\phi \times \mathbf{I}} \cdot \mathbf{C}(x_1,0,0) \\ &= e^{\phi \times \mathbf{I}} \cdot \mathbf{C}^{SR} \cdot \mathbf{C}^{Rr} \end{aligned} \tag{A-11}$$

The following represenations are equivalent:

$$\begin{aligned} \mathbf{C} &= e^{\phi \times \mathbf{I}} \cdot \mathbf{C}^{SR} \cdot \mathbf{C}^{RT} \cdot \mathbf{C}^{Tr} \\ &= e^{\phi \times \mathbf{I}} \cdot \mathbf{C}^{ST} \cdot \mathbf{C}^{Tr} \end{aligned} \tag{A-12}$$

Despite its apparent suitability for the rotational decomposition, however, if $\mathbf{b}_i^S$ were used for the orthogonal base vectors for global rotation, the overall formulation would become more complicated. The main reason for this complication stems from the close association of $\mathbf{b}_i^R$ with the reference cross section plane. This apparently motivated Simo (1985) to refer to $\mathbf{b}_i^R$ as the "intrinsic frame" rather than a mere coordinate system. An observer in this frame sees no rigid-body motion of the beam cross section—only deformation due to strain. Thus, its use facilitates a simple expression for $\mathbf{R}$. Use of the other candidates would require us to explicitly show rotations in $\mathbf{R}$ in terms of shear strain components. There might be circumstances, however, in which further decomposition of $\mathbf{C}^{Rr} = \mathbf{C}^{RT} \cdot \mathbf{C}^{Tr}$, as implied above in equation (A-8), is advantageous.

As is clear from equation (A-8), part of the global rotation is due to shear strain. There is also rotation due to shear strain associated with equation (A-10). Although both of these rotations stem from shear strain, only the one associated with global rotation influences the orientation of the cross section reference plane. The reader may wonder how to interpret a rotation component that is constant over the cross section, such as that represented by $\mathbf{C}^{SR}$, but yet does *not* represent a rigid-body rotation of the cross section. The most straightforward answer involves the relationship that, for small strain, a "simple shear" can be expressed as a combination of pure rotation and a "pure strain." For example, ignoring extension and bending from equation (A-6), we can express a simple shear of magnitude $2\gamma_{12}$ as

$$\begin{bmatrix} 1 & 0 \\ 2\gamma_{12} & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\gamma_{12} \\ \gamma_{12} & 1 \end{bmatrix} \begin{bmatrix} 1 & \gamma_{12} \\ \gamma_{12} & 1 \end{bmatrix} \tag{A-13}$$

Figure 1 shows this deformation in the exploded view. (Note that in this view, for simplicity, bending and extensional strains are not shown and shear strain is exaggerated.) The only difference between a rigidly rotated reference cross section and the final configuration of the reference cross section (shown in Fig. 1 with thickness Δx_1) is a simple shear. Both in equation (A-13) and in Fig. 1 we observe that for a simple shear, only one axis rotates. This is equivalent to a pure rotation, in which both axes rotate by the same amount in the same direction, plus a pure strain, in which the axes rotate by the same amount in the opposite directions. Furthermore, from Fig. 1 we see that $\mathbf{b}_i^S$ are those orthogonal unit vectors from which unit vectors that are parallel to the faces of the cross section parallelogram differ by a pure strain. I.e.,

$$\begin{Bmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^R \end{Bmatrix} = \begin{bmatrix} 1 & \gamma_{12} \\ \gamma_{12} & 1 \end{bmatrix} \begin{Bmatrix} \mathbf{b}_1^S \\ \mathbf{b}_2^S \end{Bmatrix} \tag{A-14}$$

When we consider the schematic in Fig. 1 in the limit as Δx_1 tends towards zero, a choice of $\mathbf{b}_i^R$ as the "natural" system becomes even clearer.

B. Poddar*

F. C. Moon
Professor,
Mem. ASME

S. Mukherjee
Professor,
Mem. ASME

Department of Theoretical and Applied Mechanics,
Cornell University,
Ithaca, NY 14853

Chaotic Motion of an Elastic-Plastic Beam

A numerical study is presented here which suggests that chaotic motion is possible from periodic excitation of an elastic-plastic beam. Poincaré maps of the motion reveal a fractal-like structure of the attractor. The results suggest that geometric and material nonlinearities in solid mechanics problems may lead to extreme sensitivity to small changes in parameters and resulting unpredictability. These results may explain the total disagreement of nine finite element codes in the analysis of the transient response of an elastic-plastic beam, that has been reported recently by Symonds and his coworkers.

Introduction

Recent research in dynamical systems has shown that the time histories of deterministic nonlinear systems can be very sensitive to initial conditions and may exhibit chaotic or random-like behavior (see, e.g., Holmes and Moon, 1983). The dynamic behavior of elastic-plastic structures is a highly nonlinear system and may, in some problems, have very sensitive solutions which make the precise time history unpredictable. In two recent papers (Symonds and Yu, 1985; Symonds et al. 1985), the authors have presented numerical results for the transient response of an elastic-plastic beam using nine different finite element codes and have obtained different time histories from all the codes. Considering all the codes, Symonds and Yu (1985) report that "12 solutions may be considered strictly comparable." For a certain narrow range of initial displacements, the subsequent free motion appears impossible to predict and the different computer programs could not even agree as to the final equilibrium configuration of the beam. Clearly, these results suggest extreme sensitivity of the system to initial conditions, such that the slightest uncertainty in initial conditions causes extreme variation in the response.

In Symonds' problem, a beam has fixed pin supports at both ends and is subjected to a transient force. In the two papers quoted above, Symonds et al. (1985) have studied this problem by both finite element analysis and also by employing a simplified Shanley (1947) model. This paper describes further analysis of this problem, again using a Shanley model, for the case of free and forced vibrations of the beam. Evidence is presented here that the "counter-intuitive" problem of Symonds represents an example of an elastic-plastic chaotic vibration problem. In such problems both the geometric and constitutive stress-strain nonlinearities play an important role. While a highly idealized model for the plasticity in the beam is used here (the Shanley model), it is expected that similar unpredictable phenomena may be a generic problem involving the dynamics of inelastic structures with geometric nonlinearities.

Governing Differential Equations

The Shanley (1947) model, as used here, approximates a pin-ended beam of length 2ℓ, with a uniform rectangular cross section of area $A = bh'$, by two rigid links, each of length ℓ (Fig. 1). These links are pinned at the ends and are joined together at the center by an elastic-plastic element. This elastic-plastic element has two short flanges placed at distances $h/2$ above and below the beam axis, respectively. Each flange is elastic-perfectly plastic with a yield stress, both in tension or compression, of σ_y. Following Symonds et al. (1985),

$$\dot{a} = \ell\dot{\phi},\ \dot{e} = \ell\phi\dot{\phi} \tag{1}$$

where ϕ is the angular deflection of a link of the beam, a is the

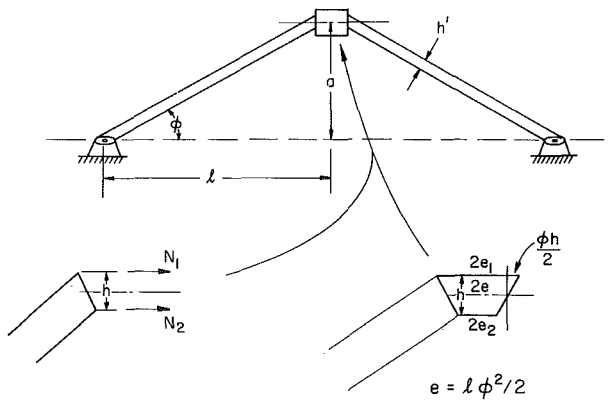


Fig. 1

*Currently Assistant Professor, Department of Mechanical Engineering, University of Miami, Coral Gables, FL 33124. Assoc. Mem. ASME

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May 16, 1986; final revision July 15, 1987.

transverse central deflection, and $2e$ is the extension of the center line of the beam. The semiextension rates of the flanges, therefore, become

$$\dot{e}_1 = (a + h/2)\dot{a}/\ell,\ \dot{e}_2 = (a - h/2)\dot{a}/\ell \tag{2}$$

the subscripts 1 and 2 denoting the upper and lower flanges, respectively. The force rates $\dot{N}_1$ and $\dot{N}_2$ in the flanges are

$$\dot{N}_1 = \begin{cases} \dfrac{E'}{\ell^2}(a + h/2)\dot{a} & \text{elastic} \\ 0 & \text{plastic} \end{cases} \tag{3}$$

and similarly for $\dot{N}_2$ with $(a + h/2)$ replaced by $(a - h/2)$ in the above. The assumption here is that the strain rates, concentrated in the flanges of the plastic element, are obtained by dividing $\dot{e}_1$ and $\dot{e}_2$ by the semilength ℓ of the beam. Also, in the above, $E' = EA/2$ in terms of the effective Shanley model Young's modulus E and cross section area A of the beam.

The governing differential equation of motion of the beam is written for the case where a is moderately large so that the membrane forces N_1 and N_2 must be included in the analysis. The resulting equation is

$$\ddot{a} + 2\zeta\omega_R\dot{a} + \frac{3}{\rho\ell^2}Na - \left(\frac{3}{\rho\ell^2}\right)M = f(t) \tag{4}$$

where $N = N_1 + N_2$ and $M = (N_2 - N_1)h/2$ are the resultant axial force and moment, respectively, on half the beam, ρ is the mass per unit length of the beam, and ζ is the damping factor. Also, $f(t) = 1.5F(t)/\rho$ in terms of $F(t)$, the transverse force per unit length applied on the beam, and

$$\omega_R = \frac{h}{\ell^2}\sqrt{\frac{1.5E'}{\rho}},$$

the natural frequency of small elastic vibrations of the beam about its initial equilibrium position.

It is convenient to write equation (4) as a system of four first order differential equations

$$\begin{aligned} \dot{y}_1 &= y_2 \\ \dot{y}_2 &= -2\zeta\omega_R y_2 - \frac{3}{\rho\ell^2}y_1 y_3 + \frac{3}{\rho\ell^2}y_4 + f(t) \\ \dot{y}_3 &= \dot{N}_1 + \dot{N}_2 \\ \dot{y}_4 &= (\dot{N}_2 - \dot{N}_1)h/2 \end{aligned} \tag{5}$$

where $a \equiv y_1$, $\dot{a} \equiv y_2$, $N \equiv y_3$, $M \equiv y_4$, and $\dot{N}_1$ and $\dot{N}_2$ are given by equation (3). It is easy to show that in this case the yield surface in force space is diamond shaped and is bounded by the pairs of lines

$$|N/N_y - M/M_y| = 1,\ |N/N_y + M/M_y| = 1$$

where $N_y = A\sigma_y$ and $M_y = Ah\,\sigma_y/2$.

It is very important to note here that the governing equation (4) has the quantities N and M *which depend on the history of motion*. Hence, it is convenient to rewrite equation (4) as a system of four first order differential equations.

Initially y_1 and y_2 are prescribed and y_3 and y_4 are determined from y_1 assuming monotonic deflection up to the initial value of y_1.

Equilibrium Points, Stability and Frequency Analysis

As mentioned before, the natural frequency of small elastic vibrations is ω_R, which is given below equation (4). A very interesting situation arises if elastic oscillations follow touching of the yield surface, as is often the case for free vibrations, especially with damping. Such a situation in force space is depicted in Fig. 2. If purely elastic oscillations follow the last

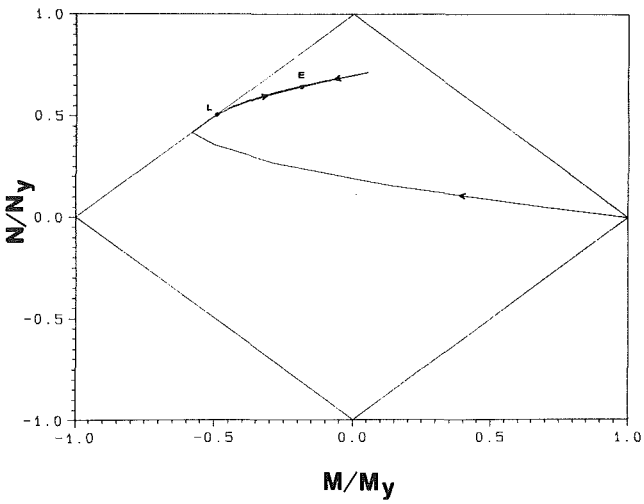


Fig. 2 Trajectory in force space for free vibrations: a_o = 0.9cm, $\dot{a}_o$ = 0, ζ = 0.2

touching of the yield surface at L, the system remembers the values M_L, N_L, and a_L at this point. Now the governing equation is (4) with the integrated values

$$\begin{aligned} N &= (E'/\ell^2)(a^2 - a_L^2) + N_L, \\ M &= -(h^2E'/2\ell^2)(a - a_L) + M_L \end{aligned} \tag{6}$$

Substituting the above into equation (4) and collecting terms, one obtains the equation

$$\ddot{a} + 2\zeta\omega_R\dot{a} + \alpha a + \beta a^3 - \delta = 0 \tag{7}$$

where $f(t)$ has been set to zero. The coefficients α, β, and δ in the above are

$$\alpha = (3/(\rho\ell^2))[N_L - E'a_L^2/\ell^2 + h^2E'/(2\ell^2)]$$

$$\beta = (3/(\rho\ell^2))(E'/\ell^2)$$

and

$$\delta = (3/(\rho\ell^2))[M_L + h^2E'a_L/(2\ell^2)] \tag{8}$$

It is clear that α and δ depend upon the situation at L, and hence on the history of the motion.

Equation (7) is a form of Duffing's equation with an extra term δ. The behavior of solutions of the standard Duffing's equation, with respect to fractal basins and chaos, have been studied by many researchers. Recent papers on the subject, of particular importance to this work, are those by Moon and Li (1985) and by Dowell and Pezeshki (1986). It is important to mention again that the coefficients α and δ, in equation (7), are history dependent. Consequently, the equilibrium points of equation (7) are history dependent.

Equation (7) has three equilibrium points. A standard analysis shows that if $\alpha^3 > -(27/4)\,\delta^2\beta$, there is only one real equilibrium point. Otherwise, there are three. Thus, for example, if $\alpha > 0$, only one equilibrium point exists (note that β is always positive, while α and δ can have either sign). The closed form expressions for these equilibrium points are quite complex. With $\zeta = 0$, however, it is easy to show that the stability of an equilibrium point a_E is determined by γ where

$$\gamma = \pm\sqrt{-(\alpha + 3a_E^2\beta)}$$

With $\alpha + 3\,a_E^2\,\beta > 0$, the two values of γ are imaginary and the equilibrium point has periodic vibrations around it. In this, case, $\zeta > 0$ leads to the equilibrium point being stable. If $\alpha + 3\,a_E^2\,\beta < 0$, the equilibrium point is unstable.

All the transient simulations considered later in this paper admit three real equilibrium points $a_E^{(1)} > a_E^{(2)} > a_E^{(3)}$. Of these, $a_E^{(1)} > 0$, $a_E^{(3)} < 0$ and both are stable, while the intermediate

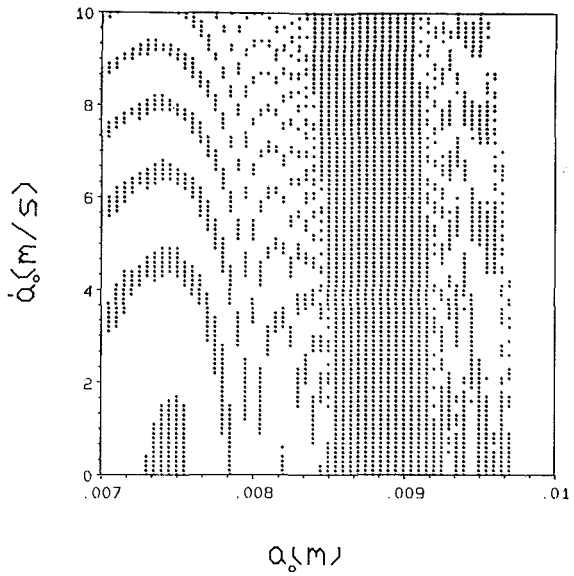


Fig. 3 Dependence of numerical simulation of the transient solution on initial conditions

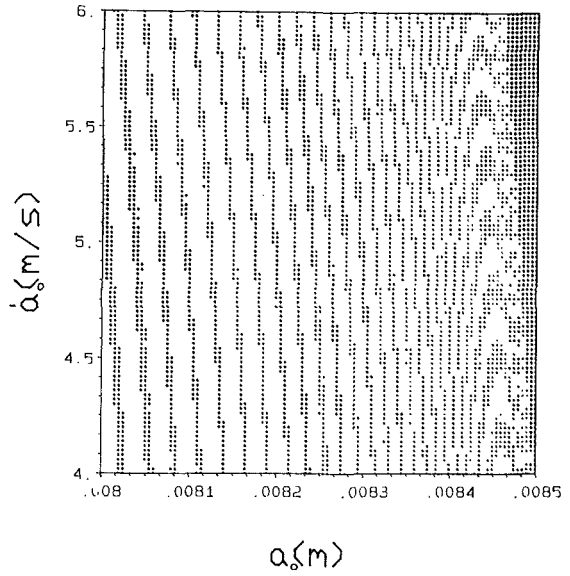


Fig. 4 Enlargement of a small region of Fig. 3

a_E is unstable. Depending on the initial conditions, the solution in the presence of damping settles down to one of the two stable equilibrium points. The phase plane trajectory corresponding to Fig. 2, for example, has the features of Fig. 1 in Dowell and Pezeshki (1986). In this case the equilibrium value $a_E < 0$. Of course, a_E can be greater than zero for other initial conditions.

The frequency of small amplitude elastic oscillations, following last touching of the yield surface at L, can be calculated from equation (7). Writing $a = a_E + \eta$ substituting the above into equation (7) and linearizing with respect to η, the damped natural frequency of small amplitude oscillations about this equilibrium point is obtained as

$$\omega_d = \omega_n \sqrt{1-(\zeta')^2} \text{ with } \omega_n = \sqrt{(3a_E^2\beta + \alpha)}$$

and

$$\zeta' = \zeta\, \omega_R/\omega_n \tag{9}$$

It is clear from equations (8) and (9) that this natural frequency is history dependent.

Numerical Simulations

Parameter Values. The parameter values used here are the same as those from Symonds et al. (1985). These are

$\ell = 10\text{cm}$, $b = 2\text{cm}$, $h' = 0.4\text{cm}$, $h = 0.68h'$

$\rho = 0.216\ \text{kg/m}$, $E = 120 \times 10^9\ \text{N/m}^2$

$\sigma_y = 0.3 \times 10^9\ \text{N/m}^2$

It should be noted that the Young's modulus of the beam material is $80 \times 10^9\ \text{N/m}^2$ and the number above is the effective Young's modulus for the Shanley model as suggested by Symonds et al. (1985).

Natural Frequencies. The equations (5) are integrated numerically by using the Runge-Kutta method with time-step control. The first check on the computer program for numerical simulations has been to calculate a_E and ω_d for two cases—(*a*) corresponding to Fig. 2 with $a_o = 0.009\text{m}$, $\dot{a}_o = 0$ where the solution settles down to $a_E < 0$, and (*b*) with $a_o = 0.011\text{m}$, $\dot{a}_o = 0$ which gives $a_E > 0$. A damping factor $\zeta = 0.2$ was included in the numerical simulations in order to get small amplitude oscillations. The numerical simulation frequencies below correspond to those obtained from the time period of a small amplitude cycle before the beam settles down at its equilibrium configuration. The results are summarized below.

	a_E (cm)		ω_d (rad/s)	
	Numerical Simulation	From equation (7) with $\ddot{a} = \dot{a} = 0$	Numerical Simulation	From equation (9)
(a)	−0.4087	−0.4112	4187.6	4238
(b)	0.8209	0.8204	9622.7	9615

It should be mentioned here that the three calculated equilibrium points for case (*a*) above are +0.5376, −0.1288, and −0.4087, of which the intermediate one is unstable and the other two are stable. Similarly, for case (*b*) the a_E values are 0.8209, −0.557, and −0.765, of which the first and last are stable.

Free Vibrations. Moon and Li (1985) have discussed the phenomenon of extreme sensitivity of solutions of Duffing's equation to the initial values of a and $\dot{a}$. In predictable problems, the boundary between two basins of attraction in initial condition space is smooth. In chaotic problems, however, the boundary becomes fractal and subsequent motion is extremely sensitive to initial conditions (see, e.g., Grebogi et al., 1983). Symonds et al. (1985) report extreme sensitivity of the solution of this beam problem to a_o (they took ζ and $\dot{a}_o$ equal to zero) in the approximate range $0.85 < a_o < 0.98$ cm. In recent work, Genna and Symonds (1987) include damping in their simulations and report anamolous behavior over a wider range of a_o (approximately for $0.70 < a_o < 0.98$ cm).

This question of extreme dependence of the transient solution to initial conditions has been studied in this work in the following way. The initial condition space $0.7 < a_o < 1$ cm, $0.0 < \dot{a}_o < 10$ m/s has been studied. In all cases the beam starts from an initial configuration above the horizontal and any nonzero initial velocity is upwards as well. For a given simulation, the following steps have been carried out.

(*a*) Choose a_o, $\dot{a}_o$. Set $\zeta = 0$.

(*b*) Compute the solution for a certain number of cycles and find $a_{\max}$ and $a_{\min}$. If $a_{\max} < 0$ then the behavior is clearly anamolous since the beam vibrates entirely below its original straight configuration.

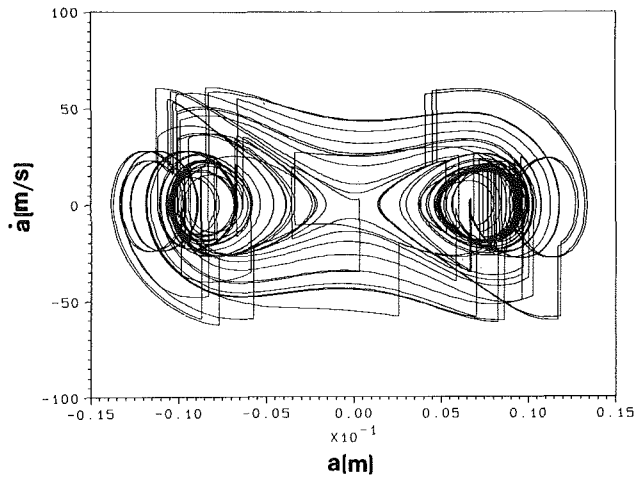


Fig. 5 Chaotic forced vibrations—first 35 forcing cycles in phase space

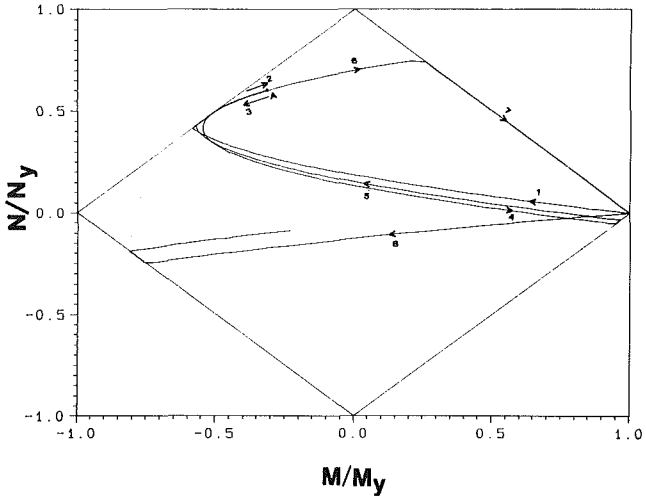


Fig. 6 Chaotic forced vibrations—first 1.5 forcing cycles in force space: $a_o = 0.009$, $\dot{a}_o = 0.0$, $\Delta\dot{a} = 38.0$ m/s, $\omega_f = 2.5\omega_R = 3926.0$ rad/s

(*c*) If $a_{\max} > 0$, introduce $\zeta = 0.10$ and let the solution converge to a_E. Anamolous behavior is obtained if $a_E < -0.1\,(a_{\max} - a_{\min})$.

Condition (*c*) is chosen to ascertain that the beam does indeed end up below its original horizontal configuration in order for the simulation to merit the anamolous classification. The results are shown in Figs. 3 and 4 where a dot denotes anamolous and a blank regular (expected) behavior. Figure 4 is an enlargement of a small region of Fig. 3. The extreme sensitivity of the simulations, to initial conditions, is evident. With $\dot{a}_o = 0$, the range of a_o, for anamolous behavior, is approximately $0.73 < a_o < 0.97$, which is in good agreement with Genna and Symonds (1987). It should be emphasized that, as expected, the details of Figs. 3 and 4 would be affected by the choice, among others, of ζ and the specific criteria used to define counterintuitive behavior. The observation of extreme sensitivity of the solution to initial conditions, however, is a salient feature of this behavior, and this feature would be retained in different types of simulations. It is clear that in this range of initial conditions, the slightest change in initial conditions can cause a drastic change in the response, and attempts at obtaining detailed numerical solutions to the problem are meaningless.

Another observation is that for all the undamped free vibration problems simulated in this work, the final limit cycle motion involves touching of the yield surface in force space without part of the trajectory moving on it. It is felt that such is the case in general. The conjecture, is, as yet, unproven but is considered highly likely.

Forced Vibrations. Two of the hallmarks of chaotic behavior are a positive Lyapunov exponent and a fractal dimension of the strange attractor on which the solution moves. One way to observe the fractal properties of an attractor is to obtain a Poincaré section of the motion. In order to explore the strange attractor one must allow the trajectory to explore many regions of the attracting set. This requires a long time simulation. For this reason, a periodically excited problem has been considered here in which the beam receives alternating positive and negative impulses. The forcing function here is a periodic series of these impulses. Each impulse is assumed to cause sudden velocity changes of $\Delta\dot{a}$ in its direction. As has been described often in the chaos literature (e.g., Dowell and Pezeshki, 1986), a suitable combination of $\Delta\dot{a}$ and forcing frequency ω_f can lead to the response being unable to decide between two equilibrium positions. Such a case is depicted in the phase plane (Fig. 5, first 35 forcing cycles) and in the force space (Fig. 6, first 1.5 forcing cycles). Physically, once the beam becomes plastic, its length increases and the beam acts like an arch with new equilibrium positions above and below the original undeformed straight shape. The shallow arch dynamics is similar to a particle in a double-well potential which is known to exhibit chaotic vibrations under periodic excitation (see Tseng and Dugundji, 1971; Moon, 1980).

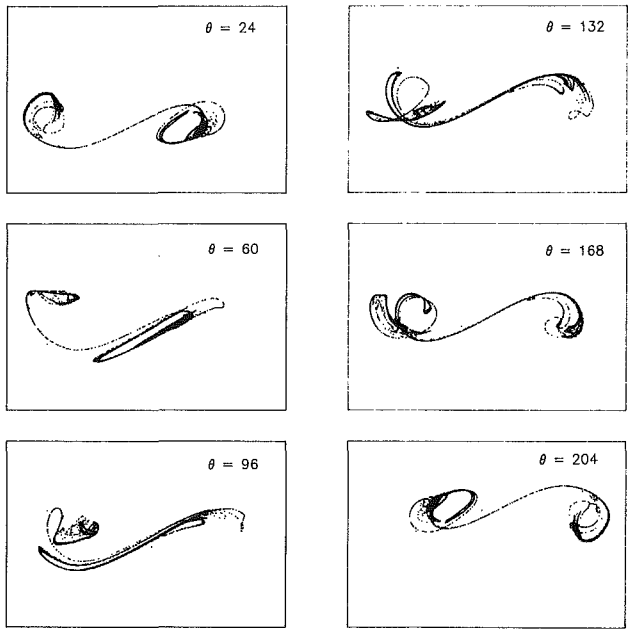


Fig. 7 Chaotic forced vibrations—Poincaré maps for 5000 forcing cycles. Same simulation as in Fig. 6.

Poincaré sections were obtained by stroboscopically looking at the motion synchronous with the periodic impulses and projecting the resulting set of points into the amplitude-velocity plane (a, $\dot{a}$). The Poincaré sections for this problem use 5000 points and show a characteristic fractal structure (Fig. 7; see also Moon, 1980). These maps correspond to initial phase intervals of $\pi/5$ and, as expected, the picture for $\theta = 204$ deg is the reversed version of that for $\theta = 24$ deg. An enlargement of one of the Poincaré sections from Fig. 7 is shown in Fig. 8. The structure of the section is more clearly apparent in this larger picture. It should be observed here that the trajectory in force space continues to touch or move on the yield surface during this entire simulation of 5000 cycles, showing that the motion, in this case, remains elasto-plastic throughout.

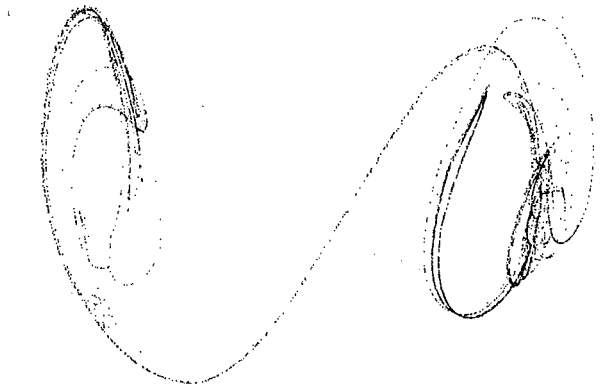


Fig. 8 Chaotic forced vibrations—Poincaré map for θ = 24 deg from Fig. 7

It should be reemphasized, in conclusion, that the chaotic motion demonstrated in this paper from the simplified Shanley model is expected to be of a generic nature. Similar unpredictable phenomena are expected in the dynamics of structures in the presence of material and/or geometric nonlinearities. Research in this general direction is currently in progress at Cornell University.

Acknowledgments

This work was supported in part by a grant from the Air Force Office of Scientific Research, Aerospace Sciences Division, to Cornell University. The computing of Poincaré maps has been carried out on the NSF supported Cornell National Supercomputer Facility.

References

Dowell, E. H., and Pezeshki, C., 1986, "On the Understanding of Chaos in Duffing's Equation Including a Comparison with Experiment," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 5–9.

Genna, F., and Symonds, P. S., 1987, Private Communications.

Grebogi, C., Ott, E., and Yorke, J. A., 1983, "Fractal Basin Boundaries, Long-Lived Chaotic Transients, and Unstable-Unstable Pair Bifurcations," *Physical Review Letters*, Vol. 50, pp. 935–938.

Holmes, P. J., and Moon, F. C., 1983, "Strange Attractors and Chaos in Nonlinear Mechanics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 1021–1032.

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

Moon, F. C., and Li, G.-X., 1985, "Fractal Basin Boundaries and Homoclinic Orbits for Periodic Motion in a Two-Well Potential," *Physical Review Letters*, Vol. 55, pp. 1439–1442.

Shanley, F. R., 1947, "Inelastic Column Theory," *Journal of the Aeronautical Sciences*, Vol. 14, pp. 261–267.

Symonds, P. S., and Yu, T. X., 1985, "Counterintuitive Behavior in a Problem of Elastic-Plastic Beam Dynamics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 517–522.

Symonds, P. S., McNamara, J. F., and Genna, F., 1985, "Vibrations of a Pin-Ended Beam Deformed Plastically by Short Pulse Excitation," *Material Nonlinearity in Vibration Problems*, Sathyamoorthy, M., ed., AMD-Vol. 71, ASME, NY, pp. 69–78.

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

D. M. Tang*
Visiting Scholar,

E. H. Dowell
Dean,
Mem. ASME

Department of Mechanical Engineering
and Materials Science,
School of Engineering,
Duke University,
Durham, NC 27706

On the Threshold Force for Chaotic Motions for a Forced Buckled Beam

The effects of higher modes on the chaotic oscillations of a buckled beam under forced external excitation are studied. Of principal interest are the threshold force required for chaotic motions and the influence of damping on the system response. A comparison is also presented of results from numerical simulations with experimental data.

Introduction

Chaotic oscillations in Duffings equation (with a negative linear stiffness) have been studied by Ueda (1980), Moon (1980), Holmes (1979), and Dowell and Pezeshki (1986). Moon (1979) and Holmes and Moon (1983) have investigated the chaotic behavior of a buckled beam for forced deterministic excitations. In their theoretical and experimental study, magnets were placed to the left and right of the free end of cantilevered beam to buckle the beam, and two or three stable static equilibrium positions were found. Chaotic motions about those multiple equilibrium positions were observed. Furthermore, the earlier, pioneering, work of Tseng and Dugundji (1971) should be noted.

Moon (1979) and Holmes and Moon (1983) presented approximate analytical threshold criteria to predict the range of parameters for which these chaotic motions will occur. These criteria were based on a one mode theoretical model of a buckled beam under the action of a harmonic external force and were in qualitative agreement with the corresponding experimental results. Dowell and Pezeshki (1986) carried out systematic numerical simulation studies which established theoretical threshold values in quantitative agreement with experiments.

In this paper, we extend the analysis of Moon (1979), Holmes and Moon (1983), and Dowell and Pezeshki (1986) to include the effects of higher modes on the chaotic response. An experimental investigation for chaotic oscillations also was carried out for a cantilevered pipe-beam. Generally good quantitative agreement between theoretical and experimental results is found by including up to three beam modes and the corresponding modal viscous damping.

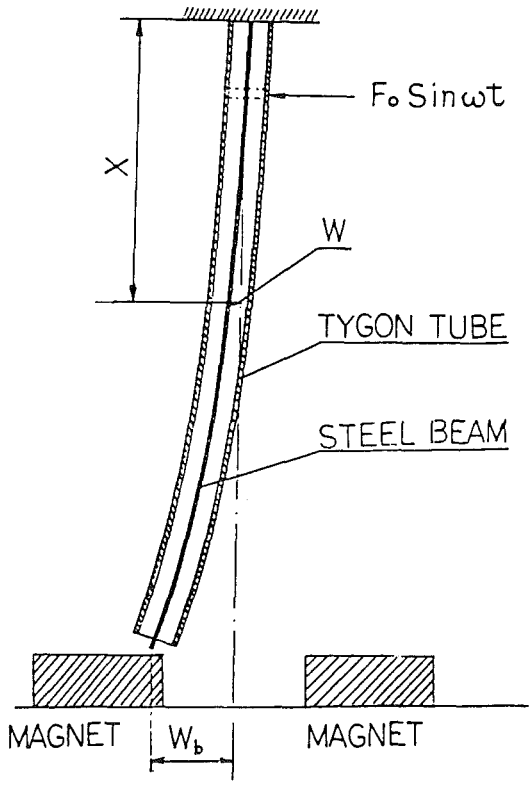


Fig. 1 Pipe-beam geometry

Problem Description and Formulation

The system under investigation is shown schematically in Fig. 1. A flexible pipe is made of Tygon material which has higher material damping, larger mass per unit length and lower stiffness than typical metals. A steel (ferromagnetic) beam is placed inside this pipe. Thus a combined pipe-beam is constructed with two cavities which has very low lateral stiffness and high longitudinal stiffness. The pipe-beam is clamped at its upper end, and free at its lower end. Two permanent

*Associate Professor, Nanjing Aeronautical Institute, China.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 10, 1986; final revision, May 26, 1987.

Paper No. 88-WA/APM-7.

magnets are placed to the left and right of the free end of the pipe-beam which is buckled by the magnetic forces.

In this paper, the equations of motion are obtained by the use of Lagrange's equations in the form

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right)-\frac{\partial T}{\partial q_i}+\frac{\partial V}{\partial q_i}+\frac{\partial D}{\partial \dot{q}_i}=Q_i \quad (i=1,2,\ldots n) \tag{1}$$

It is necessary to choose a set of convenient generalized coordinates, q_i. We assume that the beam deflection, $w(x,t)$, can be considered as the sum of displacements in the n lowest modes of a uniform cantilever elastic beam.

$$w(x,t)=\sum_{s=1}^{n} q_s(t)\phi_s(x) \tag{2}$$

where $\phi_s(x)$ is the characteristic modal function of a uniform cantilever beam.

The kinetic energy of the system is

$$T=\frac{1}{2}\,m\int_o^{\ell}\left(\frac{\partial w}{\partial t}\right)^2 dx \tag{3}$$

where m is the mass per unit length of the pipe-beam.

The vertical longitudinal displacement of the beam is

$$v=\frac{1}{2}\int_o^{x}\left(\frac{\partial w}{\partial x}\right)^2 dx \quad (0\le x\le \ell)$$

The elastic plus gravitational potential energy is

$$V_1=\frac{1}{2}\,EI\int_o^{\ell}\left(\frac{\partial^2 w}{\partial x^2}\right)^2 dx + \frac{1}{2}\,mg\int_o^{\ell}\left[\int_o^{x}\left(\frac{\partial w}{\partial x}\right)^2 dx\right]dx$$

The magnetic field acting on the system creates magnetic forces and couples which are assumed to be concentrated at the free end of the beam. A magnetic energy potential can be found in terms of the displacement at the free end (see Moon, 1979). It is

$$V_2=\frac{1}{2}\,K_1 w^2(t,\ell)+\frac{1}{4}\,K_2 w^4(t,\ell)$$

The total potential energy is

$$V=V_1+V_2 \tag{4}$$

The dissipation energy of the system induced by damping is very complex. In our system there are two possible types of damping to be considered, namely viscous and Kelvin damping. Viscous damping is often used in models of vibration systems. The corresponding dissipation function in Lagrange's equations is represented as

$$D=\frac{1}{2}\,c\int_o^{\ell}\left(\frac{\partial w}{\partial t}\right)^2 dx \tag{5}$$

where c is the viscous damping coefficient.

Kelvin damping is widely accepted as representing internal damping in a viscoelastic material. Its properties are expressed by a dynamic viscous modulus coefficient, E^*. In Lagrange's equations, the Young's modulus E will simply be replaced by the expression $(E+E^*\partial/\partial t)$.

The virtual work done by the exciting force is

$$\delta W=F\delta w \quad \text{at} \quad x=x_F$$

On substituting equation (2) into these energy and work expressions, and the results in equation (1), the equations of motion become

$$\sum_{s=1}^{n}\int_o^{\ell}\Big[m\ddot{q}_s\phi_s\phi_i+EIq_s\phi_s''\phi_i''+E^*I\dot{q}_s\phi_i''\phi_s''+c\dot{q}_s\phi_s\phi_i + mgq_s\int_o^{x}\phi_s'\phi_i'\,dx\Big]dx+K_1\Big[\sum_{s=1}^{n}q_s\phi_s(\ell)\Big]\phi_i(\ell) + K_2\Big[\sum_{s=1}^{n}q_s\phi_s(\ell)\Big]^3\phi_i(\ell)=Q_i \quad (i=1,2,\ldots n) \tag{6}$$

where $Q_i=F\phi_i(x_F)$, and the exciting force is assumed to be a concentrated force applied at point $x=x_F$.

We note that

$$\int_o^{\ell}\phi_s''\phi_i''\,dx=\int_o^{\ell}\phi_s^{iv}\phi_i\,dx$$

and

$$\phi_s'(0),\ \phi_i(0)=0.$$

Also

$$\int_o^{\ell}\int_o^{x}\Big[\phi_s'\phi_i'\,dx\Big]dx=\int_o^{\ell}[\phi_s'\phi_i-(\ell-x)\phi_s''\phi_i]dx$$

Substituting the above expressions into equation (6), it is found that

$$\sum_{s=1}^{n}\int_o^{\ell}[m\ddot{q}_s\phi_s+EIq_s\phi_s^{iv}+mgq_s\phi_s'+E^*I\dot{q}_s\phi_s^{iv}+c\dot{q}_s\phi_s - mg(\ell-x)q_s\phi_s'']\phi_i dx+K_1\Big[\sum_{s=1}^{n}q_s\phi_s(\ell)\Big]\phi_i(\ell) + K_2\Big[\sum_{s=1}^{n}q_s\phi_s(\ell)\Big]^3\phi_i(\ell)=F\phi_i(x_F) \quad (i=1,2,\ldots n) \tag{7}$$

In order to put equation (7) into a dimensionless form, we introduce the following dimensionless parameters and coordinates:

$$\bar{x}=\frac{x}{\ell},\quad \psi_s=\phi_s(\bar{x}),\quad a_s=\frac{q_s}{\ell},\quad f=\frac{F\ell^2}{EI},\quad \bar{w}=\frac{w}{\ell},$$

$$\sigma=\frac{mg}{EI}\,\ell^3,\quad \zeta=\frac{c\ell^2}{(EIm)^{1/2}},\quad \nu=\frac{E^*}{\ell^2 E}\left[\frac{EI}{m}\right]^{1/2}$$

$$\theta_1=\frac{K_1}{EI}\,\ell^3,\quad \theta_2=\frac{K_2}{EI}\,\ell^5,\quad \tau=\left[\frac{EI}{m\ell^4}\right]^{1/2}t$$

Finally, equation (7) can be rewritten in a more compact form. It is

$$\sum_{s=1}^{n}[A_{is}\ddot{a}_s+B_{is}\dot{a}_s+C_{is}a_s]+\theta_2\Big[\sum_{s=1}^{n}a_s\psi_s(1)\Big]^3\psi_i(1) = f\psi_i(\bar{x}_F) \quad (i=1,2\ldots n) \tag{8}$$

where

$$\begin{aligned}
&A_{is}=0,\ A_{ii}=1,\ B_{is}=0,\ B_{ii}=\zeta+\nu\lambda_i^4\\
&C_{is}=\sigma D_{is}+\theta_1 E_{is}\\
&C_{ii}=\lambda_i^4+\sigma D_{ii}+\theta_1 E_{ii}\\
&D_{is}=G_{is}[t_{is}^2(2+G_{is})-T_{is}^2(\sigma_i\lambda_i-\sigma_s\lambda_s)-(-1)^{i+s}(4+G_{is})]\\
&D_{ii}=\frac{1}{2}\,\sigma_i^2\lambda_i^2-\sigma_i\lambda_i+2\\
&E_{is}=\psi_i(1)\psi_s(1),\ E_{ii}=\psi_i^2(1)\\
&t_{is}=\frac{\lambda_i}{\lambda_s},\quad G_{is}=\frac{4}{t_{is}^4-1}
\end{aligned} \tag{9}$$

In equation (9) σ_i and λ_i are the constants appearing in the characteristic functions of cantilevered beam vibration, $\sigma_i=0.73409,\ 1.0184\ldots\ldots$, $\lambda_i=1.8751,\ 4.694\ldots\ldots$ If

the system is subjected to a concentrated harmonic force $F_o \sin \omega t$, the dimensionless forcing frequency is

$$\bar{\omega} = \omega \left[\frac{m\ell^4}{EI}\right]^{1/2}$$

and then

$$F = F_o \sin \bar{\omega}\tau.$$

Numerical Simulations

Let

$$\frac{da_i}{d\tau} = a_i^{\tau},$$

then equations (8) is rewritten as

$$\frac{da_i^{\tau}}{d\tau} = f\psi_i(\bar{x}_F) - B_{ii}a_i^{\tau} - \sum_{s=1}^{n} C_{is}a_s$$

$$-\theta_2\left[\sum_{s=1}^{n} a_s\psi_s(1)\right]^3 \psi_i(1) \quad (i=1,2 \ldots n) \tag{10}$$

Solutions of equation (10) can be obtained using numerical integration. The fourth-order Runge-Kutta method was used to solve the above equations. We introduce a phase plane trajectory in the $\bar{w}, \dot{\bar{w}}$ plane to represent the solutions of equation (10). The phase plane trajectory depends on the system parameters as well as on the initial conditions $a_i(0)$ and $a_i^{\tau}(0)$.

In order to make a comparison of results from a numerical integration with those from our physical experiment, some of the parameters used in the calculations are taken from measurements. The dimensions of the rectangular permanent magnets which are made of a soft iron material are 16 cm (6.3 in.) long, 7 cm (2.75 in.) wide, and 3.5 cm (1.38 in.) thick. The dimensions of the steel beam are 54.5 cm (21.46 in.) long, 9.5 mm (3/8 in.) wide, and 0.635 mm (0.025 in.) thick. The dimensions of the Tygon tube are 54.5 cm (21.46 in.) long, 12.7 mm (1/2 in.) O.D., and 9.5 mm (3/8 in.) I.D. The flexural rigidity EI is found by measuring the frequency of the pipe-beam in its fundamental mode of oscillation as a cantilever. The circular frequency depends both on gravity and elasticity and is given approximately by

$$\frac{\omega_1^2\ell}{g} = \frac{81}{52} + \frac{162}{13}\frac{EI}{mg\ell^3}$$

as shown in Gregory and Paidoussis (1966).

The magnet stiffness terms K_1 and K_2 are found by measuring the first buckled natural freqeuncy for the postbuckling state of the system and the buckled static equilibrium position, and they are given approximately by

$$K_1 = -M_1(\omega_{b1}^2 + 2\omega_1^2)/8$$

$$K_2 = M_1(\omega_{b1}^2/(8w_d^2)$$

as shown in the Appendix.

The first mode damping ratio, ξ_1, is measured by the classical logrithmic decrement technique. The higher mode damping ratios, ξ_2 and ξ_3, are determined by impact test and measuring the frequency response function of the receptance.

The parameters of the system are listed in Table 1.

Table 1

1	54.5 cm
m	1.254*E-6 kg s^2/cm^2
EI	49.338 kg cm^2
f_1	1.453 HZ
f_2	7.669 HZ
f_3	20 HZ
f_{b1}	5.469 HZ
f_{b2}	11.72 HZ
K_1	−11.508*E-3 kg/cm
K_2	2793*E-6 kg/cm^3
w_b	1.9 cm
ξ_1	.0191
ξ_2	.0339
ξ_3	.0437

where $2\xi_i\lambda_i^2 = \zeta_i$.

Strictly speaking, in the above formula λ_i^2 should be replaced by $[\lambda_i^4 + \sigma D_{ii}]^{1/2}$, see equation (9). However, the differences are small.

In this paper numerical simulations for periodic and chaotic response are conducted to determine the threshold force for chaotic motions, especially the minimum threshold force. In the calculations, equations (10) with different mode numbers $n = 1,2,3$, are considered. The initial conditions are taken to be the same for various mode numbers, $a_1(0) = 0.0176$ (corresponding to a positive buckled equilibrium position), $a_1^{\tau}(0) = 0$, and $a_i^{\tau}(0) = a_i(0) = 0$ $(i = 2,3)$. The calculated data were obtained by fixing the forcing frequency, f, and increasing the exciting force level until chaos is observed in the phase plane trajectory map. The results will be presented in the section on Results and Discussion. Phase plane trajectory data were also used for experimental determination of chaos, though simple visual observation proved to be a reliable indicator as well. The threshold force level separating periodic and chaotic motion was sharply defined both theoretically and experimentally and did not seem especially sensitive to initial conditions, etc. All numerical simulations were done on a Data General MV/8000 computer.

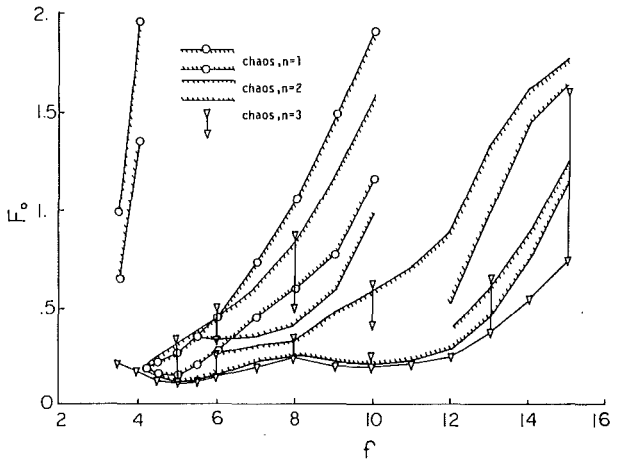


Fig. 2 Threshold force versus forcing frequency for *A*-damping model, $\xi_1 = 0.0191$

Experimental Verification

In order to check the validity of the theoretical results, a forced vibration test was performed on the pipe-beam. The geometric dimensions and the material properties of the system are as given before.

There are four major components of the vibration test: excitation, measurement, recording, and analysis. A magnetic shaker, B&K 4810, which is mechanically attached to the pipe-beam and a power amplifier, B&K 2706, are used to excite the system. It is driven by a sine generator, HP 202A. In order to reduce the effect of the additional mass of the joint components between the shaker and the system on the vibration characteristics, the position of the driving point is located near the clamped end, $x_F = 6$ cm. The force transducer, B&K 8200, displacement transducer, R30D, and accelerometer, B&K 4375 (the velocity signal is obtained through integration networks in the charge amplifier, B&K 2536), are used to measure the exciting force value and the phase plane trajectory map. The

measurement point for displacement and acceleration is at $x = 50$ cm.

A 8-channel tape recorder, HP 3968A, is used to record all signals. The signal analysis is completed by a multichannel signal processor, ZONIC 6088. The experimental results will be discussed in the following section.

Results and Discussion

The theoretical threshold forces versus forcing frequency for a given Kelvin damping $\nu = 0.0191$ and zero viscous damping are shown in Fig 2. There are three groups of the curves on this figure for mode numbers $n = 1$, 2, and 3, respectively. For $n = 1$, the threshold force for the onset of chaos increases as forcing frequency increases both for lower and upper force chaotic boundaries in the region above 5 HZ. The minimum threshold force occurs near $f/f_{b1} = 0.88$. For $n = 2$, the chaotic force boundaries are quite different with those for $n = 1$. It is very clear that the second mode strongly affects chaotic response. From Table 1, it is seen that the second natural frequency is near the first buckled natural frequency, as well as the second buckled natural frequency. For $n = 3$, in the region of lower frequencies (lower than 8 HZ) the agreement is good in comparison with those for $n = 2$. But in the region of higher frequencies the difference is larger.

Figures 3(*a*)–3(*e*) show a typical evolution process of the phase plane trajectory map for $n = 2$, and forcing frequency, $f = 6$ HZ. The interesting phenomenon that periodic and chaotic responses alternatively occur as the force level increases has been explained in Dowell and Pezeshki (1986).

Figures 4(*a*)–4(*b*) show typical patterns of the phase plane trajectory map, and its PSD, in the chaotic region for the lower forcing frequencies as obtained from experiments. A continuous spectrum of frequencies, including the forcing frequency and many subharmonics, displays one of the essential characteristics of chaotic motions.

It is observed from the experimental data that there are two kinds of vibration shape for the chaotic oscillations. In the region of lower forcing frequencies the system oscillation is dominated by the first natural mode of the cantilever beam as shown in Fig. 4. In the region of higher forcing frequencies, the system oscillation is dominated by the second natural mode. The phase plane trajectory map and its PSD are shown in Figs. 5(*a*) and 5(*b*).

The effect of damping on threshold force for the onset of chaos is very significant. Here, we only discuss the effects of damping on the minimum threshold force. At the same time, a comparison of the theory with the experimental data for chaotic motions is presented.

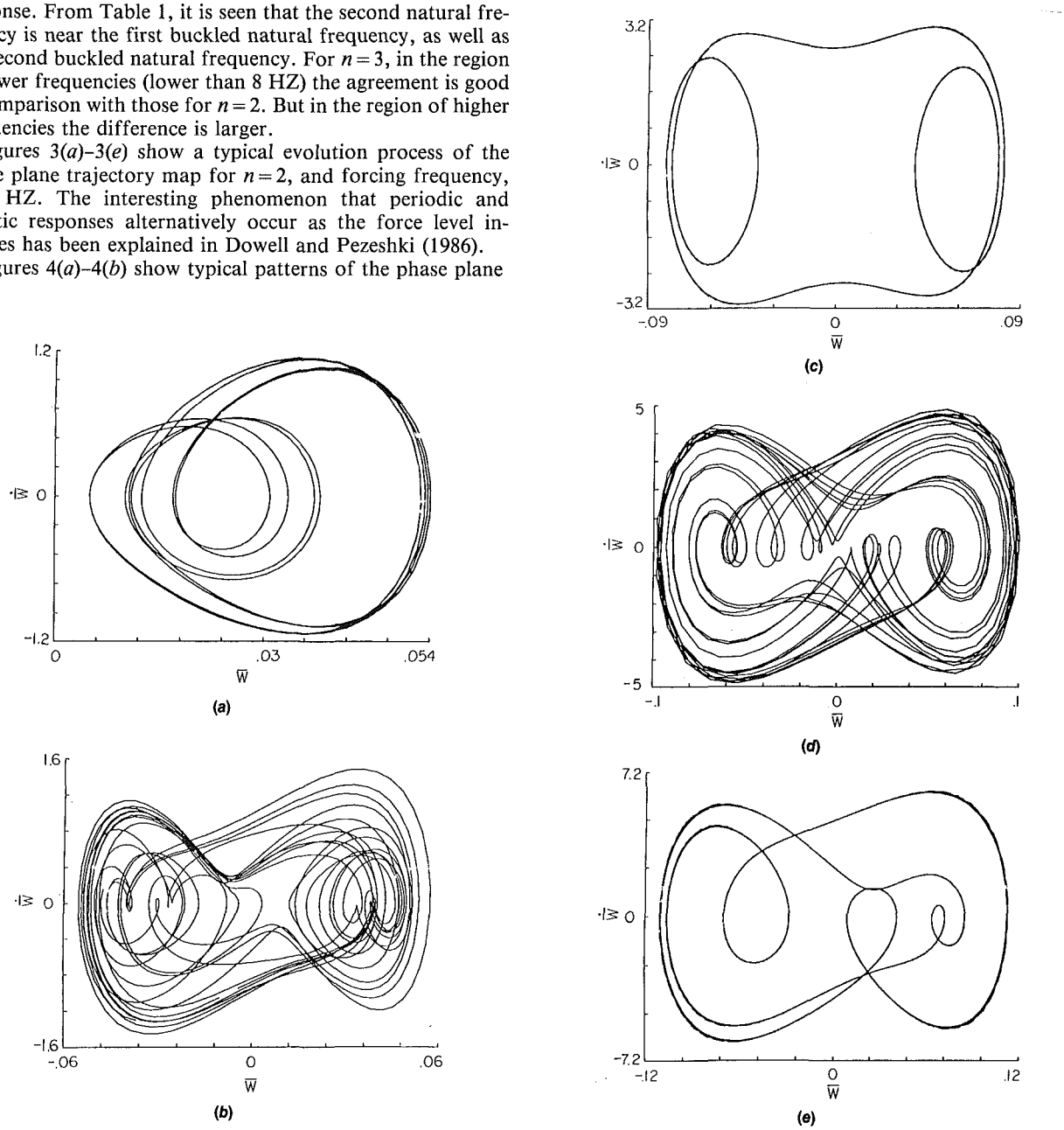


Fig. 3 Evolution of phase plane trajectory in case of $n = 2$, $f = 6$ HZ: (*a*) For $F_0 = 0.12$ kg; (*b*) For $F_0 = 0.25$ kg; (*c*) For $F_0 = 0.31$ kg; (*d*) For $F_0 = 0.40$ kg; (*e*) For $F_0 = 0.50$ kg

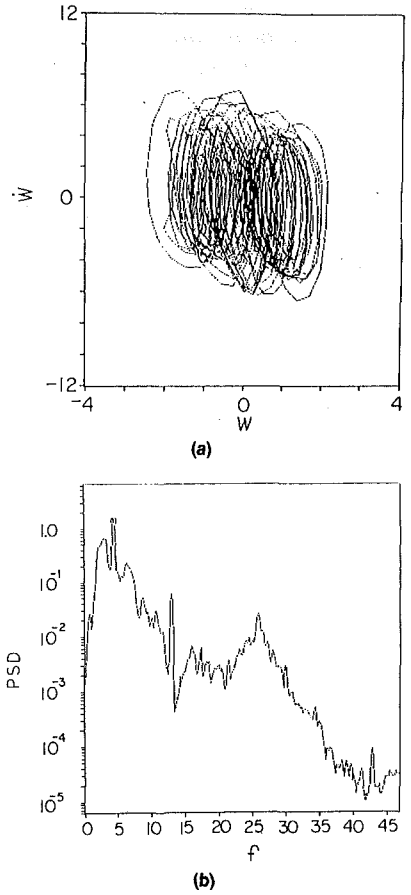


Fig. 4 Experimental results for $F_0 = 0.13$ kg, $f = 4.22$ HZ: (*a*) phase plane trajectory; (*b*) PSD

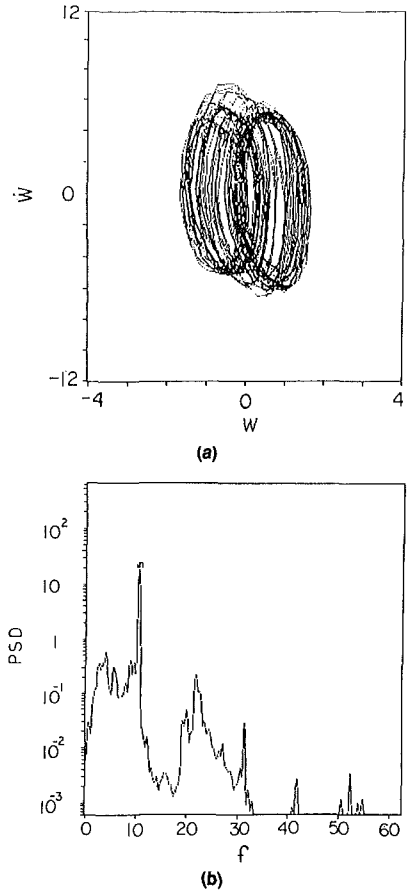


Fig. 5 Experimental results for $F_0 = 0.12$ kg, $f = 11$ HZ: (*a*) phase plane trajectory; (*b*) PSD

In equation (8), three types of damping models are considered, respectively. They are called *A*, *B*, and *C*-damping models (DM)

$$\text{for } A\text{-DM}, \qquad B_{ii} = \nu\mu_i^4 = \xi_1\lambda_i^4$$

$$\text{for } B\text{-DM}, \qquad B_{ii} = \zeta_i = 2\xi_1\lambda_i^2$$

$$\text{for } C\text{-DM}, \qquad B_{ii} = \zeta_i = 2\xi_i\lambda_i^2$$

The results of numerical integration are plotted in Fig. 6 and Fig. 7 for *B* and *C*-DM, respectively. The result for *A*-DM is shown in Fig. 2. Comparison of calculated results for $n = 3$ with experimental data in the case of different damping models are shown in Fig. 8. It is interesting to note that the agreement between theory and the experiment is best for *C*-DM, less good for *B*-DM, and least good for A-DM. This suggests that the pipe beam is not a real viscoelastic material. The assumption of Kelvin damping is not suitable for the present system. The damping distribution of this system is more nearly a viscous type damping.

It is also found that the agreement between theory and experiment becomes better as the mode number, n, increases. For the present system, taking $n = 3$ is sufficient. This improved agreement as n increases would also appear to explain the (small) differences between the one mode ($n = 1$) theory of Dowell and Pezeshki (1986) and the Moon experiment (1980).

Simplified Threshold Criteria. The threshold criterion presented by Holmes-Melnikov is (Holmes, 1979)

$$f > f_1 \equiv \frac{\gamma\sqrt{2}}{3\pi\omega^*} \cosh\left(\frac{\pi\omega^*}{\sqrt{2}}\right) \tag{11}$$

and by Moon (1980) is

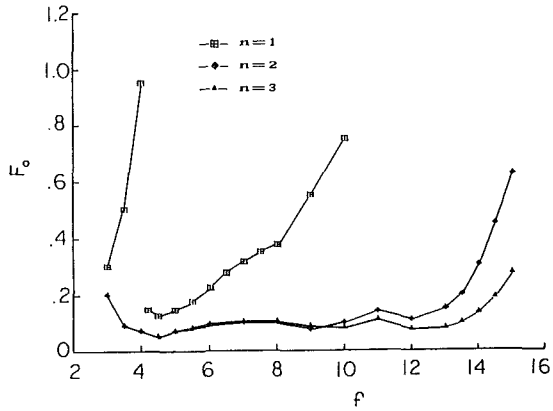


Fig. 6 Minimum threshold force versus forcing frequency for *B*-DM

$$f > f_1 \equiv \frac{\alpha}{2\omega^*}\left\{\left[(1-\omega^{*2}) - \frac{3}{8}\frac{\alpha^2}{\omega^{*2}}\right]^2 + \gamma^2\omega^{*2}\right\}^{1/2} \tag{12}$$

where

$$\omega^* = \omega/\omega_{b1}.$$

Moon's threshold criterion is based on base displacement excitation of the cantilever beam. In order to apply the above criterion to the present model, the following correspondence of the parameters must be satisfied:

$$f = \frac{1}{\omega_{b1}^2 w_b}\frac{F_0\phi_1(x_F)}{M_1}, \quad \gamma = \frac{2\xi_1\omega_1}{\omega_{b1}} \tag{13}$$

These are obtained by comparing equation (7) for the case of $n = 1$ with equation (4) and equation (5) in Moon (1980).

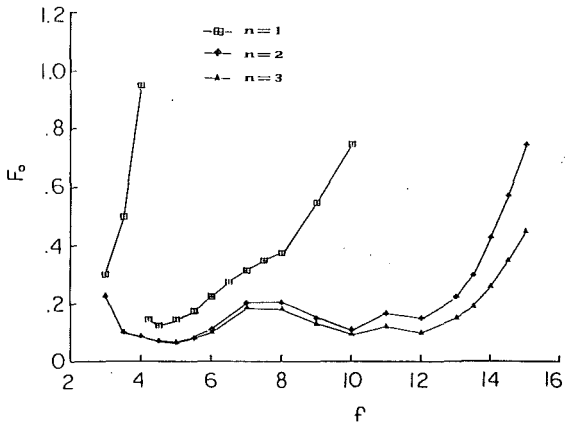


Fig. 7 Minimum threshold force versus forcing frequency for C-DM

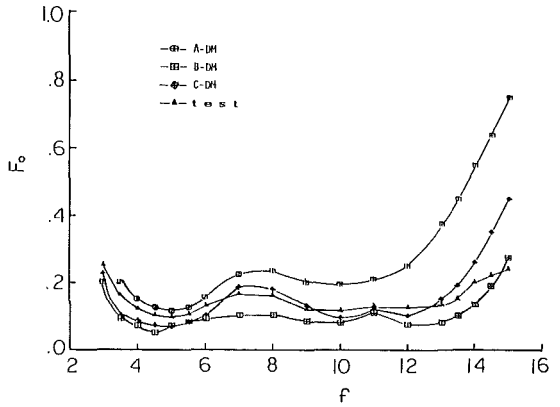


Fig. 8 Minimum threshold force versus forcing frequency for $n = 3$

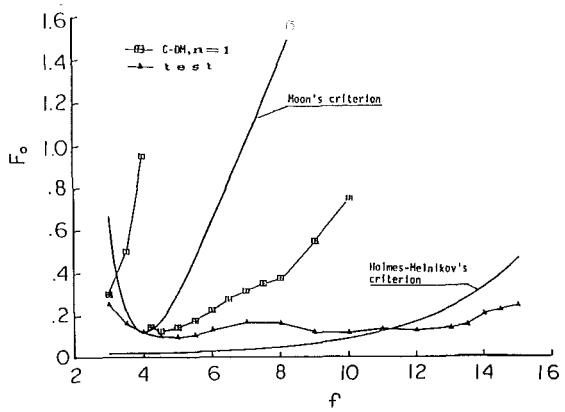


Fig. 9 Experimental and theoretical thresholds for $\xi_1 = 0.0191$, $\alpha = 0.86$

Fig. 9 shows the comparison of the simplified threshold criteria with the theoretical and experimental results for $\xi_1 = 0.0191$ and $\alpha = 0.86$. From Fig. 9 it is seen that the Holmes-Helnikov criterion is only approximate. Moon's threshold criterion was somewhat better for his experiment with a cantilever beam (Moon, 1980) because the second natural frequency (26.6 HZ) in his experiment was much higher than the first buckled natural frequency (9.34 HZ). Both simplified criteria are based upon a one-mode model and the assumption that the excitation frequency is near the (first) beam natural frequency.

Conclusions

1. When the forced response of a buckled beam is studied, it is generally necessary to consider the effects of higher beam modes on the response. The importance of the effect depends on the proximity of the lowest buckled natural frequency to the natural frequency of the second mode.

2. Various types of damping and the modal damping levels will influence significantly the threshold force for the onset of chaos. A careful evaluation of damping is important in predicting and understanding chaos.

3. Good quantitative agreement has been obtained between theory and experiment when a sufficient number of modes is included in the theoretical model and the damping is carefully evaluated.

Acknowledgment

This work was supported by the National Science Foundation under Grant No. MSM-85-04105 and The Army Research Office under Grant DAA-87-K-0023. Drs. Elbert Marsh and Gary Anderson are the technical monitors.

References

Dowell, E. H., and Pezeshki, C., (1986), "On the Understanding of Chaos in Duffing's Equation Including a Comparison With Experiment," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 5–9.

Gregory, R. W., and Paidoussis, M. P., (1966), "Unstable Oscillation of Tubular Cantilevers Conveying Fluid: II. Experiments," *Proc. of the Royal Soc. of London*, Vol. 293 (series A), pp. 512–542.

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

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

Moon, F. C., (1980), "Experiments on Chaotic Motions of a Forced Nonlinear Oscillator: Strange Attractors," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 683–644.

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

Ueda, Y., (1980), "Steady Motions Exhibited by Duffing's Equation: A Picture Book of Regular and Chaotic Motions," *New Approaches to Nonlinear Problems in Dynamics*, Holmes, P. J., ed., SIAM, PA.

APPENDIX

Determination of Magnet Stiffness Terms

The magnet stiffness terms K_1 and K_2 are very important parameters for the chaotic motions of a buckled beam. Moon (1980) used a direct measurement method to determine these values. It is not always convenient. Here, a more convenient method is suggested to obtain K_1 and K_2.

Consider a one-mode vibration analysis without damping and external forces in equation (7). The equation is simplified to

$$M_1[\ddot{q}_1 + \omega_1^2 q_1] + K_1 q_1 \phi_1^2(\ell) + K_2 q_1^3 \phi_1^4(\ell) = 0 \tag{1}$$

where

$$M_1 = m\int_o^\ell \phi^2(x)dx$$

Considering the buckled beam to be at rest in one of its stable static equilibria, equation (1) becomes

$$[M_1\omega_1^2 + K_1\phi_1^2(\ell)]q_{b1} + K_2 q_{b1}^3 \phi_1^4(\ell) = 0 \tag{2}$$

Solving

$$q_{b1}^2 = -\frac{[M_1\omega_1^2 + K_1\phi_1^2(\ell)]}{K_2\phi_1^4(\ell)} \tag{3}$$

where q_{b1} is the generalized coordinate of the buckled equilibrium position. Now $w_b = q_{b1}\phi_1(\ell)$

and

$$w_b^2 = -\frac{[M_1\omega_1^2 + K_1\phi_1^2(\ell)]}{K_2\phi_1^2(\ell)} \tag{3a}$$

Let

$$q_1 = q_{b1} + \hat{q}_1 \tag{4}$$

where $\hat{q}_1$ is a small perturbation value of the postbuckled beam. Substituting equation (4) into equation (1), it is found that

$$M_1[\ddot{\hat{q}}_1 + \omega_1^2(q_{b1} + \hat{q}_1)] + K_1(q_{b1} + \hat{q}_1)\phi_1^2(\ell) + K_2\phi_1^4(\ell)\,(q_{b1}^3 + 2q_{b1}^2\hat{q}_1 + \ldots\ldots) = 0 \tag{5}$$

Substituting equation (2) into equation (5) one obtains

$$M_1\ddot{\hat{q}}_1 + \omega_1^2\hat{q}_1 + K_1\hat{q}_1\phi_1^2(\ell) + 3K_2\hat{q}_1 q_{b1}^2\phi_1^4(\ell) = 0 \tag{6}$$

Let

$$\hat{q}_1 = \hat{q}_o e^{i\omega_{b1}t} \tag{7}$$

Substituting equation (3) and equation (7) into equation (6) gives

$$K_1 = -\frac{M_1(\omega_{b1}^2 + 2\omega_1^2)}{2\phi_1^2(\ell)} \tag{8}$$

and from equation (3*a*)

$$K_2 = -\frac{[M_1\omega_1^2 + K_1\phi_1^2(\ell)]}{w_b^2\phi_1^2(\ell)} \tag{9}$$

Now $\phi_1^2(\ell) = 4$, and thus

$$K_1 = -M_1(\omega_{b1}^2 + 2\omega_1^2)/8$$
$$K_2 = M_1\omega_{b1}^2/(8w_b^2) \tag{10}$$

Equations (10) are used to determine K_1 and K_2 with the several parameters in the right-hand sides determined from measurement.

Approximate Nonstationary Random Vibration Analysis for MDOF Systems

C. G. Bucher
Institute of Engineering Mechanics,
University of Innsbruck,
A-6020 Innsbruck, Austria

An approximate method for nonstationary random vibration analysis is presented. This method utilizes properties of the stationary solution for simplifying the analysis. This approach has previously been applied by the author to linear and nonlinearly damped SDOF systems. In the present paper the concept is extended to linear MDOF systems and applied to nonstationary earthquake-type loading. Comparisons with available exact solutions show very good agreement in numerical results with the additional benefit of reducing computer time by more than one order of magnitude.

Introduction

Load processes encountered in engineering analysis may exhibit strongly nonstationary features (e.g., earthquake loading). It is important to determine the nonstationary stochastic properties of the response of structural systems to this type of loading, especially for realistic reliability estimates. In the case of linear MDOF system subjected to Gaussian loading, the stochastic analysis may be confined to the covariance matrix.

In general, two ways are open, i.e., the analysis in the time domain or in the frequency domain. Frequency domain analyses require the evaluation of an evolutionary spectral density function (Hammond, 1968). In this case, computer time may become prohibitively long due to many numerical integrations for each time step (e.g., Yang and Iwan, 1972). CPU reduction may be achieved through methods avoiding modal analysis, e.g., transfer matrix methods (Yang et al., 1980), but the problem of integration remains. If the excitation can be represented by white noise (or nonstationary shot noise), time domain method utilizing a state vector (Markov vector) formulation are more suitable. Computational procedures were developed for classically damped systems (Gasparini and DebChaudhury, 1980), as well as nonclassically damped systems (Der Kiureghian and Igusa, 1985). These methods use $2N \times 2N$ matrices, if N is the number of DOF of the system.

An approximate solution for linear SDOF systems, as suggested by Solomos and Spanos (1984), utilizes the Fokker-Planck Equation and stochastic averaging procedures (Stratonovich, 1967). Due to it complexity, however, this approach does not seem to be applicable to MDOF systems. The approximate method discussed in this paper is an improvement of existing approximations utilizing basic relations of stationary covariance analysis. For SDOF systems this method yields the same result as the approach used by Solomos and Spanos (1984), but as shown it is, moreover, easily applicable to nonlinearly damped systems (Bucher, 1986a) and linear MDOF systems. Application to the MDOF case leads to the advantage that only $N \times N$ matrices have to be used in the analysis, which reduces computer time considerably and which, together with its versatility, seems to be a major advantage of the presented approach.

Statement of Problem

The equation of motion for a N-DOF linear system is

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{f}(t) \tag{1}$$

or in first order form

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{f} \end{bmatrix} \tag{2}$$

in which $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are the mass, damping and stiffness matrices of order $N \times N$; $\mathbf{x}$, $\mathbf{y} = \dot{\mathbf{x}}$ and $\mathbf{f}$ are the displacement, velocity and load vectors, respectively. Introducing the state vector $\mathbf{z} = [\mathbf{x}, \mathbf{y}]^T$, equation (2) becomes

$$\dot{\mathbf{z}} = \mathbf{G}\mathbf{z} + \mathbf{f}^* \tag{3}$$

in which $\mathbf{G}$ is a $2N \times 2N$ matrix. The definitions of $\mathbf{G}$, $\mathbf{f}^*$ may be seen from comparing equations (2), (3). In the following shot noise (evolutionary white noise) properties of the load process $\mathbf{f}^*(t)$ are assumed, i.e.

$$\mathbf{R}_f^*(t, t+\tau) = \mathbf{B}^*(t)\delta(\tau) \tag{4}$$

in which $\mathbf{B}^*(t)$ is the time dependent cross intensity matrix and $\delta(.)$ is Dirac's delta function. In that case the covariance

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, September 8, 1986.

Paper No. 88-WA/APM-9.

matrix $\mathbf{S}$ of the state vector $\mathbf{z}$ is governed by the differential equation (Lin, 1967)

$$\dot{\mathbf{S}} = \mathbf{GS} + \mathbf{SG}^T + \mathbf{B}^* \tag{5}$$

Colored load processes of the Kanai-Tajimi type may be modeled by a stable second-order filter, so that the equation of motion for base acceleration becomes

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = -\mathbf{M}\mathbf{r}a(t) = -\mathbf{M}\mathbf{r}(2D_g\omega_g\dot{x}_a + \omega_g^2 x_a) \tag{6a}$$

together with the filter equation

$$\ddot{x}_a + 2D_g\omega_g\dot{x}_a + \omega_g^2 x_a = f(t) \tag{6b}$$

In equation (6*a*) $\mathbf{x}$ is the vector of displacements relative to ground, $\mathbf{r}$ is the influence vector coupling the ground motion to the degrees of freedom of the structure. The parameter ω_g and D_g may be interpreted as dominant ground frequency and damping, respectively; $f(t)$ is assumed evolutionary white noise. By augmenting the state vector $\mathbf{z}$ with the filter variables x_a and $\dot{x}_a$, equations (6) may be cast into the same form as equation (2). Thus the covariance matrix of the augmented state vector $\mathbf{z}_a$ is also governed by equation (5).

An exact solution to equation (5) is obtained by diagonalizing the matrix $\mathbf{G}$ by means of a complex eigenvalue analysis, so that the differential equations (5) become decoupled. This solution is used for reference in the numerical examples.

Approximate Analysis

For systems with classical damping, the equations of motion equation (1) may be decoupled using standard modal analysis

$$\ddot{\mathbf{x}}_p + \mathbf{C}_p\dot{x}_p + \mathbf{K}_p\mathbf{x}_p = \mathbf{f}_p(t) \tag{7}$$

in which $\mathbf{C}_p = \text{diag}(2D_i\omega_i)$ and $\mathbf{K}_p = \text{diag}(\omega_i^2)$ are diagonal matrices. ω_i are the eigenfrequencies, D_i the modal damping ratios. The modal load vector $\mathbf{f}_p$ may be described by its cross intensity matrix $\mathbf{B}_p$ if $\mathbf{f}_p$ is evolutionary white noise. Partitioning of the modal covariance matrix $\mathbf{S}_p$ into four submatrices

$$\mathbf{S}_p = \begin{bmatrix} \mathbf{S}_{pxx} & \mathbf{S}_{pxy} \\ \mathbf{S}_{pyx} & \mathbf{S}_{pyy} \end{bmatrix} \tag{8}$$

equation (5) may be written in the form of four submatrix equations

$$\dot{\mathbf{S}}_{pxx} = \mathbf{S}_{pyx} + \mathbf{S}_{pxy} \tag{9a}$$

$$\dot{\mathbf{S}}_{pxy} = \mathbf{S}_{pyy} - \mathbf{S}_{pxx}\mathbf{K}_p - \mathbf{S}_{pxy}\mathbf{C}_p \tag{9b}$$

$$\dot{\mathbf{S}}_{pyx} = \mathbf{S}_{pyy} - \mathbf{K}_p\mathbf{S}_{pxx} - \mathbf{C}_p\mathbf{S}_{pyx} \tag{9c}$$

$$\dot{\mathbf{S}}_{pyy} = -\mathbf{K}_p\mathbf{S}_{pxy} - \mathbf{S}_{pyx}\mathbf{K}_p - \mathbf{C}_p\mathbf{S}_{pyy} - \mathbf{S}_{pyy}\mathbf{C}_p + \mathbf{B}_p \tag{9d}$$

Performing some matrix algebra on equations (9) yields

$$\begin{aligned} 2\dot{\mathbf{S}}_{pyy} + \mathbf{K}_p\dot{\mathbf{S}}_{pxx} + \dot{\mathbf{S}}_{pxx}\mathbf{K}_p &= \mathbf{K}_p\mathbf{C}_p^{-1}\mathbf{S}_{pyy} - \mathbf{K}_p\mathbf{S}_{pyy}\mathbf{C}_p^{-1} \\ &- \mathbf{K}_p\mathbf{C}_p^{-1}\mathbf{K}_p\mathbf{S}_{pxx} + \mathbf{K}_p\mathbf{S}_{pxx}\mathbf{K}_p\mathbf{C}_p^{-1} - \mathbf{C}_p^{-1}\mathbf{S}_{pyy}\mathbf{K}_p \\ &+ \mathbf{S}_{pyy}\mathbf{C}_p^{-1}\mathbf{K}_p + \mathbf{C}_p^{-1}\mathbf{K}_p\mathbf{S}_{pxx}\mathbf{K}_p - \mathbf{S}_{pxx}\mathbf{K}_p\mathbf{C}_p^{-1}\mathbf{K}_p \\ &- 2\mathbf{B}_p\mathbf{S}_{pyy} - 2\mathbf{S}_{pyy}\mathbf{C}_p - 2\mathbf{B}_p + \mathbf{E} \end{aligned} \tag{10}$$

The elements E_{ij} of the matrix $\mathbf{E}$ are

$$E_{ij} = \dot{S}_{x_{pi}y_{pj}}\frac{\omega_i^2 - \omega_j^2}{D_j\omega_j} \tag{11}$$

so that these elements are zero for $i=j$. The off-diagonal terms of $\mathbf{E}$ will be neglected in the following (which is exact in the stationary state), so the equation (10) contains the submatrices $\mathbf{S}_{pxx}$ and $\mathbf{S}_{pyy}$ and their time derivatives only.

It is most important to note that an approximate relation between the elements of $\mathbf{S}_{pxx}$ and $\mathbf{S}_{pyy}$ may be obtained by utilizing the stationary form of equations (9*a*) to (9*c*) (Bucher, 1986b)

$$S_{y_{pi}y_{pj}} = \omega_i\omega_j\frac{D_i\omega_j + D_j\omega_i}{D_i\omega_i + D_j\omega_j}S_{x_{pi}x_{pj}} \tag{12}$$

Substituting equation (12) into equation (10), the differential equations for the elements of $\mathbf{S}_{pxx}$ decouple, so that

$$a_{ij}\dot{S}_{x_{pi}x_{pj}} = b_{ij}S_{x_{pi}x_{pj}} + B_{pij}(t) \tag{13}$$

in which

$$a_{ij} = \omega_i\omega_j\frac{D_i\omega_j + D_j\omega_i}{D_i\omega_i + D_j\omega_j} + \frac{\omega_i^2 + \omega_j^2}{2} \tag{14a}$$

$$\begin{aligned} b_{ij} = \omega_i\omega_j &\frac{D_i\omega_j + D_j\omega_i}{D_i\omega_i + D_j\omega_j}\left(\frac{\omega_i}{4D_i} + \frac{\omega_j}{4D_j}\right. \\ &\left. - \frac{\omega_i^2}{4D_j\omega_j} - \frac{\omega_j^2}{4D_i\omega_i} - 2D_i\omega_i - 2D_j\omega_j\right) \\ &+ \frac{\omega_i^2\omega_j}{4D_j} + \frac{\omega_i\omega_j^2}{4D_i} - \frac{\omega_i^3}{4D_i} - \frac{\omega_j^3}{4D_j} \end{aligned} \tag{14b}$$

Equation (13) may be solved in a straightforward way. The covariance matrix of the displacements in physical coordinates is obtained from retransforming by means of the modal matrix.

It is to be mentioned that for $i=j$ equation (13) reduces to

$$\dot{S}_{x_{pi}x_{pi}} + 2D_i\omega_i S_{x_{pi}x_{pi}} = \frac{B_{pii}}{2\omega_i^2} \tag{15}$$

which is identical to the SDOF result obtained by Solomos and Spanos (1984), however, it was achieved by avoiding their somewhat lengthy stochastic averaging procedure.

Another response quantity of potential interest is the covariance matrix of the absolute accelerations. Rewriting equation (7) for base acceleration in the form

$$\mathbf{b}_p + \mathbf{C}_p\dot{\mathbf{x}}_p + \mathbf{K}_p\mathbf{x}_p = \mathbf{0} \tag{16}$$

in which $\mathbf{b}_p$ is the vector of absolute modal accelerations the covariance matrix $\mathbf{S}_{pbb}$ becomes

$$\mathbf{S}_{pbb} = \mathbf{C}_p\mathbf{S}_{pyy}\mathbf{C}_p + \mathbf{C}_p\mathbf{S}_{pyx}\mathbf{K}_p + \mathbf{K}_p\mathbf{S}_{pxy}\mathbf{C}_p + \mathbf{K}_p\mathbf{S}_{pxx}\mathbf{K}_p \tag{17}$$

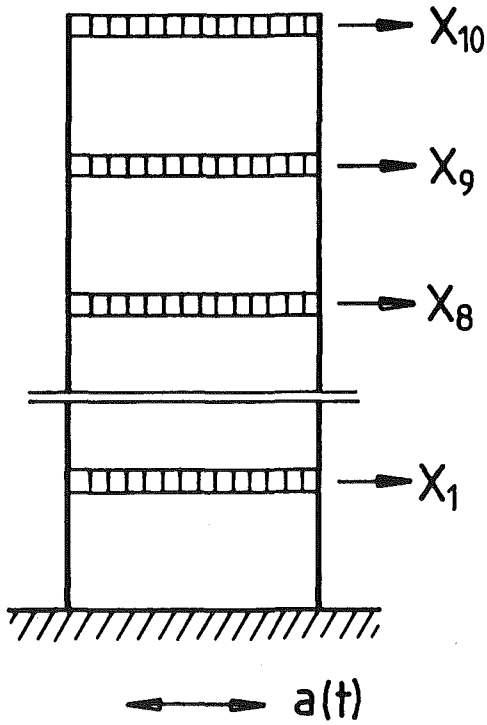


Fig. 1 Structure used for numerical examples

Table 1 System and ground parameters for numerical examples

Case	ω_1	$D_1 - D_{10}$	ω_g	D_g	Error in extreme m.s. displacement	Error in extreme m.s. abs. accel.
a	14.900	0.05	15.0	0.65	0.4%	3%
b	14.900	0.05	15.0	0.35	0.5%	3%
c	14.900	0.05	15.0	0.25	0.4%	2.5%
d	14.900	0.02	15.0	0.35	0.03%	1%
e	14.900	0.10	15.0	0.35	0.5%	5%

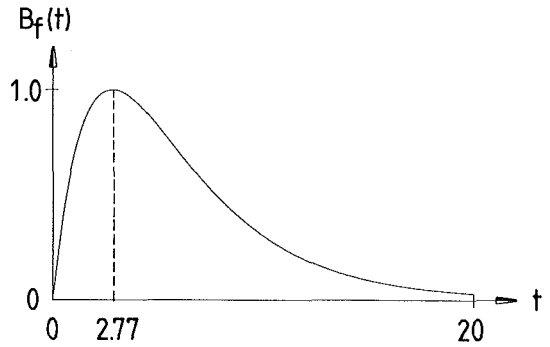


Fig. 2 Intensity of shot noise input to ground filter

Again utilizing properties of the stationary solution (Bucher, 1986b) the elements of $\mathbf{S}_{pbb}$ may be expressed by

$$S_{b_{pi}b_{pj}} = \left[(4D_iD_j\omega_i\omega_j + \omega_i^2 + \omega_j^2)\cdot\omega_i\omega_j \frac{D_i\omega_j + D_j\omega_i}{D_i\omega_i + D_j\omega_j} - \omega_i^2\omega_j^2\right] S_{x_{pi}x_{pj}} \tag{18}$$

so that only the mean square displacements are required.

So far, the derivations have been concerned with white noise loading. In the next step, this approximate method is extended to colored load processes of the Kanai-Tajimi type. For this purpose the modal cross intensity matrix $\mathbf{B}_p$ of equivalent shot noise must be defined suitably to represent essential features of the colored ground motion.

First, it is assumed that the evolutionary power spectral density of the ground acceleration $\Phi_a(\omega,t)$ may be approximated by

$$\Phi_a(\omega,t) = e(t)\Phi_{a,\infty}(\omega) \tag{19}$$

in which $e(t)$ is a modulating function and $\Phi_{a,\infty}(\omega)$ is the well known Kanai-Tajimi spectrum

$$\Phi_{a,\infty}(\omega) = \frac{1 + \left(\frac{2D_g\omega}{\omega_g}\right)^2}{\left[1 - \left(\frac{\omega}{\omega_g}\right)^2\right]^2 + \left(\frac{2D_g\omega}{\omega_g}\right)^2} \tag{20}$$

Since the variance of the ground acceleration $\sigma_a^2(t)$ can easily be evaluated exactly (Bucher, 1986b) $e(t)$ may be chosen to yield this exact variance if equation (19) is integrated:

$$\sigma_a^2(t) = \int_{-\infty}^{\infty} S_a(\omega,t)d\omega = e(t)\frac{\pi\omega_g}{2D_g}(1+4D_g^2) \tag{21}$$

Second, the colored load process is replaced by evolutionary white noise with modal intensities $B_{pii}(t)$ which are chosen as to predict the modal variances $S_{x_{pi}x_{pi}}(\infty)$ exactly in the stationary state. If, additionally, equation (21) is satisfied it follows that

$$B_{pii}(t) = \frac{8D_g}{\pi\omega_g}\frac{S_{x_{pi}x_{pi}}(\infty)D_i\omega_i^3}{1+4D_g^2}\sigma_a^2(t) \tag{22}$$

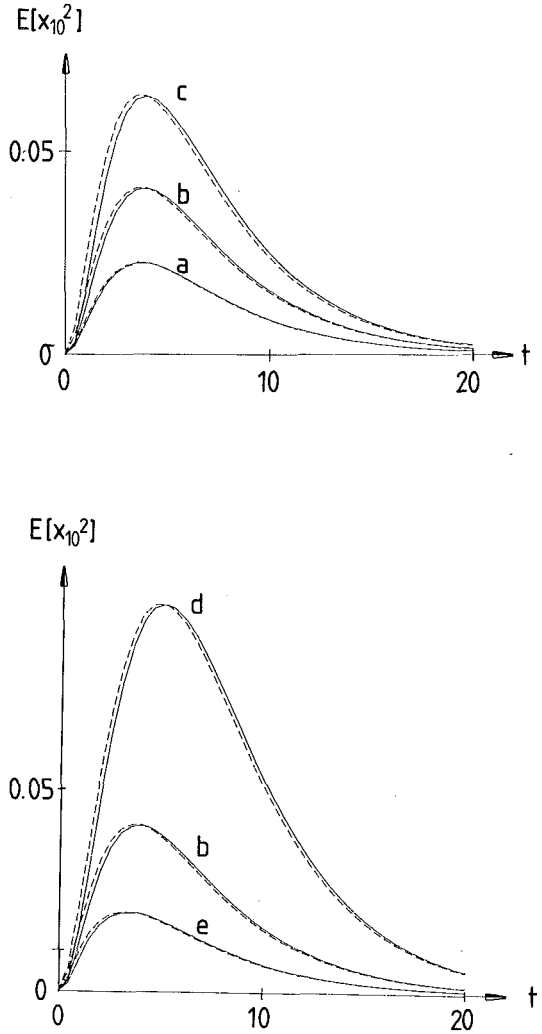


Fig. 3 Mean square displacement of top floor relative to ground (solid line indicates exact results, dashed line indicates approximation)

in which the stationary variances $S_{x_{pi}x_{pi}}(\infty)$ may easily be obtained by standard methods. Finally the off-diagonal terms $B_{pij}(t)$ are defined by

$$B_{ij}(t) = \sqrt{B_{ii}(t)B_{jj}(t)} \tag{23}$$

implying full correlation of the modal excitations (which is to be expected for earthquake loading).

Numerical Example

A ten-story building under earthquake-type loading is considered for numerical comparisons. The structure is modeled as a plane frame with identical floor masses and interstory stiffnesses (Fig. 1). Numerical values for system and ground parameters are indicated in Table 1. The intensity of the shot noise $f(t)$ driving the ground filter is assumed to be $B_f(t) = 16[\exp(-0.25t) - \exp(-0.5t)]^2$ (see Fig. 2). Figures 3(*a*) and 3(*b*) compare the exact results (solid line) to the results obtained by the approximate method as suggested here (dashed line) for the mean square displacement of the top floor relative to ground for cases *a* to *e* (cf Table 1). Figures 4(*a*) and 4(*b*) compare the mean square absolute acceleration of the top floor for the same cases. The comparison between exact and approximate method indicates that the latter shows an extremely good accuracy, even for very high values of structural damping, as well as for the case of low ground damping and ground frequency close to natural frequencies of

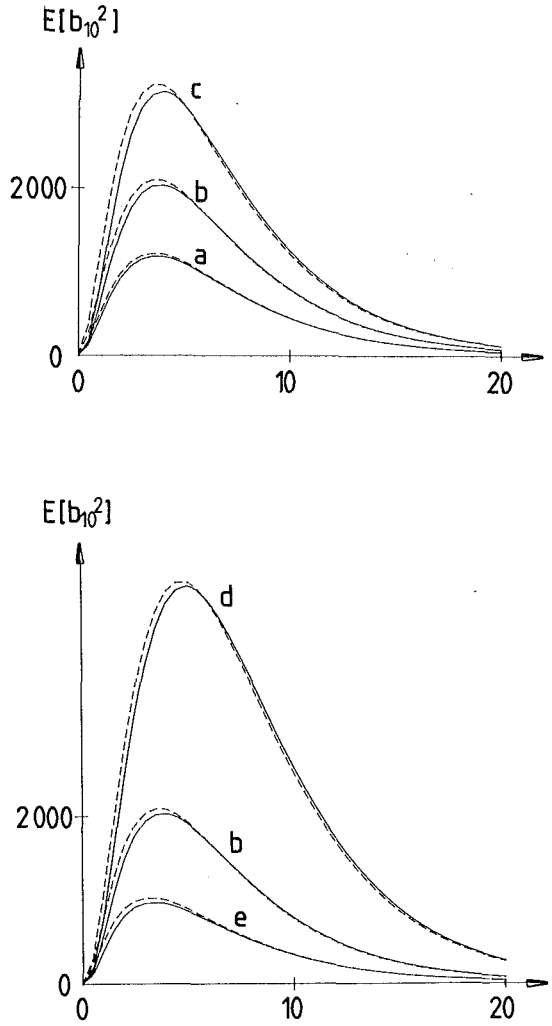


Fig. 4 Mean square absolute acceleration of top floor (solid line indicates exact results, dashed line indicates approximation)

the structure. The error in the maximum mean square displacement is found to be smaller than 1 percent for the cases presented here. For the mean square accelerations, typical errors are still in the range of 3 percent (cf Table 1).

Conclusions

This paper presents an approximate solution for nonstationary covariance analysis. It is shown that basic *assumptions of stationary relations* lead to more efficient approximate methods than existing ones. For linear SDOF systems, the result obtained by the present approach equals the result obtained by Solomos and Spanos (1984); however, it avoids their lengthy and rather limited procedure (which seems not to be applicable to nonlinear SDOF as well as linear MDOF systems). In contrast, the methods as presented here is shown to be easily applicable to the analysis of MDOF systems even if colored loading is considered. This holds not only for the mean square displacements but also for other response statistics, such as mean square absolute accelerations which can be evaluated on the basis of displacements only with considerable accuracy. Moreover, the versatility of the present approach has already been shown by applying it to nonstationary variance analysis of nonlinearly damped SDOF systems (Bucher, 1986a).

Finally, it is pointed out that due to the reduction of the matrices used in the analysis to half size, as compared to existing computational procedures, CPU time is reduced considerably. Furthermore, this proposed new method may easily be implemented into existing Finite Element Codes without significant increase of storage requirements.

Acknowledgment

Partial support of this research by the Austrian Research Association (FWF) under Contract No. S30-03 is acknowledged. Special thanks are due to G. I. Schuëller for his encouragement and helpful discussions.

References

Bucher, C. G., 1986a, "Response of Nonlinearly Damped Systems to Nonstationary Excitation," *Proc. of the 2nd Int. Workshop on Stochastic Methods in Structural Mechanics, Pavia, Italy*, F. Casciati, L. Faravelli, eds., pp. 271–282.

Bucher, C. G., 1986b, "Approximate Method for Random Vibration Analysis of MDOF-Systems Subjected to Nonstationary Non-White Excitation," Report No. 6/86, Institute of Engineering Mechanics, University of Innsbruck, Austria.

Der Kiureghian, A., and Igusa, T., 1985, "Stochastic Response of Secondary Systems to Short Duration Earthquakes," *Trans. 8th Int. Conf. Struct. Mech. in Reactor Technology*, J. Stalpaert, ed., Vol. M1/M2, North Holland Phys. Publ., pp. 93–98.

Gasparini, D. A., and Debchaudhury, A., 1980, "Dynamic Response to Nonstationary Non White Excitation," *Proc. ASCE*, Vol. 106, No. EM 6, pp. 1233–1248.

Hammond, J. K., 1968, "On the Response of Single and Multidegree of Freedom Systems to Non-Stationary Random Excitations," *J. Sound and Vibration*, Vol. 7, No. 3, pp. 393–416.

Lin, Y. K., 1976, *Probabilistic Theory of Structural Dynamics*, McGraw-Hill; 1967, repr. Robert E. Krieger Publ. Comp., Malabar.

Solomos, G. P., and Spanos, P.-T. D., 1984, "Oscillator Response to Nonstationary Excitation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 907–912.

Stratonovich, R. L., 1967, *Topics in the Theory of Random Noise*, Translations by R. A. Silverman, Gordon and Breach.

Yang, I-Min, and Iwan, W. D., 1972, "Calculation of Correlation Matrices for Linear Systems Subjected to Nonwhite Excitation," ASME JOURNAL OF APPLIED MECHANICS, pp. 559–562.

Yang, J. N., Lin, Y. K., and Sae-Ung, S., 1980, "Tall Building Response to Earthquake Excitations," *Proc. ASCE*, Vol. 106, No. EM4, pp. 801–817.

On the Dynamic Response of a Single-Degree-of-Freedom Structure Attached to the Interior of a Rotating Rigid Ring

A. G. Hernried
Assistant Professor,
Department of Civil Engineering.

G. B. Gustafson
Professor,
Department of Mathematics.

University of Utah,
Salt Lake City, UT 84112

The dynamic response of a single-degree-of-freedom structure attached to the interior of a rigid ring that is rotating with constant angular velocity is investigated. It is assumed that the deformations of the structure from the undeformed configuration are small, and that the structure exhibits linear elastic material behavior. Both undamped and viscously damped structures are considered. Inclusion of all of the essential dynamic features of the problem (notably Coriolis effects) results in a nonlinear differential equation governing the response of the attached structure. Numerical and analytical studies are performed on the nonlinear governing equation to determine the response and stability of the structure.

Introduction

The dynamic response of structures attached to rotating frames is a problem of significant importance, since such systems exist in rotating machinery, spin-stabilized satellites, centrifuges, as well as other physical systems. Thompson (1971) was among the first to investigate the dynamic effects on a structure subjected to rotation induced compression. In that work, it was concluded that extraneous dynamic effects (notably Coriolis effects) were not present in a linear analysis. Based on Thompson's initial investigation, Mostaghel and Tadjbakhsh (1973) developed linear differential equations governing the stability of rotating beams and plates. Several additional insights into the stability and dynamic response of rotating beams governed by linear differential equations that neglect Coriolis effects were offered by a variety of investigators (Rammerstorfer, 1974; Nachman, 1975; Lakin, 1976; Wang, 1976; Lakin and Nachman, 1978 and 1979; Peters and Hodges, 1980; Kammer and Schlack, 1984). A slightly different model of a structure subjected to rotation induced compression was proposed by Weber (1976). It was shown (Weber, 1976) that, even under the simplifying assumptions of linear elastic theory and small deformations, inclusion of Coriolis effects results in nonlinear differential equations governing the response of the rotating structure. The majority of Weber's (1976) work was an investigation of stability based on a linear differential equation resulting from dropping the nonlinear terms. However, a preliminary analysis of the original nonlinear differential equations was also presented.

In this paper, the dynamic response of a single-degree-of-freedom structure attached to a rigid ring that is rotating with constant angular velocity is studied. As in previous investigations, it is assumed that the structure exhibits linear elastic material behavior and that the displacements from the undeformed configuration are small. The investigation presented here, however, focuses on the response of the rotating structure as determined by the resulting nonlinear differential equations. Both undamped and viscously damped structures are considered.

Model

The structural model is a single-degree-of-freedom system (of height h) consisting of a mass m concentrated at the tip of a massless flexible elastic element (stiffness k). The nonzero displacements of the mass from its undeformed position, denoted $u(t)$, are small. The structure is attached to the inside of a rotating rigid ring (see Fig. 1). The ring of radius R is rotating with constant angular velocity Ω about a vertical axis passing through the point O. The height of the structure is less than the radius of the ring, i.e., $h \leq R$. It is convenient to introduce two frames of reference, the inertial (Newtonian) frame $[X-Y]$ and the rotating frame $[x-y]$ whose unit base vectors are $\mathbf{i}$ and $\mathbf{j}$.

Referring to Fig. 1, the position vector $\mathbf{r}_P$ is given as

$$\mathbf{r}_P = \mathbf{r}_A + \mathbf{r}_{P/A} \tag{1}$$

where

$$\mathbf{r}_A = -R\mathbf{j}; \quad \mathbf{r}_{P/A} = u(t)\mathbf{i} + h\mathbf{j} \tag{2}$$

Taking the first and second derivatives of $\mathbf{r}_P$ with respect to

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 24, 1987; final revision, June 28, 1987.
Paper No. 88-WA/APM-6.

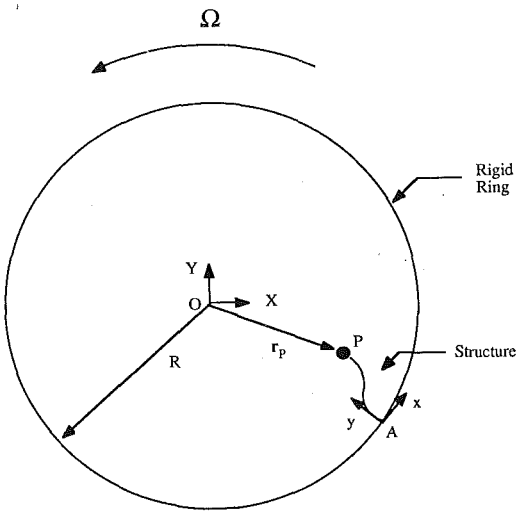


Fig. 1 Single-degree-of-freedom structure attached to rotating frame

time and noticing that $\dot{\mathbf{i}} = \Omega\mathbf{j}$ and $\dot{\mathbf{j}} = -\Omega\mathbf{i}$ (where dot indicates differentiation with respect to time) results in

$$\dot{\mathbf{r}}_P = [(R-h)\Omega + \dot{u}]\mathbf{i} + u\Omega\mathbf{j}$$
$$\ddot{\mathbf{r}}_P = (\ddot{u} - u\Omega^2)\mathbf{i} + [(R-h)\Omega^2 + 2\dot{u}\Omega]\mathbf{j} \tag{3}$$

The force on the mass m is

$$\mathbf{F}_P = F_T\mathbf{i} + F_N\mathbf{j} \tag{4}$$

where, for the undamped system and including the effects of geometric stiffness (linear approximation, Clough and Penzien, 1975, p. 168),

$$F_T = -(k - F_N/h)u \tag{5}$$

In the inertial frame,

$$\mathbf{F}_P = m\ddot{\mathbf{r}}_P \tag{6}$$

Equating $\mathbf{i}$ and $\mathbf{j}$ components of equation (6) above results in the subsequent two scalar equations

$$\ddot{u} + [\omega^2 - F_N/(mh) - \Omega^2]u = 0 \tag{7}$$
$$F_N = m[(R-h)\Omega^2 + 2\dot{u}\Omega] \tag{8}$$

where the natural frequency squared of the structure is $\omega^2 = k/m$. The normal force given by equation (8) would be transmitted through the structure and act on the rotating ring. It is interesting to note that, in contrast to the traditional problem of structural dynamics where the normal force is not a function of time, the normal force in the model presented here is time-varying. This fact may be significant for purposes of control, if the system under consideration is a spin-stabilized satellite or an orbiting space station.

Substituting equation (8) into equation (7) results in the equation

$$\ddot{u} + \left[\omega^2 - \frac{R}{h}\Omega^2 - \frac{2}{h}\Omega\dot{u}\right]u = 0 \tag{9}$$

which governs the free vibration response of the attached structure. The product of the rotational angular velocity Ω with the structural velocity $\dot{u}$, i.e., $2\Omega\dot{u}/h$ in the above equation, is the Coriolis term. Equation (9) can be rewritten by introducing the transformation

$$v(\omega t) = u(t) \tag{10}$$

and the nondimensional quantities $\xi = R/h$ and $\alpha = \Omega/\omega$. The result is

$$v'' + \left[1 - \xi\alpha^2 - \frac{2}{R}\xi\alpha v'\right]v = 0 \tag{11}$$

or more simply

$$v'' + a^2 v - bvv' = 0 \tag{12}$$

with $a^2 = 1 - \xi\alpha^2$ and $b = (2/R)\xi\alpha$. If the structure is viscously damped, with damping factor $c = 2\beta m\omega$, then equation (12) becomes

$$v'' + 2\beta v' + a^2 v - bvv' = 0 \tag{13}$$

The nonlinear term bvv' in equations (12), (13) results from the inclusion of Coriolis effects.

Analytical Studies for the Damped System

The equivalent autonomous form of equation (13) is

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}) \tag{14}$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} v \\ v' \end{bmatrix}; \quad \mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \tag{15}$$

where

$$f_1 = y_2; \quad f_2 = (by_1 - 2\beta)y_2 - a^2 y_1 \tag{16}$$

The function $\mathbf{f}(\mathbf{y})$ of equation (14) can be written

$$\mathbf{f}(\mathbf{y}) = \mathbf{A}\mathbf{y} + \mathbf{g}(\mathbf{y}) \tag{17}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -a^2 & -2\beta \end{bmatrix}; \quad \mathbf{g} = \begin{bmatrix} 0 \\ by_1y_2 \end{bmatrix} \tag{18}$$

It can be established that the zero solution of equation (14) is asymptotically stable. This results from the use of the following theorem (Theorem 4.3, Brauer and Nohel, 1969, p. 161). For the system

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g}(t,\mathbf{y}), \tag{19}$$

suppose that all the eigenvalues of $\mathbf{A}$ have negative real parts and $\mathbf{g}(t, \mathbf{y})$ and $(\partial\mathbf{g}/\partial y_j)(t,\mathbf{y})(j = 1, \ldots, n)$ are continuous in $(t, \mathbf{y})$ for $0 \le t < \infty$, $|\mathbf{y}| < k$ where $k > 0$ is a constant. Suppose also that $\mathbf{g}$ is small in the sense that

$$\lim_{|\mathbf{y}|\to 0} \frac{|\mathbf{g}(t,\mathbf{y})|}{|\mathbf{y}|} = 0 \tag{20}$$

uniformly with respect to t on $0 \le t < \infty$. Then the solution $\mathbf{y} \equiv \mathbf{0}$ of equation (19) is asymptotically stable. It is a simple matter to show that for a^2, $\beta > 0$ the requirements of this theorem are satisfied for the system of equations (14). This indicates that for "small" initial disturbances $[v(0), v'(0)]$ from the initial state, the structure eventually returns to its undeformed, rest configuration. Stability regions cannot be computed using the above theorem. Lyapunov theory, however, can be invoked to determine a stability region.

An appropriate Lyapunov function is

$$V(\mathbf{y}) = \frac{1}{2}(a^2 y_1^2 + y_2^2) \tag{21}$$

which satisfies $V^*(\mathbf{y}) = y_2^2\,(by_1 - 2\beta) \le 0$ in the region

$$S = \left\{\mathbf{y} \,\middle|\, y_1 < \frac{2\beta}{b}, -\infty < y_2 < \infty\right\}.$$

It can be shown that the zero solution of equation (14) is asympotically stable (Theorem 5.2, Brauer and Nohel, 1969, p. 195)

Let us construct a region of asymptotic stability for equation (14) using Lyapunov function (21). Define $E = \{\mathbf{y} \epsilon S \,|\, V^*(\mathbf{y}) = 0\}$ and let M be the largest positively invariant subset of E. It is easily checked that $M = \{(0,0)\}$. Let $C_\gamma = \{\mathbf{y} \,|\, V(\mathbf{y}) \le \gamma\}$ where

$$\gamma = \frac{2\beta^2 a^2}{b^2} \ge 0.$$

All solutions of equation (14) which start in C_γ approach M as $t \to \infty$ (Theorem 5.5, Brauer and Nohel, 1969, p. 214).

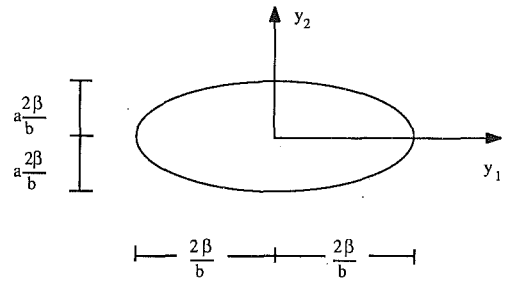


Fig. 2 An elliptical stability region predicted by Lyapunov theory

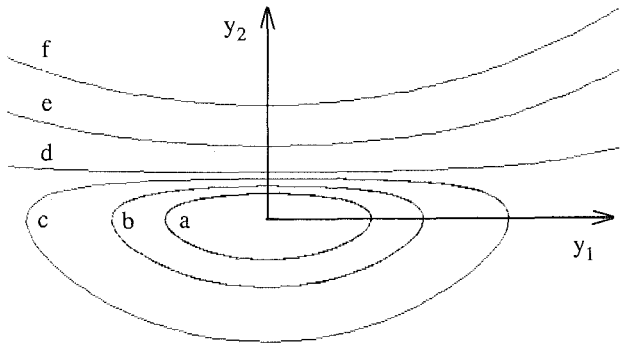


Fig. 3 Phase plane solutions for $a^2 = 0.1$, $b = 0.006$, $\beta = 0$, $\lambda = 16.7$. Initial values: $a = (0, -15)$, $b = (0, -25)$, $c = (0,15)$, $d = (-95,20)$, $e = (-95,40)$, $f = (-95,60)$.

Therefore, C_γ is a region of asymptotic stability for system (14). The above result guarantees that the structure returns to its original undeformed configuration for all initial disturbances $y_1(0)$, $y_2(0)$ within the elliptical region depicted in Fig. 2. Figure 2 indicates that for a given value of damping β, Coriolis effects represented by the parameter b control the extent of the stability region. As b becomes large, this stability region shrinks.

Analytical Studies for the Undamped System

Let $y_1 = v$, $y_2 = v'$; then from equation (12)

$$dy_2/dy_1 = -(a^2 - by_2)y_1/y_2. \tag{22}$$

Since equation (22) has variables separable, a quadrature results in

$$y_2 + (a^2/b)\ln|a^2 - by_2| = (by_1^2/2) + C \tag{23}$$

where C is a constant to be evaluated. For the initial data

$$v(0) = 0, \quad v'(0) = \kappa \neq \lambda = a^2/b \tag{24}$$

from equation (23) it results that trajectories in phase space of equation (12) satisfy the energy relation

$$v' + \frac{a^2}{b}\ln\frac{|a^2 - bv'|}{|a^2 - b\kappa|} = \frac{b}{2}v^2 + \kappa \tag{25}$$

As is established in the appendix, stable, periodic response of the undamped structure is guaranteed for initial disturbances $[v(0), v'(0)]$ from the rest (undeformed) configuration that fall within the region $\{(v,v') \mid -\infty < v < \infty, -\infty < v' < \lambda\}$. The period of oscillation $2T$ is such that

$$0 < T \leq \frac{\pi}{2(a^2 - b\kappa)^{1/2}} + \frac{\pi}{2a}. \tag{26}$$

Since periodic solutions satisfy $\kappa < \lambda$, the amplitude of a periodic solution of equation (12), with the initial data of equation (24), results from equation (25) and is given as

$$v_{\max} = \left\{\frac{2}{b}\left(\frac{a^2}{b}\ln\left[\frac{a^2}{a^2 - b\kappa}\right] - \kappa\right)\right\}^{1/2} \tag{27}$$

The system is unstable if the initial disturbances fall within the region of instability $\{(v,v') \mid -\infty < v < \infty, v' > \lambda\}$. To establish this result, by symmetry it suffices to show that $v(0) = 0$, $v'(0) = \kappa > \lambda$ implies $\lim_{t\to\infty} v(t) = \infty$. For this initial data, $v''(t) \geq 0$ which implies $v(t) \geq \kappa t$ for $t \geq 0$.

A separatrix, which "separates" the regions of stability and instability, is given by $\{(v,v') \mid -\infty < v < \infty, v' = \lambda\}$. Solutions starting on the separatrix follow, in phase space, the line given above. This results from equation (12), noting that $v' = \lambda$ implies $v = \lambda t$.

Numerical Studies

Several numerical studies were performed to obtain further insight into the response of the structure. It is necessary to determine reasonable ranges for the parameters of the system; a^2, b, β. For common structures, $0 \leq \beta \leq 0.1$. The analytical studies given above indicate that in order to insure stability of the system, a^2 must be greater than zero. This implies that $0 < a^2 \leq 1$ since $\xi \geq 1$, $\alpha \geq 0$. A necessary condition for stability of the system, $0 \leq \Omega < \omega(h/R)^{1/2}$, immediately results. Assume that $1 \leq \xi \leq \xi_m$ where ξ_m is the maximum value of ξ, and that R varies from one to maximum value of R_m (i.e., $1 \leq R \leq R_m$). Then $b_{\min} \leq b \leq b_{\max}$ where

$$b_{\min} = \frac{2}{R_m}\frac{1}{\sqrt{\xi_m}}(1 - a^2)^{1/2}, \quad b_{\max} = 2\xi_m(1 - a^2)^{1/2}. \tag{28}$$

Values of $\xi_m = 10$ and $R_m = 100$ were considered reasonable. Notice that if $a^2 = 1.0$ then $b \equiv 0$ and the system is linear. This is as one would expect since $a^2 = 1.0$ implies $\Omega = 0$. In this instance the ring is not rotating, and the nonlinearity resulting from rotation is necessarily absent.

Phase plane solutions, based on the nonlinear differential equation solution procedure discussed by Hindsmarsch (1980), were generated for a variety of initial conditions and system parameters within the ranges discussed above. As is illustrated by Fig. 3, the results of the analytical studies for the undamped system given in the preceding section are substantiated by numerical experiments. The influence of damping on the response of this system is shown in Fig. 4. This figure indicates that there exists a region of stable attraction that is significantly different from that predicted by Lyapunov theory. A value of b near the maximum reasonable value for the particular choice of a^2 was then chosen, i.e., $a^2 = 0.1$, $b = 5.0$. For the undamped system, the separatrix in this case was very close to zero. Results similar to those discussed above were obtained for this case as well; however, the computational effort required to solve this system was significant, indicating that the relatively large value of b introduces computational difficulties.

In contrast to the above experiments where the value of a^2 was near the minimum allowable value of zero, numerical experiments were also performed for a^2 near the maximum allowable value of one, i.e., $a^2 = 0.9$. Figure 5 indicates that for a value of b near the minimum reasonable value for the particular choice of a^2, the system behaves in essentially a linear manner over a wide range of initial values. This is not the case, however, as b approaches the maximum reasonable value for the particular choice of a^2, as is evidenced by Fig. 6.

Closure

The dynamic response of a single-degree-of-freedom structure attached to the interior of a rigid ring that is rotating with constant angular velocity is presented. Small deformation theory and linear elastic material behavior are assumed. The structure may also be viscously damped.

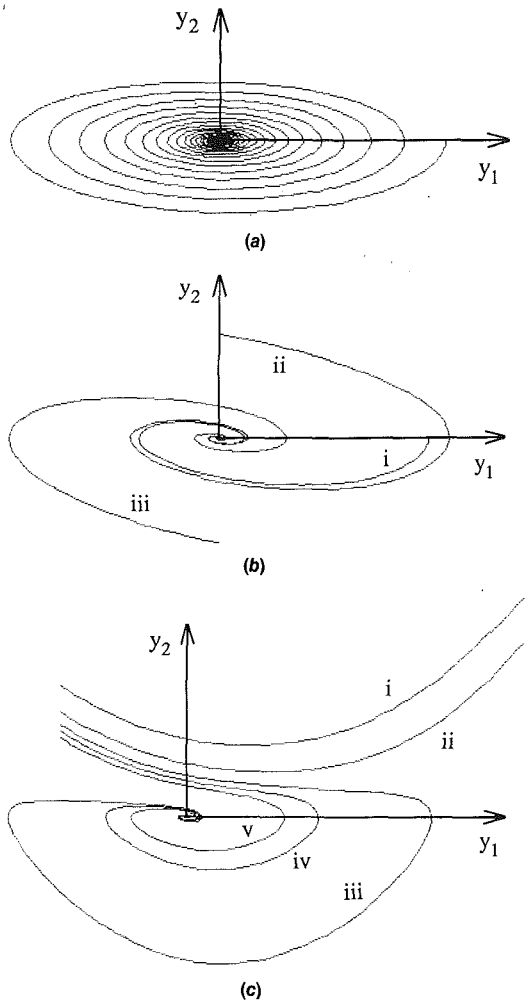


Fig. 4 Influence of damping on the response for $a^2 = 0.1$, $b = 0.006$. (a) $\beta = 0.01$. Initial Value: (10,0); (b) $\beta = 0.1$. Initial Values: $i = (10,0)$, $ii = (0,5)$, $iii = (0.-5)$; (c) $\beta = 0.1$. Initial Values: $i = (-95,100)$, $ii = (-95,80)$, $iii = (-95,70)$, $iv = (-95,65)$, $v = (-95,60)$.

If Coriolis effects are ignored *a priori*, then a linear differential equation results. The stability of the system then depends solely on the physical parameters, i.e., stable response is guaranteed for $\Omega < \omega(h/R)^{1/2}$ whether the system is damped or undamped, and not on the initial conditions. For stable systems, the dynamic response of the structure is then similar to that of a simple harmonic oscillator.

However, Coriolis effects must be included if one considers all of the essential dynamic characteristics of the problem. This results in a nonlinear differential equation governing the response of the structure. Although the requirement $\Omega < \omega(h/R)^{1/2}$ is a necessary condition for stability of the nonlinear system, it is not a sufficient condition. Stability of the nonlinear system depends on the initial conditions as well. For the undamped system, it is shown that stable response is guaranteed for an initial structural velocity less than the separatrix constant. For stable systems, the dynamic response, although in most instances markedly different from the simple harmonic oscillator, is periodic. An expression for the amplitude and a bound on the period of oscillation of the nonlinear system in terms of the system parameters is given. When the system is damped, Lyapunov theory can be employed to generate a region of asymptotic stability in terms of the parameters of the system which, in the phase plane, is elliptical in nature. Also presented are the results of several numerical experiments on the governing nonlinear differential equation for ranges of parameter values of interest.

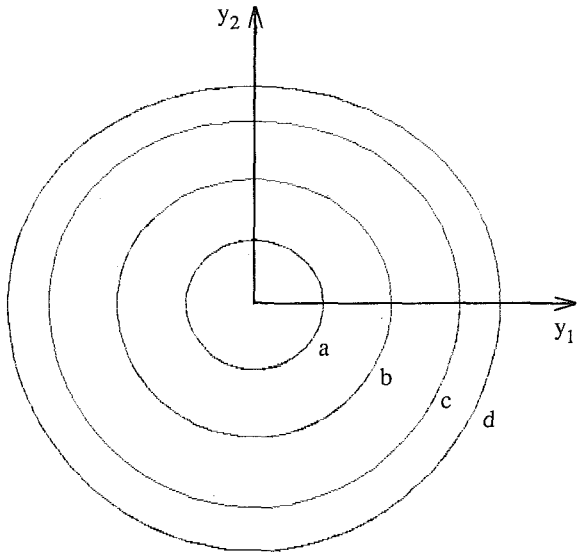


Fig. 5 Phase plane solutions for $a^2 = 0.9$, $b = 0.002$, $\beta = 0$, $\lambda = 450$. Initial Values: $a = (25,0)$, $b = (50,0)$, $c = (75,0)$, $d = (90,0)$.

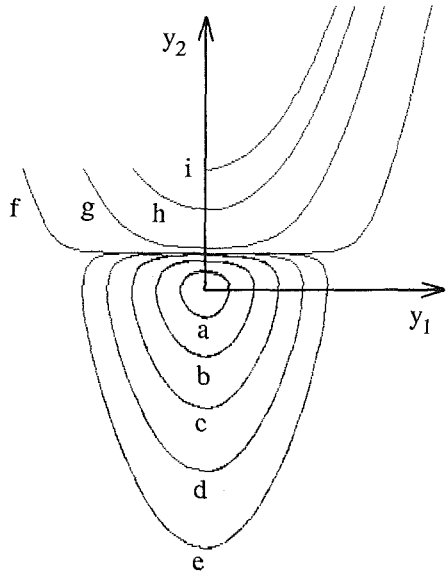


Fig. 6 Phase plane solutions for $a^2 = 0.9$, $b = 0.6$, $\beta = 0$, $\lambda = 1.5$, Initial values: $a = (1,0)$, $b = (2,0)$, $c = (3,0)$, $d = (4,0)$, $e = (5,0)$, $f = (-7.5,5)$, $g = (-5,5)$, $h = (-3,5)$, $i = (0,5)$.

In conclusion, Coriolis effects are extremely important since, for a large variety of physical systems that would occur in practice, they significantly alter the stability and dynamic response of the system from that predicted by linear theory. This is the case for both damped and undamped systems. Damping, however, serves to increase the stability of the particular system in question.

Acknowledgment

The first author would like to acknowledge the financial support of his research by the U.S. National Science Foundation.

References

Brauer, F., and Nohel, J. A., 1969, *The Qualitative Theory of Ordinary Differential Equations*, W. A. Benjamin.

Clough, R. W., and Penzien, J., 1975, *Dynamics of Structures*, McGraw-Hill, New York.

Hindmarsch, A. C., 1980, "LSODE and LSODI, Two New Initial Value Or-

dinary Differential Equation Solvers," *ACI-SIGNUM Newsletter*, Vol. 15, pp. 10-11.

Kammer, D. C., and Schlack, A. L., 1984, "Liapunov's Direct Method Applied to the Buckling of Rotating Beams," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 941-942.

Lakin, W. D., 1976, "Vibrations of a Rotating Flexible Rod Clamped Off the Axis of Rotation," *J. Eng. Math.*, Vol. 10, pp. 313-321.

Lakin, W. D., and Nachman, A., 1978, "Unstable Vibrations and Buckling of Rotating Flexible Rods," *Quart. Appl. Math.*, Vol. 35, pp. 479-493.

Lakin, W. D., and Nachman, A., 1979, "Vibration and Buckling of Rotating Flexible Rods at Transitional Parameter Values," *J. Eng. Math.*, Vol. 13, pp. 339-346.

Mostaghel, N., and Tadjbakhsh, I., 1973, "Buckling of Rotating Rods and Plates," *Int. J. Mech. Sci.*, Vol. 15, pp. 429-434.

Nachman, A., 1975, "The Buckling of Rotating Rods," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 222-224.

Peters, D. A., and Hodges, D. H., 1980, "In-Plane Vibration and Buckling of a Rotating Beam Clamped Off the Axis of Rotation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 398-402.

Rammerstorfer, F. G., 1974, "Comment on Buckling of Rotating Rods and Plates," *Int. J. Mech. Sci.*, Vol. 16, pp. 515-517.

Thompson, J. M. T., 1971, "On the Simulation of a Gravitational Field by a Centrifugal Field," *Int. J. Mech. Sci.*, Vol. 13, pp. 979-986.

Waltman, P., 1986, *A Second Course in Elementary Differential Equations*, Academic Press, New York.

Wang, J. T. S., 1976, "On the Buckling of Rotating Rods," *Int. J. Mech. Sci.*, Vol. 18, pp. 407-411.

Weber, H. I., 1976, "A Note on the Stability of a Rod Subjected to Compression by Centrifugal Force," *J. Sound and Vibr.*, Vol. 46, pp. 105-111.

APPENDIX

Stability and Periodicity of $v'' + (a^2 - bv')v = 0$

Lemma 1. Let $v(t)$ be a solution of

$$v'' + (a^2 - bv')v = 0, \quad 0 \le t \le T,$$
$$v(0) = v(T) = 0, \quad T > 0. \tag{A1}$$

Define $p(t) = (v(t), \ v'(t))$. Let $q(t) = (-v(2T-t), v'(2T-t))$, $T \le t \le 2T$. Then p and q are phase solutions of

$$y_1' = y_2$$
$$y_2' = -(a^2 - by_2)y_1 \tag{A2}$$

and

$$r(t) = \begin{cases} p(t) & 0 \le t \le T \\ q(t) & T \le t \le 2T \end{cases}$$

is a periodic orbit of equations (A2) of period $2T$. The fundamental period is $2T$ provided $v(t) > 0$ on $0 < t < T$.

Proof of Lemma 1. Clearly $p(t)$ is a phase solution. To prove the same is true about $q(t)$, let $y_1(t) = -v(2T-t)$, $y_2(t) = v'(2T-t)$. A short computation shows that (y_1, y_2) satisfies equations (A2) for $T \le t \le 2T$. Since $p(T) = q(T)$, $r(t)$ is a solution of equations (A2) on $0 \le t \le 2T$. Furthermore, $r(t)$ is closed due to $p(0) = q(2T)$, therefore $r(t)$ is a periodic orbit of period $2T$. If $v(t) > 0$ on $0 < t < T$, then $v'(0) > 0$ and $v'(T) < 0$. Therefore, $p(t)$ flows clockwise through quadrants *I* and *II*, while $q(t)$ flows clockwise through quadrants *III* and *IV*. Consequently, the exact period of $r(t)$ is $2T$.

Lemma 2. The problem

$$v'' + (a^2 - bv')v = 0,$$
$$v(0) = v(T) = 0, \tag{A3}$$
$$v(t) > 0 \ \text{on} \ 0 < t < T$$
$$a^2 - b\kappa > 0, \ v'(0) = \kappa > 0,$$

has a solution $v(t)$ for some

$$T > 0, \ T \le \frac{\pi}{2(a^2 - b\kappa)^{1/2}} + \frac{\pi}{2a}.$$

Proof of Lemma 2. Let $v(t)$ be the solution of problem (A3) with $v(0) = 0$, $v'(0) = \kappa > 0$. It is claimed that $v'(t) = 0$ for some first $t = t_1 > 0$ and $t_1 \le t_2 \equiv \pi/[2(a^2 - b\kappa)^{1/2}]$.

Suppose not. Then $v'(t) > 0$ and $v''(t) \le 0$ on $[0, t_2]$. Let $\bar{\omega}^2 = a^2 - b\kappa$ and $z(t) = \kappa \sin(\bar{\omega}t)/\bar{\omega}$, a solution of $z'' + \bar{\omega}^2 z = 0$, $z(0) = 0$, $z'(0) = \kappa$.

Introduce polar coordinates θ_1, r_1, and θ_2, r_2 such that $z(t) = r_1(t)\sin\theta_1(t)$, $z'(t) = r_1(t)\cos\theta_1(t)$ and $v(t) = r_2(t)\sin\theta_2(t)$, $v'(t) = r_2(t)\cos\theta_2(t)$. The polar equations of interest are

$$\theta_1' = \cos^2\theta_1 + Q_1(t)\sin^2\theta_1$$
$$\theta_1(0) = 0 \tag{A4}$$
$$\theta_2' = \cos^2\theta_2 + Q_2(t)\sin^2\theta_2$$
$$\theta_2(0) = 0 \tag{A5}$$

where $Q_1(t) = a^2 - b\kappa \equiv \bar{\omega}^2$ and $Q_2(t) = a^2 - bv'(t)$.

Since v' decreases on $[0, t_2]$, $v'(t) \le \kappa$ and $Q_2(t) \ge Q_1(t)$. By standard comparison theory (Waltman, 1986, p. 209) applied to equation (A4) and (A5), it follows that $\theta_1(t) \le \theta_2(t)$ on $[0, t_2]$. However, $r_1(t)\sin\theta_1(t) = z(t) = \kappa\sin(\bar{\omega}t)/\bar{\omega}$ gives $\theta_1(t_2) = \pi/2$, therefore $\theta_2(t_2) \ge \pi/2$. But $v'(t) = r_2(t)\cos\theta_2(t) > 0$ on $[0, t_2]$, a contradiction. This completes the proof of the claim that $v'(t_1) = 0$, $t_1 \le t_2$.

The last claim is that $v(t) = 0$ for some first $t = t_3 \ge t_1$ and furthermore $t_3 \le t_4 \equiv t_2 + \pi/(2a)$. Suppose not. Then $v(t) > 0$ on $[0, t_4]$. Define $z(t) = v(t_1)\sin[a(t - t_1) + \pi/2]$, $Q_1(t) = a^2$, $Q_2(t) = a^2 - bv'(t)$, so that $z'' + Q_1 z = 0$, $v'' + Q_2 v = 0$.

Since $v'(t_1) = 0$ and $v''(t_1) < 0$, it follows that $v'(t) < 0$ and $v''(t) \le 0$ on $t_1 < t \le t_4$. Therefore, $Q_2(t) \ge Q_1(t)$ on $[t_1, t_4]$. Finally, $\theta_1(t_1) = \pi/2$ and $\theta_2(t_1) = \pi/2$ is a consistent set of initial values, so that $\theta_1(t) \le \theta_2(t)$ by the comparison theorem mentioned previously. However, $\theta_1(t_1 + \pi/(2a)) = \pi$, hence $\theta_2(t) = \pi$ for some $t_1 \le t \le t_4$, a contradiction to $v(t) = r_2(t)\sin\theta_2(t) > 0$ on $[t_1, t_4]$.

By the two claims, the trajectory of $v(t)$ in phase winds clockwise from $(0, \kappa)$ in (v, v')-space to a point $(0, -\delta)$ in time t, $0 < t \le t_4$. This completes the proof.

Theorem 1. For $\pi/a < T < \infty$, there exists a periodic solution $v(t)$ of

$$v'' + (a^2 - bv')v = 0$$

with period $2T$ satisfying $v(0) = 0$, $v'(0) < a^2/b$. Furthermore, each starting condition $v(0) = 0$, $0 < v'(0) < a^2/b$ corresponds to a periodic solution.

Proof of Theorem 1. Let $T > \pi/a$. Solutions of $v'' + (a^2 - bv')v = 0$ with phase data $(0, \kappa)$, $\kappa < a^2/b$, satisfy $v(t_\kappa) = 0$, $v'(t_\kappa) < 0$ for some first $t_\kappa > 0$. It is a consequence of basic existence theorems that t_κ is a continuous function of κ.

As $\kappa \to a^2/b$, the solution must follow the separatrix $t \to a^2 t/b$, therefore $t_\kappa \to \infty$ as $\kappa \to a^2/b$. Given $T > \pi/a$, there is a value of κ near a^2/b such that $t_\kappa > T$.

Choose γ so that $T > \pi/\gamma > \pi/a$. By stability at $v = v' = 0$, $a^2 - bv'(t) > \gamma^2$ on $0 \le t \le t_\kappa$ for all small $\kappa > 0$. By Sturm comparison, $v(t)$ must vanish between zeros of $\sin(\gamma t)$, therefore $t_\kappa < \pi/\gamma$. So $t_\kappa < T$ for all small $\kappa > 0$.

Since the continuous mapping $\kappa \to t_\kappa$ reaches values greater and smaller than T, $t_\kappa = T$ for some κ. This completes the proof.

D. J. Inman
Professor,
Department of Mechanical and Aerospace Engineering.
Mem. ASME

C. L. Olsen
Associate Professor,
Department of Mathematics.

State University of New York at Buffalo,
Buffalo, NY 14260

Dynamics of Symmetrizable Nonconservative Systems

This work examines a subclass of nonconservative and nonself-adjoint linear distributed parameter systems modeled by partial differential equations subject to various boundary conditions. It has been previously illustrated that certain nonself-adjoint operators arising in mechanics can be shown to be self-adjoint with respect to a particular self-adjoint operator. The work here extends and formalizes this approach to include systems with viscous damping. The extension presented here follows the lumped parameter case presented earlier and places emphasis on the existence of eigenfunctions for use in performing modal analysis.

Introduction

Recently Inman (1983) illustrated that certain asymmetric lumped parameter systems have dynamical properties very similar to the properties of related symmetric systems. In fact, for certain systems there exists a similarity transformation that transforms the asymmetric system into an equivalent symmetric one (i.e., one having the same eigenvalues). The work here intends to extend this idea to nonself-adjoint distributed parameter systems described by partial differential equations. Leipholz (1974a, 1974b) has shown that certain distributed parameter systems such as Pflüger's column have a stiffness operator which is the sum of a self-adjoint operator and a nonself-adjoint operator. With certain boundary conditions this operator is shown to be self-adjoint with respect to an assigned self-adjoint operator. In this sense the system is self-adjoint in a generalized way. The physical interpretation also predicts the existence of this generalized self-adjointness inasmuch as the nonself-adjoint systems that satisfy Leipholz's condition exhibit instability via divergence rather than by flutter, just as one would expect from a self-adjoint system.

Walker (1973, 1977) has shown that Liapunov functions for nonself-adjoint systems can sometimes be constructed by writing down energy functionals with respect to an assigned symmetric operator. In this way, he was able to extend the current Liapunov stability analysis for self-adjoint systems to a larger class of systems. Walker also considered damping mechanisms, but was not concerned with symmetry of the revised damping operator. While Walker's work focuses on asymptotic stability, the existence of symmetrizing Liapunov functionals provides an alternative motivation for the following work.

The work presented here extends the above-mentioned works to include distributed parameter systems with velocity dependent forces. In particular, the question of the existence of eigenfunction expansions for nonself-adjoint systems is addressed. Unlike previous work in nonself-adjoint problems, the work presented here focuses on the existence of eigenfunctions and the existence of oscillation conditions for such systems. This has significance because of the recent interest in modal control methods in the large flexible space structure research area and because of the popularity of modal testing and design in structural dynamics research.

Background

The class of systems considered here are those systems successfully described by a partial differential equation or set of equations, of the form

$$u_{tt}(x,t) + L_1 u_t(x,t) + L_2 u(x,t) = f(x,t) \text{ on } \Omega \tag{1}$$

plus boundary conditions of the form $Bu = 0$ along $\partial\Omega$ where:

- x denotes the position in Ω, a bounded open region in R^n with boundary $\partial\Omega$;
- $u(x, t)$ denotes the deflection in Ω;
- the subscript t denotes partial differentiation with respect to the time, (usually taken as positive);
- $f(x, t)$ denotes external forces;
- L_1 and L_2 denote linear spatial (time invariant) differential operators of order n_1 and n_2, respectively;
- B denotes a linear spatial operator reflecting the boundary conditions.

Let $\mathcal{L}^2(\Omega)$ denote the Hilbert space of all square integrable functions in the Lebesgue sense defined by the inner product and norm denoted by:

$$\langle u,v\rangle = \int_\Omega u(x)v(x)dx \tag{2}$$

$$\|u(x)\| = \langle u,u\rangle^{1/2} \tag{3}$$

respectively. The operators L_1 and L_2 are defined on a domain consisting of a dense subset of the set of functions $u(\cdot, t)$ in $\mathcal{L}^2(\Omega)$ such that all partial derivatives with respect to x of order

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, June 30, 1986; final revision June 8, 1987.

Paper No. 88-WA/APM-5.

up to and including k are in $\mathfrak{L}^2(\Omega)$, where $k = \max(2n_1, n_2)$ and such that $Bu = 0$ along $\partial\Omega$ for all t. The above formulation can be used as a generic description of most linear vibrating systems, including those with odd boundary conditions and physical configurations.

Leipholz (1974a, 1974b) considered the special case of equation (1) with $L_1 = 0$, i.e., the case with no velocity dependent forces. In the problem considered by Leipholz, the operator L_2 is nonself-adjoint but is such that there exists a weighted inner product defined by:

$$\langle u,v\rangle_T = \int_\Omega uTv dx \tag{4}$$

where T is an operator (unbounded) and $\langle u, L_2 v\rangle_T = \langle L_2 u, v\rangle_T$.

Walker considered equation (1) with $L_1 = 0$ and L_2 such that there exists a bounded operator G with GL_2 symmetric. Under this condition he proves that the stability of equation (1) can be determined from a Liapunov function. In this work we take a slightly different approach to examining these results by considering the factors of the operator L_2, and mimicking the procedure used in the matrix case (Inman, 1983). In addition the nonconservative cases with $L_1 \neq 0$ are also considered with the emphasis on the existence of eigenfunction expansions and oscillation conditions.

To review briefly the results in Inman (1983), an asymmetric real matrix can always be factored into the product of two symmetric matrices. If one of the factors is positive definite, then the dynamic system associated with the matrix behaves as if the system were symmetric and is called symmetrizable. This statement is based on the theory of symmetric matrix factors (Taussky, 1968). A similar theory is presented here for unbounded operators as groundwork for the vibration theory of symmetrizable nonself-adjoint distributed parameter systems. The mathematics here in our theory for unbounded operators is also motivated by similar ideas developed for bounded operators by Lax (1954) and Dieudonne' (1961).

Results

Following the results obtained for multidegree of freedom systems given by Inman (1983), a theory for the vibration of a subclass of nonself-adjoint distributed parameter systems can be developed. The partial differential equations and boundary conditions used to describe such systems require the use of unbounded operators, making the mathematics more difficult and restrictive. In this first section, operator properties will be considered and then used to transform the original nonself-adjoint problem into an "equivalent" self-adjoint problem. In subsequent sections results regarding the vibrations of such systems will be presented and the relation to stability theory will be discussed.

Operator Products

The theory of unbounded operators is far more complicated than the theory of matrices. However, several useful results for unbounded operators can be stated based on theories developed for matrices and bounded operators.

Definition 1. Let $\mathfrak{H}$ be an infinite-dimensional separable Hilbert space. A set of vectors in $\mathfrak{H}$ whose linear span is dense in $\mathfrak{H}$ will be called a *total set* in $\mathfrak{H}$.

Theorem 1. Let L be a nonsingular unbounded operator defined on a domain $D(L)$ dense in a Hilbert space $\mathfrak{H}$. Assume there exists a positive definite self-adjoint operator T such that the product $P = TL$ is symmetric on a dense domain with dense range. Assume furthermore that the operator $T^{1/2}L^{-1}T^{-1/2}$ is compact.

Then there is a total set of eigenfunctions for L, with corresponding real eigenvalues. Furthermore, these eigenfunctions are a (nonorthogonal) basis for the dense subspace, $D(L)\cap D(T^{1/2})$ of $\mathfrak{H}$.

Proof. Recall first that a positive definite operator T has a unique positive definite square root, denoted $T^{1/2}$, with $T^{1/2}T^{1/2} = T$ (see Kato, 1966). Note that since $T^{1/2}L^{-1}T^{-1/2}$ is assumed to be compact, then $T^{-1/2}$, being everywhere defined and self-adjoint, must be bounded. Thus the operator $T^{-1} = T^{-1/2}T^{-1/2}$ must be bounded. Furthermore,

$$T^{-1/2}(T^{1/2}L^{-1}T^{-1/2})T^{-1/2} = L^{-1}T^{-1} \tag{5}$$

is a product of bounded operators with a compact operator, so $L^{-1}T^{-1}$ is a compact operator. Thus, the operator P has a compact inverse, $P^{-1} = (TL)^{-1} = L^{-1}T^{-1}$, and P is symmetric on a dense domain with dense range, so that P^{-1} is also symmetric on a dense domain. Thus, the operator P^{-1} is self-adjoint, so:

$$P \text{ is self-adjoint.} \tag{6}$$

Observe that for u in $D(T^{1/2})$, then

$$\begin{aligned} T^{1/2}L^{-1}T^{-1/2}u &= T^{1/2}L^{-1}T^{-1/2}(T^{-1/2}T^{1/2})u \\ &= T^{1/2}L^{-1}T^{-1}T^{1/2}u \\ &= T^{1/2}(TL)^{-1}T^{1/2}u \\ &= T^{1/2}P^{-1}T^{1/2}u. \end{aligned} \tag{7}$$

Thus the operator $T^{1/2}P^{-1}T^{1/2}$ is defined and bounded on the dense domain $D(T^{1/2})$ and its closure must be the compact operator $T^{1/2}L^{-1}T^{-1/2}$. But also, for u in $D(T^{1/2})$,

$$\begin{aligned} (T^{1/2}P^{-1}T^{1/2}u,u) &= (P^{-1}T^{1/2}u,T^{1/2}u) \\ &= (T^{1/2}u,P^{-1}T^{1/2}u) \\ &= (u,T^{1/2}P^{-1}T^{1/2}u), \end{aligned} \tag{8}$$

so that $T^{1/2}P^{-1}T^{1/2}$ is symmetric.

This means that the closure of $T^{1/2}P^{-1}T^{1/2}$ must also be symmetric; so we can conclude that $T^{1/2}L^{-1}T^{-1/2}$ is compact and symmetric. Thus

$$T^{1/2}L^{-1}T^{-1/2} \quad \text{and} \quad T^{1/2}LT^{-1/2} \tag{9}$$

are both self-adjoint (the latter being the inverse).

To obtain the eigenfunctions for the operator L, start with an orthogonal basis $\{w_n\}$ for $\mathfrak{H}$, of eigenfunctions for the compact self-adjoint operator $T^{1/2}L^{-1}T^{-1/2}$; these have corresponding real eigenvalues $\{\lambda_n^{-1}\}$. Note here that the self-adjoint operator P will be positive definite if and only if $T^{1/2}P^{-1}T^{1/2}$ is positive definite, and this is true if and only if the self-adjoint closure $T^{1/2}L^{-1}T^{-1/2}$, is positive definite. In particular, P is positive definite if and only if all the $\{\lambda_n\}$ are positive. In any case we have

$$T^{1/2}L^{-1}T^{-1/2}w_n = \lambda_n^{-1}w_n, \tag{10}$$

so applying $T^{-1/2}$ to this gives

$$L^{-1}(T^{-1/2}w_n) = \lambda_n^{-1}(T^{-1/2}w_n). \tag{11}$$

Thus, the sequence $\{u_n\} = \{T^{-1/2}w_n\}$ consists of eigenfunctions for L^{-1} with corresponding real eigenvalues $\{\lambda_n^{-1}\}$. Then

$$L^{-1}u_n = \lambda_n^{-1}u_n, \quad \text{so} \quad Lu_n = \lambda_n u_n. \tag{12}$$

To see that the span of the eigenfunctions $\{u_n\}$ is dense in $\mathfrak{H}$, we show that they are a basis for the dense subspace $D(T^{1/2})$. For $u\in D(T^{1/2})$, then $u = T^{-1/2}w$ for some w in $\mathfrak{H}$, with expansion $w = \Sigma a_n w_n$ in terms of the basis $\{w_n\}$. Then

$$u = T^{-1/2}\left(\sum_{n=1}^{\infty} a_n w_n\right)$$

$$= T^{-1/2}\left(\lim_{k\to\infty} \sum_{n=1}^{k} a_n w_n\right)$$

$$= \lim_{k\to\infty} \left(\sum_{n=1}^{k} a_n T^{-1/2} w_n\right)$$

$$= \lim_{k\to\infty} \left(\sum_{n=1}^{k} a_n u_n\right) \tag{13}$$

since $T^{-1/2}$ is bounded. Thus the linear span of the $\{u_n\}$ is dense in $D(T^{1/2})$, which in turn is dense in $\mathcal{H}$. But also, the linear span of $\{u_n\}$ is contained in the intersection of $D(T^{1/2})$ with $D(L)$, so this, too, is dense in $\mathcal{H}$. □

The next section gives an application of this result to get information about the solution of equation (1). We note before proceeding, that having assumed $T^{1/2}L^{-1}T^{-1/2}$ is compact, we could just as well have assumed above that P was self-adjoint with a compact inverse, since this follows. Similarly, we might as well have assumed that the inverse of T is bounded.

Symmetrizable Systems

The following definition is modeled after a corresponding notion of bounded symmetrizable operator in Lax (1954).

Definition 2. Let L be an unbounded operator with domain $D(L)$ dense in a Hilbert space $\mathcal{H}$. If there exists a positive definite self-adjoint operator T with bounded inverse such that $P = TL$ is symmetric on a dense domain with dense range, then we will say that L is *symmetrizable.*

A system given by equation (1) will be called a *symmetrizable system* if there is a positive definite self-adjoint operator T with bounded inverse, and an orthonormal basis $\{w_n\}$ such that $\{u_n\} = \{T^{-1/2}w_n\}$ are common eigenfunctions for L_1 and L_2 with real eigenvalues. We will see that this allows for a theoretical modal analysis of the system.

In these terms, Theorem 1 shows that if L is a nonsingular operator symmetrizable by T where $T^{1/2}L^{-1}T^{-1/2}$ is compact, then one obtains a total set $\{u_n\} = \{T^{-1/2}w_n\}$ of eigenfunctions for L with real eigenvalues. In what follows we will use this to obtain a symmetrizable system given by equation (1). In particular, we have immediately:

Corollary 1. In equation (1) assume

$$L_1 = \alpha I + \beta L_2 \tag{14}$$

α and β real constants. Assume also that L_2 is a nonsingular unbounded operator symmetrizable by T, where $T^{1/2}L_2^{-1}T^{-1/2}$ is compact. Then equation (1) is a symmetrizable system.

Proof. This follows at once from Theorem 1 by the preceding remarks, since the eigenfunctions $\{u_n\}$ for L_2 will also be eigenfunctions for L_1, with real eigenvalues. □

This Corollary includes the conservative case where $L_1 = 0$. Systems without damping where $L_2 = T^{-1}P$ is symmetrizable are called conservative systems of the second kind by Leipholtz (1974a).

For systems with dissipative viscous damping and gyroscopic forces, the operator L_1 in equation (1) is not zero, and is also potentially nonself-adjoint. The Corollary addresses the case of proportional damping, which is illustrated by the example given in the final section of this paper. A result for a more general class of nonself-adjoint systems with velocity dependent forces is obtained in Theorem 2 below. We relax the hypotheses that L_1 is proportional to L_2, and assume, more generally, that L_1 and L_2 commute. We are motivated by the result for the self-adjoint case stated by Caughey and O'Kelly (1965), i.e., that L_1 and L_2 commute on an appropriate domain D if and only if they share the same set of eigenfunctions and share a compatible set of boundary conditions (in fact, although the authors do not say this explicitly, some additional hypotheses must be assumed for their result: one sufficient hypotheses would be that L_1 and L_2 have compact resolvents).

Later on (in Theorem 3) we will produce a theoretical modal analysis of equation (1) for the symmetrizable systems of Theorem 2 and of Corollary 1 and Corollary 2 (below).

Theorem 2. In equation (1), assume that L_1 and L_2 are nonsingular unbounded operators, both symmetrizable by T, and assume that both $T^{1/2}L_1^{-1}T^{-1/2}$ and $T^{1/2}L_2^{-1}T^{-1/2}$ are compact. Let $\{u_n'\} = \{T^{-1/2}w_n'\}$ denote the total set of eigenfunctions for L_2 obtained in Theorem 1. Assume that $L_1L_2 = L_2L_1$ on a dense subspace D_0 where the range of $L_1L_2D_0$ contains the $\{u_n'\}$. Then equation (1) is a symmetrizable system.

Before proceding with the proof, we remark that we could just as well have taken $\{u_n'\}$ to be the eigenfunctions for L_1. Note also that the hypothesis required for D_0 is satisfied in the important case where L_1 is an arbitrary polynominal in L_2, with no constant term.

Proof. As in equation (9), the compact operators $T^{1/2}L_2^{-1}T^{-1/2}$ and $T^{1/2}L_1^{-1}T^{-1/2}$ are both self-adjoint. We now show that they commute on the orthonormal basis $\{w_n\}$. Observe first that there is a v_n in D_0 with

$$u_n' = L_1L_2v_n = L_2L_1v_n. \tag{15}$$

Thus premultiplying by $L_2^{-1}L_1^{-1}$ and $L_1^{-1}L_2^{-1}$ gives

$$L_2^{-1}L_1^{-1}u_n' = v_n = L_1^{-1}L_2^{-1}u_n'. \tag{16}$$

Since $u_n' = T^{-1/2}w_n'$, then

$$L_2^{-1}L_1^{-1}T^{-1/2}w_n' = L_1^{-1}L_2^{-1}T^{-1/2}w_n'. \tag{17}$$

Since $T^{1/2}L_j^{-1}T^{-1/2}$ is compact, then $L_j^{-1}T^{-1/2}w_n'$ is in $D(T^{1/2})$, so

$$L_2^{-1}(T^{-1/2}T^{1/2})L_1^{-1}T^{-1/2}w_n' = L_1^{-1}(T^{-1/2}T^{1/2})L_2^{-1}T^{-1/2}w_n'. \tag{18}$$

Finally for the same reason, we can premultiply by $T^{1/2}$ to get

$$(T^{1/2}L_2^{-1}T^{-1/2})(T^{1/2}L_1^{-1}T^{-1/2})w_n' = (T^{-1/2}L_1^{-1}T^{-1/2})(T^{1/2}L_2^{-1}T^{-1/2})w_n'. \tag{19}$$

Thus, these compact, self-adjoint operators commute on all of the Hilbert space. This implies that they have a common orthonormal basis of eigenfunctions $\{w_n\}$ (von Neumann, 1929–1930, p. 401, or see the discussion in Dunford and Schwartz, 1958, p. 927). Note that if all the $\{\lambda_n\}$ are distinct, one could show that $\{w_n\} = \{w_n'\}$ but in general these bases will likely differ. Then $\{w_n\}$ is also a common basis of eigenfunctions for the self-adjoint inverses $T^{1/2}L_1T^{-1/2}$ and $T^{1/2}L_2T^{-1/2}$ with corresponding real eigenvalues $\{\beta_n\}$ and $\{\lambda_n\}$, respectively. As in equation (13), we get that $\{u_n\} = \{T^{-1/2}w_n\}$ is a total set of eigenfunctions for L_1 and L_2, with real eigenvalues $\{\beta_n\}$ and $\{\lambda_n\}$, respectively. Thus, with the coefficients L_1 and L_2, equation (1) becomes a symmetrizable system.

Corollary 2. In equation (1), assume that $L_1 = L_1' + \alpha I$ and $L_2 = L_2' + \beta I$, where α and β are real constants and L_1' and L_2' satisfy the hypotheses of Theorem 2. Then equation (1) is a symmetrizable system. □

This is immediate from Theorem 2. Note that this

hypotheses on L_1 and L_2 is equivalent to assuming in Theorem 2 that $T^{1/2}L_1T^{-1/2}$ and $T^{1/2}L_2T^{-1/2}$ have compact resolvants $R(\alpha, L_1)$ and $R(\beta, L_2)$ for some real α and β, rather than assuming that $T^{1/2}L_1^{-1}T^{-1/2}$ and $T^{1/2}L_2^{-1}T^{-1/2}$ are compact.

We will now argue that the symmetrizable systems of Theorem 2 and of Corollaries 1 and 2 can be analyzed using modal analysis on a dense domain of solutions. In particular, let

$$D_{00} = \{u \text{ in } D(L_1L_2): L_1L_2u = L_2L_1u, \text{ where } u, L_1u \text{ and } L_2u \text{ are in } D(T^{1/2})\}. \tag{20}$$

Note that D_{00} contains the dense span of the common eigenvectors $\{u_n\}$ for L_1 and L_2.

Theorem 3. Under the hypotheses of Theorem 2, we can exhibit a modal analysis for the symmetrizable system (1); that is, for

$$u_{tt} + L_1u_t + L_2u = f, \tag{21}$$

where f is in $D(T^{1/2})$ and u is in the dense domain D_{00} of equation (20).

Proof. For u in D_{00} and f in $D(T^{1/2})$, we can set $u = T^{-1/2}w$ and $f = T^{-1/2}F$ to get

$$T^{-1/2}w_{tt} + L_1T^{-1/2}w_t + L_2T^{-1/2}w = T^{-1/2}F. \tag{22}$$

Premultiplying by $T^{1/2}$ gives the self-adjoint system

$$w_{tt} + T^{1/2}L_1T^{-1/2}w_t + T^{1/2}L_2T^{-1/2}w = F, \tag{23}$$

with common eigenfunction basis $\{w_n\}$ for the coefficient operators, as at the end of the proof of Theorem 1. Hence, $F(x, t)$ and the solution $w(x, t)$ can be expanded as

$$w(x,t) = \sum_{n=1}^{\infty} a_n(t)w_n(x), \quad \text{and} \quad F(x,t) = \sum_{n=1}^{\infty} b_n(t)w_n(x). \tag{24}$$

Substitution of equation (24) into equation (25) and taking the inner product with $w_n(x)$ then yields

$$(w_n, w_n)\ddot{a}_n(t) + (w_n, T^{1/2}L_1T^{-1/2}w_n)\dot{a}_n(t) + (w_n, T^{1/2}L_2T^{-1/2}w_n)a_n(t) = b_n(t). \tag{25}$$

The temporal functions for this self-adjoint system can be calculated from

$$\ddot{a}_n(t) + \beta_n\dot{a}_n(t) + \lambda_n a_n(t) = b_n(t) \tag{26}$$

and the given initial conditions, where $\{\beta_n\}$ and $\{\lambda_n\}$ are the real eigenvalues of L_1 and L_2, respectively. But these also give the temporal functions for the original system (1); since applying the bounded operator $T^{-1/2}$ to equation (24) gives expansions

$$u(x,t) = \sum_{n=1}^{\infty} a_n(t)u_n(x), \quad \text{and} \quad f(x,t) = \sum_{n=1}^{\infty} b_n(t)u_n(x), \tag{27}$$

for any solution $u(x, t)$ to equation (1) where $u(x, t)$ is in D_{00} and f is in $D(T^{1/2})$.

Corollary 3. Under the hypotheses of Corollary 1 or Corollary 2, we obtain a modal analysis for the symmetrizable system (1). □

Equations (26) and (27) constitute the modal expansion and solution for the nonself-adjoint case with commuting symmetrizable coefficient operators, under the given compactness and domain hypotheses. In the terms of the language used by Caughey and O'Kelly (1965), this yields a "normal mode system" and the real eigenfunctions $\{u_n\}$ are referred to as "classical normal modes." Equations (26) and (27) also constitute a theoretical modal analysis for the damped nonself-adjoint case. The modal testing community refers to this simply as the "real mode case." A physical interpretation is that the eigenfunctions $u_n(x)$ are the mode shapes, the corresponding eigenvalues λ_n are the squares of the natural frequencies (recall $\lambda_n > 0$ for all n, whenever P is positive definite, as noted in the discussion following equation (9)), and the temporal coefficients $a_n(t)$ are referred to as the modal participation factors.

Oscillation Conditions

The concepts of underdamping and critical damping commonly used in discussing single degree of freedom systems have been extended to self-adjoint distributed parameter systems by Inman and Andry (1982). Here, these results are extended to include normal mode symmetrizable systems.

Definition 3. To review the definitions given by Inman and Andry (1982), we define a system such as equation (1) to be *underdamped* if there exists a solution to equation (1) of the form:

$$u(x,t) = \sum_{n=1}^{\infty} a_n(t)u_n(x), \tag{28}$$

such that each of the temporal functions $a_n(t)$ is an underdamped function of time. That is, the system is underdamped if each $a_n(t)$ is of the form:

$$a_n(t) = A_ne^{-\zeta_n\omega_nt}\sin(\omega_{dn}t + \emptyset_n), \tag{29}$$

where A_n and $\emptyset_n$ are constants determined by initial conditions, ζ_n is the nth damping ratio, ω_n the nth undamped natural frequency, and ω_{dn} is the nth damped natural frequency. Likewise, the system defined by equation (1) is said to be *critically damped* if equation (28) holds and each of the temporal coefficient $a_n(t)$ is a critically damped function of time of the form:

$$a_n(t) = (A_n + B_nt)e^{-\omega_nt}, \tag{30}$$

where again A_n and B_n are constants determined by the initial conditions. Furthermore, the system is defined to be *overdamped* if each $a_n(t)$ is overdamped, i.e., of the form:

$$a_n(t) = A_ne^{-\lambda_1t} + B_ne^{-\lambda_2t}, \tag{31}$$

and *mixed damped* if there is at least one $a_n(t)$ in equation (28) that oscillates and at least one $a_n(t)$ that does not oscillate.

Theorem 4. In equation (1) assume that L_1 and L_2 are nonsingular unbounded operators, both symmetrizable by T, and assume that both $T^{1/2}L_1^{-1}T^{-1/2}$ and $T^{1/2}L_2^{-1}T^{-1/2}$ are compact, and that L_1 and L_2 commute on D_0 giving a modal analysis for any u in D_{00} (20). Let $\{u_n\}$ be the resulting total set of eigenfunctions common to L_1 and L_2 with eigenvalues β_n and λ_n, respectively. Then the eigenvalues of the operator $L_1^2 - 4L_2$ determine the oscillation properties of equation (1), provided $\beta_n > 0$ and $\lambda_n > 0$ for all n.

Proof. The validity of this result follows from examining the eigenstructure of the operator $L_1^2 - 4L_2$ and equation (26) with $b_n(t) = 0$ for all indices n. Since L_1 and L_2 have common eigenfunctions $\{u_n\}$, the operator $L_1^2 - 4L_2$ also has these eigenfunctions and the corresponding eigenvalues for $L_1^2 - 4L_2$ can be calculated from:

$$(L_1^2 - 4L_2)u_n = L_1^2 u_n - 4L_2 u_n = L_1(L_1 u_n) - 4\lambda_n u_n$$
$$= \beta_n L_1 u_n - 4\lambda_n u_n = (\beta_n^2 - 4\lambda_n)u_n. \quad (32)$$

Thus the eigenvalues of the operator $L_1^2 - 4L_2$ are

$$\gamma_n = \beta_n^2 - 4\lambda_n. \quad (33)$$

The discriminant associated with equation (26) determines the nature of the temporal coefficient $a_n(t)$ for each n and is also given by $\beta_n^2 - 4\lambda_n$, provides that $\beta_n > 0$ and $\lambda_n > 0$ for all n. Hence, the sign of the eigenvalues of $L_1^2 - 4L_2$ determines the oscillatory nature of the temporal functions. Specifically, if the eigenvalues of $L_1^2 - 4L_2$ are all zero ($L_1^2 = 4L_2$ on D_{00}), then the system is *critically damped* and if they are all positive numbers the system is *overdamped*. In either case there is no oscillation. If, on the other hand, the eigenvalues are all negative, the system is *underdamped* and if at least one is positive and at least one is negative the system is *mixed damped*. In the last two cases the system exhibits oscillation. □

These oscillation results become more useful if instead of being stated in terms of operator eigenvalues the results can be extended so that they can be stated in terms of definiteness conditions of the coefficient operators. This is especially true in the context of the control theory (Gibson, 1981). Definiteness conditions allow a direct formulation of oscillation conditions in terms of physical parameters (Inman, 1982, 1984). For example, in the self-adjoint case, simple integration by parts of the operator $L_1^2 - 4L_2$ can be used to derive relationships between a structure's mass density, elastic modulus, and loss factor that specify the oscillation behavior of each $a_n(t)$ (see the examples in Inman and Andry, 1982; Gibson, 1981).

These results can be extended to the nonself-adjoint case under consideration. Equation (6) demonstrates that both products $P_1 = TL_1$ and $P_2 = TL_2$ are self-adjoint. We will show that TL_1^2 must also be self-adjoint. Then we can show that the definiteness of the symmetric operator

$$TL_1^2 - 4TL_2 \quad (34)$$

determines the oscillation behavior of the solutions to equation (1) (provided $\beta_n > 0$ and $\lambda_n > 0$, for all n). A particular class of operators L_1 satisfying this condition is of course the proportional damping case given in equation (14). To develop this result we need to know (or assume) that the total set of $\{u_n\}$ of eigenfunctions obtained in Theorem 1 is contained in $D(T)$. One can show that this holds in the (likely) event that either L_1 or L_2 has a compact inverse (Olsen and Inman, 1987).

Theorem 5. Let L be an unbounded nonsingular operator symmetrizable by T, where $T^{1/2}L^{-1}T^{-1/2}$ is compact. Assume that the total set of eigenfunctions for L in Theorem 1 is contained in $D(T)$. Then L^2 is also symmetrizable by the same T, $T^{1/2}L^{-2}T^{-1/2}$ is compact, and TL^2 is self-adjoint.

Proof. In order to show L^2 is symmetrizable, observe first that since Tu_n is defined for the eigenfunctions $\{u_n\} = \{T^{-1/2}\omega_n\}$ with $Lu_n = \beta_n u_n$, β_n real; then

$$\beta_n^2 Tu_n = T\beta_n^2 u_n = TL^2 u_n \quad (35)$$

is defined. Since all the β_n are real,

$$(TL^2 u_n, u_n) = (\beta_n^2 Tu_n, u_n)$$
$$= (u_n, \beta_n^2 Tu_n)$$
$$= (u_n, TL^2 u_n). \quad (36)$$

Thus TL^2 is symmetric on the span of the $\{u_n\}$, which is dense. Furthermore, the range of TL^2 on this domain is dense, for

$$TL^2 u_n = \beta_n^2 Tu_n = \beta_n^2 TT^{-1/2}w_n = \beta_n^2 T^{1/2}w_n, \quad (37)$$

and we can show that the set $\{T^{1/2}w_n\}$ is total. That is, if $(v, T^{1/2}w_n) = 0$, then $(T^{1/2}v, w_n) = 0$, for all n. Hence $T^{1/2}v = 0$, since $\{w_n\}$ is an orthonormal basis, so $v = 0$, since $T^{1/2}$ is nonsingular. Thus $\{T^{1/2}w_n\}$ is total, and we conclude L^2 is symmetrizable by T. Also for any u in $\mathcal{H}$,

$$(T^{1/2}L^{-1}T^{-1/2})(T^{1/2}L^{-1}T^{-1/2})u = T^{1/2}L^{-1}L^{-1}T^{-1/2}u$$
$$= T^{1/2}L^{-2}T^{-1/2}u. \quad (38)$$

Thus $T^{1/2}L^{-2}T^{-1/2}$, as the square of a compact self-adjoint operator, is also compact and self-adjoint. Thus $P = TL^2$ is self-adjoint, as in equation (6). Furthermore, the inverse operator $T^{1/2}L^2T^{-1/2}$ is also self-adjoint.

Theorem 6. Let L_1 and L_2 be nonsingular unbounded operators symmetrizable by T satisfying the hypotheses of Theorem 4. Let D denote the dense span of the resulting eigenfunctions for L_1 and L_2 with real eigenvalues β_n and λ_n, respectively, and assume $D \subset D(T)$. Then the definiteness properties of the symmetric operator $TL_1^2 - 4TL_2$ on D determine the oscillation properties of equation (1), provided $\beta_n > 0$ and $\lambda_n > 0$, for all n. To be precise: the system given by equation (1) is *underdamped* if and only if $4TL_2 - TL_1^2$ is positive definite on D, *critically damped* if and only if $L_1^2 = 4L_2$ on D, *overdamped* if and only if $TL_1^2 - 4TL_2$ is positive definite on D, and *mixed damped* if and only if $TL_1^2 - 4TL_2$ is indefinite on D.

Proof. We have that $TL_1^2 - 4TL_2$ is defined on D, since $D \subset D(T)$. And $TL_1^2 - 4TL_2$ is symmetric on D, since TL_2 is self-adjoint as in equation (6), and TL_1^2 is self-adjoint by Theorem 5. Using $u_n = T^{-1/2}w_n$,

$$((TL_1^2 - 4TL_2)u_n, u_m) = (T(\beta_n^2 - 4\lambda_n)u_n, u_m)$$
$$= (\beta_n^2 - 4\lambda_n)(T^{1/2}u_n, T^{1/2}u_m)$$
$$= (\beta_n^2 - 4\lambda_n)(w_n, w_m)$$
$$= \begin{cases} \beta_n^2 - 4\lambda_n, & \text{if } n = m \\ 0, & \text{if } n \neq m, \end{cases} \quad (39)$$

since $\{w_n\}$ is an orthonormal basis. Thus, if the symmetric operator

$$TL_1^2 - 4TL_2 = 0 \quad (40)$$

on D, then $\beta_n^2 - 4\lambda_n = 0$, i.e., the eigenvalues of $L_1^2 - 4L_2$ are all zero so that, as in Theorem 4, the system is critically damped. Conversely, if $\beta_n^2 - 4\lambda_n = 0$ for all n, then $TL_1^2 - 4TL_2 = 0$ on D.

As in equations (10)–(12), the $\{w_n\}$ are a complete orthonormal system of eigenvectors for the self-adjoint operators $T^{1/2}L_1T^{-1/2}$ and $T^{1/2}L_2T^{-1/2}$ with eigenvalues β_n and λ_n, respectively. Thus, arguing as in equation (32), $\{w_n\}$ is also a complete set of eigenvectors for the symmetric operator

$$S = (T^{1/2}L_1T^{-1/2})^2 - 4T^{1/2}L_2T^{-1/2}, \quad (41)$$

with eigenvalues $\beta_n^2 - 4\lambda_n$, for each n.

If $TL_1^2 - 4TL_2 > 0$ on D, then equation (39) shows the eigenvalues of $L_1^2 - 4L_2$ are all positive, so the system is overdamped. Conversely, suppose that $\beta_n^2 - 4\lambda_n > 0$, for all n. Then S is positive definite. For any nonzero u in D, we have $T^{1/2}u = w$ for some w in span $\{w_n\}$, so Sw is defined and

$$((TL_1^2 - 4TL_2)u, u) =$$
$$= (T^{1/2}L_1^2T^{-1/2} - 4T^{1/2}L_2T^{-1/2})T^{1/2}u, T^{1/2}u$$
$$= ((T^{1/2}L_1T^{-1/2}T^{1/2}L_1T^{-1/2} - 4T^{1/2}L_2T^{-1/2})w, w)$$
$$= (Sw, w) > 0. \quad (42)$$

Thus, $TL_1^2 - 4TL_2$ is positive definite on D. Similarly, if $TL_1^2 - 4TL_2 < 0$ on D, then the eigenvalues of $L_1^2 - 4L_2$ are all negative, so the system is underdamped. And, as above, if β_n^2

$- 4\lambda_n < 0$ for all n, then $TL_1^2 - 4TL_2$ is negative definite on D.

Suppose now that $TL_1^2 - 4TL_2$ is indefinite on D. Then for some u, v in D,

$$((TL_1^2 - 4TL_2)u,u) > 0 \quad \text{and} \quad ((TL_1^2 - 4TL_2)v,v) < 0. \tag{43}$$

Write $u = T^{-1/2}w$, $v = T^{-1/2}y$, where w and y are in the span of $\{w_n\}$. Then for S as in equation (41) we have, as in equation (42) that

$$((TL_1^2 - 4TL_2)u,u) = (Sw,w). \tag{44}$$

Thus $(Sw, w) > 0$, and similarly $(Sy, y) < 0$. So the symmetric operator S is indefinite, where we know S has an orthonormal basis of eigenfunctions $\{w_n\}$ with eigenvalues $\beta_n^2 - 4\lambda_n$; at least one of these eigenvalues must be positive and one negative. Hence, in this case, the system given by equation (1) is mixed damped.

Conversely, if $\beta_k^2 - 4\lambda_k > 0$ and $\beta_m^2 - 4\lambda_m < 0$, then $(Sw_k, w_k) > 0$ and $(Sw_m, w_m) < 0$. Arguing as in equation (44) for $u = u_m = T^{-1/2}w_m$ and $v = u_k = T^{-1/2}w_k$, we see that $TL_1^2 - 4TL_2$ must be indefinite on $D = \text{span}\{u_n\}$. □

Recall that these oscillation results require that $\lambda_n > 0$ and $\beta_n > 0$ for all indices n. This assumption is equivalent to requiring that the self-adjoint operators $P_1 = TL_1$ and $P_2 = TL_2$ be positive definite, as noted in the discussion following equation (9).

Example

The existence of nonself-adjoint structures and systems may be unfamiliar to some. Hence, a standard example of a damped distributed parameter system, Pflüger's column (Bolotin, 1963), is presented here. Other examples come from control problems with state variable feedback, motors with hydrodynamics, journal bearings, labyrinth-seal effects in motors, and gyroscopic systems (Zhang and Lallement, 1985).

Consider equation (1) with:

$$L_1 = 2\gamma I$$
$$L_2 = \alpha\partial^4/\partial x^4 + p(\ell - x)\partial^2/\partial x^2 \tag{45}$$

and boundary conditions $u(0) = u(\ell) = u''(0) = u''(\ell) = 0$. Here I is the identity operator, ℓ is the length of the rod, γ is a damping coefficient, α is a positive constant reflecting the stiffness, mass, etc., and p is a positive loading parameter. Equation (1) with these operators represents a beam, subject to a uniformly distributed follower force and external damping (viscous fluid). The boundary conditions represent a pinned-pinned configuration.

The operator $T = -\partial^2/\partial x^2$, defined on a domain smooth enough with boundary conditions $u(0) = u(\ell) = 0$, allows this problem to be transformed into a related self-adjoint problem. In particular, T is positive definite and self-adjoint on the domain D of absolutely continuous u with u' absolutely continuous and u, u', u'' in $\mathcal{L}^2[0, \ell]$, with the given boundary conditions. The inverse of T is a compact integral operator with a Green's function kernel. Note that

$$TL_1 = -2\gamma\partial^2/\partial x^2 \tag{46}$$

and

$$TL_2 = -[\alpha\partial^6/\partial x^6] - p\partial^2/\partial x^2[(\ell - x)\partial^2/\partial x^2]. \tag{47}$$

Integration by parts demonstrates that both of these operators are symmetric. The domains are dense; the domain of TL_2 being all smooth enough u with

$$u(0) = u(\ell) = u''(0) = u''(\ell) = u^{(4)}(0) = u^{(4)}(\ell) = 0. \tag{48}$$

Note also that L_1 is of the proportional form $\alpha I + \beta L_2$. Rather than attempt to verify directly that the operator $T^{1/2}L_2^{-1}T^{-1/2}$ is compact, we will show that L_2^{-1} is compact. It then follows (as shown in Olsen and Inman, 1987) that since L_2^{-1} is compact, then $T^{1/2}L_2^{-1}T^{-1/2}$ is also compact. This will also show that the range of TL_2 is dense, so that L_2 is symmetrizable by T.

To see that L_2^{-1} is compact, we first regard L_2 as a product of operators,

$$L_2 = [\alpha\partial^6/\partial x^6 + p(\ell - x)]\partial^2/\partial x^2. \tag{49}$$

The factor $\alpha^2/\partial x^2$ is negative definite and self-adjoint on the domain D, with a compact inverse. The factor $\alpha\partial^2/\partial x^2 + p(\ell - x)$ is also self-adjoint on D, being the sum of a self-adjoint (unbounded) operator with a bounded self-adjoint operator. This factor can be shown to have a compact inverse, except for certain values of p and α ($p/\alpha = \lambda_n$, $n = 1, 2, \ldots$, $\{\lambda_n\}$, a discrete set). To see this, note that if u is in D with

$$[\alpha\partial^2/\partial x^2 + p(\ell - x)]u = 0, \tag{50}$$

then

$$\alpha u'' = -p(\ell - x)u, \tag{51}$$

so u is in the domain of the unbounded operator $-(1/\ell - x)\partial^2/\partial x^2$, and

$$[-(1/\ell - x)\partial^2/\partial x^2]u = \lambda u, \quad \lambda = p/\alpha. \tag{52}$$

A computation of the Green's function for this operator shows that it has a compact inverse, and hence at most countably many discrete eigenvalues $\{\lambda_n\}$. Thus the operator $\alpha\partial^2/\partial x^2 + p(\ell - x)$ has an inverse for $p/\alpha \neq \lambda_n$, $n = 1, 2, \ldots$. Furthermore, this inverse is compact: $\alpha\partial^2/\partial x^2 + p(1 - x)$ is the nonsingular sum of a self-adjoint operator with compact inverse plus a bounded operator, and any such sum has a compact inverse (see Appendix). Thus L_2 is the product of self-adjoint operators, each having a compact inverse, so L_2^{-1} is compact. Thus the coefficient operators in this example satisfy the conditions of our theory (in particular, of Corollaries 1, 2, and 3), so that this problem can be analyzed by modal methods.

Discussion

Several results on the dynamics of nonconservative, nonself-adjoint systems have been presented. Using symmetric operator products, a subclass of nonself-adjoint systems, called symmetrizable systems, has been defined. This subclass roughly coincides with the class of nonself-adjoint problems considered by Leipholz (1980, 1984) and Walker (1973, 1977). In his early work Leipholz (1974a, 1974b) examined nonself-adjointness by defining a weighted inner product (equation (4)). The weighting operator is equivalent to the operator products used in the approach presented here, which reproduces and extends this result. In his later work, Leipholz (1980, 1984) used perturbation methods to derive extensions of Mercer's theorem and the Hilbert Schmidt theorem to gain the existence of a complete set of eigenfunctions to be used for a modal expansion. Here, the operator products and compactness are used to produce the expansions. As pointed out by Leipholz (1980) these expansions are of great practical importance because they furnish a means of solving certain nonself-adjoint problems.

Walker's work (1973, 1977) concentrates on stability analysis of general systems of the form specified in equation (1), including the effects of damping. His approach is similar inasmuch as it involves a search for a symmetric operator G that causes a Liapunov functional to become symmetric. The difference between his approach and the one taken here is that G is bounded. Carr and Malhardeen (1979) also examined symmetrizing operators for a specific operator, as part of a proof indicating the exact value of its first buckling mode for the operator $(\partial^4/\partial x^4) + p(\partial^2/\partial x^2)$. All of the previous work has been primarily concerned with dynamic stability. The work presented here, however, has as its primary concern the

existence of modal expansions and oscillation conditions. However, it is an easy exercise to see that the symmetrizable operators, as defined here, can be used to generate Liapunov functionals. These Liapunov functionals can then be used to reproduce some of the stability results contained in the works by Walker (1973, 1977) and Leipholz (1974a, 1974b).

The results presented above have physical relevance inasmuch as they extend and define the simple and very useful concepts of critical damping, overdamping and underdamping to a class of nonself-adjoint distributed parameter systems.

Acknowledgments

This work was supported in part by National Science Foundation Grant number MSM-8351807, and in part by Air Force Office of Scientific Research, grant number AFOSR 85-0220.

References

Bolotin, V. V., 1963, *Nonconservative Problems of the Theory of Elastic Stability*, Pergamon Press, New York.

Carr, J., and Malhardeen, M. Z. M., 1979, "Beck's Problem," *SIAM Journal of Applied Mathematics*, Vol. 37, No. 2, pp. 261–262.

Caughey, T. K., and O'Kelly, M. E. J., 1965, "Classical Normal Modes in Damped Linear Dynamic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 538–588.

Dieudonnè, J., 1961, "Quasi-Hermitian Operators," *Proceedings of the International Symposium on Linear Spaces*, Jerusalem Academic Press, Jerusalem, pp. 115–122.

Dunford, N., and Schwartz, J. T., 1958, *Linear Operators*, Part II, Wiley Interscience, New York.

Gibson, J. S., 1981, "An Analysis of Optimal Modal Regulation: Convergence and Stability," *SIAM Journal of Control and Optimization*, Vol. 19, pp. 686–706.

Inman, D. J., 1984, "Finite Control of Underdamped Distributed Parameter Systems," *Proceedings of the 1984 JPL Workshop on Identification and Control of Flexible Space Structures*, Vol. 11, pp. 101–109.

Inman, D. J., 1983, "Dynamics of Asymmetric Nonconservative Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, No. 1, pp. 199–203.

Inman, D. J., 1982, "Oscillatory Damped Distributed Parameter Systems," *Mechanics Research Communications*, Vol. 29, No. 2, pp. 101–107.

Inman, D. J., and Andry, A. N., Jr., 1982, "The Nature of the Temporal Solutions of Damped Distributed Systems wtih Classical Normal Modes," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 867–870.

Kato, T., 1966, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York.

Lax, P. D., 1954, "Symmetrizable Linear Transformations," *Communications on Pure and Applied Mathematics*, Vol. VII, pp. 633–647.

Leipholz, H. H. E., 1984, "Extension of the Hilbert Schmidt Theorem to a Certain Class of Problems with Unsymmetric Kernels," *Acta Mechanica*, Vol. 51, pp. 69–84.

Leipholz, H. H. E., 1980, "An Extension of Mercer's Theorem to a Class of Nonsymmetric Integral Equations," *Acta Mechanica*, Vol. 39, pp. 99–110.

Leipholz, H. H. E., 1974a, "On Conservative Elastic Systems of the First and Second Kind," *Ingenieur-Archiv*, Vol. 43, pp. 255–271.

Leipholz, H. H. E., 1974b, "On a Generalization of the Concepts of Self Adjointness and of Rayleigh's Quotient," *Mechanics Research Communications*, Vol. 1, pp. 67–72.

Olsen, C. L., and Inman, D. J., 1987, "Some Results on Symmetrizable Operators in a Hilbert Space," Mechanical Systems Laboratory Report No. 87-11, State University of New York at Buffalo, Buffalo, NY.

Taussky, O., 1968, "Positive Definite Matrices and Their Role in the Study of the Characteristics Roots of General Matrices," *Advances in Mathematics*, Vol. 2, pp. 175–186.

von Neumann, J., 1920–1930, "Zur Algebra der Funktionaloperationen und Theorie der Normalen Operatoren," *Mathematische Annalen 102*, pp. 370–427.

Walker, J. A., 1977, "Liapunov Analysis of the Generalized Pflüger Problem," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 935–938.

Walker, J. A., 1973, "Energy-Like Liapunov Functionals for Linear Elastic Systems on a Hilbert Space," *Quarterly of Applied Mathematics*, Vol. 30, No. 4, pp. 465–480.

Zhang, Q., and Lallement, G., 1985, "Experimental Determination of Eigensolutions of Nonself-adjoint Mechanical Structures," *Proceedings of the 3rd International Modal Analysis Conference*, Vol. 11, pp. 768–774.

APPENDIX

Theorem

Let A be a self-adjoint operator on a Hilbert space $\mathcal{H}$ with a compact inverse, and let B be a bounded self-adjoint operator on $\mathcal{H}$. Then the spectrum of the self-adjoint operator $A+B$ is a discrete set of eigenvalues $\{u_n\}$. The dimension of the space of eigenfunctions for each u_n is finite.

Proof

Observe first that for any orthonormal sequence $\{e_n\}$ of functions in $\mathcal{H}$ with $\{e_n\} \subset D(A)$, $\|Ae_n\| \to \infty$. To see this, consider a complete set $\{w_k\}$ of eigenfunctions for A which is an orthogonal basis for $\mathcal{H}$. Let β_k be the associated eigenvalues and order the sequences so that $|\beta_k|$ converges monotonically to $+\infty$ (using the spectral theorem for compact self-adjoint operators, Kato, 1966). If $\{e_n\}$ is an arbitrary sequence of orthonormal functions in $\mathcal{H}$ with $\{e_n\}$ in $D(A)$, then the $\lim_{n\to\infty} (e_n, w_k) = 0$, for each k. For any given large M and small $e>0$, there is a large enough index K, so that $|\beta_k| \le M$ for $k \le K$ and $|\beta_k| > M$ for $k > K$. It is then possible to find a large enough N so that for $n > N$,

$$\sum_{k\le K} |(e_n, w_k)|^2 < \epsilon^2/M^2.$$

Then for $n > N$,

$$\|Ae_n\| = \|A \sum_{k\le K} (e_n, w_k) w_k + A \sum_{k>K} (e_n, w_k) w_k\|$$

$$\ge \|A \sum_{k>K} (e_n, w_k) w_k\|$$

$$\ge M \|\sum_{k>K} (e_n, w_k) w_k\|$$

$$> M(1 - \epsilon/M)$$

$$= M - \epsilon,$$

and we conclude $\|Ae_n\| \to \infty$.

It now follows that for any orthonormal sequence of functions $\{e_n\}$, in $D(A+B) = D(A)$, that $\|(A+B)e_n\| \to \infty$ since:

$$\|(A+B)e_n\| \ge \|Ae_n\| - \|Be_n\|$$

$$\ge \|Ae_n\| - \|B\|.$$

Then for $E(\lambda)$, the spectral resolution for $A+B$, the preceding implies that $E[a, b]$ is finite-dimensional for any bounded interval $[a, b]$. Hence the spectrum of $A+B$ consists of discrete eigenvalues with associated eigenspaces being finite-dimensional.

X. M. Gu
P. R. Sethna
Mem. ASME
Department of Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, MN 55455

A. Narain
Department of Mechanical Engineering, Michigan Technological University, Houghton, MI 49931

On Three-Dimensional Nonlinear Subharmonic Resonant Surface Waves in a Fluid: Part I—Theory*

Three-dimensional surface waves in a rectangular container subjected to vertical excitation are studied. The analysis includes the effects of surface tension, energy dissipation, and critical depth. Both steady state and transient phenomena are discussed.

Introduction

Study of free surface waves in a container subjected to vertical oscillations goes back to the works of Faraday (1831). More recent studies start with the work of Benjamin and Ursell (1954) who study the linearized problem. Among the other references are Skalak and Yarymovych (1962) who have done nonlinear analysis of deep water two-dimensional waves in rectangular containers. Dodge et al. (1965) study nonlinear surface waves in circular containers of finite fluid depth and Chu (1968) has done a general study for arbitrary axisymmetric containers. For studies of nonlinear two and three-dimensional unforced waves in rectangular containers we refer to the work of Tadjbakhsh and Keller (1960) and Verma and Keller (1962), respectively. In these works they show, for the first time, the role of a critical depth of the fluid in the qualitative behavior of unforced waves. The most recent contribution to this same and a related subject is by Miles (1984). Most of the above work deals with the special case of wave motions when only one of the spatial modes, corresponding to the linearized problem, is assumed to be excited.

There is now a renewed interest in the subject due to the possibility of having chaotic motions when more than one spatial mode interact. Miles (1984) gives a general Lagrangian formulation for nonlinear gravity waves, which results in a one-mode system and a two-mode system. Studies when two modes dominate and chaotic motions are possible have been done by Holmes (1986).

The work presented here is part of a larger study by Gu (1986). We present here a study of three-dimensional nonlinear waves under vertical excitations, again in the special case when only one spatial mode is strongly excited. Our work differs from and is a generalization of previous work in several respects. We present a very general formulation of the problem which gives insights into the variety of complex phenomena that are possible. As in the references, we also look for motion which is predominantly in one mode; but in our case, because of the above mentioned general formulation, it is possible to show that it is impossible to have the motion in one dominant mode unless additional necessary conditions are satisfied and we give these conditions explicitly. In fact, without the knowledge of these conditions, it would have been difficult to get good experimental results as given in Part II. Our analysis also generalizes the previous results. We include the effects of surface tension and effect of energy dissipation, and we discuss not only periodic but also transient phenomena. The results on transient motions are a generalization of Miles (1984). The results of Miles reduce to those given here when energy dissipation is ignored. We show that for a given set of parameters as many as five periodic solutions, three stable and two unstable, are possible. Furthermore, we give estimates of basins of attraction of the three stable periodic motions. Our results, when specialized for unforced motions, confirm the results on critical depth of Tadjbakhsh and Keller (1960) and by Verma and Keller (1962).

Problem Formulation

We discuss the problem formulation very briefly since it is broadly available in the literature by Whitham (1973). Consider a rigid rectangular container containing an inviscid incompressible fluid. The container is subjected to vertical motions $f(t) = -\Delta \cos \omega t$ with respect to a fixed reference frame. We assume that the flow is irrotational and thus, there exists a potential function $\phi(\mathbf{x}, t)$, where $\mathbf{x}$ is the position with respect to the container, so that $\nabla\phi = \mathbf{V}$, where $\mathbf{V}$ is the fluid velocity relative to the container, and, due to the incompressibility, ϕ, satisfies Laplace's equation. The origin of a moving Cartesian reference frame is attached to the container at the height of the undisturbed free surface of the fluid. The cross section of the container is assumed to be rectangular of sides a and b and the fluid depth is h. The free surface boundary conditions can be given in terms of the deviation from the undisturbed free surface $z = \eta(x,y,t)$.

We introduce the following dimensionless quantities

*This work was supported by Grant NSF-MEA-8310966.
Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August 5, 1986; final revision July 25, 1987.

$$\Delta = \epsilon b \Delta', \; x = bx', \; y = by', \; z = bz', \; t = \frac{1}{\omega} t',$$

$$\phi = \epsilon \omega b^2 \phi', \; \eta = \epsilon b \eta', \; G = \frac{g}{\omega^2 b}, \; \tilde{\gamma} = \frac{\gamma}{\rho \omega^2 b^3},$$

$$\lambda_1 = \frac{a}{b}, \; \lambda = \frac{h}{b}$$

where g and γ are the gravity and the surface tension constants, respectively, and ϵ is a small parameter. Dropping the primes for simplicity, we obtain the field equations and boundary conditions in dimensionless form as follows. The potential function satisfies:

$$\nabla^2 \phi = 0, \text{ in } 0 \leqq x \leqq \lambda_1, \; 0 \leqq y \leqq 1,$$
$$-\lambda \leqq z \leqq \epsilon \eta(x,y,t). \tag{1}$$

along with the boundary conditions on the rigid walls:

$$\phi_x = 0, \quad \text{on } x = 0, \; \lambda_1;$$
$$\phi_y = 0, \quad \text{on } y = 0, \; 1;$$
$$\phi_z = 0, \quad \text{on } z = -\lambda. \tag{2}$$

The dynamic boundary conditions on the free surface, $z = \epsilon\eta(x, y, t)$, takes the form:

$$\phi_t + \frac{1}{2} \epsilon (\nabla \phi)^2 + (G + \epsilon \Delta \cos t)\eta$$
$$= \tilde{\gamma} \Big[(\eta_{xx} + \eta_{yy}) - \frac{\epsilon^2}{2} (\eta_{xx}\eta_y^2 + \eta_{yy}\eta_x^2 + 3\eta_{xx}\eta_x^2$$
$$+ 3\eta_{yy}\eta_y^2 + 4\eta_x\eta_y\eta_{xy}) \Big] + 0(\epsilon^2). \tag{3}$$

The kinematic boundary condition on the free surface is

$$\phi_z = \epsilon[\phi_x \eta_x + \phi_y \eta_y] + \eta_t, \tag{4}$$

and from the incompressibility of the fluid, we have the condition

$$\int_0^1 \int_0^{\lambda_1} \eta(x, y, t) dx \, dy = 0. \tag{5}$$

Asymptotic Analysis

We assume the nth order asymptotic expansion of the solution of equations (1) through (5), asymptotic as $\epsilon \to 0$, of the form

$$\phi = \phi(x, y, z, \mathbf{a}, \theta, t, \epsilon) = \sum_{n=0}^{N} \sum_{ij=0}^{\infty} \epsilon^n q_{ij}^{(n)}(t) S_{ij}(x, y) g_{ij}(Z),$$

$$\eta = \eta(x, y, \mathbf{a}, \theta, t, \epsilon) = \sum_{n=0}^{N} \sum_{ij=0}^{\infty} \epsilon^n p_{ij}^{(n)}(t) S_{ij}(x, y), \tag{6}$$

where

$$S_{ij}(x, y) = \cos \frac{i\pi x}{a} \cos j\pi y,$$
$$g_{ij}(z) = \cosh K_{ij}(\lambda + z) \cosh^{-1} K_{ij}\lambda, \tag{7}$$
$$K_{ij}^2 = \pi^2 \Big[\frac{i^2}{a^2} + j^2 \Big].$$

When $\epsilon = 0$ we have

$$p_{ij}^{(0)}(t) = a_{ij} \cos(\mu_{ij} t + \theta_{ij}),$$
$$q_{ij}^{(0)}(t) = -\frac{\mu_{ij}}{\alpha_{ij}} a_{ij} \sin(\mu_{ij} t + \theta_{ij}), \tag{8}$$

where

$$\mu_{ij} = [(G + \gamma K_{ij}^2)\alpha_{ij}]^{1/2}, \; \alpha_{ij} = k_{ij} \tanh k_{ij}\lambda$$

and a_{ij} and θ_{ij} are arbitrary constants. For $i = j = 0$, (1.5) yields $p_{00}^{(0)} = 0$ and $q_{00}^{(0)} = 0$.

When $\epsilon \neq 0$, the method of analysis is an adaptation of the method of averaging as given by Mitropolsky (1965) and as employed by Bajaj and Sethna (1980). The solution is assumed to depend on two infinite-dimensional vectors $\mathbf{a}$ and θ, the time derivative of which will be quantities of $0(\epsilon)$ as follows:

$$\dot{a}_{ij} = \sum_{n=1}^{N} \epsilon^n A_{ij}^{(n)}(\mathbf{a}, \theta); \; i, j = 0, 1, 2, \ldots, \; i = j \neq 0, \tag{9}$$

$$\dot{\theta}_{ij} = \sum_{n=1}^{N} \epsilon^n B_{ij}^{(n)}(\mathbf{a}, \theta); \; i, j = 0, 1, 2, \ldots, \; i = j \neq 0. \tag{10}$$

Substituting equation (6) into equations (3) and (4), taking into account equations (9) and (10), equating equal powers of ϵ, and taking projections, we have (for details see Gu, 1986)

$$\frac{\partial p_{mn}^{(1)}}{\partial t} - \alpha_{mn} q_{mn}^{(1)} = -\Big[\frac{\partial p_{mn}^{(0)}}{\partial a_{mn}} A_{mn}^{(1)} + \frac{\partial p_{mn}^{(0)}}{\partial \theta_{mn}} B_{mn}^{(1)} \Big]$$
$$- G_{uvrs}^{(1)mn} p_{uv}^{(0)} q_{rs}^{(0)}, \tag{11}$$

$$\frac{\partial q_{mn}^{(1)}}{\partial t} + \frac{\mu_{mn}^2}{\alpha_{mn}} p_{mn}^{(1)} = -\Delta \cos t \, p_{mn}^{(0)}$$
$$- \Big[\frac{\partial q_{mn}^{(0)}}{\partial a_{mn}} A_{mn}^{(1)} + \frac{\partial q_{mn}^{(0)}}{\partial \theta_{mn}} B_{mn}^{(1)} \Big]$$
$$- \Big[\frac{1}{2} G_{uvrs}^{(2)mn} q_{uv}^{(0)} q_{rs}^{(0)} + G_{uvrs}^{(3)mn} p_{uv}^{(0)} \frac{\partial q_{rs}^{(0)}}{\partial t} \Big] \tag{12}$$

for terms of order ϵ and

$$\frac{\partial p_{mn}^{(2)}}{\partial t} - \alpha_{mn} q_{mn}^{(2)} =$$
$$- \Big[\frac{\partial p_{mn}^{(0)}}{\partial a_{mn}} A_{mn}^{(2)} + \frac{\partial p_{mn}^{(0)}}{\partial \theta_{mn}} B_{mn}^{(2)} + \frac{\partial p_{mn}^{(1)}}{\partial a_{mn}} A_{mn}^{(1)}$$
$$+ \frac{\partial p_{mn}^{(1)}}{\partial \theta_{mn}} B_{mn}^{(1)} \Big] - G_{uvrs}^{(1)mn} (p_{uv}^{(1)} q_{rs}^{(0)}$$
$$+ p_{uv}^{(0)} q_{rs}^{(1)}) + H_{uvrskl}^{(1)mn} p_{uv}^{(0)} p_{rs}^{(0)} q_{kl}^{(0)}, \tag{13}$$

$$\frac{\partial q_{mn}^{(2)}}{\partial t} + \frac{\mu_{mn}^2}{\alpha_{mn}} p_{mn}^{(2)} = -F \cos t \, p_{mn}^{(1)}$$
$$- \Big[\frac{\partial q_{mn}^{(0)}}{\partial a_{mn}} A_{mn}^{(2)} + \frac{\partial q_{mn}^{(0)}}{\partial \theta_{mn}} B_{mn}^{(2)} + \frac{\partial q_{mn}^{(1)}}{\partial a_{mn}} A_{mn}^{(1)}$$
$$+ \frac{\partial q_{mn}^{(1)}}{\partial \theta_{mn}} B_{mn}^{(1)} \Big] - \Big[G_{uvrs}^{(2)mn} q_{uv}^{(0)} q_{rs}^{(1)}$$
$$+ H_{uvrskl}^{(2)mn} q_{uv}^{(0)} q_{rs}^{(0)} p_{kl}^{(0)} + G_{uvrs}^{(3)mn} \Big(p_{uv}^{(0)} \frac{\partial q_{rs}^{(1)}}{\partial t}$$
$$+ p_{uv}^{(1)} \frac{\partial q_{rs}^{(0)}}{\partial t} \Big) + G_{uvrs}^{(3)mn} \Big[p_{uv}^{(0)} \frac{\partial q_{rs}^{(0)}}{\partial a_{rs}} A_{rs}^{(1)}$$
$$+ p_{uv}^{(0)} \frac{\partial q_{rs}^{(0)}}{\partial \theta_{rs}} B_{rs}^{(1)} \Big] + H_{uvrskl}^{(3)mn} p_{uv}^{(0)} p_{rs}^{(0)} \frac{\partial q_{kl}^{(0)}}{\partial t}$$
$$+ H_{uvrskl}^{(4)mn} p_{uv}^{(0)} p_{rs}^{(0)} p_{kl}^{(0)} \Big] \tag{14}$$

for terms of order ϵ^2, where the repeated subscripts, except m

and n, imply summation from zero to ∞. The definitions of the coupling coefficients $G^{(p)mn}_{uvrs}$ and $H^{(p)mn}_{uvrskl}$ are complicated expressions, as given in the appendix, and are not zero when the equations of the involved modes are spatially coupled with the (m,n)th mode.

The general procedure involves the determination of quantities $A^{(p)}_{mn}$ and $B^{(p)}_{mn}$ in equations (11) through (14). We note that equations (11) through (14) are linear. We seek solutions $p^{(p)}_{mn}$ and $q^{(p)}_{mn}$ for all m, n as periodic functions of t, regarding a_{mn} and θ_{mn} as constants. The solvability condition, since the equations are in first order form, is the Fredholm Alternative (Hale, 1969), which is used to determine the quantities $A^{(p)}_{mn}$ and $B^{(p)}_{mn}$.

Dissipation

Energy dissipation causes the attenuation of the free motion amplitude of the surface waves. The main sources of dissipation are due to the viscous dissipation in the boundary layers, free-surface contamination, and capillary hysteresis at the contact line of the free surface. These effects are discussed by Ursell and Benjamin (1954) and Miles (1967). The effect of capillary hysteresis is known to be relatively small for a clean, hydrophilic wall; we, therefore, do not include this effect in our discussion.

Based on considerations by Miles (1967) we can calculate the logarithmic decrement D_{mn}, the "damping factor" ζ_{mn}, and the "damping coefficient" d_{mn} (to be explained later) of the (m,n)th mode and their relationship with each other as follows:

$$\zeta_{mn}=\frac{D_{mn}}{2\pi} \text{ and } \epsilon d_{mn}=\omega^{-1}(\omega_{mn}\zeta_{mn})$$

$$=\frac{1}{2\pi}\mu_{mn}D_{mn}=\left(\frac{2\nu\mu_{mn}}{\omega}\right)^{1/2}H_{mn} \tag{15}$$

where H_{mn} takes the form

$$H_{mn}=\frac{1}{2b}\left\{2^{-\delta on}+\frac{2^{-\delta om}}{\lambda_1}+\frac{2^{-\delta on}m^2+2^{-\delta om}\lambda_1 n^2}{m^2+\lambda_1^2 n^2}\right.$$

$$+\left[1-2\lambda\left(2^{-\delta on}+\frac{2^{-\delta om}}{\lambda_1}-\frac{2^{-\delta on}m^2+2^{-\delta om}\lambda_1 n^2}{m^2+\lambda_1^2 n^2}\right)\right]$$

$$\left.\cdot K_{mn}(\sinh 2K_{mn}\lambda)^{-1}+\frac{C}{2}K_{mn}(\tanh K_{mn}\lambda)^{-1}\right\}. \tag{16}$$

where C is a parameter, $0 \leqq C < 2$, where the equality holds only for a clean surface, and $C = 1$ for an inextensible film. In equation (16) the Kronecker deltas are introduced to distinguish two-dimensional waves from three-dimensional waves.

Coupling Phenomena

The system of equation (11) through (14) and those of higher order can be studied under a variety of assumptions on the physical parameters. Qualitatively different phenomena occur depending on the coupling between the infinity of equations at each approximation. The coupling itself is of two kinds. It can be described as "spatial coupling" or "dynamic coupling."

In a first order theory, we see that in the system (11) and (12), coupling between separate modes occurs in a definite pattern and this pattern is determined entirely by the mode numbers as they affect the coupling coefficients $G^{(p)ij}_{uvrs}$ for $p=1,2,3$, and $H^{(p)ij}_{uvrskl}$ for $p=1,2,3,4$, An examination of these quantities shows (see appendix) that for modes with mode numbers (i, j) and (r,s) to couple, and, therefore, to affect the mode with mode numbers (m,n), the following conditions have to be satisfied:

$$i \pm r \pm m = 0 \text{ and } j \pm s \pm n = 0 \tag{17}$$

where i, j, r, s, n, and m are all non-negative integers.

There is, however, dissipation in the system and unless a mode gets energy at a steady rate, a steady (periodic) motion is not possible. This energy transfer can happen through internal resonance phenomena which are critically dependent on the frequencies of the motion of the modes, their ratios, and their relationship to the excitation frequency.

Any mode can get energy from the external excitation if the frequency of the mode is near half the frequency of the excitation. We will call the phenomena associated with the system when this happens "subharmonic external resonance." An examination of equations (11) and (12) shows that, except in very special situations, only one mode can get excited in this manner and our analysis is valid for this case.

Motion in a Single Mode

We discuss here the motion when it is predominantly in the mode that is in subharmonic resonance with the external excitation and determine the conditions under which this occurs. Restricting the discussion for now to terms of order ϵ, let the (m,n)th mode be excited by the external excitation ($\mu_{mn} \simeq 1/2$). Let all the other modes with mode numbers $(i, j) \neq (m,n)$ have motions that are of order ϵ, i.e., let $p^{(0)}_{ij}$ and $q^{(0)}_{ij} = 0$, $(i, j) \neq (m, n)$. Then examining the coupling coefficients $G^{(p)mn}_{ijrs}$, $p=1$, 2, and 3, it is clear that the modes with mode numbers $(2m, 2n)$, $(2m, 0)$ and $(0, 2n)$ are spatially coupled with the strongly excited mode with mode numbers (m, n). These modes are excited by periodic excitation with a frequency that is twice the frequency of the (m, n)th mode. For these modes not to get into resonance, it is necessary that their own natural frequencies not be near the excitation frequency. From this we obtain the desired necessary conditions for motions to occur in a single dominant externally excited mode. These conditions are that μ_{2m2n}, μ_{2m0}, and μ_{02n} not be near $2\mu_{mn}$, i.e., they must not be near the integer one in our dimensionless variable. A similar set of conditions can be obtained at higher approximations.

We now give the equations that determine the motion in this single dominant mode (for details, see Gu, 1986). Let the external excitation detuning parameter be σ so that

$$\mu^2_{mn}=\left(\frac{1}{2}\right)-\epsilon\sigma. \tag{18}$$

Then, following the general procedure outlined above, we have

$$A^{(1)}_{ij}=B^{(1)}_{ij}=0, \quad (i, j)\neq(m, n)$$

and if $\Delta_{mn}=1/2\ \Delta\alpha_{mn}$, then

$$A^{(1)}_{mn}=\Delta_{mn}A_{mn}\sin 2\theta_{mn},$$

$$B^{(1)}_{mn}=-\sigma+\Delta_{mn}\cos 2\theta_{mn}. \tag{19}$$

and in the second approximation

$$A^{(2)}_{ij}=B^{(2)}_{ij}=0, \quad (i, j)\neq(m, n)$$

and

$$A^{(2)}_{mn}=0$$

$$B^{(2)}_{mn}=M_{mn}A^2_{mn}+\left(\frac{3}{2}\Delta^2_{mn}-\sigma^2\right) \tag{20}$$

where

$$M_{mn}=-\left[\frac{1}{2}S^{(1)}_{mn}+\alpha_{mn}c^{(2)}_{mn}\right] \tag{21}$$

and

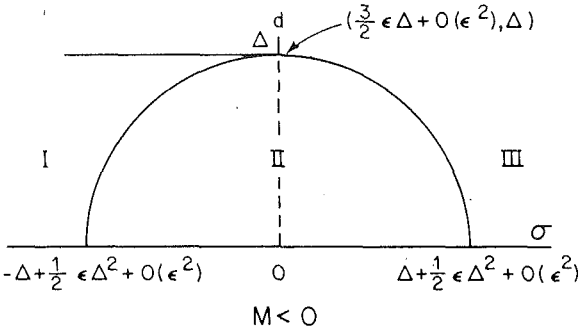


Fig. 1 Bifurcation diagram for $M<0$

$$s^{(1)}_{mn}=-\frac{1}{64\alpha^3_{mn}\mu^2_{ij}(\mu^2_{ij}-1)}\,G^{(1)mn}_{ijmn}\Big[\Big(-2\alpha_{ij}\alpha_{mn}G^{(3)ij}_{mnmn}+3\alpha_{ij}G^{(2)ij}_{mnmn}+4\alpha_{mn}G^{(1)ij}_{mnmn}\Big)\mu^2_{ij}+4\alpha_{ij}\alpha_{mn}G^{(3)ij}_{mnmn}-2\alpha_{ij}G^{(2)ij}_{mnmn}\Big]$$

$$+\frac{G^{(1)mn}_{mnij}}{32\alpha^2_{mn}(\mu^2_{ij}-1)\alpha_{ij}}\Big[4\alpha_{mn}\mu^2_{ij}G^{(1)ij}_{mnmn}+\alpha_{ij}G^{(2)ij}_{mnmn}+2\alpha_{mn}\alpha_{ij}G^{(3)ij}_{mnmn}\Big]-\frac{1}{8\alpha_{mn}}H^{(1)mn}_{mnmnmn},\tag{22}$$

$$c^{(2)}_{mn}=\frac{1}{128\alpha^3_{mn}(\mu^2_{ij}-1)}\Big\{G^{(3)ij}_{mnmn}\Big[\alpha_{ij}\alpha_{mn}\Big(6-\frac{4}{\mu^2_{ij}}\Big)G^{(3)mn}_{ijmn}-4\alpha_{mn}G^{(2)mn}_{mnij}+8\alpha^2_{mn}G^{(3)mn}_{mnij}\Big]$$

$$+G^{(2)ij}_{mnmn}\Big[\alpha_{ij}\Big(-1+\frac{2}{\mu^2_{ij}}\Big)G^{(3)mn}_{ijmn}-2G^{(2)mn}_{mnij}+4\alpha_{mn}G^{(3)mn}_{mnij}\Big]$$

$$+G^{(1)ij}_{mnmn}\Big[4\alpha_{mn}G^{(3)mn}_{ijmn}-8\mu^2_{ij}\frac{\alpha_{mn}}{\alpha_{ij}}G^{(2)mn}_{mnij}+\frac{16\alpha^2_{mn}}{\alpha_{ij}}\mu^2_{ij}G^{(3)mn}_{mnij}\Big]\Big\}$$

$$-\frac{1}{16\alpha^2_{mn}}H^{(2)mn}_{mnmnmn}-\frac{3}{4}H^{(4)mn}_{mnmnmn}+\frac{3}{16\alpha_{mn}}H^{(3)mn}_{mnmnmn},\tag{23}$$

Subharmonic Periodic Solutions and Their Stability

Our study of system behavior is based on the solutions of equations (9) and (10) with $A^{(\gamma)}_{mn}$ and $B^{(\gamma)}_{mn}$, $\gamma=1$ and 2 as in equations (19) and (20). At this point, we introduce a small "damping factor" ϵd_{mn} which depends on the mode number and which is a measure of the exponential decay rate of the motion due to dissipation. Specifically we add a term, $-\epsilon d_{mn}a_{mn}$, where d_{mn} is the same as in equations (15) to the right side of equations (9) to have

$$\dot{a}_{mn}=\epsilon(-d_{mn}+\Delta_{mn}\sin 2\theta_{mn})a_{mn}+0(\epsilon^3),$$

$$\dot{\theta}_{mn}=\epsilon(-\sigma+\Delta_{mn}\cos 2\theta_{mn})+\epsilon^2\Big[M_{mn}a^2_{mn}+\Big(\frac{3}{2}\Delta^2_{mn}-\sigma^2\Big)\Big]+0(\epsilon^3).\tag{24}$$

Let $\epsilon t=\tau$, and still use a dot to denote $d/d\tau$. Then equation (24) becomes

$$\dot{a}=(-d+\Delta\sin 2\theta)a,$$

$$\dot{\theta}=-\sigma+\Delta\cos 2\theta+\epsilon\Big[\mathrm{Ma}^2+\Big(\frac{3}{2}\Delta^2-\sigma^2\Big)\Big],\tag{25}$$

where we have deleted the subscripts mn for simplicity.

The nontrivial constant solutions of equations (25) represent periodic motions and their amplitudes can be shown to satisfy the following quadratic equation in a^2

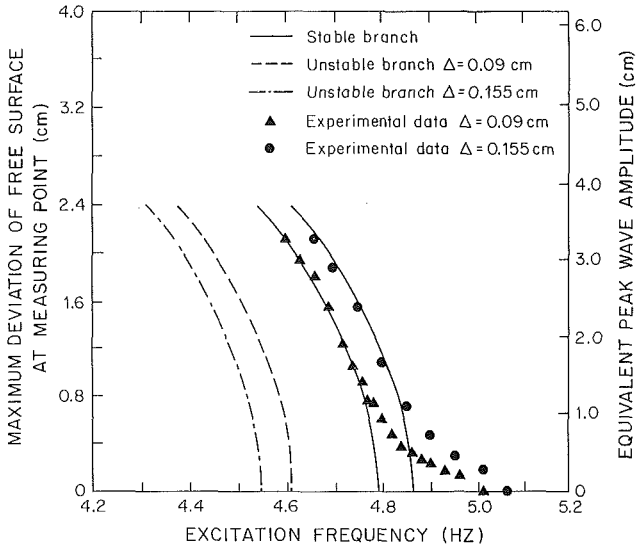


Fig. 2 Response curves compared with experiment data for a (1, 1) mode using tap water. For definition of measuring point see Part II. Container dimensions: 17.78 cm. by 22.86 cm.; depth = 11.43 cm.

$$\epsilon^2\Delta^2M^2a^4-2\Big[\sigma_1+\epsilon\Big(\sigma^2-\frac{3}{2}\Delta^2\Big)\Big]\epsilon\Delta\mathrm{Ma}^2+\Big[\sigma+\epsilon\Big(\sigma^2-\frac{3}{2}\Delta^2\Big)\Big]^2+d^2-\Delta^2=0\tag{26}$$

and the phase angle of the periodic solutions is determined by

$$\tan 2\theta=\frac{d}{\sigma-\epsilon\Big[\mathrm{Ma}^2-\Big(\sigma^2-\frac{3}{2}\Delta^2\Big)\Big]}.\tag{27}$$

Solving equation (26) gives

$$\epsilon\Delta a^2=\frac{1}{M}\Big[\sigma+\epsilon\Big(\sigma^2-\frac{3}{2}\Delta^2\Big)\pm(\Delta^2-d^2)^{1/2}\Big].\tag{28}$$

Inserting equation (28) into (27) yields

$$\theta=\begin{cases}\mp\dfrac{1}{2}\tan^{-1}\dfrac{d}{(\Delta^2-d^2)^{1/2}}\\[2ex]\pi\mp\dfrac{1}{2}\tan^{-1}\dfrac{d}{(\Delta^2-d^2)^{1/2}}\end{cases}\tag{29}$$

and thus there are four nontrivial periodic solutions. For a fixed Δ we study the solutions in (σ, d) parameter plane. The plane is divided by bifurcation curves into three open subsets of the (σ, d) plane. We first indicate the number of solutions that occur in each of these open subsets, and then discuss the nature of the bifurcations that occur on the boundaries.

The (σ, d) parameter plane is divided by a curve S_1: $[\sigma+\epsilon(\sigma^2-3/2\Delta^2)]^2+d^2-\Delta^2=0$, which represents the condition for equation (26) to have at least one zero root and a line S_2: $d-\Delta=0$ which determines the condition for a double root of equation (26) for a^2.

We note that S_1: $\sigma^2+d^2=\Delta^2+0(\epsilon)$ and since $d\geqq 0$, S_1 is approximately a semicircle of radius Δ. Furthermore the line S_2 is tangent to S_1 at

$$(\sigma,d)=\Big(\frac{-1+\sqrt{1+6\epsilon^2\Delta^2}}{2\epsilon\Delta},\ \Delta\Big)\approx\Big(\frac{3}{2}\epsilon\Delta,\Delta\Big)$$

In view of the fact that we seek real roots for a, the boundary

is determined by S_2: $d-\Delta=0$ for $\sigma<3/2\epsilon\Delta$ when $M<0$ and $d-\Delta=0$ for $\sigma>3/2\ \epsilon\Delta$ when $M>0$.

The curves S_1 and S_2 divide the (σ,d) plane into three open subsets as shown in Fig. 1 (for $M<0$). In subset I there is a zero solution and four nontrivial solutions; in subset II there is a zero solution and two nontrivial solutions; and in subset III there is only a zero solution. A typical response curve relating the amplitude of the periodic motion as a function of excitation frequency is given in Fig. 2. Also shown are experimental results which are discussed in Part II.

If we recall the definition of Δ_{mn}, we can conclude by examining equation (28) that there exists a critical value of the amplitude of excitation Δ (not Δ_{mn} without subscripts), $\Delta^*=2d_{mn}/\alpha_{mn}$, and no motion is possible unless $\Delta>\Delta^*$.

Critical Depth

We discuss here the critical role of the fluid depth on the qualitative nature of the solutions. In the absence of excitation and damping, the dimensionless amplitude is given by

$$\epsilon a_{mn}^2=\frac{1}{M_{mn}}\ (\sigma+\epsilon\sigma^2)$$

and if $(a_{mn})_p$ is the physical amplitude then

$$(a_{mn})_p^2=\frac{b^2}{M_{mn}}\ (\epsilon\sigma+\epsilon^2\sigma^2).$$

We note that M_{mn} may be regarded as a function of three variables:

$$M_{mn}=M_{mn}\left(\frac{m}{\lambda_1},n,\lambda\right)$$

and that

$$M_{mn}\left(\frac{m}{\lambda_1},n,\lambda\right)=n^2M_{mn}\left(\frac{m}{n\lambda_1},1,n\lambda\right)$$

and thus in physical terms

$$(a_{mn})_p^2=\frac{b^2}{n^2M_{mn}\left(\frac{mb}{na},1,\frac{nh}{b}\right)}\ (\epsilon\sigma+\epsilon^2\sigma^2).$$

The amplitude, for a given detuning and height, therefore, depends on the aspect ratio of a "modal cell" mb/na and the basic size of the cell b/n. The sign of M_{mn} determines whether the frequency of the unforced motion increases or decreases with amplitude and thus can be shown, for given aspect ratio and cell size, to depend on the depth. To determine the critical depth at which the quantity M_{mn} changes sign, we have

$$\frac{n^2}{b^2}\ M_{mn}\left(\frac{mb}{na},1,\frac{n}{b}\ h^*\right)=0$$

We note that the expression represents the "back bone" of the forced motion and the forced motion also changes qualitatively depending on whether the depth of the fluid is less than or greater than h^*. Response curves are given in Fig. 2 when $M<0$.

For two-dimensional gravity waves, M_{mn} has a relatively simpler expression

$$M_{on}(o,n,\lambda)=-\frac{K_{on}^2(2\alpha^3+3\alpha^2+12\alpha-9)}{128\alpha^2},\qquad(30)$$

where $\alpha=(\tanh\ Kon\lambda)^2$. The critical depth for two-dimensional gravity waves can be obtained simply by equating the numerator of equation (30) to zero. The critical dimensionless depth can be shown to be h^* and $nh^*/b\approx0.3368$.

The above effect of the depth of the fluid has been discussed by Takjbakhsh and Keller (1960) and Verma and Keller (1962). Our results agree with their results, though not as well as with the experiments discussed in Part II.

Stability and Local Bifurcation Analysis

In order to study the stability of the zero solution, we introduce Cartesian variables as follows. Let $x=a\ \cos\ \theta$ and $y=-a\ \sin\ \theta$, we then have

$$\dot{x}=-dx-(\Delta+\sigma)y+\epsilon y\left[M(x^2+y^2)+\left(\frac{3}{2}\ \Delta^2-\sigma^2\right)\right],$$

$$\dot{y}=-dy-(\Delta-\sigma)x-\epsilon x\left[M(x^2+y^2)+\left(\frac{3}{2}\ \Delta^2-\sigma^2\right)\right],\qquad(31)$$

The stability of the zero solution is determined by the eigenvalues of the linearized problem:

$$\lambda_{1,2}=-d\pm\left\{\Delta^2-\left[\sigma-\epsilon\left(\sigma^2-\frac{3}{2}\ \Delta^2\right)\right]^2\right\}^{1/2}.\qquad(32)$$

We first note that for $d\neq0$ the system cannot have pure imaginary eigenvalues, and thus Hopf bifurcations are impossible. All co-dimension one bifurcations are, therefore, due to a zero eigenvalue which occur when $\sigma=\sigma_0$ where σ_0 satisifes

$$\Delta^2-\left[\sigma+\epsilon\left(\sigma_0^2-\frac{3}{2}\right)\right]^2=d^2.\qquad(33)$$

We note that expression (33) is the same as condition $s_1=0$. The zero solution of (31) is stable when $\Delta^2-[\sigma+\epsilon(\sigma^2-3/2\Delta^2)]^2<d^2$ and a pitch-fork bifurcation occurs as the parameters cross S_1.

To study the stability of the nonzero solutions, we construct the Jacobian matrix of equations (25) and eliminate the variable θ to obtain the characteristic equation

$$\lambda^2+2d\lambda+4\epsilon\mathrm{Ma}^2\left[\sigma+\epsilon\left(\mathrm{Ma}^2+\frac{3}{2}\ \Delta^2-\sigma^2\right)\right]=0.\qquad(34)$$

We again note that for $d\neq0$, λ cannot be pure imaginary and thus Hopf bifurcations are impossible. Inserting (28) into the square brackets gives

$$\lambda^2+2d\lambda\pm4\epsilon\mathrm{Ma}^2(\Delta^2-d^2)^{1/2}=0,\qquad(35)$$

where the + and − signs of the third term correspond to the upper branch and the lower branch, respectively. Therefore, for given $0<d<\Delta$, the upper branch is always stable and the lower branch is always unstable. For $d=\Delta$ we observe from equation (35) that two *different* solutions have zero eigenvalues and as d decreases from $d>\Delta$ to $d<\Delta$ across curve S_2, it can be shown that two simultaneous saddle-node bifurcations occur. Thus the system goes from one stable periodic zero solution to five periodic solutions including the zero solution and two unstable solutions.

Transient Behavior

We refer back to equations (9), (10), (19), (20), and (25). In physical terms, the amplitude of the small modal motions corresponding to spatial modes that are not in resonance decay to zero, and thus the motion decays to that of the dominant mode. On the planar subspace corresponding to the (m, n)th mode, we have the motion described by equations (25). If the damping is zero, equations (25) have the following integral

$$a^2\ \cos\ 2\theta-\frac{1}{\Delta}\ a^2\left[\sigma-\epsilon\left(\frac{3}{2}\ \Delta^2-\sigma^2\right)\right]+\frac{\epsilon}{2\Delta}\ \mathrm{Ma}^4=\text{const.}\qquad(36)$$

Expression (36) gives closed level curves in the plane in the polar coordinates $(a_{mn},\ \theta_{mn})$. For the specific case when $\lambda_1=7/9$, $\lambda=1/2$, $M_{mn}=-0.8863$, $\Delta=0.09$ cm, $\epsilon=0.00394$, and $\epsilon\sigma=-0.02$, equations (25) has five fixed points, three

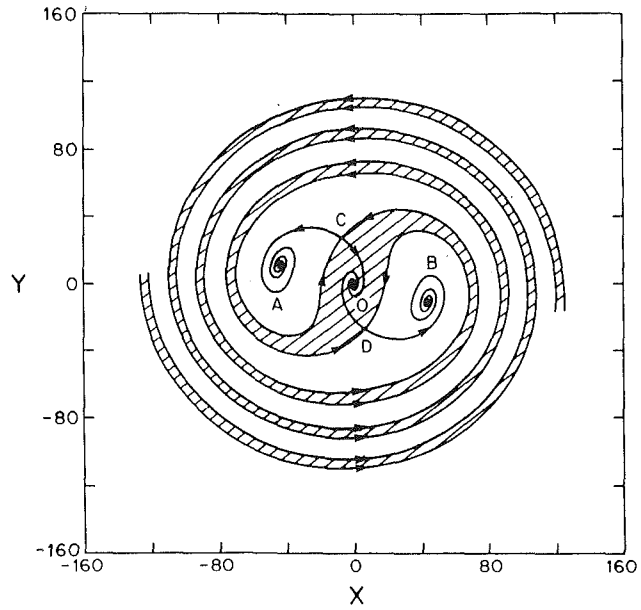


Fig. 3 Basins of attraction of the three stable periodic solutions (see text)

centers, and two saddles with the origin a center and closed level curves around the centers; homoclinic orbits connecting the saddles are obtained from the level curves of equation (36) as discussed by Miles (1984). With dissipation the three stable centers of equations (25) become foci and the homoclinic loops connecting saddle points break. Now, it can be shown, as by Sethna and Moran (1968), that any solution of the equation that approaches a stable fixed point is a good approximation of the amplitude and phase. Basing our conclusion on this result, we see that the stable and unstable mainfolds of the two saddle points separate the space into basins of attraction of the stable periodic solutions represented by the fixed points. Figure 3 represents the solutions of equations (25) in the damped case for the same values of the constants as mentioned above and the damping coefficient calculated from equations (15) for $C=2$ which gives the value of $\epsilon d_{mn}=0.0038$. This figure was obtained numerically by starting the solutions on the stable and unstable linear subspaces of points C and D for negative and positive times, respectively. It is important to note the complex behavior of the transient motion that can occur depending on the initial conditions. Initial conditions in the basin of attraction of 0 leads to an undisturbed liquid surface, while initial conditions in the basin of attraction of A or B lead to steady periodic motions that are apart in phase by π. We note that it is possible to have undisturbed motion as $t\to\infty$ even when the initial conditions are *not* small if they have the appropriate amplitude and phases. If the detuning is such that the origin is unstable, the situation is simpler. There are only two stable fixed points and the origin now is a saddle and its stable and unstable manifolds determine the basin of attraction of the stable fixed points. These solutions again represent solutions that are π apart in phase. Referring back to bifurcation diagram Fig. 1, the pitch-fork bifurcations on the semicircle $d^2+\sigma^2=\Delta^2$, $d>0$ are bifurcations involving 0, A, and B while the bifurcation on the line $d=\Delta$ are saddle node bifurcations when A and C and B and D coalesce in pairs simultaneously.

References

Bajaj, A. K., and Sethna, P. R., 1980, "Hopf Bifurcation Phenomena in Tubes Carrying a Fluid," *SIAM Journal of Applied Mathematics*, Vol. 39, No. 2, pp. 213-230.

Benjamin, T. B., and Ursell, F., 1954, "The Stability of the Plane Free Surface of a Liquid in Vertical Periodic Motion," *Proceedings of the Royal Society of London*, A 225, pp. 505-517.

Chu, W. H., 1968, "Subharmonic Oscillations in Arbitrary Axisymmetric Tank Resulting from Axial Excitation," ASME JOURNAL OF APPLIED MECHANICS, Series E, Vol. 35, pp. 148-150.

Dodge, F. T., Kana, D. D., and Abramson, H. N., 1965, "Liquid Surface Oscillations in Longitudinally Excited Rigid Cylindrical Containers," *AIAA Journal*, Vol. 3, pp. 1052-1059.

Faraday, M., 1831, "On the Forms and States Assumed by Fluids in Contact with Vibrating Elastic Surfaces," *Philosophical Transactions of the Royal Society of London*, Vol. 121, pp. 319-346.

Gu, X. M., 1986, "Nonlinear Surface Waves of a Fluid in Rectangular Containers Subjected to Vertical Periodic Excitations," Ph.D. Thesis, University of Minnesota, Minneapolis, MN.

Hale, J. K., 1969, *Ordinary Differential Equations*, Wiley-Interscience, New York.

Holmes, P., 1986, "Chaotic Motions in a Weakly Nonlinear Model for Surface Waves," *Journal of Fluid Mechanics*, Vol. 162, pp. 365-388.

Miles, J. W., 1984, "Nonlinear Resonance," *Journal of Fluid Mechanics*, Vol. 146, pp. 285-302.

Miles, J. W., 1967, "Surface-Wave Damping in Closed Basins," *Proceedings of the Royal Society of London*, A 297, pp. 459-475.

Sethna, P. R., and Moran, T. J., 1968, "Some Nonlocal Results for Weakly Nonlinear Dynamical Systems," *Quarterly Journal of Applied Mathematics*, XXVI, No. 2, pp. 175-185.

Skalak, R., and Yarymovych, M. I., 1962, "Forced Large-Amplitude Surface Waves," *Fourth U. S. National Congress of Applied Mechanics*, ASME, Rosenberg, R. M., ed., ASME United Engineering Center, New York, pp. 1411-1418.

Tadjbakhsh, I., and Keller, J. B., 1960, "Standing Surface Waves of Finite Amplitude," *Journal of Fluid Mechanics*, Vol. 8, pp. 442-451.

Verma, G. H., and Keller, J. B., 1963, "Three-Dimensional Standing Waves of Finite Amplitude," *Physics of Fluids*, Vol. 5, pp. 52-56.

Whitman, G. B., 1973, Linear and Nonlinear Waves, Wiley, New York.

APPENDIX

Definitions and Calculations of Coupling Coefficients

It was shown ealier that the eigenfunctions evaluated at the free surface are a complete set:

$$S_{ij}=\cos\frac{i\pi x}{\lambda_1}\cos j\pi y,\quad i,j=0,\ 1,\ 2,\ \ldots, \tag{A.1}$$

The inner product of any two elements of (A.2) is defined as

$$<S_{ij},S_{rs}>\overset{\text{def}}{=}\frac{4}{\lambda_1}2^{-(\delta_{or}+\delta_{os})}\int_0^{\lambda_1}\int_0^1 S_{ij}S_{rs}\,dx\,dy, \tag{A.2}$$

where δ_{or} and δ_{os} are the Kronecker deltas, and a direct result from (A.2) is

$$<S_{ij},S_{rs}>=\delta_{ir}\delta_{js}, \tag{A.3}$$

where δ_{ir} and δ_{js} are the Kronecker deltas.

Listed below are the definitions of the coupling coefficients:

$$G^{(1)mn}_{ijrs}=<\frac{\partial S_{ij}}{\partial x}\frac{\partial S_{rs}}{\partial x},S_{mn}>+<\frac{\partial S_{ij}}{\partial y}\frac{\partial S_{rs}}{\partial y},S_{mn}>-K^2_{rs}<S_{ij}S_{rs},S_{mn}>,$$

$$G^{(2)mn}_{ijrs}=<\frac{\partial S_{ij}}{\partial x}\frac{\partial S_{rs}}{\partial x},S_{mn}>+<\frac{\partial S_{ij}}{\partial y}\frac{\partial S_{rs}}{\partial y},S_{mn}>+\alpha_{ij}\alpha_{rs}<S_{ij}S_{rs},S_{mn}>,$$

$$G^{(3)mn}_{ijrs}=\alpha_{rs}<S_{ij}S_{rs},S_{mn}>,$$

$$H^{(1)mn}_{ijrskl}=\frac{1}{2}K^2_{kl}\alpha_{kl}<S_{ij}S_{rs}S_{kl},S_{mn}>-\alpha_{kl}\left[<S_{ij}\frac{\partial S_{rs}}{\partial x}\frac{\partial S_{kl}}{\partial x},S_{mn}>+<S_{ij}\frac{\partial S_{rs}}{\partial y}\frac{\partial S_{kl}}{\partial y},S_{mn}>\right],$$

$$H^{(2)mn}_{ijrskl} = \alpha_{rs}\left[< \frac{\partial S_{ij}}{\partial x} \frac{\partial S_{rs}}{\partial x} S_{kl}, S_{mn} > \right.$$

$$\left. + < \frac{\partial S_{ij}}{\partial y} \frac{\partial S_{rs}}{\partial y} S_{kl}, S_{mn} > \right]$$

$$+ K^2_{rs}\alpha_{ij} < S_{ij}S_{rs}S_{kl}, S_{mn} >,$$

$$H^{(3)mn}_{ijrskl} = \frac{1}{2} K^2_{kl} < S_{ij}S_{rs}S_{kl}, S_{mn} >,$$

$$H^{(4)mn}_{ijrskl} = \frac{\tilde{r}}{2}\left[< \frac{\partial^2 S_{ij}}{\partial x^2} \frac{\partial S_{rs}}{\partial y} \frac{\partial S_{kl}}{\partial y}, S_{mn} > \right.$$

$$+ < \frac{\partial^2 S_{ij}}{\partial y^2} \frac{\partial S_{rs}}{\partial x} \frac{\partial S_{kl}}{\partial x}, S_{mn} >$$

$$+ 3 < \frac{\partial^2 S_{ij}}{\partial x^2} \frac{\partial S_{rs}}{\partial x} \frac{\partial S_{kl}}{\partial x}, S_{mn} >$$

$$+ 3 < \frac{\partial^2 S_{ij}}{\partial y^2} \frac{\partial S_{rs}}{\partial y} \frac{\partial S_{kl}}{\partial y}, S_{mn} >$$

$$\left. + 4 < \frac{\partial S_{ij}}{\partial x} \frac{\partial S_{rs}}{\partial y} \frac{\partial^2 S_{kl}}{\partial x \partial y}, S_{mn} > \right].$$

In the above expressions, α_{ij} is defined by

$$\alpha_{ij} = K_{ij} \tanh K_{ij}\lambda.$$

On Three-Dimensional Nonlinear Subharmonic Resonant Surface Waves in a Fluid: Part II—Experiment

J. C. Virnig

A. S. Berman

P. R. Sethna
Mem. ASME

Department of Aerospace Engineering and Mechanics,
University of Minnesota,
Minneapolis, MN 55455

A series of experiments have been performed to investigate wave motion in a rectangular container partially filled with water and shaken in a direction normal to the horizontal free water surface. For a (1-1) wave mode and for several water depths, the dependence of wave amplitude on both excitation frequency and excitation amplitude was found to be in good agreement with the predictions of the nonlinear theory in Part I. In general, at low wave amplitudes better agreement was observed when the surface tension was lowered by the addition of a surfactant. The experiments imply the existence of a critical water depth below which the frequency dependence of wave amplitude is qualitatively different than for depths greater than this critical.

Introduction

In Part II of this study we present results from experiments. The study, as in Part I, is restricted to the case of excitation of a single spatial mode of the surface waves by vertical sinusoidal motions of a container of rectangular cross section.

The study has two major objectives. The first is the comparison of experimental observations with the theoretical predictions of the behavior of three-dimensional waves. The second is the investigation of the predictions regarding the effect of the depth of the fluid on the dependence of wave amplitude on excitation frequency.

Among the previous experimental work on the response of such systems, a limited study of two-dimensional waves in a container of rectangular cross section by Skalak and Yarymovych (1962), and a study by Dodge et al. (1965), presents results for the case of a cylindrical container. Critical depth experiments have been done by Fultz (1962) and Hayama et al. (1983). These studies, however, are for waves produced by excitations in a direction parallel to the free surface.

In general, our experimental results verify the predictions of the theory of Part I regarding the behavior of the wave motions when the wave amplitudes are large and nonlinear effects are dominant. These experimental results include the dependence of wave amplitude on excitation frequency and excitation amplitude, as well as the dependence of the wave motion on water depth. We also present phenomena not predicted by the theory for experiments in which the amplitude of the motions is small and nonlinear effects are negligible. These phenomena are shown to be associated with surface tension.

Experimental Equipment and Measurement Techniques

The rectangular container used in the study is constructed of plexiglass. Its interior is 17.78 cm wide, 30.48 cm long, and 25.40 cm high. An additional partition can be fitted into the container so that the 30.48 cm dimension can be adjusted. This partition is sealable so that no leakage past it can occur. The partition is held securely by the downward pressure supplied by two braces which are attached to the permanent walls of the container. The container is firmly attached to one of two machines used to produce the sinusoidal excitation. Special care is taken to assure the container is level, and the center of mass of the container assembly coincides with the center of the shaking mechanisms. It is estimated that the container was never more than 0.05 deg out of level.

The wave amplitude of the water surface is measured with the aid of a depth micrometer. The micrometer is held in place by a device which is firmly attached to the two braces on top of the container. The micrometer moves with the motion of the container. Two measurements with the micrometer are necessary to determine the wave amplitude. With no excitation present, the micrometer rod is lowered until contact is made with the surface of the water. This is then repeated during steady excitation of the container with the wave form present. The difference between these micrometer readings represents the amplitude of the wave. Because the micrometer has a range of 2.54 cm, this is the largest wave amplitude which can

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, August 5, 1986; final revision July 25, 1987.

be measured. The smallest division on the micrometer is 0.0254 mm, and this imposes a limit on the precision of the measurement of wave amplitude.

To determine exactly when the micrometer rod touches the surface of the water, a series circuit containing the micrometer, a battery, and a oscilloscope are set up. The circuit is made complete through the liquid in the container by a very small diameter metal rod which runs down one corner of the container. When the micrometer rod makes contact with the water in the container, the circuit is closed and the oscilloscope clearly indicates this moment of closure. For experiments in which the wave amplitude does not vary too much with frequency, amplitude measurements are reproducible to about 0.025 mm. In some experiments with strong dependence of amplitude on frequency, successive amplitude measurement might vary by as much as 0.250 mm. In the latter case an average of three or four amplitude measurements is used.

Either an MB electromagnetic shaker table (model T-19M) or an MTS machine (model 810) is used to provide the vertical sinusoidal excitation. The advantage of using the shaker table is that it is capable of producing relatively large excitation amplitudes (upwards of 1.3 cm peak to peak for frequencies below seven Hertz). However, because the excitation amplitude is sensitive to changes in the excitation frequency, the shaker is used only for experiments where excitation frequency is held constant. The MTS machine has a built in amplitude controller and a digital amplitude display and so it is much easier to perform constant excitation amplitude experiments. Because the driving piston of the MTS machine is hydraulically driven, excitation amplitude is limited at some of the frequencies of interest.

The frequency and amplitude of both machines are controlled by an HP 3310B function generator driving a feedback power amplifier. An HP 5315A universal counter is used to display the frequency. After approximately one half hour for warm-up, the system frequency stability is excellent and in accord with specifications for the function generator. These specifications are: frequency drift of 0.1 percent of the setting for any ten minute interval and 0.5 percent of the setting for an eight hour interval.

The amplitude of the sinusoidal excitation for the MTS machine is obtained directly from the built-in digital display which can be read to 0.1 mm. The shaker has no such display, so excitation amplitude is determined using an accelerometer, as will be discussed later.

The harmonic content of the motion of the driving platforms of both machines is determined using an accelerometer, a Tektronix Type 3C66 carrier amplifier with Type 561A oscilloscope, and an HP 3582A spectrum analyzer. The accelerometer is calibrated using the earth's gravitational field. The accelerometer is clamped to the side of the water container, and its output is displayed on the oscilloscope after passing through a strain gage amplifier. The signal from the amplifier is also supplied to the spectrum analyzer, from which the harmonics present in the signal are determined as a percentage of the fundamental. At operating frequencies of three, five, and seven Hertz and peak to peak excitation amplitudes below about five millimeters, the vertical acceleration of the electromagnetic shaker has no harmonics which exceed four percent of the fundamental, and most harmonics are less than two percent of the fundamental. As the peak to peak amplitude is raised beyond five millimeters, the harmonic content becomes increasingly more prominent, reaching levels as high as thirteen percent of the fundamental. In experiments with the electromagnetic shaker, excitation amplitudes are always below 5 mm.

The MTS machine was checked for harmonic content at frequencies of three and five Hertz. When operating at peak to peak amplitudes above 50 percent of maximum (5.8 mm is the maximum amplitude at three Hertz and 3.4 mm is the maximum at five Hertz), the first and second harmonics are not greater than four percent of the fundamental, whereas the third harmonic sometimes reaches eight percent. As the excitation amplitude is reduced to approximately 25 percent of maximum, the first and second harmonics reach levels as high as eleven percent of the fundamental. All quantitative experiments using the MTS driver were performed using excitation amplitudes greater than about 50 percent of maximum.

For both machines, the greatest amount of acceleration in the horizontal direction occurs at the fundamental driving frequency, yet this amount never exceeds about one percent of the vertical acceleration and hence has little effect on the experimental results.

The peak to peak amplitude of the excitation motion of the electromagnetic shaker is determined using the same equipment and setup described above. The maximum acceleration associated with the excitation motion is read in digital form from the spectrum analyzer. The spectrum analyzer output is divided by the square of the excitation frequency and with the accelerometer calibration factor gives the excitation amplitude. When this method of obtaining excitation amplitude is compared with the amplitude obtained optically with the use of a cathetometer and stroboscopic lighting, agreement is within three percent. The cathetometer check was also made on the amplitude displayed by the MTS machine, and the same order of agreement was found.

Experimental Procedure

A typical experiment proceeds as follows. The container partition is adjusted for the desired container size and the container is filled with tap water (in most cases Kodak Photo-Flo is used to reduce surface tension) at near room temperature. The depth is measured along an outer wall of the container using a rule subdivided in increments of 0.254 mm. It is felt this measurement is accurate to within one small subdivision of the rule. The position of the micrometer as measured from two edges of the container is then determined. The same rule is used and the same level of accuracy is achieved. The micrometer is positioned far enough from the container walls to avoid capillary effects. Thus the wave amplitude is measured at a point which is not necessarily the maximum wave amplitude.

In all experiments done using the electromagnetic shaker, the excitation frequency is held constant and the excitation amplitude is varied. This is necessary because the amplitude of the shaker motion depends strongly on the forcing frequency. Due to the way in which the amplitude is measured (the averaging performed by the spectrum analyzer produces a time delay in amplitude measurement), it is inconvenient to duplicate the same shaker amplitude for different driving frequencies.

The system is excited at a frequency near double the natural frequency of the chosen mode and the induced wave is brought to nearly breaking amplitude by increasing the excitation amplitude. The excitation amplitude is then decreased in a stepwise manner, and the wave amplitudes at the micrometer location are recorded at each excitation amplitude. This procedure is carried out until the wave suddenly decays or steadily decreases to zero amplitude. The wave amplitude measurements at each excitation amplitude are made only after steady motion of the wave is achieved, as determined by successive wave amplitude measurements. Once steady motion is present, a final fine adjustment to the driving frequency is made to insure that the wave amplitude measurement can be made at a frequency known to about 0.001 Hz. For each mode this experiment is repeated for several values of frequency. Occasionally an experiment is started at small shaker amplitude which is then increased. This provides information

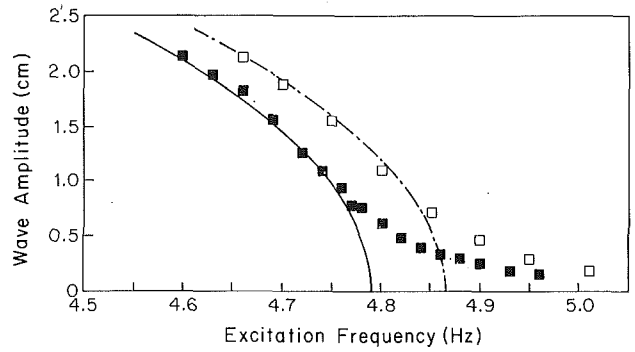


Fig. 1 Wave amplitude for a (1—1) mode versus excitation frequency—without additive. Excitation amplitudes: 0.90 mm (— predicted stable branch, ■ observed); 1.55 mm (-·- predicted stable branch, □ observed). Container dimensions: 17.78 cm by 22.86 cm, depth = 11.43 cm.

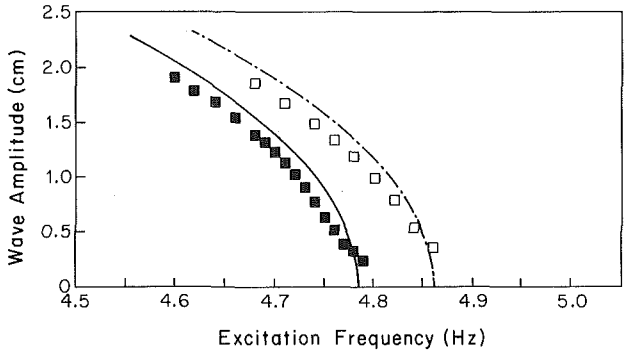


Fig. 2 Wave amplitude for a (1—1) mode versus excitation frequency—with additive. Excitation amplitudes: 0.90 mm (— predicted stable branch, ■ observed); 1.55 mm (-·- predicted stable branch, □ observed). Container dimensions: 17.78 cm by 22.86 cm, depth = 11.43 cm.

about any hysteresis effects. The results of these experiments will be described later.

The MTS machine provides a better means of maintaining a constant excitation amplitude when the frequency is varied. Also, because the MTS machine has a digital amplitude readout and control, it is easy to correct for slight amplitude drifts. Once a given free surface mode is chosen, an excitation amplitude is selected which drives the desired mode. The excitation frequency is varied in the neighborhood of double the natural frequency of the mode, and the steady wave amplitudes at the micrometer location are recorded at several frequencies. This procedure is repeated at several excitation amplitudes for each of the modes investigated.

Results and Discussion

Three-Dimensional Wave Mode Amplitude. Experiments to investigate the frequency dependence of the wave amplitude of three-dimensional wave modes induced by vertical, sinusoidal excitation of the container were performed. These experiments are qualitatively similar to those performed by Dodge et al. (1965), except that the basic container geometry is different. The observed response of the system is compared to the predictions of the second order analysis given by equation (43) in Part I. The theoretical results are calculated with the appropriate value of the surface tension constant for each experiment and the disssipation constant is computed from equation (24) in Part I with $C = 1$ throughout. Generally, the choice of C between zero and two does not significantly alter the results with one exception, which will be noted later. Equation (43) predicts the wave amplitude for a given mode as a function of excitation frequency and excitation amplitude. All the experiments are done for $(1-1)$ and $(0-1)$ spatial modes. These modes are chosen because at these mode numbers it is easiest to satisfy the dual requirements of the lack of space and dynamic coupling as discussed in Part I. An additional reason for the choice of these modes along with the choice of the size of the container is that we can use the shaking devices at frequencies at which they have desirable characteristics (high excitation amplitude and low harmonic content.) Finally, the theory predicts the largest critical depth for the $(0-1)$ mode. The experiments were performed for this mode so as to avoid small water depths, where viscous boundary layer effects are important.

On Figure 1 are shown theory and experimental results for $(1-1)$ mode oscillations. In this case, the variable container dimension is adjusted so as to have a rectangular cross section of sides 17.78 cm and 22.86 cm, and the container is filled to a water depth of 11.43 cm. The micrometer is positioned 3.68 cm from the container edge associated with the 22.86 cm dimension and 4.19 cm from the edge associated with the 17.78 cm dimension. Excitation at constant amplitude is provided by the MTS machine. The observed wave amplitudes and the theoretical results are shown in Fig. 1 for two excitation amplitudes as a function of excitation frequency. Note that only the stable branches of the theory curves are shown. At relatively large wave amplitudes, agreement between theory and experiment, when viewed along lines of constant wave amplitude, is within 0.4 percent for an excitation amplitude of 0.9 mm and within 0.2 percent for an excitation amplitude of 1.55 mm. At low wave amplitudes, however, a "tailing" effect is observed which is not predicted by the theory. The relation between wave amplitude and frequency is not subject to hysteresis effects. This is determined by data taken with both increasing and decreasing frequency at constant excitation amplitude.

In order to study the cause of the tailing effect and to determine if it is due to surface tension, Kodak Photo-Flo 200 solution is added to the water at a concentration of one part in 1000 by volume to reduce the surface tension of the water to approximately 30 dynes/cm (see Fultz, 1962; pure water has a surface tension of 72 dynes/cm). The effects that this additive has on the experimental results for the same operating conditions as in Fig. 1 are given in Fig. 2. In the figure the lowest measured wave amplitude at an excitation amplitude of 0.9 mm is 0.23 cm at an excitation frequency of 4.790 Hz. When the excitation frequency is raised to 4.800 Hz, the wave amplitude decays to zero. The lowest measured wave amplitude at an excitation amplitude of 1.55 mm is 0.35 cm at an excitation frequency of 4.860 Hz. When the frequency is raised to 4.870 Hz, the wave amplitude decays to zero. The "tailing" has been virtually removed with the addition of the surfactant. However, the same wave amplitudes found with no additive occur at a frequency of about 0.025 Hz less than those found with additive present. The maximum difference between the predicted and observed system response increases to roughly 0.5 percent for both excitation amplitudes. Although the analysis leading to the predicted amplitude includes an estimate of the effects of surface tension (the natural frequency of the mode is reduced when surface tension is reduced), the predicted effect is not as great as that observed in the experiment.

Experiments are also performed in which the excitation frequency is held constant and the excitation amplitude is varied. The electromagnetic shaker is used for these experiments. Figure 3 shows some data from these experiments. The container dimensions and water depth and the micrometer position are the same as described above. The observed and predicted response of the system, again for a $(1-1)$ mode, at four different excitation frequencies are shown in Fig. 3. For

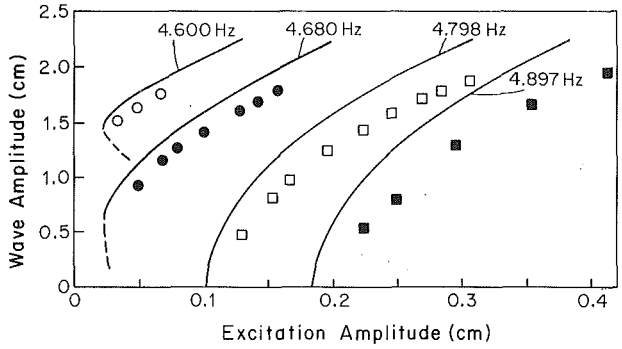


Fig. 3 Wave amplitude for a (1—1) mode versus excitation amplitude—with additive. Wave amplitudes: (—) predicted: (○) 4.600 Hz, (●) 4.680 Hz, (□) 4.798 Hz, (■) 4.897 Hz observed. Container dimensions: 17.78 cm by 22.86 cm, depth = 11.43 cm.

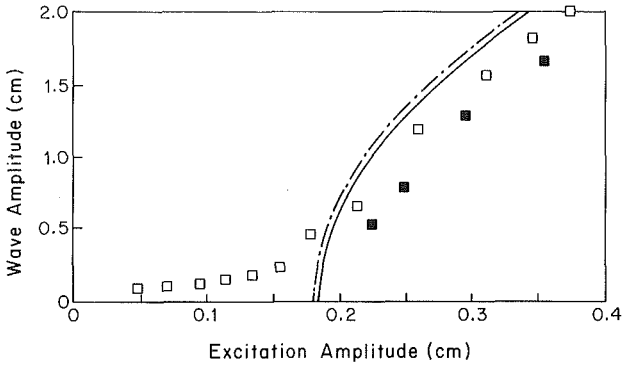


Fig. 4 Wave amplitude for a (1—1) mode versus excitation amplitude—with and without additive. Wave amplitudes with additive (— predicted stable branch, ■ observed), and without additive (- - - predicted stable branch, □ observed). Excitation frequency = 4.897 Hz. Container dimensions: 17.78 cm by 22.86 cm, depth = 11.43 cm.

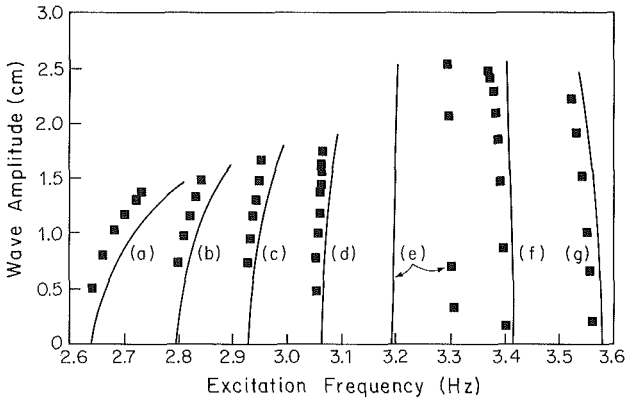


Fig. 5 Wave amplitude for a (0—1) mode versus excitation frequency for several fluid depths—with additive. Wave amplitudes (- - - predicted stable branch, ■ observed) for water depths and excitation amplitudes of (*a*) 4.27 cm and 2.5 mm (*b*) 4.98 cm and 2.5 mm, (*c*) 5.69 cm and 2.5 mm, (*d*) 6.60 cm and 2.5 mm, (*e*) 7.47 cm and 1.75 mm, (*f*) 8.18 cm and 2.3 mm, (*g*) 10.16 cm and 2.3 mm. Container dimensions: 17.78 cm by 22.86 cm.

each frequency the theory predicts larger wave amplitudes than is found experimentally. Slightly better agreement between theory and experiment is observed at frequencies below twice the natural frequency of the system (4.692 Hz). Although the difference between the predicted and observed wave amplitudes is as high as 25 percent (this difference occurs at high wave amplitude and excitation frequency), this represents only a 2.1 percent difference in the predicted and observed excitation frequencies, i.e., a 2.1 percent shift in frequency would bring theory and experiment into agreement.

For excitation frequencies less than twice the natural frequency, the theory predicts a saddle node bifurcation to occur at a critical excitation amplitude. At this amplitude, the wave amplitude versus excitation amplitude curve has a vertical tangent, and a stable and an unstable branch are predicted. The unstable branch is indicated by a dashed line in Fig. 3. Physically one expects a "jump" at this point, i.e., the wave amplitude should precipitously decrease from a finite amplitude to zero as the excitation amplitude is reduced by a small amount. What is remarkable is that the location of this jump is predicted to depend solely on the dissipation which depends on the physical parameters and the frequency of motion. In the experiments the frequencies of excitation are relatively close so that the jumps are predicted to occur at approximately the same excitation amplitude. This prediction is qualitatively verified as indicated in Fig. 3.

In the laboratory, a rapid decay of the excited wave to zero amplitude is observed when the amplitude is lowered from a value above to a value below the critical excitation amplitude. At an excitation frequency of 4.600 Hz, and at an excitation amplitude of 0.35 mm, a finite amplitude steady wave is observed. A decrease in excitation amplitude to 0.26 mm results in a rapid decay to zero wave amplitude. Although the theory curves in general are not significantly affected by the value of C, this is not true in the case of critical excitation amplitude where the excitation amplitudes are very small. For the case of an excitation frequency of 4.600 Hz, the predicted critical excitation amplitude ranges from 0.13 mm to 0.33 mm for $C = 0$ and $C = 2$, respectively. For the theory curve shown in Fig. 1, $C = 1$ was used and the resulting critical excitation amplitude is 0.22 mm. A choice of a different value of C would have given better agreement. At an excitation frequency of 4.680 Hz steady finite amplitude waves are observed for excitation amplitudes from about two millimeters to 0.49 mm. At 0.30 mm the wave amplitude slowly decays to zero amplitude. Again the critical amplitude is very sensitive to the choice of C. For $C = 1$ the predicted critical amplitude is 0.223 mm for the excitation frequency of 4.680 Hz.

To confirm the striking effect of surface tension, an experiment with and without the additive is repeated, this time with the shaker frequency fixed and shaker amplitude as the independent parameter. The experiment is done with an excitation frequency of 4.897 Hz with and without the additive. Again it is found that the addition of Photo-Flo results in a reduction of "tailing" (see Fig. 4) and a decrease in wave amplitude for a given excitation amplitude. In the figure the lowest measured wave amplitude with the additive is 0.53 cm at an excitation amplitude of about 2.2 mm. When the excitation amplitude is lowered to 2.0 mm, the wave amplitude decays to zero. These phenomena are similar to those discussed previously in connection with Figs. 1 and 2. There were no hysteresis effects observed in the portion of the response curves shown in Fig. 4.

Critical Depth. The analysis in Part I predicts, within the limitations of a second order theory, that the dependence of wave amplitude on excitation frequency for a given container size is qualitatively different for water depths above and below a certain critical depth. When the water level in the container is above this depth, the amplitude decreases with increasing frequency. When the water level is below this depth, the theory predicts increasing amplitude with increasing frequency. The results of experiments to examine these predictions are described below.

For a container of the same dimensions as given previously and for a (0 – 1) mode, the critical depth computed from equation (45) in Part I is 7.70 cm. Figure 5 shows the predicted stable branches of the frequency response curve and also gives

the observed wave amplitudes for seven fluid depths as a function of excitation frequency at a position on the wave 3.68 cm from the container edge (corresponding to the 22.86 cm dimension). It should be noted that the excitation amplitude is not the same at every fluid depth and that the data points shown in Fig. 5 for the water depth of 7.47 cm are actually an average of at least three points taken at the same frequency (this was necessary due to the rapid variation of wave amplitude with frequency at this depth).

At the two largest depths tested (curves *f* and *g*), observations indicate that increasing the frequency decreases the wave amplitude, while at the four smallest depths (curves *a*, *b*, *c*, and *d*) increasing the frequency increases the wave amplitude. This behavior is in accord with the predictions of the analysis. At the depth of 7.47 cm (curve *e*), experimental observations indicate that increasing the frequency decreases the wave amplitude. Since the theory predicts this depth to be below the critical depth, the opposite behavior is predicted, resulting in bad agreement between the theory and the experiment at this depth. Generally, however, agreement is good and the difference between the theory and the experiment when viewed at constant wave amplitude on Fig. 5 is smaller at larger depths than at smaller depths (there is a maximum difference of 1.0 percent above the predicted critical depth as compared with a maximum difference of 1.8 percent below).

These experiments suggest that the critical depth for this container size is between 6.60 cm and 7.47 cm, which is, respectively, fourteen and three percent below the predicted critical depth of 7.70 cm. Attempts were made to obtain frequency response curves at depths of 6.90 cm and 7.20 cm. However, for these depths a steady wave form with amplitude below 2.54 cm could not be maintained. Steady wave forms with amplitudes between 2.54 cm and 4.5 cm were obtainable, and qualitative measurements made at these amplitudes indicate that increasing the frequency decreases the amplitude of the wave at both of these fluid depths. With such large wave amplitudes the phenomena observed may well be determined by higher order effects not covered in the second order theory presented in Part I.

References

Dodge, F. T., Kana, D. D., and Abramson, H., 1965, "Liquid Surface Oscillations in Longitudinally Excited Rigid Cylindrical Containers," *AIAA Journal*, Vol. 3, No. 4, pp. 685–695.

Fultz, D., 1962, "An Experimental Note on Finite-Amplitude Standing Gravity Waves," *Journal of Fluid Mechanics*, Vol. 13, No. 21, pp. 193–212.

Hayama, S., Aruga, K., and Watanabe, T., 1983, "Nonlinear Response of Sloshing in Rectangular Tanks," *Bulletin of JSME*, Vol. 26, No. 219, pp. 1641–1648.

Skalak, R., and Yarymovych, M. I., 1962, "Forced Large Amplitude Surface Waves," *Proceedings of the Fourth US National Congress on Applied Mechanics*, Rosenberg, R. M., ed., ASME United Engineering Center, New York, pp. 1411–1418.

Some Aspects of the Behavior of High Order Mode Internal Gravity Waves on Interfaces With Finite Thickness

M. R. Muller
Department of Mechanical and Aerospace Engineering,
Rutgers University,
New Brunswick, NJ 08903

P. C. Shang
David W. Taylor Naval Ship Research and Development Center,
Bethesda, MD 20084

The role of interface thickness on the propagation of high order internal wave modes is examined using simple analytical models for an unmixed, thick interface and one that has been vigorously mixed. It is shown that for the unmixed case, a local maximum in wave speed occurs as the interface grows from zero thickness, but that only for mode-one waves does the maximum point occur within areas of practical significance. When these models are applied to three wave resonant interactions, it is found that vigorous mixing will augment energy transfer rates when the interface thickness surpasses a certain value.

1 Introduction

Internal waves within fluids have historically been described as propagating either in continuously stratified fluids or on interfaces. In the former case waves can propagate in both the vertical and horizontal while in the latter case waves are constrained to propagate only in the horizontal. One can look at these two descriptions as limits of the nondimensional parameter $k\epsilon$, where k is the magnitude of the wave number vector of waves of interest and ϵ is a measure of the interface thickness. When $\epsilon \to 0$, one recovers purely interfacial waves which have received a great deal of attention in the literature. We are concerned in this paper with the intermediate range, $(0.1 < k\epsilon < \pi)$ where waves are still constrained to propagate in the horizontal, but the thickness of the interface affects the propagation of the waves. We consider both the impact of interfacial thickness on wave propagation for a single, thick interface and that of a double interface (which is viewed as resulting from vigorous mixing at the center of a single, thick interface).

Internal wave motion can be described by linearizing the equations of motion, assuming wave-like solutions whose amplitude varies only as a function of vertical position, and solving the resulting ordinary differential equation by any of several means (see Thorpe, 1978, or Shang, 1984, for details). The equation governing the amplitude distribution of the wave is called the Taylor-Goldstein equation and can be written as:

$$\left\{\frac{d^2}{dz^2} - \left(k^2 - \frac{N^2(z)}{(U-c)^2} + \frac{U''(z)}{(U-c)}\right)\right\}\psi = 0 \tag{1}$$

where ψ is the amplitude function, $N(z)$ is the buoyancy frequency, k is the wave number, and c is the phase velocity of a particular wave mode. U'' is the vertical gradient of the shear which will be neglected in this analysis. The shape of the density profile is described by $N(z)$ where:

$$N^2(z) = -\frac{g}{\rho}\frac{d\rho}{dz}. \tag{2}$$

An interface is defined as a position where $N(z)$ has a local maximum. The lowest order solution of this equation will provide for the largest wave amplitude to occur at the same vertical position as the largest density gradient. The eigenfunctions for higher mode waves have "zero-crossings" within the domain of the fluid, but as long as there are $n+1$ local maxima in $N(z)$ (and therefore $n+1$ interfaces) the points of maximum amplitude for a mode-N wave will correspond with points of maximum buoyancy. The mode-$(n+1)$ and higher eigenfunctions have maximum amplitudes away from points of maximum density gradient and often have minimum amplitudes at local maxima in $N(z)$.

For the case of a single thick interface, the amplitude function for the mode-zero wave will have its maximum in the center of the interface while the mode-one amplitude function will vanish at the center. If the interface is vigorously mixed, say by a shear layer instability coincident with the interface, the resulting buoyancy frequency distribution will have two separate maxima located at the top and bottom of the old interface and a zero point at the old center. The mode-one amplitude function will now have amplitude maxima corresponding to buoyancy maxima. The mode-zero maximum will move to the stronger of the two new interfaces (Muller

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of the American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, May 19, 1986; final revision, June 1, 1987.

Paper No. 88-WA/APM-4.

and Shang, 1985). These changes will have a profound effect on the propagation of waves which we wish to model using well established approaches along with examining the role of increasing overall thickness.

2 Single Interfaces

We consider first the case of an infinite fluid with a single, thick interface. In 1948, P. Groen published results (Groen, 1948) where interfaces of finite thickness were modelled mathematically using hyperbolic tangents and solutions to the Taylor-Goldstein equation were arrived at with the use of hypergeometric series. Expressing his results in our notation provides a dispersion relation:

$$\Omega^2/\beta k = \frac{\epsilon k}{\epsilon^2 k^2 + (2n+1)2\epsilon k + 4n(n+1)} \tag{3}$$

where Ω is the angular frequency, k is the wave number, n is the mode number of the internal wave, and ϵ is the interface thickness. This solution applies only where the wavelength of the waves in question are long compared to the interface thickness.

For the lowest internal mode ($n=0$), equation (3) reaches a maximum value of 1/2 for any particular wave number as $\epsilon k \to 0$, and the wave speed (Ω/k) is a monotonically decreasing function of interface thickness. This is not the case when higher mode waves are considered. For $n>1$, the limit of equation (3) as $\epsilon k \to 0$ is zero wave speed. However, the wave speed also tends to zero in the limit as $\epsilon k \to \infty$. Therefore, for a given wave number there must be a maximum wave speed as the interface grows in thickness from zero. After that point, the wave speed will decrease with increasing interface thickness as in the case of the mode-zero wave. The location of the maximum in the dispersion relation can easily be found by setting the derivatives of equation (3) to zero. The maximum occurs at:

$$\epsilon^2 k^2 = 4n(n+1). \tag{4}$$

Again, for the model to be valid, it is necessary for the wavelength to be at least somewhat larger (say twice as large) than the thickness of the interface. Therefore,

$$k\epsilon = 2\epsilon\pi/\lambda < \pi \tag{5}$$

where λ is the wavelength of the internal wave. For modes greater than $n=1$, this criteria is never met and the local maximum in the dispersion relation occurs at a point where the internal waves are no longer constrained to propagate horizontally. This is not the case for mode-one waves, however, where the maximum occurs at $k\epsilon = 2\sqrt{2}$.

An explanation for the predicted increase in wave speed with interface thickness can be obtained by looking at the mode-one case. Here, the wave amplitude has a zero crossing at the center of the interface. One can imagine placing a "slippery" wall at exactly this point with no change in the wave. As the interface thickness increases, the point of maximum amplitude (in the vertical) moves farther away from the imaginary wall. It is well known (see Phillips, 1977) that the presence of a wall slows internal waves. Therefore, an increase in interface thickness should result in a speeding up of the wave. When $k\epsilon > 2\sqrt{2}$, the effect of the "wall" along the centerline of the interface is overbalanced by the decreasing local density gradient and the wave slows down again.

3 Double Interfaces

A double interface created by the vigorous mixing of a single interface can be characterized by a zero in buoyancy frequency surrounded by two sharp peaks which are in general larger than would appear in the single interface.

When two interfaces are present, both mode-zero and mode-one waves can be resolved by treating the interfaces as discontinuous density jumps. This was carried out following the method of Neumann and Pierson (1966, Chapter 10). Solving the problem for two interfaces separated by a distance ϵ results in the following dispersion relation:

$$\Omega^2/\beta k = \frac{(1+\Gamma) \pm \sqrt{\Gamma^2 - \delta}}{2(2\Gamma + 1 + \delta)} \tag{6}$$

where

$$\delta = 4\rho_1\rho_2/(\rho_1\rho_2)^2$$

and

$$\Gamma = \coth(k\epsilon).$$

The parameter δ is the ratio between the harmonic mean and the algebraic mean of the densities of the upper and lower layers. Variations in δ would be expected only to change the quantitative nature of the phenomena and, therefore, will be held at 0.96 (representing weak interfaces) in calculations reported here.

Neumann and Pierson considered the case when $\rho_1 = 0$ which is useful when modeling the interactions between an air/water interface with zero surface tension and a thermocline at depth ϵ. The two roots to the dispersion relation for $\rho_1 = 0$ are:

$$\Omega^2/\beta k \begin{cases} = 1/2 & \text{for mode-zero} \\ = 0.5/(2\Gamma+1) & \text{for mode-one} \end{cases}$$

Note that the mode-zero wave (surface wave) shows no variation in velocity even as $\epsilon \to 0$ while the mode-one wave will increase in velocity as the interfaces separate until Γ becomes close to one.

It is useful to examine the limits of equation (6) for Γ very large ($\epsilon \to 0$) and Γ equal to one ($\epsilon \to \infty$). In tabular form we find:

Γ	n	$\Omega^2/\beta k$
∞	0	1/2
∞	1	1/2
1	0	$(1-\rho_1)/(1+\rho_1)$
1	1	$(\rho_2-1)/(\rho_2+1)$

When the interfaces are close together, the limits are identical to the single interface model. When the interfaces are far apart they reach limiting wave speeds and propagate independently from one another. In a sense, one wave doesn't know that the other wave exists. This means also that there is no local maximum in wave speed for the mode-one wave.

The effect of interface thickness on these two models can be seen by holding the buoyancy constant and observing how the wave speed depends on $k\epsilon$. Figure 1 shows a plot of such an analysis for the four waves that have been discussed with $\beta = 0.2g$ selected as a representative buoyancy. One should note that for $k\epsilon < 1$, there is little difference between models for the mode-zero wave. However, there is a marked difference in the mode-one waves for $k\epsilon$ relatively small.

4 Effect on Resonance

One interesting question concerning interfacial waves is the vertical transport of energy through resonant wave interactions. Most recently Phillips (1980) has written a fascinating historical account of progress made in understanding resonant interactions of many kinds. Resonant interactions can occur in a fluid whenever the quadric forcing term resulting from the nonlinear superposition of two waves also satisfies the dispersion relation for the system resulting in the standard resonance criteria:

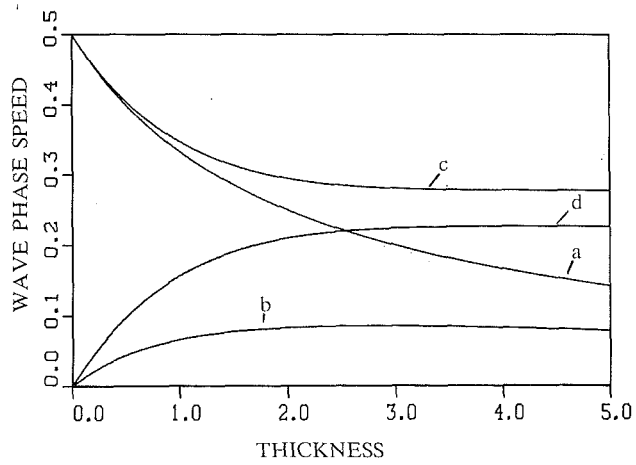


Fig. 1 Dispersion relations for the first two modes using the single and double interface models: (a) single interface, Mode 0; (b) single interface, Mode 1; (c) double interface, Mode 0 (d) double interface, Mode 1

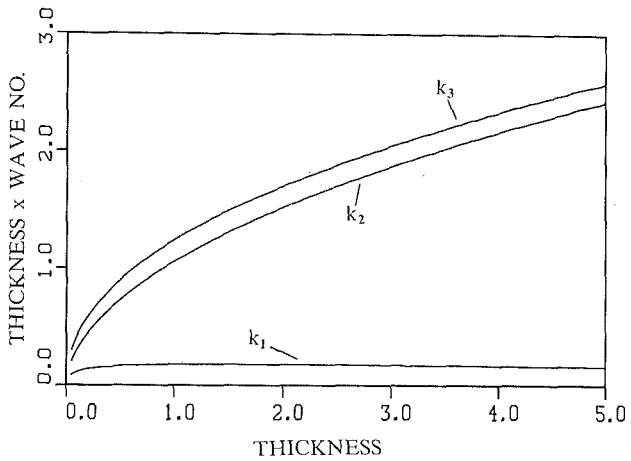


Fig. 2 Resonant triads for a single interface. k_1 is the rapidly traveling Mode 0 wave, k_2 is the slowly traveling Mode 0 wave, and k_3 is the slowly traveling Mode 1 wave.

$$k_1 \pm k_2 = k_3 \quad \text{and} \quad \Omega_1 \pm \Omega_2 = \Omega_3$$

We consider here a special subset of possible resonant triads which can occur in moving fluids. Earlier work (Muller and Karlsson, 1980; Muller and Karlsson, 1982; Muller and Shang, 1985) has shown that if a mean flow exists which travels more rapidly than the lowest internal wave modes of a system, a nontranslating wave-maker will generate two families of waves for each wave mode, one traveling faster than the mean flow and one traveling more slowly. All of the generated waves will have the same frequency, namely the frequency of the wave-maker and therefore the dispersion relation will become double-valued. We then consider whether two mode-zero waves so generated can resonantly interact and thus transfer energy vertically through a layered fluid. The additional complexity of the algebra required to study resonant interactions in a moving fluid is detailed in Shang (1984). The interest in studying resonance in moving fluids is two-fold. Firstly, the presence of a supercritical mean flow provides, under special conditions, for "explosive" resonant interactions to occur as discussed by Cairn (1979) and Craik and Adam (1979). This aspect will not be addressed here. The second reason for interest is that, in a moving fluid, a single wave source can produce both waves necessary for resonant energy transfer.

Since the two primary waves have the same frequency, one can rewrite the conditions for resonance as:

$$k_1 + k_2 = k_3 \quad \text{and} \quad \Omega_1 = \Omega_2 = \Omega_3/2$$

where subscripts 1 and 2 refer to mode-zero waves and subscript 3 refers to the mode-one wave.

Using the dispersion relations for both the double interface and the thick single interface cases, possible occurrences of resonance were determined using the following algorithm. A guess was made for the wave number of the rapidly traveling mode-zero wave and the mode-zero dispersion relation was used to find the frequency of the wave and the other corresponding mode-zero wave. The two wave numbers were summed and then placed into the mode-one dispersion relation and the resultant frequency was compared with that of the mode-zero waves. The value of the initial wave number was then iterated until a resonance was found.

These calculations were carried out using both models for various interface thicknesses. Resonance was found for each interface thickness using the two different models indicating that such a resonance is possible in nature. The results are shown in Fig. 2 for the thick single interface and in Fig. 3 for the double interface case. In each plot, the three wave

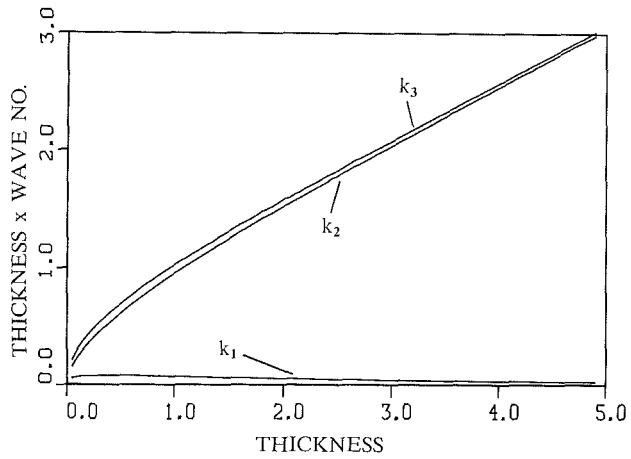


Fig. 3 Resonant triads for a double interface. k_1 is the rapidly traveling Mode 0 wave, k_2 is the slowly traveling Mode 0 wave, and k_3 is the slowly traveling Mode 1 wave.

numbers making up the triad are plotted as a function of interface thickness. With a mean velocity, U, the parameters of interest can be nondimensionalized as follows:

$$k' = kU^2/\beta, \Omega' = \Omega U/\beta, \quad \text{and} \quad \epsilon' = \epsilon\beta/U^2$$

The mean velocity can be viewed as either that of the flow or that of the wave-maker in a still fluid.

Resonant interactions stem from the nonlinearities in the equations of motion and occur at second order for internal gravity waves. Since the nonlinear terms contain spatial derivatives, the growth rate of the growing mode-one wave scales with the wave number. That is, more rapid energy transfer will be expected when the wave number is larger. A careful examination of Figs. 2 and 3 show that for $k\epsilon < 2\sqrt{2}$, resonant interactions will occur with larger wave numbers on the non-mixed interface, while above $k\epsilon = 2\sqrt{2}$, larger wave numbers will be obtained when the interface is vigorously mixed.

5 Discussion of Results

Many previous studies of interfacial waves have been concerned with $k\epsilon << 1$ and certainly this is an area of great importance. However, in the intermediate range, $(0.1 < k\epsilon < \pi)$ this study has shown a local maximum in wave speed which occurs on interfaces of finite thickness for waves of mode-one.

Consequently for waves of mode-zero, the wave speed is a monotonically decreasing function of interface thickness while for mode-one waves the change in wave speed with interface thickness depends on the relative size of $k\epsilon$.

It was also observed that the vigorous mixing of an interface will greatly increase the speed of higher mode waves while leaving the mode-zero waves relatively untouched. The importance of this is demonstrated by the calculation of resonant triads. As the interface thickness becomes very large, the mode-one wave of the triad propagating on the single interface will have a smaller wave number than in the vigorously mixed case and, therefore, energy will transfer between modes more slowly. For very thin interfaces the opposite is true.

We have considered here a moving fluid and a non-translating wave-maker, a case which would occur, for example, in the ocean when strong currents flow about an undersea mount. An obvious and important counter application is where currents could be ignored and the wave-source is translating. An application of this would be the possible identification of isolated gravity currents or other undersea motions by wave motions excited at the sea surface.

References

Cairns, R. A., 1979, "The Role of Negative Energy Waves in Some Instabilities of Parallel Flow," *J. Fluid Mech.*, Vol. 92, pp. 1–14.

Craik, A. D. D., and Adam, J. A., 1979, "Explosive Resonant Interactions in a Three-Layer Fluid Flow," *J. Fluid Mech.*, Vol. 92, pp. 15–33.

Groen, P., 1948, "Contributions to the Theory of Internal Waves," *Mededeligen en Verhandelingen*, Series B, Deel II, No. 11, Koninklijk Nederlands Meteoro logisch Institut de Bilt, pp. 23.

Muller, M. R., and Karlsson, S. K. F., 1981, "Experiments on the Propagation of Composite Internal Wave Trains," *Physics of Fluids*, Vol. 24, No. 11, pp. 1937–1943.

Muller, M. R., and Karlsson, S. K. F., 1982, "A Shear Instability Induced by a Composite Internal Wave," *Physics of Fluids*, Vol. 25, No. 10, pp. 1705–1714.

Muller, M. R., and Shang, P. C., 1985, "The Generation of Internal Gravity Waves in Multi-Layered Fluids by a Flap-Type Wave-Maker," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 247–252.

Neumann, G., and Pierson, Jr., W. J., 1966, *Principles of Physical Oceanography*, Prentice Hall, Inc., Englewood Cliffs, NJ.

Phillips, O. M., 1981, "Wave Interactions: The evolution of an Idea," *J. Fluid Mech.*, Vol. 106, pp. 215–227.

Shang, P. C., 1984, "Some Aspects of Internal Gravity Waves in a Multi-Layered Moving Stratified Fluid," Ph.D. Dissertation, Rutgers University.

Thorpe, S. A., 1978, "On the Shape and Breaking of Finite Amplitude Internal Gravity Waves in a Shear Flow," *J. Fluid Mech.*, Vol. 85, pp. 7–31.

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.

Effect of Nominal Strain-Rates on the Initiation and Growth of Adiabatic Shear Bands in Steels

R. C. Batra[1]

Adiabatic shear banding, i.e., the phenomenon of shear strain localization in high strain-rate deformations of ductile materials, has received considerable attention recently. We refer the reader to Clifton et al. (1984), Wright (1987), Batra (1987a), Fressengeas and Molinari (1987), and Grady and Kipp (1987) for references on various aspects of adiabatic shear bands. Herein we focus on studying the effect of the nominal strain-rate on the initiation and development of shear bands. For this purpose we use a special case of the thermo-viscoplasticity theory which Wright and Batra (1987) derived by including the strain-rate effect in the general thermoplasticity theory due to Green et al. (1968).

Previous works (Clifton et al., 1984; Batra, 1987a) studied initial-boundary value problems that simulated a material defect by superposing a perturbation, usually in the temperature field, on the body previously deformed to a point on the stress-strain curve that was just before the peak stress. Here we add a temperature perturbation to the configuration of the body in which it just starts deforming plastically.

Formulation of the Problem

Referrring the reader to Wright and Batra (1987) for details, we note that equations, in nondimensional variables, that govern the simple shearing deformations of a semi-infinite viscoplastic body bounded by the planes $y = \pm 1$ are

$$\dot{v} = \frac{1}{\rho} s_{,y}, \quad \dot{\theta} = k\,\theta_{,yy} + \Lambda\, s^2, \tag{1}$$

$$\dot{s} = \mu(v_{,y} - \Lambda\, s), \quad \dot{\Psi} = \Lambda\, s^2 \Big/ \left(1 + \frac{\Psi}{\Psi_o}\right)^n, \tag{2}$$

$$\Lambda = \max\left[0, \left\{\left(\frac{s}{\left(1+\frac{\Psi}{\Psi_o}\right)^n (1 - a\theta)}\right)^{\frac{1}{m}} - 1\right\}\right] / bs, \tag{3}$$

with boundary conditions

$$v(\pm 1, t) = \pm 1, \quad \theta_{,y}(\pm 1, t) = 0, \tag{4}$$

and a suitable set of initial conditions. Equations $(1)_1$ and $(1)_2$ express, respectively, the balance of linear momentum and internal energy. Here v is the x-velocity of a material particle, ρ its mass density, s the shearing stress at a material particle, θ its temperature change from that in the unstressed reference configuration, μ is the shear modulus, parameters n and m describe the strain and strain-rate sensitivity of the material, a superimposed dot indicates material time differentiation, and a comma followed by a y signifies partial differentiation with respect to y. Of these nondimensional variables only ρ, b, and k depend upon the nominal strain-rate $\dot{\gamma}_o$; ρ varies as $\dot{\gamma}_0^2$, b is proportional to $\dot{\gamma}_o$, and k to $1/\dot{\gamma}_o$.

The constitutive relations (2) and (3) give one possible model of viscoplastic materials. Equations $(2)_1$ and (3) imply that the plastic part ($\Lambda\, s$) of the strain rate vanishes when

$$|s| \le \left(1 + \frac{\Psi}{\Psi_o}\right)^n (1 - a\theta). \tag{5}$$

For the initial conditions we take

$$\begin{aligned} &v(y,0) = y, \; s(y,0) = (1 - \theta(y,0))(1+b)^m, \\ &\theta(y,0) = 0.1(1-y^2)^9 e^{-5y^2}. \end{aligned} \tag{6}$$

We seek solutions of equations (1)–(4) and (6) such that v is antisymmetric, and s and θ are symmetric in y. Thus the problem is to be studied over the spatial domain [0, 1] and the boundary conditions become

$$v(0,t) = 0, \; \theta_{,y}(0,t) = 0, \; v(1,t) = 1, \; \theta_{,y}(1,t) = 0. \tag{7}$$

Computation and Discussion of Results

The reader is referred to Batra (1987a) for details of integrating the preceding equations. We took the following values of various parameters that correspond to a typical hard steel when $\dot{\gamma}_o = 500\ \text{s}^{-1}$ in computing the numerical results presented and discussed below.

[1] Department of Engineering Mechanics, University of Missouri-Rolla, Rolla, MO 65401-0249. Mem. ASME.

Manuscript received by ASME Applied Mechanics Division, April 20, 1987; final revision August 15, 1987.

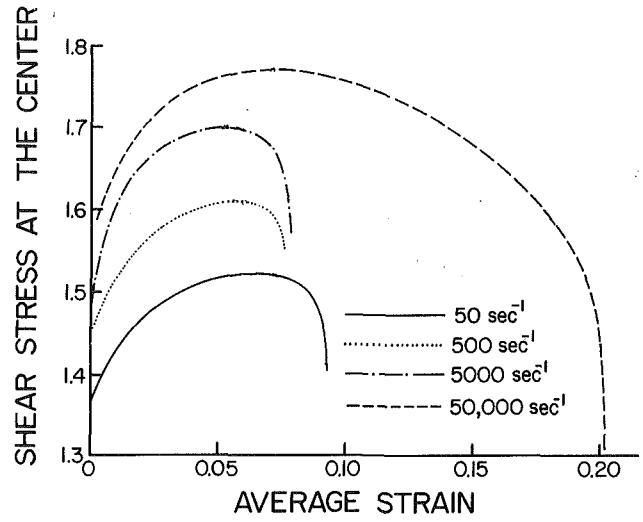


Fig. 1 Shear stress at the center versus average shear strain

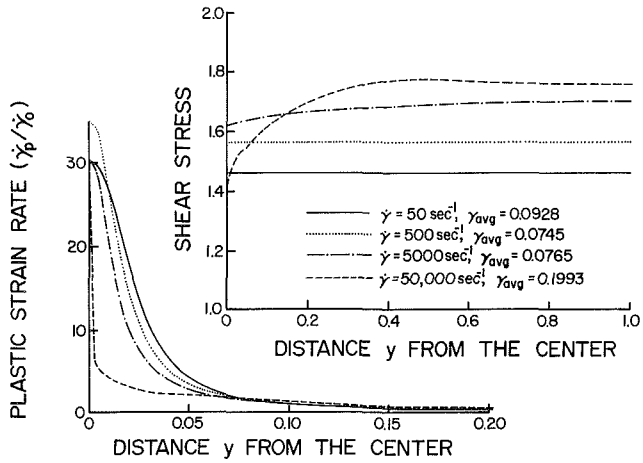


Fig. 2 Shear stress and plastic strain-rate versus distance from the center

$\rho = 3.928 \times 10^{-5}$, $k = 3.978 \times 10^{-3}$, $a = 0.4973$, $\mu = 240.3$

$n = 0.09$, $\psi_o = 0.017$, $b = 5 \times 10^6$, $m = 0.025$

At other nominal strain-rates the values of ρ, b, and k were appropriately scaled.

Figure 1 depicts the shear stress at the center versus the average strain γ_o in the specimen for $\dot{\gamma}_o$ = 50 s^{-1}, 500 s^{-1}, 5000 s^{-1}, and 50,000 s^{-1}. Note that the peak in the stress-strain curve occurs at a lower value of the average strain as $\dot{\gamma}_o$ is increased from 50 s^{-1} to 5000 s^{-1} but occurs at a relatively larger value of γ_o when $\dot{\gamma}_o$ is increased from 5000 s^{-1} to 50,000 s^{-1}. Prior to the peak in the stress, the strain and strain-rate hardening predominate but beyond the peak these effects are overcome by the thermal softening. In Fig. 2 the distribution of the shear stress and the plastic strain-rate $\dot{\gamma}_o$ is plotted through the thickness of the slab at values of the average strain when the rate of change of plastic strain-rate at the center becomes 10^7 s^{-2} or higher. The values of average strains at which results are plotted in Fig. 2 are listed therein. For $\dot{\gamma}_o$ = 50 s^{-1} and 500 s^{-1} the shear stress is essentially uniform throughout the thickness of the slab but for $\dot{\gamma}_o$ = 5000 s^{-1} and 50,000 s^{-1}, the shear stress in the center is less than what it is in the interior of the slab. Since the nondimensional mass density is proportional to $\dot{\gamma}_o^2$, at high strain rates the inertia effects become more predominant. That the shear strain localization has occurred is clear from the plots of $\dot{\gamma}_p$ versus y in Fig. 2.

Following Wright (1987) we define the half band-width d as the distance from the center of the band to the point where the plastic strain rate has fallen to 1/10th of its central value. The values of d so determined from Fig. 2(b) equal 158 μm, 136 μm, 122 μm, and 9 μm at $\dot{\gamma}_o$ = 50 s^{-1}, 500 s^{-1}, 5000 s^{-1} and 50,000 s^{-1}, respectively. These values of d and those computed by Wright (1987) who studied the steady state problem are of the same order of magnitude and exhibit the same trend. As $\dot{\gamma}_o$ varies from 50 s^{-1} to 50,000 s^{-1} the value of k decreases from 3.978×10^{-2} to 3.978×10^{-5} and that of b increases from 5×10^5 to 5×10^8. Numerical computations described by Batra (1987b) who used the constitutive relations (2) and (3), kept $\dot{\gamma}_o$ = 500 s^{-1} but changed k and b by varying their dimensional values reveal that the values of the semi-band width d are unaffected for $0 \le k \le 0.063$ and $5 \times 10^2 \le b \le 5 \times 10^6$. Finally, we note that for $b = 5 \times 10^8$ and $\dot{\gamma}_o$ = 500 s^{-1} the value of d was computed to be 85 μm.

Conclusions

In the simple shearing of the viscoplastic block studied herein at $\dot{\gamma}_o$ = 50 s^{-1}, 500 s^{-1}, 5000 s^{-1}, and 50,000 s^{-1}, the inertia forces start playing a significant role at $\dot{\gamma}_o$ = 5000 s^{-1}. For other constitutive relations this may not be true. At lower strain rates the shear stress stays essentially uniform throughout the specimen. The computed velocity fields reveal that after the shear strain localization has occurred, most of the specimen away from the boundaries of the band moves as a rigid block at relatively low strain rates, but such is not the case at strain rates of 5000 s^{-1} or higher. Narrower bands are formed at higher strain rates.

Acknowledgments

This work was supported by the U. S. Army Research Office through Contract DAAG29-85-K-0238 to the University of Missouri-Rolla.

References

Batra, R. C., 1987a, "The Initiation and Growth of, and the Interaction Among Adiabatic Shear Bands in Simple and Dipolar Materials," *Int. J. Plasticity*, Vol. 3, pp. 75–89.

Batra, R. C., 1987b, "Effect of Material Parameters on the Initiation and Growth of Adiabatic Shear Bands," *Int. J. Solids and Structures*, Vol. 23, pp. 1435–1446.

Clifton, R. J., Duffy, J., Hartley, K. A., and Shawki, T. G., 1984, "On Critical Conditions for Shear Band Formation at High Strain Rates," *Scripta Metallurgica*, Vol. 19, pp. 443–448.

Fressengeas, C., and Molinari, A., 1987, "Instability and Localization of Plastic Flow in Shear at High Strain Rates," *J. Mech. Phys. Solids*, Vol. 35, pp. 185–211.

Grady, D. E., and Kipp, M. E., 1987, "The Growth of Unstable Thermoplastic Shear With Application to Steady-Wave Shock Compression in Solids," *J. Mech. Phys. Solids*, Vol. 35, pp. 95–118.

Green, A. E., McInnis, B. C., and Naghdi, P. M., 1968, "Elastic-Plastic Continua with Simple Force Dipole," *Int. J. Engng. Sci.*, Vol. 6, pp. 373–394.

Wright, T. W., 1987, "Some Aspects of Adiabatic Shear Bands," in *Metastability and Incompletely Posed Problems*, Antman, S., Ericksen, J. L., Kinderlehrer, D., and Muller, I., eds., Springer-Verlag, Berlin, pp. 353–372.

Wright, T. W., and Batra, R. C., 1987, "Adiabatic Shear Bands in Simple and Dipolar Plastic Materials," in *Proceedings of IUTAM Symposium on Macro- and Micro-Mechanics of High Velocity Deformation and Fracture*, Kawata, K., ed., Springer-Verlag, Berlin, in press.

Effects of Coriolis Acceleration on the Forced Vibration of Rotating Cylindrical Shells

S. C. Huang[1,3] and W. Soedel[2,3]

Introduction

The dynamic behavior of rotating cylindrical shells has been of interest for nearly a century. Bryan (1890) first investigated a revolving cylinder and discovered the traveling modes. After that, many authors have steadily and systematically improved our knowledge of rotating cylinders (DiTaranto and Lessen, 1964; Mizoguchi, 1964; Srinivasan and Lauterbach, 1971; Padovan, 1973). However, all the papers mentioned above concentrated on the effects of rotation on natural frequencies and modes. Lately, Fox and Hardie (1985) applied a harmonic point load to a rotating cylinder as a rotation measuring device, and the authors developed a new method to solve for the forced vibrations of rotating rings and cylindrical shells (Huang and Soedel, 1987). In the following, a rotating simply-supported cylindrical shell acted upon by a constant, or harmonic point load is treated. An inversion of the rotating shell problem to the case of a stationary shell with rotating load is also investigated since such inversions are often used in applied engineering. The Coriolis effect on the forced vibration response is extracted through the comparison of these two cases.

I Formulation

The equations of motion are

$$L_x(u_x,u_\theta,u_z)+\rho h\ddot{u}_x=q_x(\xi,\theta,t) \tag{1a}$$

$$L_\theta(u_x,u_\theta,u_z)+\rho h(\ddot{u}_\theta+2\Omega\dot{u}_z-\Omega^2u_\theta)=q_\theta(\xi,\theta,t) \tag{1b}$$

$$L_z(u_x,u_\theta,u_z)+\rho h(\ddot{u}_z-2\Omega\dot{u}_\theta-\Omega^2u_z)=q_z(\xi,\theta,t) \tag{1c}$$

where θ and $\xi=x/\ell$ are shell-fixed coordinates. Also

$$L_x(u_x,u_\theta,u_z)=-\frac{(K+N_x^i)}{\ell^2}u_{x,\xi\xi}-\frac{K(1+\nu)}{2a\ell}u_{\theta,\xi\theta}-\frac{1}{a^2}\left[N_\theta^i+\frac{K(1-\nu)}{2}\right]u_{x,\theta\theta}+\frac{\nu K}{a\ell}u_{z,\xi} \tag{2a}$$

$$L_\theta(u_x,u_\theta,u_z)=\frac{D}{a^4}(u_{z,\theta\theta\theta}-u_{\theta,\theta\theta})-\frac{K}{a^2}(u_{\theta,\theta\theta}+u_{z,\theta})+\frac{D}{a^2\ell^2}u_{z,\xi\xi\theta}-\frac{(1+\nu)K}{2a\ell}u_{x,\xi\theta}-\frac{1}{\ell^2}\left[N_x^i+\frac{(1-\nu)}{2}\left(K+\frac{D}{a^2}\right)\right]u_{\theta,\xi\xi}+k_\theta u_\theta-\frac{N_\theta^i}{a^2}(2u_{z,\theta}+u_{\theta,\theta\theta}-u_\theta) \tag{2b}$$

$$L_z(u_x,u_\theta,u_z)=\frac{D}{a^4}(u_{z,\theta\theta\theta\theta}-u_{\theta,\theta\theta\theta})+\frac{K}{a^2}(u_{\theta,\theta}+u_z)-\frac{D}{a^2\ell^2}u_{\theta,\xi\xi\theta}+\frac{\nu K}{a\ell}u_{x,\xi}+\frac{2D}{a^2\ell^2}u_{z,\xi\xi\theta\theta}+\frac{D}{\ell^4}u_{z,\xi\xi\xi\xi}-\frac{N_x^i}{\ell^2}u_{z,\xi\xi}+k_zu_z-\frac{N_\theta^i}{a^2}(-2u_{\theta,\theta}+u_{z,\theta\theta}-u_z) \tag{2c}$$

where h = transverse thickness, D = bending stiffness, K = membrane stiffness, E = Young's modulus, ν = Poisson's ratio, ρ = density, k_θ = tangential foundation stiffness, k_z = radial foundation stiffness, q_x = applied load in axial direction, q_z = applied load in radial direction, q_θ = applied load in tangential direction, N_x^i = initial tension in axial direction = 0., N_θ^i = initial tension in circumferential direction = $\rho ha^2\Omega^2+ap_0$, p_0 = internal pressure.

Traveling modes are of the form

$$\begin{Bmatrix}u_x\\u_\theta\\u_z\end{Bmatrix}=\begin{Bmatrix}C_{mn}\cos(m\pi\xi)\cos(n\theta+\omega_{mn}t)\\D_{mn}\sin(m\pi\xi)\sin(n\theta+\omega_{mn}t)\\\sin(m\pi\xi)\cos(n\theta+\omega_{mn}t)\end{Bmatrix} \tag{3a}$$

or

$$\begin{Bmatrix}u_x\\u_\theta\\u_z\end{Bmatrix}=\begin{Bmatrix}C_{mn}\cos(m\pi\xi)\sin(n\theta+\omega_{mn}t)\\-D_{mn}\sin(m\pi\xi)\cos(n\theta+\omega_{mn}t)\\\sin(m\pi\xi)\sin(n\theta+\omega_{mn}t)\end{Bmatrix} \tag{3b}$$

where C_{mn} and D_{mn} are the displacement ratios, and m, $n=1$, 2, ...

The solution of the general forced vibration problem is of the form

$$u_x=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{i=1}^{6}C_{mni}p_{mni}(t)\cos(m\pi\xi)\cos(n\theta+\omega t) \tag{4a}$$

$$u_\theta=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{i=1}^{6}D_{mni}p_{mni}(t)\sin(m\pi\xi)\sin(n\theta+\omega t) \tag{4b}$$

$$u_z=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{i=1}^{6}p_{mni}(t)\sin(m\pi\xi)\cos(n\theta+\omega t) \tag{4c}$$

where $p_{mni}(t)$ is the modal participation factor. The forced solution is obtained in terms of infinite series sums of different mn combinations, and each combination is characterized by

$$\begin{Bmatrix}\ddot{\mathbf{x}}_1\\\ddot{\mathbf{x}}_2\end{Bmatrix}+\begin{bmatrix}\mathbf{C}&-\mathbf{G}\\\mathbf{G}&\mathbf{C}\end{bmatrix}\begin{Bmatrix}\dot{\mathbf{x}}_1\\\dot{\mathbf{x}}_2\end{Bmatrix}+\begin{bmatrix}\mathbf{K}&\mathbf{O}\\\mathbf{O}&\mathbf{K}\end{bmatrix}\begin{Bmatrix}\mathbf{x}_1\\\mathbf{x}_2\end{Bmatrix}=\begin{Bmatrix}\mathbf{Q}_1\\\mathbf{Q}_2\end{Bmatrix} \tag{5}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the generalized coordinate vectors, and $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are the generalized force vectors. They are of the forms

$$\mathbf{x}_1=\begin{Bmatrix}\alpha_{mn1}\\\beta_{mn1}\\\gamma_{mn1}\end{Bmatrix}=\sum_{i=1}^{6}\cos(\omega_{mni}t)\begin{Bmatrix}C_{mni}p_{mni}(t)\\D_{mni}p_{mni}(t)\\p_{mni}(t)\end{Bmatrix} \tag{6a}$$

$$\mathbf{x}_2=\begin{Bmatrix}\alpha_{mn2}\\\beta_{mn2}\\\gamma_{mn2}\end{Bmatrix}=\sum_{i=1}^{6}\sin(\omega_{mni}t)\begin{Bmatrix}C_{mni}p_{mni}(t)\\D_{mni}p_{mni}(t)\\p_{mni}(t)\end{Bmatrix} \tag{6b}$$

$$\mathbf{Q}_1=\begin{Bmatrix}Q_1\\Q_2\\Q_3\end{Bmatrix}=\frac{2}{\rho h\pi}\int_0^1\int_0^{2\pi}\begin{Bmatrix}q_x\cos(m\pi\xi)\cos(n\theta)\\q_\theta\sin(m\pi\xi)\sin(n\theta)\\q_z\sin(m\pi\xi)\cos(n\theta)\end{Bmatrix}d\theta d\xi \tag{7a}$$

$$\mathbf{Q}_2=\begin{Bmatrix}Q_4\\Q_5\\Q_6\end{Bmatrix}=\frac{2}{\rho h\pi}\int_0^1\int_0^{2\pi}\begin{Bmatrix}-q_x\cos(m\pi\xi)\sin(n\theta)\\q_\theta\sin(m\pi\xi)\cos(n\theta)\\-q_z\sin(m\pi\xi)\sin(n\theta)\end{Bmatrix}d\theta d\xi \tag{7b}$$

The matrices $\mathbf{G}$, $\mathbf{K}$, and $\mathbf{C}$ are

[1]Graduate Assistant.
[2]Professor.
[3]Ray W. Herrick Laboratory, School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907.

Manuscript received by ASME Applied Mechanics Division, January 20, 1987; final revision June 29, 1987.

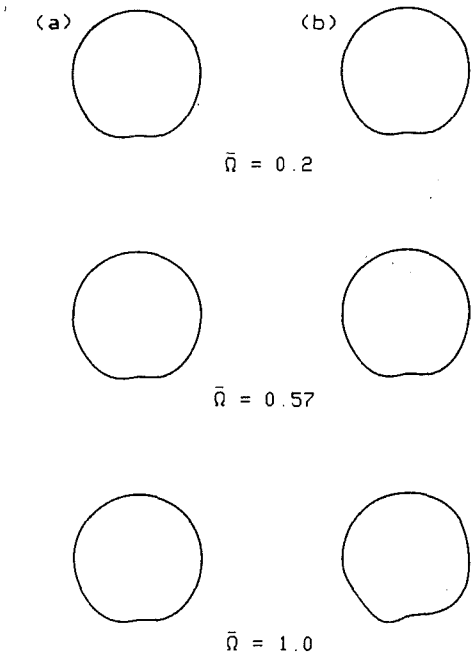


Fig. 1 Steady state response at $\xi = \xi^*$ for three rotational speeds for (*a*) rotating shell and (*b*) rotating load

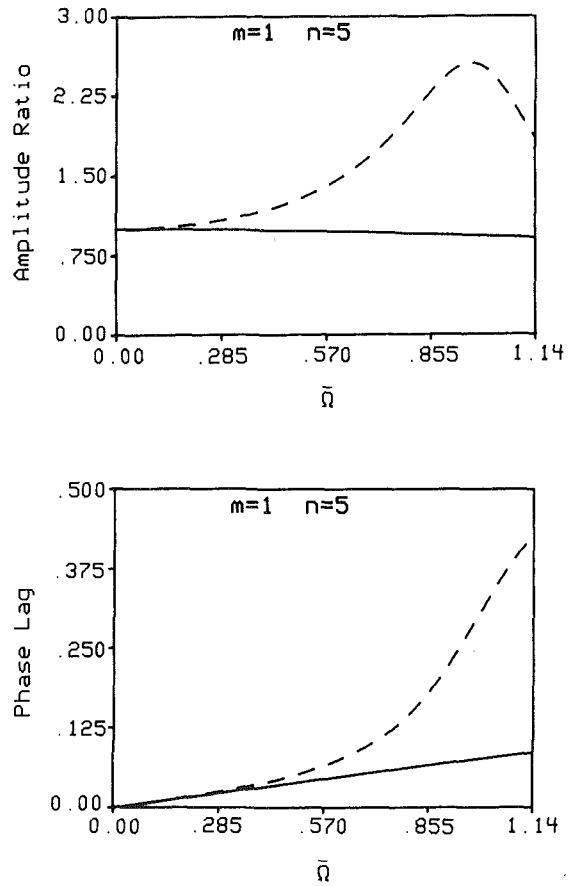


Fig. 2 Steady state displacement ratios and phase lags for $m = 1$, $n = 5$ mode for rotating shell: _____, and rotating load: ·····

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2\Omega \\ 0 & 2\Omega & 0 \end{bmatrix} \qquad \mathbf{K} = \begin{bmatrix} \hat{a} & -\hat{b} & -\hat{c} \\ -\hat{b} & \hat{d} & \hat{e} \\ -\hat{c} & \hat{e} & \hat{f} \end{bmatrix} \tag{8a,b}$$

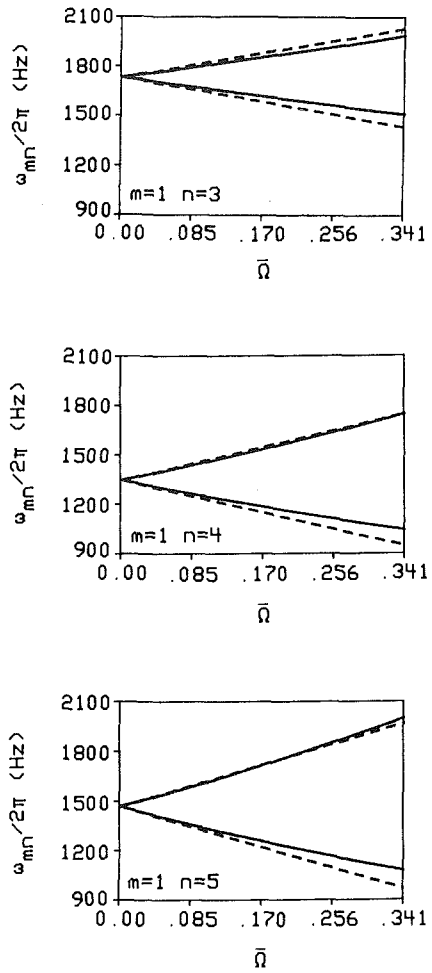


Fig. 3 The variations of resonance frequencies as functions of rotation for rotating shell: _____, and rotating load: ·····

$$\mathbf{C} = \begin{bmatrix} \frac{\lambda_{mn}}{\rho h} & 0 & 0 \\ 0 & \frac{\lambda_{mn}}{\rho h} & 0 \\ 0 & 0 & \frac{\lambda_{mn}}{\rho h} \end{bmatrix} \tag{8c}$$

where the λ_{mn} are equivalent modal damping constants. $\hat{a}$, ... , $\hat{f}$ are system parameters given by

$$\hat{a} = \frac{1}{\rho h}\left[K\left(\frac{m\pi}{\ell}\right)^2 + K\frac{(1-\nu)}{2}\left(\frac{n}{a}\right)^2 + n^2\left(\rho h\Omega^2 + \frac{p_0}{a}\right)\right] \tag{9a}$$

$$\hat{b} = \frac{1}{\rho h}\left[\frac{(1+\nu)nK}{2a}\left(\frac{m\pi}{\ell}\right)\right] \tag{9b}$$

$$\hat{c} = \frac{1}{\rho h}\left[\frac{\nu K}{a}\left(\frac{m\pi}{\ell}\right)\right] \tag{9c}$$

$$\hat{d} = \frac{1}{\rho h}\left[n^2\left(\frac{D}{a^4} + \frac{K}{a^2} + \rho h\Omega^2 + \frac{p_0}{a}\right) + \frac{(1-\nu)}{2}\left(K + \frac{D}{a^2}\right)\left(\frac{m\pi}{\ell}\right)^2 + k_\theta + \frac{p_0}{a}\right] \tag{9d}$$

$$\hat{e} = \frac{n}{\rho h}\left[n^2\frac{D}{a^4} + \frac{K}{a^2} + 2\rho h\Omega^2 + \frac{D}{a^2}\left(\frac{m\pi}{\ell}\right)^2\right] \tag{9e}$$

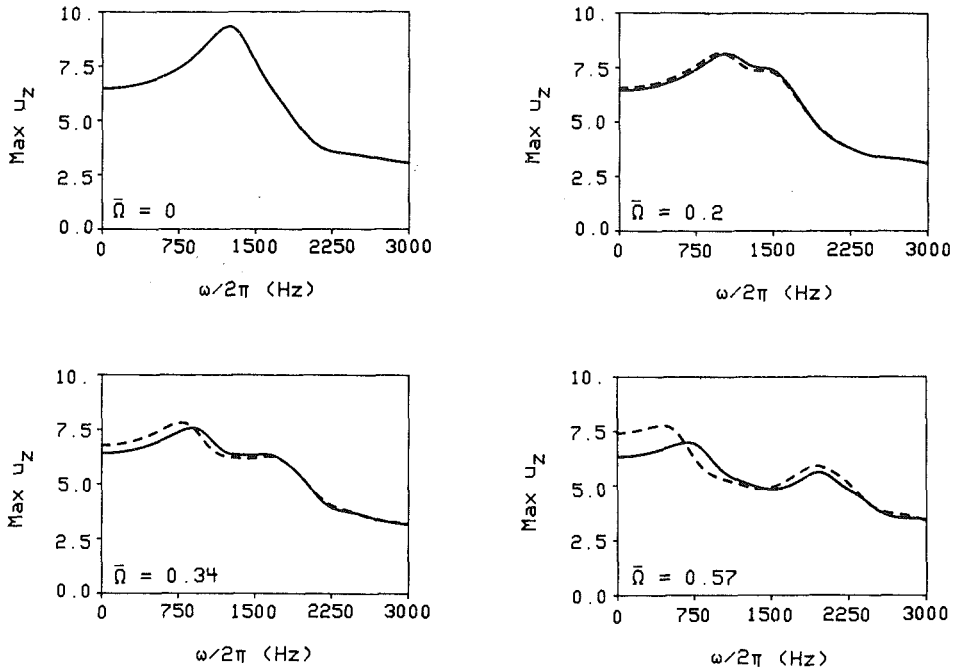


Fig. 4 Frequency response at the load location for the rotating shell: _____, and rotating load: ·····, for four rotational speeds

$$\hat{f}=\frac{1}{\rho h}\left[n^4\frac{D}{a^4}+\frac{K}{a^2}+n^2\left(\rho h\Omega^2+\frac{p_0}{a}\right)+\frac{2n^2D}{a^2}\left(\frac{m\pi}{\ell}\right)^2+D\left(\frac{m\pi}{\ell}\right)^4+k_z+\frac{p_0}{a}\right] \tag{9f}$$

II Constant Point Load Response

A constant point load in transverse direction is described by the following:

$$\begin{aligned} q_x(\xi,\theta,t)&=0\\ q_\theta(\xi,\theta,t)&=0\\ q_z(\xi,\theta,t)&=-\frac{F_z}{a\ell}\delta(\xi-\xi^*)\delta(\theta-\theta^*+\Omega t)\end{aligned} \tag{10}$$

where ξ^* and θ^* locates the load position at a time equal to zero with respect to the θ coordinate. F_z is the force magnitude.

The steady state response to the above loading is shown in Fig. 1(*a*). Figure 1(*b*) shows the results of the corresponding rotating load on stationary shell problem. The difference between the two problems is the Coriolis effect. $\bar{\Omega}$ is a normalized rotational speed with respect to the first critical speed of the undamped shell for the rotating load case.

The results show that the solutions of the two cases differ only slightly at lower $\bar{\Omega}$ values. However, a dramatic difference is found as $\bar{\Omega}$ approaches unity. The difference is attributed to a critical speed which is reached in the rotating load case but not in the rotating shell case. Figure 2 further illustrates the displacement and phase response of the specific mode at which the first critical speed will occur. In that figure the ordinate is normalized with respect to the value for zero rotation.

III Harmonic Point Load Response

For a harmonic point load in radial direction, the loading is described as follows:

$$\begin{aligned} q_x(\xi,\theta,t)&=0\\ q_\theta(\xi,\theta,t)&=0\\ q_z(\xi,\theta,t)&=-\frac{F_z}{a\ell}\delta(\xi-\xi^*)\delta(\theta-\theta^*+\Omega t)\cos(\omega t-\phi)\end{aligned} \tag{11}$$

where ω is the excitation frequency and ϕ is the phase angle.

Rotation splits the original resonance into two. Figure 3 illustrates the resonances as functions of rotational speed for both rotating shell and rotating load cases for the lowest three modes, i.e., $m=1$, $n=3$, 4, and 5.

Figure 4 shows the frequency response at the load location for four different speeds, $\bar{\Omega}=0.$, 0.2, 0.34, 0.57. It is seen that at $\bar{\Omega}=0$ both problems yield the same response curve as expected. This is a partial check of the solution. As rotation increases, the frequency responses of these two cases differ from each other.

References

Bryan, G. H., 1890, "On the Beats in the Vibrations of a Revolving Cylinder or Bell," *Proceedings of the Cambridge Philosophical Society*, pp. 101–111.

DiTaranto, R. A., and Lessen, M., 1964, "Coriolis Acceleration Effect on the Vibration of a Rotating Thin-Walled Circular Cylinder," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 700–701.

Fox, C. H., and Hardie, D. J. W., 1985, "Harmonic Response of Rotating Cylindrical Shells," *Journal of Sound and Vibration*, Vol. 101, pp. 495–510.

Herrmann, G., and Armenakas, A. E., 1962, "Dynamic Behavior of Cylindrical Shells Under Initial Stress," *Proceedings the Fourth U.S. National Congress of Applied Mechanics*, ASME, pp. 203–213.

Huang, S. C., and Soedel, W., 1987, "Effects of Coriolis Acceleration on the Free and Forced In-Plane Vibrations and Rotating Rings on Elastic Foundation," *Journal of Sound and Vibration*, Vol. 115, pp. 253–274.

Huang, S. C., and Soedel, W., 1987, "Response of Rotating Rings to Harmonic and Periodic Loading and Comparison with the Inverted Problem," *Journal of Sound and Vibration*, to appear.

Huang, Shyh-Chin, 1987, "Effects of Coriolis Acceleration on the Free and Forced Vibrations of Spinning Structures," Ph.D. dissertation, Purdue University.

Padovan, J., 1973, "Natural Frequencies of Rotating Prestressed Cylinders," *Journal of Sound and Vibration*, Vol. 31, No. 4, pp. 469–482.

Srinivasan, A. V., and Lauterbach, G. F., 1971, "Traveling Waves in Rotating Cylindrical Shells," ASME *Journal of Engineering for Industry*, Vol. 93, pp. 1229–1232.

On the Order of the Stress Singularity for an Antiplane Shear Crack at the Interface of Two Bonded Inhomogeneous Elastic Materials

L. Schovanec[1] and J. R. Walton[2]

1 Introduction

A number of investigations of crack problems in nonhomogeneous media have been undertaken in which the elastic moduli vary continuously with spatial coordinates. In all of these studies special forms of inhomogeneities have been assumed in order to insure a tractable problem for which the asymptotic form of the stress field near the crack tip could be calculated. One of the primary objectives of these works was to determine the effect of spatial inhomogeneity upon the known singular field quantities associated with the corresponding homogeneous problem.

In the case of a Mode I crack embedded in a nonhomogeneous medium, symmetric about the plane of the crack, the specific models of inhomogeneities examined thus far show that the usual square root singular crack tip stress field associated with the homogeneous medium occurs for the nonhomogeneous problem provided the shear modulus does not vanish in the plane of the crack (see, e.g., Delale and Erdogan, 1983; Gerasoulis and Srivastav, 1980; Schovanec, 1986; Schovanec and Walton, 1987.

Studies of a Mode III crack located on the interface of two bonded materials have shown that the assumption of spatial inhomogeneity in either component part does not alter the known square root singular behavior present in the piecewise homogeneous bimaterial problem. One of the more general models utilized, which includes the homogeneous and bimaterial mediums as special cases, was investigated by Delale (1985) who took the shear modulus in the form $\mu(x, y) = \mu^{\pm} e^{\alpha_x + \beta^{\pm} u}$ with the crack situated in the plane $y = 0$ and μ^+, β^+ (μ^-, β^-) constants specified in the half plane $y > 0$ $(y < 0)$ (see also Clements et al., 1978; Dhaliwal and Singh, 1978).

In the event that the crack tip terminates at a bimaterial interface between two homogeneous mediums, it is known that the stress field singularity is of the form $r^{-\gamma}$, $0 < \gamma < 1$, where r is the distance from the crack tip (Erdogan and Cook, 1972). As discussed by Atkinson (1977), such a result is unstaisfactory from a fracture mechanics point of view in that it results in a stress intensity factor and energy release rate of zero or infinity whenever $\gamma \neq 0.5$. To alleviate this inadequacy in the "ideal" interface model, Atkinson introduced a nonhomogeneous layer at the interface in which the shear modulus varied continuously, matching the constant values at the two outer mediums. For the specific forms of the modulus considered it was then shown that the stress at the crack tip displayed an $r^{-1/2}$ singularity. Assuming that the elastic moduli could be expanded in a Taylor series about $r = 0$, Atkinson suggested that the square root behavior would always result. Somewhat later, the antiplane shear crack perpendicular to an interface was also considered by Erdogan (1985). To investigate the effect of a continuous but nondifferentiable elastic modulus at the interface, a shear modulus of the form $\mu(x) = \mu_0 e^{\beta^{\pm} x}$ was assumed with β^+ (β^-) corresponding to $x > 0$ $(x < 0)$. The main focus of that study was to show that for such a model a square root singularity at the crack tip resulted. Erdogan proposed that the same conclusion should be valid for any continuous but nondifferentiable modulus $\mu(x)$.

This note addresses the conjecture raised in the two previous studies. In particular, the nature of the stress singularity is investigated for a semi-infinite Mode III crack normally incident to the interface of two bonded nonhomogeneous elastic halfspaces. Without assuming any particular form for the shear modulus, only that it be continuous throughout the composite medium, symmetric about the crack plane and differentiable everywhere except along the interface at which the crack tip terminates, it is shown that the crack tip stresses have a square root singularity, thus verifying the conjecture of Erdogan. This note does not consider the general question of existence of solutions for such a class of boundary value problems, which does not seem to be completely settled. Rather, the approach taken here is to show that when there exists a physically meaningful solution, a notion to be made precise later, then that solution exhibits a square root singular stress field.

2 The Boundary Value Problem

The specific problem considered here is that of an infinite, bimedia, nonhomogeneous isotropic elastic solid containing a semi-infinite Mode III crack normally incident to the interface of the two bonded materials. Referred to Cartesian coordinates x, y, z, the crack is assumed to be in the plane $y = 0, x < 0$ and the nonzero out of plane displacement $\omega(x, y)$ is related to the nonzero shear stresses by $\sigma_{xz} = \mu \partial\omega/\partial_x$, $\sigma_{yz} = \mu \partial\omega/\partial y$ where the shear modulus $\mu(x, y)$ is assumed to be continuous throughout the composite medium, differentiable everywhere but along the interface $x = 0$ and symmetric about the crack plane $y = 0$. Symmetry permits consideration of the following boundary value problem for the half-plane $y > 0$:

$$\operatorname{div}(\mu \nabla \omega) = 0 \quad |x| < \infty,\ y > 0$$

$$\mu(x,0)\frac{\partial \omega}{\partial y}(x,0) = P(x) \quad x < 0$$

$$\omega(x,0) = 0 \quad x > 0.$$

The existence of a solution with locally finite strain energy, i.e., $\int_\Omega (\mu/2)\ |\nabla\omega|^2\ dA < \infty$ where Ω is any bounded measurable subset of the upper half-plane, will be assumed. Note that a sufficient condition to insure ω have locally finite strain energy is that near the crack tip, $|\nabla\omega| = 0(r^{-\gamma})$, $0 \leq \gamma < 1$, where r is the radial component in a local polar coordinate system. The point here is to show $\gamma = 1/2$. A growth condition on ω of the form $\int_{|x|>R} \log |\mathbf{x}|\ ((\nabla\mu \cdot \nabla\omega)/\mu)\ (\mathbf{x}/)\ d\mathbf{x} < \infty$ for $R > 0$ is also assumed. For a large class of moduli (e.g., $\mu(x) = (\alpha_x + \beta)$ as considered in Atkinson, 1977), this condition is satisfied if $|\nabla\omega| = 0\ (r^{-1-\epsilon})$, $r \to \infty$, which, is indeed the case if the applied tractions are statically self-equilibrating (Cook and Erdogan, 1977). Again it should be mentioned that the question of determining conditions on the applied traction $S(x)$ and the curve C which assure the existence of such a solution ω is not taken up here.

Define the Green's Γ function by

$$\Gamma(s,t;x,y) = \frac{1}{2\pi} \log \sqrt{[(t-x)^2 + (s-y)^2][(t-x)^2 + (s+y)^2]} + \psi(s,t;x,y)$$

where

$$\frac{\partial^2 \psi}{\partial_t^2} + \frac{\partial^2 \psi}{\partial_s^2} = 0, \quad |s| < \infty,\ t > 0$$

[1]Department of Mathematics, Texas Tech University, Lubbock, TX 79409.
[2]Department of Mathematics, Texas A&M University, College Station, TX 77843. Mem. ASME. Supported by the Office of Naval Research under Contract No. N00014-86-K0298.
Manuscript received by ASME Applied Mechanics Division, October 27, 1986; final revision September 2, 1987.

$$\psi(s,0)=\frac{-1}{2\pi}\log[(s-x)^2+y^2] \quad s>0 \tag{1}$$

$$\frac{\partial\psi}{\partial t}(s,0)=0 \qquad s<0. \tag{2}$$

Applying Green's second identity

$$\int_\Omega v_1\Delta v_2 - v_2\Delta v_1 dA=\int_{\partial\Omega} v_1\frac{\partial v_2}{\partial n}-v_2\frac{\partial v_1}{\partial n}\,ds$$

in each medium separately with $v_1=\omega$, $v_2=\Gamma$, and using continuity of the displacement ω and stress σ_{xz} across the interface, there results the representation formula

$$\omega(x,y)=-\int_{-\infty}^{\infty}\int_0^{\infty}\Gamma(s,t;x,y)\left(\frac{\nabla\mu\cdot\nabla\omega}{\mu}\right)(s,t)dsdt$$

$$+\int_{-\infty}^{0}\Gamma(s,0;x,y)\frac{P(s)}{\mu(s,0)}\,ds. \tag{3}$$

With the above assumptions on ω the first integral exists and the subsequent expressions to be derived from it are also well defined.

It is convenient to seek the harmonic function ψ in the form

$$\psi(s,t;x,y)=\frac{1}{\pi}\int_{-\infty}^{\infty}f(\tau;x,y)\frac{s}{(\tau-s)^2+t^2}\,d\tau.$$

Since

$$\frac{\partial\psi}{\partial t}(s,0;x,y)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\partial f}{\partial\tau}(\tau;x,y)\frac{d\tau}{\tau-s},$$

inversion of this relation and use of equations (2) and (1) gives

$$\frac{-\partial f}{\partial s}(s;x,y)=\frac{1}{\pi}\int_0^{\infty}\frac{\partial\psi}{\partial t}(\tau,0;x,y)\frac{d\tau}{\tau-s}, \quad -\infty<s<\infty \tag{4}$$

and

$$\frac{(s-x)}{(s-x)^2+y^2}=\int_0^{\infty}\frac{\partial\psi}{\partial t}(\tau,0;x,y)\frac{d\tau}{\tau-s}, \quad s>0. \tag{5}$$

Methods for solving the Cauchy singular integral equation (5) are well known (see e.g., Gakov, 1966) and it can be shown that

$$\frac{\partial\psi}{\partial t}(s,0;x,y)=\frac{-1}{\pi^2\sqrt{s}}\int_0^{\infty}\frac{(u-x)\sqrt{u}}{(u-x)^2+y^2}\frac{du}{u-s}, \quad s>0. \tag{6}$$

Substitution of equation (6) into (4), an interchange in the order of integration and the observation that for $u>0, s<0$

$$\int_0^{\infty}\frac{1}{\sqrt{t}}\frac{dt}{(u-t)(t-s)}=\frac{\pi}{\sqrt{|s|}}\frac{1}{u-s}$$

leads to

$$\frac{\partial f}{\partial s}(s;x,y)=\frac{1}{\pi^2\sqrt{|s|}}\int_0^{\infty}\frac{\sqrt{u}(u-x)}{(u-x)^2+y^2}\frac{du}{u-s}, \quad s<0.$$

An integration by parts gives

$$\frac{\partial f}{\partial s}(s;x,y)=\frac{1}{4\pi^2\sqrt{|s|}}\int_0^{\infty}\frac{(u+s)}{\sqrt{u}(u-s)^2}$$

$$\log[(u-x)^2+y^2]du, \quad s<0. \tag{7}$$

At this point it is convenient to observe that if equation (7) is differentiated with respect to y and then integrated with respect to s, keeping in mind that $f(\cdot,x,y)$ must vanish at infinity, then one has for $s<0$.

$$\frac{\partial f}{\partial y}(s;x,0)=\begin{cases}-\dfrac{1}{\pi}\dfrac{1}{\sqrt{x}}\dfrac{\sqrt{|s|}}{x-s} & x>0\\ 0 & x<0\end{cases}. \tag{8}$$

In order to illustrate the singular nature of the stress σ_{yz} and to determine the stress intensity factor, only $\partial\omega/\partial y\ (x,0)$ is required. The analysis of σ_{xz} is similar and omitted.

3 The Stress Singularity and the SIF

From the definition of Γ, equation (1), and the representation of ω given by equation (3), one has that

$$\frac{\partial\omega}{\partial y}(x,0)=\lim_{y\to0^+}\left\{\int_{-\infty}^{\infty}\int_0^{\infty}\left\{\int_{-\infty}^{\infty}\frac{\partial f}{\partial y}(\tau;x,y)\right.\right.$$

$$\left.\left.\frac{t}{(\tau-s)^2+t^2}\,d\tau\right\}\left(\frac{\nabla\mu\cdot\nabla\omega}{\mu}\right)(s,t)dtds\right. \tag{9}$$

$$+\frac{1}{\pi}\int_{-\infty}^{0}\frac{y}{(s-x)^2+y^2}\frac{p(s)}{\mu(s,0)}\,ds$$

$$\left.+\int_{-\infty}^{0}\frac{\partial f}{\partial y}(s;x,y)\frac{P(s)}{\mu(s,0)}\,ds\right\}.$$

Making use of equation (8) it easily follows that the $y\to0+$ limit of the single integrals in equation (9) is given by

$$\begin{cases}\dfrac{P(x)}{\mu(x,0)} & x<0\\ \dfrac{1}{\pi}\dfrac{1}{\sqrt{x}}\displaystyle\int_{-\infty}^{0}\sqrt{|s|}\,\frac{P(s)}{\mu(s,0)}\frac{ds}{s-x} & x>0.\end{cases} \tag{10}$$

By utilizing equations (8) and (1) it can be shown in a similar manner that the $y\to0+$ limit of the triple integral in equation (9) is

$$\begin{cases}\dfrac{1}{\pi}\displaystyle\int_{-\infty}^{\infty}\int_0^{\infty}\left\{\frac{t}{(x-s)^2+t^2}+\frac{1}{\pi}\frac{1}{\sqrt{x}}\int_{-\infty}^{0}\frac{\sqrt{|\tau|}}{x-\tau}\frac{t}{(\tau-s)^2+t^2}\,d\tau\right\}\frac{\nabla\mu\cdot\nabla\omega}{\mu}\,dtds, & x>0\\ 0 & x<0.\end{cases} \tag{11}$$

Combining equations (10) and (11) it follows that for $x>0$,

$$\sigma_{yz}(x,0)=\frac{\mu(x,0)}{\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\frac{t}{(x-s)^2+t^2}\frac{\nabla\mu\cdot\nabla\omega}{\mu}\,dtds$$

$$+\frac{\mu(x,0)}{\pi\sqrt{x}}\int_{-\infty}^{\infty}\int_0^{\infty}\frac{\sqrt{|\tau|}}{x-\tau}\frac{t}{(\tau-s)^2+t^2}\frac{\nabla\mu\cdot\nabla\omega}{\mu}\,d\tau dtds$$

$$+\frac{\mu(x,0)}{\pi\sqrt{x}}\int_{-\infty}^{0}\sqrt{|s|}\,\frac{P(s)}{\mu(s,0)}\frac{ds}{s-x}. \tag{12}$$

From equation (12) the square root singular behavior of σ_{yz} is apparent and, in the case that μ is constant, the known result for the homogeneous medium (Willis, 1967),

$$\sigma_{yz}(x,0)=\frac{1}{\pi\sqrt{x}}\int_{-\infty}^{0}\sqrt{|s|}P(s)\frac{ds}{s-x}$$

is easily recovered. Since a square root singularity is always present at the crack tip, the stress intensity factor may be defined as

$$K=\lim_{x\to 0^+}\ \mu(x,0)\sqrt{x}\,\frac{\partial\omega}{\partial y}(x,0)$$

and from equation (12) one has

$$K=\frac{-\mu_0}{\pi}\left\{\int_{-\infty}^{0}\frac{1}{\sqrt{|s|}}\,\frac{P(s)}{\mu(s,0)}\,ds\right.$$

$$\left.+\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{|\tau|}}\,\frac{t}{(\tau-s)^2+t^2}\,\frac{\nabla\mu\cdot\nabla\omega}{\mu}\,d\tau dt ds\right\}.$$

As pointed out above, the calculation of σ_{xz} is carried out in a similar manner and by changing to polar coordinates it is not difficult to show that in general $\sigma_{ij}(r,\theta) = 0\,(1/\sqrt{r})$, $r \to 0$. It is also worth noting that a similar derivation can be carried out for a finite domain. In that case the appropriate boundary conditions must be prescribed and the analogous Green's function computed. The growth condition imposed on ω, however, would no longer be necessary.

A number of extensions and generalizations of the above result are possible. Clearly, the assumed symmetry of the shear modulus about the crack plane is not essential for the stress singularity to have order 1/2. Rather, it merely simplifies the analysis by permitting reduction to a half plane problem for which a simple, explicit expression for the Green's function, Γ, is easily constructed. When symmetry is not assumed it is still possible to represent the displacement $\omega(x, y)$ by integrals of the form (3) in each of the half spaces $y \gtrless 0$, only now the Green's function Γ must satisfy more complicated boundary conditions derived from imposing continuity of stresses and the displacement along $y = 0, x > 0$ and the applied crack face traction $P(x)$ on $y = 0, x < 0$. Once the result has been established for a nonsymmetric shear modulus, the extension to a possibly curved (smooth) crack normally incident to a possibly curved (smooth) interface is easy, since consideration of an infinitessimally small neighborhood about the crack tip produces a straight crack normally incident to a straight interface. Finally, a major modification of the analysis presented here seems required to study the case of a crack that is obliquely incident to the interface.

References

Atkinson, C., 1977, "On Stress Singularities and Interfaces in Linear Elastic Fracture Mechanics," *International Journal of Fracture*, Vol. 13, pp. 807–820.

Clements, D. L., Atkinson, C., and Rogers, C., 1978, "Antiplane Crack Problems for an Inhomogeneous Elastic Material," *Acta Mechanica*, Vol. 20, pp. 199–211.

Cook, T. S., and Erodgan, F., 1972, "Stresses in Bonded Materials with a Crack Perpendicular to the Interface," *International Journal of Engineering Science*, Vol. 10, pp. 667–697.

Delale, F., 1985, "Mode III Fracture of Bonded Nonhomogeneous Materials," *Engineering Fracture Mechanics*, Vol. 22, pp. 213–226.

Delale, F., and Erdogan, F., 1983, "The Crack Problem for a Nonhomogeneous Plane," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 609–614.

Dhaliwal, R. S., and Singh, B. M., 1978, "On the Theory of Elasticity of a Nonhomogeneous Medium," *Journal of Elasticity*, Vol. 8, pp. 211–219.

Erdogan, F., 1985, "The Crack Problem for Bonded Nonhomogeneous Materials Under Antiplane Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 823–828.

Gakov, F. D., 1966, *Boundary Value Problems*, Pergamon, London.

Gerasoulis, A,. and Srivastav, R. O., 1980, "A Griffith Crack Problem for a Nonhomogeneous Medium," *International Journal of Engineering Science*, Vol. 18, pp. 239–247.

Schovanec, L., 1986, "A Griffith Crack Problem for an Inhomogeneous Elastic Material," *Acta Mechanica*, Vol. 58, pp. 67–80.

Schovanec, L., and Walton, J. R., 1987, "The Quasi-Static Propagation of a Plane Strain Crack in a Power-Law Inhomogeneous Linearly Viscoelastic Body," *Acta Mechanica*, Vol. 66, No. 1, pp.

Willis, J. R., 1967, "Crack Propagation in Viscoelastic Media," *Journal of Mechanics and Physics of Solids*, Vol. 15, pp. 229–240.

Perforation of Aluminum Plates with Conical-Nosed Rods – Additional Data and Discussion

Z. Rosenberg[1] and M. J. Forrestal[2]

Introduction

Backman and Goldsmith (1978) discuss many analytical and experimental methods used to study the broad field of penetration mechanics. For plate perforation, they show eight possible mechanisms that depend on the geometry and material of the projectiles and targets. Because many perforation mechanisms are possible, experimental observations usually precede and guide engineering models. Our study is limited to high strength, conical-nosed projectiles that perforate low strength, ductile targets. For the experiments conducted in this study and reported in a recent paper (Forrestal et al., 1987), targets were 6061-T6 aluminum plates and projectiles were machined from T-200 maraging steel. As shown in our recent paper (1987), the targets experienced ductile hole-growth, and post-perforation, x-ray photographs showed that the projectile remained undeformed.

By contrast to our studies, most work on perforation by long rods has been conducted with high strength, steel targets. For example, Grabarek (1973)[3] launched steel rods against rolled-homogeneous, armor plates and showed, via x-ray photographs, massive nose deformation and rod fracturing. Clearly, the experimental work presented in our study with rigid rods and lower strength targets contains fewer mechanism and is easier to model.

In our recent paper (Forrestal et al., 1987), we developed an engineering model to predict residual velocity and ballistic-limit for nondeforming, conical-nosed rods that perforate ductile plates at normal incidence. This model is in good agreement with one set of experiments conducted at impact velocities between 0.2–0.7 km/s. To add confidence to our model, we conducted additional experiments with smaller-scale projectiles at impact velocities between 0.3–1.6 km/s. In addition, we cast our model in a form that more easily explains the dominant mechanisms.

Perforation Model

In our previous paper (Forrestal et al., 1987), we developed an engineering model to predict residual velocity V_r and ballistic limit V_{bl} for targets that exhibit ductile hole-growth (Backman and Goldsmith, 1978). This model is derived from a kinetic energy-work balance for a pointed, rigid projectile and only considers work done in the target from radial cavity-expansion. That is, ductile hole-growth is approximated by the cylindrical, cavity-expansion approximation that idealizes the target as thin, independent layers normal to the penetration direction (see Forrestal, 1986). From Forrestal et al., 1987,

$$(m/2)(V_i^2 - V_r^2) = \pi a^2 h \sigma_r \qquad (1)$$

where V_i is the impact velocity. The conical nose projectile has shank radius a and mass m; the target is described by thickness h, Poisson's ratio ν, bulk modulus K, and yield stress Y from the Tresca criterion. Dynamic, radial stress σ_r on the nosetip is derived from a cavity-expansion analysis (Forrestal, 1986) and is available in closed-form.

[1]University of Dayton Research Institute, Dayton, OH 46469; currently, Armament Development Authority, P.O. Box 2250 (24), Haifa, Israel.

[2]Sandia National Laboratories, Albuquerque, NM 87185. Fellow ASME.

Manuscript received by ASME Applied Mechanics Division, October 27, 1986; final revision September 24, 1987.

[3]Also summarized by Zukas (1982).

$$K=\lim_{x\to 0^+}\ \mu(x,0)\sqrt{x}\,\frac{\partial\omega}{\partial y}(x,0)$$

and from equation (12) one has

$$K=\frac{-\mu_0}{\pi}\left\{\int_{-\infty}^{0}\frac{1}{\sqrt{|s|}}\,\frac{P(s)}{\mu(s,0)}\,ds\right.$$

$$\left.+\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{|\tau|}}\,\frac{t}{(\tau-s)^2+t^2}\,\frac{\nabla\mu\cdot\nabla\omega}{\mu}\,d\tau dt ds\right\}.$$

As pointed out above, the calculation of σ_{xz} is carried out in a similar manner and by changing to polar coordinates it is not difficult to show that in general $\sigma_{ij}(r,\theta) = 0\,(1/\sqrt{r})$, $r \to 0$. It is also worth noting that a similar derivation can be carried out for a finite domain. In that case the appropriate boundary conditions must be prescribed and the analogous Green's function computed. The growth condition imposed on ω, however, would no longer be necessary.

A number of extensions and generalizations of the above result are possible. Clearly, the assumed symmetry of the shear modulus about the crack plane is not essential for the stress singularity to have order 1/2. Rather, it merely simplifies the analysis by permitting reduction to a half plane problem for which a simple, explicit expression for the Green's function, Γ, is easily constructed. When symmetry is not assumed it is still possible to represent the displacement $\omega(x, y)$ by integrals of the form (3) in each of the half spaces $y \gtrless 0$, only now the Green's function Γ must satisfy more complicated boundary conditions derived from imposing continuity of stresses and the displacement along $y = 0, x > 0$ and the applied crack face traction $P(x)$ on $y = 0, x < 0$. Once the result has been established for a nonsymmetric shear modulus, the extension to a possibly curved (smooth) crack normally incident to a possibly curved (smooth) interface is easy, since consideration of an infinitessimally small neighborhood about the crack tip produces a straight crack normally incident to a straight interface. Finally, a major modification of the analysis presented here seems required to study the case of a crack that is obliquely incident to the interface.

References

Atkinson, C., 1977, "On Stress Singularities and Interfaces in Linear Elastic Fracture Mechanics," *International Journal of Fracture*, Vol. 13, pp. 807–820.

Clements, D. L., Atkinson, C., and Rogers, C., 1978, "Antiplane Crack Problems for an Inhomogeneous Elastic Material," *Acta Mechanica*, Vol. 20, pp. 199–211.

Cook, T. S., and Erodgan, F., 1972, "Stresses in Bonded Materials with a Crack Perpendicular to the Interface," *International Journal of Engineering Science*, Vol. 10, pp. 667–697.

Delale, F., 1985, "Mode III Fracture of Bonded Nonhomogeneous Materials," *Engineering Fracture Mechanics*, Vol. 22, pp. 213–226.

Delale, F., and Erdogan, F., 1983, "The Crack Problem for a Nonhomogeneous Plane," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 609–614.

Dhaliwal, R. S., and Singh, B. M., 1978, "On the Theory of Elasticity of a Nonhomogeneous Medium," *Journal of Elasticity*, Vol. 8, pp. 211–219.

Erdogan, F., 1985, "The Crack Problem for Bonded Nonhomogeneous Materials Under Antiplane Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 823–828.

Gakov, F. D., 1966, *Boundary Value Problems*, Pergamon, London.

Gerasoulis, A,. and Srivastav, R. O., 1980, "A Griffith Crack Problem for a Nonhomogeneous Medium," *International Journal of Engineering Science*, Vol. 18, pp. 239–247.

Schovanec, L., 1986, "A Griffith Crack Problem for an Inhomogeneous Elastic Material," *Acta Mechanica*, Vol. 58, pp. 67–80.

Schovanec, L., and Walton, J. R., 1987, "The Quasi-Static Propagation of a Plane Strain Crack in a Power-Law Inhomogeneous Linearly Viscoelastic Body," *Acta Mechanica*, Vol. 66, No. 1, pp.

Willis, J. R., 1967, "Crack Propagation in Viscoelastic Media," *Journal of Mechanics and Physics of Solids*, Vol. 15, pp. 229–240.

Perforation of Aluminum Plates with Conical-Nosed Rods – Additional Data and Discussion

Z. Rosenberg[1] and M. J. Forrestal[2]

Introduction

Backman and Goldsmith (1978) discuss many analytical and experimental methods used to study the broad field of penetration mechanics. For plate perforation, they show eight possible mechanisms that depend on the geometry and material of the projectiles and targets. Because many perforation mechanisms are possible, experimental observations usually precede and guide engineering models. Our study is limited to high strength, conical-nosed projectiles that perforate low strength, ductile targets. For the experiments conducted in this study and reported in a recent paper (Forrestal et al., 1987), targets were 6061-T6 aluminum plates and projectiles were machined from T-200 maraging steel. As shown in our recent paper (1987), the targets experienced ductile hole-growth, and post-perforation, x-ray photographs showed that the projectile remained undeformed.

By contrast to our studies, most work on perforation by long rods has been conducted with high strength, steel targets. For example, Grabarek (1973)[3] launched steel rods against rolled-homogeneous, armor plates and showed, via x-ray photographs, massive nose deformation and rod fracturing. Clearly, the experimental work presented in our study with rigid rods and lower strength targets contains fewer mechanism and is easier to model.

In our recent paper (Forrestal et al., 1987), we developed an engineering model to predict residual velocity and ballistic-limit for nondeforming, conical-nosed rods that perforate ductile plates at normal incidence. This model is in good agreement with one set of experiments conducted at impact velocities between 0.2–0.7 km/s. To add confidence to our model, we conducted additional experiments with smaller-scale projectiles at impact velocities between 0.3–1.6 km/s. In addition, we cast our model in a form that more easily explains the dominant mechanisms.

Perforation Model

In our previous paper (Forrestal et al., 1987), we developed an engineering model to predict residual velocity V_r and ballistic limit V_{bl} for targets that exhibit ductile hole-growth (Backman and Goldsmith, 1978). This model is derived from a kinetic energy-work balance for a pointed, rigid projectile and only considers work done in the target from radial cavity-expansion. That is, ductile hole-growth is approximated by the cylindrical, cavity-expansion approximation that idealizes the target as thin, independent layers normal to the penetration direction (see Forrestal, 1986). From Forrestal et al., 1987,

$$(m/2)(V_i^2 - V_r^2) = \pi a^2 h \sigma_r \qquad (1)$$

where V_i is the impact velocity. The conical nose projectile has shank radius a and mass m; the target is described by thickness h, Poisson's ratio ν, bulk modulus K, and yield stress Y from the Tresca criterion. Dynamic, radial stress σ_r on the nosetip is derived from a cavity-expansion analysis (Forrestal, 1986) and is available in closed-form.

[1]University of Dayton Research Institute, Dayton, OH 46469; currently, Armament Development Authority, P.O. Box 2250 (24), Haifa, Israel.

[2]Sandia National Laboratories, Albuquerque, NM 87185. Fellow ASME.

Manuscript received by ASME Applied Mechanics Division, October 27, 1986; final revision September 24, 1987.

[3]Also summarized by Zukas (1982).

Table 1 Data summary for the projectile shown in Fig. 1(*a*). The nosetip for shots 849 and 945 deformed. The projectile for shot 963 was made from C-300 maraging steel and this nosetip did not deform

Shot Number	Mass (g)	V_i (km/s)	V_r (km/s)
994	25.1	0.314	0
992	25.1	0.315	0
995	25.1	0.327	0
993	25.1	0.383	0.175
991	25.1	0.419	0.260
990	25.1	0.515	0.399
971	25.1	0.886	0.827
848	24.0	1.394	1.334
849	24.0	1.442	1.397
945	24.0	1.516	1.445
963	25.3	1.575	1.509

Table 2 Data summary for the arrowhead projectile shown in Fig. 1(*b*)

Shot Number	Mass (g)	V_i (km/s)	V_r (km/s)
988	23.4	0.301	0
989	23.4	0.315	0.040
985	23.6	0.360	0.175
987	23.5	0.421	0.282
983	23.7	0.532	0.433

The ballistic limit is the impact velocity below which the projectile fails to perforate the target. The ballistic limit V_{bl} is obtained from equation (1) with $V_r = 0$ and the quasi-static, radial stress σ_s (the value of σ_r for the radial, cavity-expansion velocity approaching zero). Thus, an estimate for the ballistic limit is given by

$$V_{bl} = (2\pi a^2 h \sigma_s / m)^{1/2} \tag{2a}$$

$$\sigma_s = \frac{Y}{2}\left\{1 + \ell n \left[\frac{6(1-2\nu)K}{(5-4\nu)Y}\right]\right\} \tag{2b}$$

where σ_s was derived by Bishop, Hill and Mott (1945).[4] From equations (1) and (2), we obtain an expression for residual velocity

$$V_r = V_{bl}\left[\left(\frac{V_i}{V_{bl}}\right)^2 - \frac{\sigma_r}{\sigma_s}\right]^{1/2} \tag{3}$$

Recht and Ipson (1983) propose an empirical equation for V_r that requires an experimentally determined value for V_{bl}. Their equation is similar in form to equation (3), and when $\sigma_r = \sigma_s$, their equation is identical to equation (3). In this study, we present data for $V_i \leq 1.6$ km/s, and for $V_i = 1.6$ km/s, $\sigma_r/\sigma_s = 3$. Thus, the dynamic, radial stress from an elastic-plastic, wave-propagation calculation can be much larger than the quasi-static value. However, as σ_r/σ_s increases from one to three, $(V_i/V_{bl})^2$ increases from one to 35. Thus, the increase in σ_r/σ_s has a small effect on the magnitude[5] of V_r. As we show later, the predicted value of V_{bl} is in good agreement with measurements, so equations (2*b*) and (2*c*) and the Recht-Ipson equation provide a predictive tool prior to experimentation. In summary, to close approximation σ_r/σ_s can be replaced with unity in equation (3), and residual velocity is predicted in terms of the ballistic limit (2*a*), (2*b*) and the impact velocity.

[4] Also in the text by Hill (1950).

[5] Forrestal et al. (1987) present a procedure to calculate V_r that takes into account the dynamic, radial stress σ_r; so, the effect of σ_r/σ_s on V_r can be calculated easily for other parameters.

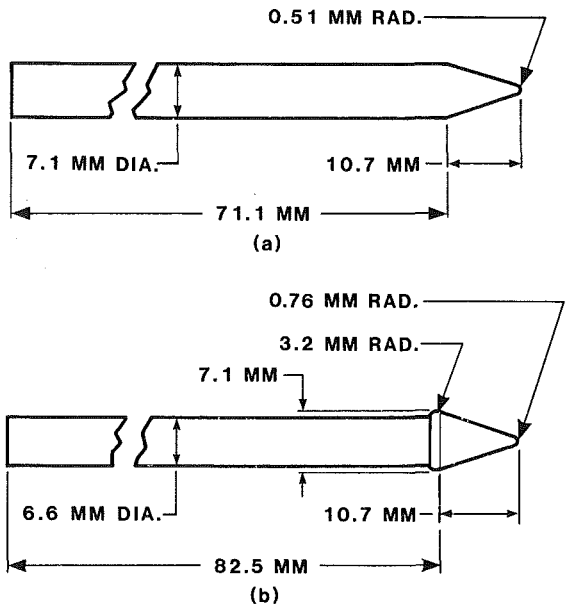


Fig. 1 Projectile geometries

Experiments

We describe the experimental arrangement and measurements in our report (Forrestal et al., 1986). Briefly, a powder-gun launched conical nose rods at aluminum target-plates. Laser beams determined incident velocity, and flash x-ray photography measured residual velocity and yaw angles.

Targets. The square targets were 0.305-m-wide (12 in.), 25.4-mm-thick (1.0 in.), 6061-T6 aluminum plates. Tension and compression tests were conducted on samples in the plane of the plate. For strain rates between 10^{-2} and 10^{0}, we measured an average flow stress, $Y = 300$ MPa (44.1 ksi).

Projectiles. We tested the projectile shapes shown in Figs. 1(*a*) and 1(*b*) that were machined from T-200 and C-300 maraging steel[6] and had nominal mass, $m = 25$ g. For impact velocities below about 0.6 km/s, the target rebounded after nose entry and created a friction drag on the shank of the rod shown in Fig. 1(*a*). Since our model predicts only forces on the nose, we avoided shank drag by using the arrowhead projectile shown in Fig. 1(*b*). Residual velocities are found for both projectiles, so the magnitude of shank resistance was determined experimentally.

Results. Table 1 lists the data for the projectile shown in Fig. 1(*a*). For shot numbers 849 and 945 with T-200 projectiles, post-perforation shadowgraphs (Fig. 2) showed blunting of the conical nosetips; whereas, the shadowgraph for shot number 848, with a slightly lower impact velocity, showed no permanent nosetip deformation. One experiment (shot number 963) was conducted at $V_i = 1.575$ km/s with a C-300 maraging steel projectile that had a yield stress of about 300 ksi. The post-perforation shadowgraph for this experiment showed no permanent nosetip deformation. Table 2 lists the data for the arrowhead projectile shown in Fig. 1(*b*).

Figure 3 compares residual-velocity predictions from equations (2*a*) and (3) with σ_r/σ_s taken as unity and the data obtained in this study. Although our model considers the projectiles as rigid, data from the projectiles with slight, permanent nosetip deformations were included in Fig. 3. However, for $V_i > 1.6$, the entire nose could deform such that its maximum deformed diameter would be larger than the shank diameter

[6] Vasco Pacific, 707 West Olympic Blvd., Montebello, CA 90640. T-200 and C-300 have nominal yield stresses of 200 ksi and 300 ksi, respectively.

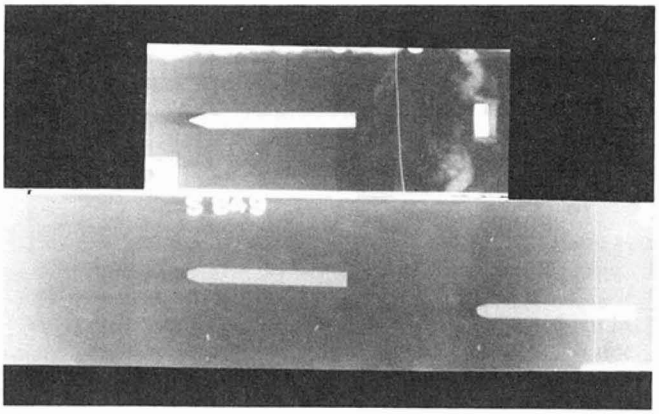


Fig. 2 Top photograph shows projectile before perforation; bottom photograph shows projectile after perforation (shot number 849)

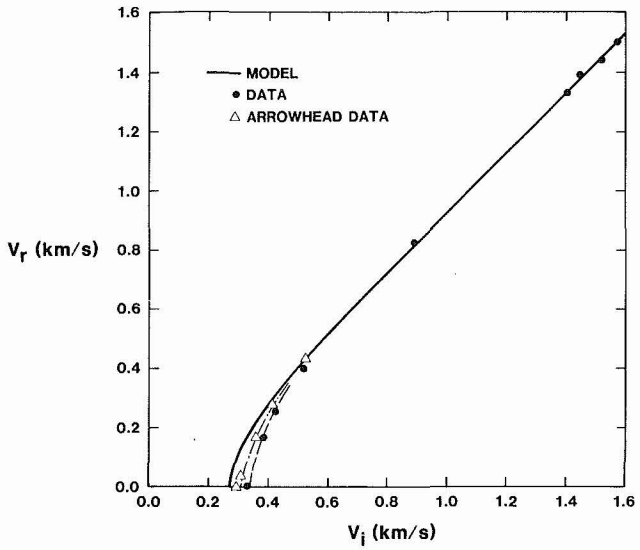


Fig. 3 Predicted and measured residual velocity

(see Zukas, 1982). At these higher impact velocities, one could expect a departure in agreement between our model and residual-velocity measurements. The calculated ballistic limit is 0.266 km/s, and the ballistic limit determined by curve fitting the arrowhead data is 0.315 km/s. As mentioned, the prediction used the Tresca yield criterion; the predicted ballistic limit for the Mises[7] yield criterion is 0.282 km/s. Figure 3 also shows that the arrowhead projectiles have slightly less resistance for $V_i < 0.60$ km/s.

Summary

We performed experiments with the projectiles shown in Fig. 1 to impact velocities of 1.6 km/s. Figure 3 shows good agreement with our engineering model and also shows the effect of shank drag for $V_i < 0.60$ km/s. In addition, post-perforation shadowgraphs showed that the T-200 maraging steel penetrators exhibited permanent, nosetip deformation between $V_i = 1.394$–1.442 km/s, whereas the C-300 maraging steel penetrator showed no permanent deformation for $V_i = 1.575$ km/s.

Acknowledgment

The authors thank Mark Laber for assisting with the terminal-ballistic experiments. This work was supported by the US Department of Energy.

[7] As discussed by Hill (1950), the Mises criterion increases the yield stress by 15 percent.

References

Backman, M. E., and Goldsmith, W., 1978, "The Mechanics of Penetration of Projectiles into Targets," *International Journal of Engineering Science*, Vol. 16, p. 5.

Bishop, R. F., Hill, R., and Mott, N. F., 1945, "The Theory of Indentation and Hardness Tests," *The Proceedings of the Physical Society*, Vol. 57, Part 3, pp. 147–155.

Forrestal, M. J., 1986, "Penetration into Dry Porous Rock," *International Journal of Solids and Structures*, Vol. 22, No. 12, pp. 1485–1500.

Forrestal, M. J., Rosenberg, Z., Luk, V. K., and Bless, S. J., 1986, "Perforation of Aluminum Plates with Conical-Nosed Rods," SAND86-0292J, Sandia National Laboratories, Albuquerque, NM.

Forrestal, M. J., Rosenberg, Z., Luk, V. K., and Bless, S. J., 1987, ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 230–232.

Hill, R., 1950, *The Mathematical Theory of Plasticity*, Oxford, London.

Recht, R. F., and Ipson, T. W., 1963, "Ballistic Perforation Dynamics," ASME JOURNAL OF APPLIED MECHANICS, Sept., pp. 384–390.

Zukas, J. A., et al., 1982, *Impact Dynamics*, Wiley, New York, pp. 202–205.

Effect of an Elastic Foundation on the Buckling and Postbuckling of Delaminated Composites Under Compressive Loads

G. A. Kardomateas[1]

Introduction

Consider a composite boxed beam filled with a soft elastic medium such as foam or a sandwich beam consisting of two fiber-reinforced sheets separated by a low stiffness core. A bending load on these structures is equivalent to a compressive force on one face and a tensile force on the other; furthermore, a delamination may be present on the compressively loaded composite face. In those cases the composite face rests on an elastic "foundation" which imposes reaction forces on the beam that are proportional to the deflection of the "foundation."

Studies of the delamination problem have been undertaken for the usual case without the elastic foundation (e.g., Chai et al., 1981; Yin et al., 1986; Bottega and Maewal, 1983). In a recent study by Kardomateas and Schmueser (1987) the perturbation technique was used to obtain an analytical expression for the initial postbuckling deflections. This solution is extended here for the case of a beam/plate on an elastic foundation. Analytical solutions for the critical load and the initial postbuckling behavior will be derived.

Analysis

The configuration consists of a homogeneous, orthotropic beam-plate of thickness T, length L and unit width, containing a single delamination of length $\ell = 2a$ and at depth H from the top surface of the plate. The plate has a permanently attached Winkler-type elastic foundation. Over the delamination region, the laminate consists of two parts, the part above the delamination, of thickness H, referred to as the "upper" part, and the part below the delamination, of thickness $T-H$, referred to as the "lower" part. The remaining laminate outside the delamination interval and of thickness T is referred to as the "base" laminate. Local coordinate systems with the origin at the left end of each part are assumed. These parts have a common section referred to as the "interface section." The corresponding axial and shearing forces and moments at this section for the different parts are denoted by P_i, V_i, M_i.

Although the differential equation for the deflections of the

[1] General Motors Research Laboratories, Engineering Mechanics Department, Warren, MI 48090-9055.

Manuscript received by ASME Applied Mechanics Division, February 27, 1987; final revision August 10, 1987.

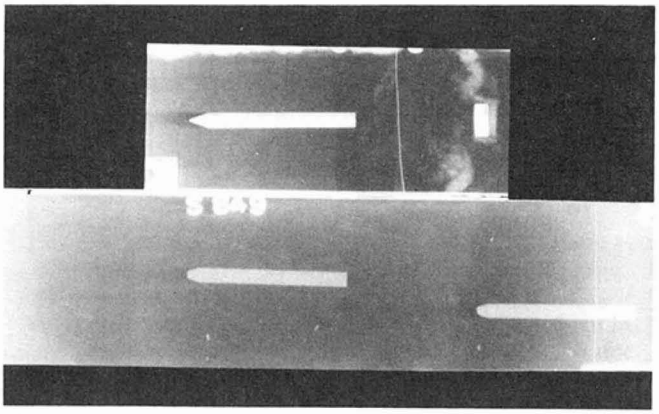


Fig. 2 Top photograph shows projectile before perforation; bottom photograph shows projectile after perforation (shot number 849)

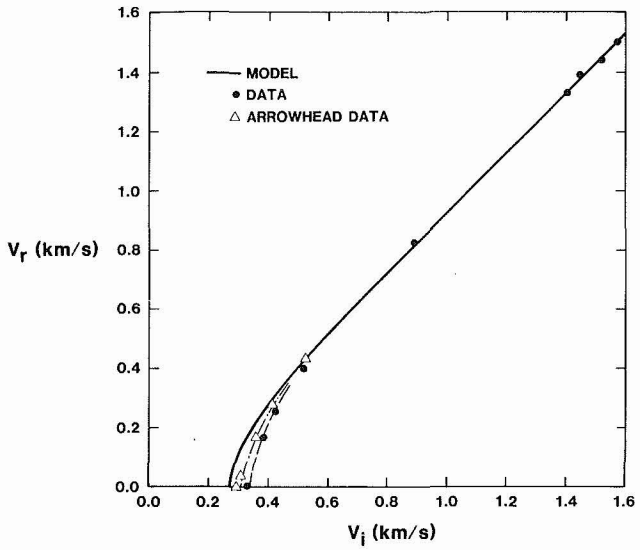


Fig. 3 Predicted and measured residual velocity

(see Zukas, 1982). At these higher impact velocities, one could expect a departure in agreement between our model and residual-velocity measurements. The calculated ballistic limit is 0.266 km/s, and the ballistic limit determined by curve fitting the arrowhead data is 0.315 km/s. As mentioned, the prediction used the Tresca yield criterion; the predicted ballistic limit for the Mises[7] yield criterion is 0.282 km/s. Figure 3 also shows that the arrowhead projectiles have slightly less resistance for $V_i < 0.60$ km/s.

Summary

We performed experiments with the projectiles shown in Fig. 1 to impact velocities of 1.6 km/s. Figure 3 shows good agreement with our engineering model and also shows the effect of shank drag for $V_i < 0.60$ km/s. In addition, post-perforation shadowgraphs showed that the T-200 maraging steel penetrators exhibited permanent, nosetip deformation between $V_i = 1.394$–1.442 km/s, whereas the C-300 maraging steel penetrator showed no permanent deformation for $V_i = 1.575$ km/s.

Acknowledgment

The authors thank Mark Laber for assisting with the terminal-ballistic experiments. This work was supported by the US Department of Energy.

[7] As discussed by Hill (1950), the Mises criterion increases the yield stress by 15 percent.

References

Backman, M. E., and Goldsmith, W., 1978, "The Mechanics of Penetration of Projectiles into Targets," *International Journal of Engineering Science*, Vol. 16, p. 5.

Bishop, R. F., Hill, R., and Mott, N. F., 1945, "The Theory of Indentation and Hardness Tests," *The Proceedings of the Physical Society*, Vol. 57, Part 3, pp. 147–155.

Forrestal, M. J., 1986, "Penetration into Dry Porous Rock," *International Journal of Solids and Structures*, Vol. 22, No. 12, pp. 1485–1500.

Forrestal, M. J., Rosenberg, Z., Luk, V. K., and Bless, S. J., 1986, "Perforation of Aluminum Plates with Conical-Nosed Rods," SAND86-0292J, Sandia National Laboratories, Albuquerque, NM.

Forrestal, M. J., Rosenberg, Z., Luk, V. K., and Bless, S. J., 1987, ASME Journal of Applied Mechanics, Vol. 54, pp. 230–232.

Hill, R., 1950, *The Mathematical Theory of Plasticity*, Oxford, London.

Recht, R. F., and Ipson, T. W., 1963, "Ballistic Perforation Dynamics," ASME Journal of Applied Mechanics, Sept., pp. 384–390.

Zukas, J. A., et al., 1982, *Impact Dynamics*, Wiley, New York, pp. 202–205.

Effect of an Elastic Foundation on the Buckling and Postbuckling of Delaminated Composites Under Compressive Loads

G. A. Kardomateas[1]

Introduction

Consider a composite boxed beam filled with a soft elastic medium such as foam or a sandwich beam consisting of two fiber-reinforced sheets separated by a low stiffness core. A bending load on these structures is equivalent to a compressive force on one face and a tensile force on the other; furthermore, a delamination may be present on the compressively loaded composite face. In those cases the composite face rests on an elastic "foundation" which imposes reaction forces on the beam that are proportional to the deflection of the "foundation."

Studies of the delamination problem have been undertaken for the usual case without the elastic foundation (e.g., Chai et al., 1981; Yin et al., 1986; Bottega and Maewal, 1983). In a recent study by Kardomateas and Schmueser (1987) the perturbation technique was used to obtain an analytical expression for the initial postbuckling deflections. This solution is extended here for the case of a beam/plate on an elastic foundation. Analytical solutions for the critical load and the initial postbuckling behavior will be derived.

Analysis

The configuration consists of a homogeneous, orthotropic beam-plate of thickness T, length L and unit width, containing a single delamination of length $\ell = 2a$ and at depth H from the top surface of the plate. The plate has a permanently attached Winkler-type elastic foundation. Over the delamination region, the laminate consists of two parts, the part above the delamination, of thickness H, referred to as the "upper" part, and the part below the delamination, of thickness $T-H$, referred to as the "lower" part. The remaining laminate outside the delamination interval and of thickness T is referred to as the "base" laminate. Local coordinate systems with the origin at the left end of each part are assumed. These parts have a common section referred to as the "interface section." The corresponding axial and shearing forces and moments at this section for the different parts are denoted by P_i, V_i, M_i.

Although the differential equation for the deflections of the

[1] General Motors Research Laboratories, Engineering Mechanics Department, Warren, MI 48090-9055.

Manuscript received by ASME Applied Mechanics Division, February 27, 1987; final revision August 10, 1987.

upper delaminated layer have the usual form (e.g., Kardomateas and Schmueser, 1987), for the lower part and the base plate there is an additional term due to the elastic foundation. In terms of the modulus of the foundation, β:

$$D_i \frac{d^4 y_i}{dx^4} + P_i \frac{d^2 y_i}{dx^2} = -\beta y_i, \tag{1}$$

where $D_i = E_1 t_i^3/[12(1-\nu_{13}\nu_{31})]$ are the bending stiffnesses (E_1 is the modulus of elasticity in the axial $\equiv 1$ direction; ν_{13}, ν_{31} are Poisson's ratio's where 3 is the in-plane transverse direction and t_i is the thickness of the corresponding part). A condition of common deflection ζ at the interface section should be satisfied:

$$y_u|_{x=0,\ell} = y_l|_{x=0,\ell} = y_b|_{x=\ell_0} = \zeta. \tag{2}$$

Force and moment equilibrium at this section give:

$$P_u + P_l = P_b = P;\ V_u + V_l = V_b;\ M_u + M_l + P_u\left(\frac{T-H}{2}\right) - P_l\left(\frac{H}{2}\right) = M_b. \tag{3}$$

Furthermore, the shortening due to the deflections of the upper and lower parts should be geometrically compatible, which is expressed as:

$$(1-\nu_{13}\nu_{31})\frac{P_u \ell}{A_u E_1} - (1-\nu_{13}\nu_{31})\frac{P_l \ell}{A_l E_1} + \frac{1}{2}\int_0^\ell y_u'^2 dx - \frac{1}{2}\int_0^\ell y_l'^2 dx = T y_u'|_{x=0}, \tag{4}$$

where A_u, A_l are the cross-sectional areas of the upper and lower part.

Buckling. The deflection and load quantities at each part are developed into ascending perturbation series with respect to the angle at the interface section, ϕ.

$$y_i(x) = \phi y_{i,1}(x) + \phi^2 y_{i,2}(x) + \ldots;\ P_i = P_{i,0} + \phi P_{i,1} + \phi^2 P_{i,2} + \ldots$$
$$M_i = \phi M_{i,1} + \phi^2 M_{i,2} + \ldots;\ V_i = \phi V_{i,1} + \phi^2 V_{i,2} + \ldots \tag{5}$$

By this definition, at the interface section:

$$y_{i,1}' = 1;\ y_{i,2}' = y_{i,3}' = \ldots = 0. \tag{6}$$

Substituting equations (5) into the differential equation and equating like powers of ϕ leads to a set of linear differential equations and boundary conditions for each part. In the first approximation the terms in the first power of ϕ are equated. For the resulting equation, assuming a solution of the form $y = e^{\gamma x}$ gives for γ either a purely imaginary number or a complex one with a real and an imaginary part, rendering the solution in terms of trigonometric only or a combination of trigonometric and hyperbolic functions depending on the magnitude of the modulus of the foundation.

Define

$$k_{l,0}^2 = P_{l,0}/D_l;\ \lambda_l = \beta/D_l. \tag{7}$$

The solution for the lower part for $k_{l,0}^4 > 4\lambda_l$ is given as:

$$y_{l,1} = \sum_{j=1,2} c_{1j}\cos(\gamma_j a - \gamma_j x);\quad \gamma_{1,2} = \sqrt{|(-k_{l,0}^2 \pm \sqrt{k_{l,0}^4 - 4\lambda_l})/2|}. \tag{8}$$

For $k_{l,0}^4 < 4\lambda_l$, the solution to the first order equation is

$$y_{l,1} = c_{11}\cosh(\gamma_1 a - \gamma_1 x)\cos(\gamma_2 a - \gamma_2 x) + c_{12}\sinh(\gamma_1 a - \gamma_1 x)\sin(\gamma_2 a - \gamma_2 x), \tag{9}$$

where γ_1 and γ_2 are defined as follows

$$r = \sqrt{\lambda_l};\ \theta = \arccos[-k_{l,0}^2/(2r)];\ \gamma_1 = \sqrt{r}\cos(\theta/2);\ \gamma_2 = \sqrt{r}\sin(\theta/2). \tag{10}$$

The constants c_{1j} are found from equations (6) in terms of the common deflection ζ_1 from equation (2). The first order end shear can be expressed in the form

$$V_{l,1} = -D_l(y_{l,1}''' + k_{l,0}^2 y_{l,1}')|_{x=0} = V_{l,1}^c + \zeta_1 V_{l,1}^f, \tag{11}$$

and the first order end moment is

$$M_{l,1} = -D_l y_{l,1}''|_{x=0} = M_{l,1}^c + \zeta_1 M_{l,1}^f, \tag{12}$$

where the quantities $V_{l,1}^c$, $V_{l,1}^f$, $M_{l,1}^c$, $M_{l,1}^f$ are given in terms of γ_j. Analogous quantities are found for the base plate (subscript b).

The condition of shear equilibrium (3), written for the first order terms, produces an expression for ζ_1 (note that $V_{u,1} = 0$),

$$\zeta_1 = (V_{l,1}^c - V_{b,1}^c)/(V_{b,1}^f - V_{l,1}^f). \tag{13}$$

Writing the associated moment equilibrium and compatibility equations (3), (4), for the terms in ϕ^1, substituting the end moments from equation (12) and the quantity ζ_1 from equation (13), and eliminating the quantity $P_{l,1}H/2 - P_{u,1}(T-H)/2$, gives an equation for the critical load, P_0 (characteristic equation), as follows

$$D_u k_{u,0}\cot k_{u,0} a + M_{l,1}^c - M_{b,1}^c + (V_{l,1}^c - V_{b,1}^c)(M_{l,1}^f - M_{b,1}^f)/(V_{b,1}^f - V_{l,1}^f) = -TE_1 H(T-H)/[4a(1-\nu_{13}\nu_{31})]. \tag{14}$$

Postbuckling. When the terms in ϕ^2 are equated, the following differential equation for the lower part and the base plate is obtained from equation (1):

$$D_i y_{i,2}^{(4)} + P_{i,0} y_{i,2}'' + \beta y_{i,2} = -P_{i,1} y_{i,1}''. \tag{15}$$

For $k_{l,0}^4 > 4\lambda_l$, with the definition (8) for γ_j, the solution for the lower part is:

$$y_{l,2} = \sum_{j=1,2} c_{2j}\frac{P_{l,1}}{D_l}\cos(\gamma_j a - \gamma_j x) + b_{2j}\frac{P_{l,1}}{D_l}(x-a)\sin(\gamma_j a - \gamma_j x);\ b_{2j} = \frac{(-1)^j c_{1j}\gamma_j}{2(\gamma_2^2 - \gamma_1^2)}. \tag{16}$$

For $k_{l,0}^4 < 4\lambda_l$, the solution to equation (15), and with the definition (10) for γ_j, is found to be:

$$y_{l,2} = c_{21}\frac{P_{\ell,1}}{D_\ell}\cosh(\gamma_1 a - \gamma_1 x)\cos(\gamma_2 a - \gamma_2 x) + c_{22}\frac{P_{\ell,1}}{D_\ell}\sinh(\gamma_1 a - \gamma_1 x)\sin(\gamma_2 a - \gamma_2 x) + b_{21}\frac{P_{\ell,1}}{D_\ell}(x-a)\cosh(\gamma_1 a - \gamma_1 x)\sin(\gamma_2 a - \gamma_2 x) + b_{22}\frac{P_{\ell,1}}{D_\ell}(x-a)\sinh(\gamma_1 a - \gamma_1 x)\cos(\gamma_2 a - \gamma_2 x),$$

$$b_{2j} = \frac{(\gamma_2^2 - \gamma_1^2)[c_{12}\gamma_{3-j} + (-1)^j c_{11}\gamma_j] + 2\gamma_1\gamma_2[c_{11}\gamma_{3-j} + (-1)^{3-j}c_{12}\gamma_j]}{8\gamma_1\gamma_2(\gamma_1^2 + \gamma_2^2)}. \tag{17}$$

The constants c_{2j} can be found from equation (6) in terms of ζ_2. Furthermore, the second order end shear can be expressed as:

$$V_{l,2} = -D_l(y_{l,2}''' + k_{l,0}^2 y_{l,2}')|_{x=0} = P_{l,1}V_{l,2}^g + \zeta_2 V_{l,2}^f, \tag{18}$$

and the second order moment is written in the form

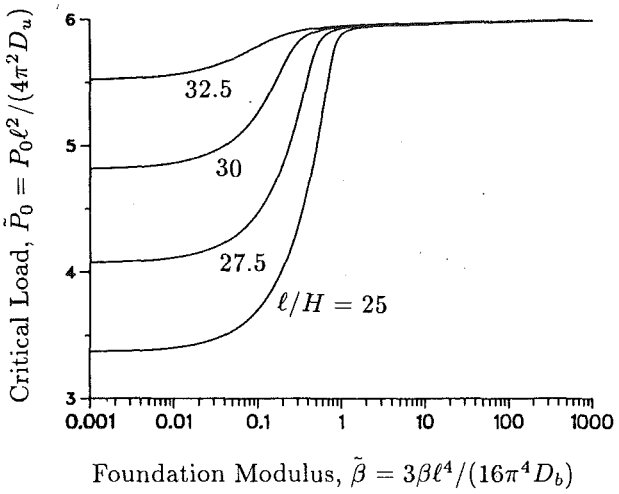


Fig. 1 Critical (buckling) force versus foundation modulus for a set of delamination lengths

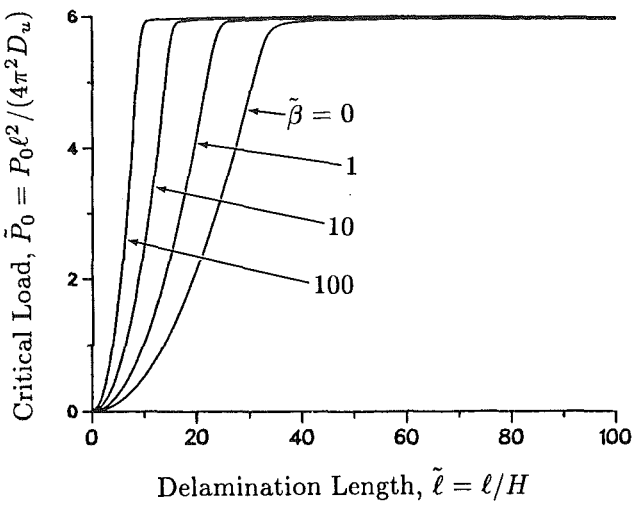


Fig. 2 Critical load versus delamination length for a set of foundation moduli

$$M_{l,2} = -D_l y''_{l,2}|_{x=0} = P_{l,1} M^g_{l,2} + \zeta_2 M^f_{l,2}, \tag{19}$$

where the quantities $V^g_{l,2}$, $V^f_{l,2}$, $M^g_{l,2}$, $M^f_{l,2}$ are given in terms of γ_j. Analogous quantities are found for the base plate (subscript b). An additional quantity needed to solve the problem is the shortening due to the first order deflections, which apears when equation (4) is written for the terms in ϕ^2, $S_l = 1/2 \int_0^\ell y'^2_{l,1} dx$, and which is already known since the first order solution has been obtained.

Now the shear condition (3) allows determining ζ_2 in terms of the first order (yet unknown) forces (note that $V_{u,2} = P_{u,1}$) as follows:

$$\zeta_2 = \frac{V^g_{b,2} - 1}{V^f_{l,2} - V^f_{b,2}} P_{u,1} + \frac{V^g_{b,2} - V^g_{l,2}}{V^f_{l,2} - V^f_{b,2}} P_{l,1}. \tag{20}$$

Writing the moment equilibrium equation (3) and the geometric compatibility equation (4) for the terms in ϕ^2, expressing the second order moments at the interface from equation (19), in terms of the (yet undetermined) first order end forces, taking into account equation (20), and eliminating the quantity $P_{l,2}H/2 - P_{u,2}(T-H)/2$ gives the following equation for the first order forces $P_{u,1}$ and $P_{l,1}$:

$$\left[M^g_{u,2} - M^g_{b,2} + \frac{V^g_{b,2} - 1}{V^f_{l,2} - V^f_{b,2}} (M^f_{l,2} - M^f_{b,2})\right] P_{u,1}$$
$$+ \left[M^g_{l,2} - M^g_{b,2} + \frac{V^g_{b,2} - V^g_{l,2}}{V^f_{l,2} - V^f_{b,2}} (M^f_{l,2} - M^f_{b,2})\right] P_{l,1} =$$
$$= \left(\frac{2k_{u,0}a - \sin 2k_{u,0}a}{4k_{u,0}\sin^2 k_{u,0}a} - S_l\right) \frac{E_1 H(T-H)}{4a(1-\nu_{13}\nu_{31})}. \tag{21}$$

The second equation needed for finding $P_{u,1}$, $P_{l,1}$ is the first order moment equilibrium equation at the interface, namely,

$$P_{l,1}H/2 - P_{u,1}(T-H)/2 = F_0(P_{u,0}, P_{l,0}), \tag{22}$$

where F_0 is the left-hand side of equation (14) which depends only on the zero order quantities. The above system of linear equations allows finding $P_{u,1}$ and $P_{l,1}$ and hence the first order applied end force $P_1 = P_{u,1} + P_{l,1}$. The solution to higher order terms can be obtained in a similar fashion.

The postcritical characteristics are studied next. For this purpose we use the expression derived by Yin and Wang (1984) for the energy release rate of a one-dimensional delamination in terms of the axial forces and bending moments acting across the various cross sections adjacent to the tip of the delamination (these quantities are directly determined from the above postbuckling solution).

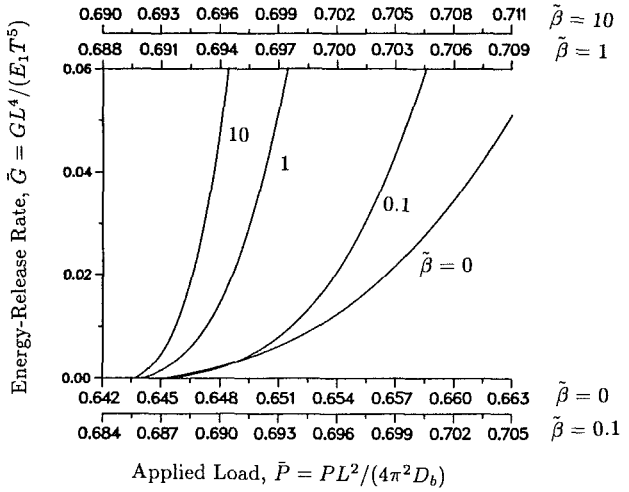


Fig. 3 Strain energy release rate versus applied compressive force during the initial postbuckling stage for a delamination length $\ell/H = 40$ and a set of foundation moduli. The different load scales are of the same length and correspond to the different initial buckling loads.

Discussion of Results

Numerical examples are presented for the case of a clamped-clamped plate with $H/T = 1/6$, $L/H = 200$. Figure 1 shows the variation of the critical load, normalized with respect to the Euler load for the delaminated layer, $\tilde{P}_0 = P_0\ell^2/(4\pi^2 D_u)$, with the foundation modulus which is normalized as $\tilde{\beta} = 3\beta\ell^4/(16\pi^4 D_b)$ (on a logarithmic scale). As $\tilde{\beta}$ increases, the critical load increases, the effect being bigger on the smaller delaminations. Notice that there is a small range of values for the foundation modulus, for which the critical load undergoes a rather significant increase. Two limiting cases are the "global" buckling, characterized by buckling of the composite beam-plate as a whole, and "local" buckling, characterized by deflections of only the delaminated layer, the rest of the plate remaining flat (typical of long and thin delaminations). Figure 2 shows the variation of the critical load with delamination length for a set of values for $\tilde{\beta}$. As $\tilde{\beta}$ increases, the curves are shifted to the left, indicating the attainment of loads similar in magnitude to the local buckling ones for smaller delamination lengths.

The variation during the initial postbuckling stage of the normalized strain energy release rate, $\tilde{G} = G/(E_1T^5/L^4)$, and applied load normalized with respect to the Euler load for the entire beam with no elastic foundation, $\tilde{P} = PL^2/(4\pi^2 D_b)$, is

plotted in Fig. 3 for the case of delamination length $\ell/H = 40$ and for a set of values of the foundation modulus. Since the critical load changes with $\tilde{\beta}$, different scales (of the same length) are used on the load axis. The important thing to observe is that the curves are steeper for a larger $\tilde{\beta}$. This increased slope means that delamination growth will occur sooner and that there will be potentially more energy absorbed since the energy released per unit applied load is larger.

References

Bottega, W. J., and Maewal, A., 1983, "Delamination Buckling and Growth in Laminates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 183–189.

Chai, H., Babcock, C. D., and Knauss, W. G., 1981, "One-Dimensional Modelling of Failure in Laminated Plates by Delamination Buckling," *International Journal of Solids and Structures*, Vol. 17, pp. 1069–1083.

Kardomateas, G. A., Schmueser, D. W., 1987, "Effect of Transverse Shearing Forces on Buckling and Postbuckling of Delaminated Composites Under Compressive Loads," *Proceedings, 28th SDM AIAA/ASME/ASCE/AHS Conference*, Monterey, CA, pp. 757–765, April 1987, also to appear in *AIAA Journal*.

Yin, W. -L., Sallam, S. N., and Simitses, G. J., 1986, "Ultimate Axial Load Capacity of a Delaminated Beam-Plate," *AIAA Journal*, Vol. 24, pp. 123–128.

Yin, W. -L., and Wang, J. T. S., 1984, "The Energy-Release Rate in the Growth of a One-Dimensional Delamination," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 939–941.

An Approximate Solution of the Axisymmetric von Karman Equations for a Point-Loaded Circular Plate

A. T. Dolovich,[1] G. W. Brodland,[1] and A. B. Thornton-Trump[2]

Introduction

A closed-form solution of the axisymmetric von Karman equations or their equivalent for a point-loaded circular plate continues to evade researchers. Various kinds of approximate solutions now exist. See Chia (1980). The most popular of these (Vol'mir, 1956 or Timoshenko and Woinowsky-Krieger, 1959) assumes that the *shape* of the plate does not change from that predicted by linear theory even when deflections are large. The load required to produce a given deflection, however, becomes higher as membrane effects become important. Various series-form solutions have also been employed, including perturbation solutions in terms of load (Stippes and Hausrath, 1952; Cherepy, 1960) or central deflection (Chien and Yeh, 1954; Schmidt, 1968). Frakes and Simmonds (1985) used the symbol manipulating program MACSYMA to generate asymptotic solutions, the convergence of which was then improved using Aitken-Shanks transformations. Berger (1955) proposed a modification to the von Karman equations which led to numerous new plate solutions. The axisymmetric point-loaded plate problem was solved by Basuli (1961) using Berger's approach. A numerical technique developed by Brodland (1987) does not rely on the simplifications inherent in the von Karman equations, is valid for arbitrarily large strains and rotations, and thus allows a highly accurate reference solution to be calculated. Unfortunately, those of the above solutions which are for clamped plates have drawbacks which limit their usefulness. Many involve long and complicated mathematical expressions, some produce spurious results when loads are high, and others require considerable computation.

[1]Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

[2]Department of Mechanical Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2.

Manuscript received by ASME Applied Mechanics Division, June 9, 1987; final revision August 31, 1987.

The approximate solution presented here consists of general equations which are easy to use and which compare well with other solutions including the previously unpublished numerical results of Brodland. The transverse deflection is given by an assumed expression which contains a single parameter associated with plate shape. This parameter, β, is determined by minimizing a shear-related residual. Thus, the present analysis takes advantage of the ease with which assumed-form solutions can be used, while ensuring that the *shape change* which occurs with increasing load can be accommodated.

Analysis

Consider a thin, clamped plate of radius a and thickness h made from an isotropic, linearly elastic material (Young's modulus E, Poisson's ratio ν) and subjected to a central point load p. When transverse deflections w are of the order of the plate thickness, deformation is governed by the Kirchhoff nonlinear plate theory as embodied in the two, fourth-order von Karman equations. When deformations are axisymmetric, it is sometimes convenient to use the mathematically equivalent set of three equations given by Way (1934). These equations of axial equilibrium, radial equilibrium and compatibilty can be written, respectively, in dimensionless form as

$$\frac{d}{dR}\left[\frac{1}{R}\frac{d}{dR}\left(R\frac{dW}{dR}\right)\right] = \frac{P}{2\pi R} + S_r^m \frac{dW}{dR}$$

$$\frac{d}{dR}(R\,S_r^m) - S_t^m = 0 \tag{1}$$

$$R\frac{d}{dR}(S_r^m + S_t^m) + 6(1-\nu^2)\left(\frac{dW}{dR}\right)^2 = 0$$

where dimensionless radial position, transverse deflection and centric load are given by

$$R = \frac{r}{a},\quad W = \frac{w}{h},\quad P = \frac{pa^2}{Dh}, \tag{2}$$

dimensionless radial and circumferential (hoop) stresses are

$$S_r^m = \frac{\sigma_r^m a^2 h}{D},\quad S_t^m = \frac{\sigma_t^m a^2 h}{D}, \tag{3}$$

and

$$D = \frac{Eh^3}{12(1-\nu^2)}. \tag{4}$$

For a clamped plate, the associated boundary conditions representing zero dimensionless transverse deflection, slope, and radial deflection at the clamped edge are

$$W_{|R=1} = 0,\ \frac{dW}{dR}_{|R=1} = 0 \text{ and } U_{|R=1} = 0 \tag{5}$$

where

$$U = \frac{ua}{h^2}. \tag{6}$$

The expression chosen for the transverse deflection is

$$W(R) = \frac{W_o}{\beta - 2}(2R^\beta - \beta R^2 + \beta - 2) \tag{7}$$

where W_o is the center deflection and β is a parameter which affects the plate shape. A related form was used by Nadai (1925) to analyze uniformly-loaded plates. The expression given in equation (7) satisfies the boundary conditions $(5)_1$ and $(5)_2$ associated with the transverse deflection W for all values of W_o and β. In addition, it is easy to show using l'Hopital's rule that it approaches the well known linear solution as $\beta \to 2$; i.e.,

plotted in Fig. 3 for the case of delamination length $\ell/H = 40$ and for a set of values of the foundation modulus. Since the critical load changes with $\tilde{\beta}$, different scales (of the same length) are used on the load axis. The important thing to observe is that the curves are steeper for a larger $\tilde{\beta}$. This increased slope means that delamination growth will occur sooner and that there will be potentially more energy absorbed since the energy released per unit applied load is larger.

References

Bottega, W. J., and Maewal, A., 1983, "Delamination Buckling and Growth in Laminates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 183–189.

Chai, H., Babcock, C. D., and Knauss, W. G., 1981, "One-Dimensional Modelling of Failure in Laminated Plates by Delamination Buckling," *International Journal of Solids and Structures*, Vol. 17, pp. 1069–1083.

Kardomateas, G. A., Schmueser, D. W., 1987, "Effect of Transverse Shearing Forces on Buckling and Postbuckling of Delaminated Composites Under Compressive Loads," *Proceedings, 28th SDM AIAA/ASME/ASCE/AHS Conference*, Monterey, CA, pp. 757–765, April 1987, also to appear in *AIAA Journal*.

Yin, W. -L., Sallam, S. N., and Simitses, G. J., 1986, "Ultimate Axial Load Capacity of a Delaminated Beam-Plate," *AIAA Journal*, Vol. 24, pp. 123–128.

Yin, W. -L., and Wang, J. T. S., 1984, "The Energy-Release Rate in the Growth of a One-Dimensional Delamination," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 939–941.

An Approximate Solution of the Axisymmetric von Karman Equations for a Point-Loaded Circular Plate

A. T. Dolovich,[1] G. W. Brodland,[1] and A. B. Thornton-Trump[2]

Introduction

A closed-form solution of the axisymmetric von Karman equations or their equivalent for a point-loaded circular plate continues to evade researchers. Various kinds of approximate solutions now exist. See Chia (1980). The most popular of these (Vol'mir, 1956 or Timoshenko and Woinowsky-Krieger, 1959) assumes that the *shape* of the plate does not change from that predicted by linear theory even when deflections are large. The load required to produce a given deflection, however, becomes higher as membrane effects become important. Various series-form solutions have also been employed, including perturbation solutions in terms of load (Stippes and Hausrath, 1952; Cherepy, 1960) or central deflection (Chien and Yeh, 1954; Schmidt, 1968). Frakes and Simmonds (1985) used the symbol manipulating program MACSYMA to generate asymptotic solutions, the convergence of which was then improved using Aitken-Shanks transformations. Berger (1955) proposed a modification to the von Karman equations which led to numerous new plate solutions. The axisymmetric point-loaded plate problem was solved by Basuli (1961) using Berger's approach. A numerical technique developed by Brodland (1987) does not rely on the simplifications inherent in the von Karman equations, is valid for arbitrarily large strains and rotations, and thus allows a highly accurate reference solution to be calculated. Unfortunately, those of the above solutions which are for clamped plates have drawbacks which limit their usefulness. Many involve long and complicated mathematical expressions, some produce spurious results when loads are high, and others require considerable computation.

[1]Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

[2]Department of Mechanical Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2.

Manuscript received by ASME Applied Mechanics Division, June 9, 1987; final revision August 31, 1987.

The approximate solution presented here consists of general equations which are easy to use and which compare well with other solutions including the previously unpublished numerical results of Brodland. The transverse deflection is given by an assumed expression which contains a single parameter associated with plate shape. This parameter, β, is determined by minimizing a shear-related residual. Thus, the present analysis takes advantage of the ease with which assumed-form solutions can be used, while ensuring that the *shape change* which occurs with increasing load can be accommodated.

Analysis

Consider a thin, clamped plate of radius a and thickness h made from an isotropic, linearly elastic material (Young's modulus E, Poisson's ratio ν) and subjected to a central point load p. When transverse deflections w are of the order of the plate thickness, deformation is governed by the Kirchhoff nonlinear plate theory as embodied in the two, fourth-order von Karman equations. When deformations are axisymmetric, it is sometimes convenient to use the mathematically equivalent set of three equations given by Way (1934). These equations of axial equilibrium, radial equilibrium and compatibilty can be written, respectively, in dimensionless form as

$$\frac{d}{dR}\left[\frac{1}{R}\frac{d}{dR}\left(R\frac{dW}{dR}\right)\right] = \frac{P}{2\pi R} + S_r^m \frac{dW}{dR}$$

$$\frac{d}{dR}(R\,S_r^m) - S_t^m = 0 \tag{1}$$

$$R\frac{d}{dR}(S_r^m + S_t^m) + 6(1-\nu^2)\left(\frac{dW}{dR}\right)^2 = 0$$

where dimensionless radial position, transverse deflection and centric load are given by

$$R = \frac{r}{a},\quad W = \frac{w}{h},\quad P = \frac{pa^2}{Dh}, \tag{2}$$

dimensionless radial and circumferential (hoop) stresses are

$$S_r^m = \frac{\sigma_r^m a^2 h}{D},\quad S_t^m = \frac{\sigma_t^m a^2 h}{D}, \tag{3}$$

and

$$D = \frac{Eh^3}{12(1-\nu^2)}. \tag{4}$$

For a clamped plate, the associated boundary conditions representing zero dimensionless transverse deflection, slope, and radial deflection at the clamped edge are

$$W_{|R=1} = 0,\ \frac{dW}{dR}_{|R=1} = 0 \text{ and } U_{|R=1} = 0 \tag{5}$$

where

$$U = \frac{ua}{h^2}. \tag{6}$$

The expression chosen for the transverse deflection is

$$W(R) = \frac{W_o}{\beta - 2}(2R^\beta - \beta R^2 + \beta - 2) \tag{7}$$

where W_o is the center deflection and β is a parameter which affects the plate shape. A related form was used by Nadai (1925) to analyze uniformly-loaded plates. The expression given in equation (7) satisfies the boundary conditions $(5)_1$ and $(5)_2$ associated with the transverse deflection W for all values of W_o and β. In addition, it is easy to show using l'Hopital's rule that it approaches the well known linear solution as $\beta \to 2$; i.e.,

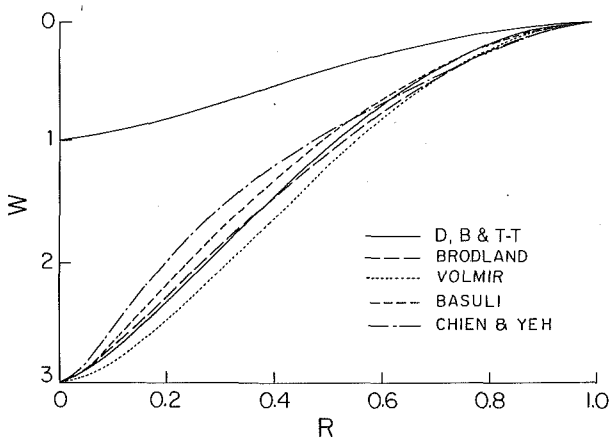


Fig. 1 Transverse deflection profiles

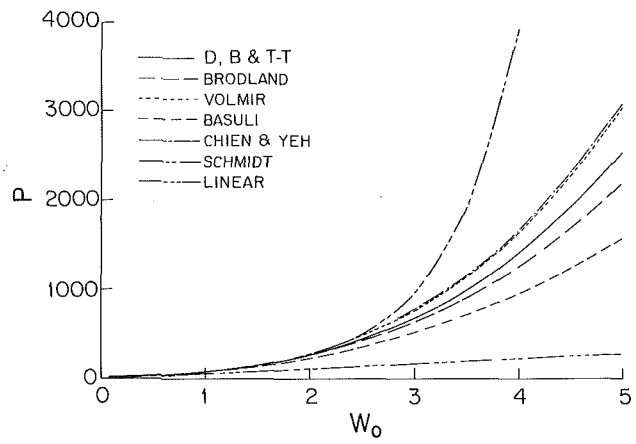


Fig. 2 Load versus center deflection

$$\lim_{\beta \to 2} W = W_o(2R^2 \ln R - R^2 + 1). \tag{8}$$

The parameter β depends on the membrane stresses. To find the radial membrane stress S_r^m, equation $(1)_2$ is solved for S_t^m and this result, together with W from equation (7), is substituted into equation $(1)_3$ to give a second-order, linear differential equation in S_r^m which is easily solved (within the range of interest $1.4 < \beta < 2.0$) to yield

$$S_r^m = -\frac{6(1-\nu^2)W_o^2}{(\beta-2)^2}\left(\frac{\beta}{\beta-1}R^{2\beta-2} - \frac{8\beta}{\beta+2}R^{\beta} + \frac{\beta^2}{2}R^2\right)$$

$$-\frac{C_1}{2R^2} + C_2 \tag{9}$$

where C_1 and C_2 are arbitrary constants which are chosen so that the radial displacement is finite at the plate center ($C_1 = 0$) and zero at the plate edge, whence

$$C_2 = \frac{6(1+\nu)W_o^2}{(\beta-2)^2}\left[\frac{\beta(2\beta-1-\nu)}{\beta-1} - \frac{8\beta(\beta-\nu+1)}{\beta+2} + \frac{\beta^2(3-\nu)}{2}\right]. \tag{10}$$

Now, deflections and membrane stresses are explicitly defined in terms of β. Bending stresses are easily obtained using the classical (linear) stress-strain and strain-displacement relations given in Brodland (1986).

The principle of virtual work allows the dimensionless load P to be determined from the total strain energy V of the plate, noting that although $V = V(W_o, \beta(W_o))$ for a given geometry, the dependence of V on β is small (Dolovich, 1986) and hence

$$P = \frac{dV}{dW_o} \approx \frac{\partial V}{\partial W_o}. \tag{11}$$

Now, all deflections and stresses, as well as the value of the centric load P, are known in terms of parameter β. This parameter is found by substituting equations (7), (9), and (11) into equation $(1)_1$ and minimizing the residual with respect to β. This residual X is a measure of the imbalanced shear at radius R. The order of the metric has little effect on the calculated value of β (Dolovich, 1986) and hence the functional

$$F(X) = \int_0^1 |X| dR \tag{12}$$

was used because it is easy to evaluate numerically. Values for β thus obtained are described, approximately, by the expressions

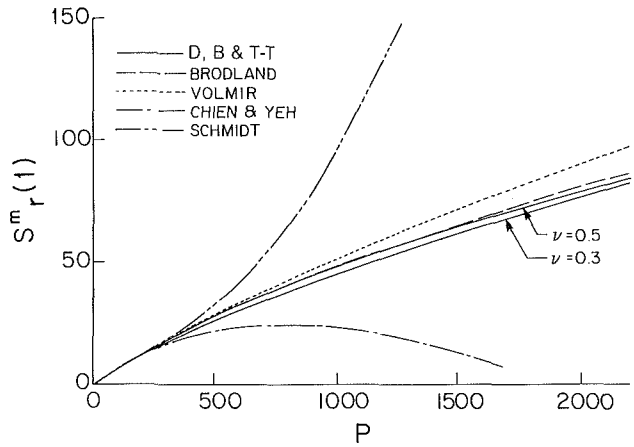


Fig. 3 Radial membrane edge stress versus load

$$\beta = \begin{cases} 2.0 - 0.12W_o, & 0 < W_o \le 1.0 \\ 2.035 - 0.171W_o + 0.0116W_o^2, & 1.0 < W_o. \end{cases} \tag{13}$$

Similar expressions can be constructed to relate other quantities. The relationship between P and W_o, for example, takes the form

$$P = \begin{cases} 50.41W_o + (14.79 + 16.52\nu - 11.21\nu^2)W_o^3, & 0 < W_o \le 1 \\ 51.88W_o + (14.36 + 17.49\nu - 11.16\nu^2)W_o^3, & 1 < W_o \le 2 \\ 56.55W_o + (13.61 + 16.00\nu - 11.23\nu^2)W_o^3, & 2 < W_o. \end{cases} \tag{14}$$

The above load-deflection relations agree well with corresponding expressions derived by others such as Schmidt (1968) and Chien and Yeh (1954).

Results

The clamped plate relationships illustrated *here* are essentially independent of Poisson's ratio (Brodland, 1986; Dolovich, 1986) except as noted. Figure 1 shows deflection profiles for center deflections W_o of 1 and 3 times the plate thickness. The analytical results of Chien and Yeh (1954), Basuli (1961), and Vol'mir (1956) and previously unpublished numerical results calculated using a method presented by Brodland (1987) are compared with the solution developed above. For $W_o = 1$ the solutions are nearly indistinguishable. However, at larger deflections such as $W_o = 3$, the plate *shape* is significantly altered from that predicted by linear theory (Vol'mir, 1956). It is clear from the figure that the solution developed here is closest to the most accurate solution, namely that predicted numerically.

Load-deflection curves are shown in Fig. 2. Here again, the new solution is seen to agree well with the numerical solution. Figure 3 shows the edge, radial, membrane stress as a function of load. All solutions are in excellent agreement for $P < 250$, however the power-series solutions diverge from the others above this load. More detailed calculations and comparisons, which show that the new solution is particularly valuable in the range $1 < W_o < 2$, are contained in Dolovich (1986).

The solution presented herein models the plate shape change more accurately than any of the other analytical solutions. It is also simple in form and easy to use, general and complete, valid for all values of Poisson's ratio, and accurate over a large range of deflections. Although each of these advantages can be found in one or more of the other analytical solution, no other solution has them all.

Acknowledgments

The authors gratefully acknowledge the help provided by D. W. Trim, Department of Applied Mathematics, University of Manitoba, and the financial support of the Natural Sciences and Engineering Research Council of Canada.

References

Basuli, S., 1961, "Note on the Large Deflection of a Circular Plate Under a Concentrated Load," *Zeitschrift fuer angewandtke Mathematik und Physik*, Vol. 12, pp. 357–362.

Berger, H. M., 1955, "A New Approach to the Analysis of Large Deflections of Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 22, pp. 465–472.

Brodland, G. W., 1986, "Nonlinear Deformation of Uniformly Loaded Circular Plates," *Solid Mechanics Archives*, Vol. 11, No. 4, pp. 219–256.

Brodland, G. W., 1987, "Highly Nonlinear Deflection of Uniformly-Loaded Circular Plates," *International Journal of Solids and Structures*, to appear.

Cherepy, R. D., 1960, "Large Deflection of a Circular Plate Subjected to a Concentrated Load at the Center," Thesis, University of Arizona, Tucson, AZ.

Chia, C. Y., 1980, *Nonlinear Anlaysis of Plates*, McGraw-Hill, New York.

Chien, W. Z., and Yeh, K. Y., 1954, "On the Large Deflection of Circular Plates," *Scientia sin.*, Vol. 3, pp. 405–436.

Dolovich, A. T., 1986, "Large Deflection of Point Loaded Circular Plates," Thesis, University of Manitoba, Winnipeg, Manitoba.

Frakes, J. P., and Simmonds, J. G., 1985, "Asymptotic Solutions of the von Karman Equations for a Circular Plate Under a Concentrated Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 326–330.

Nadai, A., 1925, *Die Elastichen Platten*, Springer, Berlin.

Schmidt, R., 1968, "Large Deflections of a Clamped Circular Plate," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 6, EM6, pp. 1603–1606.

Stippes, M., and Hausrath, A. H., 1952, "Large Deflections of Circular Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 19, No. 3, pp. 287–292.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, McGraw-Hill, New York.

Vol'mir, A. S., 1956, *Gibkie Plastinki i Obolchki*, Moscow; Engl. Trans. by Dept. of Eng. Sci. and Mech., University of Florida, 1967, *Flexible Plates and Shells*, AFFDL-TR-66-216.

Way, S., 1934, "Bending of Circular Plates with Large Deflection," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 627–636.

The Formulation of Dynamical Equations of Holonomic Mechanical Systems Using a Natural Orthogonal Complement

Jorge Angeles[1] and Sang Koo Lee[1]

1 Introduction

The subject of this note is the formulation of the dynamical equations of mechanical systems composed of rigid bodies coupled by holonomic constraints. Several approaches to the problem have been proposed, aimed at separating independent from dependent generalized coordinates (Nikravesh and Haug, 1982; Kammam and Huston, 1984; Singh and Likins, 1985; Kim and Vanderploeg, 1986-1. 2). A comprehensive discussion of the problem at hand, and various formulations proposed to solve it appear in Jerkovsky (1978), whereas Casey (1983) presents a tensor treatment of rigid-body dynamics that is particularly useful in this context.

Proposed in this note is a method that allows the elimination of the constraint forces by multiplication of the unconstrained dynamical equations by a suitably defined orthogonal complement of the matrix of linear velocity constraints. Although the method presented here bears many items in common with the formulation presented by Jerkovsky (1978), the said method differs from the latter in various respects, as described next.

2 Preliminary Definitions

A system of r holonomically-coupled rigid bodies is the subject of this study. The *twist* of the system's ith rigid body, undergoing an arbitrary motion in the three-dimensional space, $\mathbf{t}_i$, is defined here in terms of its angular velocity, ω_i, and the velocity of its mass center, $\dot{\mathbf{c}}_i$:

$$\mathbf{t}_i \equiv [\omega_i^T, \dot{\mathbf{c}}_i^T]^T \tag{1}$$

in which all vectors involved are considered to be column arrays. Moreover, $\mathbf{I}_i$ denotes the *inertia tensor* of the rigid body about its mass center, and this, as well as all vector quantities involved, are referred to a coordinate system fixed to the body. Furthermore, let $\mathbf{n}_i$ and $\mathbf{f}_i$ denote the resultant torque about the resultant force acting at the mass center of the body. Then, the *wrench*, $\mathbf{w}_i$, acting on the body is defined, in accordance with the definition of $\mathbf{t}_i$, as

$$\mathbf{w}_i \equiv [\mathbf{n}_i^T, \mathbf{f}_i^T]^T \tag{2}$$

and this is decomposed into an *external* wrench, $\mathbf{w}_i^E$, of forces and moments applied to the body under study by the system's environment, and a *constraint* wrench, $\mathbf{w}_i^C$, of nonworking forces and moments of constraint, i.e.,

$$\mathbf{w}_i = \mathbf{w}_i^E + \mathbf{w}_i^C \tag{3}$$

Now, tensor $\mathbf{\Omega}_i$, which many an author denote by $\tilde{\omega}_i$ is defined as

$$\mathbf{\Omega}_i \equiv \mathbf{1} \times \omega_i \equiv \frac{\partial (\omega_i \times \mathbf{x})}{\partial \mathbf{x}} \tag{4}$$

whereas the 6×6 matrices of *extended angular velocity*, $\mathbf{W}_i$, and of *extended mass*, $\mathbf{M}_i$, are now defined as

$$\mathbf{W}_i \equiv \begin{pmatrix} \mathbf{\Omega}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{M}_i \equiv \begin{pmatrix} \mathbf{I}_i & \mathbf{0} \\ \mathbf{0} & m_i \mathbf{1} \end{pmatrix} \tag{5}$$

and $\mathbf{0}$ denotes the zero tensor. Thus, the Newton-Euler equations of the ith rigid body can be written in compact form as

$$\mathbf{M}_i \dot{\mathbf{t}}_i = -\mathbf{W}_i \mathbf{M}_i \mathbf{t}_i + \mathbf{w}_i^E + \mathbf{w}_i^C \tag{6}$$

which will be used in the following Sections to derive the constrained system of dynamical equations sought.

3 Unconstrained Dynamical Equations

In what follows, it is assumed that the bodies comprising the mechanical system at hand move with respect to an inertial frame, which is labelled 0, the system's bodies being numbered from 1 to r.

Next, the $6r \times 6r$ matrices of *generalized mass*, $\mathbf{M}$, and *generalized angular velocity*, $\mathbf{W}$, as well as the $6r$-dimensional vectors of *generalized twist*, $\mathbf{t}$, *generalized external wrench*, $\mathbf{w}^E$, and *generalized—nonworking—constraint wrench*, $\mathbf{w}^C$, are defined as

[1]Department of Mechanical Engineering, Robotic Mechanical Systems Laboratory, McGill Research Centre for Intelligent Machines, McGill University, Montreal, Quebec, Canada H3A 2A7.

Manuscript received by ASME Applied Mechanics Division, June 1, 1987; final revision September 2, 1987.

Load-deflection curves are shown in Fig. 2. Here again, the new solution is seen to agree well with the numerical solution. Figure 3 shows the edge, radial, membrane stress as a function of load. All solutions are in excellent agreement for $P < 250$, however the power-series solutions diverge from the others above this load. More detailed calculations and comparisons, which show that the new solution is particularly valuable in the range $1 < W_o < 2$, are contained in Dolovich (1986).

The solution presented herein models the plate shape change more accurately than any of the other analytical solutions. It is also simple in form and easy to use, general and complete, valid for all values of Poisson's ratio, and accurate over a large range of deflections. Although each of these advantages can be found in one or more of the other analytical solution, no other solution has them all.

Acknowledgments

The authors gratefully acknowledge the help provided by D. W. Trim, Department of Applied Mathematics, University of Manitoba, and the financial support of the Natural Sciences and Engineering Research Council of Canada.

References

Basuli, S., 1961, "Note on the Large Deflection of a Circular Plate Under a Concentrated Load," *Zeitschrift fuer angewandtke Mathematik und Physik*, Vol. 12, pp. 357–362.

Berger, H. M., 1955, "A New Approach to the Analysis of Large Deflections of Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 22, pp. 465–472.

Brodland, G. W., 1986, "Nonlinear Deformation of Uniformly Loaded Circular Plates," *Solid Mechanics Archives*, Vol. 11, No. 4, pp. 219–256.

Brodland, G. W., 1987, "Highly Nonlinear Deflection of Uniformly-Loaded Circular Plates," *International Journal of Solids and Structures*, to appear.

Cherepy, R. D., 1960, "Large Deflection of a Circular Plate Subjected to a Concentrated Load at the Center," Thesis, University of Arizona, Tucson, AZ.

Chia, C. Y., 1980, *Nonlinear Anlaysis of Plates*, McGraw-Hill, New York.

Chien, W. Z., and Yeh, K. Y., 1954, "On the Large Deflection of Circular Plates," *Scientia sin.*, Vol. 3, pp. 405–436.

Dolovich, A. T., 1986, "Large Deflection of Point Loaded Circular Plates," Thesis, University of Manitoba, Winnipeg, Manitoba.

Frakes, J. P., and Simmonds, J. G., 1985, "Asymptotic Solutions of the von Karman Equations for a Circular Plate Under a Concentrated Load," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 326–330.

Nadai, A., 1925, *Die Elastichen Platten*, Springer, Berlin.

Schmidt, R., 1968, "Large Deflections of a Clamped Circular Plate," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 6, EM6, pp. 1603–1606.

Stippes, M., and Hausrath, A. H., 1952, "Large Deflections of Circular Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 19, No. 3, pp. 287–292.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, McGraw-Hill, New York.

Vol'mir, A. S., 1956, *Gibkie Plastinki i Obolchki*, Moscow; Engl. Trans. by Dept. of Eng. Sci. and Mech., University of Florida, 1967, *Flexible Plates and Shells*, AFFDL-TR-66-216.

Way, S., 1934, "Bending of Circular Plates with Large Deflection," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 627–636.

The Formulation of Dynamical Equations of Holonomic Mechanical Systems Using a Natural Orthogonal Complement

Jorge Angeles[1] and Sang Koo Lee[1]

1 Introduction

The subject of this note is the formulation of the dynamical equations of mechanical systems composed of rigid bodies coupled by holonomic constraints. Several approaches to the problem have been proposed, aimed at separating independent from dependent generalized coordinates (Nikravesh and Haug, 1982; Kammam and Huston, 1984; Singh and Likins, 1985; Kim and Vanderploeg, 1986-1. 2). A comprehensive discussion of the problem at hand, and various formulations proposed to solve it appear in Jerkovsky (1978), whereas Casey (1983) presents a tensor treatment of rigid-body dynamics that is particularly useful in this context.

Proposed in this note is a method that allows the elimination of the constraint forces by multiplication of the unconstrained dynamical equations by a suitably defined orthogonal complement of the matrix of linear velocity constraints. Although the method presented here bears many items in common with the formulation presented by Jerkovsky (1978), the said method differs from the latter in various respects, as described next.

2 Preliminary Definitions

A system of r holonomically-coupled rigid bodies is the subject of this study. The *twist* of the system's ith rigid body, undergoing an arbitrary motion in the three-dimensional space, $\mathbf{t}_i$, is defined here in terms of its angular velocity, ω_i, and the velocity of its mass center, $\dot{\mathbf{c}}_i$:

$$\mathbf{t}_i \equiv [\omega_i^T, \dot{\mathbf{c}}_i^T]^T \tag{1}$$

in which all vectors involved are considered to be column arrays. Moreover, $\mathbf{I}_i$ denotes the *inertia tensor* of the rigid body about its mass center, and this, as well as all vector quantities involved, are referred to a coordinate system fixed to the body. Furthermore, let $\mathbf{n}_i$ and $\mathbf{f}_i$ denote the resultant torque about the resultant force acting at the mass center of the body. Then, the *wrench*, $\mathbf{w}_i$, acting on the body is defined, in accordance with the definition of $\mathbf{t}_i$, as

$$\mathbf{w}_i \equiv [\mathbf{n}_i^T, \mathbf{f}_i^T]^T \tag{2}$$

and this is decomposed into an *external* wrench, $\mathbf{w}_i^E$, of forces and moments applied to the body under study by the system's environment, and a *constraint* wrench, $\mathbf{w}_i^C$, of nonworking forces and moments of constraint, i.e.,

$$\mathbf{w}_i = \mathbf{w}_i^E + \mathbf{w}_i^C \tag{3}$$

Now, tensor $\mathbf{\Omega}_i$, which many an author denote by $\tilde{\omega}_i$ is defined as

$$\mathbf{\Omega}_i \equiv \mathbf{1} \times \omega_i \equiv \frac{\partial (\omega_i \times \mathbf{x})}{\partial \mathbf{x}} \tag{4}$$

whereas the 6×6 matrices of *extended angular velocity*, $\mathbf{W}_i$, and of *extended mass*, $\mathbf{M}_i$, are now defined as

$$\mathbf{W}_i \equiv \begin{pmatrix} \mathbf{\Omega}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{M}_i \equiv \begin{pmatrix} \mathbf{I}_i & \mathbf{0} \\ \mathbf{0} & m_i \mathbf{1} \end{pmatrix} \tag{5}$$

and $\mathbf{0}$ denotes the zero tensor. Thus, the Newton-Euler equations of the ith rigid body can be written in compact form as

$$\mathbf{M}_i \dot{\mathbf{t}}_i = -\mathbf{W}_i \mathbf{M}_i \mathbf{t}_i + \mathbf{w}_i^E + \mathbf{w}_i^C \tag{6}$$

which will be used in the following Sections to derive the constrained system of dynamical equations sought.

3 Unconstrained Dynamical Equations

In what follows, it is assumed that the bodies comprising the mechanical system at hand move with respect to an inertial frame, which is labelled 0, the system's bodies being numbered from 1 to r.

Next, the $6r \times 6r$ matrices of *generalized mass*, $\mathbf{M}$, and *generalized angular velocity*, $\mathbf{W}$, as well as the $6r$-dimensional vectors of *generalized twist*, $\mathbf{t}$, *generalized external wrench*, $\mathbf{w}^E$, and *generalized—nonworking—constraint wrench*, $\mathbf{w}^C$, are defined as

[1]Department of Mechanical Engineering, Robotic Mechanical Systems Laboratory, McGill Research Centre for Intelligent Machines, McGill University, Montreal, Quebec, Canada H3A 2A7.

Manuscript received by ASME Applied Mechanics Division, June 1, 1987; final revision September 2, 1987.

$$\mathbf{M} \equiv \text{diag}(\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_r),\ (\mathbf{W} \equiv \text{diag}(\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_r) \tag{7a}$$

$$\mathbf{t} \equiv [\mathbf{t}_1^T, \mathbf{t}_2^T, \ldots, \mathbf{t}_r^T]^T \tag{7b}$$

with similar definitions for $\mathbf{w}^E$ and $\mathbf{w}^C$. Hence, the r dynamical equations (6) can now be expressed in compact form as follows:

$$\mathbf{M}\dot{\mathbf{t}} = -\mathbf{WMt} + \mathbf{w}^E + \mathbf{w}^C \tag{8}$$

which is formally identical to equation (6), and constitutes a set of $6r$ *unconstrained* dynamical equations. Next, the kinematic constraints produced by the holonomic couplings are introduced.

4 Kinematic Holonomic Constraints and Constraint Forces

The kinematic constraints produced by holonomic couplings are usually represented in integral form, i.e., as a set of—most likely nonlinear—algebraic equations constraining the generalized coordinates. In the present formulation, it will prove convenient to represent the said constraints as a system of linear homogeneous equations on the twists. Indeed, it is well known that the holonomic constraints can be written as linear equations on the mass-center velocities and angular velocities. In the present formulation, this is equivalent to the following linear homogeneous system on the vector of generalized twist:

$$\mathbf{At} = \mathbf{0} \tag{9}$$

In the present formulation, however, rather than expressing the velocity constraint equations (9) as an independent system, it will prove advantageous to express them as a system of $6r$ linearly dependent equations, out of which, only $p\,(<6r)$ are independent. Thus, matrix $\mathbf{A}$ will be, in this formulation, a $6r \times 6r$ matrix of rank p. Since $6r$ components of generalized twist have been introduced, and p independent constraints have been assumed, the degree of freedom of the system is $q = 6r - p$. Hence, a set of q independent linear combinations of the aforementioned components exists that can play the role of *independent generalized speeds* (Kane, 1968). Let this set be denoted by $\dot{\mathbf{q}}$. Then, the vector of generalized twist can be represented as the following linear transformation of $\dot{\mathbf{q}}$:

$$\mathbf{t} = \mathbf{T}\dot{\mathbf{q}} \tag{10}$$

where $\mathbf{T}$ is a $6r \times q$ matrix. Upon substitution of $\mathbf{t}$, as given by equation (10), into equation (9), the following is readily derived:

$$\mathbf{AT} = \mathbf{0} \tag{11}$$

which shows that $\mathbf{T}$ is an *orthogonal complement* of $\mathbf{A}$. In what follows, such an orthogonal complement is termed *the natural orthogonal complement* of matrix $\mathbf{A}$.

It is known—see, e.g., (Nikravesh and Haug, 1982)—that, if matrix $\mathbf{A}$ of equation (9) is of $p \times q$ and of full rank, i.e., rank $(\mathbf{A}) = p$, then the constraint wrench is the product of $\mathbf{A}^T$ by a p-dimensional vector of *Lagrange multipliers*, λ. The same holds within the present formulation, i.e.,

$$\mathbf{w}^C = \mathbf{A}^T\lambda \tag{12a}$$

Furthermore, $\mathbf{w}^E$ is decomposed as follows:

$$\mathbf{w}^E = \mathbf{w}^G + \mathbf{w}^J + \mathbf{w}^D \tag{12b}$$

where $\mathbf{w}^G$, $\mathbf{w}^J$, and $\mathbf{w}^D$ represent the wrenches due to gravity, applied torques and forces, and dissipative torques and forces, respectively. Notice that $\mathbf{w}^J$ comprises driving wrenches from the motors, whereas $\mathbf{w}^D$ accounts for friction torques and forces, i.e., actions arising from Coulomb and viscous friction.

5 Constrained Dynamical Equations

Now, a system of q constrained independent dynamical equations is derived systematically, that is equivalent to the system's Euler-Lagrange equations, making use of the natural orthogonal complement. To this end, the vector of generalized nonworking constraint wrench, as given by equation (12*a*), is substituted into equation (8), and then both sides of the resulting equation are multiplied by $\mathbf{T}^T$, which produces the following:

$$\mathbf{T}^T\mathbf{M}\dot{\mathbf{t}} = -\mathbf{T}^T\mathbf{WMt} + \mathbf{T}^T\mathbf{w}^E \tag{13}$$

where equation (11) has been taken into account. Now, both sides of equation (10) are differentiated with respect to time, and the arising expression for $\dot{\mathbf{t}}$ is substituted into equation (13), which produces the following system of q independent constrained dynamical equations:

$$\mathbf{T}^T\mathbf{MT}\ddot{\mathbf{q}} = -\mathbf{T}^T\mathbf{M}\dot{\mathbf{T}}\dot{\mathbf{q}} - \mathbf{T}^T\mathbf{WMT}\dot{\mathbf{q}} + \mathbf{T}^T\mathbf{w}^E \tag{14}$$

Now, the $q \times q$ matrices of *generalized inertia*, $\mathbf{I}$, and of *convective inertia terms*, $\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$, the q-dimensional vector of *generalized driving force*, τ, the q-dimensional vector of *generalized dissipative force*, δ, and the q-dimensional vector of *generalized force due to gravity*, γ, are defined as follows:

$$\mathbf{I} = \mathbf{T}^T\mathbf{MT},\ \mathbf{C}(\mathbf{q},\dot{\mathbf{q}}) = -\mathbf{T}^T(\mathbf{M}\dot{\mathbf{T}} + \mathbf{WMT}) \tag{15a}$$

$$\tau = \mathbf{T}^T\mathbf{w}^J,\ \delta = \mathbf{T}^T\mathbf{w}^D,\ \gamma = \mathbf{T}^T\mathbf{w}^G \tag{15b}$$

Hence, equation (14) can be rewritten in the form

$$\mathbf{I}(\mathbf{q})\ddot{\mathbf{q}} = \mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} + \tau + \delta + \gamma \tag{16}$$

where, clearly, the first term of the right-hand side is the one known popularly as the vector of *generalized centrifugal and Coriolis forces*. It is readily identified because of its quadratic nature in the generalized independent speeds. From the foregoing, then, it becomes apparent that equation (16) represents the system's Euler-Lagrange dynamical equations which appear, moreover, *free of constraint forces*.

Acknowledgments

The research work reported here was possible under NSERC (Natural Sciences and Engineering Research Council of Canada) Grant No. A4532 and FCAR (Funds pour la formation de chercheurs et l'aide á la recherche, of Québec) Grant No. EQ3072.

References

Casey, J., 1983, "A Treatment of Rigid Body Dynamics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 905–907.

Jerkovsky, W., 1978, "The Structure of Multibody Dynamical Equations," *J. of Guidance and Control*, Vol. 1, No. 3, pp. 173–182.

Kamman, J. W., and Huston, R. L., 1984, "Dynamics of Constrained Multibody Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 899–903.

Kane, T. R., 1968, *Dynamics*, Holt, Rinehart and Winston, New York.

Kim, S. S., and Vanderploeg, M. J., 1986a, "A General and Efficient Method for Dynamic Analysis of Mechanical Systems Using Velocity Transformations," *J. of Mech., Transm., and Auto. in Design*, Vol. 108, pp. 176–182.

Kim, S. S., and Vanderploeg, M. J., 1986b, "QR Decomposition for State Space Representation of Constrained Mechanical Dynamic Systems," *J. of Mech. Transm., and Auto. in Design*, Vol. 108, pp. 183–188.

Nikravesh, P. E., and Huag, E. J., 1982, "Generalized Coordinate Partitioning for Analysis of Mechanical Systems with Nonholonomic Constraints," *J. of Mech. Design*, Vol. 104, No. 1, p. 247–255.

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.

Thermoelastic Green's Functions for Plane Problems in General Anisotropy

F. A. Sturla[1] and J. R. Barber[1]

1 Introduction

A major contribution to the study of general anisotropic elasticity was made by Eshelby et al. (1953), who showed that, by making an appropriate linear transformation of the coordinate axes, the governing equations of the plane problem could be reduced to Laplacian form, thus permitting solutions to be written down in terms of analytic functions of the complex variable.

This method was elaborated by Stroh (1958), who also gave the solutions for a plane dislocation and for a Griffith crack perturbing an otherwise uniform tensile stress.

Eshelby and Stroh's method was extended to plane thermoelasticity by Clements (1973). The results were used by Atkinson and Clements (1977) to treat the Griffith crack problem with heat flow.

2 Solutions Using Green's Functions

The above authors used Fourier transforms to represent the stress and displacement fields, but an alternative method, which has been extensively used in isotropic crack and contact problems, is to represent the elastic fields in terms of Green's functions (see, e.g., Comninou, 1977). Green's functions for isothermal problems have been discussed by Atkinson (1966) and Sinclair and Hirth (1975).

The Green's function method has some important advantages over transform methods. In particular, we note that the solution is expressed in terms of physical variables so that it is easier to determine at intermediate stages whether the solution is physically reasonable. Also, it is usually possible to express the solution in terms of distributions of Green's functions over a finite range, with bounded or integrable singular behavior at the end points. This ensures that the resulting integral equations will have regular solutions. In contrast, Fourier transform methods require that we pay very careful attention to the way the representation behaves at infinity, to avoid possible divergent integrals. This is particularly important in thermoelastic problems (Barber, 1983).

Green's functions have not been extensively used for thermoelastic problems, but appropriate functions for the frictionless contact of two isotropic half-planes are given by Dundurs and Comninou (1979a,b) and used by Comninou and Dundurs (1980) to treat a problem in which separation between such bodies occurs in a central region. It is the purpose of the present paper to develop the appropriate plane thermoelastic Green's functions for the generally anisotropic material.

3 The Temperature Field

The thermal analog of a dislocation is a temperature field which exhibits a constant discontinuity along the half-line $x_2=0$, $x_1>0$ in Cartesian coordinates x_1, x_2, x_3, i.e.,

$$T(x_1,0^+)=0; \quad x_1>0 \tag{1}$$

$$T(x_1,0^-)=T_0; \quad x_1>0 \tag{2}$$

and which involves no net source at the origin. A related isotropic Green's function was introduced by Dundurs and Comninou (1979a), who described it as a "heat vortex," so we shall retain this terminology here.

The temperature distribution, T, must satisfy the heat conduction equation

$$K_{ij}\frac{\partial^2 T}{\partial x_i \partial x_j}=0 \tag{3}$$

where $K_{ij}=K_{ji}$ are the coefficients of heat conduction. By a suitable linear transformation of coordinates, equation (3) can be reduced to plane harmonic form. We define the modified complex variable $z_t=x_1+\tau x_2$, where τ is the root with positive imaginary part of the equation

$$K_{11}+2K_{12}\tau+K_{22}\tau^2=0 \tag{4}$$

With this notation, it follows that any distribution of the form $T(x_1,x_2)=f(z_t)$ satisfies equation (3) where $f(z_t)$ is an analytic function of z_t.

A suitable function satisfying the boundary conditions (1) and (2) can be written as

$$T(x_1,x_2)=T_0(\log z_t-\log\bar{z}_t)/4\pi i \tag{5}$$

where $\bar{z}_t$ denotes the complex conjugate $x_1+\bar{\tau}x_2$. The heat flux is

$$q_i=-K_{ij}\frac{\partial T}{\partial x_j}=(-T_0/4\pi i)\{K_{i1}(1/z_t-1/\bar{z}_t) + K_{i2}(\tau/z_t-\bar{\tau}/\bar{z}_t)\} \tag{6}$$

In particular, on the plane $x_2=0$ we have

$$q_2=-(T_0K_{22}/4\pi i)\{\tau-\bar{\tau}\}/x_1 \tag{7}$$

4 Particular Thermoelastic Solution

Following Clements (1973), we can develop a particular solution of the equilibrium equation in the form

$$u_k=C_k f(z_t)+\bar{C}_k\bar{f}(\bar{z}_t) \tag{8}$$

$$T=Af'(z_t)+\bar{A}\bar{f}'(\bar{z}_t) \tag{9}$$

where C_k are obtained from the equation

$$D_{ik}C_k=A\Gamma_i \tag{10}$$

with D_{ik} and Γ_i given by

$$D_{ik}=c_{i1k1}+\tau(c_{i1k2}+c_{i2k1})+\tau^2 c_{i2k2} \tag{11}$$

$$\Gamma_i=\beta_{i1}+\tau\beta_{i2} \tag{12}$$

and c_{ijkl}, β_{ij} are the elastic constants and the stress-temperature coefficients, respectively. From equation (10) we have

$$C_k=AP_{ki}\Gamma_i \tag{13}$$

where

$$P_{ji}D_{ik}=\delta_{jk} \tag{14}$$

The stress components σ_{ij} can be obtained by substituting for u_k, T from equations (8) and (9) into the constitutive relation giving

$$\sigma_{ij}=A(N_{ij}-\beta_{ij})f'(z_t)+\bar{A}(\bar{N}_{ij}-\beta_{ij})\bar{f}'(\bar{z}_t) \tag{15}$$

(Clements, 1973), where

$$N_{ij}=c_{ijk1}P_{kl}\Gamma_l+\tau c_{ijk2}P_{kl}\Gamma_l \tag{16}$$

The temperature distribution of equation (5) can be expressed in the form of equation (9) by defining $A=T_0/4\pi i$ and

$$f(z_t)=z_t\log z_t-z_t \tag{17}$$

This function is continuous everywhere except on the real axis $x_2=0$, $x_1>0$ and hence u_k, σ_{ij} are continuous except on this line. We can, therefore, regard equation (17) as defining an

[1]Department of Mechanical Engineering and Applied Mechanics, University of Michigan, Ann Arbor, MI 48109.

Manuscript received by ASME Applied Mechanics Division, August 8, 1986; final revision September 27, 1987.

appropriate particular solution for the temperature distribution (5), provided we make a cut on the real axis. The stresses σ_{ij} and displacements u_k then have different values on each side of the cut, the corresponding expressions being

$$u_k = \frac{T_0 E_k}{4\pi i}(x_1 \log x_1 - x_1)$$

$$-\frac{T_0 \bar{E}_k}{4\pi i}(x_1 \log x_1 - x_1);\ x_2 = 0^+,\ x_1 > 0 \tag{18}$$

$$= \frac{T_0 E_k}{4\pi i} x_1 (\log x_1 - 1 + 2\pi i)$$

$$-\frac{T_0 \bar{E}_k}{4\pi i} x_1 (\log x_1 - 1 - 2\pi i);\ x_2 = 0^-,\ x_1 > 0 \tag{19}$$

$$\sigma_{ij} = \frac{T_0}{4\pi i}(N_{ij} - \beta_{ij}) \log x_1$$

$$-\frac{T_0}{4\pi i}(\bar{N}_{ij} - \beta_{ij}) \log x_1;\ x_2 = 0^+,\ x_1 > 0 \tag{20}$$

$$= \frac{T_0}{4\pi i}(N_{ij} - \beta_{ij})(\log x_1 + 2\pi i)$$

$$-\frac{T_0}{4\pi i}(\bar{N}_{ij} - \beta_{ij})(\log x_1 - 2\pi i);\ x_2 = 0^-,\ x_1 > 0 \tag{21}$$

where

$$E_k = P_{ki}\Gamma_i \tag{22}$$

It follows that there is a discontinuity in displacement

$$\Delta u_k \equiv u_k(x_1, 0^+) - u_k(x_1, 0^-)$$

$$= -(E_k + \bar{E}_k) T_0 x_1 / 2;\ x_1 > 0 \tag{23}$$

and in stress

$$\Delta\sigma_{ij} \equiv \sigma_{ij}(x_1, 0^+) - \sigma_{ij}(x_1, 0^-)$$

$$= -\{(N_{ij} + \bar{N}_{ij})/2 - \beta_{ij}\} T_0 \tag{24}$$

From these equations we note that the stress discontinuity, $\Delta\sigma_{ij}$, is constant across the cut $x_2 = 0$, $x_1 > 0$, but the discontinuity in displacement varies linearly with x_1.

These discontinuities are unacceptable if the Green's function is to be used for the solution of problems involving one or more closed cracks in an infinite medium. We, therefore, seek a corrective isothermal solution for the cut plane which when superposed on the particular thermoelastic solution will restore continuity of stress and displacement across the cut.

5 The Isothermal Wedge Problem

The infinite body with a cut on the positive x_1 axis constitutes an infinite 360 deg wedge, defined by $0 < \theta < 2\pi$, $r > 0$ in polar coordinates (see Fig. 1). The corrective solution is required to satisfy boundary conditions which are equal and opposite to equations (23) and (24).

To solve this problem, we make use of the solution due to Stroh (1958), in which the displacement is written

$$u_k = \sum_\alpha A_{k\alpha}\phi_\alpha(z_\alpha) + \bar{A}_{k\alpha}\bar{\phi}_\alpha(\bar{z}_\alpha) \tag{25}$$

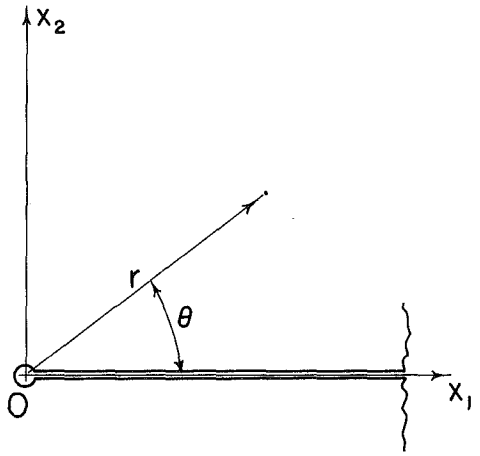
where

$$z_\alpha = x_1 + p_\alpha x_2 \tag{26}$$

and the p_α are the three roots with positive imaginary part of the equation

$$|c_{i1k1} + p_\alpha(c_{i1k2} + c_{i2k1}) + p_\alpha^2 c_{i2k2}| = 0 \tag{27}$$

The stress components σ_{ij} are defined in terms of the function



$$\Omega_i = \sum_\alpha L_{i\alpha}\phi_\alpha(z_\alpha) + \bar{L}_{i\alpha}\bar{\phi}_\alpha(\bar{z}_\alpha) \tag{28}$$

(Stroh, 1958), where

$$L_{i\alpha} = (c_{i2k1} + p_\alpha c_{i2k2}) A_{k\alpha} \tag{29}$$

in particular, we have

$$\sigma_{i2} = \frac{\partial\Omega_i}{\partial x_1} = \sum_\alpha L_{i\alpha}\phi'_\alpha(z_\alpha) + \sum_\alpha \bar{L}_{i\alpha}\bar{\phi}'_\alpha(\bar{z}_\alpha) \tag{30}$$

For the 360 deg wedge with uniform tractions on the faces, we take.

$$\phi_\alpha = D_\alpha [z_\alpha \log z_\alpha - z_\alpha] \tag{31}$$

where D_α is a set of three complex constants which can be chosen to satisfy the six boundary conditions. Substituting into equations (25) and (30), we find

$$u_k = \sum_\alpha A_{k\alpha} D_\alpha x_1 (\log x_1 - 1)$$

$$+ \bar{A}_{k\alpha}\bar{D}_\alpha x_1 (\log x_1 - 1);\ x_2 = 0^+,\ x_1 > 0 \tag{32}$$

$$= \sum_\alpha A_{k\alpha} D_\alpha x_1 (\log x_1 - 1 + 2\pi i)$$

$$+ \bar{A}_{k\alpha}\bar{D}_\alpha x_1 (\log x_1 - 1 - 2\pi i);\ x_2 = 0^-, x_1 > 0 \tag{33}$$

$$\sigma_{i2} = \sum_\alpha L_{i\alpha} D_\alpha \log x_1 + \bar{L}_{i\alpha}\bar{D}_\alpha \log x_1;$$

$$x_2 = 0^+,\ x_1 > 0 \tag{34}$$

$$= \sum_\alpha L_{i\alpha} D_\alpha (\log x_1 + 2\pi i) + \bar{L}_{i\alpha}\bar{D}_\alpha (\log x_1 - 2\pi i);$$

$$x_2 = 0^-,\ x_1 > 0 \tag{35}$$

We conclude that this solution will cancel the discontinuity of equations (23) and (24) provided D_α satisfies the equations

$$2\pi i \sum_\alpha (A_{k\alpha} D_\alpha - \bar{A}_{k\alpha}\bar{D}_\alpha) = -T_0(E_k + \bar{E}_k)/2;\ k = 1,2,3 \tag{36}$$

$$2\pi i \sum_\alpha (L_{k\alpha} D_\alpha - \bar{L}_{k\alpha}\bar{D}_\alpha)$$

$$= -T_0\{(N_{k2} + \bar{N}_{k2})/2 - \beta_{k2}\};\ k = 1,2,3 \tag{37}$$

Equations (36) and (37) can be solved, giving

$$D_\alpha = -(T_0/4\pi i)$$

$$\frac{\{(L_{k\alpha}(E_k+\bar{E}_k)/2+A_{k\alpha}[(N_{k2}+\bar{N}_{k2})/2-\beta_{k2}]\}}{L_{j\alpha}A_{j\alpha}} \tag{38}$$

Where we have used the following identities

$$L_{i\alpha}\bar{A}_{i\beta}+\bar{L}_{i\beta}A_{i\alpha}=0 \tag{39}$$

$$L_{i\alpha}A_{i\beta}+L_{i\beta}A_{i\alpha}=2L_{i\alpha}A_{i\alpha}\delta_{\alpha\beta} \tag{40}$$

Equation (39) was proved by Stroh (1958, p. 628), and equation (40) can be proven using a similar procedure.

6 The Thermoelastic Green's Function

The required thermoelastic Green's function can now be obtained by superposing the particular solution of Section 4 and the corrective solution of Section 5. In particular, we find that the stress on the plane $x_2=0$ is given by

$$\sigma_{i2}=T_0G_i\log|x_1| \tag{41}$$

where

$$G_i=\frac{1}{4\pi i}\Big\{\sum_\alpha\Big(\frac{L_{i\alpha}A_{k\alpha}}{L_{j\alpha}A_{j\alpha}}-\frac{\bar{L}_{i\alpha}\bar{A}_{k\alpha}}{\bar{L}_{j\alpha}\bar{A}_{j\alpha}}\Big)\Big(\frac{N_{k2}+\bar{N}_{k2}}{2}-\beta_{k2}\Big)+\sum_\alpha\Big(\frac{L_{i\alpha}L_{k\alpha}}{L_{j\alpha}A_{j\alpha}}-\frac{\bar{L}_{i\alpha}\bar{L}_{k\alpha}}{\bar{L}_{j\alpha}\bar{A}_{j\alpha}}\Big)\Big(\frac{E_k+\bar{E}_k}{2}\Big)+N_{i2}-\bar{N}_{i2}\Big\} \tag{42}$$

More general thermoelastic stress fields can be constructed by distributing Green's functions of the above form with an appropriate weight function. This method is particularly useful for representing the perturbation in the thermoelastic stress field in an infinite plane due to a closed crack. Green's functions are distributed along the line of the crack and the boundary conditions lead to integral equations for the unknown weight function. The problem of a plane crack in an infinite anisotropic plane is considered in a companion paper (Sturla and Barber, 1987).

Acknowledgment

The authors are pleased to acknowledge the support of the National Science Foundation under the grant number MSM-8419324.

References

Atkinson, C., 1966, "The Interaction Between a Dislocation and a Crack," *Int. J. Fracture Mech.*, Vol. 2, pp. 567–575.

Atkinson, C., and Clements, D. L., 1977, "On Some Crack Problems in Anisotropic Thermoelasticity," *Int. J. Solids Structures*, Vol. 13, pp. 855–864.

Barber, J. R., 1983, "The Solution of Elasticity Problems for the Half-Space by the Method of Green and Collins," *Appl. Sci. Res.*, Vol. 40, pp. 135–157.

Clements, D. L., 1973, "Thermal Stress in an Anisotropic Elastic Half-Space," *SIAM J. Appl. Math.*, Vol. 3, pp. 332–337.

Comninou, M., and Dundurs, J., 1980, "On Lack of Uniqueness in Heat Conduction Through a Solid to Solid Contact," *ASME J. Heat Transfer*, Vol. 102, pp. 319–323.

Comninou, M., 1977, "The Interface Crack," ASME Journal of Applied Mechanics, Vol. 44, pp. 631–636.

Dundurs, J., and Comninou, M., 1979a, "Green's Functions for Planar Thermoelastic Contact Problems—Exterior Contact," *Mech. Res. Comm.*, Vol. 6, No. 5, pp. 309–316.

Dundurs, J., and Comninou, M., 1979b, "Green's Functions for Planar Thermoelastic Contact Problems—Interior Contact," *Mech. Res. Comm.*, Vol. 6, No. 5, pp. 317–321.

Eshelby, J. D., Read, W. T., and Shockley, W., 1953, "Anisotropic Elasticity With Applications to Dislocation Theory," *Acta Met.*, Vol. 1, pp. 251–259.

Sinclair, J. E., and Hirth, J. P., 1975, "Two Dimensional Elastic Green Function for a Cracked Anisotropic Body," *J. Phys. F:Metal Phys.*, Vol. 5, pp. 236–246.

Stroh, A. N., 1958, "Dislocations and Cracks in Anisotropic Elasticity," *Phil. Mag.*, Vol. 3, pp. 625–646.

Sturla, F. A., and Barber, J. R., 1988, "Thermal Stresses Due to a Plane Crack in a General Anisotropic Material," ASME Journal of Applied Mechanics, Vol. 55, to appear.

Journal of Applied Mechanics

Discussion

Periodic Array of Cracks in a Half-Plane Subjected to Arbitrary Loading[1]

N. I. Ioakimidis.[2] The writer noted with much interest the results of this work by the author and took into account the difficulties having been faced, so successfully, during the solution of the title problem. The author should be congratulated for his important results and, particulary, for the use of *finite-part integrals* and *hypersingular integral equations* in a very interesting problem of fracture mechanics. It is the third time, recently, that the writer notices the use of these techniques in crack problems in the ASME JOURNAL OF APPLIED MECHANICS and further contributions are not only welcomed, but expected as well. Therefore, it seems that some additional literature in the field, beyond the very important results by Kaya and Erdogan (appropriately referenced in the paper), should be added and this is really useful. Because of space restrictions, this literature is, obviously, somewhat incomplete.

Finite-part integrals have been repeatedly used in the past in fluid dynamics (Ashley and Landahl, 1985) and additional fields of engineering. In fracture mechanics, they seem to be introduced by the writer (Ioakimidis, 1982a) both in two and in three-dimensional elasticity. The terms *hypersingularities* and *hypersingular integral equations* (with the use of the Greek word "ὑπέρ") seem to also have been introduced by the writer (Ioakimidis, 1984a, for the first term, and Ioakimidis, 1985a, 1986, 1987a, 1987b, for the second term). The writer was very happy to see the adoption of the term *hypersingular integral equation* by the author. On the other hand, additional results by the writer on the hypersingular integral equations for crack problems in two-dimensional elasticity can be found in subsequent papers by him (Ioakimidis, 1983a, 1983b, 1984a).

Furthermore, for the numerical evaluation of finite-part integrals with a second-order singularity, as is exactly the case in the present paper, reference can be made to the following papers by the writer (Ioakimidis, 1981, 1983c, 1985b). Moreover, for the numerical solution of the corresponding hypersingular integral equations, one can use either the quadrature method (Ioakimidis and Theocaris, 1979b; Ioakimidis, 1984b), perhaps slightly modified, or the collocation or even the Galerkin methods, applied to the hypersingular integral equation of crack problems by the writer (Ioakimidis, 1982b). The author of the paper chose the collocation method (as is clear) by comparing the formulas of Section 3 of his paper with the formulas of Section 2 of the last reference (Ioakimidis, 1982b). Even the selection of the collocation points was the same (roots of the Chebyshev polynomial of the first kind). As far as the theoretical proof of the convergence both of the collocation and of the Galerkin methods is concerned, reference is made to the results by Golberg (1983, 1985).

Finally, it can be added that the most general problem of a periodic array of curvilinear cracks in a half-plane (or in two bonded half-planes) under completely arbitrary loading was reduced to one complex Cauchy-type singular integral equation by the writer (Ioakimidis and Theocaris, 1979a) by a direct method, not requiring integral transform techniques.

It is believed that the above comments and the literature reported will be of some interest to the author, as well as to the casual reader of the paper, and will significantly improve its undoubted importance and originality.

References

Ashley, H., and Landahl, M., 1985, *Aerodynamics of Wings and Bodies*, Dover, New York, originally published by Addison-Wesley, Reading, MA, 1965.

Golberg, M. A., 1983, "The Convergence of Several Algorithms for Solving Integral Equations with Finite-Part Integrals," *Journal of Integral Equations*, Vol. 5, pp. 329–340.

Golberg, M. A., 1985, "The Convergence of Several Algorithms for Solving Integral Equations with Finite-Part Integrals. II," *Journal of Integral Equations*, Vol. 9, pp. 267–275.

Ioakimidis, N. I., 1981, "On the Numerical Evaluation of Derivatives of Cauchy Principal Value Integrals," *Computing*, Vol. 27, pp. 81–88.

Ioakimidis, N. I., 1982a, "Application of Finite-Part Integrals to the Singular Integral Equations of Crack Problems in Plane and Three-Dimensional Elasticity," *Acta Mechanica*, Vol. 45, pp. 31–47.

Ioakimidis, N. I., 1982b, "Two Methods for the Numerical Solution of Bueckner's Singular Integral Equation for Plane Elasticity Crack Problems," *Computer Methods in Applied Mechanics and Engineering*, Vol. 31, pp. 169–177.

Ioakimidis, N. I., 1983a, "A New Singular Integral Equation for the Classical Crack Problem in Plane and Antiplane Elasticity," *International Journal of Fracture*, Vol. 21, pp. 115–122.

Ioakimidis, N. I., 1983b, "On the Application of the Generalized Plemelj Formulas to Crack Problems," *Revue Roumaine des Sciences Techniques-Série de Mécanique Appliquée*, Vol. 28, pp. 475–478.

Ioakimidis, N. I., 1983c, "A Direct Method for the Construction of Gaussian Quadrature Rules for Cauchy Type and Finite-Part Integrals," *L'Analyse Numérique et la Théorie de l'Approximation*, Vol. 12, pp. 131–140.

Ioakimidis, N. I., 1984a, "Hypersingularities and Cracks in Plane and Three-Dimensional Elasticity," *Theoretical and Applied Mechanics*, Bulgarian Academy of Sciences, Year XV, pp. 35–40.

Ioakimidis, N. I., 1984b, "A Natural Interpolation Formula for Prandtl's Singular Integrodifferential Equation," *International Journal for Numerical Methods in Fluids*, Vol. 4, pp. 283–290.

Ioakimidis, N. I., 1985a, "Determination of the Order of Singularity at the Apex of a Wedge-Shaped Crack," *Engineering Fracture Mechanics*, Vol. 22, pp. 369–373.

Ioakimidis, N. I., 1985b, "On Uniform Convergence of Gaussian Quadrature Rules for Cauchy Principal Value Integrals and Their Derivatives," *Mathematics of Computation*, Vol. 44, pp. 191–198.

Ioakimidis, N. I., 1986, "A One-Dimensional Hypersingular Integral Equation for Axisymmetric Planar Cracks in Three-Dimensional Isotropic Elasticity," *Revue Roumaine des Sciences Techniques-Série de Mécanique Appliquée*, Vol. 31, pp. 391–395.

[1] By H. F. Nied and published in the September, 1987, issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, No. 3, pp. 642–648.

[2] Professor, Division of Applied Mathematics and Mechanics, School of Engineering, University of Patras, GR-261.10 Patras, Greece, Assoc. Mem. ASME.

Ioakimidis, N. I., 1987a, "Validity of the Hypersingular Integral Equation of Crack Problems in Three-Dimensional Elasticity Along the Crack Boundaries," *Engineering Fracture Mechanics*, Vol. 26, pp. 783–788.

Ioakimidis, N. I., 1987b, "Application of Betti's Reciprocal Work Theorem to the Construction of the Hypersingular Integral Equation of a Plane Crack in Three-Dimensional Elasticity," *Journal of Elasticity*, Vol. 18, pp. 165–171.

Ioakimidis, N. I., and Theocaris, P. S., 1979a, "A System of Curvilinear Cracks in an Isotropic Elastic Half-Plane," *International Journal of Fracture*, Vol. 15, pp. 299–309.

Ioakimidis, N. I., and Theocaris, P. S., 1979b, "On the Numerical Solution of Singular Integrodifferential Equations," *Quarterly of Applied Mathematics*, Vol. 37, pp. 325–331.

Author's Closure

The author appreciates the additional information and references provided by Professor Ioakimidis. In the journal article under discussion there was insufficient space to provide a review of the development of the theory of finite-part integrals and hypersingular integral equations. If this had been done, the publications of Professor Ioakimidis would have figured prominently. Since several of the formulas used for numerical computation were taken directly from Kaya and Erdogan (1984, 1985), these references were cited, with the knowledge that Professor Ioakimidis' publications were in turn adequately referenced in those sources.

In addition to the references provide by Professor Ioakimidis, the following references are of historical importance concerning the development and application of finite-part integrals in mechanics: Hadamard (1923), Mangler (1951), and Kutt (1975a, 1975b).

References

Hadamard, J., 1923, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press.

Kaya, A. C., and Erdogan, F., 1984, "On the Solution of Integral Equations with Strong Singularities," *Numerical Solutions of Singular Integral Equations*, Proceedings of IMACS Symposium, Gerasoulis, A., and Vichnevetsky, R., eds., Rutgers University, pp. 54–57.

Kaya, A. C., and Erdogan, F., 1987, "On the Solution of Integral Equations with Strongly Singular Kernels," *Quarterly of Applied Mathematics*, Vol. 45, No. 1, pp. 105–122.

Kutt, H. R., 1975, "The Numerical Evaluation of Principal Value Integrals by Finite-Part Integration," *Numer. Math.*, Vol. 24, pp. 205–210.

Kutt, H. R., 1975, "On the Numerical Evaluation of Finite-Part Integrals Involving an Algebraic Singularity," Special Report WISK 179, Pretoria, National Research Institute for Mathematical Sciences.

Mangler, K. W., 1951, "Improper Integrals in Theoretical Aerodynamics," Royal Aircraft Establishment, Farnborough, Report No. 2424.

A Sixth-Order Plate Theory—Derivation and Error Estimates[3]

C. W. Bert.[4] The author is to be congratulated for making three advances in the theory of small deflections of plates: (1) improvement on Levinson's theory by incorporation of thickness normal strain and transverse isotropy; (2) provision of an *à priori* error estimate for the resulting theory; and (3) reducing the coupled equations of the theory to a single sixth-order equation in the midplane deflection.

In connection with the error estimate, the author attributed the improvement in the error estimate for transverse shear stress over that of the Reissner (1944, 1945) theory to *both* the z^4 term in the deflection and the z^3 term in the in-plane normal stresses. It would be interesting to determine the relative importance of each of the two effects. Also, one would expect that, if the z^4 term in the deflection plays a role, then the z^2 term would also.

Table 1 Coefficient C for isotropic plates according to various sixth-order theories

Theory	Reduced by	C
Reissner (1944, 1945)	Speare and Kemp (1977)	$(2/5)(2-\nu)/(1-\nu)$
	Shirakawa (1983)	$-2/5$
	Discusser (Bert, 1986)	$-2/5$
Ambartsumyan (1970)	Shirakawa (1983)	$-1/(3K)$
Schmidt (1969)	Schmidt (1969)	$(2/3)/[(1-\nu)K]$
Levinson (1980)	Discusser (Bert, 1986)	$-2/5$
Rychter (1987)	Rychter (1987)	$(4/5)/(1-\nu)$

The author's theory still suffers from one small deficiency, from a physical viewpoint, in that the thickness normal stress is neglected. This could be remedied in the case of normal-pressure loaded plates in the way that has been used by Reissner (1944, 1945) and Baluch et al. (1984), for instance.

In the absence of normal pressure, apparently all sixth-order uniform-thickness plate theories, when reduced to a single equation in the midplane deflection, are found to have the following form, using the author's geometry (total plate thickness = $2h$) and other notation:

$$D(1+Ch^2\Delta)\Delta\Delta w=0 \tag{1}$$

The only difference among them is the expression for C, as listed in Table 1.

It is interesting to note that, if one takes the shear correction factor K in the Schmidt *ad hoc* theory to be 5/6, the result is a value exactly equal to that of the present author. However, as first enunciated by Shirakawa (1983) for the Reissner and Ambartsumyan plate theories and recently verified by the discusser for the Reissner and the Levinson theories, the value of C is identical to the coefficient of the second term in Reissner's (1945) second fundamental equation, his equation (22), for an auxiliary stress function ψ,

$$[1-Ch^2\Delta](\psi)=0 \tag{2}$$

If the shear correction K for Ambartsumyan's theory is taken to be 5/6, then the value of C derived by Shirakawa for that theory and for Reissner's (1945) theory coincides with that obtained for Levinson's theory by the discusser. Also, if K in Schmidt's theory is taken to be 5/6, the resulting C value coincides with that of Rychter.

It is noted that the degrees of polynomials in z for the in-plane displacements and the normal deflection used in the author's theory has been derived by the discusser (Bert, 1986) using a Saint-Venant inverse approach that is a generalization of the 1883 work of Clebsch, as presented by Timoshenko and Woinowsky-Krieger (1959).

References

Ambartsumyan, S. A., 1970, *Theory of Anisotropic Plates*, Technomic, Westport, Conn.

Baluch, M. H., Voyiadjis, G. Z., and Azad, A. K., 1984, "A Refined Theory for Isotropic Plates," *Trans., Canadian Soc. of Mech. Engrg.*, Vol. 8, No. 1, pp. 21–27.

Bert, C. W., 1986, "Moderately Thick Plates," in *Theory of Elastic Plate Structures*, Chapter 7, unpublished notes for Course AMNE 5053, University Oklahoma, Norman.

Reissner, E., 1944, "On the Theory of Bending of Elastic Plates," *Journal of Mathematics and Physics*, Vol. 23, pp. 184–191.

Reissner, E., 1945, "The Effect of Transverse Shear Deformation on the Bending of Elastic Plates," ASME Journal of Applied Mechanics, Vol. 12, pp. A69–A77.

Schmidt, R., 1969, "On the Accuracy of Timoshenko's Beam Theory," *Journal of Engrg. Mechanics Division*, Proc. ASCE, Discussion, Vol. 95, No. EM6, pp. 1439–1441.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, 2nd Ed., McGraw-Hill, New York, p. 102.

[3] By Z. Rychter and published in the June, 1987, issue of the ASME Journal of Applied Mechanics, Vol. 54, pp. 275–279.

[4] Perkinson Chair Professor of Engineering, School of Aerospace, Mechanical and Nuclear Engineering, University of Oklahoma, Norman, OK 73019. Fellow ASME.

Ioakimidis, N. I., 1987a, "Validity of the Hypersingular Integral Equation of Crack Problems in Three-Dimensional Elasticity Along the Crack Boundaries," *Engineering Fracture Mechanics*, Vol. 26, pp. 783-788.

Ioakimidis, N. I., 1987b, "Application of Betti's Reciprocal Work Theorem to the Construction of the Hypersingular Integral Equation of a Plane Crack in Three-Dimensional Elasticity," *Journal of Elasticity*, Vol. 18, pp. 165-171.

Ioakimidis, N. I., and Theocaris, P. S., 1979a, "A System of Curvilinear Cracks in an Isotropic Elastic Half-Plane," *International Journal of Fracture*, Vol. 15, pp. 299-309.

Ioakimidis, N. I., and Theocaris, P. S., 1979b, "On the Numerical Solution of Singular Integrodifferential Equations," *Quarterly of Applied Mathematics*, Vol. 37, pp. 325-331.

Author's Closure

The author appreciates the additional information and references provided by Professor Ioakimidis. In the journal article under discussion there was insufficient space to provide a review of the development of the theory of finite-part integrals and hypersingular integral equations. If this had been done, the publications of Professor Ioakimidis would have figured prominently. Since several of the formulas used for numerical computation were taken directly from Kaya and Erdogan (1984, 1985), these references were cited, with the knowledge that Professor Ioakimidis' publications were in turn adequately referenced in those sources.

In addition to the references provide by Professor Ioakimidis, the following references are of historical importance concerning the development and application of finite-part integrals in mechanics: Hadamard (1923), Mangler (1951), and Kutt (1975a, 1975b).

References

Hadamard, J., 1923, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press.

Kaya, A. C., and Erdogan, F., 1984, "On the Solution of Integral Equations with Strong Singularities," *Numerical Solutions of Singular Integral Equations*, Proceedings of IMACS Symposium, Gerasoulis, A., and Vichnevetsky, R., eds., Rutgers University, pp. 54-57.

Kaya, A. C., and Erdogan, F., 1987, "On the Solution of Integral Equations with Strongly Singular Kernels," *Quarterly of Applied Mathematics*, Vol. 45, No. 1, pp. 105-122.

Kutt, H. R., 1975, "The Numerical Evaluation of Principal Value Integrals by Finite-Part Integration," *Numer. Math.*, Vol. 24, pp. 205-210.

Kutt, H. R., 1975, "On the Numerical Evaluation of Finite-Part Integrals Involving an Algebraic Singularity," Special Report WISK 179, Pretoria, National Research Institute for Mathematical Sciences.

Mangler, K. W., 1951, "Improper Integrals in Theoretical Aerodynamics," Royal Aircraft Establishment, Farnborough, Report No. 2424.

A Sixth-Order Plate Theory—Derivation and Error Estimates[3]

C. W. Bert.[4] The author is to be congratulated for making three advances in the theory of small deflections of plates: (1) improvement on Levinson's theory by incorporation of thickness normal strain and transverse isotropy; (2) provision of an *à priori* error estimate for the resulting theory; and (3) reducing the coupled equations of the theory to a single sixth-order equation in the midplane deflection.

In connection with the error estimate, the author attributed the improvement in the error estimate for transverse shear stress over that of the Reissner (1944, 1945) theory to *both* the z^4 term in the deflection and the z^3 term in the in-plane normal stresses. It would be interesting to determine the relative importance of each of the two effects. Also, one would expect that, if the z^4 term in the deflection plays a role, then the z^2 term would also.

Table 1 Coefficient C for isotropic plates according to various sixth-order theories

Theory	Reduced by	C
Reissner (1944, 1945)	Speare and Kemp (1977)	$(2/5)(2-\nu)/(1-\nu)$
	Shirakawa (1983)	$-2/5$
	Discusser (Bert, 1986)	$-2/5$
Ambartsumyan (1970)	Shirakawa (1983)	$-1/(3K)$
Schmidt (1969)	Schmidt (1969)	$(2/3)/[(1-\nu)K]$
Levinson (1980)	Discusser (Bert, 1986)	$-2/5$
Rychter (1987)	Rychter (1987)	$(4/5)/(1-\nu)$

The author's theory still suffers from one small deficiency, from a physical viewpoint, in that the thickness normal stress is neglected. This could be remedied in the case of normal-pressure loaded plates in the way that has been used by Reissner (1944, 1945) and Baluch et al. (1984), for instance.

In the absence of normal pressure, apparently all sixth-order uniform-thickness plate theories, when reduced to a single equation in the midplane deflection, are found to have the following form, using the author's geometry (total plate thickness = $2h$) and other notation:

$$D(1+Ch^2\Delta)\Delta\Delta w=0 \tag{1}$$

The only difference among them is the expression for C, as listed in Table 1.

It is interesting to note that, if one takes the shear correction factor K in the Schmidt *ad hoc* theory to be 5/6, the result is a value exactly equal to that of the present author. However, as first enunciated by Shirakawa (1983) for the Reissner and Ambartsumyan plate theories and recently verified by the discusser for the Reissner and the Levinson theories, the value of C is identical to the coefficient of the second term in Reissner's (1945) second fundamental equation, his equation (22), for an auxiliary stress function ψ,

$$[1-Ch^2\Delta](\psi)=0 \tag{2}$$

If the shear correction K for Ambartsumyan's theory is taken to be 5/6, then the value of C derived by Shirakawa for that theory and for Reissner's (1945) theory coincides with that obtained for Levinson's theory by the discusser. Also, if K in Schmidt's theory is taken to be 5/6, the resulting C value coincides with that of Rychter.

It is noted that the degrees of polynomials in z for the in-plane displacements and the normal deflection used in the author's theory has been derived by the discusser (Bert, 1986) using a Saint-Venant inverse approach that is a generalization of the 1883 work of Clebsch, as presented by Timoshenko and Woinowsky-Krieger (1959).

References

Ambartsumyan, S. A., 1970, *Theory of Anisotropic Plates*, Technomic, Westport, Conn.

Baluch, M. H., Voyiadjis, G. Z., and Azad, A. K., 1984, "A Refined Theory for Isotropic Plates," *Trans., Canadian Soc. of Mech. Engrg.*, Vol. 8, No. 1, pp. 21-27.

Bert, C. W., 1986, "Moderately Thick Plates," in *Theory of Elastic Plate Structures*, Chapter 7, unpublished notes for Course AMNE 5053, University Oklahoma, Norman.

Reissner, E., 1944, "On the Theory of Bending of Elastic Plates," *Journal of Mathematics and Physics*, Vol. 23, pp. 184-191.

Reissner, E., 1945, "The Effect of Transverse Shear Deformation on the Bending of Elastic Plates," ASME Journal of Applied Mechanics, Vol. 12, pp. A69-A77.

Schmidt, R., 1969, "On the Accuracy of Timoshenko's Beam Theory," *Journal of Engrg. Mechanics Division*, Proc. ASCE, Discussion, Vol. 95, No. EM6, pp. 1439-1441.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, 2nd Ed., McGraw-Hill, New York, p. 102.

[3] By Z. Rychter and published in the June, 1987, issue of the ASME Journal of Applied Mechanics, Vol. 54, pp. 275-279.

[4] Perkinson Chair Professor of Engineering, School of Aerospace, Mechanical and Nuclear Engineering, University of Oklahoma, Norman, OK 73019. Fellow ASME.

A Sixth-Order Plate Theory—Derivation and Error Estimates[3]

R. Schmidt.[5] The theory of plates to which the author refers to as "Levinson's theory" and cites Levinson (1980) as his reference is nothing else but a linearized version of the nonlinear theory published by Schmidt (1977). A statement to that effect can be found in *Applied Mechanics Reviews*, Vol. 34, No. 7, July 1981, Review 6280, p. 952. Moreover, Dr. Levinson has acknowledged the identicalness of the two theories in a private letter to the discusser.

The theory of Schmidt (1977) has not been completely unknown. It has been referred to in several journals, including this one (e.g., Sathyamoorthy and Chia, 1980).

In the name of fairness, this oversight should be acknowledged and corrected.

References

Levinson, M., 1980, "An Accurate, Simle Theory of the Statics and Dynamics of Elastic Plates," *Mechanics Research Communications*, Vol. 7, No. 6, pp. 343-350.

Sathyamoorthy, M., and Chia, C. Y., 1980, "Effect of Transverse Shear and Rotatory Inertia on Large Amplitude Vibration of Anisotropic Skew Plates; Part 1—Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, No. 1, pp. 128-132.

Schmidt, R., 1977, "A Refined Nonlinear Theory of Plates with Transverse Shear Deformation," *Industrial Mathematics*, Vol. 27, Part 1, pp. 23-38.

A Sixth-Order Plate Theory—Derivation and Error Estimates[3]

E. Reissner.[6] The unqualified acceptance of results in Speare and Kemp (1977), in conjunction with the author's derivation of analogous results based on an analysis in Levinson (1980), suggests the following observation.

The essence of the results in Speare and Kemp (1977) is that is should be appropriate to reduce the sixth order theory of shear deformable plates in Reissner (1945) to one equivalent sixth order equation

$$D\left\{1+\frac{2-\nu}{1-\nu}\frac{h^2}{10}\right\}\nabla^4 w=q, \tag{1}$$

for the deflection w of a homogeneous isotropic plate, with h designating plate thickness and ∇^2 for the Laplace operator, in place of the author's $2h$ and Δ. Unfortunately, the derivation of equation (1) depended on a fundamental oversight which invalidates this equation as well as its consequences.

To describe what is involved in this matter we depart, as in Spear and Kemp (1977), from the following three equations for shear stress resultants Q_x and Q_y and a load intensity function q,

$$\frac{\partial Q_x}{\partial x}+\frac{\partial Q_y}{\partial y}=-q, \tag{2}$$

$$\left[1-\frac{h^2}{10}\nabla^2\right][Q_x,\ Q_y]=-\left[\frac{\partial}{\partial x},\ \frac{\partial}{\partial y}\right]\left[D\nabla^2 w+\frac{h^2 q}{10(1-\nu)}\right]. \tag{3}$$

In order to deduce equation (1) from equations (2) and (3), the assumption was made that the terms $h^2\nabla^2(Q_x,\ Q_y)$ and $h^2 q$ in equation (3) were small compared with the terms $(Q_x,\ Q_y)$ and $D\nabla^2 w$, respectively, in such a way that it was appropriate to set, *in these terms*, $(Q_x,\ Q_y)=-D(\partial\nabla^2 w/\partial x,\ \partial\nabla^2\partial y)$, with q as in equation (2), so as to have, in place of equation (3)

$$(Q_x,\ Q_y)=-D\left[\frac{\partial}{\partial x},\ \frac{\partial}{\partial y}\right]\left[\nabla^2 w+\frac{h^2}{10}\left[1+\frac{1}{1-\nu}\right]\nabla^4 w\right], \tag{4}$$

with the introduction of equation (4) into (2) leading to equation (1).

While it should be evident that the validity of equation (1) should be suspect in view of the facts that (*i*) the indicated reduction procedure necessarily ceases to be valid for plates acted upon by edge loads only, and (*ii*) the positiveness of the coefficient of the term $h^2\nabla^2$ is associated with physically unreasonable behavior of portions of the solution function w, one can, going beyond this, state the reason for these unacceptable results as follows. It is the essence of shear-deformable plate theory that portions of the terms $h^2\nabla^2(Q_x,\ Q_y)$ in equation (3) are *not* small compared to the corresponding portions of the terms $(Q_x,\ Q_y)$, consistent with the fact that the solution of the sixth order theory of shear deformable plates involves a boundary layer solution contribution which is absent in the classical fourth order Kirchhoff theory.

References

Levinson, M., 1980, "An Accurate Simple Theory of the Statics and Dynamics of Elastic Plates," *Mechanics Research Communications*, Vol. 7, pp. 343-350.

Reissner, E., 1945, "The Effect of Transverse Shear Deformations on the Bending of Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 12, pp. A69-A77.

Speare, P. R. S., and Kemp, K. G., 1977, "A Simplified Reissner Theory for Plate Bending," *International Journal of Solids and Structures*, Vol. 13, pp. 1073-1079.

A Method of Eliminating Lagrangian Multipliers from the Equation of Motion of Interconnected Mechanical Systems[7]

J. G. Papastavridis.[8] Although its final result, i.e., equations (14) are correct, the Note contains a number of erroneous statements:

(*i*) The imposition of the *N catastatic* nonholonomic constraints (1) on the system does *not* affect the number of its (independent or unconstrained or minimal) generalized coordinates; the latter, as we infer from equations (10), are still M in number and not M-N as the Note states.

(*ii*) The constraint reactions, in an ideal system, produce no work for any *virtual* displacement, i.e., displacement compatible with the instantaneous/frozen constraints (in the form (1), but with $\dot{q}$ replaced with δq), and not just ". . . any displacement compatible with the constraints," as that author states (near the top of p. 236). Thus the following "zero reaction power" condition (between his equations (7′) and (8)) holds *only for catastatic* (i.e., homogeneous) nonholonomic constraints *such as* (*1*), but *not for general acatastatic* (i.e., nonhomogeneous) ones; for the relevant definitions see Rosenberg (1977).

[3] See footnote 3, p. 249.

[5] Professor of Engineering Mechanics, College of Engineering and Science, University of Detroit, Detroit, MI 48221. Mem. ASME

[6] Department of Applied Mechanics and Engineering Sciences, University of California at San Diego, La Jolla, CA 92093.

[7] By S. Vlase, published in the March, 1987, issue of ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 235-237.

[8] School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

A Sixth-Order Plate Theory—Derivation and Error Estimates[3]

R. Schmidt.[5] The theory of plates to which the author refers to as "Levinson's theory" and cites Levinson (1980) as his reference is nothing else but a linearized version of the nonlinear theory published by Schmidt (1977). A statement to that effect can be found in *Applied Mechanics Reviews*, Vol. 34, No. 7, July 1981, Review 6280, p. 952. Moreover, Dr. Levinson has acknowledged the identicalness of the two theories in a private letter to the discusser.

The theory of Schmidt (1977) has not been completely unknown. It has been referred to in several journals, including this one (e.g., Sathyamoorthy and Chia, 1980).

In the name of fairness, this oversight should be acknowledged and corrected.

References

Levinson, M., 1980, "An Accurate, Simle Theory of the Statics and Dynamics of Elastic Plates," *Mechanics Research Communications*, Vol. 7, No. 6, pp. 343-350.

Sathyamoorthy, M., and Chia, C. Y., 1980, "Effect of Transverse Shear and Rotatory Inertia on Large Amplitude Vibration of Anisotropic Skew Plates; Part 1—Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, No. 1, pp. 128-132.

Schmidt, R., 1977, "A Refined Nonlinear Theory of Plates with Transverse Shear Deformation," *Industrial Mathematics*, Vol. 27, Part 1, pp. 23-38.

A Sixth-Order Plate Theory—Derivation and Error Estimates[3]

E. Reissner.[6] The unqualified acceptance of results in Speare and Kemp (1977), in conjunction with the author's derivation of analogous results based on an analysis in Levinson (1980), suggests the following observation.

The essence of the results in Speare and Kemp (1977) is that is should be appropriate to reduce the sixth order theory of shear deformable plates in Reissner (1945) to one equivalent sixth order equation

$$D\left\{1+\frac{2-\nu}{1-\nu}\,\frac{h^2}{10}\right\}\nabla^4 w=q, \tag{1}$$

for the deflection w of a homogeneous isotropic plate, with h designating plate thickness and ∇^2 for the Laplace operator, in place of the author's $2h$ and Δ. Unfortunately, the derivation of equation (1) depended on a fundamental oversight which invalidates this equation as well as its consequences.

To describe what is involved in this matter we depart, as in Spear and Kemp (1977), from the following three equations for shear stress resultants Q_x and Q_y and a load intensity function q,

$$\frac{\partial Q_x}{\partial x}+\frac{\partial Q_y}{\partial y}=-q, \tag{2}$$

$$\left[1-\frac{h^2}{10}\nabla^2\right][Q_x,\ Q_y]=-\left[\frac{\partial}{\partial x},\ \frac{\partial}{\partial y}\right]\left[D\nabla^2 w+\frac{h^2 q}{10(1-\nu)}\right]. \tag{3}$$

In order to deduce equation (1) from equations (2) and (3), the assumption was made that the terms $h^2\nabla^2(Q_x,\ Q_y)$ and $h^2 q$ in equation (3) were small compared with the terms $(Q_x,\ Q_y)$ and $D\nabla^2 w$, respectively, in such a way that it was appropriate to set, *in these terms*, $(Q_x,\ Q_y)=-D(\partial\nabla^2 w/\partial x,\ \partial\nabla^2\partial y)$, with q as in equation (2), so as to have, in place of equation (3)

$$(Q_x,\ Q_y)=-D\left[\frac{\partial}{\partial x},\ \frac{\partial}{\partial y}\right]\left[\nabla^2 w+\frac{h^2}{10}\left[1+\frac{1}{1-\nu}\right]\nabla^4 w\right], \tag{4}$$

with the introduction of equation (4) into (2) leading to equation (1).

While it should be evident that the validity of equation (1) should be suspect in view of the facts that (*i*) the indicated reduction procedure necessarily ceases to be valid for plates acted upon by edge loads only, and (*ii*) the positiveness of the coefficient of the term $h^2\nabla^2$ is associated with physically unreasonable behavior of portions of the solution function w, one can, going beyond this, state the reason for these unacceptable results as follows. It is the essence of shear-deformable plate theory that portions of the terms $h^2\nabla^2(Q_x,\ Q_y)$ in equation (3) are *not* small compared to the corresponding portions of the terms $(Q_x,\ Q_y)$, consistent with the fact that the solution of the sixth order theory of shear deformable plates involves a boundary layer solution contribution which is absent in the classical fourth order Kirchhoff theory.

References

Levinson, M., 1980, "An Accurate Simple Theory of the Statics and Dynamics of Elastic Plates," *Mechanics Research Communications*, Vol. 7, pp. 343-350.

Reissner, E., 1945, "The Effect of Transverse Shear Deformations on the Bending of Elastic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 12, pp. A69-A77.

Speare, P. R. S., and Kemp, K. G., 1977, "A Simplified Reissner Theory for Plate Bending," *International Journal of Solids and Structures*, Vol. 13, pp. 1073-1079.

A Method of Eliminating Lagrangian Multipliers from the Equation of Motion of Interconnected Mechanical Systems[7]

J. G. Papastavridis.[8] Although its final result, i.e., equations (14) are correct, the Note contains a number of erroneous statements:

(*i*) The imposition of the *N catastatic* nonholonomic constraints (1) on the system does *not* affect the number of its (independent or unconstrained or minimal) generalized coordinates; the latter, as we infer from equations (10), are still M in number and not M-N as the Note states.

(*ii*) The constraint reactions, in an ideal system, produce no work for any *virtual* displacement, i.e., displacement compatible with the instantaneous/frozen constraints (in the form (1), but with $\dot{q}$ replaced with δq), and not just ". . . any displacement compatible with the constraints," as that author states (near the top of p. 236). Thus the following "zero reaction power" condition (between his equations (7′) and (8)) holds *only for catastatic* (i.e., homogeneous) nonholonomic constraints *such as* (*1*), but *not for general acatastatic* (i.e., nonhomogeneous) ones; for the relevant definitions see Rosenberg (1977).

[3] See footnote 3, p. 249.

[5] Professor of Engineering Mechanics, College of Engineering and Science, University of Detroit, Detroit, MI 48221. Mem. ASME

[6] Department of Applied Mechanics and Engineering Sciences, University of California at San Diego, La Jolla, CA 92093.

[7] By S. Vlase, published in the March, 1987, issue of ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 235-237.

[8] School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

(*iii*) The method proposed by the author, in particular his equations (12)–(14), are neither "new" nor "natural". Instead they constitute a special application of the general reaction-free equations of Maggi (presented by him first in 1896, and then elaborated in 1901—see, e.g., Neimark and Fufaev, 1967), *when one chooses the last M-N generalized velocities* (author's $\dot{q}_c$) *as the independent parameters or quasivelocities*. This particular choice of quasivelocities ("elimination" or "embedding" of constraints) dates back from Chaplygin (1895, 1897) and Voronets (1901); these two also formulated reaction-free equations of motion that are special cases of the general nonholonomic Boltzmann (1902)/Hamel (1903, 1094) equations. That special case of Maggi's equations has been derived and discussed by such mechanicians as Hamel (1924), and Lur'e (1961) in his monumental monograph (see also Appell 1953, or earlier editions, and Rosenberg, 1977). The same special case was also independently rediscovered by Passerello and Huston (1973) in a rather ad hoc and unmotivated fashion, as an isolated result; on the contrary, both Hamel and Lur'e discuss the relation of Maggi's equations with other sets of nonholomonic systems equations such as those by Appell, Chaplygin, Voronets, Tzenov. It should be added here that these special Maggi equations hold for general rheonomic systems (i.e., nonstationary in their holonomic and nonholonomic constraints), and not only for scleronomic (i.e., stationary constraint) ones, as the author's equations (9) and (1) imply.

As Prange (1935) puts it, *Maggi's equations are the projection of the general Lagrange/Routh/Voss equations (10) onto the M-N dimensional "nonholonomic manifold" or "virtual hyperplane" of the system at the "point" (q, t) in configuration space* (that plane is defined by the Note's constraint equations (1) with $\dot{q}$ replaced by the virtual δq). Since the system constraint-reaction vector is perpendicular to that hyperplane, Maggi's equations are reaction-free! Several other authors have presented their nonholonomic system discussions in simple and fruitful geometrical/tensorial language (see, e.g., Dobronravov, 1970).

References

Appell, P., 1953, *Traité de Mécanique Rationelle*, Tome Deuxiéme, Sixth Edition, Gauthier-Villars, Paris, pp. 407–408.

Dobronravov, V. V., 1970, *Fundamentals of Nonholonomic System Mechanics*, in Russian, Vischaya Shkola, Moscow, pp. 147–190.

Hamel, G., 1924, "Uber Nichtholonome Systeme," *Mathematische Annalen*, Vol. 92, pp. 33–41.

Lur'e, A. I., 1968, *Mécanique Analytique*, 2 volumes, Librairie Universitaire, Louvain, Belgium, pp. 395–398 (especially equations (8.7.6)); Russian original, 1961.

Neimark, J. I., and Fufaev, N. A., 1972, *Dynamics of Nonholonomic Systems*, American Mathematical Society, Providence, RI, pp. 119–120; Russian original, 1967.

Passerello, C. E., and Huston, R. L., 1973, "Another Look at Nonholonomic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, pp. 101–104, eqns. (24), (28).

Prange, G., 1904–1935, "Die Allgemeinen Integrationsmethoden der Analytischen Mechanik," Vol. 4, pp. 505–804; *Encyklopaedie der Mathematischen Wissenschaften*, Teubner, Leipzig, pp. 557–560.

Rosenberg, R. M., 1977, *Analytical Dynamics of Discrete Systems*, Plenum, New York, pp. 43–45, pp. 243–245.

Author's Closure

The method is the result of the author's concerns toward conceiving a computer program for the elastodynamic analysis of mechanisms. In this respect, the computing algorithm must determine the constraint reactions.

Many difficulties have been encountered in making a program that first builds the reactionless equations of motion, and then computes the reaction forces concerning its efficiency. It is therefore more convenient to generate equations (10) only once and then operate with them.

Although the basic idea of our method has *implicitly* been found in classical analytical mechanics, as documented by Papastravridis, its *explicit* form has not been used yet. The method can therefore be considered new. Dynamic systems analysis programs eliminate multipliers in equations (10) by using a coefficient matrix inversion technique. The method allows the replacement of all the abovementioned procedures by a simple matrix multiplication.

A Method of Eliminating Lagrangian Multipliers from the Equations of Motion of Interconnected Mechanical Systems[7]

L. Y. Bahar.[9] The author should be commended for an interesting approach of first adjoining constraints via Lagrange multipliers, then eliminating them.

A simplier and more natural method consists of not introducing Lagrange multipliers at all, but utilizing the well established approach of "constraint embedding," most elegantly described in a recent text by Rosenberg [1], and generally attributed to Woronetz. This deviation is elementary and straightforward, as it consists of the variational counterpart of finding constrained extrema in elementary calculus through the process of elimination of dependent differentials. It is recognized that this approach is particulary useful when implicit differentiation is involved.

The basic idea in [1] is to begin with the fundamental equation of dynamics (or the principle of virtual work) stated through Lagrange's form of d'Alembert's principle, followed by the elimination of the excess or dependent virtual displacements by expressing them in terms of the independent (minimal set) of virtual displacements through the use of nonholonomic constraints, or the variation of holonomic constraints. The vanishing of the coefficients then yields the equations governing the motion. For algebraic details, as well as an illustrative example, [1] may be consulted.

It should be pointed out that the above method is in keeping with the spirit of analytical dynamics, where the concept of virtual work is utilized in order to eliminate the total virtual work performed by all the reaction forces undergoing virtual displacements compatible with the constraints. It is thus possible to obtain equations of motion that reflect only the impressed forces. Once these equations of motion have been obtained, it is a simple procedure to determine the constraint forces at a later stage, through the introduction of Lagrange multipliers. In contrast with the Newtonian approach in which all the reactions and external forces appear in the equations of motion (through the free-body diagram method, for example), in Lagrangian dynamics the problem is divided into two logically separate steps. In the first step, the reactionless equations of motion are formulated and solved in terms of the known impressed forces (i.e., generalized coordinates and their time derivatives are obtained in terms of the impressed forces by analytical and/or numerical means). The dynamical quantities thus arrived at as outputs of the first stage of the problem are now used as inputs to determine the constraint reactions that arise in the second stage of the problem.

Reference

1 Rosenberg, R. M., *Analytical Dynamics of Discrete Systems*, Plenum Press, New York, pp. 243–246.

[7] See footnote 7, p. 250.

[9] Department of Mechanical Engineering and Mechanics, Drevel University, Philadelphia, PA 19104.

Journal of Applied Mechanics

Book Reviews

Mathematical Models in Applied Mechanics, by A. B. Tayler. Oxford University Press, New York, 1986. 280 Pages. Price: $29.95/cloth; $14.95/paper.

REVIEWED BY J. B. KELLER[1]

This very readable book is concerned primarily with the formulation of mathematical models of physical phenomena, especially in mechanics. It teaches modeling by example, using twenty-two typical problems which have arisen from industrial research. In each case it explains the physical background of the situation and then shows how to formulate a mathematical problem to model it. This is followed by some mathematical analysis of the problem, with the purpose of illustrating some important features of the solution.

All of the mathematical models involve partial differential equations. They are grouped into three chapters on hyperbolic, elliptic, and parabolic problems, respectively. These chapters are preceded by one which describes nondimensionalization, some perturbation procedures, some methods for solving ordinary differential equations, and some basic equations of fluid dynamics and elasticity. A final chapter presents asymptotic methods for nonlinear problems, such as boundary layer and ray methods. Each chapter ends with a set of exercises.

The variety of problems is great as the following selection indicates: Pantograph dynamics, coal seam exploration, flow of granular material, electrochemical painting, percolation in a sand dune, hot rolling of steel, smoke dispersion from a chimney, welding of two steel plates, spreading of an oil film, injection moulding, seismic detection.

The book can be read by anyone with some knowledge of differential equations and mechanics. It could be used as a text for a course on modeling or as a supplementary text for a course on partial differential equations. Since the presentation is lively and informal, it is a pleasure to read just to learn about a large number of interesting applications.

[1]Professor, Mathematics Department, Stanford University, Stanford, CA 94305.

Solid Mechanics Research for Quantitative Non-Destructive Evaluation, edited by J. D. Achenbach and Y. Rajapakse. Martinus Nijhoff Publishers, The Netherlands, 1987. 438 Pages price: $122.00.

REVIEWED BY A. K. GAUTESEN[2]

This book contains the proceedings of the ONR Symposium on Solid Mechanics Research for Quantitative Non-Destructive Evaluation (QNDE) which was held at Northwestern University in September, 1985. Here, the state of the art in QNDE was reviewed along with the methods and techniques being investigated at the time.

With the advent of modern technology it has become possible to detect very small inhomogeneities in materials by nondestructive evaluation procedures. Since some of the detected inhomogeneities will not affect the serviceable life of the material, the need has arisen for more quantitative information about the inhomogeneity—such as location, size, shape, and orientation. This need has given birth to a more fundamental and rigorous approach called QNDE. Disciplines spanned by this approach include mechanics of solids, materials science, electrical engineering, applied physics, applied mathematics, and computer science. Although significant progress has been made, QNDE is still in its infancy and fundamental deficiencies exist in many areas. This field continues to be the focus of intense research.

This book is organized into nine chapters. In Chapter 1, R. deNale and D. E. Chimenti describe requirements in QNDE for the Navy and Air Force, respectively. In Chapter 2, H. N. G. Wadley discusses the use of acoustic emission as a QNDE technique for the study of fracture; and W. Sachse summarizes the application of quantitative acoustic emission measurements for the investigation of dynamic fracture processes and for the characterization of materials and transducers. In Chapter 3, J. D. Achenbach reviews some aspects of flaw characterization by ultrasonic scattering

[2]Senior Mathematician and Professor of Mathematics, Dept. of Mathematics and Ames Laboratory, Iowa State University, Ames, IA 50011.

methods, and L. Adler discusses related experimental results. In Chapter 4, J. N. Yang describes the application of QNDE results to retirement-for-cause analysis; R. B. Thompson, O. Buck, and D. K. Rehbein review the present state of theoretical understanding of elastic wave interactions with partially contacting interfaces; and A. Vary advances three conceptual models for interrelating ultrasonic attention, microstructure, and fracture toughness. In Chapter 5, G. A. D. Briggs and M. G. Somekh review recent developments in acoustic microscopy of surface cracks; and B. T. Khuri-Yakub and P. A. Reinholdtsen present experimental results obtained by an acoustic microscope. In Chapter 6, Y. Bar-Cohen reviews the ultrasonic NDE of composites; and in separate articles by I. M. Daniel and S. C Wooh, and by E. G. Henneke, II, J. C. Duke and R. C. Stiffler, the characterization of damage in composite laminates is studied. In Chapter 7, R. L. Thomas, L. D. Favro, and P. K. Kuo review various techniques for making thermal wave images of defects and maps of material characteristics of opaque solids. Chapter 8 begins with a review by Y. H. Pao of the theories of acoustoelasticity and of acoustoplasticity, which is followed by three papers (by H. Fukuoka; M. Namkung, D. Utrata, J. S. Heyman, and S. G. Allison; and C. M. Sayers) on the measurement of residual stresses. The final chapter contains eight contributed papers.

This book offers a valuable collection of expository articles and is to be recommended to the reader for its excellent overview of current directions of research and latest results in the field of QNDE.

Theory of Plasticity, by J. Chakrabarty. McGraw-Hill Book Company, New York, 1987. 791 pages.

REVIEWED BY W. J. DRUGAN[3]

Since the pioneering work of Tresca in 1864, many talented engineers and applied mathematicians have contribuited to the theory of metal plasticity, developing an elegant framework capable of describing many nonlinear physical phenomena, and devising clever solutions to numerous technological problems. This book describes certain parts of this theory and its applications. On the topics it covers, the book is extraordinarily thorough, providing detailed analyses and discussions, and numerous references to the technical literature. It is well-written, makes generous use of figures, and reports many sets of experimental results. There is also a long list of problems at the end of each chapter. My main difficulty with the book is that, despite its substantial length (791 pages), it is too narrow in scope: many important advances of the past 25 years or so go unmentioned. Perhaps this is to be rectified by a second volume that the Preface implies is forthcoming, but the author has permitted the book's limited scope to affect the perspective presented in the present volume. For example, on Page 1, the author states "The development of anisotropy with progressive cold work and the resulting strain-hardening are too complex to be successfully incorporated in the theoretical framework. In the mathematical theory of plasticity, it is generally assumed that the material remains isotropic throughout the deformation irrespective of the degree of cold work." Another example occurs on Page 55, in a general discussion of the yield criterion concept: "In developing a mathematical theory, it is necessary to take into account a number of idealizations at the outset. Firstly, it is assumed that the conditions of loading are such that all strain rate and thermal effects can be neglected. Secondly, the Bauschinger effect and the hysteresis loop, which arise from nonuniformity on the microscope scale, are disregarded. Finally, the material is assumed to be isotropic..." These statements and others may be misleading to a graduate student or researcher attempting to learn the current status of the theory.

The book if organizaed as follows. Chapter 1 briefly describes the physical mechanisms of plastic deformation, provides a fairly detailed discussion of uniaxial stress-strain behavior, and reviews the analysis of stress and strain, including objective and nominal stress rates. Some of the discussion of stress and strain could be improved by using tensor notation or making more use of index notation. Chapter 2 is the best in the book. Although restricted to initially isotropic response of rate-independent materials, this chapter gives a clear, well-referenced, relatively concise presentation of isotropic yield criteria and stress-strain equations, with comparisons of their predictions to experimental data, isotropic and kinematic hardening, general; plastic flow rules and Drucker's postulate, Hencky's "deformation" theory and Budiansky's juistification of its use at yield surface vertices; theorems of limit analysis, and uniqueness theorems and extremum principles for small and large deformations.

The remainder of the book gives applications of the isotropic elastic-plastic theory characterized by the Mises yield condition and Prandtl-Reuss flow rule (or Tresca yield condition and associated flow rule), or its rigid-plastic simplification, assuming nonhardening response for the most part. Except for Chapter 4, which is a quite comprehensive account of the plastic analysis of beams and frames (including limit analysis of beams, frames, arches, columns, variable repeated loading, and minimum weight design), the remainder (indeed, most) of the book is largely patterned after Rodney Hill's superb *The Mathematical Theory of Plasticity* (Oxford, 1950; 1983 paperback). This contributes greatly to the book's narrowness of focus, since many new areas and perspectives within plasticity have developed in the thirty-seven years since Hill organized the then-extant body of research. Chapters 3 and 5, then, are lengthy (one is tempted to say exhaustive) collections of analytical solutions to elastic-plastic problems having simple geometries, such as problems involving bending, torsion, spherical and axial symmetry. Most of the work reported in Chapters 1–5 was done prior to 1960, although Chapter 5 does have a nine-page subsection on the finite element method. Chapter 6 lays out plane strain rigid-ideally plastic slip-line theory, the application of which to numerous problems of steady and nonsteady flow is reported in detail in Chapters 7 and 8. One of the book's greatest strengths is the completeness and clarity with which it presents and interprets the solutions of the specific problems reported in Chapters 3–5 and 7–8, together with the references provided both to the original sources and to more recent work related to these problems, such as experimental results, analyses including work-hardening, and numerical analyses.

As noted, perhaps the major drawback of the book is its omission of several important advances and areas that have developed in the past 25 years or so. To be specific, I would like to have seen included treatments of: general anisotropic elastic-plastic response, including anisotropic initial yield; modern constitutive formulations for large elastic-plastic deformations; thermodynamics of elastic-plastic materials; the micromechanical foundations of macroscopic constitutive equations; plastic flow localization; constitutive equations that account for void formation and growth, and those that exhibit non-normality; elastic-plastic wave propagation; rate-dependent plasticity; plasticity of nonmetals, such as soil, ice, etc.; and a much more expansive coverage of numerical methods for plasticity. This being understood, the book is recommended for its treatment of the topics it does cover.

[3]Associate Professor, Department of Engineering Mechanics, University of Wisconsin, Madison, WI 53706. Assoc. Mem., ASME.

Turbulence in Fluids, by M. Lesieur. Martinus Nijhoff Publishers, The Netherlands, 1987. 286 pages. Price: $68.50.

REVIEWED BY REDA R. MANKBADI[4]

There are two different approaches to the closure problem of turbulence. One, which is traditionally called "Semiempirical theories of turbulence," is based on the closure of the equations of the mean fluid dynamics fields. The other is the statistical approach which is appropriate for the study of isotropic turbulence and is based on special assumptions about spectral characteristics of turbulent fluctuations. Professor Lesieure's book falls into the second approach. The author presents his experience in studying stochastic models of turbulence as applied to various situations. The author attempts to reconcile the statistical point of view and the basic concepts of fluid mechanics which determine the evolution of flows arising in various engineering fields.

The book presents a synthesis between two graduate courses given in Grenoble by the author, namely a "Tubulence" course and a "Geophysical Fluid Dynamics" course. This book is written for an audience already familiar with the field and interested more in the stochastic modelling approach to the turbulence problem.

The book consists of thirteen chapters. The first two chapters discuss the definition of turbulent flows and present the fundamental laws governing the development of turbulent flows. The objective of chapter three is to give the physical insight to various ways through which a flow can become turbulent, with emphasis on the instabilities which develop from a basic shear flow. In chapter four the various Fourier representations of statistically homogeneous turbulent flows are presented and the Navier-Stokes equation is projected in Fourier Space. This is followed in chapter five by discussing the kinematics of homogenous turbulence. Several of the basic ideas of the closure problem in statistical turbulence are presented in chapter six.

The first six chapters of the book serve as an introduction to chapter seven which can be considered as the core chapter. This chapter gives the reader a good understanding of the analytical theories and stochastic models. The Eddy-Damped-Quasi-Normal Morkovian approximations are mainly used. But several other theories are presented for the stochastic model point of view. The results of these theories concerning passive scalar diffusion is given in chapter eight. Chapter nine can serve as an introduction to the application of stochastic models to two-dimensional Quasi-Geostrophic turbulence. Chapter ten discusses the absolute equilibrium ensembles and the two-dimensional turbulence over topography.

Chapter eleven is concerned with the closure of the Eddy-Damped-Quasi-Normal Morkovian for two-and three-dimensional isotropic turbulence. Large-Eddy- Simulation of turbulence have allowed us to study the dynamics of stably stratified turbulence, where internal gravity waves and turbulence interact. This is of great importance for the mesoscale atmosphere and the mixing process in the ocean. Therefore, chapter twelve is devoted to this subject with emphasis on two-dimensional isotropic turbulence. In the previous chapters isotropy is always assumed. In chapter thirteen two practical turbulent flow problems, that do not fall into the group of isotropic turbulence, are considered. The first problem is that of a three-dimensional homogeneous turbulence submitted to a stable stratification. The second problem is that of a two-dimensional turbulence resulting from the development of the inflexional shear instability in the case of the mixing layer.

The book presents a limited number of topics and techniques which are not usually encountered in text books in this field. The treatment of those topics is quite clear and skillful. The stochastic modelling tools are extremely effective to understand the energy transfers between the different scales of motion. This is a very useful book and most specialists in turbulence will benefit and enjoy reading it as well.

[4]Department of Mechanical Engineering, Cairo University, Cairo, Egypt.

ERRATA

Errata on "A Mixture Model for Unidirectionally Fiber-Reinforced Composites," by H. Murakami and G. A. Hegemier, and published in the December 1986 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 765–773.

Equation (42*b*) should be read as

$$\beta_2 = \beta_3 = 1/\left\{ \sum_{\alpha=1}^{2} h^{(\alpha)} (\lambda + 3\mu)^{(\alpha)} / 8\mu^{(\alpha)} (\lambda + \mu)^{(\alpha)} \right\}. \tag{42b}$$

Similarly, equation (43*g*) should be corrected to

$$\overset{2}{M}_{31} = \overset{3}{M}_{12} = 0, \quad \overset{2}{M}_{23} = \frac{h}{4} P_3, \quad \overset{3}{M}_{23} = \frac{h}{4} P_2. \tag{43g}$$