Transactions of the ASME®

Technical Editor, LEON M. KEER (1993)

The Technological Institute
Northwestern University
Evanston, IL 60201

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.) for $120 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Second class postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.

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

PRICES: To members, $29.00, annually; to nonmembers, $120.00. Add $15.00 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 56 • NUMBER 3 • SEPTEMBER 1989

(Contents continued on Inside Back Cover)

CONTENTS (CONTINUED)

BRIEF NOTES

DISCUSSION

BOOK REVIEWS

M. Rides
National Physical Laboratory,
Teddington, Middlesex, England

A. C. F. Cocks
Department of Engineering,
Leicester University,
Leicester, England

D. R. Hayhurst
Department of Mechanical Engineering,
Sheffield University,
Sheffield, England

The Elastic Response of Creep Damaged Materials

A series of stress change experiments on a batch of tough pitch copper are presented which were devised to evaluate the variation of Young's modulus with creep damage. Kachanov's model is used to describe the creep response and a model orginally proposed by Chaboche (1979) is found to adequately represent the elastic response of the material. A simple two-bar structure is analyzed to assess the effect of including the variation of elastic properties with creep damage in structural analysis. In most practical situations the added complexity involved in incorporating this effect does not strongly affect the structural response.

1 Introduction

Constitutive equations for the response of creep damaging materials have either been developed phenomenologically (Kachanov, 1958; Rabotnov, 1969; Leckie and Hayhurst, 1975) or from an understanding of the internal microscopic processes that take place as the material creeps (Cocks and Ashby, 1982; Tvergaard, 1986). Despite the differences of these two approaches, the general structure of constitutive laws that are obtained are very similar. In each case the strain rate at a given instant in time is expressed in terms of the stress σ and a number of state variables, ω_i, which measure the extent of damage in the material. The formalism is completed by developing laws for the rate of increase of the damage in terms of σ and ω_i. In each of these approaches it is generally assumed that the damage does not affect the elastic properties of the material.

The objective of this paper is to investigate the effect of creep damage on the elastic properties and examine how this influences the predicted response of component behavior. We limit our attention to material behavior in uniaxial tension, and describe the material response using a single state variable ω. Following the philosophy of the modeling of the creep response, we express the compliance at a given instant in time in terms of ω. We could follow the mechanistic approach and assign a physical interpretation to ω, such as the volume fraction of voids or the proportion of cracked grain boundaries, and use the theoretical results of Duva and Hutchinson (1984) and Budiansky and O'Connell (1976) for the elastic response of porous and cracked media to construct appropriate constitutive relationships. Here we prefer the phenomenological approach and define the state variable from the shape of the creep curve for the material. Young's modulus for a given value of ω is then determined from a series of elastic unloading and reloading tests. The measured response is compared with the proposal of Chaboche (1979) for the variation of modulus with creep damage which is based on the physical interpretation of the damage variable proposed by Odqvist and Hult (1961).

The proposed material model is used to assess the response of a simple two-bar structure to a constant load in Section 4, to give an indication of the likely effect of including the variation of Young's modulus in component analysis.

2 A Phenomenological Description of Creep Failure in Uniaxial Tension

The formalism adopted here was first proposed by Kachanov (1958) and later generalized by Rabotnov (1969). For simplicity we ignore the effects of primary creep; the inelastic strain rate at a given instant in time is then given by

$$\dot{\epsilon} = \frac{\dot{\epsilon}_o (\sigma/\sigma_o)^n}{(1-\omega)^\phi} \tag{1}$$

where n and ϕ are material constants and $\dot{\epsilon}_o$ is the steady-state strain rate at a suitably chosen reference stress σ_o. The state variable ω varies from a value of 0 at the start of a test to 1 at failure. The rate of growth of damage at a given instant can be obtained from the relationship

$$\dot{\omega} = \frac{\dot{\omega}_o (\sigma/\sigma_o)^\nu}{(1-\omega)^\psi} \tag{2}$$

where ν and ψ are material constants and $\dot{\omega}_o$ is the initial damage rate at a stress σ_o. The tests described in this paper were conducted at constant load and the stress used in equations (1) and (2) is the nominal stress (load/original cross-sectional area). The damage, ω, in these equations then has two contributions: from the reduction of cross-sectional area and the growth of internal voids and cracks.

During steady-state creep, $\omega = 0$ and the material creeps at a rate $\dot{\epsilon}_{ss}$, where

$$\dot{\epsilon}_{ss} = \dot{\epsilon}_o (\sigma/\sigma_o)^n. \tag{3}$$

Figure 1 summarizes the steady-state response over a range

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 20, 1987; final revision, April 6, 1988.

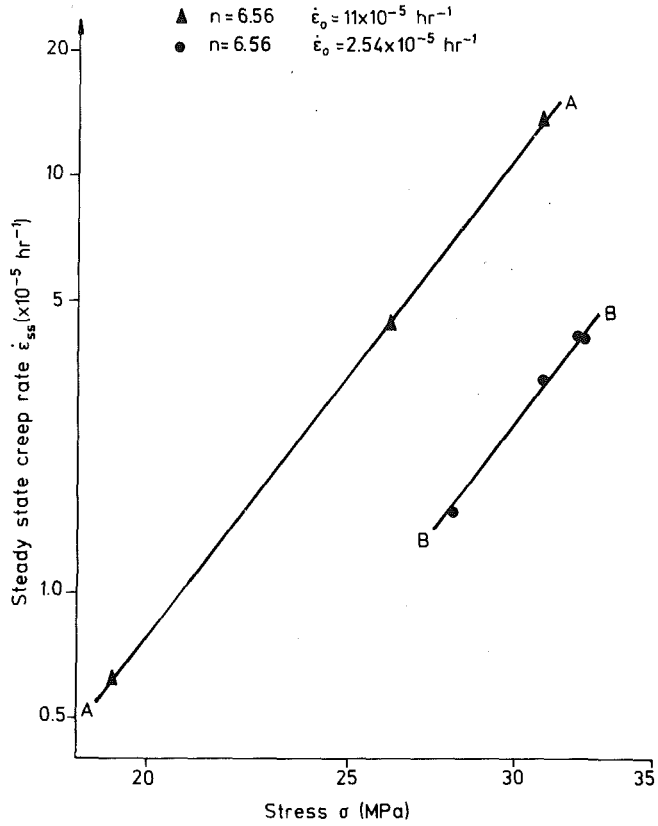


Fig. 1 Logarithmic plots of steady-state strain, rate versus stress for the two batches of copper

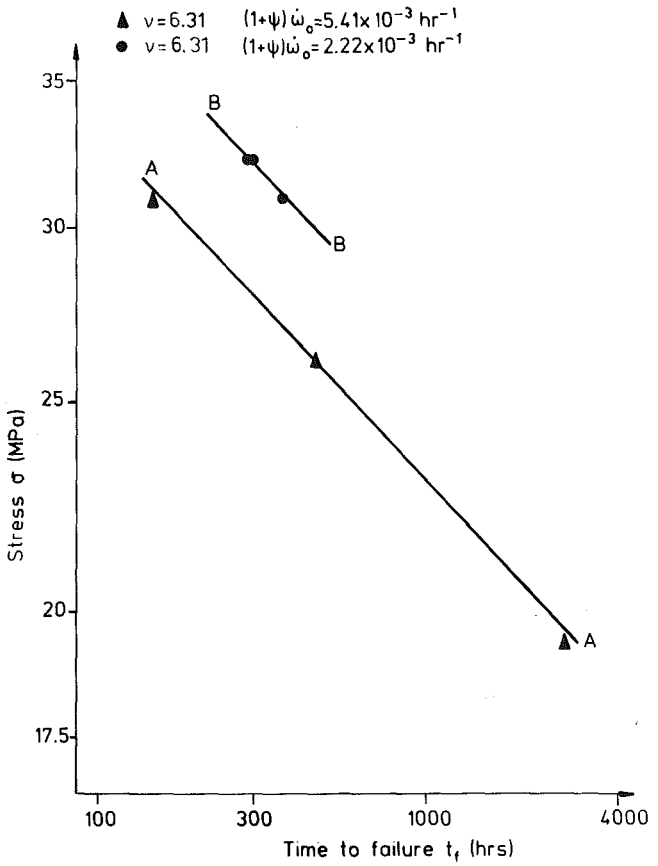


Fig. 2 Logarithmic plots of stress versus time to failure for the two batches of copper

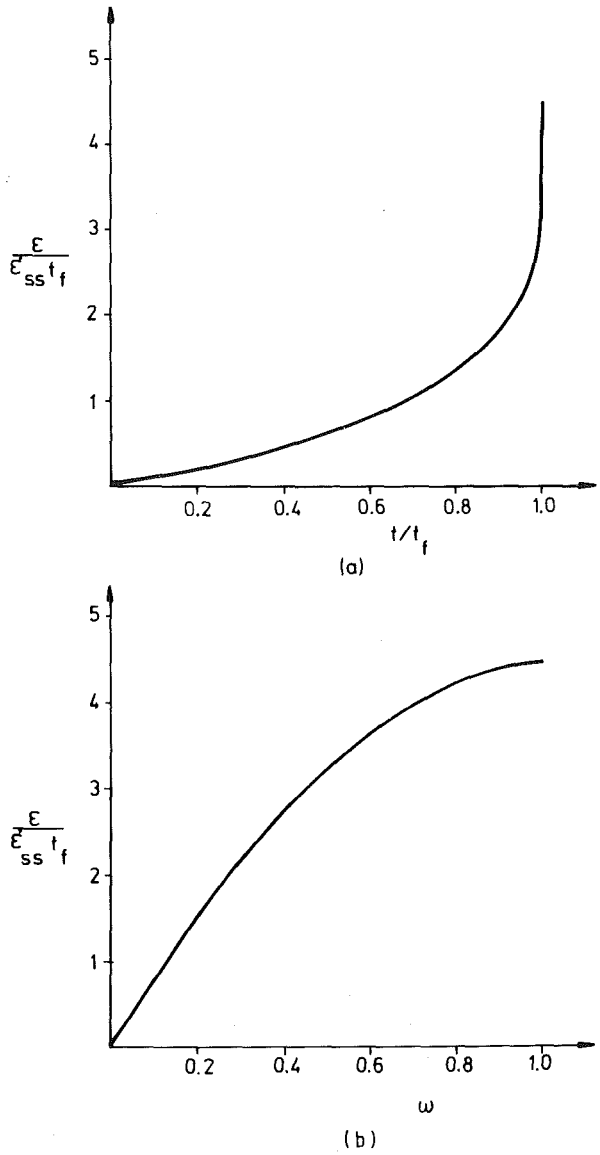


Fig. 3 Normalized plots of (a) strain against time and (b) strain against damage

of stress states for two batches of tough pitch copper tested at 300°C (designated batch A and batch B). Also indicated on Fig. 1 are the values of n and $\dot{\epsilon}_o$ obtained from the plots, where $\dot{\epsilon}_o$ was evaluated using a value of $\sigma_o = 30$ MPa.

Integration of equation (2) between the limits $\omega = 0$ at $t = 0$ and $\omega = 1$ at $t = t_f$ at constant stress gives

$$t_f = \frac{1}{(1+\psi)\dot{\omega}_o}\left(\frac{\sigma_o}{\sigma}\right)^{\nu} \tag{4}$$

Stress-life curves for the two batches of copper are given in Fig. 2 where the values of $(1+\psi)\,\dot{\omega}_o$ and ν obtained by fitting equation (4) are indicated. A value of $\sigma_o = 30$ MPa was again used in these calculations.

Integration of equations (1) and (2) between the limits $\omega = \epsilon = 0$ at $t = 0$ and $\omega = \omega$ and $\epsilon = \epsilon$ at $t = t$ gives

$$\omega = 1 - \left(1 - \frac{t}{t_f}\right)^{\frac{1}{1+\psi}} \tag{5}$$

and

$$\frac{\epsilon}{\epsilon_{ss} t_f} = \left(\frac{1+\psi}{1+\psi-\phi}\right)\left[1 - \left(1 - \frac{t}{t_f}\right)^{\frac{1+\psi-\phi}{1+\psi}}\right]. \tag{6}$$

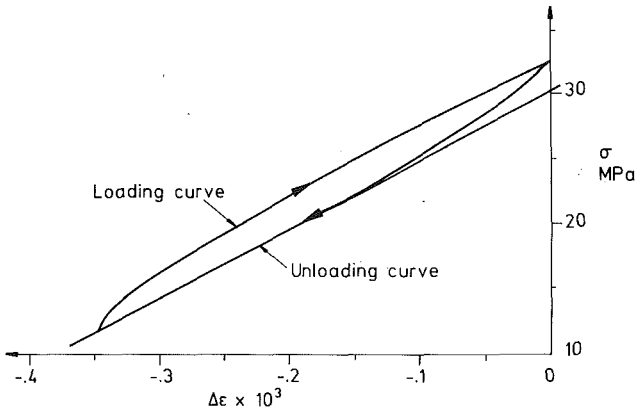


Fig. 4 Plot of stress σ against change in strain $\Delta\epsilon$ for a typical unloading/reloading test

Substituting equation (5) into equation (6) gives the relationship between creep strain and damage

$$\frac{\epsilon}{\dot{\epsilon}_{ss} t_f} = \left(\frac{1+\psi}{1+\psi-\phi}\right)[1-(1-\omega)^{1+\psi-\phi}]. \tag{7}$$

When $t = t_f$, equation (6) becomes

$$\frac{\epsilon_f}{\dot{\epsilon}_{ss} t_f} = \left(\frac{1+\psi}{1+\psi-\phi}\right)$$

where $\epsilon = \epsilon_f$ is the strain at failure. Rabotnov (1969) proposed that this equation can be used to obtain appropriate values of ψ and ϕ. Use of this equation, however, proved unsuitable for determining ϕ and ψ for the two batches of copper due to the scatter in the values of ϵ_f. For the present material a more consistent result was obtained by fitting the shape of the creep curve by adjusting the values of ϕ and ψ in equation (6). The best fit was obtained by use of $\phi = 6.3$ and $\psi = 7.1$. Equations (6) and (7) are plotted in Fig. 3 using these values of ϕ and ψ.

3 The Variation of Young's Modulus With Creep Damage

A relationship between creep strain and the damage parameter ω was obtained in the last section, equation (7), for conditions of constant stress. We make use of this relationship here to develop an experimental procedure for evaluating the way in which the modulus varies with damage.

3.1 Experimental Determination of Young's Modulus. A constant load creep testing machine was adapted to allow the load to be increased and decreased at a controlled rate. This was achieved by connecting a servometer to the jacking system located beneath the loading train. At intervals during the creep test the specimen was unloaded to approximately 40 percent of the full load at a rate of 0.6 MPa/s. and then reloaded at the same rate. During this loading cycle the load was measured using a load cell situated beneath the specimen and the change in length of the specimen was measured by a ridge located extensometer fitted with an LVDT.

A series of constant-load creep tests were performed on samples of copper from batch B at a temperature of 300°C. A stress of 32.4 MPa was selected for these tests to give a failure time of the order of 15 days. At regular intervals during the test the load was cycled as just described. A typical stress/strain loop obtained from a cycle is given in Fig. 4. Young's modulus was determined from the slope of the unloading curve as indicated in Fig. 4. This was found to give the most consistent values for the modulus. It was observed that if the specimen was held in the partially unloaded state, negative creep occurred. This type of behavior is consistent

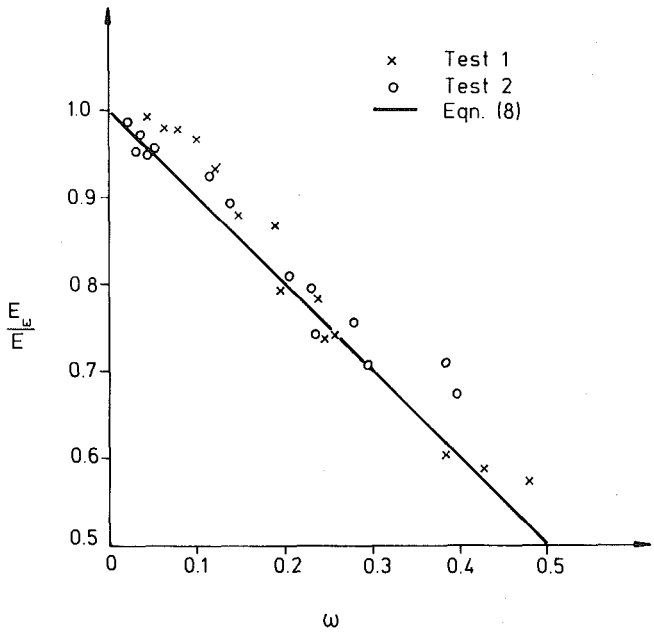


Fig. 5 The variation of Young's modulus with creep damage

with the observation of Davies et al. (1973) who have conducted controlled stress drop tests on a number of materials. Analysis of the negative creep rate suggests that the effect could result in a maximum decrease of approximately 8 percent in the measured value of Young's modulus. This effect was most pronounced just before failure of the specimen. Measurements of the modulus taken over a short period of time suggests that the method of measurement was accurate to ±8 percent.

The calculations of Young's modulus did not take into account any change in cross-sectional area of the specimen during the course of a constant-load creep test. This is consistent with our interpretation of equations (1) and (2) and the definition of damage as described in Section 2. In the tests used to obtain the variation of E with damage, the strain accumulated at $t/t_f = 0.95$ was typically 3 percent. At this time Young's modulus had reduced to 70 percent of its original value and, therefore, the major contribution to this reduction of modulus results from the growth of internal damage rather than from the reduction of cross-sectional area of the specimen.

After completion of a test $\dot{\epsilon}_{ss}$ and t_f were evaluated and the results were presented in the form of Fig. 3(*a*). These normalizations were chosen to minimize, as far as possible, the effects of any variation between the different test specimens. The instances during the test when the load was cycled were then identified on this plot and the value of damage ω, was determined from Fig. 3(*b*). The way in which Young's modulus, E_ω, varies with ω is shown in Fig. 5. The modulus has been normalized by the modulus of the damage-free material E measured at the beginning of the test at the test temperature of 300°C.

3.2 Correlation of Experimental Results. The material model of equations (1) and (2) is often simplified by assuming that $\phi = n$ and $\psi = \nu$. In this instance Odqvist and Hult (1961) assign a physical interpretation to ω. They assume that ω represents the effective cracked area fraction of material on any plane normal to the direction of loading and $\sigma/1-\omega$ is the mean stress in the uncracked material, so that the strain rate is evaluated by using this mean stress, equation (7), rather than the applied stress. Chaboche (1979) have extended this interpretation to the elastic response and propose that the elastic strain, ϵ^e, should be evaluated using the mean stress in the undamaged material and not the applied stress, i.e.

Table 1 Reference stresses for the two-bar structure of Fig. 6, evaluated for $n=5$, $\nu=4$, $\beta=1$, and $\lambda=2$ assuming (*a*) that Young's modulus is constant and (*b*) that the modulus is a function of creep damage.

	(*a*)	(*b*)
$\bar{\epsilon}$	σ_R	σ_R
1	1.05 σ_o	1.04 σ_o
3	1.02 σ_o	1.01 σ_o
10	1.02 σ_o	1.01 σ_o
30	1.01 σ_o	1.01 σ_o

$$\epsilon^e = \frac{\sigma}{(1-\omega)\mathrm{E}}.$$

Young's modulus for the damaged material is then given by

$$\mathrm{E}_\omega = (1-\omega)\mathrm{E} \tag{8}$$

This equation is plotted in Fig. 5 and compares favorably with the experimental results.

Micromechanical models of void growth during creep (Cocks and Ashby; 1982) result in a different form of creep constitutive law to that proposed by Kachanov (1958). The physical interpretation of ω offered by Odqvist and Hult (1961) as used by Chaboche (1979) is therefore not strictly valid. If we accept that Kachanov equations as an adequate description of the creep curve then ω should be interpreted as a parameter that provides a measure of the state of the material at a given instant in time which is determined from the shape of the creep curve. For a given creep response equation (8) provides a measure of the elastic response of the material. Since we have not assigned a physical interpretation to ω, it is not possible to directly compare the result of equation (8) with the predictions of Duva and Hutchinson (1984) and Budiansky and O'Connell (1976) for the elastic response of voided and cracked media.

4 Computed Failure Times for a Two-Bar Structure

In this section the time to failure of the simple structure of Fig. 6 is computed for a material that creeps and damages according to equations (1) and (2). The structure consists of two bars of lengths l and λl and cross-sectional areas A and βA, which are constrained to suffer the same extension under the action of a constant load P. Details of the calculations are presented in the Appendix, where, for simplicity, it is assumed that $\phi = n$ and $\psi = \nu$ in equations (1) and (2). In the computations it is convenient to adopt the following normalizations

$$\sigma_o = \frac{P}{A(1+\beta)}$$

and

$$t_f^o = \frac{1}{\dot{\omega}_o(1+\nu)}$$

for stress and time, respectively, where t_f^o is the time to failure is a uniaxial test at a constant stress σ_o, and perform the calculations for different values of the quantity

$$\bar{\epsilon} = \frac{\dot{\epsilon}_o t_f^o}{\epsilon_o^e}$$

where $\dot{\epsilon}_o t_f^o$ is the Monkman Grant constant (Monkman and Grant, 1956) and ϵ_o^e is the elastic strain at a stress σ_o

$$\epsilon_o^e = \sigma_o/\mathrm{E}.$$

Small values of $\bar{\epsilon}$ correspond with low strains to failure, where the elastic response becomes increasingly important in determining the overall response of the structure.

In the Appendix the effect of the variation of Young's modulus with creep damage is assessed by performing two sets of calculations. The modulus is either assumed to remain constant or vary according to equation (8). The results for these two situations are presented in Table 1 for a range of values of $\bar{\epsilon}$, in terms of the structural reference stress

$$\sigma_R = \alpha\sigma_o$$

which is the stress required in a uniaxial test to give the same time to failure as the structure.

The general trend for the two sets of results are similar, with the reference stress decreasing with increasing values of $\bar{\epsilon}$. For small values of $\bar{\epsilon}$ the assumption that Young's modulus remains constant gives the higher reference stresses, while for large $\bar{\epsilon}$ the two sets of calculations yield the same value of σ_R.

These results can best be explained by examining the bounding results of Ponter (1977) for the life of creeping structures. Ponter (1977) proposed a global damage measure

$$\Omega = \frac{1}{V}\int_V \frac{(1-\omega)^{\nu+1}}{(\nu+1)}dV \tag{9}$$

where V is the volume of the structure, which grows at a rate

$$\dot{\Omega} = -\frac{\dot{\omega}_o}{V}\int_V \left(\frac{\sigma}{\sigma_o}\right)^\nu dV \tag{10}$$

when the loading is uniaxial. Initially a structure responds elastically and the stresses redistribute with time due to creep and damage accumulation in the structure. This changing stress field must be taken into account in the evaluation of Ω, equation (10). Ponter (1977), however, identified an equilibrium stress field that results in the absolute minimum value of $\dot{\Omega}$. The appropriate stress field is the steady-state field for a creeping material with creep exponent equal to $(\nu-1)$. Assuming that the damage always grows at this minimum gives an upper bound to the life of a component, and thus a lower bound to the reference stress σ_R.

Ponter (1977) has also demonstrated that use of a limit-load stress distribution for a perfectly-plastic material in equation (10) also provides an upper bound to the time to failure. More recently Cocks and Ponter (1989) have shown that the rate of change of Ω remains virtually constant during the life of a component, and extrapolation of the initial rate of growth provides a good approximation of the time to failure. In this study the effects of elastic deformation were ignored and the stress field employed by Cocks and Ponter (1989) in equation (10) was the steady-state distribution.

The calculations which led to the results of Table 1 were conducted using $\nu=4$ and $n=5$. For large values of $\bar{\epsilon}$ the stresses have redistributed to the steady-state distribution before there is any significant growth of the damage, and the stresses are at or near the steady-state values for the bulk of the life. The time to failure is then close to that obtained by using the steady-state stress distribution in equation (10). For small values of $\bar{\epsilon}$, however, significant damage accumulation occurs as the peak elastic stresses relax and the steady-state stress distribution may never be achieved. In this instance the damage always accumulates faster than the minimum rate and a higher value of σ_R results. When the variation of Young's modulus with damage is included in the analysis the stresses are caused to relax faster in the initial highly stressed regions as the material becomes more compliant. This smoothing out of the stress field results in a lower global damage rate and a lower reference stress.

5 Conclusions

The variation of Young's modulus with a phenomenological measure of creep damage has been obtained for a batch of tough pitch copper. It was found that the relationship proposed by Chaboche (1979) (equation (8)) adequately describes the results.

The effect of including the variation of modulus with damage in the analysis of creeping structures was assessed by

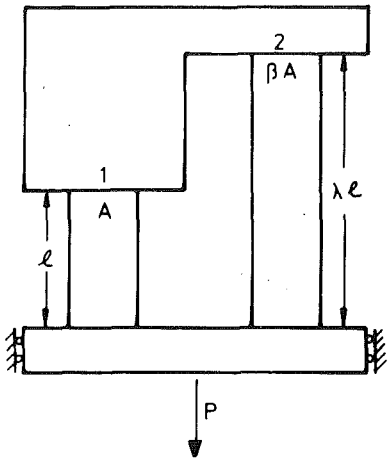


Fig. 6 Two-bar structure analyzed in Section 4. Bar 1 is a length of *l* and cross-sectional area A and bar 2 is of length λl and cross-sectional area βA

analyzing the simple two-bar structure of Fig. 6. It was found that its inclusion only affected the results for materials of low creep ductility. In these situations a conservative estimate of the life is obtained by assuming that the modulus remains constant. We would expect this result to apply, in general, to structures composed of materials which damage according to an effective stress criterion under multiaxial states of stress (Leckie and Hayhurst, 1974). It is not possible, however, to extend these results to materials which fail according to a maximum principal stress criterion, as the stress field that gives the minimum global damage rate (Ponter, 1977) can differ significantly from the steady-state creep stress distribution (Cocks and Ponter, 1989).

Monkman and Grant (1956) have analyzed the creep data for a range of engineering materials and find values of $\dot{\epsilon}_o t^o{}_f$ in the range 0.02 to .40. In practical situations the reference elastic strain $\epsilon_o{}^e$ is of the order of .001, giving values of $\bar{\epsilon}$ in the range 20 to 400. For these values of $\bar{\epsilon}$ inclusion of the variation of Young's modulus with creep damage in structural analysis has an insignificant effect on the predicted response of the structure.

Acknowledgments

The help of Mr. C. J. Morrison with the experimental work is gratefully acknowledged.

References

Budiansky, B., and O'Connell, R. J., 1976, "Elastic Moduli of a Cracked Solid," *Int. Jnl. Solids Struct.*, Vol. 12, p. 81.

Chaboche, J. L., 1979, "The Concept of Effective Stress Applied to Elasticity and to Plasticity in the Presence of Anisotropic Damage," ONERA Report 1979-77.

Cocks, A. C. F., and Ashby, M. F., 1982, "On Creep Fracture by Void Growth," *Prog. Mat. Sci.*, Vol. 27, p. 189.

Cocks, A. C. F., and Ponter, A. R. S., "Creep Deformation and Failure Under Cyclic Thermal Loading," to appear in *Nucl. Eng. Design*.

Davies, P. W., Nelmes, G., Williams, K. R., and Wilshire, B., 1973, "Stress Change Experiments During High-Temperature Creep of Copper, Iron, and Zinc," *Met. Sci.*, Vol. 7, p. 87.

Duva, J. M., and Hutchinson, J. W., 1984, "Constitutive Potentials for Dilutely Voided Non-Linear Materials," *Mech. Mater.*, Vol. 3, pp. 41–54.

Kachanov, L. M., 1958, "The Time to Failure Under Creep Conditions," *Izv. Akad. Nauk SSSR*, Otdel. Tekh. Nauk, No. 8, pp. 26–31.

Leckie, F. A., and Hayhurst, D. R., 1974, "Creep Rupture of Structures," *Proc. Roy. Soc.*, Vol. A340, p. 323.

Monkman, F. C., and Grant, N. J., 1956, "An Empirical Relationship Between Rupture Life and Minimum Creep-Rate in Creep-Rupture Tests," *Proc. ASTM.*, Vol. 56.

Odqvist, F. K. G., and Hult, J., 1961, "Some Aspects of Creep Rupture," *Arkiv for Fysik*, Vol. 19, p. 379.

Rabotnov, Yu, M., 1969, *Creep Problems in Structural Members*, (English Translation by F. A. Leckie), North-Holland, Amsterdam.

Ponter, A. R. S., 1977, "Upper Bounds on the Creep Rupture Life of Structures, Subjected to Variable Load and Temperature," *Int. J. Mech. Sci.*, Vol. 19, p. 79.

Tvergaard, V., 1986, "On the Stress State Dependence of Creep Rupture," *Acta Metallurgica*, Vol. 34, p. 243.

APPENDIX

Failure Analysis for a Two-Bar Structure

When analyzing the two-bar structure of Fig. 6, it is convenient to use the following normalized variables:

$$\Sigma = \frac{\sigma}{\sigma_o},\ S = \sigma/\sigma_o(1-\omega) \text{ and } \tau = \frac{t}{t_f^o}$$

where

$$\sigma_o = \frac{P}{A(1+\beta)} \text{ and } t_f^o = \frac{1}{\dot{\omega}_o(1+\nu)}.$$

Equations (1) and (2) for the creep response then become

$$\dot{\epsilon}_c = \frac{d\epsilon^c}{d\tau} = \dot{\epsilon}_o t_f^o S^n \tag{A1}$$

and

$$\dot{\omega} = \frac{d\omega}{d\tau} = \frac{S^v}{(1+\nu)} \tag{A2}$$

where the dots now indicate differentiation $w\bullet r\bullet t\bullet \tau$. The total strain rate at any instant is a combination of the elastic and creep responses

$$\dot{\epsilon} = \dot{\epsilon}^c + \dot{\epsilon}^e. \tag{A3}$$

The compatibility requirement for the two-bar structure is that the two bars must extend at the same rate, i.e.,

$$\dot{\epsilon}_1 = \lambda \dot{\epsilon}_2 \tag{A4}$$

where a subscript refers to the bar under consideration. The equilibrium condition can be expressed as

$$\Sigma_1 + \beta\Sigma_2 = 1. \tag{A5}$$

Here we analyze the situations where the modulus E remains constant and varies according to equation (8).

A.1 Analysis for Constant E. When the modulus remains constant the elastic response can be expressed as

$$\dot{\epsilon}^e = \epsilon_o^e \dot{\Sigma} \tag{A6}$$

where

$$\epsilon_o^e = \sigma_o/\mathrm{E}.$$

Combining equations (A4), (A3), (A1), and (A6) then gives

$$\lambda\dot{\Sigma}_2 = \dot{\Sigma}_1 = \bar{\epsilon}(S_1^n - \lambda S_2^n) \tag{A7}$$

where

$$\bar{\epsilon} = \frac{\dot{\epsilon}_o t_f^o}{\epsilon_o^e}.$$

$\dot{\Sigma}_1$ can be expressed in terms of $\dot{\Sigma}_2$ using the rate form of equation (A5). Equation (A7) then becomes

$$\dot{\Sigma}_2 = \bar{\epsilon}(S_1^n - \lambda S_2^n)/(\lambda+\beta). \tag{A8}$$

This equation must be integrated along with

$$\left.\begin{aligned} \dot{\omega}_1 &= \frac{S_1^v}{(1+\nu)} \\ \text{and} \qquad \dot{\omega}_2 &= \frac{S_2^v}{(1+\nu)} \end{aligned}\right\} \tag{A9}$$

to give the variation of stress and damage with time.

An Eulerian integration scheme was adopted in the present analysis with the variation of a quantity, such as ω_1, during a time increment $\Delta\tau$ being given by

$$\Delta\omega_1 = \dot{\omega}_1 \Delta\tau \tag{A10}$$

where the quantities $\dot{\omega}_1$, $\dot{\omega}_2$, and $\dot{\Sigma}_2$ are evaluated using the values of S_1 and S_2 at the start of the increment. A bar is taken to fail when $\omega = 0.9$, and the time to failure of the structure, τ_f, is evaluated by summing all the time increments $\Delta\tau$ up to the instant when the second bar fails. The normalized reference stress is then given by

$$\Sigma_R = \tau_f^{-\frac{1}{\nu}}. \tag{A11}$$

A.2 Analysis When E is a Function of ω. When E varies with ω according to eqn. (8), the elastic response becomes

$$\dot{\epsilon}_e = \epsilon_o^e \dot{S} \tag{A12}$$

Combining this with equations (A4), (A3), and (A4) gives

$$\lambda \dot{S}_2 - \dot{S}_1 = \bar{\epsilon}(S_1^n - \lambda S_2^n).$$

Substituting for $\dot{S}_1$ using the relationship of equation (A5) gives

$$\dot{S}_2 = \frac{(\bar{\epsilon}(S_1^n - \lambda S_2^n) + (\beta S_2^{\nu+1} + S_1^{\nu+1})/(1+\nu)(1-\omega_1))}{1+\beta(1-\omega_2)/(1-\omega_1)}.$$

This equation was integrated along with equations (A9) using the Eulerian scheme described in the last section. The reference stress was again evaluated using equation (A11).

Computations were performed using $n = 5$, $\nu = 4$, $\beta = 1$ and $\lambda = 2$ for a range of values of $\bar{\epsilon}$. The results of these calculations are summarized in Table 1.

An Anisotropic Hardening Rule for Elastoplastic Solids Based on Experimental Observations

Fernand Ellyin
Department of Mechanical Engineering,
University of Alberta,
Edmonton, Alberta,
Canada T6G 2G8

A hardening rule is described based on yield and memory surfaces. A memory surface indicates the extent of loading, and a yield surface is the locus of the elastic region. We define a hardening modulus curve which relates the change in size of the yield and memory surfaces to the tangent modulus of the material at the maximum load. The evolution of the yield surface is described for both the proportional and nonproportional loading paths. Both quasi-static and stable cyclic loading is considered. An attractive feature of this nonlinear hardening law is that the material constants associated with it are limited—three in all—and they can be easily determined from a simple test. The predictions of the proposed hardening law are compared with the experimental data for proportional and nonproportional loading paths, and are found to be in good agreement.

Introduction

Mechanical/structural systems in modern industrial plants are generally subject to complex loadings. Because of the functional requirements, design details, and manufacturing process, local plastic flow takes place in almost all of the important mechanical/structural components. Most design codes recognize the inevitability of such an occurrence and allow for the local stresses to exceed the elastic limit of the material under operational conditions (see, for example, ASME Boiler and Pressure Code (1983)). Furthermore, vital components, such as safety devices, have to be designed to withstand overload or prescribed accident/emergency conditions. Generally, elastic-plastic analysis is carried out to demonstrate compliance with the requirements. However, the predictions of this type of analysis are only as good as the model assumed for the material behavior.

The unprecedented advances in the computational methods and increases in the computer memory, now permit us to use more realistic hardening models, to some extent, indicative of the true material response. It is the objective of this paper to present a hardening model which is based on the observations made from the experimental data. An attractive feature of this hardening law is that the previously proposed classical hardening rules can be derived as particular cases. Furthermore, the parameters required to specify the proposed model can be easily obtained from a simple test.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

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 on the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, January 28, 1988; final revision, November 1, 1988.

Paper No. 89-WA/APM-4.

It is not within the scope of this paper to review the contributions made by numerous investigators on the yield surfaces. Review articles such as Paul (1968), Ikegami (1975), Hecker (1976), and Phillips (1986) could be consulted for a list of such contributions.

Background

The constitutive relations of solids in the multiaxial stress states are generally extensions of observations made from the

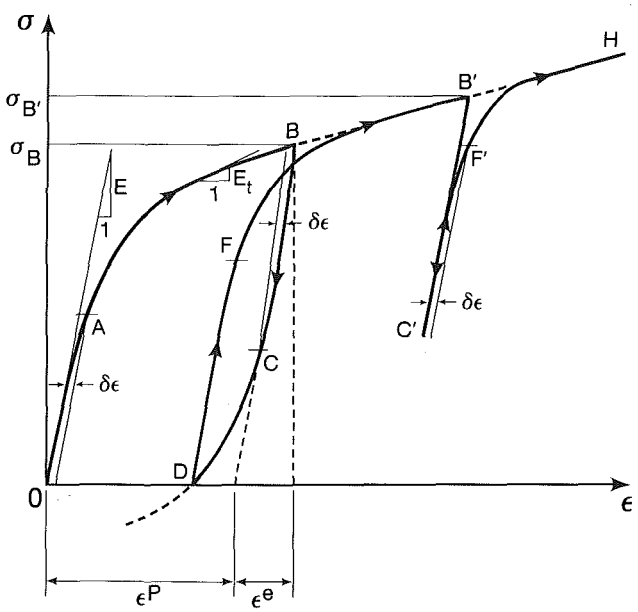


Fig. 1 Uniaxial stress-strain curve of an elastic-plastic material

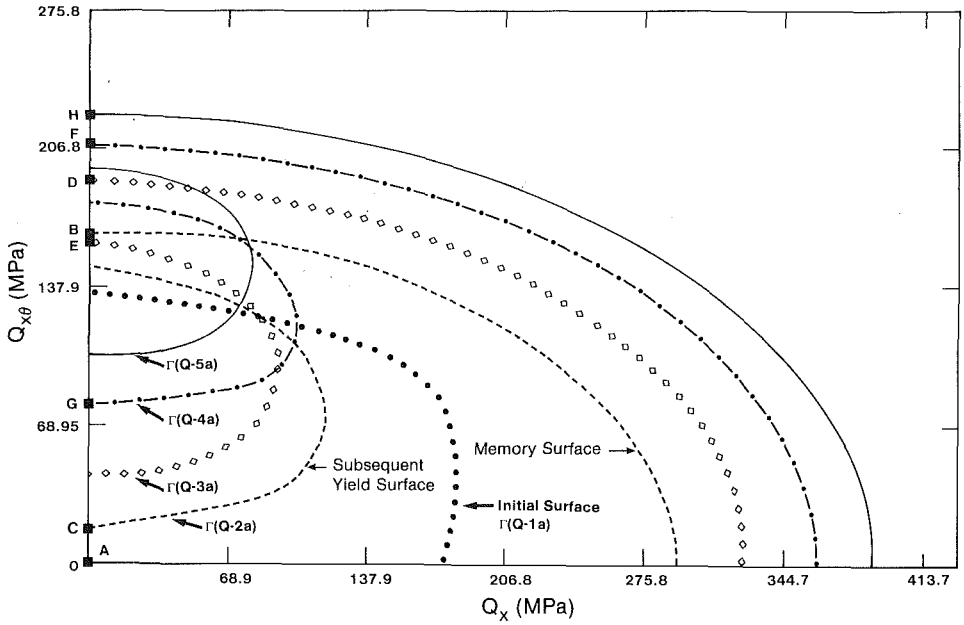


Fig. 2 Yield and memory surfaces in two-dimensional stress space for shear prestrain of commercially-pure titanium, specimen Q

uniaxial tension (or compression) tests. It may therefore be instructive to review briefly the material response to simple loading and unloading in the uniaxial stress state. Figure 1 shows a typical stress-strain curve of a strain-hardening material. Up to 0A the response is linear and the process is reversible. Upon further loading, permanent deformations set in as a result of the irreversible plastic flow. For example, when the loading is reduced at point B, the initial response is linear up to point C, and it is essentially parallel to 0A. When the unloading is further continued, an irreversible domain is entered (CD). Reloading from D we again observe a linear portion $DF \neq 2(0A)$ followed by a nonlinear response FB′. Two important observations are: (*a*) the curve following reloading does not pass through the unloading point B and (*b*) the linear portion BC following the unloading is not equal to the sum of the linear portions in tension and compression, i.e., $BC \neq 2(0A)$. The latter is generally known as the Bauschinger effect. Another observation to note is that the linear portion upon unloading and reloading, C′F′, depends on the maximum loading point B′.

To define the linear response regions (0A, BC, DF, etc.) experimentally, one has to load beyond it, and enter into the irreversible region. The imposed amount of plastic strain, $\delta\epsilon$, has to be larger than the zone of measuring accuracy and small enough not to cause a change in the orientation of the crystalline structure. In experiments performed by Ellyin and Grass (1975), this limit was set equal to 20×10^{-6} cm/cm (0.002 percent). The points A, F, C, etc., are then obtained from the back extrapolation of the tangent and the linear response (Phillips et al. (1972); Ellyin and Grass (1975)).

The generalization to the biaxial stress state is shown in Fig. 2. Two types of surfaces are shown in this figures. One which corresponds to point B in Fig. 1, hereinafter termed the *memory surface*; it indicates the extent of the maximum loading. The other surfaces corresponding to points A, C, F′, in Fig. 1, are called *yield surfaces*, which indicate the locus of the elastic response. The surface corresponding to point A is the *initial yield surface* whose shape in the stress space is generally elliptical.

Existing incremental theories of plasticity differ from one another in the choice of the hardening rule which specifies the change of the yield surfaces during plastic deformation. In an effort to better describe the experimentally observed evolution of the yield surfaces than the classical isotropic or kinematic hardening rules, "multi-surface" theories have been proposed. For example, in Mroz's model (1967) a family of concentric hypersurfaces were introduced in the stress space corresponding to constant hardening moduli. The evolution equation was derived by assuming that the initial and subsequent surfaces were translated along the loading path without intersecting one another.

The above model was subsequently simplified by Kreig (1975) and Dafalis and Popov (1975). In these models the family of constant moduli surfaces of Mroz were replaced by two surfaces: an outer bounding surface and an inner yield surface. The evolution equations then describe the motion of yield and bounding surface. For example, in Dafalis and Popov's (1975) model the motion of the yield and bounding surfaces is coupled, and is of a Ziegler (1959) type. A bonding or limiting surface simplifies the analysis, but its interpretation from a physical point of view can only be made at a limiting case.

Interpreting the experimental observations, Phillips and Lee (1979), and Ellyin et al. (1975, 1983) arrived at similar conclusions with some variation in details. It is to be noted that Phillips' experiments were in the *small* deformation range, while those of Ellyin are extended into the *finite deformation* domain. In Phillips and Lee's experiments on commercially-pure aluminum, the evolution of the yield surface was explained in terms of a loading surface expanding isotropically in the stress space and passing through the maximum loading point. The motion of the yield surface was deduced to be in the direction of the prestress, $d\sigma$, when the angle with the normal $\mathbf{n}$ to the yield surface is small. If the angle between the two is large, then the motion is between $d\sigma$ and $\mathbf{n}$ with $d\sigma$ being the predominant factor. The experimental investigation on commercially-pure titanium by Ellyin et al. will be described next.

Experimental Observations

As mentioned earlier, experimental determination of yield surfaces requires special care and instrumentation. Generally, thin-walled tubes are subjected to combined axial force and torsion (or internal pressure). A strict determination of the yield point (limit of the material's elastic response) at various

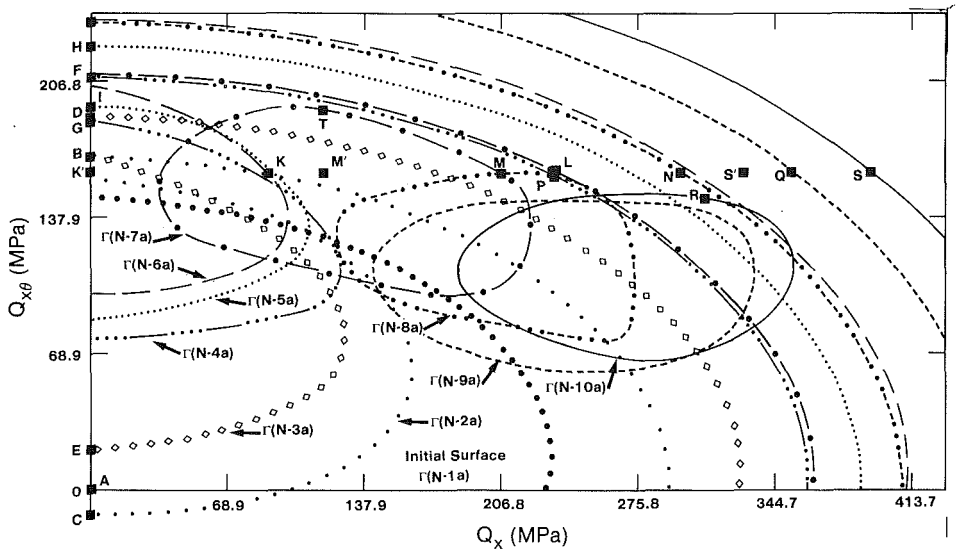


Fig. 3 Yield and memory surfaces for shear prestrain followed by an axial prestrain, specimen N

stress paths from a point inside the yield surface, is an essential prerequisite. The residual strain (accumulated plastic strain along various paths during determination of the yield surface) must be small enough to insure a *closed* yield surface corresponding to a given prestrain or prestress.

Figure 2 shows the initial and subsequent yield surfaces corresponding to various torsional prestrains. Each surface is identified by a set of letters and numbers, Γ (Q-ia) where Q identifies the specimen and i = 2-5 denotes the subsequent yield surfaces; Γ (Q-1a) being the initial yield surface. The sequence of loading for this test was ABCDEFGH. Two types of surfaces can be identified in this figure. A set of self-similar surfaces passing through the maximum loading point which are termed *memory surfaces.*[1] The memory surface at first coincides with the initial yield surface and expands uniformly and remains self-similar with the increased deformation; thus it obeys the isotropic hardening model. On the other hand, each subsequent yield surface (locus of the elastic domain) changes in shape and moves in the direction of loading, but it always remains inside the corresponding memory surface. In the case of quasi-static loading, the subsequent yield loci gradually contract and tend towards a limiting value. Therefore they do not, in general, follow the kinematic hardening rule. In Figs. 2, 3, and 5, the mean normal stress is identified by Q_x and the mean shear stress by $Q_{x\theta}$. In the case of thin-walled tubes, $Q_x \approx \sigma_x$ and $Q_{x\theta} \approx \tau_{xy}$.

To illustrate further the evolution of the yield surface, let us examine specimen "N" which is pretorqued initially, following a path similar to that of the specimen Q (path ABCDEFGHIJK′), Fig. 3. The last surface corresponding to the torsional path J is denoted by Γ(N-6a). Note that in both cases, specimen Q and N, there is a strong cross effect, i.e., decrease in the elastic limit in shear is accompanied by a similar decrease in tensile elastic limit. In Phillips' investigations (1986) no such cross effect was observed. The reason is most probably due to the small magnitude of the prestrain in Phillips et al.'s experiments. In Figs. 2 and 3, the proportional part of the strain path is up to plastic shearing strains: $(\gamma^p_{z\theta})_N \approx 8.5$ percent, and $(\gamma^p_{z\theta})_Q \approx 13$ percent.

When the specimen N is further subjected to a nonproportional strain path K′KMM′LNPQRSS′, the corresponding yield loci are denoted by Γ(N-7a) to Γ(N-10a). It is seen that the subsequent yield surfaces change in shape and rotate, but always remain inside the memory surface. However, once the loading exceeds the previous maximum effective stress, (point J) the subsequent yield loci Γ(N-9a) and Γ(N-10a) move almost parallel to the prestress path, QS.

It is to be noted that the experimental data in Fig. 2 and 3 include whatever effect shear deformation (large or small) has on the material behavior.[2]

Expressions for Memory and Yield Surfaces. As mentioned earlier, the memory surface obeys the isotropic hardening rule, and we use the stress as a parameter of loading history. Thus, the loading function is given by

$$\phi_m = f_m(\sigma_{ij}) - c(\sigma^*) = 0, \tag{1}$$

where $f_m(\sigma_{ij})$ is the yield function, and $c(\sigma^*)$ is the maximum value achieved by the function $f_m(\sigma_{ij})$ during the plastic deformation. In particular, in the case of the von Mises yield condition we get:

$$f_m(\sigma_{ij}) = J_2 \text{ and } c(\sigma^*) = (J_2)_{\max}, \tag{2}$$

where $J_2 = s_{ij}s_{ij}/2$, and $s_{ij} = \sigma_{ij} - \sigma_{kk}\delta_{ij}/3$, is the deviatoric stress.

For the subsequent yield surfaces, we use a combination of the kinematic and isotropic hardening models, i.e.,

$$\phi_y = f_y(\sigma_{ij} - \alpha_{ij}) - q = 0, \tag{3}$$

where α_{ij}'s are the coordinates of the centers of the subsequent yield surfaces, and q is a measure of the yield surface contraction.

In most two or multi-surface theories, the function q in (3) is taken to be proportional to the length of the plastic strain trajectory,

$$\bar{\epsilon}^p = \int \left(\frac{2}{3} d\epsilon^p_{ij} d\epsilon^p_{ij}\right)^{1/2}, \tag{4}$$

e.g., Mroz (1967), Dafalias and Popov (1975). In Tseng and Lee (1983) q is related to the plastic strain trajectory and plastic work, but no explicit relation is given. McDowell (1987a) advocates a constant value of q based on experimental data which is more appropriate for the stable cyclic loading. Note that a constant q indicates that the elastic regions does change with the imposed plastic strain history which generally may not be the case (e.g., see Figs. 2 and 3).

[1] The term memory surface was used by Ellyin (1983) and independently by Tseng and Lee (1983), as opposed to Phillips loading surface.

[2] See Drucker (1985) for a discussion of treatment of the continuum rotation associated with the shear deformation.

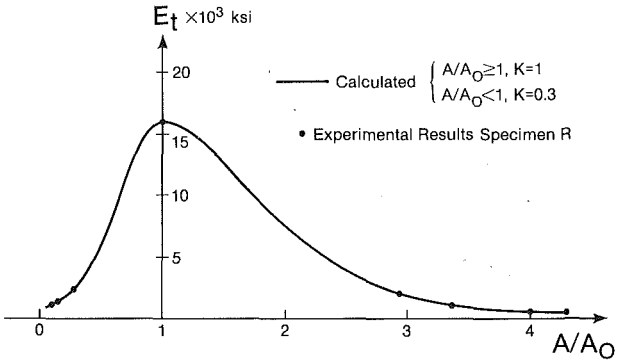


Fig. 4 Hardening modulus curve; tangent modulus versus ratio of the area of yield or memory surface to the initial one

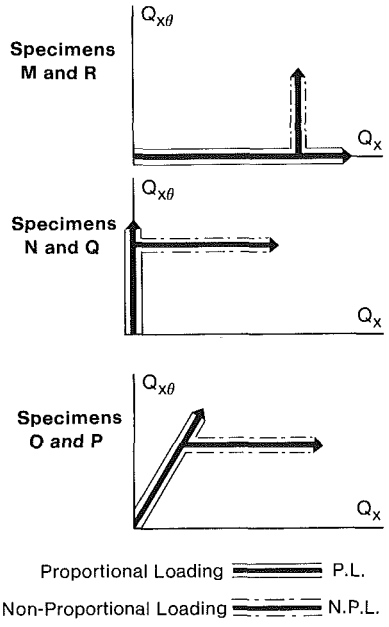


Fig. 5 Loading path for various specimens

An investigation of the experimental results reported by Ellyin et al. (1975, 1983) indicates that there exists a relationship between the tangent modulus, E_t, of the material and the ratio of the yield surface area (or memory surface) to that of the initial yield surface (see Fig. 4). This relationship will be termed the "hardening modulus curve."

Expression for the Hardening Modulus Curve

As discussed earlier, experimental results for a number of metallic materials indicate that the shape and size of subsequent yield surfaces do not generally remain constant with the increased plastic deformation. Therefore, it is suitable to use the area of the yield surface in the Π-plane to indicate its rate of change. In Fig. 4, A is the area of the memory surface or the subsequent yield surface: A_o is the area of the initial yield surface (loading from a virgin state), and E_o is the tangent modulus at the point of the initial yield. When $A/A_o > 1$, the curve represents the relationship between the memory surface and the tangent modulus at the maximum load, E_t. For $A/A_o < 1$, the curve depicts the relationship of the subsequent yield surface with the same tangent modulus. Generally speaking, the slope of the curve at the $A/A_o < 1$, is steeper than that at $A/A_o > 1$. Note that the tangent modulus is constant when moving on a given memory surface.

For the hardening modulus curve, the following relationship would apply (Ellyin and Wu (1987))

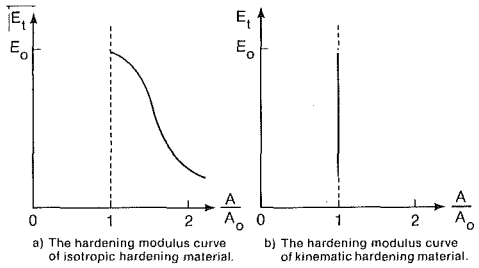


Fig. 6 Special cases of the hardening modulus curve

Table 1 Material Properties of Commercially-Pure Titanium 50A

Young's Modulus E_o (GPa)	Shear Modulus G (GPa)	Equation (18) K (MPa)	Equation (18) $1/n$	Equation (6) k
103.6 ±1.2	40.7 ±1.3	607 ±70	6.936 ±0.9	0.3 ±0.05

$$E_t(\omega, k) = E_o e^{-(\omega-1)^2/2k^2}, \tag{5}$$

where $\omega = A/A_o$ represents the area ratio of the memory or yield surface relative to the initial value, and k is a parameter indicating the extent of hardening. In general, $0 < k \leq 1$, and it controls the slope of the hardening modulus curve. Two different values for k can be chosen to represent the dissymmetry of the curve about the $A/A_o = 1$ axis. For a given material, k is constant and it can be determined through a suitable experiment involving loading and unloading (see Fig. 1).

The loading (strain) path program of the commercially-pure titanium is shown in Fig. 5 and the material properties are given in Table 1. The experimental results of specimen R (axial loading followed by partial unloading and subsequent torsional loading) are depicted in Fig. 4, as well as the curve predicted by equation (5). For this case, $k = 1$ for $A/A_o \geq 1$ and $k = 0.3$ for $A/A_o < 1$.

It would be useful to analyze further the hardening modulus curve. Equation (5) can be written as,

$$\omega = 1 \pm k\sqrt{\ln(E_o/E_t)^2}. \tag{6}$$

If the stress-strain relation of the material is known, then at any given stress, we can calculate E_t, and determine the RHS of (6). The positive sign in (6) is for the memory surfaces and the negative sign and, in general, is associated with the subsequent yield surfaces. For expanding yield surfaces one can use the positive sign with a different k value.

It is clear from equation (6) that the rate of change of the subsequent yield and memory surfaces depend on the material parameter k. For smaller values of k, there would be smaller dispersion of the hardening modulus curve, with respect to the vertical line, $\omega = A/A_o = 1$. In the limit when $k \to 0$, the hardening modulus curve collapse to line $A/A_o = 1$, and we obtain the kinematic hardening rule, Fig. 6(b). On the other hand, when the hardening modulus curve is symmetric with respect to line $\omega = 1$, we then recover the isotropic hardening rule, Fig. 6(a). It is, therefore, clear that the hardening modulus curve Fig. 4, (equation (6)) can be viewed as a general description of the material hardening process, and the isotropic and kinematic hardening laws are but two particular cases of it. Whether equation (6) applies to more complicated strain paths than those shown in Fig. 5, remains to be verified. One limiting case could be a class of materials whose stable cycle curve is not unique (see section on cyclic loading).

Evolution of Yield and Memory Surfaces

The hardening modulus curve described in the previous sec-

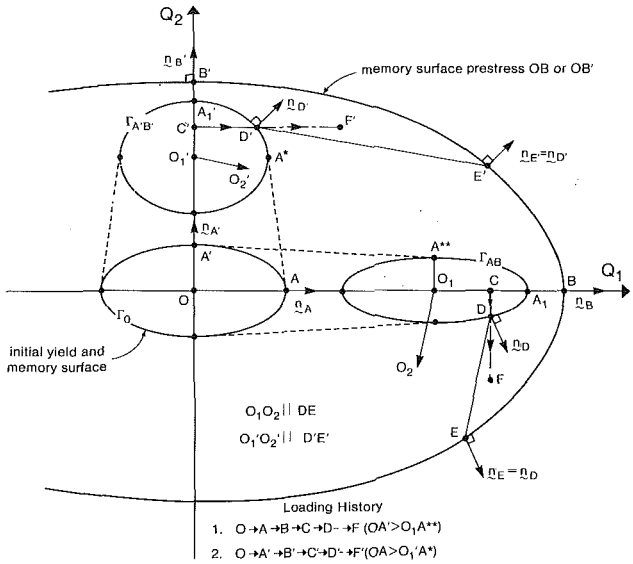


Fig. 7 Schematic representation of the motion of subsequent yield surfaces when remaining inside the memory surface

tion indicates the relationship between the size of the memory or yield surfaces and the tangent modulus at the maximum load, E_t, of the stress-strain curve. It does not, however, give the change of the position of the yield surface during the loading process.

It was indicated earlier that subsequent yield surfaces move along the prestressing path for the proportional loading, Fig. 2. For the nonproportional loading, the motion of the subsequent yield surface is more complex (see Fig. 3). The direction of its motion deviates from that of the loading path. Ellyin (1983), based on the analysis of the experimental results, deduced that the motion of the subsequent yield surface is as follows: The center of the yield surface moves parallel to the direction of the vector, connecting the starting point of the stress path (on the current yield locus) to a point on the corresponding memory surface whose exterior normal is parallel to the outward normal of the instantaneous yield surface at the start of loading. The above rule applies as long as the yield surface remains inside the memory surface. Figure 7 is a schematic representation of the motion of the subsequent yield surfaces with various prestressing paths (Q_1 and Q_2 are averaged principal stresses, $Q_1 \approx \sigma_1$ and $Q_2 \approx \sigma_2$ in a thin-walled section). For example, the motion of the subsequent yield surface Γ_{AB} when prestressed along DF is obtained by finding point E on the memory surface whose exterior normal $\mathbf{n}_E$, is parallel to the outward normal at D, $\mathbf{n}_D$. The center of the subsequent yield surface then will move in the direction parallel to DE and will remain inside the memory surface.

The evolution law thus described is, in a sense, a generalization of that of Phillips. That is, in the case of monotonic loading, the yield surface moves along the stress path and remains tangent to the memory surface. However, once there is a reversed loading, the motion is according to the rule described by Fig. 7, which is of the Mroz type. The necessity of describing two hardening rules is well described in the concluding part of Phillips (1986) review of his experimental work. This, of course, could not be done by plasticity theories with either a single surface or those with limit surfaces. To describe the aforementioned evolution equation we note, from Fig. 7, that when the loading path remains inside the memory surface, the motion of the center of the subsequent yield surface may be expressed as,

$$d\alpha_{ij} = d\mu[(\sigma_{ij})_m - (\sigma_{ij})_y], \tag{7}$$

where $(\sigma_{ij})_y$ is the stress point on the yield surface at the beginning of the loading increment, e.g., point D in Fig. 7 and $(\sigma_{ij})_m$ is the corresponding point on the memory surface, e.g., point E in Fig. 7.

For a monotonic or reloading path exceeding the current memory surface, the memory surface will expand and the yield surface will move in tandem with it, having the same tangent at the loading point. This condition is mathematically expressed as:

$$d\left[\frac{\partial f_y(\sigma_{ij}-\alpha_{ij})}{\partial \sigma_{ij}} - \mu \frac{\partial f_m(\sigma_{ij})}{\partial \sigma_{ij}}\right] = 0 \tag{8}$$

where μ is a proportionality factor.

Returning to Fig. 7 and relation (7) we note that the direction of outward normals at points D and E must be the same. Since the starting point of loading, i.e., $(\sigma_{ij})_D$ is known, our task therefore is to find the corresponding point $(\sigma_{ij})_E$ on the memory surface. The point on the memory surface is obtained from the solution of the following equations:

$$\frac{\partial f_m(\sigma_{ij})}{\partial \sigma_{ij}}\bigg|_{\sigma_{ij}=(\sigma_{ij})_m} = \lambda \frac{\partial f_y(\sigma_{ij}-\alpha_{ij})}{\partial \sigma_{ij}}\bigg|_{\sigma_{ij}=(\sigma_{ij})_y}, \tag{9}$$

and

$$f_m(\sigma_{ij})\,|_{\sigma_{ij}=(\sigma_{ij})_m} - c(\sigma^*) = 0. \tag{10}$$

Equation (9) specifies the condition that the gradient of the memory surface at E is proportional to that of the subsequent yield surface at D (Fig. 7). The coefficient λ in (9) is a proportionality constant. The supplementary condition (10) ensures that $(\sigma_{ij})_m$ is on the memory surface and thus satisfies equation (1). Note that in the aforementioned equations $(\sigma_{ij})_y$, $f_m(\sigma_{ij})$, $f_y(\sigma_{ij}-\alpha_{ij})$ are known, and thus we have ten equations in ten unknowns $(\sigma_{ij})_m$ and λ.

To determine the scalar $d\mu$ in (7), we use the consistency condition, i.e., when the stress point D remains on the yield surface during the loading process, then $d\phi_y = 0$. From (3),

$$d\phi_y = \frac{\partial f_y}{\partial \sigma_{ij}}(d\sigma_{ij} - d\alpha_{ij}) - dq = 0. \tag{11}$$

Substituting from (7) we get,

$$d\mu = \frac{\dfrac{\partial f_y}{\partial \sigma_{ij}} d\sigma_{ij} - dq}{[(\sigma_{kl})_m - (\sigma_{kl})_y]\partial f_y/\partial \sigma_{kl}}. \tag{12}$$

The description of the hardening rule when the stress path remains inside the memory surface, is now complete through specification of four relations (7), (9), (10), and (12).

Similarly, from (8) and (11), we get $d\alpha_{ij}$ and μ which completes the description of the hardening rule when loading exceeds the current memory surface. Note that in the previous formulation the function q is kept general, and a special form of it will be discussed in the following section.

Particular Case: Plane Stress Condition

Let us now consider the plane-stress case in more detail. The motivation for this stems from the fact that most experimental investigations are carried out on thin-walled tubular specimens loaded by a combination of tension-torsion. A state of plane stress then exists at the exterior surface of the tube. For the sake of illustration, we will assume that the material obeys von Mises' condition, then the memory surfaces (1) with condition (2) are given by,

$$\phi_m = (\sigma_x^2 + 3\tau_{xy}^2) - (\bar{\sigma}^2)_{\max} = 0 \tag{13}$$

where $\bar{\sigma}^2 = 3J_2$. The above relation indicates a family of ellipses similar to von Mises yield criterion.

The subsequent yield surfaces have a slightly distorted elliptical shape. We may represent them by an anisotropic yield

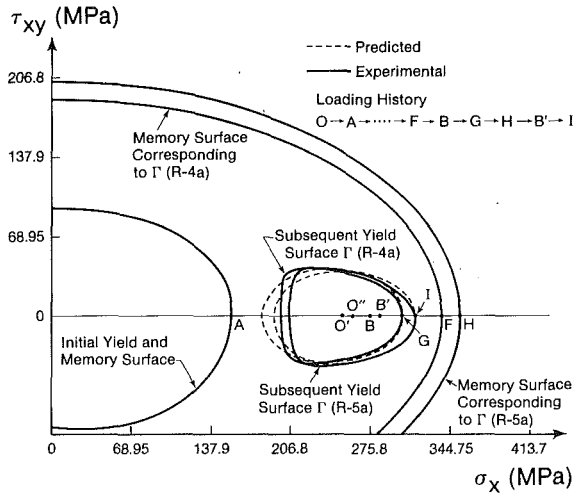


Fig. 8 Comparison between experimental and calculated yield surfaces, and the corresponding point on the memory surface for an axial (tensile) loading path, specimen R

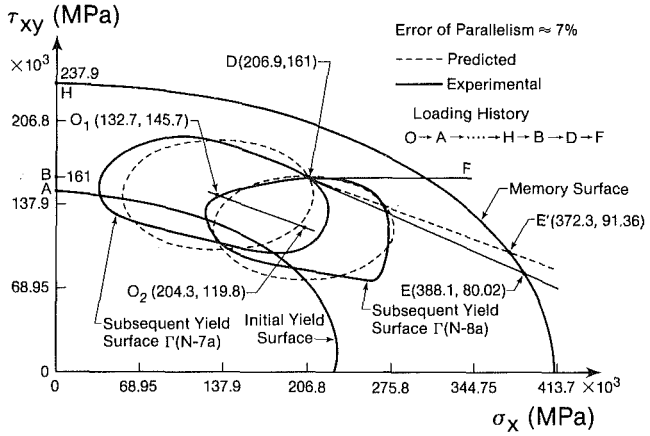


Fig. 9 Comparison between experimental and calculated yield surfaces, and the corresponding point on the memory surface for a torsional loading path followed by a nonproportional tensile path, specimen N

function, or use a number of other proposals, for example, Phillips and Weng (1975), Eisenberg and Yen (1984), and Budinansky (1984). Since the purpose of this section is to illustrate the proposed theory, for the sake of simplicity we will use von Mises-type yield criterion as a first approximation.

In this case, (3) will be represented by,

$$\phi_y = (\sigma_x - \sigma_x)^2 + 3(\tau_{xy} - \alpha_{xy})^2 - q = 0. \tag{14}$$

It can be shown that in the case of both the von Mises and Tresca yield conditions,

$$\omega = \frac{A(q)}{A_o(\bar{\sigma})} = \frac{q}{\sigma_o^2}, \tag{15}$$

and thus $q = q(E_t)$ can be found from the hardening modulus curve (6), i.e.,

$$q(E_t) = \sigma_o^2[1 - k\sqrt{\ln(E_o/E_t)^2}]. \tag{16}$$

Substituting from (16) into (14), the subsequent yield loci are given by,

$$\phi_y = (\sigma_x - \alpha_x)^2 + 3(\tau_{xy} - \alpha_{xy})^2 - \sigma_o^2[1 - k\sqrt{\ln(E_o/E_t)^2}] = 0. \tag{17}$$

The value of the tangent modulus, E_t, for a monotonically increasing load can be obtained, for example, from Ramberg-Osgood type relation,

$$\bar{\epsilon} = (\bar{\sigma}/E_o) + (\bar{\sigma}/K)^{1/n} \tag{18}$$

which results in:

$$E_t = E_o/[1 + (E_o/nK)(\bar{\sigma}/K)^{(n-1)/n}]. \tag{19}$$

In the previous equations n and K are parameters chosen to provide the best fit to the stress-strain curve of the material, E_o is the elastic modulus.

To find the corresponding point of the memory surface, we use conditions (9) and (10). Noting that the gradient on the instantaneous yield surface, ϕ_y, is taken at a definite point (D), substitution from (13) and (17) into (9) yields,

$$(\sigma_x)_m = \lambda(\sigma_x - \alpha_x)_y,$$
$$(\tau_{xy})_m = \lambda(\tau_{xy} - \alpha_{xy})_y. \tag{20}$$

The coordinates of the loading point, D, and the center of the current yield surface (see Fig. 7) are known and consequently, the magnitude of the terms in the parentheses on the RHS of (20) can be calculated. Substituting from (20) into (13), the value of the λ is then determined; thus, the coordinates of the

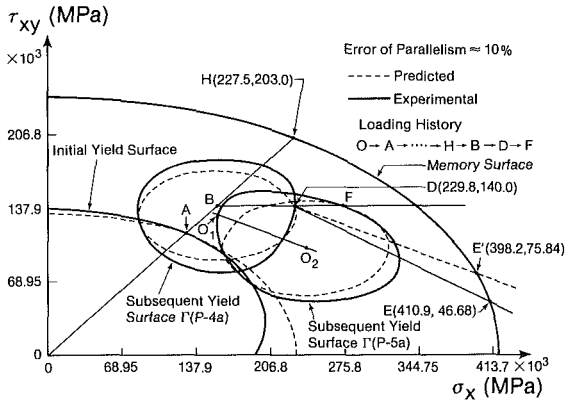


Fig. 10 Comparison between experimental and calculated yield surfaces, and the corresponding point on the memory surface for an oblique loading path followed by a nonproportional tensile path, specimen P

corresponding point on the memory surface are obtained from (20).

Comparison with Experimental Results

Let us now examine some of the experimental results for the paths shown in Fig. 5. The results for the axial prestrain path, (specimen R) are shown in Fig. 8. As to be expected, the direction of the motion is along the prestress and the predicted coordinates from equation (20) are exactly the same as points F and H in Fig. 8.

In the next two examples, the loading path is first proportional and then a nonproportional path is imposed. For example, in Fig. 9, the initial loading path is along the torsional axis (pretwist) followed by a nonproportional tensile prestrain (c.f., Fig. 3). The first subsequent yield surface in the nonproportional path is Γ(N-7a) with its center at 0_1. Now the loading path is from D to F. The experimental yield surface is Γ(N-8a) with its center at 0_2. $0_1 0_2$ is the actual direction of motion. If a line parallel to $0_1 0_2$ is drawn from D (loading point on yield surface Γ(N-7a)), it will intersect the memory surface at E. The point obtained from equations (20) is E′; the maximum deviation in parallelism in this case is about seven percent. Similarly, in the case of combined (oblique) proportional loading followed by a nonproportional pretension path, Fig. 10, the maximum deviation in the parallelism is about ten percent. We therefore observe the agreement between the observed motion and the predicted direction is fairly good, notwithstanding the von Mises shape approximation.

Plastic Flow Rule

In this section, we will derive an incremental law relating the increment of the strain components to those of the stress and stress increment for the case of small deformations. The usual assumption regarding the total strain increment decomposition will be made, i.e.,

$$d\epsilon_{ij} = d\epsilon^e_{ij} + d\epsilon^p_{ij}. \tag{21}$$

The increment of the elastic components of strain are related to the stress increment through the generalized Hooke's law,

$$d\epsilon^e_{ij} = \frac{1}{E}\left[(1+\nu)\delta_{ik}\delta_{jl} - \nu\delta_{ij}\delta_{kl}\right]d\sigma_{kl}. \tag{22}$$

For the increment of the plastic strain components, the consistency condition will impose that they be proportional to the quantity (Kachanov, 1971):

$$df_y = \frac{\partial f_y}{\partial \sigma_{kl}} d\sigma_{kl}, \tag{23}$$

where the increment df_y is carried out while the plastic strain is held constant. In addition, we can use von Mises flow rule which states that the increment of the plastic strain rate lies along the exterior normal of the yield surface, i.e.,

$$\begin{aligned} d\epsilon^p_{ij} &= g\frac{\partial f_y}{\partial \sigma_{ij}} df_y \quad \text{for} \frac{\partial f_y}{\partial \sigma_{kl}} d\sigma_{kl} \geq 0 \text{ and } \phi_y = 0, \\ d\epsilon^p_{ij} &= 0 \qquad\qquad \text{for} \frac{\partial f_y}{\partial \sigma_{kl}} d\sigma_{kl} < 0, \text{ or } \phi_y < 0, \end{aligned} \tag{24}$$

where the proportionality factor "g" is a function of the material hardening. The relations in equations (24) along with the yield surface (3) completely determine the plastic strain increments.

To determine the hardening function g, we proceed as follows: For metals, the plastic strain increment is not normally influenced by the hydrostatic pressure, thus:

$$d\epsilon^p_{ii} = 0, \quad \frac{\partial f_y}{\partial \sigma_{ij}} = \frac{\partial f_y}{\partial s_{ij}}. \tag{25}$$

Substituting from (25) into (24) we obtain,

$$d\epsilon^p_{ij} = g\frac{\partial f_y}{\partial s_{ij}} \frac{\partial f_y}{\partial s_{kl}} ds_{kl}. \tag{26}$$

By specializing (26) to the uniaxial tension case, and noting that

$$f_y(\sigma_{ij} - \alpha_{ij}) = q, \text{ and } d\bar{\sigma}/d\bar{\epsilon}^p = E^p_t, \text{ we get,}$$

$$g = \frac{1}{4qE^p_t} = \frac{1}{4}\left(\frac{1}{E_t} - \frac{1}{E_o}\right)q^{-1} \tag{27}$$

where the value of the tangent modulus E_t can be obtained from (19).

An alternative and more general expression can be found by multiplying (26) by itself, and using definition (4) which yields

$$g = d\bar{\epsilon}^p \Big/ \left[\frac{\partial f_y}{\partial \sigma_{ij}} \frac{\partial f_y}{\partial \sigma_{ij}} \left(\frac{\partial f_y}{\partial \sigma_{kl}} d\sigma_{kl}\right)^2\right]^{1/2}. \tag{28}$$

It is to be noted that the hardening function (27) or (28) contains the loading history through the change of q, and E_t is a function of the current stress level. Thus, g is a function of the loading history, and the current stress. A more restrictive assumption was made by Ziegler (1959) and retained by Dafalias and Popov (1975) in relating the stress and strain increments. It was assumed that the projection of $d\sigma_{ij}$ and of $d\alpha_{ij}$ on the normal are the same. Substituting from (27) or (28) into (26) the increment of the plastic strain components is completely determined.

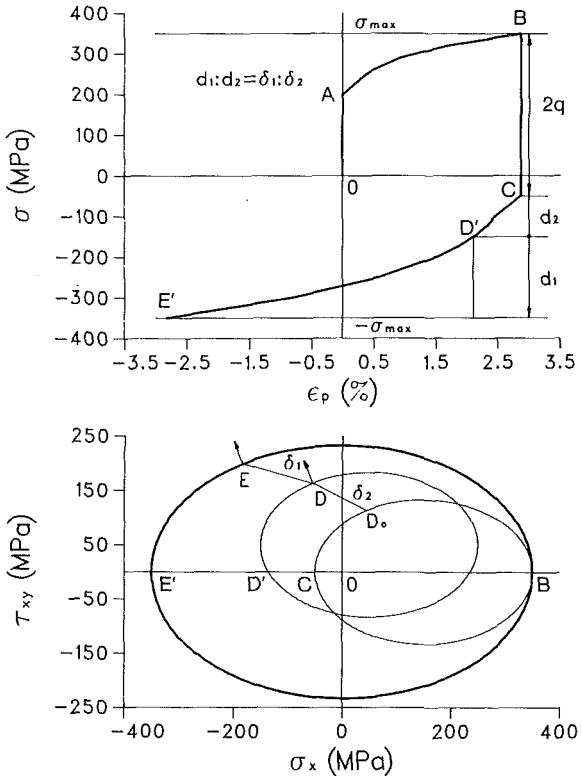


Fig. 11 Change of plastic modulus during reversed (cyclic) loading

If we employ (28) for the hardening function g, then function q in (3) could also be related to the length of the plastic strain trajectory, $\bar{\epsilon}^p$. However, using (27) and (16) simplifies calculation considerably.

Combining (26) and (22) we obtain,

$$d\epsilon_{ij} = \left\{\frac{1}{E}[(1+\nu)\delta_{ik}\delta_{jl} - \nu\delta_{ij}\delta_{kl}] + cg\frac{\partial f_y}{\partial \sigma_{ij}} \frac{\partial f_y}{\partial \sigma_{kl}}\right\}d\sigma_{kl} \tag{29}$$

where,

$$\begin{aligned} c &= 1 \text{ for } \frac{\partial f_y}{\partial \sigma_{kl}} d\sigma_{kl} \geq 0 \text{ and } \phi_y = 0, \\ c &= 0 \text{ for } \frac{\partial f_y}{\partial \sigma_{kl}} d\sigma_{kl} < 0, \text{ or } \phi_y < 0. \end{aligned} \tag{30}$$

It may be more convenient to express (29) in its inverse form, i.e.,

$$d\sigma_{ij} = L_{ijkl} d\epsilon_{kl} \tag{31}$$

where L_{ijkl} can be found after certain algebraic manipulations (Ellyin and Xia (1987)). By suitable definition of stress and strain rates, (31) can be generalized to the finite deformations (see Neale (1981)). However, this is beyond the scope of this paper and will not be discussed further.

Reversed and Stable Cyclic Loading

The hardening modulus "g" given by (27) requires calculation of the tangent modulus, E_t. As pointed out earlier for the monotonically increasing load, this can be calculated from (19). Also, when the stress path is inside the yield surface, the tangent modulus is E_o. Thus, for any stress path outside the yield surface and inside the memory surface, the tangent modulus must have a value between the two aforementioned values. For example, we note from Fig. 11 that for every point on the yield surface, there is a definite distance δ between it and the memory surface, e.g., D_oE. We seek a relationship between δ in the stress space and d of the uniaxial stress-strain

curve, (c.f., Fig. 11(*a*) and (*b*)). Referring to Fig. 11(*a*), an elastic unloading will take place from B to C. Continuing the reversed loading, the plastic flow will occur and the yield surface will move as shown in Fig. 11(*b*). We note that the ratio,

$$r=\frac{d_1}{d_2}=\frac{\delta_1}{\delta_2}, \tag{32}$$

uniquely determines the corresponding point on the reversed loading path. At C, $d_2=0$ and $r=\infty$, thus $E_t=E_o$ and at E', $d_1=0$, $r=0$ and $E_t=E_t^{(E')}$. Thus, r will be chosen to correlate the tangent modulus (and E_t for any loading path inside the memory surface). From (18) we get,

$$E_t=\frac{E_o}{1+\dfrac{E_o}{nK}\left[\dfrac{\sigma_{\max}+r\sigma_o}{(1+r)K}\right]^{(n-1)/n}}. \tag{33}$$

In the multiaxial stress state, δ_1 is the distance between the current stress point D and the corresponding stress point on the memory surface whose outward normal is parallel to that at point D. The distance δ_2 is measured from the point of the onset of the plastic flow D_o, to the point D (see Fig. 11(*b*)). The distance between two points is measured in the stress space and its projection in the deviatoric space, δ, is given by the Euclidian metric,

$$\delta=[(s_{ij}^{(2)}-s_{ij}^{(1)})(s_{ij}^{(2)}-s_{ij}^{(1)})]^{1/2}. \tag{34}$$

The rule of determining the tangent modulus has certain advantages to those proposed by Dafalias and Popov (1975), and Tseng and Lee (1983), i.e.,

$$E^p=E_o^p\left(1+h\frac{\delta}{\delta_{in}-\delta}\right) \tag{35}$$

where δ_{in} is the value of δ at the initiation of yield for each plastic loading. Since δ_{in} changes for plastic loading and unloading, then one may encounter a situation whereby $\delta>\delta_{in}$. In addition, the value of h has to be found through fitting by a number of material parameters. McDowell (1987) has discussed various proposals for the hardening modulus h, and concludes that one should normalize δ with a parameter which is a function of history of loading. The parameter r, as defined by equation (32), satisfies such a requirement. It is to be noted that the proposed theory is an incremental formulation and, as a result, no jump discontinuities are experienced during any elastic unloading process where $dq=0$.

Figure 12 is an example of the stable hysteresis loops for a proportional biaxial cyclic loading. In this test a circular cylindrical specimen was subjected to cyclic axial load and constant external and fluctuating internal pressure. (For the description of the experimental setup, see Ellyin and Valaire, 1982). The specimen was subjected to equi-biaxial strain ($\epsilon_a=\epsilon_t$) under deformation-controlled mode with zero mean strain. The material is ASTM A-516 Gr. 70 steel and all the pertinent mechanical and cyclic properties are given elsewhere (Ellyin (1984)). It is to be noted that the stable uniaxial cyclic curve is used to predict the biaxial stable response. The experimental hysteresis loops are very similar to those predicted by the theory and only the maximum and minimum values are indicated on the figure.

It may be pertinent to mention that the motion of yield surfaces for stable proportional loading is similar to that of the monotonic loading (see, for example, Ellyin and Neale (1979)). For the cyclic loading, the monotonic stress-strain curve has to be replaced by stable cyclic loading curve whereby the strain hardening or softening of the material is taken into account. In the case of nonproportional cyclic loading, some materials show additional hardening (e.g., see Lamba and Sidebottom (1978)). While the evolution of the yield surface

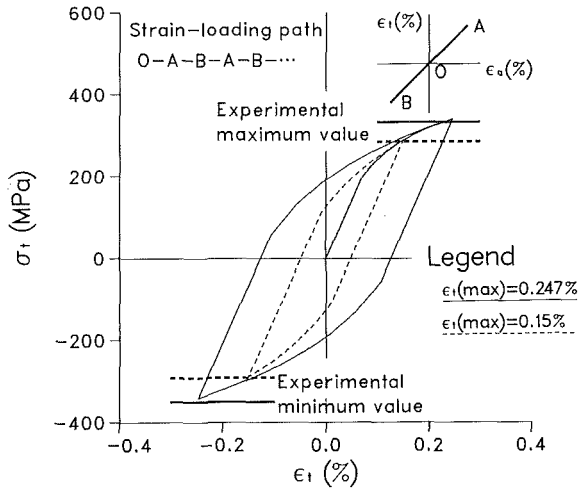


Fig. 12 Stable hysteresis loops for an equibiaxial proportional strain-controlled loading and comparison with experimental values

can be described by (7) and (8), the transient response cannot be predicted from the uniaxial cyclic curve alone. This type of behavior will require specification of additional material parameters (see, for example, Kremple (1984), McDowell (1985), and Ellyin and Xia (1988).

Conclusions

The elastoplastic response of solids is modeled based on two types of surfaces, memory and yield loci. A memory surface, ϕ_m, indicates the extent of the loading, and a yield surface, ϕ_y, portrays the locus of the elastic region. Based on the experimental observations, a hardening modulus curve is defined which establishes a relationship between the memory and yield surfaces in the Π-plane with the tangent modulus of the material at the maximum load. Thus, it becomes possible to predict change in size of the subsequent yield surface from the stress-strain relation of the material.

The evolution of the yield surfaces is predicted from the position of the loading on the yield surface. The motion of the center of the yield locus is parallel to the direction connecting the starting point of the stress path to the point on the memory surface whose exterior normal is parallel to the outward normal of the instantaneous yield surface at the start of plastic deformation (see Fig. 7) as long as the stress path remains inside the memory surface. This evolution is described mathematically and the hardening rule of the solid is given by three equations (7), (9), and (12). When the stress path moves outside the memory surface, then the motion of the center of the yield surface is given by (8) and (11).

An attractive feature of the proposed hardening rule is that it contains only three material constants which can be easily obtained from a simple test. The prediction of the proposed theory are compared with test results for various proportional and nonproportional loading paths, and the agreement is found to be fairly good (see Figs. 8–10).

The change of tangent modulus during reversed loading is described by relations (32)–(34) and an example of stable cyclic loading is given in Fig. 12.

Based on the proposed hardening rule, an explicit relationship is given relating the increment of the plastic strain to the stress and stress increments, equations (29)–(31).

Acknowledgment

The work reported herein is part of a general investigation into the material behavior under multiaxial states of stress and adverse environments. The research is supported, in part, by

the Natural Sciences and Engineering Research Council of Canada (NSERC Grant No. A-3808). We would also like to thank Z. Xia for his comments on the manuscript and his aid in numerical calculations. Also, appreciation is expressed to J. Wu for his work on the plane stress case. The author wishes to thank the University of Alberta for awarding him a McCalla Professorship (1986–1987) which enabled him to devote concentrated effort on his research work.

References

ASME Boiler and Pressure Vessel Code, 1983, Alternative Rules, American Society of Mechanical Engineers, New York, Section VIII, Div. 2.

Budinasky, B., 1984, "Anisotropic Plasticity of Plane-Isotropic Sheets," *Mechanics of Material Behavior*, Elsevier Science Publishers, Amsterdam.

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

Drucker, D. C., 1985, "Appropriate Simple Idealizations for Finite Plasticity," *Plasticity Today, Modelling, Methods and Applications*, A. Sawczuk and G. Bianchi eds., Elsevier, New York, pp. 47–54.

Eisenberg, M. A., and Yen, C. F., 1984, "The Anisotropic Deformation of Yield Surfaces," ASME *Journal of Engineering Materials and Technology*, Vol. 106, pp. 355–360.

Ellyin, F., and Grass, J. P., 1975, "Determination experimental de surface d'ecoulement plastique du titane 50A," Parts I & II, *Trans. Canadian Society for Mechanical Engineers*, CSME, Vol. 3, pp. 156–169.

Ellyin, F., and Neale, K. W., 1979, "Effect of Cyclic Loading on the Yield Surface," ASME *Journal of Pressure Vessel Technology*, Vol. 101, pp. 59–63.

Ellyin, F., and Valaire, B., 1982, "High Strain Biaxial Test Facility," *Proc. 1982 Joint Conf. on Experimental Mechanics*, Part I, SESA-JSME, Hawaii, pp. 136–143.

Ellyin, F., 1983, "On the Concept of Initial and Subsequent Yield Loci," *Failure Criteria of Structured Media*, J. P. Boehler ed., A. A. Balkema Publishers, Rotterdam, The Netherlands.

Ellyin, F., 1985, "Effect of Tensile-Mean-Strain on Plastic Strain Energy and Cyclic Response," ASME *Journal of Engineering Materials and Technology*, Vol. 107, pp. 119–125.

Ellyin, F., and Wu, J., 1987, "A Description of Elastoplastic Behavior Based on Memory and Yield Surface in Plane Stress State," *Constitutive Laws for Engineering Materials—Theory and Application*, C. S. Desai et al., eds., Elsevier, New York, pp. 531–538.

Ellyin, F., and Xia, Z., 1987, "Elastoplastic Stress-Strain Relation Based on a New Anisotropic Hardening Model," *Proc. IUTAM/ICM Symp. on Yielding, Damage and Failure of Anisotropic Solids*, Villard-de-Lans, France.

Ellyin, F., and Xia, Z., 1989, "A Rate-Independent Constitutive Model for Transient Non-Proportional Loading," *J. Mechanics and Physics of Solids*, Vol. 37, No. 1, pp. 71–91.

Hecker, S. S., 1976, "Experimental Studies of Yield Phenomena in Biaxially Loaded Materials," *Constitutive Equations in Viscoplasticity: Computational and Engineering Aspects*, ASME, New York, Vol. 20, pp. 1–32.

Ikegami, K., 1975, "A Historical Perspective of the Experimental Study of Subsequent Yield Surfaces for Metal. Parts 1 and 2, *Japan Soc. for Materials Science*, Vol. 24, pp. 491–505 and 709–719 (BISI Trans. 14420, Sept. 1976).

Kachanov, L. M., 1971, *Foundation of the Theory of Plasticity*, North-Holland Publishing Co., Amsterdam, The Netherlands.

Kremple, E., and Lu, H., 1984, "The Hardening and Rate-Dependent Behavior of Fully Annealed AISI Type 304 Stainless Steel Under Biaxial In-Phase and Out-of-Phase Strain Cycling at Room Temperature," ASME *Journal of Engineering Materials and Technology*, Vol. 106, pp. 376–382.

Krieg, R. D., 1975, "A Practical Two Surface Plasticity Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 641–646.

Lamba, H. S., and Sidebottom, O. M., 1978, "Cyclic Plasticity for Nonproportional Paths," Parts 1 and 2, ASME *Journal of Engineering Materials and Technology*, Vol. 100, pp. 96–111.

McDowell, D. L., 1985, "A Two Surface Model for Transient Nonproportional Cyclic Plasticity," Parts 1 and 2, ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 298–308.

McDowell, D. L., 1987a, "Simple Experimentally Motivated Cyclic Plasticity Model," *ASCE Journal of Engineering Mechanics*, Vol. 113, No. 3, pp. 378–397.

McDowell, D. L., 1987b, "An Evaluation of Recent Developments in Hardening and Flow Rules for Rate-Independent, Nonproportional Cyclic Plasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 65, pp. 323–334.

Mroz, Z., 1967, "On the Description of Anisotropic Work Hardening," *Journal of Mechanics and Physics of Solids*, Vol. 15, pp. 163–175.

Neale, K. W., 1981, "Phenomenological Constitutive Laws in Finite Plasticity," *SM Archives*, Vol. 6, pp. 79–128.

Paul, B., 1968, "Macroscopic Criteria for Plastic Flow and Brittle Fracture," *Fracture*, H. Liebowitz, ed., Vol. 2, Mathematical Foundation, Academic Press, New York, pp. 313–496.

Phillips, A., Liu, C. S., and Justusson, J. W., 1972, "An Experimental Investigation of Yield Surfaces at Elevated Temperature," *Acta Mechanica*, Vol. 14, pp. 119–146.

Phillips, A., and Weng, G. J., 1975, "An Analytical Study of an Experimentally Verified Hardening Law," ASME JOURNAL OF APPLIED MECHANICS, pp. 375–378.

Phillips, A., and Lee, C. W., 1978, "Yield Surfaces and Loading Surfaces, Experiments and Recommendations," *Int. Journal of Solids and Structures*, Vol. 15, pp. 715–729.

Phillips, A., 1986, "A Review of Quasistatic Experimental Plasticity and Viscoplasticity," *International Journal of Plasticity*, Vol. 2, pp. 315–328.

Tseng, N. T., and Lee, G. C., 1983, "Simple Plasticity Model of the Two Surface Type," *Journal of Engineering Mechanics*, ASCE, New York, Vol. 109, pp. 795–810.

Ziegler, H., 1959, "A Modification of Prager's Hardening Rule," *Quarterly of Applied Mathematics*, Vol. 17, pp. 55–65.

On the Inversion of Residual Stresses From Surface Displacements

Zhanjun Gao

T. Mura
Fellow, ASME

Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208

When plastic damage regions are accumulated in a material, there exist residual displacements on the surface of the material after all the loadings are removed. The residual displacements are defined as the difference between before and after loading, and can be measured experimentally without destruction of the material. This paper addresses the problem of evaluating the residual stress field caused by the accumulation of the plastic damage regions in a subdomain of the material. The problem is formulated as a system of integral equations relating the surface displacements to the unknown plastic strains. The damage domain, which appears as the domain of integration of the integral equations, is also unknown. Determination of the shape of the damage domain, together with the plastic strains, is a very complicated nonlinear problem. In addition to the residual surface displacement data, it requires more information about the loading history or other restrictive assumptions. However, the residual stress field in the vicinity of the damage domain is obtained after the equivalent damage domain and the equivalent plastic strains are introduced. The problem is an inverse problem, which is substantially different from the conventional forward analysis of structural mechanics. Special attention is given to the uniqueness and stability of the solution.

1 Introduction

The term "inverse" or "ill-posed" problems includes a huge variety of problems of different sorts and origins. Usually it implies identification of inputs from outputs, or determination of unknown causes from known consequences.

Inverse problems are becoming more and more important in many fields of science and engineering. Stanitz (1988) studied the problem of designing a channel for arbitrarily prescribed velocity distribution along the channel walls. Dulikravich (1988) presented a methodology to determine shapes, sizes, and location of coolant flow passage in an internally cooled turbine if temperature and heat fluxes are specified. The problem of determining the coordinates of an airfoil for a given surface pressure distribution was investigated by Sobieczky (1988). The recent progress of inverse problems related to structural mechanics has been reviewed by Kubo (1988). Other typical examples of inverse problems include reconstruction of a tomographic image from X-ray shadow pictures, inverse scattering for detecting defects of materials, determination of mass distribution of mechanical structures by their natural frequencies, estimation of the properties of the earth, identification of heat conductivity from boundary temperature distribution, etc.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLED MECHANICS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, October 13, 1988; final revision, February 6, 1989.

In our previous papers (Mura, Cox, and Gao (1986), Gao and Mura (1988a), and Gao and Mura (1988b)), we have considered the inverse problem of damage evaluation from the residual surface displacements. A numerical scheme has been developed to solve the corresponding ill-posed integral equations of the first kind. The problem is essentially a nonlinear problem since the shape of the damage domain is unknown to us. The concept of equivalent damage domain and equivalent plastic strains was introduced to avoid the nonlinearity of the problem. The equivalent damage domain can be chosen as any domain covering the real damage domain. The equivalent plastic strains are the fictitious plastic strains inside the chosen damage domain, which generate the measured residual surface displacements. This enables us to compute some characteristic quantities associated with the damage of the material. These quantities include lower bounds of the strain energy or any other quadratic functions of the plastic strains, etc. In this paper, we extend the previous results to the inversion of the residual stresses caused by the plastic damages. In the engineering problems of damage evaluation, we are very much interested in the residual stress field in the vicinity of the damage domain because it determines whether the damage domain is going to be stable or begin to propagate.

2 Fundamental Equations

Plastic strains ϵ^p_{kl} are accumulated in a subdomain Ω of the given body D (see Fig. 1) after a series of unknown loading. The subdomain Ω is called the damage domain, and its shape

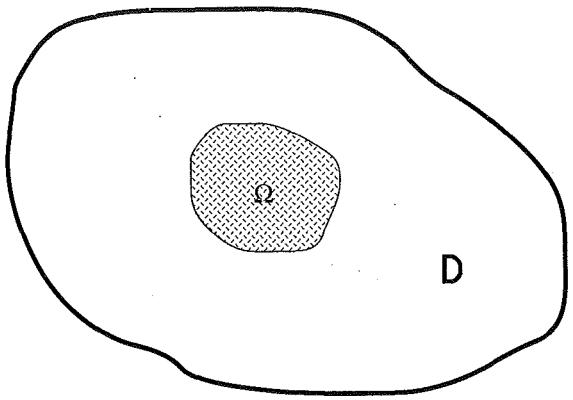


Fig. 1 The domain occupied by the material is denoted by D. Plastic strains are accumulated in Ω, a subdomain of D

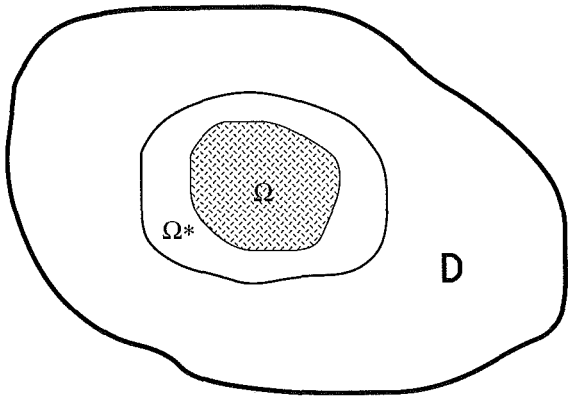


Fig. 2 An equivalent damage domain Ω^* is chosen to cover the actual damage domain Ω

and location are generally unknown. We use the residual surface displacement data to evaluate the residual stresses caused by the unknown plastic strains ϵ^p_{kl}. The residual surface displacements are relative and are defined as the difference between before and after loading.

The displacements and stresses due to ϵ^p_{kl} are denoted by u_i and σ_{ij}, respectively. The equations for equilibrium are

$$\sigma_{ij,j}=0, \text{ in } D. \tag{1}$$

The boundary conditions for free surface forces are

$$\sigma_{ij}\, n_j=0, \text{ on } \partial D, \tag{2}$$

where n_j is the exterior unit normal vector on ∂D, the boundary of D. The continuity conditions of tractions on $\partial\Omega$, the boundary of Ω, are

$$[\sigma_{ij}]n_j=0, \text{ on } \partial\Omega, \tag{3}$$

where [ ] = (out) − (in).

By using Hooke's law

$$\sigma_{ij}=C_{ijkl}(u_{k,l}-\epsilon^p_{kl}), \tag{4}$$

we rewrite (1), (2), and (3) as

$$C_{ijkl}\, u_{k,lj}=C_{ijkl}\, \epsilon^p_{kl,j} \text{ in } D, \tag{5}$$

$$C_{ijkl}\, (u_{k,l}-\epsilon^p_{kl})\, n_j=0, \text{ on } \partial D \tag{6}$$

and

$$[C_{ijkl}\, (u_{k,j}-\epsilon^p_{kl})]n_j=0, \text{ on } \partial\Omega, \tag{7}$$

respectively. C_{ijkl} in the previous equations are the elastic modulus tensor of the material.

We will find a relation between the residual surface displacements and the plastic strains.

Due to the symmetry of the elastic moduli C_{ijkl}, the Betti-Maxwell reciprocal relation holds,

$$C_{ijkl}\, G_{km,l}\, (\mathbf{x}-\mathbf{x}')(u_{i,j}(\mathbf{x})-\epsilon^p_{ij}(\mathbf{x})) = C_{ijkl}\, (u_{k,l}(\mathbf{x})-\epsilon^p_{kl}(\mathbf{x}))\, G_{im,j}(\mathbf{x}-\mathbf{x}'), \tag{8}$$

where $G_{km}\,(\mathbf{x}-\mathbf{x}')$ is Green's function for an infinite elastic medium and satisfies the equation of equilibrium for a point force with the unit magnitude,

$$C_{ijkl}\, G_{km,lj}\, (\mathbf{x}-\mathbf{x}')=-\delta_{im}\delta(\mathbf{x}-\mathbf{x}'). \tag{9}$$

δ_{im} is the Kronecker delta and $\delta(\mathbf{x}-\mathbf{x}')$ is Dirac's delta function.

When (8) is integrated in the domain D with respect to $\mathbf{x}$ and the integrations by parts are applied, we have

$$\begin{aligned}
&\int_{\partial D} C_{ijkl}\, G_{km,l}(\mathbf{x}-\mathbf{x}')\, u_i(\mathbf{x})n_j\, ds\ (\mathbf{x}) \\
&\quad-\int_D C_{ijkl}\, G_{km,lj}\ (\mathbf{x}-\mathbf{x}')u_i(\mathbf{x})d\mathbf{x} \\
&\quad-\int_D C_{ijkl}\, G_{km,l}\ (\mathbf{x}-\mathbf{x}')\ \epsilon^p_{ij}(\mathbf{x})d\mathbf{x} \\
&\quad\int_{\partial D} C_{ijkl}(u_{k,l}(\mathbf{x})-\epsilon^p_{kl}(\mathbf{x}))\, G_{im}(\mathbf{x}-\mathbf{x})n_j ds(\mathbf{x}) \\
&\quad-\int_D C_{ijkl}\, (u_{k,lj}(\mathbf{x})-\epsilon^p_{kl,\,j}\ (\mathbf{x}))\, G_{im}(\mathbf{x}-\mathbf{x}')\, d\mathbf{x}
\end{aligned} \tag{10}$$

where ds is the surface element of ∂D and $d\mathbf{x}$ is the volume element of D. The boundary integrals on $\partial\Omega$ disappear due to the continuity of u_i and $\sigma_{ij}n_j$. Equation (9) is substituted into the second integral in the left-hand side (10). The right-hand side in (10) is zero due to (5) and (6). Finally, (10) becomes

$$\begin{aligned}
&\int_\Omega C_{ijkl}\, G_{km,\,l}\ (\mathbf{x}-\mathbf{x}')\ \epsilon^p_{ij}(\mathbf{x})d\mathbf{x}= \\
&\qquad\int_{\partial D} C_{ijkl}\, G_{km,\,l}\ (\mathbf{x}-\mathbf{x}')u_i(\mathbf{x})n_j ds(\mathbf{x})+\beta\ u_m(\mathbf{x}')
\end{aligned} \tag{11}$$

since $\epsilon^p_{ij}=0$ in $D-\Omega$, where $\beta=1$ for $\mathbf{x}'\in D$ and $\beta=1/2$ for $\mathbf{x}'\in\partial D$ (see Kinoshita and Mura (1956)).

When $\mathbf{x}'$ is considered on ∂D, the equation (11) is the Fredholm integral equation of the first kind for unknown ϵ^p_{ij} under given $u_i(\mathbf{x}')$ on ∂D. Since plastic strains are incompressible, the condition

$$\epsilon^p{}_{kk}=0 \tag{12}$$

is imposed to (11).

3 Discussion on Uniqueness

When the surface displacements are measured experimentally, the left-hand side in (11) is known. The equations (11) are, therefore, equations for unknown plastic strains ϵ^p_{ij}. The solution is, however, not unique. This nonuniqueness is easily shown by the existence of the impotent eigenstrains (see Furuhashi and Mura (1979)). If compatible plastic strains are defined in Ω and the corresponding displacements vanish on $\partial\Omega$, these plastic strains are impotent eigenstrains which do not cause any elastic field in D. Therefore, the addition of these plastic strains to ϵ^p_{ij} provides the same surface displacements on ∂D.

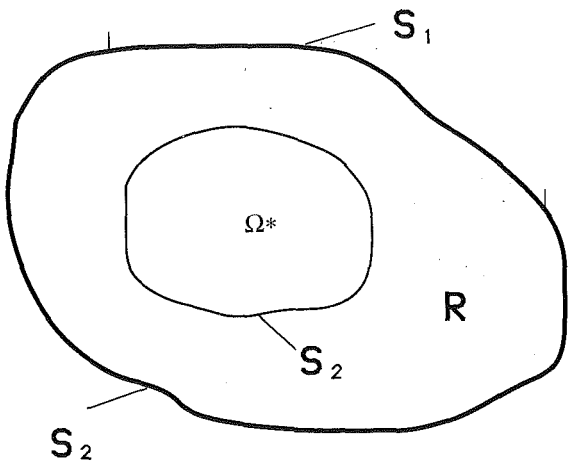


Fig. 3 Displacements and surface tractions are zero on S_1, a part of the boundary of R

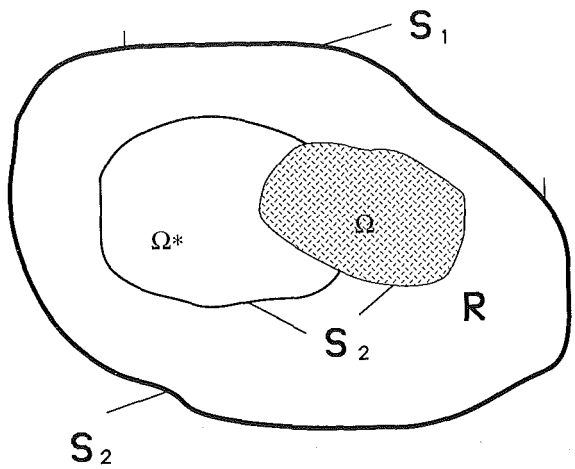


Fig. 4 The equivalent damage domain Ω^* are partially intersected with the actual damage domain Ω

Although determination of ϵ_{ij}^p is not unique, it can be shown that the stress field caused by ϵ_{ij}^p is uniquely determined in some parts of the material from the surface displacement data.

Choose a domain Ω^* such that Ω^* covers Ω (see Fig. 2). Find proper plastic strains $\epsilon_{ij}^p(\Omega^*)$ in Ω^* which provide the same surface displacements as those caused by $\epsilon_{ij}^p(\Omega)$ in Ω. Then, the displacement and stress fields caused by $\epsilon_{ij}^p(\Omega)$ and those caused by $\epsilon_{ij}^p(\Omega^*)$ are identical in the domain $D-\Omega^*$.

This identity is based upon the following lemma.

Lemma: Let S_1 be a part of boundary of an elastic body R (see Fig. 3). If surface displacements and force tractions are zero on S_1, the displacements and stresses in R are identically zero.

The difference between the elastic field caused by $\epsilon_{ij}^p(\Omega)$ in Ω and that caused by $\epsilon_{ij}^p(\Omega)$ in Ω^* is an example of this lemma, where $R = D-\Omega^*$ and $S_1 = \partial D$.

In the proposed inverse problem, the unknowns are not only ϵ_{ij}^p but also the location and size of Ω where ϵ_{ij}^p are caused. However, the identity of the stress field mentioned previously leads to a new idea to evaluate the residual stresses in $D-\Omega^*$ uniquely if Ω^* can be chosen properly.

The lemma has been proved by Gao and Mura (1988b) for the two-dimensional case. The lemma may be proved for the three-dimensional case as follows.

Let u_i and σ_{ij} be the displacements and the stresses which give the zero displacements and the zero force tractions on S_1. Consider another set of displacements u_i' and stresses σ_{ij}' which give the zero displacements on S_2 but arbitrary force tractions on S_1. Assume that both of σ_{ij} and σ_{ij}' are in equilibrium. Then, the Betti-Maxwell reciprocal relation yields

$$\int_{S_1+S_2} \sigma_{ij}' u_i n_j ds = \int_{S_1+S_2} \sigma_{ij} u_i' n_j ds \tag{13}$$

where n_j is the outward normal vector on the boundaries S_1 and S_2 and $S_1+S_2=\partial R$. It is obvious that the right-hand side in (13) is zero from the conditions of the problem. Therefore, (13) leads to

$$\int_{S_2} \sigma_{ij}' u_j n_j ds = 0. \tag{14}$$

The integral on S_1 vanishes since $u_i = 0$ on S_1. By choosing proper $\sigma_{ij}' n_j$ on S_1, we can assign any value of $\sigma_{ij}' n_j$ on S_2. It means that $\sigma_{ij}' n_j$ on S_2 can be chosen arbitrarily. Then, it is concluded that $u_i = 0$ on S_2. Consequently, u_i is zero everywhere in R.

This lemma can be extended to the case when Ω^* and Ω are partially intersected as shown in Fig. 4. In this case R is the domain bounded by ∂D and $\partial(\Omega+\Omega^*)$. The lemma can be further extended to the case of steady-state elastic wave, where the Green's function $g_{km}(\mathbf{x}-\mathbf{x}')$ is defined as (see Mura (1982)):

$$C_{ijkl} g_{km,lj}(\mathbf{x}-\mathbf{x}') + \rho\omega^2 g_{im}(\mathbf{x}-\mathbf{x}') + \delta_{im}\delta(\mathbf{x}-\mathbf{x}') = 0 \tag{15}$$

where ρ is the density and ω is the frequency.

Accordingly, the uniqueness theory for the residual stresses may be properly modified.

Any domain covering the damage domain Ω can be chosen as an equivalent damage domain Ω^*. However, we want to choose Ω^* as close to Ω as possible so that the stresses near the boundary of Ω^* provide a good representation of the stresses in the vicinity of the actual damage domain Ω. The residual stresses in the vicinity of the damage domain are very important quantities in the engineering problems of damage evaluation. They determine whether the damage domain is going to be stable or begin to propagate.

4 The Variational Problem

We have proposed to solve the integral equation

$$\int_{\Omega^*} C_{ijkl} G_{km,l}(\mathbf{x}-\mathbf{x}') \epsilon_{ij}^p(\mathbf{x}) d\mathbf{x} = \int_{\partial D} C_{ijkl} G_{km,l}(\mathbf{x}-\mathbf{x}') u_i(\mathbf{x}) n_j ds(\mathbf{x}) + \frac{1}{2} u_m(\mathbf{x}') \tag{16}$$

for $\mathbf{x}' \in \partial D$, where Ω^* is chosen properly and $\epsilon_{kk}^p = 0$. The equation (16) is written as follows for the two-dimensional case:

$$\int_{\Omega^*} \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{V}(\mathbf{x}) d\mathbf{x} = \mathbf{U}(\mathbf{x}'),\ \mathbf{x}' \in \partial D \tag{17}$$

where

$$\mathbf{K}(\mathbf{x},\mathbf{x}') = \begin{Bmatrix} C_{ij12} G_{i1,j} & , & C_{ij22} G_{i1,j} - C_{ij11} G_{i1,j} \\ C_{ij12} G_{i2,j} & , & C_{ij22} G_{i2,j} - C_{ij11} G_{i2,j} \end{Bmatrix} \tag{18}$$

$$\mathbf{V}(\mathbf{x}) = [\gamma_{12}^p,\ \epsilon_{22}^p]^T,\ \gamma_{12}^p = 2\epsilon_{12}^p$$

$$\mathbf{U}(\mathbf{x}') = \begin{Bmatrix} -\int_{\partial D} C_{ijkl}\, G_{k1,l}(\mathbf{x}-\mathbf{x}')u_i(\mathbf{x})n_j(\mathbf{x})ds(\mathbf{x}) + 0.5u_1(\mathbf{x})' \\ -\int_{\partial D} C_{ijkl}\, G_{k2,l}(\mathbf{x}-\mathbf{x}')u_i(\mathbf{x})n_j(\mathbf{x})ds(\mathbf{x}) + 0.5u_2(\mathbf{x}') \end{Bmatrix} \tag{19}$$

$$G_{ij}(\mathbf{x}-\mathbf{x}') = \frac{1}{8\pi\mu(1-\nu)}\left\{-(3-4\nu)\delta_{ij}\,\mathrm{Log}(R) + \frac{(x_i-x'_i)(x_j-x'_j)}{R^2}\right\},$$

$$R^2 = (x_1-x'_1)^2 + (x_2-x'_2)^2.$$

Further simplication leads to

$$\mathbf{K}(\mathbf{x},\mathbf{x}') = \frac{1}{4\pi(1-\nu)R^4} \times \begin{Bmatrix} \bar{x}_2[(2\nu-3)\bar{x}_1^2 + (2\nu-1)\bar{x}_2^2],\ x_1[4(1-\nu)\bar{x}_1^2 - 4\nu\bar{x}_2^2] \\ \bar{x}_1[(2\nu-3)\bar{x}_2^2 + (2\nu-1)\bar{x}_1^2],\ \bar{x}_2[4\nu\bar{x}_1^2 - 4(1-\nu)\bar{x}_2^2] \end{Bmatrix} \tag{20}$$

where

$$\bar{x}_1 = x_1 - x'_1 \quad \bar{x}_2 = x_2 = x'_2.$$

Of all the $\mathbf{V}(\mathbf{x})$ satisfying (17), we look for the one which has the smallest L_2 norm, i.e., we pose a new minimal problem

$$\begin{cases} \text{Min} & ||\mathbf{V}(\mathbf{x})||^2 \\ \text{subject to} & ||\int_{\Omega^*} \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{V}(\mathbf{x})d\mathbf{x} - \mathbf{U}(\mathbf{x}')||^2 = 0, \end{cases} \tag{21}$$

where $||\bullet||^2$ is the square of the L_2 norm, i.e., the inner product of the function with itself on the domain it is defined,

$$||\mathbf{V}(\mathbf{x})||^2 = \int_{\Omega^*} \mathbf{V}^T(\mathbf{x})\mathbf{V}(\mathbf{x})d\mathbf{x}$$

and

$$||\int_{\Omega^*} \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{V}(\mathbf{x})d\mathbf{x} - \mathbf{U}(\mathbf{x}')||^2 = \int_{\partial D} \left\{\left[\int_{\Omega^*} \mathbf{K}(\mathbf{x},\mathbf{x}')V(x)d\mathbf{x} - \mathbf{U}(\mathbf{x}')\right]^T \times \left[\int_{\Omega^*} \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{V}(\mathbf{x})d\mathbf{x} - \mathbf{U}(\mathbf{x}')\right]\right\} d\mathbf{x}',$$

where the superscripts "T" indicate the transposes of the matrices.

For an engineering problem, our experimental measurements always contain error. $\mathbf{U}(\mathbf{x}')$ is only an approximation of the exact displacements $\mathbf{U}^0(\mathbf{x}')$ such that

$$||\mathbf{U}(\mathbf{x}') - \mathbf{U}^0(\mathbf{x}')||^2 \le \epsilon,$$

where ϵ is a small, positive number. Therefore, the problem (21) is changed into

$$\begin{cases} \text{Min} & ||\mathbf{V}(\mathbf{x})||^2 \\ \text{subject to} & ||\int_{\Omega^*} \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{V}(\mathbf{x})d\mathbf{x} - \mathbf{U}(\mathbf{x})||^2 = \epsilon, \end{cases} \tag{22}$$

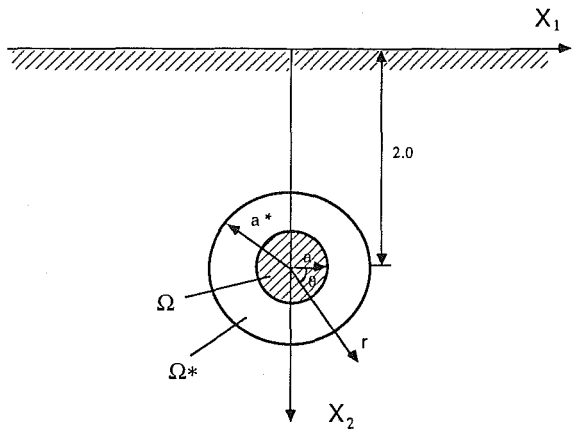


Fig. 5 Unit plastic strains (see (28)), are distributed in the damage domain Ω. The equivalent damage domain Ω* is chosen to recover the residual stress field outside Ω*

The problem (22) has the unique solution and the solution is stable to the small perturbation of the input data $\mathbf{U}(\mathbf{x}')$. The detailed discussion can be found in Gao and Mura (1988a).

The problem (22) is equivalent to its Euler equations

$$\int_{\Omega^*} \mathbf{K}^*(\mathbf{x},\mathbf{s})\mathbf{V}(\mathbf{x})d\mathbf{x} + \alpha\mathbf{V}(\mathbf{s}) = \mathbf{U}^*(\mathbf{s})$$

$$\mathbf{s} \in \Omega^* \tag{23}$$

and

$$f(\alpha) = 0, \tag{24}$$

where

$$\mathbf{K}^*(\mathbf{x},\mathbf{s}) = \int_{\partial D} \mathbf{K}^T(\mathbf{s},\mathbf{x}')\mathbf{K}(\mathbf{x},\mathbf{x}')d\mathbf{x}'$$

$$\mathbf{s} \in \Omega^*, \tag{25}$$

$$\mathbf{U}^*(\mathbf{s}) = \int_D K^T(\mathbf{s},\mathbf{x}')\mathbf{U}(\mathbf{x}')d\mathbf{x}'$$

$$\mathbf{s} \in \Omega^* \tag{26}$$

and

$$f(\alpha) = ||\int_{\Omega^*} \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{V}^{(\alpha)}(\mathbf{x})d\mathbf{x} - \mathbf{U}(\mathbf{x}')||^2 - \epsilon. \tag{27}$$

$\mathbf{V}^{(\alpha)}(\mathbf{x})$ in the equation (27) is the solution of (23).

5 An Example

Numerical results are given for the problem shown in Fig. 5. Domain D is the half space $x_2 \ge 0$ and ∂D is the free surface $x_2 = 0$. After a series of unknown loading, plastic strains are accumulated in domain

$$\Omega = \{\mathbf{x} | r \le 0.2\}, \tag{28}$$

where

$$r = \sqrt{x_1^2 + (x_2 - 2.0)^2}.$$

The rest of the material remains elastic. The distribution of the plastic strains in Ω is

$$\begin{cases} \epsilon_{22}^p = -\epsilon_{11}^p = 1, \\ \gamma_{12}^p = 1. \end{cases} \tag{29}$$

Our goal is to solve (23) and (24) for ϵ_{kl}^p by using the surface displacement data caused by the plastic strains (29). Once ϵ_{kl}^p

Table 1 Comparison of the computed plastic strains with the actual plastic strains from (29). Ω is given in (28) and $\Omega^* = \{x | r \le 0.3\}$, where $r^2 = x_1^2 + (x_2 - 2.0)^2$.

ELEMENT Coordinate	Element #	COMPUTED γ^p_{12}	COMPUTED ε^p_{22}	ACTUAL γ^p_{12}	ACTUAL ε^p_{22}
$0 \le \theta \le \pi/3$; $0 \le r \le 0.1$	1	1.0	1.0	1	1
$\pi/3 \le \theta \le 2\pi/3$; $0 \le r \le 0.1$	2	1.0	1.0	1	1
$2\pi/3 \le \theta \le \pi$; $0 \le r \le 0.1$	3	1.0	1.1	1	1
$\pi \le \theta \le 4\pi/3$; $0 \le r \le 0.1$	4	1.0	1.1	1	1
$4\pi/3 \le \theta \le 5\pi/3$; $0 \le r \le 0.1$	5	1.0	1.1	1	1
$5\pi/3 \le \theta \le 2\pi$; $0 \le r \le 0.1$	6	1.0	1.0	1	1
$0 \le \theta \le \pi/6$; $0.1 \le r \le 0.2$	7	0.68	0.64	1	1
$\pi/6 \le \theta \le \pi/3$; $0.1 \le r \le 0.2$	8	0.82	0.64	1	1
$\pi/3 \le \theta \le \pi/2$; $0.1 \le r \le 0.2$	9	0.90	0.74	1	1
$\pi/2 \le \theta \le 2\pi/3$; $0.1 \le r \le 0.2$	10	0.81	0.82	1	1
$2\pi/3 \le \theta \le 5\pi/6$; $0.1 \le r \le 0.2$	11	0.65	0.79	1	1
$5\pi/6 \le \theta \le \pi$; $0.1 \le r \le 0.2$	12	0.60	0.75	1	1
$\pi \le \theta \le 7\pi/6$; $0.1 \le r \le 0.2$	13	0.71	0.81	1	1
$7\pi/6 \le \theta \le 4\pi/3$; $0.1 \le r \le 0.2$	14	0.80	0.89	1	1
$4\pi/3 \le \theta \le 3\pi/2$; $0.1 \le r \le 0.2$	15	0.77	0.83	1	1
$3\pi/2 \le \theta \le 5\pi/3$; $0.1 \le r \le 0.2$	16	0.80	0.76	1	1
$5\pi/3 \le \theta \le 11\pi/6$; $0.1 \le r \le 0.2$	17	0.78	0.80	1	1
$11\pi/6 \le \theta \le 2\pi$; $0.1 \le r \le 0.2$	18	0.67	0.75	1	1
$0 \le \theta \le \pi/8$; $0.2 \le r \le 0.3$	19	-0.18	-0.07	0	0
$\pi/8 \le \theta \le \pi/4$; $0.2 \le r \le 0.3$	20	0.18	-0.14	0	0
$\pi/4 \le \theta \le 3\pi/8$; $0.2 \le r \le 0.3$	21	0.53	0.04	0	0
$3\pi/8 \le \theta \le \pi/2$; $0.2 \le r \le 0.3$	22	0.66	0.29	0	0
$\pi/2 \le \theta \le 5\pi/8$; $0.2 \le r \le 0.3$	23	0.49	0.40	0	0
$5\pi/8 \le \theta \le 3\pi/4$; $0.2 \le r \le 0.3$	24	0.12	0.29	0	0
$3\pi/4 \le \theta \le 7\pi/8$; $0.2 \le r \le 0.3$	25	-0.20	0.03	0	0
$7\pi/8 \le \theta \le \pi$; $0.2 \le r \le 0.3$	26	-0.24	-0.16	0	0
$\pi \le \theta \le 9\pi/8$; $0.2 \le r \le 0.3$	27	0.06	-0.06	0	0
$9\pi/8 \le \theta \le 5\pi/4$; $0.2 \le r \le 0.3$	28	0.39	0.27	0	0
$5\pi/4 \le \theta \le 11\pi/8$; $0.2 \le r \le 0.3$	29	0.25	0.35	0	0
$11\pi/8 \le \theta \le 3\pi/2$; $0.2 \le r \le 0.3$	30	-0.05	0	0	0
$3\pi/2 \le \theta \le 13\pi/8$; $0.2 \le r \le 0.3$	31	0.13	-0.18	0	0
$3\pi/8 \le \theta \le 7\pi/4$; $0.2 \le r \le 0.3$	32	0.36	0.18	0	0
$\pi/4 \le \theta \le 15\pi/8$; $0.2 \le r \le 0.3$	33	0.08	0.46	0	0
$15\pi/8 \le \theta \le 2\pi$; $0.2 \le r \le 0.3$	34	-0.24	0.25	0	0

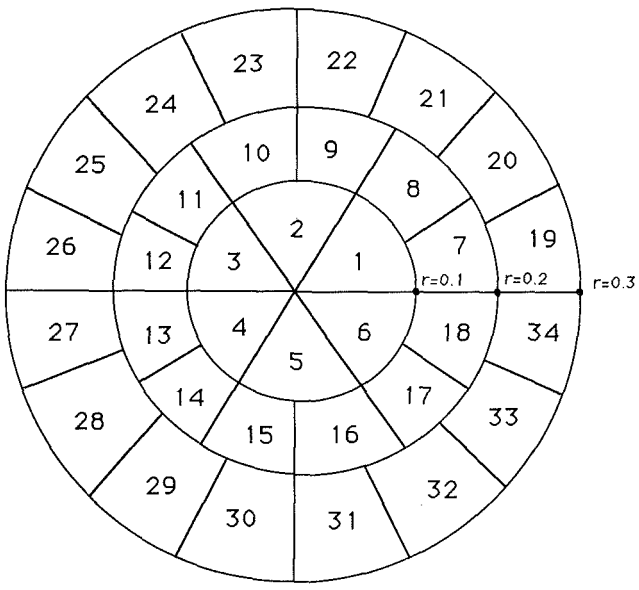


Fig. 6 Grid pattern and element numbers of the equivalent damage domain Ω^*

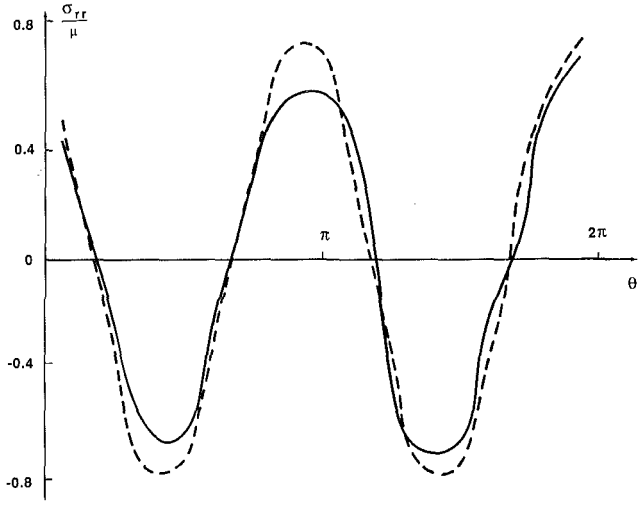


Fig. 7 Residual stress at σ_{rr} at $r/a^* = 1.1$, where $r = \sqrt{x_1^2 + (x_2 - 2)^2}$. The solid line is the computed result and the dashed line is the one induced by the plastic strains in (29).

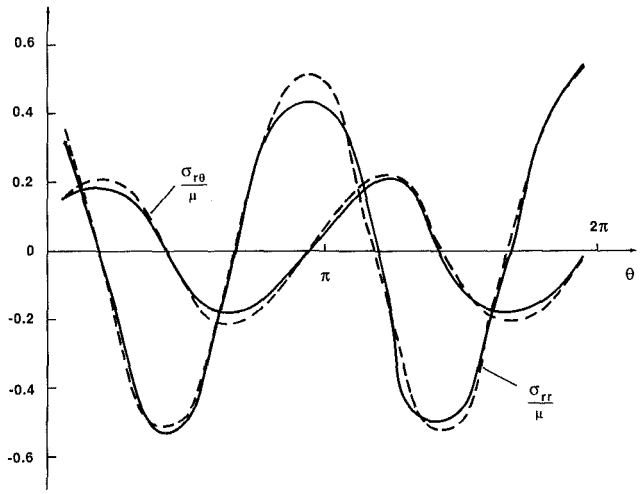


Fig. 8 Residual stresses σ_{rr} and $\sigma_{r\theta}$ and $r/a^* = 1.5$

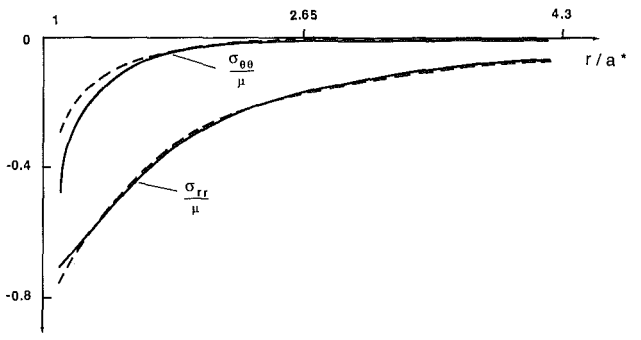


Fig. 9 Residual stresses σ_{rr} and $\sigma_{\theta\theta}$ at $\theta = \pi/2$

are obtained, we can compute the stress field on the boundary $\partial\Omega^*$ and outside of Ω^*.

In applying our method to an engineering problem of damage evaluation, we need to measure the residual surface displacements experimentally. There are many techniques available for this purpose. For example, the sterescopic analysis of optical micrographs (Cox and Morris (1986, 1988)), determines the relative displacement fields by comparing a pair of optical micrographs, one taken before the loadings are applied, and the other is taken afterwards. The objective of this paper, however, is to demonstrate the numerical inversion of the stresses if the experimental data are provided. Therefore, instead of measuring the surface displacements experimentally, we plug (29) into (16) and solve (16) for the surface displacements $u_m(\mathbf{x}')$. This is a well-posed conventional forward analysis. The domain Ω^* in (16) is chosen as Ω given in (28). The computed surface displacements

should be the same as the one obtained from experiments. We use this computed surface displacement data in our inverse computation. That is, we solve (17) for unknown ϵ_{kl}^p by using the computed surface displacement data.

The equation (16) is equivalent to (23) and (24). Therefore, the equations we need in the inverse computations are (23) and (24).

For any chosen positive α, the equation (23) is a well-posed Fredholm integral equation of the second kind, and can be solved by using conventional techniques. An interation procedure is utilized to make α satisfy the equation (24).

It has been proved (Gao and Mura (1988a) that $f(\alpha)$, defined in (27) is an increasing function and

$$\lim_{\alpha \to +\infty} f(\alpha) > 0, \quad \lim_{\alpha \to 0^+} f(\alpha) < 0.$$

Therefore, the following algorithm is employed to solve (23) and (24):

1 ϵ is chosen from our knowledge on the accuracy of the displacement data. δ is a small positive number for convergence criterion.

2 Choose positive numbers α_1 and α_2 such that $\alpha_1 < \alpha_2$, $f(\alpha_1) < 0$ and $f(\alpha_2) > 0$.

3 Let $\alpha = \alpha_1 + \alpha_2/2$ and solve (23). If $|f(\alpha)| < \delta$ then taken $\mathbf{V}^{(\alpha)}(\mathbf{x})$ as the solution and stop. Otherwise, go to the next step.

4 If $f(\alpha) > 0$, let $\alpha_1 = \alpha_2$, $\alpha_2 = \alpha$. Otherwise, let $\alpha_1 = \alpha$, $\alpha_2 = \alpha_2$.

In any case, go to Step 3.

Figure 6 shows the meshes of discretization for Ω^*. Table 1 shows the actual and computed plastic strains. The actual plastic strains are those from (29), while the computed ones are obtained by solving (23).

Figures 7–9 illustrate the residual stresses. The dashed lines are the actual stresses. The solid lines are the computed stresses.

This example indicates that the equivalent plastic strains in the equivalent damage domain Ω^* may be substantially different from the actual plastic strains in Ω. However, the stresses induced by the equivalent plastic strains are equal to the actual ones outside Ω^* (the deviations between solid and dashed lines in Figs. 7–9 are due to the errors of numerical computations). This confirms our prediction in Section 3 of this paper.

6 Conclusion

In this paper, the surface displacement data are utilized to evaluate the residual stress field in the vicinity of the damage domain caused by a series of unknown loading. It has been shown that the equivalent plastic strains, though different from the actual ones, induce the actual stresses outside of the equivalent damage domain.

The plastic strains ϵ_{kl}^p in this paper can be interpreted as eigenstrains. Eigenstrain (Mura (1982)), is a generic term representing nonelastic strains caused by thermal expansion, geometric misfit, phase transformation, etc. The existence of cracks and inhomogeneities can be simulated by appropriate distributions of eigenstrains in the homogeneous materials. Therefore, the results of this paper can be extended to the inversion of residual stresses caused by cracks and inhomogeneities, etc.

Acknowledgment

This research was supported under U.S. Army Research Office Contract Number DAAL03-88-C-0027 through a subcontract with Rockwell International Science Center.

References

Cox, B., Morris, W., and James, M., 1986, "High Sensitivity, High Spatial Resolution Strain Measurements in Alloys and Composites," *Proc. Conf. on Nondestructive Testing and Evaluation of Advanced Materials and Composites*, Colorado Spring, Co., pp. 25–39.

Dulikravich, G., 1988, "Inverse Design and Active Control Concepts in Strong Unsteady Heat Conduction," *Appl. Mech. Rev.*, Vol. 41, pp. 270–277.

Furuhashi, R., and Mura, T., 1979, "On the Equivalent Inclusion Method and Impotent Eigenstrains," *J. Elasticity*, Vol. 9, pp. 263–270.

Gao, Z., and Mura, T., 1988a, "Evaluation of Plastic Damages from Surface Displacement Data," submitted to *Int. J. Solids Structures*.

Gao, Z., and Mura, T., 1988b, "Nondestructive Evaluation of Interfacial Damages in Composite Materials," *Int. J. Solids Structures*, in press.

James, M. R., Morris, W., and Cox, B., 1988, "High Accuracy Ultimate Strain Field Mapper," submitted to *Experimental Mechanics*.

Kinoshita, N., and Mura, T., 1956, "On the Boundary Value Problem of Elasticity," *Res. Rep.*, Faculty of Engng., Meiji Univ. No. 8, pp. 1–7.

Kubo, S., 1988a, "Inverse Problems Related to the Mechanics and Fracture of Solids and Structures," *JSME International Journal*, Vol. 31, pp. 157–166.

Mura, T., 1982 and 1987, *Micromechanics of Defects in Solids*, Martinus Nijhoff Publ.

Mura, T., Cox, B., and Gao, Z., 1986, "Computer-Aided Nondestructive Measurements of Plastic Strains from Surface Measurements," *Proc. Int. Conf. on Computational Mechanics*, Tokyo, Vol. 2, pp. 43–48.

Sobieczky, H., 1988, "Research on Inverse Design and Optimization in Germany," *Appl. Mech. Rev.*, Vol. 41, pp. 239–246.

Stanitz, H., 1988, "A Review of Certain Inverse Problem Methods for the Design of Ducts with 2- or 3-Dimensional Potential Flow," *Appl. Mech. Rev.*, Vol. 41, pp. 217–238.

Application of an Internal Variable Constitutive Model to Predict Creep Response of an Aluminum Alloy Under Multiaxial Loading

I. U. Mahmood
The University of Oklahoma,
Norman, Ok. 73019

M. O. Faruque
University of Petroleum and Minerals,
Dahran, Saudi Arabia

M. M. Zaman
School of Civil Engineering and
Environmental Science,
The University of Oklahoma,
Norman, Ok. 73019

This paper discusses the application of an internal variable, creep constitutive model, where the concept of piecewise linearity in the effective stress-creep strain rate relationship is utilized. Since the concept of piecewise linearity is assumed, an explicit functional form for creep strain rate at all levels of stress and temperature is not required. The aforementioned constitutive model is used to predict the creep response of an aluminum alloy (2618-T61) at 200°C and subjected to multiaxial loading. The results are compared with available experimental results. The model shows excellent agreement in the trend of creep response. The quantitative values also agree quite good with the experimental results.

Introduction

Many researchers (Besseling, 1953; Garaflo, 1965; Henderson, 1979; Kocks, 1976; Krans, 1980; Krempl, 1974; Laften and Stouffer, 1978; Odqvist, 1974; Ostrom and Lagneborg, 1976; Rabotonov, 1969; Robinson, 1978) in the past have presented phenomenological theories to describe transient and steady-state creep behavior of materials. These models generally assume the creep strain rate to be a specified function of the applied stress, the time elapsed after the application of stress and the current temperature, if the model is temperature-dependent, as such they can not satisfactorily predict the instantaneous increase in creep strain rate caused by stress reversal. It has been experimentally observed that stress changes cause transient noncoaxiality between the stress tensor and the creep rate tensor. Reliable prediction of creep response needs constitutive models with the desired internal variables related to physical mechanisms which govern the behavior of the material under stress.

Kujawski and Mroz (1980), Leckie and Ponter (1974), Malinin and Khadjinsky (1972), Hart (1976), Miller (1976), Larsson and Storakes (1978), Chaboche (1977), and Faruque (1985) proposed phenomenological theories in which the creep strain rate is assumed to be dependent on the applied stress as well as on a number of hardening parameters.

Mroz and Trampczynski (1984) proposed a constitutive model based on the concept of kinematic hardening and taking account of the memory of minimal prestress on the back stress space. In their formulation, they assumed the existence of a creep potential whose gradient with respect to the stress tensor yields the creep rate tensor.

A constitutive model based on the two-surface kinematic and isotropic hardening concept of the plasticity theory was proposed by Chaboche et al. (1979). This concept was later extended by Murakami and Ohno (1982) and Ohno et al. (1985). Certain models developed based on the visco-plasticity theory are proposed by Kujawski and Mroz (1980), Bodner and Merzer (1978), and Cernocky and Krempl (1980). Viscoelastic models were proposed by Cho and Findley (1980, 1981, 1982, 1983a, 1983b, 1984), Ding and Findley (1984a, 1984b, 1985), and by Lai and Findley (1980, 1982). Krieg (1977), Hart (1976), Miller (1976), Paslay and Wells (1976), Ponter and Leckie (1976), and Robinson et al. (1976) have proposed unified creep and plasticity models. Krieg et al. (1978), Swearengen and Rhode (1977), Swearengen et al. (1976), and Lagneborg (1971, 1972) considered thermally-activated mechanisms such as dislocation climb and mass diffusion in their models.

In most of the creep models described, it is assumed that a single creep strain-rate equation is applicable over all levels of temperature and stresses. Hence, the predicted response does not always agree with the experimental results.

Faruque and Zaman (1988) recently proposed an internal variable creep constitutive model in which the concept of piecewise linearity in the effective stress-creep strain-rate relationship was utilized. Since the piecewise linearity concept is assumed, the creep strain rate need not be expressed by an explicit functional form. The present paper is concerned with the application of this model to predict the creep response of an aluminum alloy (2618-T61) under multiaxial loading. The results are compared with the available experimental results.

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

Discussion on this paper should be addressed to the 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 5, 1988; final revision, September 7, 1988.

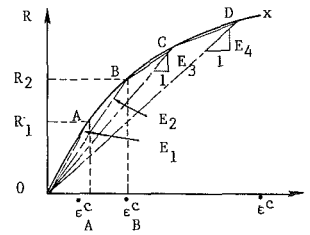


Fig. 1(a) Typical $\dot{\epsilon}^c - R$ response curve showing piecewise linear approximation ($\dot{\epsilon}^c$ = equivalent uniaxial creep strain rate)

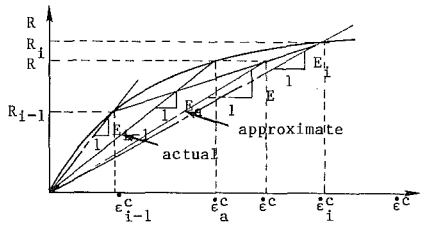


Fig. 1(b) Actual and approximate creep moduli between two consecutive modes ($\dot{\epsilon}^c$ = equivalent uniaxial creep strain rate)

Description of the Model

The general description and the formulation of the model are presented in a recent paper by Faruque and Zaman (1988). A summary of the model is presented in this section.

The creep strain-rate tensor $\dot{\epsilon}^c$ is expressed as follows:

$$\dot{\epsilon}^c = \frac{1}{E}\,\lambda\, R \tag{1}$$

where, E is a secant creep modullus associated with the creep surface $F=0$ and λ is a unit tensor defined by

$$\lambda = \frac{\text{Grad } F}{||\text{Grad } F||}\,. \tag{2}$$

$F=0$ defines the creep surface in the effective stress space as

$$F = \frac{3}{2}\, tr(\tau^2) - R^2 = 0. \tag{3}$$

The effective stress tensor τ is defined by

$$\tau = \mathbf{S} - \alpha \tag{4}$$

where $\mathbf{S}$ and α are deviatoric stress tensor and back stress tensors, respectively.

An important feature of the proposed model is the piecewise linear approximation of the actual (equivalent uniaxial) effective stress (equivlaent uniaxial) creep strain-rate relationship as shown in Fig. 1(a). Referring to Fig. 1(b), secant creep modulus, E, for any (equivalent uniaxial) effective stress (R) can now be written as

$$\frac{1}{E} = \frac{1}{R}\left[\frac{R_{i-1}}{E_{i-1}} + \frac{(R-R_{i-1})}{(R_i-R_{i-1})}\left(\frac{R_i}{E_i} - \frac{R_{i-1}}{E_{i-1}}\right)\right]$$

where R_i and R_{i-1} are (equivalent uniaxial) effective stresses and E_i and E_{i-1} are the secant creep modulii at nodes i and $i-1$, respectively.

In equations (1) and (3) R is the radius of the creep surface at any instant t and is defined by

$$R^2 = \frac{3}{2}\, tr\,(\tau^2). \tag{5}$$

The time rate of back stress α may be expressed as

$$\dot{\alpha} = \frac{d\alpha}{dt} = B\,(\alpha_s - \alpha) \tag{6}$$

where, α_s is the saturated back stress tensor defined by

$$\alpha_s = \nu\, h(||\mathbf{s}||,\, T) \tag{7}$$

T being the absolute temperature and ν being a unit tensor defined by

$$\nu = \mathbf{S}/||\mathbf{S}||. \tag{8}$$

In equations (6) and (7) the specific form of h for polycrystalline materials are chosen as follows:

$$h = H(||\mathbf{S}||)\,[1 - G(T)] \tag{9}$$

where, H is an unknown function of the applied stress, and $G(T)$ is temperature-dependent function having the expressions:

$$H = A(||\mathbf{S}||)^m \tag{10}$$

$$G(T) = \exp(-Q/RT). \tag{11}$$

In equations (10) and (11), A and m are material constants, Q is the activation energy, R is the universal gas constant, and T is the absolute temperature.

Evaluation of Model Parameters

For uniaxial creep, the constitutive equations for α_s and α can be written as:

$$\alpha_s = A\,|\sigma|^m[1 - \exp(-Q/RT)]\text{sgn}(\sigma) \tag{12}$$

$$\dot{\alpha} = B(\alpha_s - \alpha) \tag{13}$$

where A, m, Q, R, and B are unknown material constants.

For a set of uniaxial creep tests at the same temperature

$$\alpha_s = \bar{A}\,|\sigma|^m\text{sgn}(\sigma) \tag{14}$$

where $\bar{A} = A[1 - \exp(-Q/RT)]$ is a modified parameter which combines A, Q, and R. To predict creep response at a temperature other than the test temperature, the parameters, A, Q, and R needs to be determined explicitly.

The parameters $\bar{A}$ and m can be determined directly if the back stresses at saturation, α_s, are known for a number of uniaxial creep test. By using the data reported by Krieg et al. (1978) and least-square fit of equation (14), the material constants were determined to be $\bar{A}=0.6$ and $m=1.0$. Note that because experimental values for the α_s was not available for 2618-T61 aluminum, the data (15) for $<111>$ aluminum was used here. The secant modulii values, E, obtained from $\tau - \dot{\epsilon}^c$ curve for aluminum 2618-T61 and reported by Faruque and Zaman (1988) are used to predict the creep strain rate.

Discussion of Results

The model discussed in the preceding section is used to back predict the creep strain of aluminum alloy 2618-T61. The model-predicted creep strain rates are integrated to obtain the creep strain at any instant of time. The time-independent elastic strains are added to the creep strain to obtain the total axial or shear strain. Two sets of results are reported:

(1) multiaxial creep under proportional loading and
(2) multiaxial creep under side steps of tension and torsion and stress reversal.

The results are compared with the experimental results reported by Ding and Findley (1984) and Findley and Lai (1981). Loading histories or sequences are shown in each figure. Figure 2 shows the temporal variation of axial strain under tensile loading. In the experiment, the direction of loading remained constant and the magnitudes were changed abruptly. The tensile load was applied instantaneously and maintained for 48 hours. The load was withdrawn abruptly and a period of recovery followed. The load was then taken

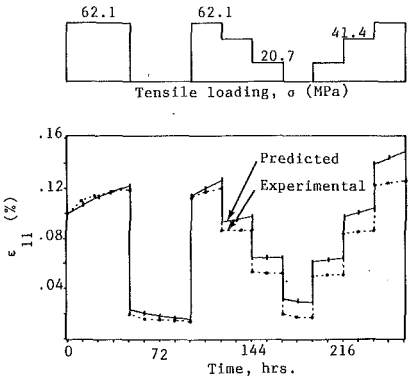


Fig. 2 Tensile strain under tensile loading

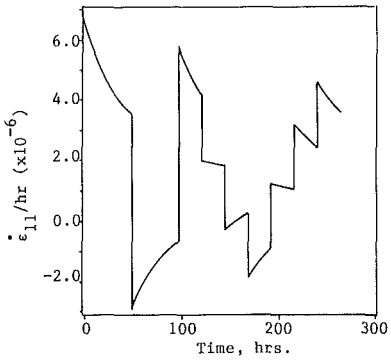


Fig. 3 Axial creep strain-rate variation under tensile loading

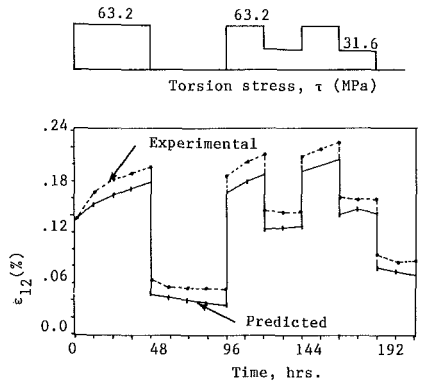


Fig. 4 Shearing strain under torsion loading

abruptly to original level. The next steps comprised of gradually lowering the tensile load until all loads were removed. In the next steps the load was gradually increased to the original level. The results compare favorably with the experimental result. The trend is in excellent agreement. Quantitatively, the predicted values deviate in the successive loading steps. The maximum difference between the predicted and observed response was approximately 13 percent. The deviation may be due to the effect of isotropic hardening which was not included in the model.

Figure 3 presents the variation of creep strain-rate with time for the loading case shown in Fig. 2. It is observed that any change in the state of stress causes a jump discontinuity in the creep strain-rate. This is similar to experimental observations.

Figure 4 shows shear strain variation with time under pure torsional loading. Here, also, the experimental results and the predicted values show consistent trend. Figure 5 shows the creep strain-rate under the same torsional loading. In both Figs. 2 and 4, the removal of load was followed by a period of creep recovery. The reloading steps always resulted in tran-

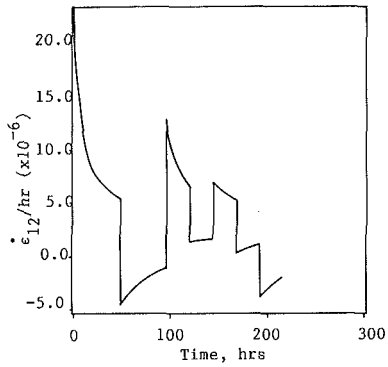


Fig. 5 Shearing creep strain-rate variation with time under torsional loading

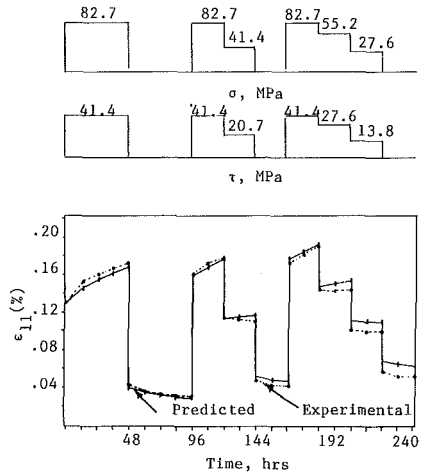


Fig. 6 Tensile strain under proportional combined tensile and torsion stress

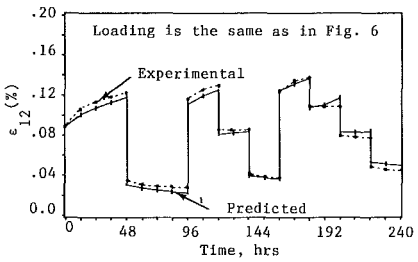


Fig. 7 Shearing strain under proportional combined tensile and torsion stress

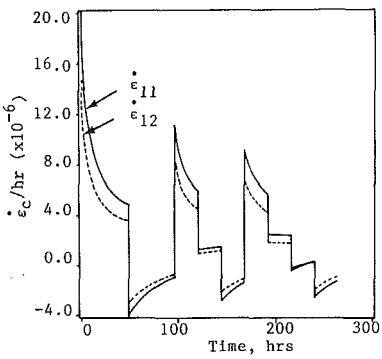


Fig. 8 Creep strain-rate variation with time under proportional combined tensile and torsional stress

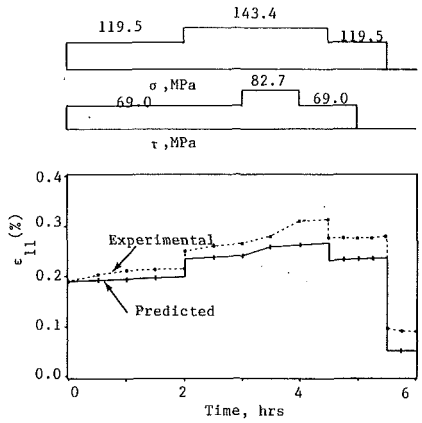


Fig. 9 Tensile strain for combined tension and torsion creep under side steps of loading, unloading, and recovery

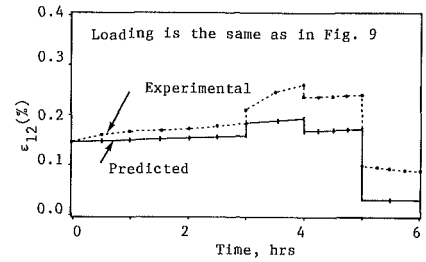


Fig. 10 Shearing strain for combined tension and torsion creep under side steps of loading, unloading, and recovery

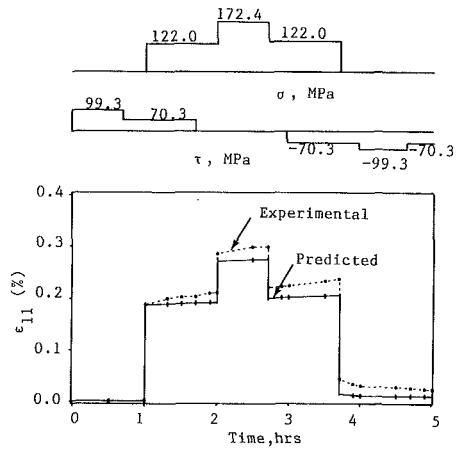


Fig. 11 Tensile strain for combined tension and torsion creep under side steps, partial and complete reversal of torsion

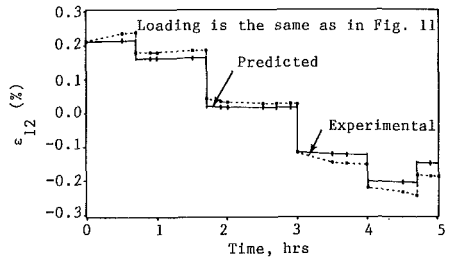


Fig. 12 Shearing strain for combined tension and torsion creep under side steps, partial, and complete reversal of torsion

sient creep strain. Partial unloading resulted in creep recovery if the unloading was substantial, otherwise it resulted in transient creep as may be seen in the fourth load step in Fig. 2. The experimental results also showed similar trends.

Figures 6 and 7 show the axial strain versus time and shear strain versus time relationships, respectively, for the combined proportional tension and torsional loading. In both cases the predicted and experimental results agree very well both qualitatively and quantitatively. Figure 8 shows the variation of axial and shear creep strain with time under the same loading.

Figures 9 through 12 show creep behavior of aluminum alloy 2618-T61 under side steps of tension, torsion, and stress reversal. Figure 9 shows the total strain versus time response. In Step 2, when the axial stress σ was increased keeping the torsional stress, τ constant, there was a small increase in creep rate; however, increase in τ at constant σ resulted in a large increase of creep rate. Reducing τ at constant σ decreases the creep rate significantly. In the next step, the decrease of σ at the constant τ did not significantly reduce the creep rate. This behavior is similar to experimental observations. Creep recovery occurred when σ was completely removed.

Figure 10 shows the result of the same test with plots of shear strain versus time. The creep rate in shear increased significantly when τ was increased at a constant σ. Partial unloading of τ did not result in creep recovery. Full unloading of τ resulted in creep recovery of shear strain component. Another observation in both the figures is that the complete unloading of one component keeping the other component constant did not affect the behavior of creep associated with the constant component. This is exactly the trend observed in experimental results as well.

Figures 11 and 12 show the strain (axial or shear) versus time relationship for combined tension and torsion under side steps, partial and complete reversal of torsion. It is observed that initial torsion with no tension induced only shear creep strain. Reversal of torsion did not affect the recovery of axial component after the tensile loading is fully removed. In Fig. 12 it is observed that partial unloading did not result in creep recovery. Full unloading resulted in creep recovery with creep rate decreasing with time. Unloading of the tensile component did not affect the creep recovery process in shear. Reverse torsion induced a larger creep rate decreasing with time. Step up of shear load increased the creep rate. The trends are in all cases similar to experimental results.

Conclusions

An internal variable creep-constitutive model is used to predict the response of aluminum alloy 2618-T61. Piecewise linearity of effective stress-creep strain-rate is assumed and the results are compared with experimental results. The model shows excellent agreement with experimental trends and results. The quantitative results show good agreement and may be improved by considering another internal variable, D (Drag stress) which represents isotropic hardening and by incorporating creep strain-rate explicitly into the evolution law for the back stress and drag stress.

References

Besseling, J. F., 1953, "A Theory of Elastic Plastic and Creep Deformation of an Initially Isotropic Material Showing Anisotropic Strain Hardening, Creep, Recovery, and Secondary Creep," ASME JOURNAL OF APPLIED MECHANICS, Vol. 25, pp. 529–536.

Bodner, S. R., and Merzer, A., 1978, "Viscoplastic Constitutive Equations for Copper With Strain History and Temperatures Effects," ASME *Journal of Engineering Materials and Technology*, Vol. 100, pp. 338–394.

Cernocky, E. P., and Krempl, E., 1980, "A Theory of Thermoviscoplasticity Based on Infinitesimal Total Strain," *Int. J. Solids and Structures*, Vol. 16, pp. 723–741.

Chaboche, J. L., 1977, "Viscoelastic Constitutive Equations for the Description of Cyclic and Anisotropic Behavior of Metals," *Bull. Acad. Polon. Sci., Ser Sci. Tech.*, Vol. 25, p. 33.

Chaboche, J. L., Dang Van, K., and Cordier, G., 1979, "Modelization of the Strain Memory Effect on the Cyclic Hardening of 316 Stainless Steel," *Trans. 5th Int. Conf. Struc. Mech. Reactor Tech.*, T. A. Jaeger, and B. A. Boley, eds., North-Holland, Amsterdam.

Cho, U. W., and Findley, W. N., 1980, "Creep and Creep Recovery of 304 Stainless Steel Under Combined Stress With a Representation by a Viscous-Viscoelastic Model," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 755-761.

Cho., U. W., and Findley, W. N., 1981, "Creep and Creep Recovery of 304 Stainless Steel at Low Stresses With Effects of Aging on Creep and Plastic Strains," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 785-790.

Cho., U. W., and Findley, W. N., 1982, "Creep and Plastic Strains of 304 Stainless Steel at 593°C Under Step Stress Changes, Considering Aging," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 297-304.

Cho., U. W., and Findley, W. N., 1983a, "Creep and Plastic Strains Under Side Steps of Tension and Torsion for 304 Stainless Steel at 593°C," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 580-586.

Cho., U. W., and Findley, W. N., 1983b, "Creep and Plastic Strains Under Stress Reversal in Torsion With and Without Simultaneous Tension for 304 Stainless Steel at 593°C," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 587-592.

Cho, U. W., and Findley, W. N., 1984, "Creep and Creep Recovery of 2618-T61 Aluminum Under Variable Temperature," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 816-820.

Ding, J. L., and Findley, W. N., 1984a, "48 Hour Multiaxial Creep and Creep Recovery of 2618 Aluminum Alloy at 200°C," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 125-132.

Ding, J. L., and Findley, W. N., *1984b*, "Multiaxial Creep of 2618 Aluminum Under Proportional Loading Steps," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 133-140.

Ding, J. L., and Findley, W. N., 1985, "Nonproportional Loading Steps in Multiaxial Creep of 2618 Aluminum," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 621-628.

Faruque, M. O., 1985, "On the Description of Cyclic Creep and Rate Dependent Plastic Deformation," *Acta Mechanica*, Vol. 55, pp. 123-136.

Faruque, M. O., and Zaman, M. M., 1988, "On Modelling Steady State and Transient Creep of Polycrystalline Solids," *Acta Mechanica*, Vol. 71, pp. 115-136.

Findley, W. N., Cho, U. W., and Ding, J. L., 1979, "Creep of Metals and Plastics Under Combined Stresses, A Review," *J. Engng. Mat. Tech.*, Vol. 101, pp. 365-368.

Findley, W. N., and Lai, J. S., 1978, "Creep and Recovery of 2618 Aluminum Alloy Under Combined Stress With a Representation by a Viscous-Viscoelastic Model," ASME JOURNAL OF APPLIED MECHANICS, Vol. 45, pp. 507-514.

Findley, W. N., and Lai, J. S., 1981, "Creep of 2618 Aluminum Under Side-Steps of Tension, Torsion and Stress Reversal Predicted by a Viscous-Viscoelastic Model," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 47-54.

Garaflo, F., 1965, *Fundamentals of Creep and Creep-Rupture in Metals*, Macmillan, New York.

Gilman, J. J., 1966, "Progress in the Microdynamical Theory of Plasticity," *Proc. of the Fifth U. S. Congress of Applied Mech.*, ASME, New York, pp. 385-403.

Gittus, J. H., 1978, "Dislocation Creep Under Cyclic Stressing: Physical Model and Theoretical Equations," *Acta Met.*, Vol. 26, pp. 305-317.

Hart, E. W., 1976, "Constitutive Relations for the Nonelastic Deformation of Metals," ASME *Journal of Engineering Materials and Technology*, Vol. 98, pp. 193-202.

Henderson, J., 1979, "An Investigation of Multi-Axial Creep Characteristics of Metals," ASME *Journal of Engineering Materials and Technology*, Vol. 101, pp. 356-364.

Kocks, U. F., 1976, "Laws for Work-Hardening and Low-Temperature Creep," Trans. ASME *Journal of Engineering Materials and Technology*, Vol. 98, pp. 76-85.

Krans, H., 1980, *Creep Analysis*, John Wiley and Sons, Inc., New York.

Krempl, E., 1974, "Cyclic Creep—An Interpretive Literature Survey," *WRC Bull.*, No. 195, Welding Research Council, p. 63.

Krieg, R. D., 1977, "Numerical Integration of Some New Unified Plasticity-Creep Formulations," *Proc. 4th International Conference on Structural Mechanics in Reactor Technology*, Comm. of the European Communities, Luxembourg.

Krieg, R. D., Swearengen, J. C., and Rhode, R. W., 1978, "A Physically-Based Internal Variable Model for Rate-Dependent Plasticity," *Inelastic Behavior of Pressure Vessel and Piping Components*, T. Y. Chang, E. Krempl, eds., pp. 15-28.

Kujawski, D., and Mroz, Z., 1980, "A Viscoplastic Material Model and its Application to Cyclic Loading," *Acta Mechanica*, Vol. 36, pp. 213-230.

Laften, J. H., and Stouffer, D. C., 1978, "An Analysis of High Temperature Metal Creep: Part II-A—Constitutive Formulation and Verification," ASME *Journal of Engineering Materials and Technology*, Vol. 100, pp. 363-380.

Lagneborg, R., 1971, "A Theoretical Approach to Creep Deformation During Intermittent Load," *ASME J. Basic Enging.*, Vol. 93, p. 205.

Lagneborg, R., 1972, "A Modified Recovery-Creep Model and its Evaluation," *Met. Sci. J.*, Vol. 6, pp. 127-133.

Lagneborg, R., 1981, "Creep: Mechanisms and Theories," *Creep and Fatigue in High Temperatures Alloys*, J. Bressers, ed., Applied Science Publishers.

Lai, J. S., and Findley, W. N., 1980, "Creep of 2618 Aluminum Under Step Stress Changes Predicted by a Viscous-Viscoelastic Model," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 21-26.

Lai, J. S., and Findley, W. N., 1982, "Simultaneous Stress Relaxation in Tension, and in Creep Torsion of 2618 Aluminum at Elevated Temperature," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 19-25.

Larsson, B., Storakes, B., 1978, "A State Variable Interpretation of Some Rate-Dependent Inelastic Properties of Steel," ASME *Journal of Engineering Materials and Technology*, Vol. 100, pp. 395-401.

Leckie, F. A., Ponter, A. R. S., 1974, "On the State Variable Description of Creeping Materials," *Ing. Archiv.*, Vol. 43, pp. 158-167.

Malinin, N. N., and Khadjinsky, G. M., 1972, "Theory of Creep With Anisotropic Hardening," *Int. J. Mech. Sci.*, Vol. 14, pp. 235-246.

Miller, A., 1976, "An Inelastic Constitutive Model for Monotonic, Cyclic and Creep Deformation," ASME *Journal of Engineering Materials and Technology*, Vol. 98, pp. 97-113.

Mroz, Z., 1981, "On Generalized Kinematic Hardening Rule With Memory of Maximal Prestress," ASME JOURNAL OF APPLIED MECHANICS, Vol. 5, pp. 241-260.

Mroz, Z., and Trampczynski, W. A., 1984, "On the Creep-Hardening Rule for Metals with a Memory of Maximal Prestress," *Int. J. Solids and Structures*, Vol. 20, pp. 467-486.

Murakami, S., and Ohno, N., 1982, "A Constitutive Equation of Creep Based on the Concept of a Creep-Hardening Surface," *Int. J. Solids and Structures*, Vol. 18, pp. 597-609.

Odqvist, F. K. G., 1974, *Mathematical Theory of Creep and Creep Rupture*, Oxford University Press.

Ohashi, Y., Ohno, N., and Kawai, M., 1982, "Evaluation of Creep Constitutive Equations for Type 304 Stainless Steel Under Repeated Multiaxial Loading," ASME *Journal of Engineering Materials and Technology*, Vol. 104, pp. 159-164.

Ohnami, M., Motoie, K., and Yoshida, N., 1969, "Study on the Influence of Strain History on Creep of Polycrystalline Metallic Materials at Elevated Temperature," *Zairyo*, Vol. 18, p. 226.

Ohno, N., Murakami, S., and Ueno, T., 1985, "A Constitutive Model of Creep Describing Creep Recovery and Material Softening Caused by Stress Reversals," ASME *Journal of Engineering Materials and Technology*, Vol. 107, pp. 1-6.

Ostrom, P., and Lagneborg, R., 1976, "A Recovery-Thermal Glide Creep Model," ASME *Journal of Engineering Materials and Technology*, Vol. 98, pp. 114-124.

Paslay, P. R., and Wells, C. H., 1976, "Uniaxial Creep Behavior of Metals Under Cyclic Temperature and Stress or Strain Variation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 98, pp. 445-449.

Ponter, A. R. S., and Leckie, F. A., 1976, "Constitutive Relationships for the Time-Dependent Deformation of Metals," ASME *Journal of Engineering Materials and Technology*, Vol. 98, pp. 47-51.

Rabotonov, Y. N., 1969, *Creep Problems of Structural Members*, North-Holland, Amsterdam.

Robinson, D. N., Pugh, C. E., and Corum, J. M., 1976, "Constitutive Equations for Describing High-Temperature Inelastic Behavior of Structural Alloys," *IAEA Int. Working Group on Fast Reactors Specialists—Meeting on High Temperature Structural Design Technology*, IAEA.

Robinson, D. N., 1978, "A Unified Creep-Plasticity Model for Structural Metals at High Temperature," ORNL/TM-5969., Oak Ridge National Laboratory, Oak Ridge, Tenn.

Swearengen, P. C., Rhode, R. W., and Hicks, D. L., 1976, "Mechanical State Relations for Inelastic Deformation on Iron: The Choice of Variables," *Acta Met.*, Vol. 24, pp. 969-975.

Swearengen, J. C., and Rhode, R. W., 1977, "Application of Mechanical State Relations at Low and High Homologous Temperatures," *Met. Trans.*, Vol. 8A, pp. 577-582.

N. Aravas
Department of Mechanical Engineering
and Applied Mechanics,
University of Pennsylvania,
Philadelphia, Penn. 19104
Assoc. Mem. ASME

R. M. McMeeking
Department of Materials and
Department of Mechanical Engineering,
University of California,
Santa Barbara, Calif. 93106
Mem. ASME

An Asymptotic Analysis of Three-Dimensional Extrusion

A new method of analysis of three-dimensional metal extrusion using asymptotic perturbation methods is presented in this paper. The plasticity model used depends on the first and second invariants of the stress tensor and covers a wide range of constitutive models commonly used for the analysis of metal-forming operations. It is shown that the three-dimensional extrusion problem can be approximated, to leading order, by a problem of generalized plane-strain. The results of the asymptotic analysis together with the finite element method are used to obtain approximate solutions for extrusions of elliptic or square cross-sections from round billets.

1 Introduction

Several solutions for extrusions of complicated shapes from cylindrical billets have been presented in the literature. These solutions involve three-dimensional finite element calculations (Boër and Webster, 1985) or are upper-bound solutions based on kinematically-admissible velocity fields (Juneja and Prakash, 1975; Nagpal and Altan, 1975; Basily and Sansome, 1976; Nagpal, 1977; Yang et al., 1978, 1979, 1984, 1986; Boër et al., 1979; Prakash and Khan, 1979; Hoshino and Gunasekera, 1980; Gunasekera and Hoshino, 1982, 1985; Cho and Yang, 1983; Kiuchi et al. 1983, 1984; Kiuchi, 1984; Han et al., 1986).

A new method of analysis of three-dimensional extrusions using asymptotic perturbation method is presented in this paper. We consider the extrusion of metal rods through lubricated dies to form a final shape with a different cross-section from the initial shape. A restriction placed on the analysis is that the cross-section of the die varies slowly down the extrusion axis. In practical situations the slope of the die is small when either the area reduction is small or the length of the die is large compared to the radius of the original cross-section. The asymptotic expansions are based on a small parameter ϵ which can be defined as the ratio of the total reduction of a characteristic dimension of the original cross-section to the length of the reduction region.

Similar asymptotic techniques have been used by Onat (1954) and recently by Johnson (1987) and Smet and Johnson (1988) for the analysis of two-dimensional problems. Onat (1954) considers a rigid perfectly-plastic material that obeys Tresca's yield condition with the associated flow rule and using a linearized version of the governing equations constructs axially-symmetric stress and velocity fields. Johnson (1987) and Smet and Johnson (1988) consider a rigid plastic material that obeys the von Mises yield criterion with the associated flow rule and analyze axisymmetric extrusion and plane-strain rolling. In these treatments, the partial differential equations of two-dimensional problems are reduced to ordinary differential equations by a regular perturbation scheme. Johnson (1987) and Smet and Johnson (1988) identify different regimes in the asymptotic analysis, controlled by the slope of the die and the coefficient of friction between the die and the workpiece, and show that, in some of those regimes, the leading order approximation involves "slab flow," as assumed by von Karman (1925) in the analysis of plane-strain rolling.

In this paper, we are concerned with the analysis of frictionless three-dimensional extrusion problems. The plasticity model used depends on the first and second invariants of the stress tensor and covers a wide range of constitutive models commonly used for the analysis of metal-forming operations. The flow rule is described in terms of a plastic potential which is, in general, different from the yield function. Using a regular perturbation method, we show that the three-dimensional problem can be approximated, to leading order, by a problem of generalized plane-strain. The leading order approximation involves slab flow, but in contrast to the common assumption of uniform stressing and deformation on each cross-section used by the so-called "slab methods," the leading order stress and deformation fields are, in general, functions of position on each cross-section. Using the results of the asymptotic analysis together with two-dimensional finite element analysis we obtain approximate solutions for extrusions of elliptic or square cross-sections from round billets.

Standard notation is used throughout. Boldface symbols denote tensors, the order of which is indicated by the context, and the summation convention is used for repeated Latin indices.

2 Description of the Method

2.1 Formulation of the Problem. We consider the extrusion of metal rods through lubricated dies to form a shape with a different cross-section. To simplify the discussion we

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 26, 1988; final revision, November 16, 1998.
Paper No. 89-WA/APM-19.

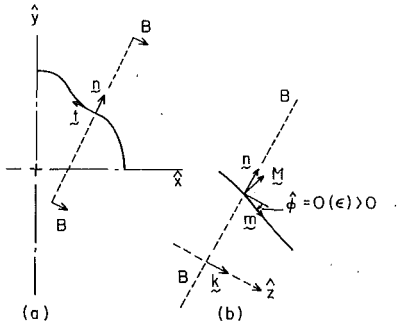


Fig. 1 *(a)* **One quadrant of the doubly-symmetric cross-section showing the unit vectors** n **and** t**; *(b)* projection BB showing the longitudinal profile of the die, the unit vectors** n, M, m **and** k, **and the small angle** $\hat{\phi}$

consider shapes which retain the same two orthogonal symmetry axes during extrusion, although the method is applicable to more general situations. Let $\hat{z}$ be the direction of extrusion and $\hat{x}$ and $\hat{y}$ be the axes of symmetry. The dies taper at a maximum angle which is of order ϵ (ϵ being a small positive number) but the taper may be nonuniform. A typical quarter cross-section is shown in Fig. 1 along with a projection of section BB illustrating the angle $\hat{\phi}$, which is defined by

$$\tan\hat{\phi} = \frac{\partial \hat{F}}{\partial \hat{z}} \left[\left(\frac{\partial \hat{F}}{\partial \hat{x}} \right)^2 + \left(\frac{\partial \hat{F}}{\partial \hat{y}} \right)^2 \right]^{-1/2} = O(\epsilon), \quad (1)$$

where $\hat{F}(\hat{x}, \hat{y}, \hat{z}) = 0$ is the equation of the die surface.

The equations for quasi-static equilibrium in the absence of body forces are

$$\frac{\partial \hat{\sigma}_{xx}}{\partial \hat{x}} + \frac{\partial \hat{\sigma}_{xy}}{\partial \hat{y}} + \frac{\partial \hat{\sigma}_{xz}}{\partial \hat{z}} = 0, \quad (2)$$

$$\frac{\partial \hat{\sigma}_{xy}}{\partial \hat{x}} + \frac{\partial \hat{\sigma}_{yy}}{\partial \hat{y}} + \frac{\partial \hat{\sigma}_{yz}}{\partial \hat{z}} = 0, \quad (3)$$

$$\frac{\partial \hat{\sigma}_{xz}}{\partial \hat{x}} + \frac{\partial \hat{\sigma}_{yz}}{\partial \hat{y}} + \frac{\partial \hat{\sigma}_{zz}}{\partial \hat{z}} = 0, \quad (4)$$

where $\hat{\sigma}$ is the stress tensor.

The material is assumed to be rigid plastic with a yield surface of the form

$$\Phi\left(\frac{\hat{I}_1}{\hat{\sigma}_0}, \frac{\hat{J}_2}{\hat{\sigma}_0^2}, H_\alpha \right) = 0 \text{ with } \frac{\partial \Phi}{\partial \hat{I}_1} \geq 0$$

$$\text{and } \frac{\partial \Phi}{\partial \hat{J}_2} \geq 0, \quad (5)$$

where Φ is a dimensionless function of dimensionless arguments, $\hat{I}_1 = \hat{\sigma}_{kk}$, $\hat{J}_2 = \hat{s}_{ij}\hat{s}_{ij}/2$, $\hat{\mathbf{s}}$ is the stress deviator, $\hat{\sigma}_0$ is the yield stress, and H_α ($\alpha = 1, 2, \ldots, n$) is a set of dimensionless scalar state variables.

The flow rule is given by

$$\hat{\mathbf{D}} = \hat{\lambda}\hat{\mathbf{N}}, \quad \hat{\mathbf{N}} = \frac{\partial g}{\partial \hat{\boldsymbol{\sigma}}}, \quad g = g\left(\frac{\hat{I}_1}{\hat{\sigma}_0}, \frac{\hat{J}_2}{\hat{\sigma}_0^2}, H_\alpha \right), \quad (6)$$

where $\hat{\mathbf{D}}$ is the deformation rate defined as the symmetric part of the spatial velocity gradient, $\hat{\lambda}$ is a non-negative scalar flow parameter, and g is a dimensionless plastic potential.

The constitutive model is completed by describing the evolution of state variables with continuing plastic straining:

$$\frac{dH_\alpha}{d\hat{t}} = \hat{h}_\alpha(\hat{\mathbf{D}}, \hat{\sigma}, H_\beta), \quad (7)$$

where $\hat{t}$ is time. For a rate-independent material, $\hat{h}_\alpha$ must be a homogeneous function of degree one in $\hat{\mathbf{D}}$, so that

$$\frac{dH_\alpha}{d\hat{t}} = \hat{\lambda}\hat{h}_\alpha(\hat{\mathbf{N}}, \hat{\sigma}, H_\beta). \quad (8)$$

The boundary conditions are that the normal velocity and the shear traction are zero on the die-metal interface, i.e.,

$$\mathbf{M} \cdot \hat{\mathbf{v}} = 0, \quad (9)$$

and

$$\mathbf{M} \cdot \hat{\sigma} \cdot \mathbf{t} = \mathbf{M} \cdot \hat{\sigma} \cdot \mathbf{m} = 0, \quad (10)$$

where $\mathbf{M}$ is the unit vector normal to the die surface, $\mathbf{t}$ is the unit vector tangent to the die profile on the $\hat{x}-\hat{y}$ plane, and $\mathbf{m}$ is the unit vector tangent to the die surface as shown in Fig. 1.

When the geometry of the die is slowly varying in one direction, a simplification of the problem can be brought about by stretching one coordinate direction with respect to others (Van Dyke, 1975). For slowly tapering dies (ϵ = small), the coordinate stretching is carried out by making $\hat{x}$ and $\hat{y}$ dimensionless by normalization of position in that plane by $\Delta\hat{R}_0$, the total reduction of a characteristic dimension of the original cross-section, i.e.,

$$x = \frac{\hat{x}}{\Delta\hat{R}_0} \text{ and } y = \frac{\hat{y}}{\Delta\hat{R}_0}. \quad (11)$$

Distance down the extrusion axis is made dimensionless by $\hat{L}$, the length of the reduction region, so that

$$z = \frac{\hat{z}}{\hat{L}} = \epsilon\frac{\hat{z}}{\Delta\hat{R}_0}, \quad (12)$$

where ϵ is defined to be $\Delta\hat{R}_0/\hat{L}$. The following non-dimensional parameters are also introduced

$$\mathbf{v} = \frac{\hat{\mathbf{v}}}{\hat{u}_0}, \quad \boldsymbol{\sigma} = \frac{\hat{\sigma}}{\hat{\sigma}_0}, \quad t = \frac{\hat{u}_0}{\hat{L}}\hat{t}, \quad \hat{\mathbf{D}} = \frac{\hat{L}}{\hat{u}_0}\mathbf{D},$$

$$\mathbf{N} = \hat{\sigma}_0\hat{\mathbf{N}}, \quad \lambda = \frac{\hat{L}}{\hat{\sigma}_0\hat{u}_0}\hat{\lambda}, \quad (13)$$

where $\hat{u}_0$ is the entrance velocity.

In terms of the non-dimensional variables, the equilibrium equations become

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \epsilon\frac{\partial \sigma_{xz}}{\partial z} = 0, \quad (14)$$

$$\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \epsilon\frac{\partial \sigma_{yz}}{\partial z} = 0, \quad (15)$$

and

$$\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} + \epsilon\frac{\partial \sigma_{zz}}{\partial z} = 0. \quad (16)$$

The flow rule is written as

$$\mathbf{D} = \lambda\mathbf{N}, \quad \mathbf{N} = \frac{\partial g}{\partial \boldsymbol{\sigma}} = \frac{\partial g}{\partial I_1}\mathbf{I} + \frac{\partial g}{\partial J_2}\mathbf{s},$$

$$g = g(I_1, J_2, H_\alpha), \quad (17)$$

where $\mathbf{I}$ is the second-order identity tensor, or

$$\frac{\partial u}{\partial x} = \epsilon\lambda N_{xx}, \quad (18)$$

$$\frac{\partial v}{\partial y} = \epsilon\lambda N_{yy}, \quad (19)$$

$$\frac{\partial w}{\partial z} = \lambda N_{zz}, \quad (20)$$

$$\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = 2\epsilon\lambda N_{xy}, \quad (21)$$

$$\frac{\partial w}{\partial x} + \epsilon\frac{\partial u}{\partial z} = 2\epsilon\lambda N_{xz}, \quad (22)$$

$$\frac{\partial w}{\partial y} + \epsilon\frac{\partial v}{\partial z} = 2\epsilon\lambda N_{yz}, \quad (23)$$

where $\mathbf{v} = (u, v, w)$ are the dimensionless velocities in the (x, y, z) directions. The evolution equations of the state variables become

$$\frac{dH_\alpha}{dt} = h_\alpha\ (\mathbf{D}, \sigma, H_\beta) \text{ where}$$

$$h_\alpha(\mathbf{D}, \sigma, H_\beta) = \hat{h}_\alpha(\mathbf{D}, \hat{\sigma}_0, \sigma, H_\beta), \tag{24}$$

and the yield condition is written as

$$\Phi(I_1, J_2, H_\alpha) = 0. \tag{25}$$

Taking into account that $\hat{\phi} = O(\epsilon)$, we can write

$$\mathbf{M} = \cos\hat{\phi}\mathbf{n} + \sin\hat{\phi}\mathbf{k} = \mathbf{n} + \tan\hat{\phi}\mathbf{k} + \mathbf{O}(\epsilon^2), \tag{26}$$

and

$$\mathbf{m} = \cos\hat{\phi}\mathbf{k} - \sin\hat{\phi}\mathbf{n} = \mathbf{k} - \tan\hat{\phi}\mathbf{n} + \mathbf{O}(\epsilon^2), \tag{27}$$

where $\mathbf{n}$ is the unit vector normal to the die profile on the $\hat{x}-\hat{y}$ plane as shown in Fig. 1, and $\mathbf{k}$ is the unit vector along the $\hat{z}$-axis. Using equations (26) and (27) we can write the boundary conditions (9) and (10) as

$$v_n + \epsilon w\tan\phi + O(\epsilon^2) = 0, \tag{28}$$

$$\sigma_{nt} + \epsilon\sigma_{zt}\tan\phi + O(\epsilon^2) = 0, \tag{29}$$

and

$$\sigma_{zn} + \epsilon(\sigma_{zz} - \sigma_{nn})\tan\phi + O(\epsilon^2) = 0, \tag{30}$$

where $\tan\phi = \tan\hat{\phi}/\epsilon = O(1)$ is the corresponding slope on the normalized (x, y, z) space.

2.2 Perturbation Expansion. We now seek a perturbation expansion in ϵ for the solution to the problem, such that

$$\mathbf{v} = \mathbf{v}^{(0)} + \epsilon\mathbf{v}^{(1)} + \ldots, \tag{31}$$

$$\sigma = \sigma^{(0)} + \epsilon\sigma^{(1)} + \ldots, \tag{32}$$

$$\mathbf{H}_\alpha = H_\alpha^{(0)} + \epsilon H_\alpha^{(1)} + \ldots, \tag{33}$$

and

$$\lambda = \lambda^{(0)} + \epsilon\lambda^{(1)} + \ldots. \tag{34}$$

Using equations (31)–(34) we write

$$I_1 = I_1^{(0)} + \epsilon I_1^{(1)} + \ldots, \tag{35}$$

$$J_2 = J_2^{(0)} + \epsilon J_2^{(1)} + \ldots, \tag{36}$$

$$\mathbf{D} = \mathbf{D}^{(0)} + \epsilon\mathbf{D}^{(1)} + \ldots, \tag{37}$$

$$\Phi = \Phi^{(0)} + \epsilon\Phi^{(1)} + \ldots, \tag{38}$$

$$h_\alpha = h_\alpha^{(0)} + \epsilon h_\alpha^{(1)} + \ldots, \tag{39}$$

$$\mathbf{N} = \mathbf{N}^{(0)} + \epsilon\mathbf{N}^{(1)} + \ldots, \tag{40}$$

where

$$I_1^{(k)} = \sigma_{ii}^{(k)},\ J_2^{(0)} = \frac{1}{2}s_{ij}^{(0)}s_{ij}^{(0)},\ J_2^{(1)} = s_{ij}^{(0)}s_{ij}^{(1)},$$

$$D_{ij}^{(k)} = \frac{1}{2}\left(\frac{\partial v_i^{(k)}}{\partial x_j} + \frac{\partial v_j^{(k)}}{\partial x_i}\right), \tag{41}$$

$$\Phi^{(0)} = (\Phi)_0,\ \Phi^{(1)} = \left(\frac{\partial\Phi}{\partial I_1}\right)_0 I_1^{(1)} + \left(\frac{\partial\Phi}{\partial J_2}\right)_0 J_2^{(1)} + \sum_{\alpha=1}^{n}\left(\frac{\partial\Phi}{\partial H_\alpha}\right)_0 H_\alpha^{(1)}, \tag{42}$$

$$h_\alpha^{(0)} = (h_\alpha)_0,\ h_\alpha^{(1)} = \left(\frac{\partial h_\alpha}{\partial\mathbf{D}}\right)_0 :\mathbf{D}^{(1)} + \left(\frac{\partial h_\alpha}{\partial\sigma}\right)_0 :\sigma^{(1)} + \sum_{\beta=1}^{n}\left(\frac{\partial h_\alpha}{\partial H_\beta}\right)_0 H_\beta^{(1)}, \tag{43}$$

$$\mathbf{N}^{(0)} = \left(\frac{\partial g}{\partial\sigma}\right)_0,\ \mathbf{N}^{(1)} = \left(\frac{\partial g}{\partial\sigma}\right)_0 :\sigma^{(1)}. \tag{44}$$

In equations (41)–(44) the scalar product of two second-order tensors is defined by $\mathbf{A}:\mathbf{B} = A_{ij}B_{ji}$, and the notation $(\text{function})_0$ indicates the value of the function at $(\sigma^{(0)}, H_\alpha^{(0)}, \mathbf{D}^{(0)})$.

Substituting the expansions into the governing equations and the boundary conditions, and collecting terms having like powers of ϵ, we obtain the following hierarchy of problems.

For the leading order problem we have

$$\frac{\partial\sigma_{xx}^{(0)}}{\partial x} + \frac{\partial\sigma_{xy}^{(0)}}{\partial y} = 0, \tag{45}$$

$$\frac{\partial\sigma_{xy}^{(0)}}{\partial x} + \frac{\partial\sigma_{yy}^{(0)}}{\partial y} = 0, \tag{46}$$

$$\frac{\partial\sigma_{xz}^{(0)}}{\partial x} + \frac{\partial\sigma_{yz}^{(0)}}{\partial y} = 0, \tag{47}$$

$$\frac{\partial u^{(0)}}{\partial x} = 0, \tag{48}$$

$$\frac{\partial v^{(0)}}{\partial y} = 0, \tag{49}$$

$$\frac{\partial w^0}{\partial z} = \lambda^{(0)}N_{zz}^{(0)}, \tag{50}$$

$$\frac{\partial u^{(0)}}{\partial y} + \frac{\partial v^{(0)}}{\partial x} = 0, \tag{51}$$

$$\frac{\partial w^{(0)}}{\partial x} = 0, \tag{52}$$

$$\frac{\partial w^{(0)}}{\partial y} = 0, \tag{53}$$

$$\frac{dH_\alpha^{(0)}}{dt} = h_\alpha^{(0)}, \tag{54}$$

and

$$\Phi^{(0)} = 0, \tag{55}$$

with boundary conditions on the die-metal interface

$$v_n^{(0)} = 0, \tag{56}$$

$$\sigma_{nt}^{(0)} = 0, \tag{57}$$

and

$$\sigma_{zn}^{(0)} = 0. \tag{58}$$

At $O(\epsilon)$ the problem is given by

$$\frac{\partial\sigma_{xx}^{(1)}}{\partial x} + \frac{\partial\sigma_{xy}^{(1)}}{\partial y} + \frac{\partial\sigma_{xz}^{(0)}}{\partial z} = 0, \tag{59}$$

$$\frac{\partial\sigma_{xy}^{(1)}}{\partial x} + \frac{\partial\sigma_{yy}^{(1)}}{\partial y} + \frac{\partial\sigma_{yz}^{(0)}}{\partial z} = 0, \tag{60}$$

$$\frac{\partial\sigma_{xz}^{(1)}}{\partial x} + \frac{\partial\sigma_{yz}^{(1)}}{\partial y} + \frac{\partial\sigma_{zz}^{(0)}}{\partial z} = 0, \tag{61}$$

$$\frac{\partial u^{(1)}}{\partial x} = \lambda^{(0)}N_{xx}^{(0)}, \tag{62}$$

$$\frac{\partial v^{(1)}}{\partial y} = \lambda^{(0)}N_{yy}^{(0)}, \tag{63}$$

$$\frac{\partial w^{(1)}}{\partial z} = \lambda^{(0)}N_{zz}^{(1)} + \lambda^{(1)}N_{zz}^{(0)}, \tag{64}$$

$$\frac{\partial u^{(1)}}{\partial y} + \frac{\partial v^{(1)}}{\partial x} = 2\lambda^{(0)}N_{xy}^{(0)}, \tag{65}$$

$$\frac{\partial w^{(1)}}{\partial x}+\frac{\partial u^{(0)}}{\partial z}=2\lambda^{(0)}N_{xz}^{(0)}, \tag{66}$$

$$\frac{\partial w^{(1)}}{\partial y}+\frac{\partial v^{(0)}}{\partial z}=2\lambda^{(0)}N_{yz}^{(0)}, \tag{67}$$

$$\frac{dH_\alpha^{(1)}}{dt}=h_\alpha^{(1)}, \tag{68}$$

and

$$\Phi^{(1)}=0, \tag{69}$$

with boundary conditions on the die-metal interface

$$v_n^{(1)}=-w^{(0)}\tan\phi, \tag{70}$$

$$\sigma_{nt}^{(1)}=-\sigma_{zt}^{(0)}\tan\phi, \tag{71}$$

and

$$\sigma_{zn}^{(1)}=-(\sigma_{zz}^{(0)}-\sigma_{nn}^{(0)})\tan\phi. \tag{72}$$

2.3 Solution of the Problem. We start with the leading order problem. Equations (48), (49), (52), and (53) imply that

$$u^{(0)}(x,y,z)=u^{(0)}(0,y,z)=0, \tag{73}$$

$$v^{(0)}(x,y,z)=v^{(0)}(x,0,z)=0, \tag{74}$$

and

$$w^{(0)}=w^{(0)}(z). \tag{75}$$

Equation (75) shows that the leading order approximation involves slab flow as assumed by von Karman (1925) in the analysis of plane-strain rolling. The slab translates down the extrusion axis, experiences a reduction in area, a change in shape, and thickens uniformly. It should be noted, however, that in contrast to the common assumption of uniform stressing and deformation on each cross-section used by the so-called slab methods, the leading order stress and deformation fields are, in general, functions of x and y on each cross-section.

Using the aforementioned results and some equations arising at $O(\epsilon)$, we find that the leading order problem becomes

$$\frac{\partial\sigma_{xx}^{(0)}}{\partial x}+\frac{\partial\sigma_{xy}^{(0)}}{\partial y}=0, \tag{76}$$

$$\frac{\partial\sigma_{xy}^{(0)}}{\partial x}+\frac{\partial\sigma_{yy}^{(0)}}{\partial y}=0, \tag{77}$$

$$\frac{\partial\sigma_{xz}^{(0)}}{\partial x}+\frac{\partial\sigma_{yz}^{(0)}}{\partial y}=0, \tag{78}$$

$$\frac{\partial u^{(1)}}{\partial x}=\lambda^{(0)}N_{xx}^{(0)}, \tag{79}$$

$$\frac{\partial v^{(1)}}{\partial y}=\lambda^{(0)}N_{yy}^{(0)}, \tag{80}$$

$$\frac{dw^{(0)}}{dz}=\lambda^{(0)}N_{zz}^{(0)}, \tag{81}$$

$$\frac{\partial u^{(1)}}{\partial y}+\frac{\partial v^{(1)}}{\partial x}=2\lambda^{(0)}N_{xy}^{(0)}, \tag{82}$$

$$\frac{\partial w^{(1)}}{\partial x}=2\lambda^{(0)}N_{xz}^{(0)}, \tag{83}$$

$$\frac{\partial w^{(1)}}{\partial y}=2\lambda^{(0)}N_{yz}^{(0)}, \tag{84}$$

$$\frac{dH_\alpha^{(0)}}{dt}=h_\alpha^{(0)}, \tag{85}$$

and

$$\Phi^{(0)}=0, \tag{86}$$

with boundary conditions on the die-metal interface

$$v_n^{(1)}=-w^{(0)}\tan\phi, \tag{87}$$

$$\sigma_{nt}^{(0)}=0, \tag{88}$$

and

$$\sigma_{zn}^{(0)}=0. \tag{89}$$

The problem now separates into an antiplane problem at each position z:

$$\frac{\partial\sigma_{xz}^{(0)}}{\partial x}+\frac{\partial\sigma_{yz}^{(0)}}{\partial y}=0, \tag{90}$$

$$\frac{\partial w^{(1)}}{\partial x}=2\lambda^{(0)}N_{xx}^{(0)}=2\lambda^{(0)}\left(\frac{\partial g}{\partial J_2}\right)_0\sigma_{xz}^{(0)}, \tag{91}$$

$$\frac{\partial w^{(1)}}{\partial y}=2\lambda^{(0)}N_{yz}^{(0)}=2\lambda^{(0)}\left(\frac{\partial g}{\partial J_2}\right)_0\sigma_{yz}^{(0)}, \tag{92}$$

with

$$\sigma_{nz}^{(0)}=0 \text{ on the perimeter}, \tag{93}$$

and a generalized plane-strain problem at each position z:

$$\frac{\partial\sigma_{xx}^{(0)}}{\partial x}+\frac{\partial\sigma_{xy}^{(0)}}{\partial y}=0, \tag{94}$$

$$\frac{\partial\sigma_{xy}^{(0)}}{\partial x}+\frac{\partial\sigma_{yy}^{(0)}}{\partial y}=0, \tag{95}$$

$$\frac{\partial u^{(1)}}{\partial x}=\lambda^{(0)}N_{xx}^{(0)}, \tag{96}$$

$$\frac{\partial v^{(1)}}{\partial y}=\lambda^{(0)}N_{yy}^{(0)} \tag{97}$$

$$\frac{\partial u^{(1)}}{\partial y}+\frac{\partial v^{(1)}}{\partial x}=2\lambda^{(0)}N_{xy}^{(0)}, \tag{98}$$

$$N_{zz}^{(0)}=\frac{1}{\lambda^{(0)}}\frac{dw^{(0)}}{dz}, \tag{99}$$

with

$$v_n^{(1)}=-w^{(0)}\tan\phi \text{ and } \sigma_{nt}^{(0)}=0 \text{ on the perimeter}. \tag{100}$$

The two problems are coupled by the common functions $\lambda^{(0)}$ and $H_\alpha^{(0)}$, and the requirements that

$$\frac{dH_\alpha^{(0)}}{dt}=h_\alpha^{(0)}, \text{ and } \Phi^{(0)}=0. \tag{101}$$

Using the principle of the virtual velocities for the antiplane problem, we find that

$$\int_{A(z)}\lambda^{(0)}\left(\frac{\partial g}{\partial J_2}\right)_0(\sigma_{xz}^{(0)2}+\sigma_{yz}^{(0)2})dA=0, \tag{102}$$

where $A(z)$ is the cross-sectional area of the die on the normalized $x-y$ plane. In general, $(\partial g/\partial J_2)_0\neq 0$, and since the whole cross-section is flowing, we infer that $\lambda^{(0)}\neq 0$, so

$$\sigma_{xz}^{(0)}=\sigma_{yz}^{(0)}=0$$

and the generalized plane-strain problem remains. Equations (91) and (92) now imply that

$$w^{(1)}=w^{(1)}(z). \tag{104}$$

The previous analysis shows that the determination of the leading order stresses of the three-dimensional problem reduces now to a two-dimensional calculation. In the following we discuss the solution of the remaining generalized plane-strain problem for the cases of incompressible ($\partial g/\partial I_1=0$) and compressible materials ($\partial g/\partial I_1\neq 0$).

2.4 Incompressible Materials. For the case of an incompressible material, $w^{(0)}(z)$ can be easily calculated by means of mass conservation. The original cross-section is distorted in a manner free of shear on the perimeter through the series of shapes along z involved in the extrusion die while the cross-section thickens uniformly in such a way that $A(z)h(z) =$ constant, where h is the thickness which is chosen arbitrarily at the outset.

At each position, z, the solution of the generalized plane-strain problem determines, to within a hydrostatic stress $p^{(0)}(z)$, the leading order stress components $\sigma_{xx}^{(0)}$, $\sigma_{yy}^{(0)}$, $\sigma_{zz}^{(0)}$ and $\sigma_{xy}^{(0)}$, as well as $u^{(1)}$ and $v^{(1)}$. Let $\Sigma_{xx}^{(0)}$, $\Sigma_{yy}^{(0)}$, $\Sigma_{zz}^{(0)}$, and $\Sigma_{xy}^{(0)}$ be a solution to the generalized plane-strain problem. Then, the actual leading order solution is

$$\sigma_{xx}^{(0)} = \Sigma_{xx}^{(0)} + p^{(0)}(z), \tag{105}$$

$$\sigma_{yy}^{(0)} = \Sigma_{yy}^{(0)} + p^{(0)}(z), \tag{106}$$

$$\sigma_{zz}^{(0)} = \Sigma_{zz}^{(0)} + p^{(0)}(z), \tag{107}$$

and

$$\sigma_{xy}^{(0)} = \Sigma_{xy}^{(0)}. \tag{108}$$

The function $p^{(0)}(z)$ can be determined as described in the following. Using the $O(\epsilon)$ equilibrium equation (61), we find that

$$\int_{A(z)} \frac{\partial \sigma_{zz}^{(0)}}{\partial z} dA = -\int_{A(z)} \left(\frac{\partial \sigma_{zx}^{(1)}}{\partial x} + \frac{\partial \sigma_{zy}^{(1)}}{\partial y} \right) dA. \tag{109}$$

Using Gauss' theorem and taking into account the boundary condition (72), we can write equation (109) as

$$\int_{A(z)} \frac{\partial \sigma_{zz}^{(0)}}{\partial z} dA = \int_{\Gamma(z)} (\Sigma_{zz}^{(0)} - \Sigma_{nn}^{(0)}) \tan\phi ds, \tag{110}$$

where $\Gamma(z)$ is the perimeter of the cross-section of the die and ds indicates infinitesimal arc length on $\Gamma(z)$. Finally, substituting equation (107) into equation (110) we find that

$$\frac{dp^{(0)}}{dz} = f(z), \tag{111}$$

where

$$f(z) = \frac{1}{A(z)} \left[\int_{\Gamma(z)} (\Sigma_{zz}^{(0)} - \Sigma_{nn}^{(0)}) \tan\phi ds - \int_{A(z)} \frac{\partial \Sigma_{zz}^{(0)}}{\partial z} dA \right] \tag{112}$$

is known from the solution of the generalized plane-strain problem. Integration of equation (111) yields

$$p^{(0)}(z) = \int_0^z f(z) dz + c, \tag{113}$$

where c is a constant. Since the net force at the exit of the die in extrusion is equal to zero, we have ($z = 1$ at the exit in normalized coordinates)

$$\int_{A(1)} \sigma_{zz}^{(0)} dA = 0 \tag{114}$$

which implies that

$$c = -\frac{1}{A(1)} \int_{A(1)} \Sigma_{zz}^{(0)} dA - \int_0^1 f(z) dz. \tag{115}$$

Finally, substituting the value of c into (113) we find

$$p^{(0)}(z) = -\frac{1}{A(1)} \int_{A(1)} \Sigma_{zz}^{(0)} dA - \int_z^1 f(z) dz, \tag{116}$$

and this completes the solution.

2.5 Compressible Materials. For the case of a compressible material, $w^{(0)}(z)$ cannot be determined from mass conservation considerations alone, and its calculation becomes part of the solution of the generalized plane-strain problem. As discussed in the previous section, the original cross-section is distorted in a manner free of shear on the perimeter through a series of shapes along z involved in the extrusion die while the cross-section thickens uniformly. At each position z, an additional boundary condition is needed for completeness, and this is the specification of the axial force

$$P^{(0)}(z) = \int_{A(z)} \sigma_{zz}^{(0)} dA, \tag{117}$$

which is determined as described in the following.

Using the definition of $P^{(0)}(z)$ and taking into account that

$$\int_{A(z)} \frac{\partial \sigma_{zz}^{(0)}}{\partial z} dA = \int_{\Gamma(z)} (\sigma_{zz}^{(0)} - \sigma_{nn}^{(0)}) \tan\phi ds, \tag{118}$$

we can easily show that

$$\frac{dP^{(0)}}{dz} = q(z), \tag{119}$$

where

$$q(z) = \int_{\Gamma(z)} (\sigma_{zz}^{(0)} - \sigma_{nn}^{(0)}) \tan\phi ds + \int_{A(s)} \frac{\sigma_{zz}^{(0)}}{J} \frac{\partial J}{\partial z} dA. \tag{120}$$

In equation (120) J is the Jacobian of the transformation that carries a material point from position (X, Y) at $z = 0$ to (x, y) at each z.

Integration of equation (119) yields

$$P^{(0)}(z) = P_0^{(0)} + \int_0^z q(z) dz, \tag{121}$$

where $P_0^{(0)}$ is the leading-order extrusion force.

With $P^{(0)}(z)$ defined by equation (121) the sequence of generalized plane-strain problems is solved, and the solution is determined in terms of the as yet unknown $P_0^{(0)}$. Finally, the extrusion force $P_0^{(0)}$ is determined by enforcing the condition of zero axial force at the die exit, i.e.,

$$\int_{A(1)} \sigma_{zz}^{(0)} dA = 0, \tag{122}$$

and this completes the solution.

The results of the asymptotic analysis can be used together with two-dimensional finite element calculations to obtain the leading order solution of three-dimensional extrusion problems. The proposed new approach for the analysis of three-dimensional extrusions is particularly attractive because the lengthy three-dimensional calculations are now avoided and the accuracy of the obtained approximate solution can be easily estimated.

Several applications of the proposed method are presented in the following.

3 Applications

In the examples presented in this section we consider a rigid perfectly-plastic incompressible material which yields according to the von Mises criterion and obeys the Lévy-Mises equations, i.e.,

$$\Phi = g = 3J_2 - 1.$$

3.1 Axisymmetric Extrusion. A simple special case of the three-dimensional problem discussed in the previous section is the extrusion of circular rods through slowly converging axisymmetric curved dies. This problem has been solved by

slab methods, but here we show that such solutions are correct to leading order. As mentioned in Section 2, the slope of the die is small when either the area reduction is small or the length of the reduction region is large compared to the radius of the original cross-section. In such cases an appropriate definition of the parameter ϵ is $\epsilon = \Delta\hat{R}/\hat{L}$, where $\Delta\hat{R}$ is the total radius reduction.

As discussed in Section 2, the leading order stresses can be determined from the solution of a generalized plane-strain problem over the cross-section of the billet at each position z with boundary conditions

$$v_r^{(1)} = -w^{(0)}\tan\phi \text{ and } \sigma_{r\theta}^{(0)} = 0 \text{ on } r = R$$

where (r,θ) are normalized polar coordinates, $R(z)$ is the normalized radius of the die,

$$\tan\phi = -\frac{dR(z)}{dz},$$

and

$$w^{(0)} = \frac{R^2(0)}{R^2(z)}.$$

The solution of the generalized plane-strain problem is discussed in the following.

Because of axial symmetry, there is no θ-dependence and the only nonzero velocity and stress components are $v_r^{(1)}$, $w^{(0)}$, and $\Sigma_{rr}^{(0)}$, $\Sigma_{\theta\theta}^{(0)}$, $\Sigma_{zz}^{(0)}$. Using the incompressibility condition

$$\frac{\partial v_r^{(1)}}{\partial r} + \frac{v_r^{(1)}}{r} + \frac{dw^{(0)}}{dz} = 0,$$

we find that

$$v_r^{(1)} = r\frac{R^2(0)}{R^3(z)}\frac{dR(z)}{dz}.$$

The flow rule, together with the yield condition, imply that the solution of the generalized strain problem to within a constant pressure is given by

$$\Sigma_{rr}^{(0)} = \Sigma_{\theta\theta}^{(0)} = -\frac{1}{3} \text{ and } \Sigma_{zz}^{(0)} = -\frac{2}{3}.$$

The function $f(z)$ defined by equation (112) is now found to be

$$f(z) = -\frac{2}{R(z)}\frac{dR(z)}{dz},$$

and the additional pressure $p^{(0)}(z)$ is given by

$$p^{(0)}(z) = -\frac{2}{3} - 2\ln\frac{R(z)}{R(1)}.$$

Finally, the leading order stresses are

$$\sigma_{rr}^{(0)} = \sigma_{\theta\theta}^{(0)} = -1 - 2\ln\frac{R(z)}{R(1)},$$

and

$$\sigma_{zz}^{(0)} = -2\ln\frac{R(z)}{R(1)}.$$

It is interesting to note that leading order stresses are constant over each cross-section and coincide with the results of the so-called slab method. The analysis shows that such results are first-order accurate and that calculations based on slab methods are suitable for slowly varying axisymmetric dies with zero friction between the die and the workpiece.

3.2 Round-to-Ellipse Extrusion. An approximate solution for extrusion of elliptic shapes from round billets through gradually tapering dies is presented in this section. The parameter ϵ can now be defined as $\epsilon = (\hat{R}_0 - \hat{b})/\hat{L}$, where $\hat{R}_0$ is the radius of the original cross-section, and $\hat{b}$ is the length of the minor semiaxis of the final elliptic cross-section.

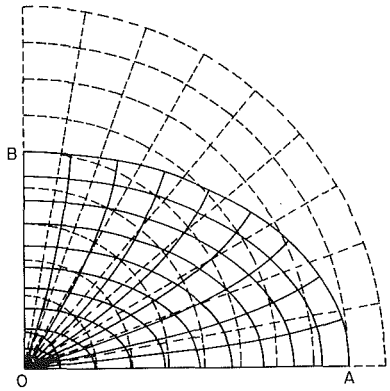


Fig. 2 Deformed finite element mesh superposed on the undeformed mesh (dashed lines)

The shape of the die on the normalized (x, y, z) space is given by

$$F(x,y,z) = \left[\frac{x}{a(z)}\right]^2 + \left[\frac{y}{b(z)}\right]^2 - 1 = 0,$$
$$a(z) = R_0 - z\Delta a,$$
$$b(z) = R_0 - z\Delta b,$$

where $R_0 - \Delta a$ and $R_0 - \Delta b$ are the lengths of the semiaxes of the elliptic cross-section at the end of the die.

The leading order stresses for this problem are obtained by solving a sequence of generalized plane-strain problems over the cross-section of the die with boundary conditions

$$v_n^{(0)} = -w^{(0)}\tan\phi \text{ and } \sigma_{nt}^{(0)} = 0, \tag{123}$$

where

$$\tan\phi = \left(\frac{x^2}{a^3}\Delta a + \frac{y^3}{b^3}\Delta b\right)\left(\frac{x^2}{a^4} + \frac{y^2}{b^4}\right)^{-1/2},$$

and

$$w^{(0)} = \frac{R_0^2}{a(z)b(z)}.$$

The solution of the generalized plane-strain problem is obtained by using the finite element method.The ABAQUS (Hibbit, 1984) general purpose finite element program is used for the computations. Because of symmetry we need to analyze only one quarter of the cross-section. Eight-node generalized plane-strain isoparametric elements with 2×2 Gauss integration are used. An elastic-perfectly-plastic model is used in the calculations; the effects of elasticity are not important in this case and the obtained solution is a very close approximation to the solution of the rigid-perfectly-plastic problem. The analysis is done incrementally using an updated-Lagrangian formulation and Newton's method is used to solve the overall discretized equilibrium equations.

The results presented in this section are for $\Delta a/R_0 = 0.1$ and $\Delta b/R_0 = 0.4$. Figure 2 shows the deformed finite element mesh at the end of the calculation superposed on the undeformed one. The analysis was completed in 80 equal increments. ABAQUS provides a general interface so that the user may introduce his own "multi-point-constraints" in a "user subroutine." The displacements of nodes A and B in Fig. 2 are prescribed and the user subroutine is used to constrain the degrees-of-freedom of the boundary nodes in such a way that they always remain on the ellipse defined by the position of points A and B. One constraint equation per boundary node is used and this is equivalent to imposing the boundary conditions (123) in a discretized manner. The finite element solution provides the stress field $\Sigma^{(0)}$; the function $f(z)$ and the additional hydrostatic stress $p^{(0)}(z)$ are obtained numerically using equations (112) and (116). Finally, the complete leading

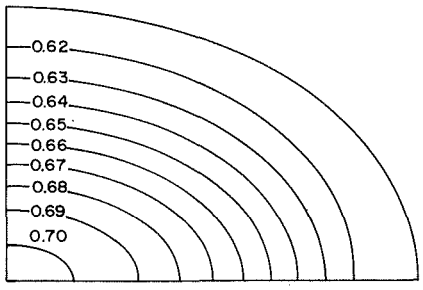


Fig. 3 Contours of the equivalent plastic strain $\bar{\epsilon}^p$ at the end of the calculations

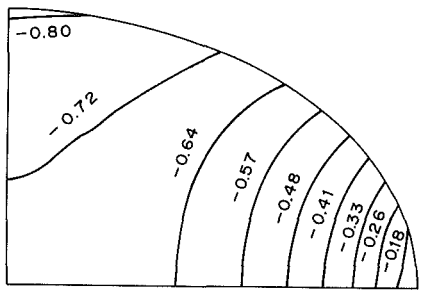


Fig. 4 Contours of the normalized hydrostatic stress component $\sigma_{kk}^{(0)}/3$ at the end of the calculation

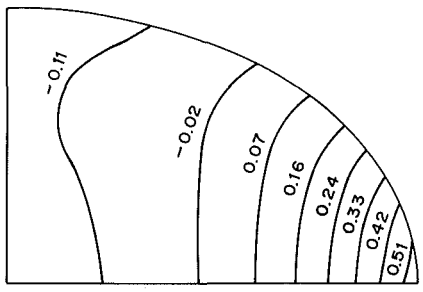


Fig. 5 Contours of the normalized axial stress component $\sigma_{zz}^{(0)}$ at the end of the calculation

order stress field $\sigma^{(0)}$ is obtained by adding $p^{(0)}(z)$ to the finite element solution $\Sigma^{(0)}$.

Figure 3 shows contours of the equivalent plastic strain $\bar{\epsilon}^p$ at the end of the calculations; the equivalent plastic strain is defined as

$$\bar{\epsilon}^p = \int_0^t \left(\frac{2}{3} D_{ij}^p D_{ij}^p\right)^{1/2} dt,$$

where $\mathbf{D}^p$ is the plastic part of the deformation rate. Figure 4 shows contours of the hydrostatic stress component $\sigma_{kk}^{(0)}/3$ at the exit of the die. It is clear from Fig. 4 that the hydrostatic stress component is compressive everywhere on the end cross-section; an examination of the complete solution reveals that it always remains so in the process of deformation. Contours of the axial stress component $\sigma_{zz}^{(0)}$ at the exit of the die are shown in Fig. 5; $\sigma_{zz}^{(0)}$ is compressive over most of the cross-section and it changes to tensile near the end of the longer semiaxis so that the net axial force over the cross-section at the exit of the die vanishes. The extrusion pressure is found to be 1.20 times the yield stress.

3.3 Round-to-Square Extrusion. An approximate solution for extrusion of square shapes from round billets through slowly varying dies is presented in this section. The cylindrical billet deforms to the final square section rod through the die defined by an envelop of a number of straight lines. The

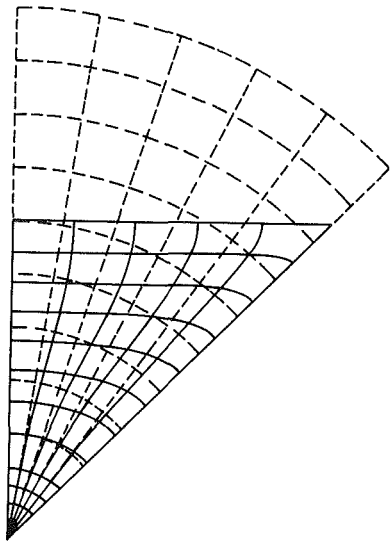


Fig. 6 Deformed finite element mesh superposed on the underformed mesh (dashed lines)

parameter ϵ can be defined as $\epsilon = (\hat{R}_0 - \hat{a})/\hat{L}$, where $\hat{R}_0$ is the radius of the original cross-section, and $2\hat{a}$ is the length of the sides of the final square cross-section.

The shape of the die on the normalized (x, y, z) space is given in this case by

$$F(x,y,z) = y - (1-z)\frac{R_0 y}{r} - az = 0,$$

where $r = (x^2+y^2)^{1/2}$, R_0 is the normalized radius of the original cross-section, and $2a$ is the length of the sides of the final square cross-section on the normalized $x-y$ plane.

The leading order stresses for this problem are obtained by solving a sequence of generalized plane-strain problems over the cross-section of the die with boundary conditions

$$v_n^{(0)} = -w^{(0)}\tan\phi \text{ and } \sigma_{nt}^{(0)} = 0,$$

where

$$\tan\phi = \left(\frac{R_0 y}{r} - a\right)\left[\left[(1-z)\frac{R_0 xy}{r^3}\right]^2 + \left[1-(1-z)\frac{R_0}{r}\left(1-\frac{y^2}{r^2}\right)\right]^2\right]^{-1/2},$$

and

$$w^{(0)} = \frac{\pi R_0^2}{A(z)}.$$

The solution of the generalized plane-strain problem is obtained in a way similar to that described in Section 3.2 by using the finite element method. Because of symmetry we need to analyze only one-eighth of the cross-section. The results presented in this section are for $a/R_0 = 0.6$.

Figure 6 shows the deformed finite element mesh at the end of the calculations superposed on the undeformed one. The analysis was completed in 80 equal increments. Contours of the equivalent plastic strain $\bar{\epsilon}^p$ are shown in Fig. 7. A strain concentration near the corners of the final square cross-section is evident. Figure 8 shows contours of the hydrostatic stress component $\sigma_{kk}^{(0)}/3$ at the exit of the die. The extrusion pressure is found to be 0.82 times the yield stress.

Near the corners of the square, tensile tractions are induced across the metal-die interface, and this is the reason for the appearance of tensile hydrostatic stresses in that region (see Fig. 8). This indicates that separation would occur and that the traction would vanish. In our calculations the boundary nodes were forced to stay on the die surface, but since tensile trac-

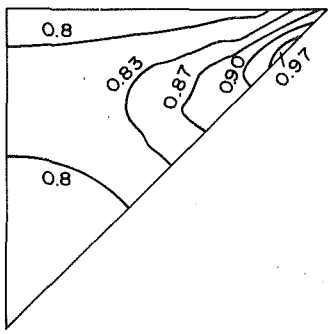


Fig. 7 Contours of the equivalent plastic strain $\bar{\epsilon}^p$ at the end of the calculations

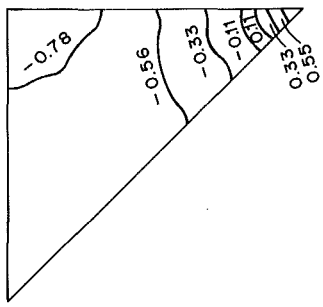


Fig. 8 Contours of the normalized hydrostatic stress component $\sigma_{kk}^{(0)}/3$ at the end of the calculation

tions appear only in a very small region near the corner, this is not expected to affect the rest of the solution significantly. The effects of such separation and any subsequent contact will be addressed in a future publication.

4 Closure

A new method of analysis of three-dimensional extrusion has been presented in this paper. Using asymptotic techniques we have shown that the determination of the leading order stresses of the three-dimensional problem reduces to a two-dimensional calculation. The constitutive model considered covers a wide range of plasticity models and the proposed method of analysis can be used to study the development of internal damage in the billet as it moves through the die. The asymptotic solution shows that the leading order approximation involves slab flow, and this is consequence of the plastic potential being a function only of the first and second invariants of the stress tensor and the frictionless conditions at the die-metal interface.

It should be noted that the proposed asymptotic method can be used for the analysis of three-dimensional metal-forming operations other than extrusion and that anisotropic plasticity models as well as friction at the die surface can be allowed for, but at the expense of a more complicated two-dimensional leading order problem requiring a different method of solution (Johnson, 1988). Frictional effects are certainly important and will be addressed in a future publication. However, in many cold-forming operations lubrication is used and the effective coefficient of friction is small (say, of order ϵ) and, therefore, a nonzero friction on the metal-die interface affects the problems of order ϵ or higher but leaves the leading order problem unchanged.

Acknowledgments

NA acknowledges the support of the National Science Foundation through a Presidential Young Investigator Award (NSF Grant No. MSM-8657860) and the support of ALCOA Foundation through a Science Support Grant. RMM acknowledges the support of the Defense Advanced Research Projects Agency through University Research Initiative Contract No. N00014-86-K-0753 at UCSB. The authors are also grateful to Hibbitt, Karlsson and Sorensen, Inc. for provision of the ABAQUS general purpose finite element program.

References

Basily, B. B., and Sansome, D. H., 1976, "Some Theoretical Considerations for the Direct Drawing of Section Rod From Round Bar," *Int. J. Mech. Sci.*, Vol. 18, pp. 201–208.

Boër, C. R., Schneider, W. R., Eliasson, B., and Avitzur, B., 1979, "An Upper-Bound Approach for the Direct Drawing of Square Section Rod from Round Bar," *20th In. Mech. Tool Des Res. Conf.*, London, England, pp. 149–156.

Boër, C. R., and Webster, Jr., W. D., 1985, "Direct Upper-Bound Solution and Finite Element Approach to Round-to-Square Drawing," ASME *Journal of Engineering for Industry*, Vol. 107, pp. 254–260.

Cho, N. S., and Yang, D. Y., 1983, "Analysis of Hydrofilm Extrusion of Elliptic Shapes Using Perturbation Method," *Int. J. Mech. Sci.*, Vol. 25, pp. 293–292.

Gunasekera, J. S., and Hoshino, S., 1982, "Analysis of Extrusion of Drawing of Polygonal Sections Through Straightly Converging Dies," ASME *Journal of Engineering for Industry*, Vol. 104, pp. 38–45.

Gunasekera, J. S., and Hoshino, 1985, "Analysis of Extrusion of Polygonal Sections Through Streamlined Dies," *ASME Journal of Engineering for Industry*, Vol. 107, pp. 229–233.

Han, C. H., Yang, D. Y., and Kiuchi, M., 1986, "A New Formulation for Three-Dimensional Extrusion and Its Application to Extrusion of Clover Sections," *Int. J. Mech. Sci.*, Vol. 28, pp. 201–218.

Hibbitt, H. D., 1984, "ABAQUS/EPGEN – A General Purpose Finite Element Coded With Emphasis on Nonlinear Applications," *Nucl. Eng. Des.*, Vol. 77, pp. 271–297.

Hoshino, S., and Gunasekera, J. S., 1980, "An Upper-Bound Solution for the Extrusion of Square Section From Round Bar Through Converging Dies," *Proc. 21st Mach. Tool Des. Res.*, pp. 97–105.

Johnson, R. E., 1987, "Conical Extrusion of a Work-Hardening Material: An Asymptotic Analysis," *J. Engng. Math.*, Vol. 21, pp. 295–329.

Juneja, B. L., and Prakash, R., 1975, "An Analysis for Drawing and Extrusion of Polygonal Sections," *Int. J. Mach. Tool Des. Res.*, Vol. 15, pp. 1–30.

von, Kármán, T., 1925, "Beitrag zur Theorie des Walzvorganges," *Z. angew. Math. Mech.*, Vol. 5, pp. 139–141.

Kiuchi, M., 1984, "Overall Analysis of Nonaxisymmetric Extrusion and Drawing," *Proc. 12th N. Am. Met. Res. Conf.*, Houghton, Mifflin, Mich. pp. 111–119.

Kiuchi, M., and Ishikawa, M., 1983, "Upper-Bound Analysis of Extrusion and/or Drawing of L-, T-, and H-Sections – Study on Nonsymmetric Extrusion and Drawing II," *Plast. Forming JSTP*, Vol. 24, pp. 721–729.

Kiuchi, M., Kishi, H., and Ishikawa, 1983, "Upper-Bound Analysis of Extrusion and/or Drawing of Square, Rectangular, Hexagonal, and Other Axisymmetric Bars and Wires – Study on Nonsymmetric Extrusion and Drawing I," *Plast. Forming JSTP*, Vol. 24, pp. 290–296.

Nagpal, V., 1977, "On the Solution of Three-Dimensional Metal-Forming Processes," ASME *Journal of Engineering for Industry*, Vol. 99, pp. 624–629.

Nagpal, V., and Altan, T., 1975, "Analysis of the Three-Dimensional Metal Flow in Extrusion of Shapes With the Use of Dual Stream Functions," *Proc. Third N. Am. Met. Res. Conf.*, Pittsburgh, Penn., pp. 26–40.

Onat, E. T., 1954, "On the Construction of Linearized Axially-Symmetric Plastic Stress and Velocity Fields," Brown University Report DA 2598/14, August 1954.

Prakash, R., and Khan, O. H., 1979, "An Analysis of Plastic Flow Through Polygonal Converging Dies With Generalized Boundaries of Zone of Plastic Deformation," *Int. J. Mach. Tool Des.*, Vol. 19, pp. 1–19.

Smet, R. P., and Johnson, R. E., 1989, "An Asymptotic Analysis of Cold Sheet Rolling," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 33–39.

Van Dyke, M., 1975, *Perturbation Methods in Fluid Mechanics*, The Parabolic Press, Stanford.

Yang, D. Y., Han, C. H., and Kim, M. U., 1986, "A Generalized Method for Analysis of Three-Dimensional Extrusion of Arbitrarily-Shaped Sections," *Int. J. Mech. Sci.*, Vol. 28, pp. 517–534.

Yang, D. Y., Kim, M. U., and Lee, C. H., 1979, "A New Approach for Generalized Three-Dimensional Extrusion of Sections From Round Billets by Conformal Transformation," *IUTAM Symposium Metal Forming Plasticity*, Tutzig, W. Germany, pp. 204–221.

Yang, D. Y., and Lange, K., 1984, "Analysis of Hydrofilm Extrusion of Three-Dimensional Shapes from Round Billets," *Int. J. Mech. Sci.*, Vol. 26, pp. 1–19.

Yang, D. Y., and Lee, C. H., 1978, "Analysis of Three-Dimensional Extrusion of Section Through Curved Dies by Conformal Transformation," *Int. J. Mech. Sci.*, Vol. 20, pp. 541–552.

Adiabatic Shear Banding in Plane Strain Problems

R. C. Batra
Mem. ASME

De-Shin Liu

Department of Mechanical and Aerospace Engineering and Engineering Mechanics, University of Missouri-Rolla, Rolla, Mo. 65401-0249

Plane strain thermomechanical deformations of a viscoplastic body are studied with the objective of analyzing the localization of deformation into narrow bands of intense straining. Two different loadings, namely, the top and bottom surfaces subjected to a prescribed tangential velocity, and these two surfaces subjected to a preassigned normal velocity, are considered. In each case a material defect, flaw, or inhomogeneity is modeled by introducing a temperature bump at the center of the specimen. The solution of the initial boundary value problem by the Galerkin-Adams method reveals that the deformation eventually localizes into a narrow band aligned along the direction of the maximum shearing strain. For both problems, bands of intense shearing appear to diffuse out from the center of the specimen.

1 Introduction

Adiabatic shear banding is the name given to a localization phenomenon that occurs during high-rate plastic deformation such as machining, explosive forming, shock-impact loading, ballistic penetration, fragmentation, ore crushing, impact tooling failure, and metal shaping and forming processes. The localization of the deformation has been observed in steels, nonferrous metals, and polymers. Practical interest in the phenomenon derives from the fact that progressive shearing on an intense shear band provides an undesirable mode of material resistance to imposed deformations, and the bands are often precursors to shear fractures. Of the many processes just stated in which adiabatic shear bands have been found to occur, flat sheet rolling and certain forging operations can be modeled as plane strain operations.

Since the time Zener and Hollomon (1944) recognized the destabilizing effect of thermal softening in reducing the slope of the stress-strain curve in nearly adiabatic deformations, there have been numerous studies aimed at delineating material parameters that enhance or retard the initiation and growth of adiabatic shear bands. Most of the effort has been concentrated in analyzing the simple shearing problem. Clifton (1980) and Bai (1981) studied the growth of infinitesimal periodic perturbations superimposed on a body deformed by a finite amount in simple shear. Burns (1985) used a dual asymptotic expansion to account for the time dependence of the homogeneous solution in the analysis of the growth of superimposed periodic perturbations. Merzer (1982) used the constitutive relation proposed by Bodner and Partom (1975) to study the problem of twisting of a thin tubular specimen having a notch in its periphery. He concluded that the band width depends upon the thermal conductivity. Wu and Freund (1984) used a different material model and studied wave propagation in an infinite medium. They concluded that the thermal conductivity has essentially no effect on the width of a shear band. Other works analyzing the initiation and growth of adiabatic shear bands include those due to Clifton et al. (1984), Wright and Batra (1985), Wright and Walter (1987), Batra (1987), and Batra and Kim, (1989). Rogers (1979, 1983) and Timothy (1987) have reviewed various aspects of adiabatic shear banding, especially from a materials point of view.

Experimental studies dealing with adiabatic shear banding include those of Zener and Hollomon (1944), Moss (1981), Costin et al. (1979), Lindholm and Johnson (1983), and Marchand and Duffy (1988). Marchand and Duffy have given a detailed history of the temperature and strain fields within a band.

Needleman (1989) has recently studied the initiation and growth of shear bands in plane strain deformations of viscoplastic materials. He studied a purely mechanical problem and approximated the effect of thermal softening by assuming that the stress-strain curve has a peak in it. He modeled a material inhomogeneity by assuming that the flow stress for a small amount of material near the center of the block was less than that of the surrounding material. We study herein the thermomechanical plane strain deformations of a thermally softening viscoplastic solid and model the material inhomogeneity by introducing a temperature bump at the center of the block. The block boundaries are assumed to be perfectly insulated. Two different deformation states, namely, that of a simple shearing of the block, and the block deformed in simple compression are analyzed. In each case a shear band develops along the direction of maximum shearing strain. Whereas the deformation localizes at an average compressive strain of 0.059 when the block is deformed in compression, the average shear strain equals 0.227 when the block is deformed in simple shear.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 29, 1988; final revision, November 18, 1988.

Paper No. 89-WA/APM-16.

2 Formulation of the Problem

We use a fixed set of rectangular Cartesian coordinate axes to describe the thermomechanical deformations of the body. In terms of the referential description the governing equations are

$$(\rho J)^{\cdot} = 0, \tag{1}$$

$$\rho_0 \dot{v}_i = T_{i\alpha,\alpha}, \tag{2}$$

$$\rho_0 \dot{e} = -Q_{\alpha,\alpha} + T_{i\alpha} v_{i,\alpha}, \tag{3}$$

and a suitable set of initial and boundary conditions. Equation (1) expresses the balance of mass, (2) the balance of linear momentum, and (3) the balance of internal energy. In these equations, ρ is the current mass density, ρ_0 the mass density in the reference configuration, J is the determinant of the deformation gradient, v_i the velocity of a material particle in the x_i direction, Q_α the heat flux, e the specific internal energy, $T_{i\alpha}$ the first Piola-Kirchoff stress tensor, a superimposed dot stands for the material time derivative, and a comma followed by an index $\alpha(j)$ implies partial differentiation with respect to $X_\alpha(x_j)$. Also $\mathbf{x}$ denotes the present position of a material particle that occupied the place $\mathbf{X}$ in the reference configuration, and a repeated index implies summation over the range of the index. For plane strain deformations, $x_3 = X_3$ and the indices i and α take on values 1 and 2.

For the constitutive relations we take

$$\boldsymbol{\sigma} = -p(\rho)\mathbf{1} + 2\mu\mathbf{D}, \; T_{i\alpha} \equiv \frac{\rho_0}{\rho} X_{\alpha,j}\sigma_{ij}, \tag{4}$$

$$2\mu = \frac{\sigma_0}{\sqrt{3}I}(1-\nu\theta)(1+bI)^m, \; 2D_{ij} = v_{i,j} + v_{j,i}, \tag{5}$$

$$I^2 = \frac{1}{2}\mathrm{tr}\tilde{\mathbf{D}}^2, \; \tilde{\mathbf{D}} = \mathbf{D} - \frac{1}{3}(\mathrm{tr}\mathbf{D})\mathbf{1}, \tag{6}$$

$$p(\rho) = B\left(\frac{\rho}{\rho_0} - 1\right), \tag{7}$$

$$Q_\alpha = -k\theta_{,\alpha}, \tag{8}$$

$$\dot{e} = c\dot{\theta} + \dot{\rho}p(\rho)/(\rho\rho_0). \tag{9}$$

Here, σ_0 is the yield stress in simple tension or compression, ν is the coefficient of thermal softening, parameters b and m represent the strain rate sensitivity of the material, B may be thought of as the bulk modulus, k is the thermal conductivity, and c the specific heat. Equation (7) is a part of the Tillotson (1962) equation wherein the dependence of the pressure upon the changes in temperature has not been considered, and equation (8) is the Fourier law of heat conduction.

Defining $\mathbf{s}$ as

$$\mathbf{s} = \boldsymbol{\sigma} + \left(p - \frac{2\mu}{3}\mathrm{tr}\mathbf{D}\right)\mathbf{1}, \tag{10}$$

$$= 2\mu\tilde{\mathbf{D}}, \tag{11}$$

equations (4) and (5) give

$$\left(\frac{1}{2}\mathrm{trs}^2\right)^{1/2} = \frac{\sigma_0}{\sqrt{3}}(1-\nu\theta)(1+bI)^m, \tag{12}$$

which can be viewed as a generalized von Mises yield surface when the flow stress (given by the right-hand side of (12)) at a material particle depends upon its strain rate and temperature. The linear dependence of the flow stress upon the temperature change has been observed by Bell (1968), Lindholm and Johnson (1983), and Lin and Wagoner (1986). A constitutive relation similar to equation (4) has been used by Zienkiewicz et al. (1981) in analyzing the extrusion problem, by Batra (1988) in studying the steady-state penetration of a viscoplastic target by a rigid cylindrical penetrator, and by Batra and Lin (1989) in studying the steady-state axisymmetric deformations of a cylindrical viscoplastic rod upset at the bottom of a hemispherical rigid cavity. Equation (4) may be interpreted as a constitutive relation for a non-Newtonian fluid whose viscosity μ depends upon the strain rate and temperature.

We introduce nondimensional variables as follows:

$$\bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma}/\sigma_0, \; \bar{p} = p/\sigma_0, \bar{\mathbf{s}} = \mathbf{s}/\sigma_0, \bar{\mathbf{v}} = \mathbf{v}/v_0, \; \bar{t} = tv_0/H, \; \bar{\mathbf{T}} = \mathbf{T}/\sigma_0,$$
$$\bar{\mathbf{x}} = \mathbf{x}/H, \; \bar{\theta} = \theta/\theta_0, \; \bar{b} = b\frac{v_0}{H}, \; \bar{\nu} = \nu\theta, \; \bar{\rho} = \rho/\rho_0, \; \bar{\mathbf{X}} = \mathbf{X}/H,$$
$$\delta = \rho_0 v_0^2/\sigma_0, \; \beta = k/(\rho_0 c v_0 H), \; \theta_0 = \sigma_0/(\rho_0 c), \; \bar{B} = B/\sigma_0. \tag{13}$$

Here, $2H$ is the height of the block, v_0 is the imposed velocity on the top and bottom surfaces, and ρ_0 is the mass density in the unstressed reference configuration. Substituting from equations (4) through (9) into the balance laws (1) through (3), rewriting these in terms of nondimensional variables, denoting the partial differentiation with respect to $\bar{x}_i(\bar{X}_\alpha)$ by a comma followed by an index $i(\alpha)$, material differentiation with respect to t by a superimposed dot, and dropping the superimposed bars, we arrive at the following set of equations:

$$\dot{\rho} + \rho v_{i,i} = 0, \tag{14}$$

$$\delta\dot{v}_i = T_{i\alpha,\alpha}, \tag{15}$$

$$\dot{\theta} = \beta\theta_{,\alpha\alpha} + [1/(\sqrt{3}I\rho)](1-\nu\theta)(1+bI)^m D_{ij}D_{ij}, \tag{16}$$

$$\boldsymbol{\sigma} = -B(\rho-1)\mathbf{1} + \frac{1}{\sqrt{3}I}(1+bI)^m(1-\nu\theta)\mathbf{D}. \tag{17}$$

It is simpler to state boundary conditions for the specific problem studied. We analyze plane strain thermomechanical deformations of an initially-square block of dimension $2H \times 2H$. The $X_1 - X_2$ plane, with the origin of the coordinate system located at the center of the block, is taken as the plane of deformation. For the simple shearing problem the boundary conditions are taken to be

$$v_1 = \pm f(t), \; v_2 = 0, \; Q_\alpha N_\alpha = 0 \text{ at } X_2 = \pm H, \tag{18}$$

$$n_i T_{i\alpha} N_\alpha = 0, \; e_i T_{i\alpha} N_\alpha = h(t), \; Q_\alpha N_\alpha = 0 \text{ at } X_1 = \pm H, \tag{19}$$

where $\mathbf{n}$ is a unit outward normal and $\mathbf{e}$ is a unit vector tangent to the surface in the present configuration and $\mathbf{N}$ is a unit outward normal in the reference configuration. Equations (18) and (19) imply that the boundaries of the block are perfectly insulated, the top and bottom faces are placed in a hard loading device and are subjected to a known velocity field. On the other two faces of the block, zero normal tractions are assigned and the tangential tractions are such as to equilibrate the ones acting on the top and bottom faces. For a known function f, the values of h depend upon the constitutive relation for the material of the block, and hence, are not known *a priori*. As discussed in Section 3, we solve the resulting system of equations iteratively and find h as a part of the solution of the problem.

For the simple compression problem, we restrict ourselves to the deformations that remain symmetric about both $X_1 = 0$ and $X_2 = 0$. The boundary conditions for the quadrant analyzed numerically are

$$v_1 = 0, \; T_{21} = 0, \; Q_1(Ku)_1 = 0, \text{ at } x_1 = X_1 = 0, \tag{20}$$

$$v_2 = 0, \; T_{12} = 0, \; Q_2 = 0, \text{ at } x_2 = X_2 = 0, \tag{21}$$

$$T_{i\alpha}N_\alpha = 0, \; Q_\alpha N_\alpha = 0, \text{ at } X_1 = H, \tag{22}$$

$$v_2 = U(t), \; e_i T_{i\alpha} N_\alpha = 0, \; Q_\alpha N_\alpha = 0, \text{ at } X_2 = H. \tag{23}$$

That is boundary conditions resulting from the assumed symmetry of deformations are applied to the left and bottom faces, the right face of the block is taken to be traction free, and a prescribed normal velocity field and zero tangential tractions are applied on the top face. All four sides of the block are assumed to be perfectly insulated.

In each of the two problems, a material inhomogeneity or flaw is modeled by adding a temperature bump at the center of the block to the temperature field that corresponds to a homogeneous deformation of the block.

3 Finite Element Formulation of the Problem

In order to avoid having to deal with a severely distorted finite element mesh within the region of localization of the deformation, we employ an updated Lagrangian formulation. Thus to find the deformed shape of the body at time $t+\Delta t$, we take the configuration at time t as the reference configuration, and denote the region occupied by the body at time t by Ω. At subsequent times the current locations of the nodes are computed and Ω equals the union of the 9-noded quadrilateral elements obtained by joining these nodes. No attempt was made to ensure that when the deformation localizes, the element sides will be aligned along the direction of the maximum shearing strain (cf., Needleman, 1989). However, for the simple shearing problem, the element sides are so aligned at the initiation of the localization of the deformation.

We first rewrite equations (14)–(16) so that terms involving the partial derivative with respect to time t only are on the left-hand side and then use the Galerkin method and the lumped mass matrix (e.g., see Hughes (1987)) to derive the following semi-discrete formulation of the problem.

$$\dot{\mathbf{d}} = \mathbf{F}(\mathbf{d}, \delta, \beta, b, m, \nu). \tag{24}$$

Here, $\mathbf{d}$ is the vector of nodal values of the mass density, two components of the velocity, and the temperature. Thus the total number of unknowns or the number of components of $\mathbf{d}$ equals four times the number of nodes. The vector-valued function $\mathbf{F}$ on the right-hand side of equation (24) is a nonlinear function of $\mathbf{d}$ and of the material parameters δ, β, b, m, and ν. For a given set of initial values of ρ, $\mathbf{v}$ and θ, one can deduce the initial conditions on $\mathbf{d}$. The nonlinear coupled set of ordinary differential equations (24) are solved by using the backward-difference Adams method included in the IMSL subroutine LSODE. During the solution of these equations, the tangential traction on the current position of the faces $X_1 = \pm H$ as determined from the immediately preceding solution, is applied. The subroutine LSODE has the option to use the modified Gear method appropriate for stiff equations. This could not be used because of the limited core storage available on the local FPS164 processor attached to IBM 4381 computer. For the Adams method, the subroutine LSODE adjusts the size of the time increment adaptively until it can compute a solution of the nonlinear equations (24) to the prescribed accuracy.

4 Computation and Discussion of Results

We took the following values of various material and geometric parameters to compute numerical results.

$$b = 10{,}000 \text{ sec}, \nu = 0.0222^\circ\text{C}^{-1}, \sigma_0 = 333\text{MPa}, m = 0.025,$$

$$k = 49.22\ \text{Wm}^{-1}{}^\circ\text{C}^{-1}, c = 473\text{Jkg}^{-1}{}^\circ\text{C}^{-1}, \rho_0 = 7{,}800\ \text{kgm}^{-3},$$

$$B = 128\text{GPa}, H = 5\text{mm}, v_0 = 25\ \text{msec}^{-1}. \tag{25}$$

For these choices, $\theta_0 = 89.6^\circ\text{C}$, the nondimensional melting temperature equals 0.5027, and the overall applied strain rate is 5000 sec^{-1}. We assigned a rather large value to the thermal softening coefficient ν to reduce the CPU time required to solve the problem.

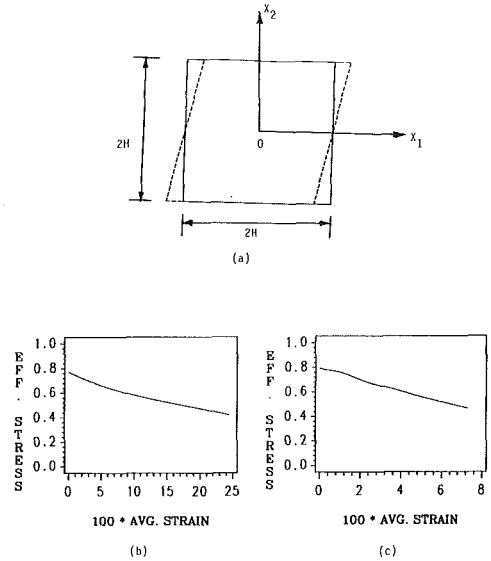


Fig. 1 (a) The shape of the block in the reference configuration and after it has been deformed uniformly in simple shear, (b) Stress-strain curve in simpler shear, and (c) Stress-strain curve in simple compression

Figure 1 depicts the block in the undeformed reference configuration and its shape after it has been deformed uniformly in simple shear. Also plotted are the stress strain curves for the material defined by parameters (25) when the block is deformed in simple shear and simple compression. It is obvious that the softening caused by the heating of the material exceeds the hardening due to strain rate effects right from the beginning. This is due to the rather high value of the thermal softening coefficient assumed for the material of the block. Once the deformation begins to localize, equations (24) become stiff and the maximum size of the time step one can use and still integrate these equations to the desired degree of accuracy becomes extremely small. Ideally, one should then use the Gear method. But, as stated previously, we could not do so because of the limited core storage available. The results presented and discussed next are up to the moment when the deformation has localized into a narrow band. Results computed earlier for the one-dimensional problem (Batra (1987), Batra and Kim (1989), and Wright and Walter (1987)) suggest that the presently computed results represent essentially all of the salient features of the localization of the deformation. We first discuss results for the simple shearing problem, and then the compression problem.

(a) Results for the Simple Shearing Problem. The square region in the configuration at time $t=0$ is divided into 16×16 uniform 9-noded square elements. The velocity field

$$v_1 = x_2, v_2 = 0 \tag{26}$$

that corresponds to steady shearing of the block, and the temperature field

$$\theta = 0 \tag{27}$$

are taken as the initial conditions at time $t=0$, and for the boundary conditions we take

$$f(t) = 1.0, t > 0.$$

Thus, the effect of initial transients is assumed to have died out. This reduces the computational effort required without altering noticeably the computed results. Subsequent calculations with zero-initial conditions for v_1, v_2, and θ have given essentially similar results.

At time $t=0$, a temperature bump given by

$$\Delta\theta = 0.2(1-r^2)^9 \exp(-5r^2),\ r^2 = X_1^2 + X_2^2 \qquad (28)$$

was introduced and the resulting initial boundary value problem solved. The temperature bump (28) simulates a material inhomogeneity or defect; the height of the bump represents, in some sense, the strength of the singularity. Without the temperature bump or some other mechanism to make the deformation nonhomogeneous, the block will undergo unlimited simple shearing deformations and no localization of the deformation will occur. We note that other ways to model an initial imperfection in the body include having a notch (Clifton et al., 1984) and a small region with a slightly lower value of the yield stress (Needleman, 1989). For strain hardening materials the introduction of a temperature bump, a notch or a softer region does not, in general, lead to the localization of the deformation. The average strain at which a shear band forms depends upon, among other factors, the amplitude and shape of the temperature bump.

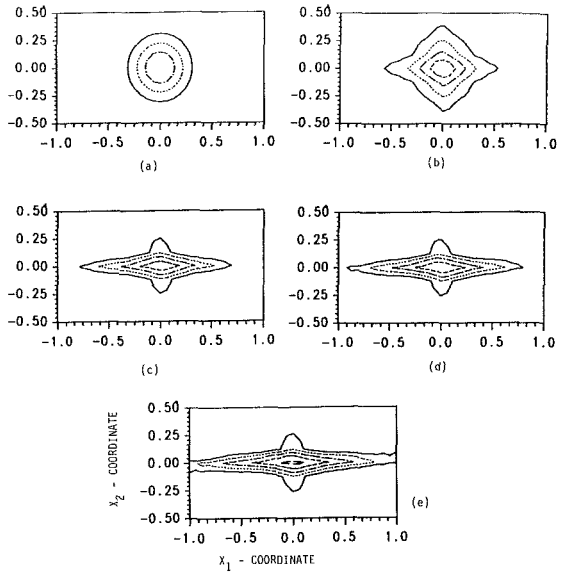


Fig. 2 Isotherms plotted in the reference configuration at different values of the average strain for simple shearing deformations of the block; (a) $\gamma_{avg}=0$, $\theta_{max}=0.2$, — 0.15, • • • • 0.10, —•—•—• 0.15, 0.05, (b) $\gamma_{avg}=0.13$, $\theta_{max}=0.344$, — 0.15, • • • • 0.20, —•—•—• 0.25, - - - - 0.30, (c) $\gamma_{avg}=0.208$, $\theta_{max}=0.441$, — 0.25, • • • • 0.30, —•—•—• 0.35, - - - - 0.40, — — — 0.45, (d) $\gamma_{avg}=0.215$, $\theta_{max}=0.449$ (see part (c) for values of θ corresponding to different curves), and (e) $\gamma_{avg}=0.227$, $\theta_{max}=0.463$ (see part (c) for values of θ corresponding to different curves)

Figure 2 shows isotherms in the reference configuration of the block at four different values of the average strain γ_{avg}. Initially, these isotherms look elliptical because of the different choice of scales along the horizontal and vertical axes. The temperature bump is symmetrical in x_1 and x_2. A reason for selecting different scales along the two axes is that the isotherms eventually flaten out and spread to the vertical boundaries of the block. Thus, larger scale is chosen along the vertical axis to decipher these isotherms. The initial temperature equals 0.20 only at the origin. At an average strain of 13 percent, the isotherms have changed shape; those for a lower temperature look like a rhombus and the ones for the higher temperature resemble closed polygons. Because of the plastic working and zero heat flux boundary conditions the temperature rises everywhere. The heat is continuously being conducted outwards from the central hotter region. Near the corners of the block deformation is nonhomogeneous (e.g., see Fig. 5) and the temperature rise there is more than that at other points except possibly near the center of the block. The nonhomogeneity of the deformation near the corners is a numerical artifact rather than due to the physics of the problem. The use of a very fine mesh should reduce the effect considerably, but a mesh finer than the one employed here could not be used because of the limited core storage available. Once the deformation begins to localize, the temperature rise within the band is significantly more than what it is elsewhere. The temperature contours at average strains of 20.8 percent, 21.5 percent, and 22.7 percent bear this out. At an average strain of 22.7 percent the maximum temperature at the center equals 92 percent of the presumed melting temperature of the material. The isotherms are quite narrow in the vertical direction and progressively become narrower as the deformation localizes.

Figure 3 depicts the v_1-velocity field in the reference configuration of the block at average strains of 0 percent, 18.5 percent, 20.8 percent, and 22.7 percent. Because of the initial temperature bump, the deformation becomes nonhomo-

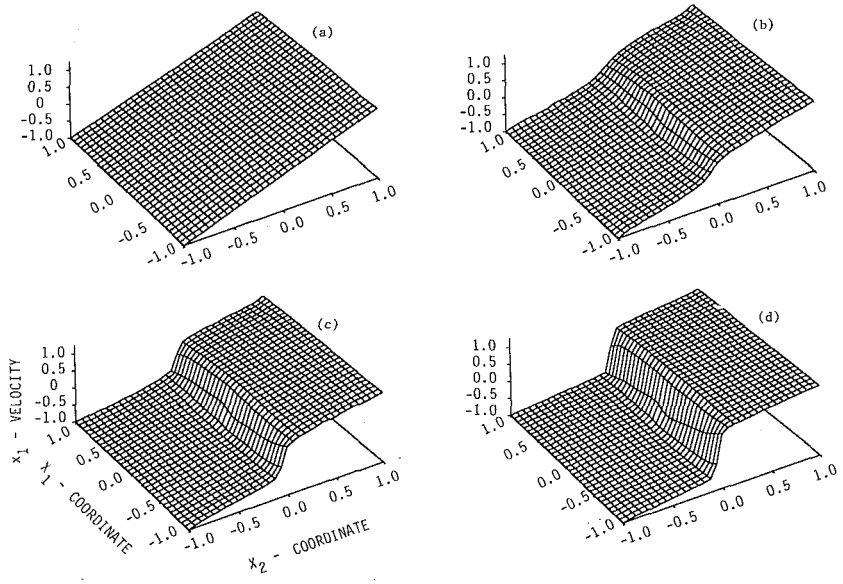


Fig. 3 Velocity field in the direction of shearing at several values of the average strain; (a) $\gamma_{avg}=0$, (b) $\gamma_{avg}=0.185$, (c) $\gamma_{avg}=0.208$, and (d) $\gamma_{avg}=0.227$

geneous. This nonuniformity becomes perceptible at an average strain of 18.5 percent and is quite noticeable when the average strain equals 20.8 percent and 22.7 percent. The nonhomogeneity in the deformation at the corners is not noticeable in these plots probably because of the scale chosen to plot the data. The v_1-velocity field appears to stay antisymmetric in x_2 even through the localization of the deformation. At an average strain of 20.8 percent the shearing strain rate at the center is noticeably higher than what it is within the region $|x_2| \geq 0.1$. During the ensuing deformations of the block, the region near the center undergoes intense straining and that outside of the domain $|x_2| \leq 0.1$ deforms at a strain rate much smaller than the imposed strain rate of 5000 $\sec^{-1}$. With a finer mesh one could sharpen a bit more the boundaries of the two domains.

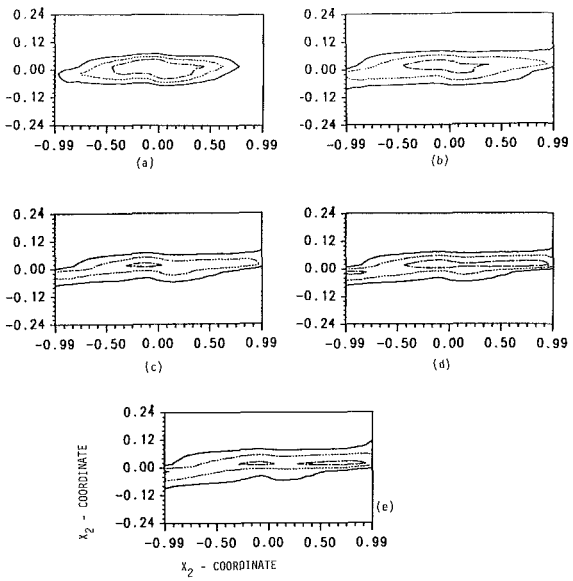


Fig. 4 Contours of the second invariant I of the deviatoric strain rate tensor at different values of the average strain; (*a*) $\gamma_{avg} = 0.185$, $I_{max} = 3.47$; — 1.5, · · · · · · 2.0, —·—·—· 3.5, (*b*) $\gamma_{avg} = 0.198$, $I_{max} = 4.45$; — 1.5, · · · · · · 2.5, —·—·—· 3.5, (*c*) $\gamma_{avg} = 0.208$, $I_{max} = 5.51$; — 2.5, · · · · · · 3.75, —·—·—· 5.0, (*d*) $\gamma_{avg} = 0.215$, $I_{max} = 6.23$; — 2.5, · · · · · · 3.75, —·—·—· 5.0, and (*e*) $\gamma_{avg} = 0.227$, $I_{max} = 8.45$; — 2.5, · · · · · · 5.0, —·—·—· 7.5

In Fig. 4 we have plotted the contours of the second invariant I of the deviatoric strain rate tensor $\tilde{\mathbf{D}}$ at different stages of the localization process. At an average strain of 18.5 percent the peak value of I equals 3.47 and it equals 4.45 when the average strain is 19.8 percent. We note that these are plotted in the reference configuration. It is clear that during the deformation of the block from 18.5 percent average strain to 19.8 percent average strain, the contour of $I = 2.5$ has spread out horizontally and become narrower in the vertical direction. The various plots in Fig. 4 give the impression that there is a kind of source term for I at the center. Once the deformation has started to localize, contours of successively higher values of I seem to originate at the center and fan out. They spread out in the direction of shearing. As noted earlier, severe deformations of the block occur now in this narrow region.

Figure 5 depicts the distribution of the effective stress s_e, defined as being equal to the right-hand side of equation (12) within the block at average strains of 0 percent, 18.5 percent, 20.8 percent, and 22.7 percent. Initially it looks like an inverted hat because every material point is assumed to lie on its yield surface. We note that for the simple shearing problem being studied, σ_{12} is the only component of stress having significant values. Because of the higher temperature at points near the center, the flow stress there is reduced. As the body continues to be deformed, the stress distribution within the block, and especially in the region surrounding the center of the block, alters. The nonhomogeneity of the deformation near the corners is now evident. The temperature rise within the block reduces the flow stress needed to deform the material. Consequently, the value of s_e drops at all points. Even though the strain rate invariant I assumes very high values at points within the region of localization, the softening caused by the temperature rise exceeds the hardening due to strain rate effects and the stress drop in the severely deforming region is enormous. For very high rate of drop of s_e, an unloading elastic wave emanates outwards from the shear band (Batra and Kim, 1989). No such unloading wave was observed in this case. It could be due to the coarseness of the mesh, the integration scheme used, or the rate of the drop of s_e was not too high.

The deformed mesh at average strain of 22.7 percent is shown in Fig. 6. The relatively severe deformations within the region of localization, and nonuniformity of deformations near the corners, is evident.

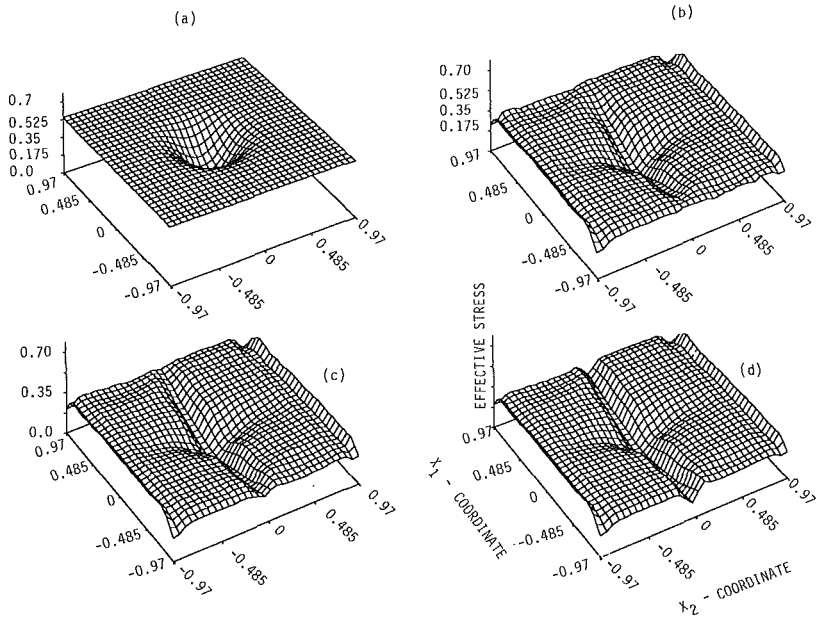


Fig. 5 Distribution of the effective stress within the block at different values of the average strain; (*a*) $\gamma_{avg} = 0$, (*b*) $\gamma_{avg} = 0.185$, (*c*) $\gamma_{avg} = 0.208$, and (*d*) $\gamma_{avg} = 0.227$

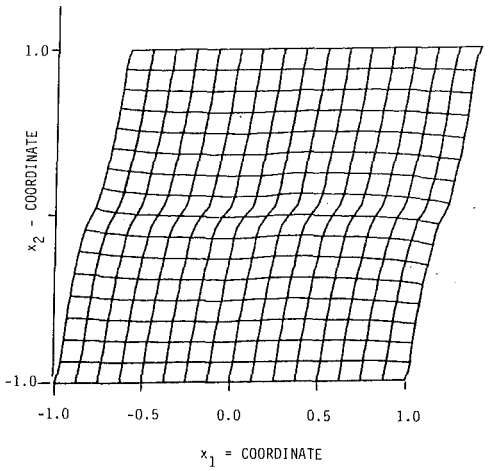


Fig. 6 Deformed mesh at an average strain of 0.227 (simple shearing deformations of the block)

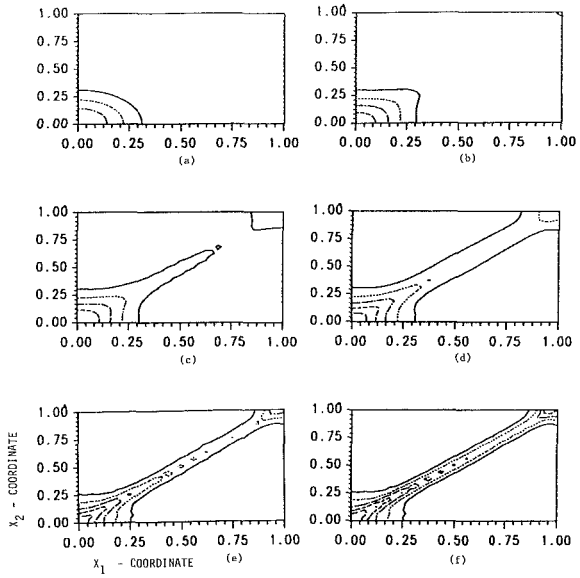


Fig. 7 Isotherms plotted in the reference configuration at different values of the average compressive strain; (*a*) $\gamma_{avg} = 0.0$, $\theta_{max} = 0.2$, — 0.05, • • • • 0.10, —•—•—• 0.15, (*b*) $\gamma_{avg} = 0.035$, $\theta_{max} = 0.286$, — 0.10, • • • • 0.15, —•—•—• 0.20, - - - - 0.25, (*c*) $\gamma_{avg} = 0.040$, $\theta_{max} = 0.313$, see part (*b*) for values of θ corresponding to different curves, (*d*) $\gamma_{avg} = 0.045$, $\theta_{max} = 0.353$, — 0.10, • • • • • 0.15, —•—•—• 0.20, - - - - 0.25, — — — 0.30, (*e*) $\gamma_{avg} = 0.055$, $\theta_{max} = 0.426$, — 0.15, • • • • • 0.20, —•—•—• 0.25, - - - - 0.30, — — — 0.35, and (*f*) $\gamma_{avg} = 0.059$, $\theta_{max} = 0.449$, (see part (*e*) for values of θ corresponding to different curves)

(*b*) ***Results for the Compression Problem.*** Because of the assumed symmetry of the deformation field, the deformations of the block within the first quadrant are analyzed. Several trial runs without introducing any temperature perturbation yielded the following values of the steady-state solution:

$$v_1 = 0.37x_1,\ v_2 = -x_2 \tag{29}$$

for an average applied strain rate of 5000 sec^{-1}. Subsequently this velocity field, and the temperature field given by equation (28), were taken as the initial conditions and the initial boundary value problem solved. A closer look at the results computed by Batra (1987a, 1987b) for the one-dimensional simple shearing problem reveals that the initial state where the perturbation is introduced has very little effect, if any, on the qualitative nature of the results. Figure 7 depicts the

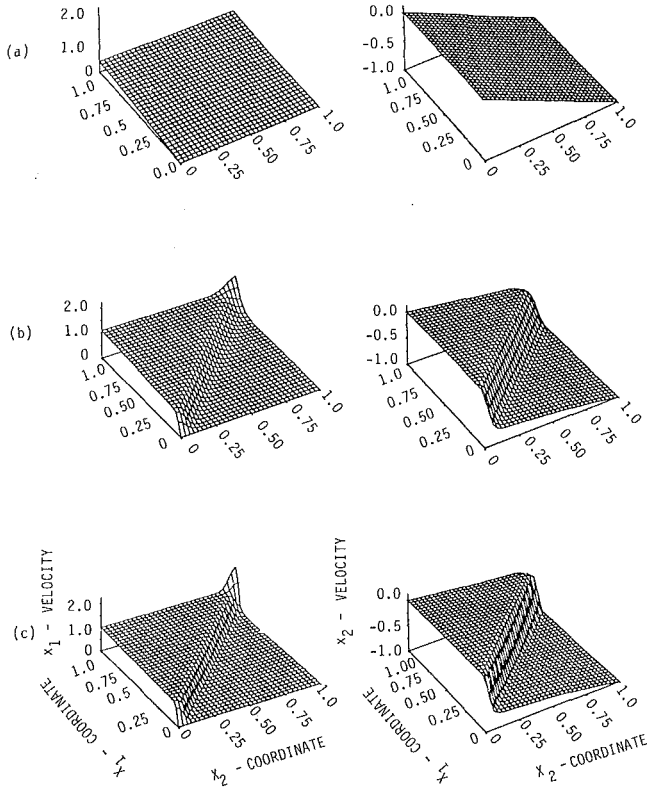


Fig. 8 Velocity field within the block at different values of the average compressive strain; (*a*) $\gamma_{avg} = 0$, (*b*) $\gamma_{avg} = 0.045$, (*c*) $\gamma_{avg} = 0.059$

temperature distribution at several values of the average compressive strain. At an average strain of 3.5 percent the isotherms have changed in shape from elliptic to rhombus and the peak temperature at the center has risen from 0.20 to 0.286. Because of the nonhomogeneous deformations near the top right corner, the temperature rise there is more than that at other points within the block except, of course, those near the center which are undergoing severe deformations. As the temperature plots at average compressive strains of 4 percent, 4.5 percent, 5.5 percent, and 5.9 percent show vividly, the isotherms spread out diagonally indicating that the material around the main diagonal is deforming severely. At these average strains the peak temperature occurs at the center and equals 0.313, 0.353, 0.426, and 0.449, respectively. Thus, the rate of temperature rise at the center is small initially, increases as the deformation begins to localize, and tapers off during the late stages of the localization. Even though heat is being conducted out of this central region the heat produced due to the plastic dissipation exceeds that lost due to conduction. Once the localization process is initiated, the heat generated due to plastic working becomes quite high and the rate of temperature rise within the central region picks up. However, the stress required to deform the material drops and thus reduces the energy dissipated due to plastic working. This and the heat conducted out of the central hotter region explains the slow rate of temperature rise during the late stages of the localization of the deformation.

In Fig. 8 we have plotted the v_1- and v_2-velocity fields at average strains of 0, 4.5 percent, and 5.9 percent. Except at points around the diagonal passing through the top right corner, both v_1 and v_2 vary slowly and nearly linearly, thereby implying that the material region within a narrow zone on both sizes of the diagonal line is undergoing severe deformations. Figure 9 shows the contours of the second invariant I of the deviatoric strain rate tensor at average compressive strains

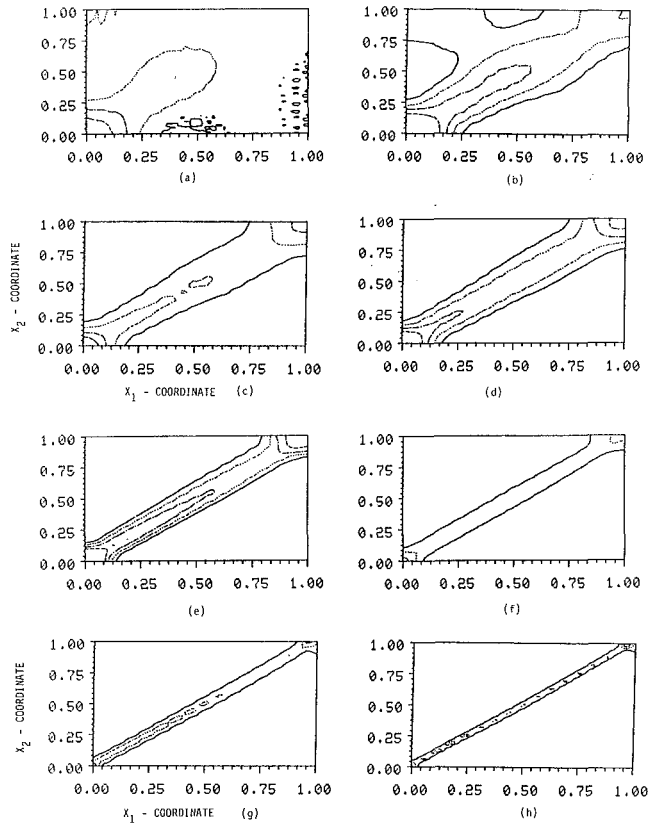


Fig. 9 Contours of the second invariant I of the deviatoric strain rate tensor at different values of the average compressive strain; (*a*) $\gamma_{avg} = 0.012$, $I_{max} = 2.0$. — 1.0, • • • • • • 1.25, —•—•—• 1.50, - - - - 1.75, (*b*) $\gamma_{avg} = 0.018$, $I_{max} = 2.53$, — 1.0, • • • • • • 1.25, —•—•—• 1.50, - - - - 1.75, (*c*) $\gamma_{avg} = 0.025$, $I_{max} = 2.95$. — 1.0, • • • • • • 1.5, —•—•—• 2.0, - - - - 2.5, (*d*) $\gamma_{avg} = 0.03$, $I_{max} = 3.70$, see part (*c*) for values of I corresponding to different curves, (*e*) $\gamma_{avg} = 0.035$, $I_{max} = 5.53$, — 1.5, • • • • • • 2.0, —•—•—• 2.5, - - - - 3.0, (*f*) $\gamma_{avg} = 0.040$, $I_{max} = 8.73$, — 2.5, • • • • • • 5.0, —•—•—• 7.5, (*g*) $\gamma_{avg} = 0.053$, $I_{max} = 16.92$, — 2.5, • • • • • • 7.5, —•—•—• 12.5, and (*h*) $\gamma_{avg} = 0.059$, $I_{max} = 20.7$, — 7.5, • • • • • • 12.5, —•—•—• 17.5

of 0.012, 0.018, 0.025, 0.03, 0.035, 0.04, 0.055, and 0.059. As for the simple shearing problem, the maximum value of I occurs at points near the center of the block and these contours seem to originate at the center and spread out along and perpendicular to the direction of maximum shearing strain; their speed probably depends upon the mesh size. Also, the width of the severely deforming region depends upon the mesh size, too.

Figure 10 depicts the distribution of the effective stress s_e at average strains of 0, 0.027, 0.045, and 0.059. Initially the stress is uniform everywhere except in a narrow region near the center where the flow stress has been reduced due to the higher value of the temperature at these points. The plot at $\gamma_{avg} = 0.027$ reveals that the flow stress has dropped everywhere due to the rise in the temperature of material particles. Still, the effective stress is uniformly distributed except at points near the center of the block. It seems that the localization of the deformation begins in earnest at $\gamma_{avg} = 0.045$. At $\gamma_{avg} = 0.059$ the material region around the main diagonal has severely deformed. The deformed mesh for $\gamma_{avg} = 0.059$ is shown in Fig. 11. That the band has formed is difficult to visualize from the deformed mesh shown. Also, the mesh is incapable of resolving sharp deformation gradients within the localized region.

5 Discussion and Conclusions

The 9-noded quadrilateral element used herein seems to have performed satisfactorily as far as the initiation and some growth of the adiabatic shear band is concerned. As for computations with one-dimensional problems (Batra, 1987a; Batra and Kim, 1989), it is probably due to the coarseness of the mesh that sharp gradients of the deformation within the region of localization could not be completely resolved. This is also supported by the recent work of Shuttle and Smith (1988) on the numerical simulation of shear band formation in soils. Both for plane strain, simple shearing deformations of the block and plane strain compression of the block, the shear band is formed along the direction of maximum shearing. For the compression problem the shear band formed at an average strain of 0.059, and for the simple shearing problem it formed when the average strain equaled 0.229. The results computed

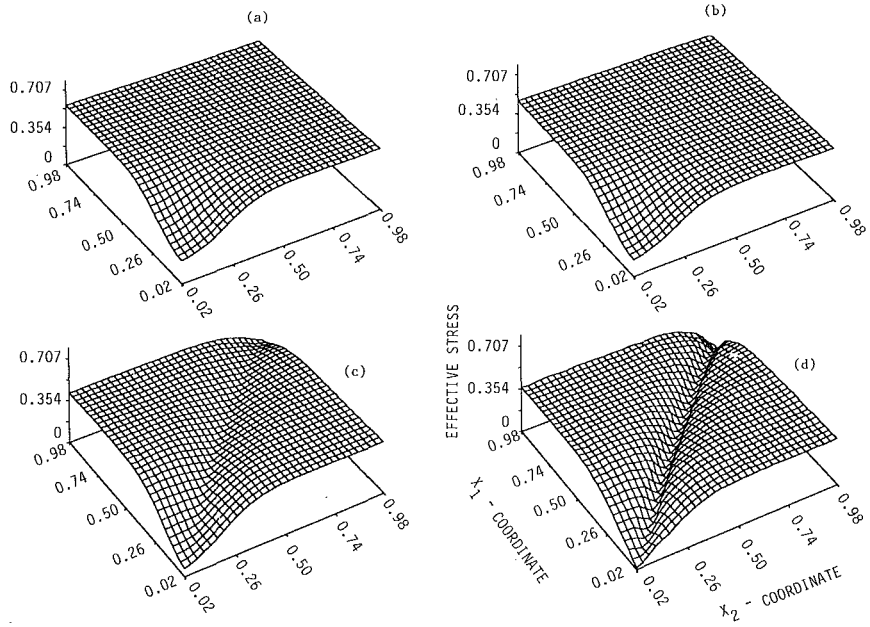


Fig. 10 Distribution of the effective stress within the block at different values of the average strain; (*a*) $\gamma_{avg} = 0$, (*b*) $\gamma_{avg} = 0.027$, (*c*) $\gamma_{avg} = 0.045$, and (*d*) $\gamma_{avg} = 0.059$

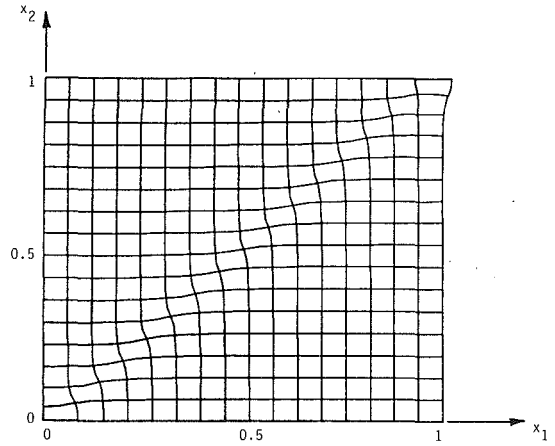


Fig. 11 Deformed mesh at an average compressive strain of 0.059

herein are in qualitative agreement with those of Needleman (1989). Because of the different constitutive assumptions made and the difference in modeling a material inhomogeneity, it is hard to make any quantitative comparisons.

Acknowledgments

This work was supported by the U.S. National Science Foundation grant MSM-8715952 and the U.S. Army Research Office Contract DAAL03-88K-0184 to the University of Missouri-Rolla.

References

Bai, Y. L., 1981, "A Criterion for Thermoplastic Shear Instability," *Shock Waves and High Strain-Rate Phenomena in Metals*, M. A. Myers, and L. E. Murr, eds., Plenum Press, New York, pp. 277–283.

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., 1988, "Steady State Penetration of Thermoviscoplastic Targets," *Comp. Mech.*, Vol. 3, pp. 1–12.

Batra, R. C., and Kim, C. H., 1989, "Adiabatic Shear Banding in Elastic-Viscoplastic Nonpolar and Dipolar Materials," *Int. J. Plasticity*, Vol. 5, to appear.

Batra, R. C., and Lin, Pei-Rong, 1989, "Steady State Axisymmetric Deformations of a Thermoviscoplastic Rod Striking a Hemispherical Rigid Cavity," *Int. J. Impact Engng.*, Vol. 8, to appear.

Bell, J. F., 1968, *Physics of Large Deformations of Crystalline Solids*, Springer-Verlag, New York.

Bodner, S. R., and Partom, Y., 1975, "Mechanical Properties at High Rate of Strain," *Inst. Phys. Conf.*, Ser. No. 21, pp. 102–110.

Burns, T. J., 1985, "Approximate Linear Stability Analysis of a Model of Adiabatic Shear Band Formation," *Q. Appl. Math.*, Vol. 43, pp. 65–84.

Clifton, R. J., 1980, "Adiabatic Shear in Material Response to Ultrahigh Loading Rates," U.S. NRC National Material Advisory Board Report NMAB-356, W. Herrman et al., eds.

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 Metall.*, Vol. 18, pp. 443–448.

Costin, L. S., Crisman, E. E., Hawley, R. H., and Duffy, J., 1979, "On the Localization of Plastic Flow in Mild Steel Tubes Under Dynamic Torsion Loading," *Inst. Phys. Conf. Ser.*, Vol. 47, pp. 90–100.

Hughes, T. J. R., 1987, *The Finite Element Method Linear Static and Dynamic Finite Element Analysis*, Prentice-Hall, Englewood Cliffs, N.J.

Lin, M. R., and Wagoner, R. H., 1986, "Effect of Temperature, Strain and Strain-Rate on the Tensile Flow Stress of I. F. Steel and Stainless Steel Type 310," *Scripta Metall.*, Vol. 20, pp. 143–148.

Lindholm, U. S., and Johnson, G. R., 1983, "Strain-Rate Effects in Metals at Large Strain-Rates," *Material Behavior Under High Stresses and Ultrahigh Loading Rates*, J. Mescal, and V. Weiss, eds., pp. 61–79.

Marchand, A., and Duffy, J., 1988, "An Experimental Study of the Formation Process of Adiabatic Shear Bands in a Structural Steel," *J. Mech. Phys. Solids*, Vol. 36, pp. 251–283.

Merzer, A. M., 1983, "Modelling of Adiabatic Shear Band Development from Small Imperfections," *J. Mech. Phys. Sol.*, Vol. 30, pp. 323–338.

Moss, G. L., 1981, "Shear Strain, Strain Rates and Temperature Changes in Adiabatic Shear Bands," *Shock Waves and High Strain-Rate Phenomenon in Metals*, M. A. Meyers and L. E. Murr, eds., Plenum Press, New York, pp. 299–312.

Needleman, A., 1989, "Dynamic Shear Band Development in Plane Strain," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 1–9.

Rogers, H. C., 1979, "Adiabatic Plastic Deformation," *Ann. Rev. Mat. Sci.*, Vol. 9, pp. 283–311.

Rogers, H. C., 1983, "A Review of Adiabatic Shearing," *Material Behavior Under High Stress and Ultrahigh Loading Rates*, J. Mescall and V. Weiss, eds., Plenum Press, New York, pp. 101–118.

Shuttle, D. A., and Smith, I. M., 1988, "Numerical Simulation of Shear Band Formation in Soils," *Int. J. for Numerical and Analytical Methods in Geomechanics*, Vol. 12, pp. 611–626.

Tillotson, J. H., 1962, General Atomic Report GA-3216.

Timothy, S. P., 1987, "The Structure of Adiabatic Shear Bands in Metals: A Critical Review," *Acta Metall.*, Vol. 35, pp. 301–306.

Wright, T. W., and Batra, R. C., 1985, "The Initiation and Growth of Adiabatic Shear Bands," *Int. J. Plasticity*, Vol. 1, pp. 205–212.

Wright, T. W., and Walter, J., 1987, "On Stress Collapse in Adiabatic Shear Bands," *J. Mech. Phys. Solids*, Vol. 35, pp. 701–716.

Wu, F. H., and Freund, L. B., 1984, "Deformation Trapping Due to Thermoplastic Instability in One-Dimensional Wave Propagation," *J. Mech. Phys. Solids*, Vol. 32, pp. 119–132.

Zener, C., and Hollomon, J. H., 1944, "Effect of Strain-Rate on Plastic Flow of Steel," *J. Appl. Phys.*, Vol. 14, pp. 22–32.

Zienkiewicz, O. C., Onate, E., and Heinrich, J. C., 1981, "A General Formulation for Coupled Thermal Flow of Metals Using Finite Elements," *Int. J. Numer. Methods Eng.*, Vol. 17, pp. 1497–1514.

A. Okuno
Mazda Motor Corporation,
Yokohama, Japan

H. B. Kingsbury
Department of Mechanical Engineering,
University of Delaware,
Newark, DE 19716
Mem. ASME

Dynamic Modulus of Poroelastic Materials

A simple mathematical formula is proposed to predict the fluid damping effects in poroelastic materials. Biot's poroelasticity equations are solved to obtain the response of poroelastic materials undergoing harmonic tension-compression and bending deformation. Complex moduli of poroelastic material are explored from the response functions on basis of mathematical models. It is shown that the effects of material parameters, geometrical parameters, and flow boundary conditions on the fluid damping are predicted by simple mathematical formulas. Numerical results are presented and compared with those of other researchers.

Introduction

A mathematical treatment to predict the fluid damping of open-cell foams was proposed by Rush (1965) and by Gent and Rusch (1966). In their analysis, a rectangular block of fluid-filled foam is considered (Fig. 1(*a*)). The specimen is harmonically compressed in the z direction, and the foam matrix deforms uniformly in the z direction. The fluid inside the foam specimen is forced to flow through the matrix in the x direction. The pressure distribution was determined and the average compressive stress in the cross-section of the specimen was calculated. This stress was added to the compressive stress of the foam matrix. Then an equivalent complex modulus of open-cell foam was derived. They predicted the frequency dependence of the complex modulus to be qualitatively as shown in Fig. 1(*b*).

As the frequency is increased, the fluid flow resistance increases, resulting in increased material stiffness and loss factor, η. At high frequencies, the iteraction force between the fluid and the solid matrix becomes so large that the solid and fluid move together and there is no fluid flow. At these frequencies the loss modulus, due to the fluid flow, becomes zero and the storage modulus becomes its maximum. Thus as the frequency increases, the loss factor η starts to increase from the matrix loss factor η_o, and reaches its maximum at a "critical" frequency ω_c, and then reduces to the matrix loss factor again. Similarly, the storage modulus E' starts to increase from the matrix storage modulus and approaches its maximum value as the frequency goes to infinity.

The Gent and Rusch model explains the effects of the material constants and the specimen geometry on the fluid damping. The validity of their analysis has been well proven by extensive experiments. However, the method cannot be applied to different deformation modes and/or flow boundary conditions. The purpose of the present work is to develop a

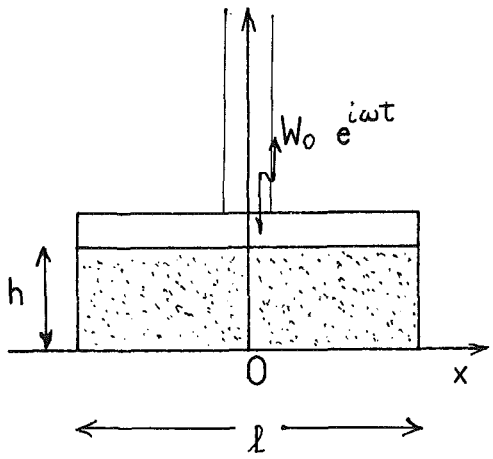


Fig. 1(*a*) Foam slab undergoing harmonic compression deformation

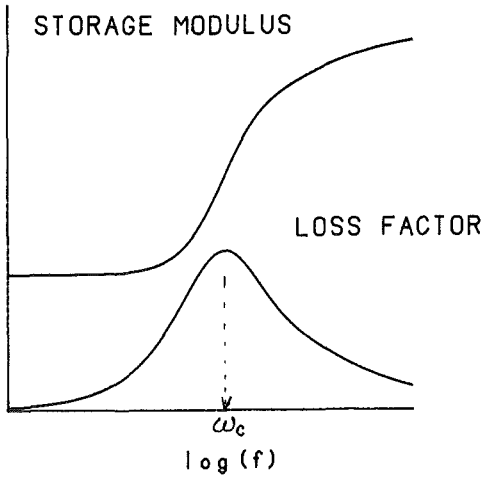


Fig. 1(*b*) Variation of storage modulus and loss factor of a fluid-filled foam with frequency

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

Discussion of this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 E. 47th Street, New York, N.Y. 10017, and will be accpeted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, October 30, 1987; final revision, June 30, 1988.

mathematical method which can account for more general flow conditions and also to get simpler expressions for the maximum loss modulus (damping) and the critical frequency ω_c, which are important for practical design of fluid damping.

To analyze the fluid damping in more general deformation modes and fluid flow conditions, it is essential to have constitutive equations for the fluid-solid system. Wijesinghe and Kingsbury (1979) developed an analytical expression for the fluid damping in poroelastic material using Biot's poroelasticity theory to derive a theoretical complex modulus for poroelastic material. In their analysis, the dynamic response of a slab, column, or disk or porous material subjected to sinusoidally-varying displacement on its surface was considered as shown in Fig. 2. The fluid flows in and out of the specimen through the rigid porous plug. The resulting theoretical complex modulus shows a frequency response similar to the one shown in Fig. 1(*b*).

In this work, a similar approach is used to develop mathematical expressions for complex modulus of poroelastic materials undergoing tension compression and bending deformation with various different flow conditions.

Poroelastic materials are a two-phase, solid-fluid system as defined by Biot (1956, 1957). The solid material forms the skeleton which has small pores filled with fluid. The skeleton is linearly elastic, and the fluid is Newtonian viscous and may be compressible or incompressible. It is assumed that bulk material is homogeneous on a macroscopic scale, and pores are all interconnected. Biot's dynamic theory of poroelasticity accounts for the effects of both fluid inertia and dissipation. In the present work, the quasi-static theory is used, which neglects the inertia effects of the fluid. Therefore the result is valid when either the density of the fluid or the flow velocity is low.

The Governing Equations

The equations governing the deformation of a poroelastic system may be phrased in terms of the average skeleton displacement components, u_i, and the pore fluid pressure, p. In the absence of body forces these equations become (Kingsbury, 1984):

$$\mu^* u_{i,kk} + (\lambda^* + \mu^*) u_{k,ki} - p_{,i} = 0 \tag{1}$$

$$p_{,kk} = \frac{\partial}{\partial t}(\beta_1 u_{i,i} - \beta_2 p) \tag{2}$$

where

$$\beta_1 = b\alpha/n^{*2}, \quad \beta_2 = b/Mn^{*2}. \tag{3}$$

In equations (1) through (3), μ^* and λ^* are the Lamé constants of the skeleton, n^* and b are the skeleton porosity and resistivity, respectively, M is a modified fluid bulk modulus, and α is a solid-phase compressibility coefficient.

The later two coefficients can be expressed in terms of more intrinsic material properties as:

$$\alpha = 1 - \delta/\kappa \tag{4}$$

$$M = \frac{1}{n^*(c-\delta) + \alpha\delta} \tag{5}$$

where δ is the compressibility of the solid comprising the skeleton, κ is the compressibility of the skeleton, and c is the fluid compressibility.

Solution I: Tension-Compression Mode

A solution is first obtained for the simple geometric configuration shown in Fig. 1(*a*). Assuming that the solid strain and the pressure are constant in the z direction, the governing equations reduce to (6) and (7). Since the strain of the x direction is restricted to zero at the boundaries between the supporting plate and the specimen, and the ratio of the specimen height, h, to the specimen width, l, is very small, the strains e_{xx} and e_{yy} are neglected. Consequently, the coupling between the solid displacement and the fluid pressure in (1) vanishes.

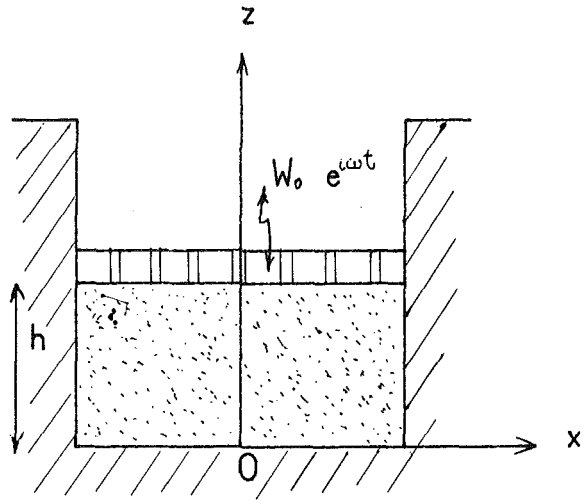


Fig. 2 Column of poroelastic material undergoing harmonic compression

The pressure boundary conditions are given by (8) and (9). The solution forms for the solid displacement and the fluid pressure are given by (10) and (11), respectively. Equation (12) is obtained from (11) and the boundary conditions (8) and (9). Substituting (10) and (11) into the governing equation (7), one obtains (13), and the pressure solution becomes (14). Now the force term, F, acting on the upper surface of the specimen is given by (15), where A is the area of the upper surface. Since the normal stress is given by (16), the force F is obtained as (17) from (15), (16), (10), and (11). Now one can define an equivalent complex modulus $\tilde{E}$ by (18). The first term in (18) represents an apparent modulus of the solid skeleton and the second term represents the effect of the fluid. It is seen that $\tilde{E}$ is a complex value and dependent on the frequency ω, the poroelastic material parameters (α, M, b, n^*), and the geometrical parameter l. The term $(\tan x)/x$ approaches zero as x goes to infinity and 1.0 as x goes to zero. Therefore, the quantity $\alpha^2 M$ represents the maximum storage modulus due to the fluid. The fluid damping is represented by the imaginary part of the complex modulus, E'', which is the loss modulus.

$$\frac{\partial^2 w}{\partial x^2} = 0 \tag{6}$$

$$\frac{\partial^2 p}{\partial x^2} = \frac{\partial}{\partial t}\left\{\beta_1\left(\frac{\partial w}{\partial z}\right) + \beta_2 p\right\} \tag{7}$$

$$\frac{\partial p}{\partial z} = 0 \text{ at } z = 0,\ h \tag{8}$$

$$p = 0 \text{ at } x = \pm 1/2 \tag{9}$$

$$w = \frac{W_o}{h} z e^{i\omega t} \tag{10}$$

$$p = (C_1 \sin\phi x + C_2 \cos\phi x + C_3) e^{i\omega t} \tag{11}$$

$$C_1 = 0,\ C_3 = C_2 \cos(\phi l/2) \tag{12}$$

$$\phi = (-i\omega\beta_2)^{1/2},\ c_2 = \frac{\beta_1 (W_o/h)}{\beta_2 \cos(\phi l/2)} \tag{13}$$

$$p = -\left(\frac{W_o}{h}\right)\frac{\beta_1}{\beta_2}\left\{1 - \frac{\cos(\phi x)}{\cos(\phi l/2)}\right\} e^{i\omega t} \tag{14}$$

$$F=\int_A \tau_{zz}|_{z=h}\,dxdy \tag{15}$$

$$\tau_{zz}=(2\mu^*+s^*)e_{zz}-\alpha p \tag{16}$$

$$F=A\,\frac{wo}{h}\left\{(2\mu^*+\lambda^*)+\alpha^2 M\left(1-\frac{\tan(\phi l/2)}{\phi l/2}\right)\right\}e^{i\omega t} \tag{17}$$

$$\bar{E}(\omega)=(2\mu^*+\lambda^*)+\alpha^2 M\left\{1-\frac{\tan(\phi l/2)}{\phi l/2}\right\}. \tag{18}$$

The merit of this solution is that the critical frequency, ω_c, and the maximum fluid damping (loss modulus) can be easily found as will be next shown.

Taking the derivative of (18) with respect to ω yields (19), where parameters (x, a) are given by (20). Equating (19) to zero, one obtains (21). The imaginary part of this equation leads to the condition shown by (22). From these results , the critical frequency ω_c and the maximum loss modulus $E''_{\max}$ are obtained as (23) and (24), respectively. In equation (23), the distance l^*, which is equal to half of the specimen width, l, is used since that is found to be convenient to generalize the flow path length. This distance l^* represents a flow path length of the fluid, which correspond to the distance between the maximum pressure point and the minimum pressure point.

$$\frac{d\bar{E}(\omega)}{d\omega}=-\frac{d}{d\omega}\left\{\frac{\tan(\phi l/2)}{\phi l/2}\right)$$

$$=\frac{a^2 i}{x}\,\frac{d}{dx}\left(\frac{\tan x}{x}\right) \tag{19}$$

$$x=\phi l/2,\ a=\frac{l}{2}\left(\frac{b}{2n^2 M}\right)^{1/2} \tag{20}$$

$$\frac{\sin(2x)-2x}{x^3\cos^2(x)}=0 \tag{21}$$

$$x=(-1+i)a',\ a'=1.127 \tag{22}$$

$$\omega_c=2.54\,\frac{n^{*2}M}{l^{*2}b},\ l^*=l/2 \tag{23}$$

$$E''_{\max}=0.41\alpha^2 M \tag{24}$$

It is seen from (24) that the maximum fluid damping is determined by only the two parameters α and M. Since the quantity $\alpha^2 M$ is equal to the maximum storage modulus due to the fluid flow, the maximum fluid damping is primarily determined by the type of fluid and the pore density of the solid, while the critical frequency ω_c depends on the geometric parameter l^* and the flow resistivity b.

This solution is comparable with the Gent and Rusch model in that the geometry and the boundary conditions of the models are identical. The advantage of this solution is the simplicity of the formulas. It is seen that equations (23) and (24) give a direct measure for the fluid damping design.

Solution II: Effect of Flow Boundary Condition

In the Solution I, if either one of the specimen side faces is sealed, the problem becomes equivalent to that in which the specimen width L is equal to twice of the original length ($L=2l$). Thus, from equations (23) and (24), the flow boundary condition is expected to change the critical frequency ω_c, but not change the maximum loss modulus nor the shape of its frequency function.

It is convenient to use a Fourier series solution for the fluid pressure to study the effect of the flow boundary conditions. For an example we consider a case for which the specimen has a finite width d in the y direction and the pressure boundary conditions given by (25). The pressure solution is assumed as (26). Then the dilatation term in equation (2) is expanded because of Fourier series as (27), where the coefficient C_{mn} is given by (28). Since the dilatation term e is constant with respect to x and y, the pressure solution is obtained as (29). The force acting on the top surface of the specimen is obtained from (15), (16), and (30). An equivalent complex modulus is obtained as (31). The second term in equation (31) represents the effect of the fluid flow ($\bar{E}_f$). Since the series is very quickly converging, the first term alone is a good approximation. The frequency function has a form shown by (32). General formulas for the maximum loss modulus and the critical frequency ω_c are given by (32) and (33). For this particular case, the maximum loss modulus is approximated by (34). It is seen that the geometrical parameters do not change the maximum value of damping. The critical frequency is expressed as (35), where g^* is an effective flow path length. Because of one term approximation of the series, the coefficients in (34) and (35) are different from those in (23) and (24). These are checked numerically later in this paper.

$$p=0 \text{ at } y=0,d \text{ and } \frac{\partial p}{\partial x}=0 \text{ at } x=0,1/2 \tag{25}$$

$$p=\sum_m\sum_n P_{mn}\sin\frac{n\pi x}{l}\sin\frac{m\pi y}{d}\,e^{i\omega t} \tag{26}$$

$$e=\sum_m\sum_n C_{mn}\sin\frac{n\pi x}{l}\sin\frac{m\pi y}{d}\,e^{i\omega t} \tag{27}$$

$$C_{mn}=\frac{4}{ld}\int_o^l\int_o^d e\sin\frac{n\pi x}{l}\sin\frac{m\pi y}{d}\,dxdy \tag{28}$$

$$p=\sum_m\sum_n\frac{-i\omega\beta_1 C_{mn}}{\left(\frac{n\pi}{l}\right)^2+\left(\frac{m\pi}{d}\right)^2+i\omega\beta_2}$$

$$\times\sin\frac{n\pi x}{l}\sin\frac{m\pi y}{d}\,e^{i\omega t} \tag{29}$$

$$F=A\,\frac{W_o}{h}\left\{(2\mu^*+\lambda^*)+\sum_{\substack{m\\ m=\text{odd}}}\sum_{\substack{n\\ n=\text{odd}}}\right.$$

$$\left.\times\frac{i\alpha\beta_1\omega\,\frac{64}{n^2m^2\pi^4}}{\left(\frac{n\pi}{l}\right)^2+\left(\frac{m\pi}{d}\right)^2+i\omega\beta_2}\right\}e^{i\omega t} \tag{30}$$

$$\bar{E}(\omega)=(2\mu^*+\lambda^*)+\sum_{\substack{m\\ m=\text{odd}}}\sum_{\substack{n\\ n=\text{odd}}}$$

$$\times\frac{i\alpha\beta_1\omega\,\frac{64}{n^2m^2\pi^4}}{\left(\frac{n\pi}{l}\right)^2+\left(\frac{m\pi}{d}\right)^2+i\omega\beta_2} \tag{31}$$

$$\bar{E}_f=\frac{ix_1\omega}{x_2+ix_3\omega} \tag{32}$$

$$\omega_c=\frac{x_2}{x_3},\quad E''_{\max}=\frac{x_1}{2x_3} \tag{33}$$

$$E''_{\max}=0.33\alpha^2 M \tag{34}$$

$$\omega_c = 2.47 \frac{n^{*2}M}{g^{*2}b},$$

$$g^* = \left(\frac{l^{*2}d^{*2}}{l^{*2}+d^{*2}}\right)^{1/2}, \quad l^* = l/2, d^* = d/2. \tag{35}$$

Solution III: Bending Deformation

The dynamic modulus of a poroelastic material undergoing bending deformation is suggested by Biot (1964). Here we consider a pure bending case shown in Fig. 3.

The pressure boundary conditions are given by (36). Neglecting strains e_{zz} and e_{yy}, the dilatation term is given by (37). The equation of motion for the classical simple beam becomes (38), where $\tilde{M}$ is bending moment. The moment term is given by (39), where E is the elastic modulus and I is the moment of inertia of the beam cross-section. The pressure solution in the poroelastic material is assumed as (40), where the flow in the x direction is ignored. Expanding the dilatation term by use of a Fourier series as (41), the pressure solution is obtained as (42). Substituting (42) and (39) into (38), one finds an equivalent complex modulus shown by (43) and (44). Taking the first term of the series, the maximum loss modulus and the critical frequency are obtained as (45) and (46). It is seen that this result is equivalent with previous solutions if one use $l = h/2$. Thus the deformation mode does not change the maximum damping obtainable from fluid flow.

$$\frac{\partial p}{\partial z} = 0 \text{ at } z = \pm\frac{h}{2} \tag{36}$$

$$e = -z\frac{\partial^2 w}{\partial x^2} \tag{37}$$

$$\frac{\partial^2 M}{\partial x^2} + \rho h \frac{\partial^2 w}{\partial t^2} = F_o e^{i\omega t} \tag{38}$$

$$\tilde{M} = \int_{h/2}^{h/2} \pi_{xx} z dz$$

$$= EI\frac{\partial^2 w}{\partial x^2} - \alpha\int_{-h/2}^{h/2} pz dz \tag{39}$$

$$p = \sum_n P_n \sin\frac{(2n-1)\pi}{h} z e^{i\omega t} \tag{40}$$

$$-z\frac{\partial^2 w}{\partial x^2} + \sum_n D_n \sin\frac{(2n-1)\pi}{h} z e^{i\omega t} \tag{41}$$

$$p = \sum_n \frac{i\omega\beta_1 \dfrac{4h}{(2n-1)^2\pi^2}(-1)^n}{\dfrac{(2n-1)^2\pi^2}{h^2} + i\omega\beta_2} \sin\frac{(2n-1)\pi}{h} z \frac{\partial^2 w}{\partial x^2} \tag{42}$$

$$\left\{E_0 + \Sigma \frac{i\alpha\omega\beta_1 \dfrac{96}{(2n-1)^4\pi^4}}{\dfrac{(2n-1)^2}{h^2} + i\omega\beta_2}\right\} I \frac{\partial^4 w}{\partial x^4}$$

$$+ h\rho \frac{\partial^2 w}{\partial t^2} = F_o e^{i\omega t} \tag{43}$$

$$\tilde{E}(\omega) = E_0 + \sum_n \frac{i96\alpha\beta_1\omega/(2n-1)^4\pi^4}{\left\{\dfrac{(2n-1)\pi}{h}\right\}^2 + i\omega\beta_2} \tag{44}$$

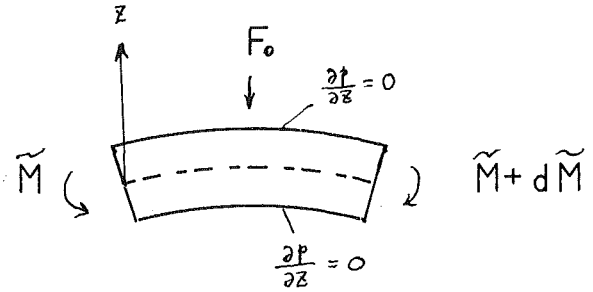


Fig. 3 Simple beam model

$$\omega_c = 2.47 \frac{n^{*2}M}{h^{*2}b} \tag{45}$$

$$E''_{\max} = 0.49\,\alpha^2 M. \tag{46}$$

Discussions

The dynamic modulus due to fluid flow can be approximated by the simple formula (47), where the maximum loss modulus E' and the critical frequency ω_c are given by (48) and (49), respectively. The idea of an effective flow path length needs to be introduced to generalize the effect of flow boundary condition. The effective flow path length is the distance between the maximum and minimum pressure points in the case of one-dimensional flow. In the case of multidimensional flow, it is approximated as (35).

$$\tilde{E}(\omega) = E_o + \frac{2i\omega E''_{\max}}{\omega_c + i\omega} \tag{47}$$

$$E''_{\max} = 0.41\,\alpha^2 M \tag{48}$$

$$\omega_c = 2.54 \frac{n^{*2}M}{l^{*2}b}. \tag{49}$$

A numerical example is chosen to compare the results obtained by equations (18), (31), and (44) with the approximate formula shown by (47). Material properties used for the sample calculation are as follows. (A polyurethane-air system was chosen as an example).

Storage modulus of skeleton E:	3.0×10^3 Pa
Porosity n:	0.94
Air bulk modulus M:	1.0×10^5 Pa
Parameter α:	0.96
Resistivity b:	1.0×10^5 Pa

The results of the sample calculation are shown in Fig. 4. The effective flow path length is the same for all models. It is seen that the four curves are fairly close to each other, although there are some differences in the maximum storage and loss moduli as expected from the analytical results.

The dynamic modulus obtained as (18) was compared with the result of Gent and Rusch in an Okuno (1986). It was shown that although there were some differences observed in the storage modulus, on the whole, the equation (18) agreed well with the Gent and Rusch model. Equation (18) results from the assumptions in equation (6) that there is no interaction between the solid and the fluid strain; i.e., the coupling term in the first governing equation was neglected. The same assumption was made by Gent and Rusch.

However, in the case shown in Fig. 2, the major solid strain occurs in the same direction as the fluid flow. Therefore, neglecting the coupling term is considered to cause significant error. It is useful to check the amount of discrepancy due to the assumption of no coupling. In Fig. 5 the result by equation (47) is compared with the result of Wijesinghe and Kingsbury (1979), which is a solution of the coupled equation.

It is seen that neglecting the coupling term results in only slight overestimation of the maximum loss modulus and

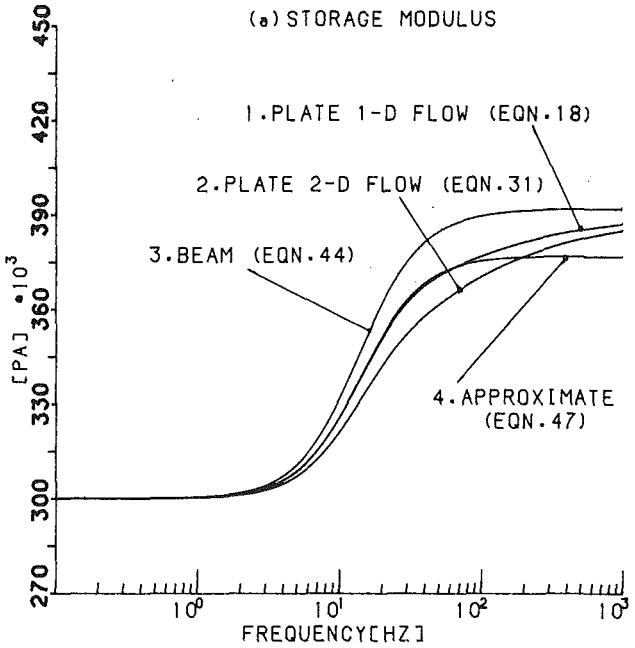


Fig. 4(*a*) Dynamic storage modulus in different deformation modes

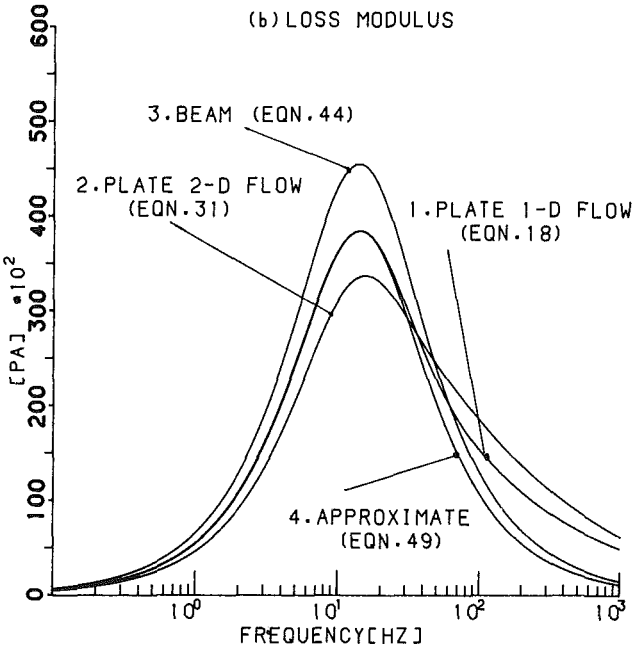


Fig. 4(*b*) Dynamic loss modulus in different deformation modes

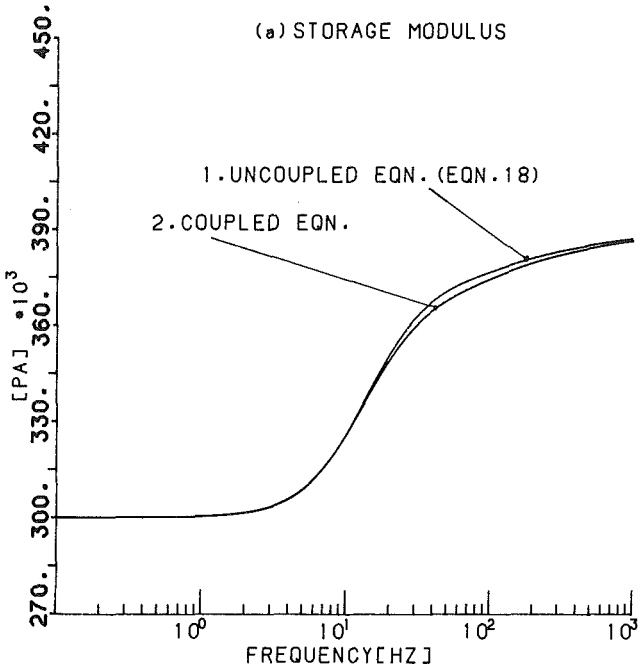


Fig. 5(*a*) Dynamic storage modulus by coupled and uncoupled equations

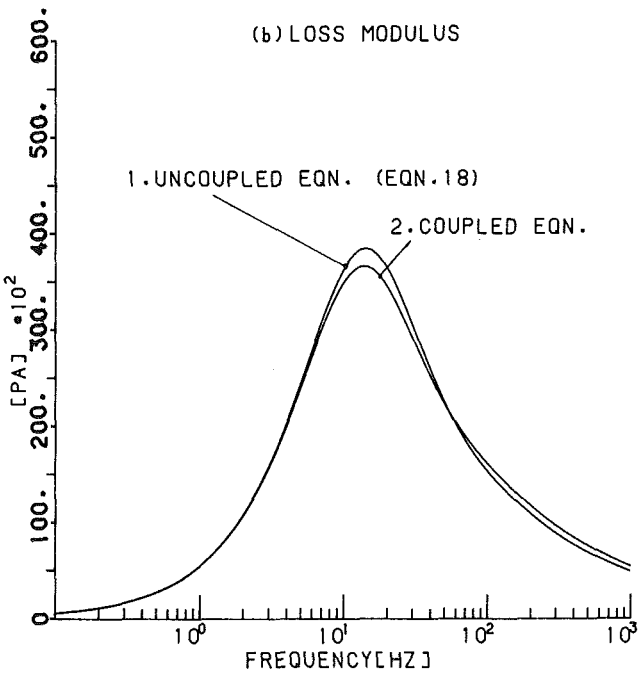


Fig. 5(*b*) Dynamic loss modulus by coupled and uncoupled equations

storage modulus, while it does not change the critical frequency significantly.

Dynamic responses of two-layer and three-layer beams with a foam layer was studied by Okuno (1986), and it was shown that the maximum loss modulus of the foam layer is about the same value as that of (48) regardless of the deformation mode. The critical frequency of the foam layer was also analyzed and it was found that the results were close to (49).

From all the analyses presented so far, it is concluded that the approximation of fluid damping, which is obtained as (47), (48), and (49), gives sufficient accuracy for practical use.

Summary

The effect of fluid flow on the dynamic modulus of a poroelastic material can be approximated by a simple complex modulus as:

$$\tilde{E}(\omega) \simeq E_o + \frac{2i\omega E'_{\max}}{\omega_c + i\omega} .$$

The loss modulus (fluid damping) shows a bell-shaped frequency dependence and its maximum value and the critical frequency are approximated as follows:

$$E'_{\max} \simeq 0.41\alpha^2 M$$

$$\omega_c = 2.54 \frac{n^{*2} M}{l^{*2} b} .$$

These formulas show that the maximum fluid damping is determined by the type of fluid, while the critical frequency can be controlled by choosing the flow resistivity of the skeleton and the flow path length. This information is quite useful for practical fluid damping design.

References

Biot, M. A., 1956, "Theory of Propagation of Elastic Waves in Fluid Saturated Porous Solid I: Low Frequency Range," *J. of Acous. Soc. of America*, Vol. 28, pp. 168–178.

Biot, M. A., and Willis, D. G., 1957, "The Elastic Coefficients and Theory of Consolidation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 24, No. 79, pp. 594–601.

Biot, M. A., 1964, "Theory of Buckling of a Porous Slab and Its Thermoelastic Analogy," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 194–198.

Gent, A. N., and Rusch, K. C., 1966, "Viscoelastic Behavior of Open-Cell Foams," *Rubber Chem. Tech.*, Vol. 39, pp. 389.

Kingsbury, H. B., 1984, "Determination of Material Parameters of Poroelastic Media," *Fundamentala of Transport Phenomena in Porous Media*," J. Bear, and M. Coropcioglu, eds., M. Nijhoff, Dordrecht, pp. 581–609.

Okuno, A., 1986, "Dynamic Response of Structures Containing Poroelastic Material," Ph.D. Thesis, The University of Delaware, Newark, Del.

Rusch, K. C., 1965, "Dynamic Behavior of Flexible Open-Cell Foams, Ph.D. Thesis, University of Akron, Akron, Ohio.

Wijesinghe, A. M., and Kingsbury, H. B., "On the Dynamic Behavior of Poroelastic Materials," *J. of Acous. Soc. of America*, Vol. 65, pp. 90–95.

Squeeze Film Behavior for Anisotropic Porous Rectangular Plates

A. B. Wheeler

R. Balasubramanyam

Department of Mathematics,
University College of Wales,
Aberystwyth SY23 3BZ,
United Kingdom

The squeeze film between two rectangular plates when one has a porous facing is analyzed taking into account the anisotropic permeability and slip velocity at the fluid and porous material interface. Modified equations for calculating the pressure, the load carrying capacity, and the film thickness and time relations are presented. The effect of the anisotropic permeability and slip velocity at the fluid and porous material interface on the squeeze film behavior is discussed and found to be important.

Introduction

Porous materials have been used in lubrication applications such as bearings (Booser, 1970) and squeeze films (Wu, 1970). They have the advantage of being self-lubricating, which overcomes the need for oil pipes, pumps, etc., and simplifies the problem of machine design. In most analyses (Wu, 1970, 1971a, 1971b; Berman, 1953; Morgan and Cameron, 1957) of flow with porous boundaries it has been customarily assumed that the conventional no-slip velocity condition remains valid at porous surfaces. Recently, however, Beavers et al. (1967, 1970) demonstrated the existence of slip velocity at the surface of a porous material in their experiments involving laminar flow of water and of oil in rectangular ducts having one porous wall. Experiments were later performed by Taylor (1971) to calculate the slip constant. Later, Sparrow et al. (1971) used the slip velocity assumption in generalizing the analysis of the squeeze films between porous annular disks (Wu, 1970) and found the effect of the slip velocity was to reduce the load-carrying capacity and the response time of the porous squeeze film. Wu (1972) extended the previous analysis of squeeze films between porous rectangular plates to include the effect of velocity slip at the porous surface and came to similar conclusions as mentioned by Sparrow et al. (1971).

The purpose of this paper is to expand the previous analysis of squeeze films between porous rectangular plates (Wu, 1972) to include the effect of anisotropic permeability and velocity slip at the porous surface and to provide the modified equations for calculating pressure, load-carrying capacity, and instantaneous film thickness of this kind of squeeze film as functions of time.

Analysis

Assumptions. The following assumptions are made in the analysis:

1 The flow in the film region is laminar.
2 The fluid is incompressible and has constant properties.
3 In the film region the inertial effects can be neglected, the pressure is independent of the z-coordinate, and the z-derivatives of the velocity components dominate.
4 The flow in the porous medium follows Darcy's law for anisotropic materials

$$\mathbf{v}^* = -\frac{\mathbf{K}}{\eta} \cdot \text{grad } P^*, \tag{1}$$

where P^* is the pressure in the porous medium, η is the isotropic viscosity of the fluid, $\mathbf{v}^* = (v_x^*, v_y^*, v_z^*)$ is the velocity vector with components referred to the Cartesian axes, shown in Fig. 1, and $\mathbf{K}$ is the anisotropic permeability tensor. It is assumed that the anisotropic properties of the porous material are such that the principal directions of $\mathbf{K}$ remain constant and parallel to the coordinate directions shown in Fig. 1. We can then write

$$\mathbf{K} = \begin{bmatrix} K_x & 0 & 0 \\ 0 & K_y & 0 \\ 0 & 0 & K_z \end{bmatrix} \tag{2}$$

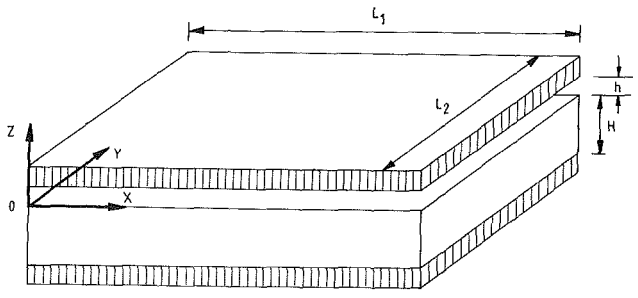


Fig. 1

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

Discussion on this paper should be addressed to the 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 22, 1988; final revision, January 3, 1989.

where K_x, K_y, K_z are the constant permeability coefficients. (Throughout, starred variables refer to the porous medium and unstarred variables to the film.)

On the previous assumptions for squeeze film flows for the quasi-static case, the Navier-Stokes equations reduce to

$$\frac{\partial^2 v_x}{\partial z^2} = \frac{1}{\eta}\frac{\partial P}{\partial x}, \qquad \frac{\partial^2 v_y}{\partial z^2} = \frac{1}{\eta}\frac{\partial P}{\partial y}. \tag{3}$$

In order to solve the equations it is necessary to specify a boundary condition on the velocity between the film and medium. In this paper, we adopt the empirical model suggested by Beavers and Joseph (1967) and, subsequently, given theoretical justification by Saffman (1971). For an anisotropic medium this takes the form

$$\left.\frac{\partial v_x}{\partial z}\right|_{z=0} = \frac{\alpha_x}{\sqrt{K_x}}(v_{0x} - v_{0x}^*), \qquad \left.\frac{\partial v_y}{\partial z}\right|_{z=0} = \frac{\alpha_y}{\sqrt{K_y}}(v_{0y} - v_{0y}^*), \tag{4}$$

where v_{0x}, v_{0y} denote the slip velocities ($v_{0x} = v_x(0)$, $v_{0y} = v_y(0)$) along the x and y directions, respectively, v_{0x}^*, v_{0y}^* the mean velocity components in the porous medium which are given by

$$v_{0x}^* = -\frac{K_x}{\eta}\frac{\partial P^*}{\partial x}, \qquad v_{0y}^* = -\frac{K_y}{\eta}\frac{\partial P^*}{\partial y}, \tag{5}$$

respectively, α_x, α_y are dimensionless constants which depend on the characteristics of the porous medium.

It can be seen that when

$$\frac{\sqrt{K_x}}{\alpha_x}, \frac{\sqrt{K_y}}{\alpha_y} \to 0,$$

equations (4) reduce to the no-slip boundary conditions appropriate to a solid wall. Integrating equations (3) twice with respect to z, and applying the slip boundary conditions at the interface and the no-slip boundary condition at $z = h$, yields

$$v_x = \frac{1}{\eta}\frac{\partial P}{\partial x}\left\{\frac{z^2(\sqrt{K_x}+h\alpha_x) - z\alpha_x(h^2-2K_x) - h\sqrt{K_x}(h+2\alpha_x\sqrt{K_x})}{2(\sqrt{K_x}+h\alpha_x)}\right\},$$
$$v_y = \frac{1}{\eta}\frac{\partial P}{\partial y}\left\{\frac{z^2(\sqrt{K_y}+h\alpha_y) - z\alpha_y(h^2-2K_y) - h\sqrt{K_y}(h+2\alpha_y\sqrt{K_y})}{2(\sqrt{K_y}+h\alpha_y)}\right\}. \tag{6}$$

The equation of continuity is

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = 0. \tag{7}$$

On substituting (6) into (7) and integrating across the film thickness h, we obtain

$$\frac{\partial}{\partial x}\left[\frac{h^3}{12\eta}\frac{\partial P}{\partial x}\sigma_x\right] + \frac{\partial}{\partial y}\left[\frac{h^3}{12\eta}\frac{\partial P}{\partial y}\sigma_y\right] = [v_x]_o^h = v_{hz} - v_{oz} = \dot{h} - v_{oz} \tag{8}$$

where $v_{hz} = \dot{h}$ is the velocity of the top plate, and σ_x and σ_y are given by

$$\sigma_x = 1 + 3\frac{\left(1 + 2\frac{\alpha_x\sqrt{K_x}}{h}\right)}{1 + \frac{h\alpha_x}{\sqrt{K_x}}}, \qquad \sigma_y = 1 + 3\frac{\left(1 + 2\frac{\alpha_y\sqrt{K_y}}{h}\right)}{1 + \frac{h\alpha_y}{\sqrt{K_y}}}.$$

It is assumed that the mean velocities in the porous medium satisfy the equation of continuity,

$$\frac{\partial v_x^*}{\partial x} + \frac{\partial v_y^*}{\partial y} + \frac{\partial v_z^*}{\partial z} = 0, \tag{9}$$

and using Darcy's law, we obtain the elliptic equation for the pressure in the porous medium

$$K_x\frac{\partial^2 P^*}{\partial x^2} + K_y\frac{\partial^2 P^*}{\partial y^2} + K_z\frac{\partial^2 P^*}{\partial z^2} = 0. \tag{10}$$

Since the velocity component in the z-direction must be continuous at the interface,

$$v_{0z} = -\frac{K_z}{\eta}\left.\frac{\partial P^*}{\partial z}\right|_{z=0}, \tag{11}$$

equation (8) becomes

$$\frac{\partial}{\partial x}\left[\frac{h^3}{12\eta}\frac{\partial P}{\partial x}\sigma_x\right] + \frac{\partial}{\partial y}\left[\frac{h^3}{12}\frac{\partial P}{\partial y}\sigma_y\right] = \dot{h} + \frac{K_z}{\eta}\left.\frac{\partial P^*}{\partial z}\right|_{z=0}. \tag{12}$$

Equation (12) represents the modified Reynolds equation for the fluid pressure in the film region for the quasi-static case and equations (10) and (12) are to be solved subject to the usual boundary conditions of zero pressure at the edges of the plates and at the lateral surfaces of the porous medium, zero-flow through the bottom plate and continuity of pressure across the interface.

$$\begin{aligned} &P(0,y) = 0, \\ &P(L_1,y) = 0, \\ &P(x,0) = 0, \\ &P(x,L_2) = 0, \\ &P^*(0,y,z) = 0, \\ &P^*(L_1,y,z) = 0, \\ &P^*(x,0,z) = 0, \\ &P^*(x,L_2,z) = 0, \\ &\left.\frac{\partial P^*}{\partial z}\right|_{z=-H} = 0. \end{aligned} \tag{13}$$

Since the pressure must be continuous at the fluid and porous material interface,

$$P(x,y) = P^*(x,y,0). \tag{14}$$

Solutions

Equation (10) can be solved by separation of variables subject to the boundary conditions (13). The solution is

$$P^*(x,y,z) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}\cosh[\gamma_{mn}(z+H)]\sin(\alpha_m x)\sin(\beta_n y), \tag{15}$$

where $\alpha_m = \frac{m\pi}{L_1}$; $\beta_n = \frac{n\pi}{L_2}$;

$$\gamma_{mn} = \left(\frac{\alpha_m^2 K_x + \beta_n^2 K_y}{K_z}\right)^{1/2}$$

and the coefficients a_{mn} are to be determined.

To solve equation (12), $P(x,y)$ is expressed as an infinite series consisting of a complete set of orthogonal functions each of which satisfies the respective boundary conditions given in (13), i.e.,

$$P(x,y) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} b_{mn}\sin(\alpha_m x)\sin(\beta_m y). \tag{16}$$

The coefficients b_{mn} are determined from the matching condition (14), i.e.,

$$b_{mn} = a_{mn}\cosh(\gamma_{mn}H). \tag{17}$$

On substituting equations (15), (16), (17) into (12), the orthogonality of the eigenfunctions gives

$$a_{mn} = \frac{-192\eta\dot{h}}{L_1 L_2 h^3 \alpha_m \beta_n \left[(\alpha_m^2\sigma_x + \beta_n^2\sigma_y)\cosh(\gamma_{mn}H) + \dfrac{12K_z\gamma_{mn}}{h^3}\sinh(\gamma_{mn}H)\right]}, \tag{18}$$

where m and n are odd.

The load-carrying capacity of the squeeze film is found by integrating the pressure over the top plate

$$W = \int_0^{L_2}\int_0^{L_1} P(x,y)\,dx\,dy, \tag{19}$$

giving in dimensionless form

$$-\frac{h^3 W}{\eta {L_1}^2 {L_2}^2 \dot{h}} = \frac{768}{\pi^4 L_1 L_2}\sum_{m,n\ \text{odd}}^{\infty}\sum \frac{\bar{F}_{mn}}{m^2 n^2}, \tag{20}$$

where

$$\bar{F}_{mn} = \left\{\alpha_m^2\bar{\sigma}_x + \beta_n^2\bar{\sigma}_y + \frac{12K_z\gamma_{mn}\tanh(\gamma_{mn}H)}{{h_o}^3\bar{h}^3}\right\}^{-1},$$

$$\bar{\sigma}_x = 1 + 3\,\frac{\left(1 + 2\,\dfrac{\alpha_x\sqrt{K_x}}{\bar{h}h_0}\right)}{1 + \bar{h}\,\dfrac{h_0\alpha_x}{\sqrt{K_x}}}, \qquad \bar{\sigma}_y = 1 + 3\,\frac{\left(1 + 2\,\dfrac{\alpha_y\sqrt{K_y}}{\bar{h}h_0}\right)}{1 + \bar{h}\,\dfrac{h_0\alpha_y}{\sqrt{K_y}}},$$

h_0 is the prescribed initial thickness and $\bar{h}$ given by

$$\bar{h} = \frac{h}{h_0}.$$

The film thickness, at any time, can now be obtained by integrating equation (20) for the given load as a function of time,

$$\int_0^{\Delta t} W(\tau)\,d\tau = -\frac{768\eta L_1 L_2}{\pi^4 {h_o}^2}\sum_{m,n\ \text{odd}}^{\infty}\sum c_{mn}, \tag{21}$$

where

$$c_{mn} = \int_1^{\bar{h}} \left\{m^2 n^2\left[\xi^2(\alpha_m{}^2\bar{\sigma}_x(\xi) + \beta_n^2\bar{\sigma}_y(\xi) + \frac{12K_z\gamma_{mn}}{{h_o}^3}\tanh(\gamma_{mn}H)\right]\right\}^{-1} d\xi.$$

Δt is the time interval. For constant load one has in dimensionless form

$$\frac{{h_o}^2 W}{L_1^2 L_2^2 \eta}\,\Delta t = \frac{-768}{\pi^4 L_1 L_2}\sum_{m,n\ \text{odd}}^{\infty}\sum c_{mn}. \tag{22}$$

The Hays (1963) result

$$\Delta t = \frac{384\eta L_1^3 L_2}{\pi^6 h_0^2 W}\left[\frac{1}{\bar{h}^2} - 1\right]\sum_{m,n\ \text{odd}}^{\infty}\sum\left\{\frac{1}{m^2 n^2\left(m^2 + \dfrac{L_1^2}{L_2^2}n^2\right)}\right\} \tag{23}$$

can be obtained from (22) by putting $K_x = K_y = K_z = 0$. Similarly, Wu's (1972) solution can be obtained by letting $K_x = K_y = K_z$, $\alpha_x = \alpha_y$ and $v_{0x}^* = 0$, $v_{0y}^* = 0$.

Results

By comparing equations (15), (20), and (22) with the corresponding equations given by Wu (1972) for the velocity slip case, it is found that the effect of anisotropy and velocity slip comes through the introduction of the parameters $\bar{\sigma}_x$, $\bar{\sigma}_y$, and γ_{mn}. Little is known about the slip constants α_x and α_y, except that they seem to be independent of the fluid viscosity, but depend on the porous surface characteristics. The smallest value of the slip constant (which yields largest slip velocity) that has ever been found (Beavers and Joseph, 1967) is 0.1. In order to show how significant the effect of velocity slip can be in certain lubrication applications the values of α_x, $\alpha_y = 0.5$ and α_x, $\alpha_y = 0.1$ are chosen.

The effect of L_1/L_2, keeping the area of the surface constant, and K_y/K_x on the load-carrying capacity for different values of α_x and α_y, are shown in Figs. 2, 3, and 4. It is seen when $K_x = K_y = K_z$ and $\alpha_x = \alpha_y$ (isotropic case, Fig. 2) the curve is symmetric with greatest load-carrying capacity occurring when the surface is square (as expected), whereas for the anisotropic case the greatest load can be carried by a surface which is off-square. The curve of maximum load-carrying capacity is also plotted showing the extent to which the ratio of K_y/K_x affects the amount off-square the surface has to be to give maximum load-carrying capacity.

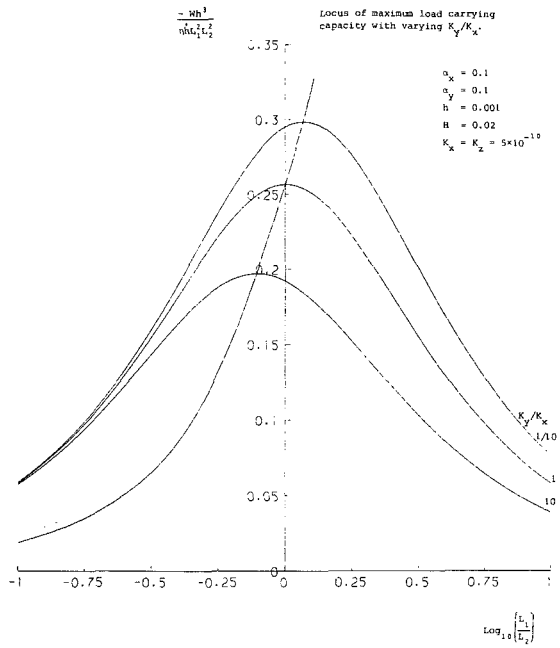


Fig. 2

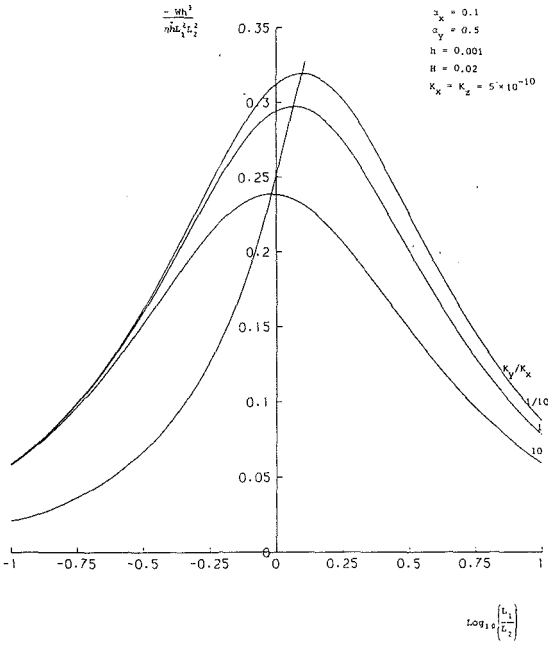


Fig. 3

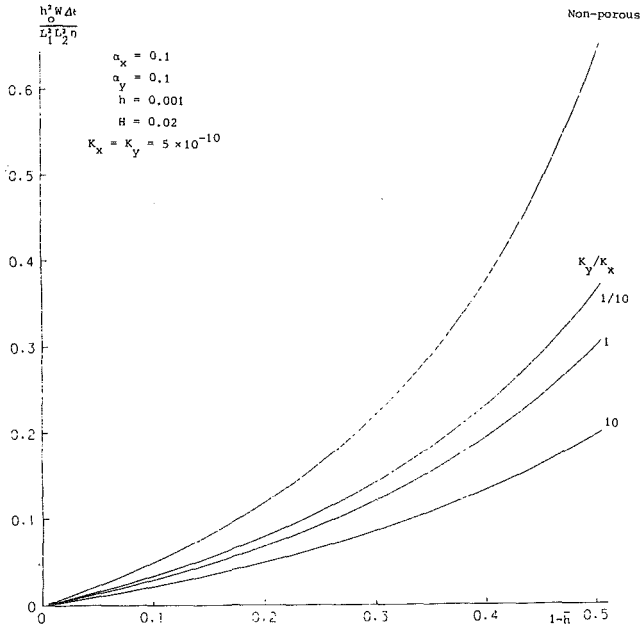


Fig. 5

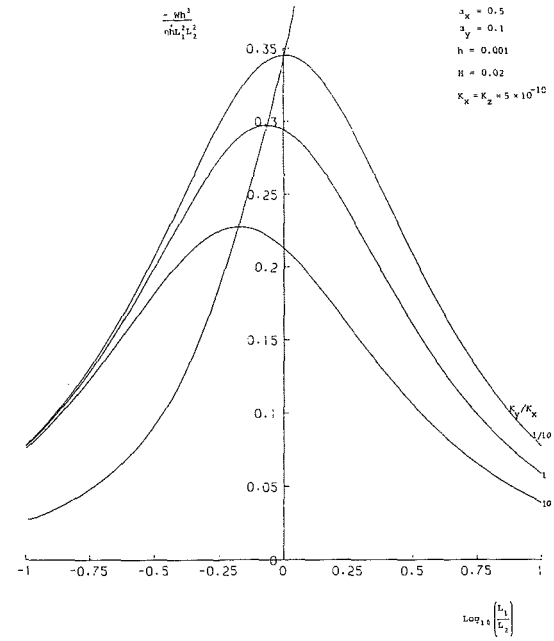


Fig. 4

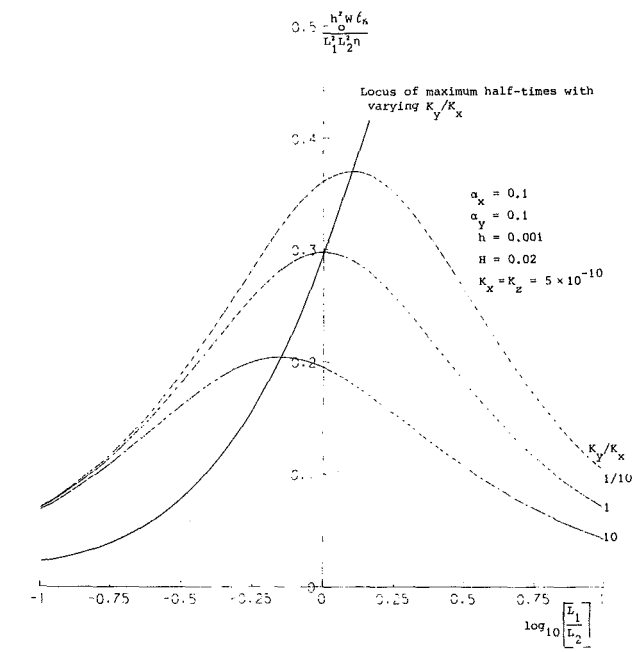


Fig. 6

Figure 5 shows that the effect of the porosity and slip constants is to alter the response time. By increasing either the ratio of K_y/K_x or the amount of slip, the response time is decreased.

The half-time of the material, $t_{1/2}$, is defined as the time required for the plates to move from a distance apart h_o to $h_o/2$. Figure 6 shows the effect of L_1/L_2 and K_y/K_x on the half-times of the material and one observes that the maximum half-time is experienced when the maximum load is carried (as expected).

Figures 3 and 4 show the significant effect of changing the slip constants α_x and α_y (which are considered as part of the anisotropy of the material) on the load-carrying capacity. Similar behavior is experienced in the case of the half-times.

Having different slip constants can either increase or decrease the amount off-square the surface must be to give maximum load-carrying capacity and corresponding maximum half-times, as well as increasing or decreasing these maximums.

In a practical situation the required shape of the bearing may be rectangular, in which case a specific choice of K_y/K_z will give maximum load-carrying capacity (remembering that in an anisotropic case the slip constants may not be identical). It must, however, also be remembered that the absolute values of the Darcy constants affect the performance of the bearing. A balance between load-carrying capacity and ability of the bearing to be self-lubricating must be reached.

Acknowledgments

The authors would like to express their gratitude to ICI and SERC for financial support and to Dr. R. S. Jones for his helpful comments.

References

Beavers, G. S., and Joseph, D. D., 1967, "Boundary Conditions at a Naturally Permeable Wall," *Journal of Fluid Mechanics*, Vol. 30, pp. 197–207.

Beavers, G. S., Sparrow, E. M., and Magnuson, J., 1970, "Experiments on a Coupled Parallel Flow in a Channel and a Bounding Porous Medium," *Journal of Basic Engineering, Trans. ASME*, Vol. 92, pp. 843–848.

Berman, A. S., 1953, "Laminar Flow in Channels with Porous Walls," *Journal of Applied Physics*, Vol. 24, pp. 1232–1235.

Booser, E. R., 1970, "Plain-Bearing Materials," *Machine Design*, Vol. 42, pp. 14–20.

Hays, D. F., 1963, "Squeeze Films for Rectangular Plates," *Journal of Basic Engineering, Trans. ASME*, Series D, Vol. 85, No. 2, pp. 243–246.

Morgan, V. T., and Cameron, A., 1957, "The Mechanism of Lubrication in Porous Metal Bearings," *Proceedings of the Conference on Lubrication and Wear*, Institute Mechanical Engineers, London, pp. 151–157.

Saffman, P. G., 1971, "On the Boundary Condition at the Surface of a Porous Medium," *Studies in Applied Mathematics*, Vol. L, pp. 93–101.

Sparrow, E. M., Beavers, G. S., and Hwang, I. T., 1971, "Effect of Velocity Slip on Porous Walled Squeeze Films," *Trans. ASME*, Paper No. 71-Lub-4, pp. 260–265.

Taylor, G. L., 1971, "A Model for the Boundary Condition of a Porous Material. Part 1," *Journal of Fluid Mechanics*, Vol. 49, pp. 319–326.

Wu, H., 1970, "Squeeze Film Behavior for Porous Annular Disks," *Journal Lubrication Tech., Trans. ASME*, Series F, Vol. 92, No. 4, pp. 593–596.

Wu, H., 1971a, "An Analysis of the Squeeze Film Between Porous Rectangular Plates," *Journal Lubrication Tech., Trans. ASME*, Paper No. 71-Lub-2.

Wu, H., 1971b, "The Squeeze Film Between Rotating Annular Porous Disks," *Wear*, Vol. 18, pp. 461–470.

Wu, H., 1972, "Effect of Velocity-Slip on the Squeeze Film Between Porous Rectangular Plates," *Wear*, Vol. 20, pp. 67–71.

Environmentally-Induced Expansion of Heterogeneous Media

Kalman Schulgasser
Pearlstone Center for Aeronautical Engineering Studies,
Department of Mechanical Engineering,
Ben-Gurion University of the Negev,
Beer Sheva, Israel
Mem. ASME

Relationships between effective expansion behavior and effective elastic constants for composite materials have been known for many years. In the present work composites are considered for which more than one environmental variable (e.g., temperature and relative humidity) cause expansions. A simple direct method to relate effective expansions due to different causes is developed. It is shown that most of the previous elasticity-expansion behavior results are gotten as corollaries, and the applicability of these relationships is broadened.

1 Introduction

Since the seminal work of Levin (1967) in which he considered the relationship between the effective elastic and expansional behavior of composite materials, other relationships of the type suggested by him have been found for a number of interesting cases (Rosen and Hashin, 1970; Laws, 1973; Hashin, 1984; Schulgasser, 1987). In obtaining such relationships, complex manipulations have often been involved which may mask the physical understanding of the problem, and which make it difficult *a priori* to know for which cases solutions can be found. It is our purpose here to consider the expansional behavior of heterogeneous media from a new point of view which leads with exceeding simplicity to new types of relationships and as a corollary to most of the results previously obtained, and makes clear under what circumstances such relationships can and cannot be found. Additionally, in Section 4, we will reestablish a result found previously by Rosen and Hashin (1970) and by Laws (1973) from an alternate point of view, which somewhat loosens the restrictions imposed in the previous developments.

Consider a homogeneous material subjected to a change in some environmental variable which results in homogeneous strains in a sample of the material. The strain resulting is

$$\epsilon_{ij} = \alpha_{ij}(T)$$

where T is the environmental variable. $\alpha_{ij}(T)$ is a characteristic function of the material. We take $\alpha_{ij}(0) = 0$. T might be temperature or humidity change, irradiation dosage, or even simply time for an aging material. We do not require a linear relationship between α_{ij} and T. The second-rank tensor function α_{ij} can always be diagonalized and we will generally henceforth describe homogeneous materials with respect to the principal expansional axes, so that the expansional behavior can be described by three functions $\alpha_1(T)$, $\alpha_2(T)$, and $\alpha_3(T)$, and we will consider only cases where the directions of the principal axes are not dependent on T. If there are two different environmental variables T and T', each of which causes expansions in a homogeneous sample of the material, we have

$$\epsilon_{ij} = \alpha_{ij}(T)$$

$$\epsilon_{ij} = \alpha'_{ij}(T').$$

Denote by $\alpha^*_{ij}(T)$ and $\alpha'^*_{ij}(T')$ the effective expansion of a composite constituted from m phases, each with expansion functions $\overset{(n)}{\alpha}_{ij}$ and $\overset{(n)}{\alpha}{}'_{ij}$ ($n = 1$ to m). We seek relationships of the form

$$\alpha'^*_{ij} = f(\alpha^*_{ij}, \overset{(n)}{\alpha}_{ij}, \overset{(n)}{\alpha}{}'_{ij}). \tag{1}$$

Having accomplished this (in those instances when such relations can indeed by found) we will see that in many cases directional compliance under pressure (S_{ijkk} in the usual notation) can replace α'_{ij} in expressions of the type (1). These are the relationships found in Levin (1967), Rosen and Hashin (1970), Laws (1973), Hashin (1984), and Schulgasser (1987).

2 Polycrystals

The utter simplicity of the proposed scheme is best illustrated for the case of the polycrystal constituted of a single species of constituent crystal. We consider crystals for which $\alpha_2 = \alpha_3$, i.e., there is expansional isotropy in the plane perpendicular to the "1" direction. Let us "disassemble" the polycrystal when $T = 0$ and $T' = 0$. Now let T take on some value other than 0. Then, in each crystal there will be expansional strain α_1 in the "1" direction and α_2 in the "2" and "3" directions. We now reassemble the polycrystal. If the constituting crystalline material is linearly elastic and the

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

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 and accepted by the Applied Mechanics Division, June 17, 1988.

Paper No. 89-WA/APM-9.

elastic constants are not dependent on T, then it is clear that the change of distance between any two points located in the polycrystal as a result of the change in environmental variable will be

$$\Delta L = a\alpha_1 + b\alpha_2 \tag{2}$$

where a and b are constants for the system depending on the geometry of the polycrystal and on the elastic constants of the constituting crystalline material. This is due simply to the linearity of the elastic boundary value problem involved in the reassembly and is true irrelevant of the crystal class of the individual crystals (as long as $\alpha_2 = \alpha_3$), and irrelevant of the nature of the macroscopic symmetry of the polycrystal. Now we perform the same experiment, this time letting T' take on some nonzero value. By the same arguments as just stated, we can write

$$\Delta L' = a\alpha_1' + b\alpha_2'. \tag{3}$$

As long as the elastic properties of the material are not dependent on T or T', a and b are the same as in (2). We can calibrate the constants a and b if we note that when $\alpha_1 = \alpha_2$, $\Delta L = \alpha_1 L$ where L is the distance between the two points. Hence,

$$a + b = L. \tag{4}$$

We now simply eliminate a and b from equations (2)–(4), and find

$$\frac{\dfrac{\Delta L}{L} - \alpha_1}{\alpha_2 - \alpha_1} = \frac{\dfrac{\Delta L'}{L} - \alpha_1'}{\alpha_2' - \alpha_1'}. \tag{5}$$

Note that (5) is valid no matter what is the elastic single crystal symmetry class, as long as $\alpha_2 = \alpha_3$ and $\alpha_2' = \alpha_3'$. Further note that it is valid no matter what is the structure of the "polycrystal." In the limit the "polycrystal" could even consist of just a few crystals with any outer boundary shape (e.g., Fig. 1). If the microstructure of the polycrystal is small compared to the sample of the polycrystal available so that statistical homogeneity can be assumed, then $\Delta L/L$ and $\Delta L'/L$ can be identified as the effective expansion functions α_d^* and $\alpha_d'^*$, and (5) becomes

$$\frac{\alpha_d^* - \alpha_1}{\alpha_2 - \alpha_1} = \frac{\alpha_d'^* - \alpha_1'}{\alpha_2' - \alpha_1'}. \tag{6}$$

If there is a linear relationship between expansion and the environmental variable, then clearly the α's in this relationship can be identified as the appropriate coefficients, e.g., thermal expansion coefficients, moisture content expansion coefficients, etc.

To identify α_1', α_2', and $\alpha_d'^*$ with compressional compliance, we simply apply a pressure p in the second experiment in place of T' after "disassembly" of the polycrystal. Then the strains in the "1", "2", and "3" directions of the individual crystals are $-S_{11kk}p$, $-S_{22kk}p$, $-S_{33kk}p$. Now reassembling while maintaining the pressure, if the principal axes of S_{ijkk} coincide with those of α_{ij} and if $S_{22kk} = S_{33kk}$ (this is true for hexagonal, tetragonal, and trigonal crystals), it is clear that equation (3) is valid with S_{11kk} and S_{22kk} replacing α_1' and α_2', respectively. Hence, equation (6) becomes

$$\frac{\alpha_d^* - \alpha_1}{\alpha_2 - \alpha_1} = \frac{S_{ddkk}^* - S_{11kk}}{S_{22kk} - S_{11kk}} \tag{7}$$

where we have identified $\Delta L'/L$ as the effective p-multiplied directional pressure compliance S_{ddkk}^* (no summation on d). This result for macroscopically isotropic polycrystals was found by Hashin (1984) and was generalized to the macroscopically anisotropic case by Schulgasser (1987). Both

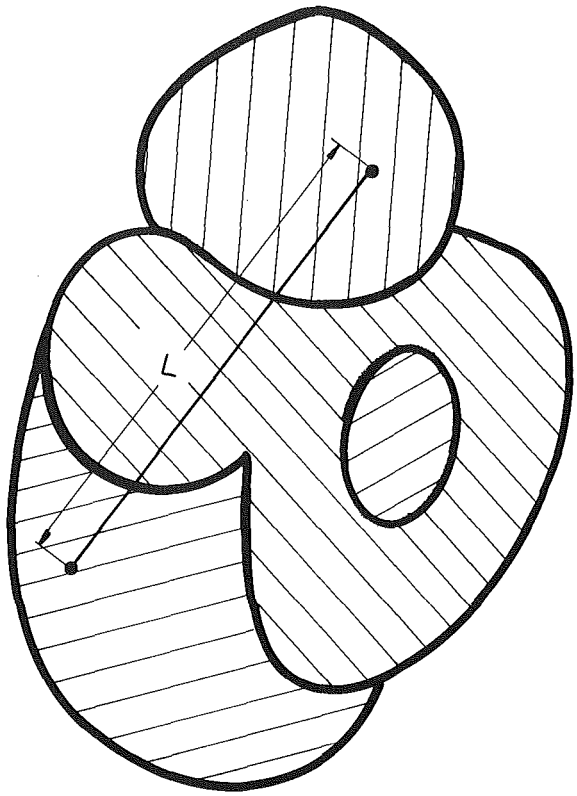


Fig. 1 A "polycrystal" composed of just a few crystals

of these works include delicate tensorial arguments and are phrased throughout in terms of the statistically homogeneous situation. The present derivation shows that (7) is simply a special case of the more general result (6), derived without recourse to any tensorial apparatus, and based simply on the linearity of the elastic boundary value problem. Indeed, harking back to the form in equation (5) and identifying α_1' and α_2' with S_{11kk} and S_{22kk} we see that change in distance between any two points in any body made up of a single constituent material with varying orientation (cf., Fig. 1) due to change of an environmental variable can be determined if the behavior of the constituent material is known, and if the distance change due to pressure applied to the outer boundary is known.

If the three principal expansions of the constituent crystal are distinct we could consider an additional environmental variable T'' resulting in the system of equations

$$\begin{aligned} \Delta L &= a\alpha_1 + b\alpha_2 + c\alpha_3 \\ \Delta L' &= a\alpha_1' + b\alpha_2' + c\alpha_3' \\ \Delta L'' &= a\alpha_1'' + b\alpha_2'' + c\alpha_3'' \\ a + b + c &= L \end{aligned} \tag{8}$$

analogous to the system (2)–(4). Then, in terms of effective expansion constants we find the relationship

$$\begin{aligned} &\frac{\alpha_d^* - \alpha_1}{\alpha_2 - \alpha_1}\left[\frac{\alpha_3' - \alpha_1'}{\alpha_2' - \alpha_1'} - \frac{\alpha_3'' - \alpha_1''}{\alpha_2'' - \alpha_1''}\right] \\ &\quad + \frac{\alpha_d'^* - \alpha_1'}{\alpha_2' - \alpha_1'}\left[\frac{\alpha_3'' - \alpha_1''}{\alpha_2'' - \alpha_1''} - \frac{\alpha_3 - \alpha_1}{\alpha_2 - \alpha_1}\right] \\ &\quad + \frac{\alpha_d''^* - \alpha_1''}{\alpha_2'' - \alpha_1''}\left[\frac{\alpha_3 - \alpha_1}{\alpha_2 - \alpha_1} - \frac{\alpha_3' - \alpha_1'}{\alpha_2' - \alpha_1'}\right] = 0. \end{aligned} \tag{9}$$

Again, α_1'', α_2'', α_3'' and $\alpha_d''^*$ could be replaced by S_{11kk}, S_{22kk}, S_{33kk}, and S^*_{ddkk}, respectively.

Considering the process used to arrive at (6) and (9), it is clear that they are valid for porous polycrystals. This is not true for (7), since the process would imply that the pressure permeates the voids.

3 Two Isotropic Phases

We consider now two-phase materials, each phase being isotropic, but the mixture may be macroscopically anisotropic. Phase expansional behavior is described for the two materials by $\overset{(1)}{\alpha}$, $\overset{(2)}{\alpha}$ and $\overset{(1)}{\alpha}{}'$, $\overset{(2)}{\alpha}{}'$ under change of the two environmental variables, T and T', respectively. Repeating the procedure of the previous section we obtain equations formally identical to (2)–(4) with $\overset{(n)}{\alpha}$, $\overset{(n)}{\alpha}{}'$, replacing α_n, α_n' in every instance. We then obtain results analogous to those in (5) and (6) with the aforementioned replacement. And analogous to equation (7) we obtain

$$\frac{\alpha_d^* - \overset{(1)}{\alpha}}{\overset{(2)}{\alpha} - \overset{(1)}{\alpha}} = \frac{S^*_{ddkk} - \dfrac{1}{\overset{(1)}{K}}}{\dfrac{1}{\overset{(2)}{K}} - \dfrac{1}{\overset{(1)}{K}}} \tag{10}$$

where $\overset{(1)}{K}$ and $\overset{(2)}{K}$ are the bulk moduli of the two phases. This result was explicity given by Rosen and Hashin (1970). The implications of equation (10) for the case of aligned isotropic fibers in an isotropic matrix, using the unique relationship found by Hill (1964) which exists between the various effective composite moduli, has been discussed by Hashin and Rosen (1970) and Dvorak and Bahei-El-Din (1981). For the macroscopically isotropic case, S^*_{ddkk} is replaced by $1/K^*$. This result was found by several investigators using various methods (Levin, 1967; Schapery, 1968; Steel, 1968; Cribb, 1968).

For a three-phase material the procedure leads to a result analogous to that in equation (9), again with the replacements as previously indicated. We can even consider a "two-phase" material, one phase being crystalline of random orientations with $\overset{(1)}{\alpha}_2 = \overset{(1)}{\alpha}_3$, $\overset{(1)}{\alpha}_2{}' = \overset{(1)}{\alpha}_3{}'$, $\overset{(1)}{\alpha}_2{}'' = \overset{(1)}{\alpha}_3{}''$ and the other phase being isotropic with expansion behaviors $\overset{(2)}{\alpha}$, $\overset{(2)}{\alpha}{}'$, $\overset{(2)}{\alpha}{}''$. Then, considering the equations of the form (8) which we would write, it is clear that a relationship analogous to (9) can be found with the superscript (1) added to variables subscripted with 2 or 3, and α_1, α_1', α_1'' replaced by $\overset{(2)}{\alpha}$, $\overset{(2)}{\alpha}{}'$, and $\overset{(2)}{\alpha}{}''$.

As pointed out for polycrystals, also the two isotropic-phase result analogous to equation (6) and the three-phase result analogous to equation (9), are valid for porous composites; however, equation (10) is not valid.

4 Two Anisotropic Phases

We consider the case of two anisotropic phases; their compliances are described by the Cartesian tensors $\overset{(1)}{S}_{ijkl}$ and $\overset{(2)}{S}_{ijkl}$, and their expansions by the functions $\overset{(1)}{\alpha}_{ij}(T)$ and $\overset{(2)}{\alpha}_{ij}(T)$. We emphasize that the orientation of each phase is fixed in space. For this case it is apparently not possible to find a relationship of the form (1). However, α_{ij}^* has been found as a function of S^*_{ijkl} and the phase elastic and expansional properties. This was accomplished first by Rosen and Hashin (1970) and later by Laws (1973). We will rederive here this result using a generalization of a method applied by Cribb (1968) to isotropic phases and later by Dvorak (1986) to anisotropic phases, which will show that certain of the restrictions imposed in Rosen and Hashin (1970) and Laws (1973) can be relaxed.

Again, we begin by disassembling the composite. Then we impose on the boundary of each phase the traction

$$T_i = \tau_{ij}^0 n_j, \tag{11}$$

where n_i is the outward unit vector normal to the phase surface, while changing the environmental constant to some nonzero value. τ_{ij}^0 is a constant tensor. These tractions result in uniform stress τ_{ij}^0 throughout, and the strain field in the first phase is

$$\overset{(1)}{\epsilon}_{ij} = \tau_{kl}^0 \overset{(1)}{S}_{ijkl} + \overset{(1)}{\alpha}_{ij}, \tag{12}$$

and in the second phase

$$\overset{(2)}{\epsilon}_{ij} = \tau_{kl}^0 \overset{(2)}{S}_{ijkl} + \overset{(2)}{\alpha}_{ij}. \tag{13}$$

If the strains in the two phases were identical, one could reassemble the composite, the interface condition would be satisfied, and no further stress is caused in the body. Equating the right-hand sides of (12) and (13) we find the condition for identical strain fields

$$\tau_{kl}^0 (\overset{(1)}{S}_{ijkl} - \overset{(2)}{S}_{ijkl}) = \overset{(2)}{\alpha}_{ij} - \overset{(1)}{\alpha}_{ij} \tag{14}$$

or

$$\tau_{kl}^0 = (\overset{(2)}{\alpha}_{ij} - \overset{(1)}{\alpha}_{ij}) P_{klij}, \tag{15}$$

where P_{ijkl} is the reciprocal of $(\overset{(1)}{S}_{ijkl} - \overset{(2)}{S}_{ijkl})$ given by

$$P_{ijkl}(\overset{(1)}{S}_{klmn} - \overset{(2)}{S}_{klmn}) = I_{ijmn}, \tag{16}$$

where I_{ijmn} is the identity tensor

$$I_{ijmn} = \frac{1}{2}(\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}). \tag{17}$$

We now have a composite body with strains throughout given by (12) or (13), with tractions $\tau_{ij}^0 n_j$ on the outer envelope. If we now apply the traction system $-\tau_{ij}^0 n_j$ to the outer envelope, and if we can solve the relevant boundary value problem, we have a solution for the deformation throughout the composite body when the environmental variable changes to T, i.e., $\overset{(1)}{\epsilon}_{ij}$ or $\overset{(2)}{\epsilon}_{ij}$ as given by equation (12) or (13) plus the strains gotten when the traction $-\tau_{ij}^0 n_j$ is applied, τ_{ij}^0 being taken from (15). Now if the body is statistically homogeneous, then the average strain in the body due to traction applied to the outer envelope is $-\tau_{kl}^0 S^*_{ijkl}$ where S^*_{ijkl} is the effective compliance of the composite (see Hashin (1983)). Adding this to the uniform strain throughout (equation (12) or (13)), we have

$$\text{Average strain} = \tau_{kl}^0 (\overset{(1)}{S}_{ijkl} - S^*_{ijkl}) + \overset{(1)}{\alpha}_{ij}. \tag{18}$$

This is the effective expansion α_{ij}^*, and using equation (15)

$$\alpha_{ij}^* = (\overset{(2)}{\alpha}_{mm} - \overset{(1)}{\alpha}_{mm}) P_{klmn} (\overset{(2)}{S}_{ijkl} - S^*_{ijkl}) + \overset{(1)}{\alpha}_{ij}. \tag{19}$$

It is clear that this equation can be written with superscripts (1) and (2) interchanged. Then, adding the two forms we get the form given in Rosen and Hashin (1970) and Laws (1973).

It should be emphasized that $\alpha_{ij}(T)$ need not be a linear function of T, and that the present technique permits calculation of any deformation in the composite body, even when statistical homogeneity cannot be assumed if that deformation due to the traction system $-\tau_{ij}^0 n_j$ can be found.

It is clear from the context of the current development that equation (19) can reasonably be applied to the case of aligned carbon fibers in isotropic matrix composites even though it is well known that the fiber itself has a decidedly nonuniform structure. However, generally the structure of such fibers is more or less radially symmetric (Hughes, 1987). Hence, the procedure implied by the equating of equations (12) and (13) will still be valid if S_{ijkl} of the fiber is interpreted as a measure

of overall radial axial compliance when we execute the disassembly-reassembly procedure. In fact, the result could have been achieved by the method of Sections 2 and 3 if the disassembly-reassembly procedure was carried out with first axial and then transverse suppression of deformation. The aforementioned comment is true also for boron and aramid fibers which, while not homogeneous, nevertheless exhibit radial symmetry.

Closure

The present approach makes it clear that in a heterogeneous material for which the description of the constituent materials requires r distinct expansion functions, a relation can be found between r effective linear expansion functions of the heterogeneous material, regardless of the macroscopic symmetry and under the condition that the constituents are linearly elastic with the elastic behavior not dependent on the environmental variable. The present approach also makes clear that the relationships which have been found are essentially independent of arguments based on statistical homogenity. The result given in equation (6) was previously reported by this author (Schulgasser, 1986) for the case of the statistically isotropic polycrystal and for the case of two isotropic phases, but was derived there as a corollary to the compliance-expansion relationships of the type (7) found previously by other researchers. Note that the validity of equation (6) for polycrystals and its analogous form for two isotropic phases implies that for a composite with constitutent material expansions not linearly dependent on an environmental variable, knowledge of the effective expansion behavior at one value of the environmental variable T immediately implies knowledge of the effective expansion for all other values of T if the constituent expansions are known as functions of T.

In the case of two anisotropic phases it is possible to derive the relationship (19), only because tractions of the form (11) applied to the outer boundary can be found which, for any given value of T, results in uniform strain throughout. However, in this case it is not possible to find a relationship between the expansions resulting from different environmental variables without explicitly involving the elastic behavior of the constitutents and of the composite.

References

Cribb, J. L., 1968, "Shrinkage and Thermal Expansion of a Two-Phase Material," *Nature*, London, Vol. 220, pp. 576–577.

Dvorak, G. J., and Bahei-El-Din, Y. A., 1979, "Elastic-Plastic Behavior of Fibrous Composites," *Journal of the Mechanics and Physics of Solids*, Vol. 27, (1979), pp. 51–72; Erratum, 1981, Vol. 29, p. 267.

Dvorak, G. J., 1986, "Thermal Expansion of Elastic-Plastic Composite Materials," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 737–743.

Hashin, Z., 1983, "Analysis of Composite Materials—A Survey," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 481–505.

Hashin, Z., 1984, "Thermal Expansion of Polycrystalline Aggregates: I. Exact Results," *Journal of the Mechanics and Physics of Solids*, Vol. 32, pp. 149–157.

Hill, R., 1964, "Theory of Mechanical Properties of Fibre-Strengthened Materials: I. Elastic Behavior," *Journal of the Mechanics and Physics of Solids*, Vol. 12, pp. 199–212.

Hughes, J. D. H., 1987, "The Evaluation of Current Carbon Fibres," *Journal of Physics, D, Applied Physics*, Vol. 20, pp. 276–285.

Laws, N., 1973, "On the Thermostatics of Composite Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 21, pp. 9–17.

Levin, V. M., 1967, "Thermal Expansion of Heterogeneous Materials," *Mekhanika Tverdoga Tela*, Vol. 2, pp. 88–94.

Rosen, B. W., and Hashin, Z., 1970, "Effective Thermal Expansion Coefficients and Specific Heats of Composite Materials," *International Journal of Engineering Science*, Vol. 8, pp. 157–173.

Schapery, R. A., 1968, "Thermal Expansion Coefficients of Composite Materials Based on Energy Methods," *Journal of Composite Materials*, Vol. 2, pp. 380–404.

Schulgasser, K., 1986, "Thermal, Hygroscopic, and Irradiation Expansion of Composite Materials," *Journal of Materials Science Letters*, Vol. 5, pp. 719–720.

Schulgasser, K., 1987, "Thermal Expansion of Polycrystalline Aggregates with Texture," *Journal of the Mechanics and Physics of Solids*, Vol. 35, pp. 35–42.

Steel, T. R., 1968, "Determination of the Constitutive Coefficients for a Mixture of Two Solids," *International Journal of Solids and Structures*, Vol. 4, pp. 1149–1160.

A Green's Function Formulation of Anticracks and Their Interaction With Load-Induced Singularities

John Dundurs
Professor of Civil Engineering and Mechanical Engineering, Northwestern University, Evanston, Ill. 60208
Fellow ASME

Xanthippi Markenscoff
Professor of Applied Mechanics, University of California, San Diego, La Jolla, Calif. 92093
Mem. ASME

This paper provides a Green's function formulation of anticracks (rigid lamellar inclusions of negligible thickness that are bonded to the surrounding elastic material). Apart from systematizing several previously known solutions, the article gives the pertinent fields for concentrated forces, dislocations, and a concentrated couple applied on the line of the anticrack. There is a reason for working out these results: In contrast to concentrated forces, a concentrated couple approaching the tip of an anticrack makes the elastic fields explode. Finite limits can be achieved, however, by appropriately diminishing the magnitude of the couple, which then leads to fields that are intimately connected with the weight functions for the anticrack. An edge dislocation going to the tip of an anticrack puts a net force on the lamellar inclusion, which is shown to be related to a previously known feature of dislocations near a bimaterial interface.

1 Introduction

The opposite of a crack, in a certain sense, is a cut in the material that is filled with a rigid lamella: A crack is a cut that transmits no tractions, but allows a displacement discontinuity. The rigid lamella transmits tractions, but prevents a displacement discontinuity. There is no uniform terminology for the latter, and we shall call them anticracks for brevity. In spite of the fact that anticracks do not have applications that are as far ranging as those of cracks, there is a considerable amount of literature on the topic. A fairly complete list of references can be compiled from the papers by Atkinson (1973), Brussat and Westmann (1975), Hasebe, Keer, and Nemat-Nasser (1984), Hasebe, Nemat-Nasser, and Keer (1984), Wang, Zhang, and Chou (1985), and Mura (1988).

The objective of this paper is to provide a direct Green's function forumulation of anticracks. Such a formulation allows one to write the governing integral equations practically at sight, and it is suitable for solution by current numerical methods. The paper also gives some new solutions. The results for edge dislocations show some unexpected features. Similarly, an anomalous behavior is discovered in considering the interaction of a concentrated couple and the anticrack.

2 Singular Nature of the Fields Induced by Concentrated Forces and Line Loads

The appropriate Green's functions for the anticrack simply are concentrated forces. However, the singular nature of their fields must be explored in some detail before they can be applied directly to the formulation. In doing this, repeated use may be made of the following theorem on the Dirac delta function in Stakgold's book (1968).

Let $f(x)$ be a non-negative, locally integrable function for which

$$\int_{-\infty}^{\infty} f(x)dx = 1. \tag{1}$$

With $\alpha > 0$, define

$$f_\alpha(x) = \frac{1}{\alpha} f\left(\frac{x}{\alpha}\right). \tag{2}$$

Then

$$\lim_{\alpha \to 0} f_\alpha(x) = \delta(x). \tag{3}$$

The displacement and stress components for a concentrated force, P_x, applied at the origin and acting in the x-direction are (Timoshenko and Goodier, 1970):

$$2\mu u_x(x,y) = \frac{P_x}{\pi(\kappa+1)}\left\{-\kappa \log r + \frac{x^2}{r^2}\right\} \tag{4}$$

$$2\mu u_y(x,y) = \frac{P_x}{\pi(\kappa+1)}\frac{xy}{r^2} \tag{5}$$

$$\sigma_{xx}(x,y) = \frac{P_x}{2\pi(\kappa+1)}\left\{-(\kappa-1)\frac{x}{r^2} - 4\frac{x^3}{r^4}\right\} \tag{6}$$

$$\sigma_{xy}(x,y) = \frac{P_x}{2\pi(\kappa+1)}\left\{-(\kappa-1)\frac{y}{r^2} - 4\frac{x^2 y}{r^4}\right\} \tag{7}$$

$$\sigma_{yy}(x,y) = \frac{P_x}{2\pi(\kappa+1)}\left\{(\kappa-1)\frac{x}{r^2} - 4\frac{xy^2}{r^4}\right\} \tag{8}$$

Contributed by the Applied Mechanics Division of The American Society of Mechanical Engineers for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the Journal of Applied Mechanics. Manuscript received by the ASME Applied Mechanics Division, May 19, 1988; final revision, March 3, 1989.

Paper No. 89-WA/APM-15.

where μ is the shear modulus and, with ν denoting Poisson's ratio, $\kappa = 3-4\nu$ for plane strain and $\kappa = (3-\nu)/(1+\nu)$ for plane stress. Of interest in the present context are the limits for $y \to 0$ of the following quantities:

$$2\mu \frac{\partial u_x(x,0)}{\partial x} = -\frac{P_x \kappa}{\pi(\kappa+1)} \frac{1}{x}, \quad 2\mu \frac{\partial u_y(x,0)}{\partial x} = 0 \tag{9}$$

$$\sigma_{xx}(x,0) = -\frac{P_x(3+\kappa)}{2\pi(\kappa+1)} \frac{1}{x} \tag{10}$$

$$\sigma_{xy}(x,0\pm) = \mp \frac{1}{2} P_x \delta(x) \tag{11}$$

$$\sigma_{yy}(x,0) = \frac{P_x(\kappa-1)}{2\pi(\kappa+1)} \frac{1}{x}. \tag{12}$$

It follows immediately from these expressions that a line load with a local density $p_x(x)$ acting on the x-axis in the x-direction gives

$$2\mu \frac{\partial u_x(x,0)}{\partial x} = \frac{\kappa}{\pi(\kappa+1)} \int_{-\infty}^{+\infty} \frac{p_x(\xi)\,d\xi}{\xi - x},$$

$$2\mu \frac{\partial u_y(x,0)}{\partial x} = 0 \tag{13}$$

$$\sigma_{xx}(x,0) = \frac{3+\kappa}{2\pi(\kappa+1)} \int_{-\infty}^{+\infty} \frac{p_x(\xi)\,d\xi}{\xi - x} \tag{14}$$

$$\sigma_{xy}(x,0\pm) = \mp \frac{1}{2} p_x(x) \tag{15}$$

$$\sigma_{yy}(x,0) = -\frac{\kappa-1}{2\pi(\kappa+1)} \int_{-\infty}^{+\infty} \frac{p_x(\xi)\,d\xi}{\xi - x} \tag{16}$$

where the integrals are to be evaluated in the sense of Cauchy principal values.

The field quantities for a concentrated force P_y acting in the y-direction can be obtained by a rotation of coordinates. The limits of interest for $y \to 0$ are

$$2\mu \frac{\partial u_x(x,0)}{\partial x} = 0, \quad 2\mu \frac{\partial u_y(x,0)}{\partial x} = -\frac{P_y \kappa}{\pi(\kappa+1)} \frac{1}{x} \tag{17}$$

$$\sigma_{xx}(x,0\pm) = \mp \frac{P_y(3-\kappa)}{2(\kappa+1)} \delta(x) \tag{18}$$

$$\sigma_{xy}(x,0) = -\frac{P_y(\kappa-1)}{2\pi(\kappa+1)} \frac{1}{x} \tag{19}$$

$$\sigma_{yy}(x,0\pm) = \mp \frac{1}{2} P_y \delta(x). \tag{20}$$

Then, a line load of intensity $p_y(x)$ acting on the x-axis in the y-direction gives

$$2\mu \frac{\partial u_x(x,0)}{\partial x} = 0, \quad 2\mu \frac{\partial u_y(x,0)}{\partial x} = \frac{\kappa}{\pi(\kappa+1)} \int_{-\infty}^{\infty} \frac{p_y(\xi)\,d\xi}{\xi - x} \tag{21}$$

$$\sigma_{xx}(x,0\pm) = \mp \frac{3-\kappa}{2(\kappa+1)} p_y(x) \tag{22}$$

$$\sigma_{xy}(x,0) = \frac{\kappa-1}{2\pi(\kappa+1)} \int_{-\infty}^{\infty} \frac{p_y(\xi)\,d\xi}{\xi - x} \tag{23}$$

$$\sigma_{yy}(x,0\pm) = \mp \frac{1}{2} p_y(x). \tag{24}$$

3 Boundary Conditions and Formulation

An anticrack, as any other rigid inclusion, restricts the displacements on the interval it occupies to be those of a rigid body, or to be of the form

$$u_x(x,y) = \alpha - \omega y, \quad u_y(x,y) = \beta + \omega x. \tag{25}$$

Consider for simplicity a single anticrack on the interval $|x| < a$, $y = 0$. In such a case, the rigid body displacements may be differentiated with respect to x without loss of generality, yielding the boundary conditions

$$\frac{\partial u_x(x,0)}{\partial x} = 0, \quad \frac{\partial u_y(x,0)}{\partial x} = \omega, \quad |x| < a. \tag{26}$$

Suppose that the applied loads induce in the body without the anticrack displacements that yield

$$\frac{2\mu(\kappa+1)}{\kappa} \frac{\partial u_x(x,0)}{\partial x} = -f_x(x),$$

$$\frac{2\mu(\kappa+1)}{\kappa} \frac{\partial u_y(x,0)}{\partial x} = -f_y(x), \quad |x| < a. \tag{27}$$

The boundary conditions (26) can be enforced by distributing line loads with densities $p_x(x)$ and $p_y(x)$ on the interval $|x| < a$. Using (13) and (21), this yields the Cauchy integral equations

$$\frac{1}{\pi} \int_{-a}^{a} \frac{p_x(\xi)\,d\xi}{\xi - x} = f_x(x), \quad |x| < a \tag{28}$$

$$\frac{1}{\pi} \int_{-a}^{a} \frac{p_y(\xi)\,d\xi}{\xi - x} = C + f_y(x), \quad |x| < a \tag{29}$$

where

$$C = [2\mu(\kappa+1)/\kappa]\omega \tag{30}$$

and ω is the rotation of the anticrack.

As it could be anticipated from (25), this system still retains three unknown constants. In addition to C in (29), they are the two free multipliers of the homogeneous solutions of the integral equations. The three constants can be determined, however, from the global equilibrium conditions of the anticrack:

$$\int_{-a}^{a} p_x(x)\,dx = F_x \tag{31}$$

$$\int_{-a}^{a} p_y(x)\,dx = F_y \tag{32}$$

$$\int_{-a}^{a} x p_y(x)\,dx = M \tag{33}$$

where F_x and F_y are the forces, and M the couple that are applied at the center of the rigid lamella by an outside agent.

4 Simple Solutions

The simplest solutions correspond to the anticrack being subjected to forces and a couple that are applied directly to the rigid lamella, and the anticrack disturbing, uniform stress fields (the anticrack induces no disturbance in a field of uniform shear). Although some of these solutions have appeared in the literature, they are compiled in Appendix A for the sake of completeness.

However, the simple anticrack solutions have remarkable features that have not been noted before.

(1) Setting $\kappa = -1$ (this corresponds to $\nu = 1$ for plane strain, and $\nu = \infty$ for plane stress) in (A18-A20) yields

$$\sigma_{xx}(x,0) = T_y \left\{ \frac{|x| H(|x|-a)}{(x^2-a^2)^{1/2}} - 1 \right\}, \quad |x| < \infty \tag{34}$$

$$\sigma_{xy}(x,0) = 0, \quad |x| < \infty \tag{35}$$

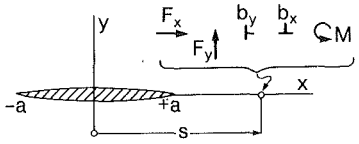


Fig. 1 Various concentrated actions near the tip of an anticrack

$$\sigma_{yy}(x,0)=T_y\frac{|x|H(|x|-a)}{(x^2-a^2)^{1/2}},\quad |x|<\infty \tag{36}$$

where H denotes the Heaviside step function. It is seen that these are precisely the results for a Griffith crack.

(2) Setting now $\kappa=-1$ in (A14–A16) gives

$$\sigma_{xx}(x,0)=T_x,\ \sigma_{xy}(x,0)=\sigma_{yy}(x,0)=0,\quad |x|<\infty \tag{37}$$

which again is true for a Griffith crack that is parallel to the direction of applied tension.

(3) Replacing $F_x/(\kappa+1)$ by $-2\mu\Delta_y/(\kappa+1)$ in (A2–A4) and setting $\kappa=-1$ elsewhere in the expressions gives

$$\sigma_{xx}(x,0)=\sigma_{yy}(x,0)=\frac{2\mu\Delta_y}{\pi(\kappa+1)}\frac{\mathrm{sgn}xH(|x|-a)}{(x^2-a^2)^{1/2}},$$

$$\sigma_{xy}(x,0)=0,\ |x|<\infty. \tag{38}$$

These are precisely the results for a Zener-Stroh crack with the Burgers vector Δ_y (Weertman, 1986).

(4) Replacing $F_y/(\kappa+1)$ by $2\mu\Delta_x/(\kappa+1)$ in (A6–A8) and again setting $\kappa=-1$ elsewhere, results in

$$\sigma_{xx}(x,0\pm)=\mp\frac{4\mu\Delta_x}{\pi(\kappa+1)}\frac{H(a-|x|)}{(a^2-x^2)^{1/2}},\ |x|<\infty \tag{39}$$

$$\sigma_{xy}(x,0)=\frac{2\mu\Delta_x}{\pi(\kappa+1)}\frac{\mathrm{sgn}xH(|x|-a)}{(x^2-a^2)^{1/2}},\ |x|<\infty \tag{40}$$

$$\sigma_{yy}(x,0)=0,\ |x|<\infty \tag{41}$$

corresponding to a Bullough-Gilman crack with the Burgers vector Δ_x (Bullough, 1964; Tucker, 1973).

The stress fields for anticracks depend on Poisson's ratio and, within the physical range $0\le\nu\le1/2$, they satisfy, of course, the boundary conditions appropriate to the anticracks. The curious result here is that these stress fields for a specific Poisson's ratio outside the physical range satisfy the boundary conditions of a different problem that is back in the physical realm. Why this is so can be traced through the equations and boundary conditions of elasticity (Dundurs 1968; Dundurs, 1970; Dundurs 1989).

5 Concentrated Force and Edge Dislocation—Symmetric Problem

The solutions for concentrated forces and edge dislocations are basic because they provide new Green's functions.

Consider a concentrated force F_x in the direction of the anticrack or an edge dislocation with the Burgers vector b_y (extra sheet of material on the x-axis in the negative direction) acting at the point $x=s$, $y=0$ (see Fig. 1). On the basis of (27)–(29), both problems lead to the same integral equation

$$\int_{-a}^{a}\frac{p_x(\xi)d\xi}{\xi-x}=\frac{L}{x-s},\quad |x|<a \tag{42}$$

where $L=F_x$ for the concentrated force, and $L=-b_y\mu(\kappa-1)/\kappa$ for the edge dislocation. Imposing the side condition

$$\int_{-a}^{a}p_x(x)dx=0 \tag{43}$$

for equilibrium of the lamella, the solution of (42) is

$$p_x(x)=\frac{L}{\pi(a^2-x^2)^{1/2}}\left\{1-\frac{(s^2-a^2)^{1/2}}{s-x}\right\},\quad |x|<a. \tag{44}$$

The integral

$$\int_{-a}^{a}\frac{p_x(\xi)d\xi}{\xi-x}=L\left\{\frac{1}{x-s}-\left[1+\frac{(s^2-a^2)^{1/2}}{x-s}\right]\frac{\mathrm{sgn}xH(|x|-a)}{(x^2-a^2)^{1/2}}\right\},\quad |x|<\infty \tag{45}$$

is needed for the computation of the stresses using (14)–(16). It should be noted that (14)–(16) give only the stresses induced by the line load, and that the stresses of the force or edge dislocation in a homogeneous material must be added to get the total stresses.

Introducing the contraction

$$g(x)=\frac{1}{|x^2-a^2|^{1/2}}\left\{1+\frac{(s^2-a^2)^{1/2}}{x-s}\right\}, \tag{46}$$

the total stresses for the concentrated force are

$$\sigma_{xx}(x,0)=-\frac{F_x(3+\kappa)}{2\pi(\kappa+1)}g(x)\mathrm{sgn}xH(|x|-a),\quad |x|<\infty \tag{47}$$

$$\sigma_{xy}(x,0\pm)=\mp\frac{F_x}{2\pi}\{\pi\delta(x-s)+g(x)H(a-|x|)\},\quad |x|<\infty \tag{48}$$

$$\sigma_{yy}(x,0)=\frac{F_x(\kappa-1)}{2\pi(\kappa+1)}g(x)\mathrm{sgn}xH(|x|-a),\quad |x|<\infty. \tag{49}$$

Similarly, the total stresses for the dislocation are

$$\sigma_{xx}(x,0)=\frac{b_y\mu}{2\pi\kappa}\left\{\frac{3-\kappa}{x-s}+\frac{(\kappa-1)(3+\kappa)}{\kappa+1}g(x)\mathrm{sgn}xH(|x|-a)\right\},\quad |x|<\infty \tag{50}$$

$$\sigma_{xy}(x,0\pm)=\mp\frac{b_y\mu(\kappa-1)}{2\pi\kappa}g(x)H(a-|x|),\quad |x|<\infty \tag{51}$$

$$\sigma_{yy}(x,0)=\frac{b_y\mu}{2\pi\kappa}\left\{\frac{\kappa+1}{x-s}-\frac{(\kappa-1)^2}{\kappa+1}g(x)\mathrm{sgn}xH(|x|-a),\quad |x|<\infty.\right. \tag{52}$$

The Peach-Koehler force on the dislocation is (Weertman and Weertman, 1964)

$$K_x=b_y\sigma_{yy}(s,0),\quad a<s, \tag{53}$$

and it acts in the x-direction. However, in (53) only the contribution to the stress by the line load must be used. The result is

$$K_x=\frac{b_y^2\mu(\kappa-1)^2}{2\pi\kappa(\kappa+1)}\frac{s-(s^2-a^2)^{1/2}}{s^2-a^2}. \tag{54}$$

Since $K_x>0$, the anticrack repells the dislocation.

Finally, consider the limit $s\to0$. For the concentrated force, (44) becomes the same as (A1), which is as expected. For the dislocation, however, (44) yields

$$p_x(x)=-\frac{b_y\mu(\kappa-1)}{\pi\kappa}\frac{1}{(a^2-x^2)^{1/2}},\quad |x|<a \tag{55}$$

and

$$\int_{-a}^{a}p_x(x)dx=-\frac{b_y\mu(\kappa-1)}{\kappa}. \tag{56}$$

This means that, in the limit $s\to a$, the dislocation with a positive Burgers vector b_y exerts a net force on the lamella that is in the negative x-direction (positive $p_x(x)$ acts on the elastic material, negative $p_x(x)$ acts on the rigid lamella). The fact that a dislocation, which is a self-equilibrated singularity, exerts a force (a real force in the sense of Newton, not a driving force in the sense of Peach and Koehler) on the rigid lamella may seem contradictory. It is explained in Appendix B why this is actually very reasonable.

6 Concentrated Force and Edge Dislocation—Antisymmetric Problem

Now the concentrated force F_y acts in a direction perpendicular to the anticrack and the Burgers vector b_x is in the x-direction (extra sheet of material perpendicular to the x-axis and on the side of positive y). The governing integral equation from (27)–(29) is then

$$\int_{-a}^{a}\frac{p_y(\xi)d\xi}{\xi-x}=N\left(C+\frac{1}{x-s}\right),\quad |x|<a. \tag{57}$$

For the force, $N=F_y$, and for the dislocation $N=b_x\mu(\kappa-1)/\kappa$. The solution of (58) under the side conditions

$$\int_{-a}^{a}p_y(x)dx=0,\quad \int_{-a}^{a}xp_y(x)dx=0 \tag{58}$$

is

$$p_y(x)=\frac{N}{\pi(a^2-x^2)^{1/2}}\left\{1+Ax-\frac{(s^2-a^2)^{1/2}}{s-x}\right\},\quad |x|<a \tag{59}$$

where

$$A=\frac{2}{a^2}[s-(s^2-a^2)^{1/2}]. \tag{60}$$

Moreover,

$$\int_{-a}^{a}\frac{p_y(\xi)d\xi}{\xi-x}=N\left\{A+\frac{1}{x-s}-\left[1+Ax+\frac{(s^2-a^2)^{1/2}}{x-s}\right]\frac{\operatorname{sgn}xH(|x|-a)}{(x^2-a^2)^{1/2}}\right\},\quad |x|<\infty. \tag{61}$$

Proceeding as in the previous case, the total stresses for the concentrated force are

$$\sigma_{xx}(x,0\pm)=\mp\frac{F_y(3-\kappa)}{2\pi(\kappa+1)}\{\pi\delta(x-s)+h(x)H(a-|x|)\},\quad |x|<\infty \tag{62}$$

$$\sigma_{xy}(x,0)=\frac{F_y(\kappa-1)}{2\pi(\kappa+1)}\{A-h(x)\operatorname{sgn}xH(|x|-a)\},\quad |x|<\infty \tag{63}$$

$$\sigma_{yy}(x,0\pm)=\mp\frac{F_y}{2\pi}\{\pi\delta(x-s)+h(x)H(a-|x|)\},\quad |x|<\infty \tag{64}$$

where

$$h(x)=\frac{1}{|x^2-a^2|^{1/2}}\left\{1+Ax+\frac{(s^2-a^2)^{1/2}}{x-s}\right\},\quad |x|<\infty. \tag{65}$$

For the dislocation,

$$\sigma_{xx}(x,0\pm)=\mp\frac{b_x\mu}{\pi(\kappa+1)}\{4\pi\delta(x-s)+\frac{(\kappa-1)(3-\kappa)}{2\kappa}h(x)H(a-|x|)\},\quad |x|<\infty \tag{66}$$

$$\sigma_{xy}(x,0)=\frac{b_x\mu}{2\pi\kappa}\left\{\frac{\kappa+1}{x-s}+\frac{(\kappa-1)^2}{\kappa+1}[A-h(x)\operatorname{sgn}xH(|x|-a)]\right\},\quad |x|<\infty \tag{67}$$

$$\sigma_{yy}(x,0\pm)=\mp\frac{b_x\mu(\kappa-1)}{2\pi\kappa}h(x)H(a-|x|),\quad |x|<\infty. \tag{68}$$

The Peach-Koehler force on the dislocation is

$$K_x=b_x\sigma_{xy}(s,0)=\frac{b_x^2\mu(\kappa-1)^2}{2\pi\kappa(\kappa+1)}k(s) \tag{69}$$

where

$$k(s)=\frac{s-(s^2-a^2)^{1/2}}{s^2-a^2}\left\{1+\frac{2}{a^2}[s^2-a^2-s(s^2-a^2)^{1/2}]\right\}. \tag{70}$$

It can be reasoned from (70) that the anticrack repels the dislocation in all positions.

In the limit $s\to a$,

$$p_y(x)=\frac{N}{\pi(a^2-x^2)^{1/2}}\left(1+\frac{2x}{a}\right),\quad x<a \tag{71}$$

which no longer satisfies (58). For the concentrated force, (71) can be reconciled with (*A*5) and (*A*9). For the dislocation the result means that, in the limit $s\to a$, the dislocation exerts a force $b_x\mu(\kappa-1)/\kappa$ that acts on the tip of the anticrack in the y-direction. Again, this result can be explained on the basis of the discussion in Appendix B and equation (*B*9).

7 Concentrated Couple

It is also interesting to consider a concentrated couple with a moment M that is acting at the point $x=s$, $y=0$ (see Fig. 1). From the known solution for a concentrated couple acting at an interior point (Timoshenko and Goodier, 1970), it follows that the couple gives

$$2\mu\frac{\partial u_y(x,0)}{\partial x}=-\frac{M}{2\pi}\frac{1}{(x-s)^2} \tag{72}$$

which, then on the basis of (29), yields the integral equation

$$\int_{-a}^{a}\frac{p_y(\xi)d\xi}{\xi-x}=\frac{M(\kappa+1)}{2\kappa}\left\{C+\frac{1}{(x-s)^2}\right\}. \tag{73}$$

The solution of (73) under the side conditions (58) is

$$p_y(x)=\frac{M(\kappa+1)}{2\pi\kappa(a^2-x^2)^{1/2}}\left\{Bx-\frac{a^2-sx}{(s^2-a^2)^{1/2}(s-x)^2}\right\},\quad |x|<a \tag{74}$$

$$B=-\frac{2[s-(s^2-a^2)^{1/2}]}{a^2(s^2-a^2)^{1/2}}. \tag{75}$$

Consequently,

$$\int_{-a}^{a}\frac{p_y(\xi)d\xi}{\xi-x}=\frac{M(\kappa+1)}{2\kappa}\left\{\frac{1}{(x-s)^2}+B-j(x)\operatorname{sgn}xH(|x|-a)\right\},\quad |x|<\infty \tag{76}$$

where

$$j(x)=\frac{1}{|x^2-a^2|^{1/2}}\left\{Bx+\frac{sx-a^2}{(s^2-a^2)^{1/2}(x-s)^2}\right\}. \tag{77}$$

Using (22)–(24) for the contributions by the line load and adding those of the couple acting alone, the total stresses in the plane of the anticrack are (terms with $\delta'(x-s)$ in the normal stresses have been omitted)

$$\sigma_{xx}(x,0\pm)=\mp\frac{M(3-\kappa)}{4\pi\kappa}j(x)H(a-|x|),\quad |x|<\infty \tag{78}$$

$$\sigma_{xy}(x,0) = -\frac{M(\kappa-1)}{4\pi\kappa}\left\{\frac{\kappa+1}{\kappa-1}\frac{1}{(x-s)^2} - B + j(x)\,\mathrm{sgn}x H(|x|-a)\right\}, \quad |x|<\infty \tag{79}$$

$$\sigma_{yy}(x,0\pm) = \mp\frac{M(\kappa+1)}{4\pi\kappa} j(x) H(a-|x|), \quad |x|<\infty. \tag{80}$$

It was recently noted by the present authors (Dundurs and Markenscoff, 1989) that the interaction of a concentrated couple with a crack involves some unexpected features. Equations (74) and (78)–(80) show that, except in one respect, the situation with the anticrack is quite similar.

(1) In the limit $s\to a$, as seen from (74) and (78)–(80), the density of the line load and the stress components become unbounded, and one does not recover (A9–A12); thus, there is a stark contrast with concentrated forces. Such an outcome is, in a way, more surprising than that for the crack: Whereas a crack can be viewed as a weakness in the material, an anticrack is a reinforcement, and yet it makes the elastic fields explode when the couple approaches the tip of the anticrack.

(2) To achieve a finite limit as $s\to a$, the magnitude of the couple must be artificially diminished so that $M(s-a)^{1/2} = Q = \text{const}$. In such a case, an $r^{-3/2}$ type singularity results at the tip of the anticrack. However, this result is not empty of meaning, as the $r^{-3/2}$ singularities are those of a weight function for the anticrack.

(3) In the vicinity of $x=+\infty$, (79) gives to the first order

$$\sigma_{xy}(x,0) \sim -\frac{M}{2\pi}\frac{1}{x^2}. \tag{81}$$

It is seen from (81) that the far-field stress of the concentrated couple is not distorted by the anticrack. This outcome is different from that for a crack involving a "magnification factor" $(s-a)^{-1/2}$ that can be arbitrarily large (Dundurs and Markenscoff, 1989). It should also be noted in this connection that the limits involved in (79) are discontinuous. Thus, the limit sequence $x\to\infty$, $s\to a$ gives a different result than the limit sequence $s\to a$, $x\to\infty$.

References

Atkinson, C., 1973, "Some Ribbon-Like Inclusion Problems," *International Journal of Engineering Science*, Vol. 11, pp. 243–266.

Barnett, D. M., and Lothe, J., 1974, "An Image Force Theorem for Dislocations in Anisotropic Bicrystals," *Journal of Physics F: Metal Physics*, Vol. 4, pp. 1618–1635.

Brussat, T. R., and Westmann, R. A., 1975, "A Westergaard-Type Stress Function for Line Inclusion Problems," *International Journal of Solids and Structures*, Vol. 11, pp. 665–667.

Bullough, R., 1964, "The Cracked Dislocation Under Tension," *Philosophical Magazine*, Vol. 9, pp. 917–925.

Comninou, M., 1977, "A Property of Interface Dislocations," *Philosophical Magazine*, Vol. 36, pp. 1281–1283.

Dundurs, J., 1968, "Analogy Between Concentrated Forces and Edge Dislocations," *Journal of Applied Physics*, Vol. 39, pp. 4152–4156.

Dundurs, J., 1970, "Some Properties of Elastic Stresses in a Composite," *Recent Advances in Engineering Science*, Vol. 5, A. C. Eringen, ed., Gordon and Breach, pp. 203–216.

Dundurs, J., 1989, "Cavities Vis-à-Vis Rigid Inclusions and Some Related General Results in Plane Elasticity," ASME JOURNAL OF APPLIED MECHANICS, in press.

Dundurs, J., and Markenscoff, X., 1989, "The Sternberg-Koiter Conclusion and Other Anomalies of the Concentrated Couple," ASME JOURNAL OF APPLIED MECHANICS, in press.

Dundurs, J., and Sendeckyj, G. P., 1965, "Behavior of an Edge Dislocation near a Bimetallic Interface," *Journal of Applied Physics*, Vol. 36, pp. 3353–3354.

Hasebe, N., Keer, L. M., and Nemat-Nasser, S., 1984, "Stress Analysis of a Kinked Crack Initiating From a Rigid Line Inclusion. Part I: Formulation," *Mechanics of Materials*, Vol. 3, pp. 131–145.

Hasebe, N., Nemat-Nasser, S., and Keer, L. M., 1984, "Stress Analysis of a Kinked Crack Initiating From a Rigid Line Inclusion. Part II: Direction of Propagation," *Mechanics of Materials*, Vol. 3, pp. 147–156.

Mura, T., 1988, "Inclusion Problems," *Applied Mechanics Reviews*, Vol. 41, ASME, pp. 15–20.

Nakahara, S., and Willis, J. R., 1973, "Some Remarks on Interfacial Dislocations," *Journal of Physics F: Metal Physics*, Vol. 3, pp. L249–L254.

Stakgold, I., 1968, *Boundary Value Problems of Mathematical Physics*, Vol. II, Macmillan, New York.

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

Tucker, M. O., 1973, "Cracked Dislocations at Boundaries in Two-Material Solids," *Philosophical Magazine*, Vol. 28, pp. 343–362.

Wang, Z. Y., Zhang, H. T., and Chou, Y. T., 1985, "Characteristics of the Elastic Field of a Rigid Line Inhomogeneity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 818–822.

Weertman, J., 1986, "Zener-Stroh Crack, Zener-Holloman Parameter, and Other Topics," *Journal of Applied Physics*, Vol. 60, pp. 1877–1887.

Weertman, J., and Weertman, J. R., 1964, *Elementary Dislocation Theory*, Macmillan, New York.

APPENDIX A

Listed are the basic solutions for the anticrack in the notation used in this article. In all cases the material is of infinite extent, and the anticrack occupies the interval $|x|<a$, $y=0$.

1 Force F_x in the x-Direction Applied Directly to the Rigid Lamella

$$p_x(x) = \frac{F_x}{\pi}\frac{1}{(a^2-x^2)^{1/2}}, \quad |x|<a \tag{A1}$$

$$\sigma_{xx}(x,0) = -\frac{F_x(3+\kappa)}{2\pi(\kappa+1)}\frac{\mathrm{sgn}x H(|x|-a)}{(x^2-a^2)^{1/2}}, \quad |x|<\infty \tag{A2}$$

$$\sigma_{xy}(x,0\pm) = \mp\frac{F_x}{2\pi}\frac{H(a-|x|)}{(a^2-x^2)^{1/2}}, \quad |x|<\infty \tag{A3}$$

$$\sigma_{yy}(x,0) = \frac{F_x(\kappa-1)}{2\pi(\kappa+1)}\frac{\mathrm{sgn}x H(|x|-a)}{(x^2-a^2)^{1/2}}, \quad |x|<\infty \tag{A4}$$

2 Force F_y in the y-Direction Applied Directly at the Center of the Rigid Lamella

$$p_y(x) = \frac{F_y}{\pi}\frac{1}{(a^2-x^2)^{1/2}}, \quad |x|<a \tag{A5}$$

$$\sigma_{xx}(x,0\pm) = \mp\frac{F_y(3-\kappa)}{2\pi(\kappa+1)}\frac{H(a-|x|)}{(a^2-x^2)^{1/2}}, \quad |x|<\infty \tag{A6}$$

$$\sigma_{xy}(x,0) = -\frac{F_y(\kappa-1)}{2\pi(\kappa+1)}\frac{\mathrm{sgn}x H(|x|-a)}{(x^2-a^2)^{1/2}}, \quad |x|<\infty \tag{A7}$$

$$\sigma_{yy}(x,0\pm) = \mp\frac{F_y}{2\pi}\frac{H(a-|x|)}{(a^2-x^2)^{1/2}}, \quad |x|<\infty \tag{A8}$$

3 Couple With Moment M Applied Directly to the Rigid Lamella

$$p_y(x) = \frac{2M}{\pi a^2}\frac{x}{(a^2-x^2)^{1/2}}, \quad |x|<a \tag{A9}$$

$$\sigma_{xx}(x,0\pm) = \mp\frac{M(3-\kappa)}{\pi(\kappa+1)a^2}\frac{xH(a-|x|)}{(a^2-x^2)^{1/2}}, \quad |x|<\infty \tag{A10}$$

$$\sigma_{xy}(x,0) = \frac{M(\kappa-1)}{\pi(\kappa+1)a^2}\left\{1-\frac{|x|H(|x|-a)}{(x^2-a^2)^{1/2}}\right\}, \quad |x|<\infty \tag{A11}$$

$$\sigma_{yy}(x,0\pm) = \mp\frac{M}{\pi a^2}\frac{xH(a-|x|)}{(a^2-x^2)^{1/2}}, \quad |x|<\infty \tag{A12}$$

4 Tension Field T_x Parallel to the Anticrack

$$p_x(x) = -\frac{T_x(\kappa+1)^2}{4\kappa}\frac{x}{(a^2-x^2)^{1/2}}, \quad |x|<a \tag{A13}$$

$$\sigma_{xx}(x,0) = T_x - \frac{T_x(\kappa+1)(3+\kappa)}{8\kappa}\left\{1-\frac{|x|H(|x|-a)}{(x^2-a^2)^{1/2}}\right\}, \quad |x|<\infty \tag{A14}$$

$$\sigma_{xy}(x,0\pm) = \mp\frac{T_x(\kappa+1)^2}{8\kappa}\left\{1-\frac{xH(a-|x|)}{(a^2-x^2)^{1/2}}\right\}, \quad |x|<\infty \tag{A15}$$

$$\sigma_{yy}(x,0) = \frac{T_x(\kappa^2-1)}{8\kappa}\left\{1-\frac{|x|H(|x|-a)}{(x^2-a^2)^{1/2}}\right\}, \quad |x|<\infty \tag{A16}$$

5 Tension Field T_y Perpendicular to the Anticrack

$$p_x(x) = \frac{T_y(\kappa+1)(3-\kappa)}{4\kappa}\frac{x}{(a^2-x^2)^{1/2}}, \quad |x|<a \tag{A17}$$

$$\sigma_{xx}(x,0) = \frac{T_x(9-\kappa^2)}{8\kappa}\left\{1-\frac{|x|H(|x|-a)}{(x^2-a^2)^{1/2}}\right\}, \quad |x|<\infty \tag{A18}$$

$$\sigma_{xy}(x,0\pm) = \mp\frac{T_y(\kappa+1)(3-\kappa)}{8\kappa}\frac{xH(a-|x|)}{(a^2-x^2)^{1/2}}, \quad |x|<\infty \tag{A19}$$

$$\sigma_{yy}(x,0) = \frac{T_y(\kappa+1)(3+\kappa)}{8\kappa} + \frac{T_y(\kappa-1)(3-\kappa)}{8\kappa}\frac{|x|H(|x|-a)}{(x^2-a^2)^{1/2}}, \quad |x|<\infty \tag{A20}$$

APPENDIX B

When a dislocation is situated in a homogeneous material, the resultant of tractions on any straight line vanishes. However, this is not so in a bi-material with an infinitely extended interface. In such a case, the image terms give a net force that is transmitted by the interface. This unusual effect was discovered by Dundurs and Sendeckyj (1965).

For a dislocation with the Burgers vector b_y in the y-direction (see Fig. 2), the distribution of normal tractions $\sigma_{xx}(0, y)$ at the interface is bell-shaped, which becomes more peaked as the dislocation approaches the interface (the total force remains constant). The following distribution of stresses at the interface can be extracted from the results given by Dundurs and Sendeckyj (1965) in the limit as $h\to 0$:

$$\sigma_{xx}(0,y) = -C\beta b_y\delta(y) \tag{B1}$$

$$\sigma_{xy}(0,y) = -\frac{C}{\pi}b_y\frac{1}{y} \tag{B2}$$

$$\sigma_{yy}(0\pm,y) = -C(\mp 2+\beta)b_y\delta(y) \tag{B3}$$

where

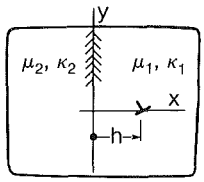


Fig. 2 Edge dislocation near an interface between two materials

$$C = \frac{2\mu_1(1+\alpha)}{(\kappa_1+1)(1-\beta^2)} = \frac{2\mu_2(1-\alpha)}{(\kappa_2+1)(1-\beta^2)} \tag{B4}$$

$$\alpha = \frac{\mu_2(\kappa_1+1)-\mu_1(\kappa_2+1)}{\mu_2(\kappa_1+1)+\mu_1(\kappa_2+1)}, \quad \beta = \frac{\mu_2(\kappa_1-1)-\mu_1(\kappa_2-1)}{\mu_2(\kappa_1+1)+\mu_1(\kappa_2+1)} \tag{B5}$$

(see also the papers by Nakahara and Willis (1973), Barnett and Lothe (1974), and Comninou (1977)). It is seen now from (*B*1) that a pair of concentrated forces

$$F_x = -C\beta b_y \tag{B6}$$

act on the solids. For positive β b_y the solids are pried apart, for negative β b_y they are pulled together. If the second phase is rigid ($\mu_2\to\infty$),

$$F_x = -\frac{b_y\mu_1(\kappa_1-1)}{\kappa_1} \tag{B7}$$

which is the same as (56). Suppose now that a dislocation with a positive b_y is placed at the right pointed end of a rigid elliptical inclusion. It is clear that the dislocation then exerts the force given by (*B*7) on the inclusion in the negative x-direction as long as the curvature at this point is finite. This force is, of course, balanced by the tractions acting on the inclusion elsewhere. The unexpected result obtained in Section 5 simply means, therefore, that nothing changes as the elliptical inclusion degenerates into an anticrack.

The counterparts of (*B*1)–(*B*3) for a dislocation with the Burgers vector b_x in the x-direction are

$$\sigma_{xx}(0,y) = -\frac{C}{\pi}b_x\frac{1}{y} \tag{B8}$$

$$\sigma_{xy}(0,y) = C\beta b_x\delta(y) \tag{B9}$$

$$\sigma_{yy}(0\pm,y) = -\frac{C}{\pi}(1\mp 2\beta)b_x\frac{1}{y}. \tag{B10}$$

The pair of concentrated forces with the magnitude

$$F_y = C\beta b_x \tag{B11}$$

in this case tend to make the interface slip. For a rigid second phase

$$F_y = \frac{b_x\mu_1(\kappa_1-1)}{\kappa_1}, \tag{B12}$$

which is the same as the force exerted by a dislocation at the tip of the anticrack.

Line Inclusions in Anisotropic Elastic Solids

Qianqian Li

T. C. T. Ting
Fellow ASME

Department of Civil Engineering, Mechanics, and Metallurgy, University of Illinois at Chicago, Chicago, Ill. 60680

A line inclusion located at $x_2 = 0$, $|x_1| < 1$ in the anisotropic elastic medium of infinite extent under uniform loading at infinity is considered. Stroh's formalism is used to find the displacement and stress fields. The inclusion can be rigid or elastic. Conditions on the loading under which the line inclusion does not disturb the homogeneous field are derived. For the rigid inclusion, a real form solution is obtained for the stress and displacement along $x_2 = 0$. When the inclusion is elastic (and anisotropic), a pair of singular Fredholm integral equations of the second kind is derived for the difference in the stress on both surfaces of the inclusion. The pair can be decoupled and asymptotic solutions of the integral equation are obtained when λ, which represents the relative rigidity of the matrix to the inclusion, is small. For the general cases, the integral equation is solved by a numerical discretization. Excellent agreements between the asymptotic and numerical solutions are observed for small λ.

1 Introduction

It is known that most materials contain some defects in the form of cracks, voids, or inclusions which can affect the load-carrying capacity of engineering structures. It is therefore important to know how the defects disturb the stress field and how the stress concentration arises due to the existence of the defects. Crack problems have received much attention and have been widely studied. Inclusions or inhomogeneity problems (sometimes called hard crack or inverse crack problems) have also aroused much interest in recent years. An extensive review on the subject has been given by Mura (1987, 1988).

Using the methods of Eshelby (1957, 1959), Muskhelishvili (1953), Chou and Wang (1983), and Wang et al. (1985, 1986) considered a rigid line inclusion in an isotropic plane elastic body. Analytical expressions of the stress fields due to uniform remote loading are derived. The same problem has been considered by Ballarini (1987) using the method of integral transform. Their results showed that the stresses near the tips of the rigid line inclusion have square root singularities. The problem of an elastic line inclusion was investigated by Sendeckyj (1970) and Selvadurai (1980). Using Muskhelishvili's method, Atkinson (1973) also considered the problem of elastic inclusions in isotropic solids. Under the assumption that the inclusion is much "harder" than the matrix, he obtained an asymptotic solution for the stress fields. Erdogan and Gupta (1972) studied the more general problem of bonded materials containing a flat inclusion which may be rigid or elastic with negligible bending rigidity. They formulated the problem into a system of singular integral equations which was then solved by expanding the solutions in a Chebyshev polynomial.

Although a great deal of work has been done for inclusions in isotropic matrices, the problem of an inclusion in anisotropic media, which is becoming more and more important with the replacement of conventional materials by varieties of composite materials, seems to have not received much attention. There are several approaches in solving anisotropic elasticity problems, but in this paper we will employ the Stroh formalism (Stroh, 1958, 1962). We consider the problem of a very thin, flat inclusion in an infinite, generally anisotropic, elastic body which is subjected to a uniform loading at infinity. The inclusion is assumed to be located at $x_2 = 0$, $|x_1| < 1$, $-\infty < x_3 < \infty$. The deformation is two-dimensional in the sense that the displacements depend on x_1 and x_2 only. In Section 2, the fundamental equations of anisotropic elasticity and a brief account of the Stroh formalism are given. The homogeneous solution, which is the solution of the infinite medium when the inclusion is absent, is given in Section 3. Conditions under which the homogeneous solution is not disturbed by the presence of the inclusion are derived here. The rigid inclusion problem is considered in Section 4 where a real form solution is obtained for the stress and displacement at $x_2 = 0$. Section 5 is devoted to the case when the inclusion is elastic and anisotropic. A pair of Fredholm integral equations is derived for the difference in the stresses on the surfaces of the inclusion. The pair can be decoupled and each of the uncoupled equations has the same Fredholm integral equation form involving the parameter λ which represents the relative rigidity of the matrix material to the inclusion. The asymptotic and numerical solutions of the in-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10-15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, September 15, 1988; final revision, January 16, 1989.

Paper No. 89-WA/APM-17.

tegral equations are presented in Section 6 and results are plotted for various values of λ. Comparisons of the numerical solutions with the asymptotic solutions show that the asymptotic solutions provide good approximations for λ up to 0.1.

2 Basic Equations

We present in this section a brief derivation of Stroh's formalism (Stroh, 1958, 1962; Barnett and Lothe, 1973; Chadwick and Smith, 1977; Ting, 1986). In a fixed rectangular coordinate system x_1, x_2, x_3, let u_i, σ_{ij}, ϵ_{ij} be, respectively, the displacement, stress, and strain of the material. The equations of equilibrium and the stress-strain laws for the material can be written as

$$\sigma_{ij,j} = 0, \tag{1}$$

$$\sigma_{ij} = C_{ijks}\epsilon_{ks} = C_{ijks}u_{k,s}, \tag{2}$$

in which a comma stands for differentiation, repeated indices imply summation, and C_{ijks} are the elastic stiffnesses which satisfy the symmetry conditions

$$C_{ijks} = C_{jiks} = C_{ksij}.$$

For two-dimensional deformations in which u_i, $i = 1, 2, 3$, depend on x_1 and x_2 only; the general solution can be written as

$$u_k = a_k f(z), \tag{3}$$

$$z = x_1 + px_2, \tag{4}$$

where p and a_k are constants and f is an arbitrary function of z. In matrix notation, p and a_k are determined by

$$\{\mathbf{Q} + p(\mathbf{R} + \mathbf{R}^T) + p^2\mathbf{T}\}\mathbf{a} = \mathbf{0}, \tag{5}$$

in which the superscript T stands for the transpose and $Q_{ij} = C_{i1k1}$, $R_{ik} = C_{i1k2}$, $T_{ik} = C_{i2k2}$. We note that $\mathbf{Q}$ and $\mathbf{T}$ are symmetric and, subject to positiveness of strain energy, positive definite. Equation (5) is obtained by substituting (3) into (2) and (1).

Introducing the new vector

$$\mathbf{b} = (\mathbf{R}^T + p\mathbf{T})\mathbf{a} = -p^{-1}(\mathbf{Q} + p\mathbf{R})\mathbf{a}, \tag{6}$$

in which the second equality comes from (5), the stress obtained by substituting (3) into (2) can be written as

$$\sigma_{i1} = -\Phi_{i,2}, \quad \sigma_{i2} = \Phi_{i,1}, \tag{7}$$

$$\mathbf{\Phi} = \mathbf{b}f(z). \tag{8}$$

Thus, $\mathbf{\Phi}$ is the stress function.

Equation (5) provides six eigenvalues, p_α, and six eigenvectors, $\mathbf{a}_\alpha$, $\alpha = 1, 2, \ldots, 6$. From (6) we obtain six $\mathbf{b}_\alpha$. Since p_α cannot be real if the strain energy is positive (Eshelby, 1953), p_α come in three pairs of complex conjugates and so do $\mathbf{a}_\alpha$ and $\mathbf{b}_\alpha$. Without loss of generality we let

$$\mathrm{Im}\{p_\alpha\} > 0, \quad \bar{p}_{\alpha+3} = p_\alpha, \tag{9}$$

$$\mathbf{a}_{\alpha+3} = \bar{a}_\alpha, \quad \mathbf{b}_{\alpha+3} = \bar{b}_\alpha, \quad \alpha = 1, 2, 3,$$

where Im stands for the imaginary part and the overbar denotes the complex conjugate. It follows from (3) and (8) that the general solution for $\mathbf{u}$ and $\mathbf{\Phi}$ can be written as

$$\mathbf{u} = \sum_{\alpha=1}^{3} \{\mathbf{a}_\alpha f_\alpha(z_\alpha) + \bar{\mathbf{a}}_\alpha f_{\alpha+3}(\bar{z}_\alpha)\}, \tag{10a}$$

$$\mathbf{\Phi} = \sum_{\alpha=1}^{3} \{\mathbf{b}_\alpha f_\alpha(z_\alpha) + \bar{\mathbf{b}}_\alpha f_{\alpha+3}(\bar{z}_\alpha)\}, \tag{10b}$$

in which $f_1, f_2, \ldots, f_6$ are arbitrary functions of their arguments. In writing (10), we have assumed that the eigenvalues p_α are all distinct or, if there is a repeated p_α, the associated eigenvectors $\mathbf{a}_\alpha$ are independent of each other. A modified expression can be found in (Ting, 1982) if this is not the case.

If we define the 3×3 matrices $\mathbf{A}$ and $\mathbf{B}$ by

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3),$$

$$\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3),$$

it can be shown that the matrices $\mathbf{H}$, $\mathbf{L}$, $\mathbf{S}$ given by

$$\mathbf{H} = 2i\mathbf{A}\mathbf{A}^T, \quad \mathbf{L} = -2i\mathbf{B}\mathbf{B}^T, \quad \mathbf{S} = i(2\mathbf{A}\mathbf{B}^T - \mathbf{I}), \tag{11}$$

where $i = \sqrt{-1}$ and $\mathbf{I}$ is the unit matrix, are real. Moreover, the matrices $\mathbf{H}$ and $\mathbf{L}$ are symmetric and positive definite, $\mathbf{H}^{-1}\mathbf{S}$ is antisymmetric, and

$$\mathbf{B}\mathbf{A}^{-1} = (\mathbf{S}^T + i\mathbf{I})\mathbf{H}^{-1} = -\mathbf{H}^{-1}(\mathbf{S} - i\mathbf{I}). \tag{12}$$

Equations (11) are valid provided the eigenvalues p_α are distinct or, if there is a repeated eigenvalue, the associated eigenvectors $\mathbf{a}_\alpha$ are independent. A modified expression in place of (11) when this is not the case can be found in (Ting and Hwu, 1988). An alternate approach using an integral formalism without determining the eigenvalues and eigenvectors was proposed by Barnett and Lothe (1973). For isotropic materials, $\mathbf{H}$ and $\mathbf{L}$ are diagonal matrices given by

$$\mathbf{H} = \mu^{-1}\mathrm{diag}\{(1+s)/2, (1+s)/2, 1\}, \tag{13}$$

$$\mathbf{L} = \mu\mathrm{diag}\{2(1-s), 2(1-s), 1\}, \tag{14}$$

$$s = (1-2\nu)/2/1-\nu),$$

where μ and ν are, respectively, shear modulus and Poisson's ratio, while $\mathbf{S}$ has only two nonzero elements

$$\mathbf{S} = \begin{bmatrix} 0 & -s & 0 \\ s & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{15}$$

Before we close this section, we list next the alternate contracted notation for the stress-strain law given by (2). By letting

$$\sigma_1 = \sigma_{11}, \sigma_2 = \sigma_{22}, \sigma_3 = \sigma_{33}, \sigma_4 = \sigma_{23}, \sigma_5 = \sigma_{13}, \sigma_6 = \sigma_{12},$$

$$\epsilon_1 = \epsilon_{11}, \epsilon_2 = \epsilon_{22}, \epsilon_3 = \epsilon_{33}, \epsilon_4 = 2\epsilon_{23}, \epsilon_5 = 2\epsilon_{13}, \epsilon_6 = 2\epsilon_{12},$$

$(2)_1$ can be written as

$$\sigma_i = C_{ij}\epsilon_j, \tag{16}$$

and the inverse of the stress-strain laws as

$$\epsilon_i = W_{ij}\sigma_j, \tag{17}$$

where W_{ij} are the elastic compliances. For the two-dimensional deformation considered here

$$\epsilon_3 = W_{3j}\sigma_j = 0.$$

Solving for σ_3 and substituting it into (17) leads to

$$\epsilon_i = \hat{W}_{ij}\sigma_j, \tag{18}$$

where

$$\hat{W}_{ij} = W_{ij} - W_{i3}W_{3j}/W_{33}.$$

Since $\hat{W}_{3j} = \hat{W}_{i3} = 0$, the elements of the third column and the third row of the 6×6 matrix $\hat{\mathbf{W}}$ vanish. Deleting the third column and the third row, it is shown in (Ting, 1988) that the reduced 5×5 matrix is positive definite.

3 Homogeneous Solution

For a line inclusion located at $x_2 = 0$, $-1 < x_1 < 1$, subject to a uniform stress σ_{ij}^∞ at infinity, the solution consists of two parts. The first part is the homogeneous solution without the presence of the inclusion. The second part is the disturbed solution due to the presence of the inclusion. In this section we consider the homogeneous solution which will be denoted by the superscript (o).

For a uniform stress σ_{ij}^∞ at infinite, the solution for displacements can be written as

$$u_i^{(o)} = (\epsilon_{ij}^{\infty} + \omega_{ij}^{(o)}) x_j, \tag{19}$$

where ϵ_{ij}^{∞} and $\omega_{ij}^{(o)}$ are constants. ϵ_{ij}^{∞} is, in contracted notations of (17),

$$\epsilon_i^{\infty} = W_{ij}\sigma_j^{\infty}. \tag{20}$$

The stresses σ_j^{∞} are prescribed in such way that $\epsilon_3^{\infty} = 0$. Alternately, one could prescribe ϵ_j^{∞} with $\epsilon_3^{\infty} = 0$. σ_j^{∞} are then determined from (16). Since displacements are assumed to depend on x_1 and x_2 only, we may choose $\omega_{ij}^{(o)}$ such that the coefficients of x_3 vanish. Also, if we let $u_2^{(o)} = 0$ along $x_2 = 0$, (19) reduces to

$$\mathbf{u}^{(o)} = \begin{bmatrix} \epsilon_1^{\infty} & \epsilon_6^{\infty} & 0 \\ 0 & \epsilon_2^{\infty} & 0 \\ \epsilon_5^{\infty} & \epsilon_4^{\infty} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

The displacement at $x_2 = 0$ where the inclusion is located is then given by

$$\mathbf{u}^{(o)}(x_1, 0) = \begin{bmatrix} \epsilon_1^{\infty} \\ 0 \\ \epsilon_5^{\infty} \end{bmatrix} x_1. \tag{21}$$

If the inclusion is rigid, and if $\epsilon_1^{\infty} = \epsilon_5^{\infty} = 0$, or

$$\begin{aligned} \hat{W}_{1j}\sigma_j^{\infty} &= 0, \\ \hat{W}_{5j}\sigma_j^{\infty} &= 0, \end{aligned} \tag{22}$$

then $\mathbf{u}^{(o)}(x_1, 0) = 0$ and the existence of a rigid inclusion at $x_2 = 0$, $-1 < x_1 < 1$ does not disturb the homogeneous solution. For given material constants, W_{ij}, (22) provide conditions on σ_j^{∞} so that the homogeneous solution is the solution for the rigid inclusion.

As an example, consider the case in which the body is subjected to a uniform tension of magnitude σ_o in the direction

$$(\cos\beta\cos\alpha, \cos\beta\sin\alpha, \sin\beta),$$

which makes an angle β with the (x_1, x_2) plane and its projection on the (x_1, x_2) plane makes an angle α with the x_1-axis, Fig. 1. To produce this stress state with $\epsilon_3^{\infty} = 0$, it can be shown that σ_i^{∞} must be given by

$$\begin{aligned} \sigma_1^{\infty} &= \sigma_o \cos^2\beta\cos^2\alpha, \\ \sigma_2^{\infty} &= \sigma_o \cos^2\beta\sin^2\alpha, \\ \sigma_4^{\infty} &= \sigma_o \sin\beta\cos\beta\sin\alpha, \\ \sigma_5^{\infty} &= \sigma_o \sin\beta\cos\beta\cos\alpha, \\ \sigma_6^{\infty} &= \sigma_o \cos^2\beta\sin\alpha\cos\alpha, \\ \sigma_3^{\infty} &= -\sum_{j\neq 3} W_{3j}\sigma_j^{\infty}/W_{33}. \end{aligned} \tag{23}$$

Equations (22) now reduce to

$$\begin{aligned} -\tan\beta &= \frac{\hat{W}_{51}\cos^2\alpha + \hat{W}_{52}\sin^2\alpha + \hat{W}_{56}\sin\alpha\cos\alpha}{\hat{W}_{54}\sin\alpha + \hat{W}_{55}\cos\alpha} \\ &= \frac{\hat{W}_{11}\cos^2\alpha + \hat{W}_{12}\sin^2\alpha + \hat{W}_{16}\sin\alpha\cos\alpha}{\hat{W}_{14}\sin\alpha + \hat{W}_{15}\cos\alpha} \end{aligned} \tag{24}$$

from which the second equality yields

$$\cot^3\alpha + a\cot^2\alpha + b\cot\alpha + c = 0, \tag{25}$$

where

$$\begin{aligned} a &= (\hat{W}_{11}\hat{W}_{54} + \hat{W}_{16}\hat{W}_{55} - \hat{W}_{14}\hat{W}_{15} - \hat{W}_{15}\hat{W}_{56})/\Delta, \\ b &= (\hat{W}_{12}\hat{W}_{55} + \hat{W}_{16}\hat{W}_{54} - \hat{W}_{14}\hat{W}_{56} - \hat{W}_{15}\hat{W}_{25})/\Delta, \\ c &= (\hat{W}_{12}\hat{W}_{54} - \hat{W}_{14}\hat{W}_{52})/\Delta, \\ \Delta &= \hat{W}_{11}\hat{W}_{55} - \hat{W}_{15}^2 > 0. \end{aligned}$$

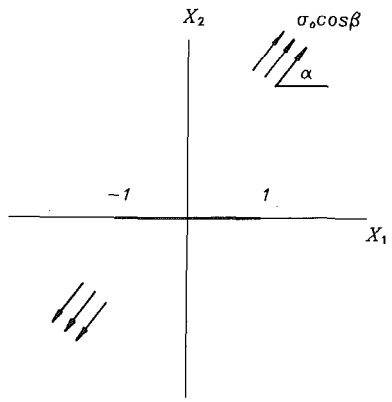


Fig. 1 Inclusion configuration

Equation (25) provides at least one real root for α. Equation (24) then furnishes β.

For isotropic materials the only nonzero $\hat{W}_{ij}$ appearing in (24) are $\hat{W}_{11}$, $\hat{W}_{12}$, $\hat{W}_{55}$ which are given by

$$\hat{W}_{11} = (1-\nu)/2\mu, \quad \hat{W}_{12} = -\nu/2\mu, \quad \hat{W}_{55} = 1/\mu.$$

Equation (24) and (25) yield $\beta = 0$ and

$$\tan^2\alpha = -\hat{W}_{11}/\hat{W}_{12}, \tag{26}$$

or

$$\alpha = \pm\tan^{-1}\sqrt{(1-\nu)/\nu}, \tag{27}$$

which agrees with the result obtained by Wang et al. (1985).

If the line inclusion is elastic with elastic compliances $W_{ij}^{(i)}$ and if

$$\begin{aligned} \epsilon_1^{\infty} &= \hat{W}_{1j}\sigma_j^{\infty} = \hat{W}_{1j}^{(i)}\sigma_j^{\infty}, \\ \epsilon_5^{\infty} &= \hat{W}_{5j}\sigma_j^{\infty} = \hat{W}_{5j}^{(i)}\sigma_j^{\infty}, \end{aligned} \tag{28}$$

assuming that σ_j $(j \neq 3)$ in the inclusion and in the matrix is identical, then the displacements at $x_2 = 0$ produced by the homogeneous solution is compatible with the deformation of the line inclusion. This means that the elastic line inclusion does not disturb the homogeneous solution and the homogeneous solution is the solution for the inclusion problem. If we define

$$[\hat{W}_{ij}] = \hat{W}_{ij} - \hat{W}_{ij}^{(i)}, \tag{29}$$

(28) can be written as

$$\begin{aligned} [\hat{W}_{1j}]\sigma_j^{\infty} &= 0, \\ [\hat{W}_{5j}]\sigma_j^{\infty} &= 0, \end{aligned} \tag{30}$$

which provide conditions on σ_j^{∞} for the homogeneous solution to be the solution for the elastic inclusion. Equations (30) are identical to (22) if we replace $\hat{W}_{ij}$ by $[\hat{W}_{ij}]$ for $i = 1, 5$. Therefore, for the loading given by (23), equations (24) through (26) remain valid for the elastic inclusion if we replace $\hat{W}_{ij}$ by $[\hat{W}_{ij}]$. If the inclusion is isotropic with shear modulus $\mu^{(i)}$ and Poisson's ratio $\nu^{(i)}$, (27) is replaced by

$$\alpha = \pm\tan^{-1}\left\{\frac{(1-\nu)\mu^{(i)} - (1-\nu^{(i)})\mu}{\nu\mu^{(i)} - \nu^{(i)}\mu}\right\}^{1/2}.$$

We see that if $\nu = \nu^{(i)}$, α is independent of μ and $\mu^{(i)}$ and is identical to (27). On the other hand, if $\mu = \mu^{(i)}$, $\alpha = \pm\pi/4$ regardless of the value of ν and $\nu^{(i)}$.

4 A Rigid Line Inclusion

In this section we assume that the inclusion, which is at $x_2 = 0$, $-1 < x_1 < 1$ is rigid, and study the disturbed solution due to the presence of the inclusion. The disturbed solution will remove the displacements at $x_2 = 0$, $-1 < x_1 < 1$, given

in (21) where is generated by the homogenous solution. The boundary conditions for the disturbed solution can be written as

$$\mathbf{u}(x_1, 0) = -\mathbf{q}x_1, \quad \text{for } |x_1| < 1, \tag{31}$$

$$\int_{-1}^{1} [\sigma_{22}^{+}(x_1, 0^+) - \sigma_{22}^{-}(x_1, 0^-)]x_1 dx_1 = 0, \tag{32}$$

$$\sigma_{ij}(x_1, x_2) = 0, \quad \text{as } |\mathbf{x}| \to \infty, \tag{33}$$

where

$$\mathbf{q} = \begin{bmatrix} \epsilon_1^{\infty} \\ -\omega \\ \epsilon_5^{\infty} \end{bmatrix}. \tag{34}$$

In (34), ω is an unknown constant which represents the rotation of the rigid inclusion. It will be determined by the condition (32) that the total moment about $x_1 = 0$ due to the surface traction on the rigid inclusion vanishes. To satisfy the boundary conditions, we employ the general solution (10) and choose the function f_α such that

$$\mathbf{u} = \mathrm{Re} \sum_{\alpha} \mathbf{a}_\alpha (\mathbf{A}^{-1})_{\alpha j} \{ (z_\alpha^2 - 1)^{1/2} - z_\alpha \} q_j, \tag{35}$$

$$\boldsymbol{\Phi} = \mathrm{Re} \sum_{\alpha} \mathbf{b}_\alpha (\mathbf{A}^{-1})_{\alpha j} \{ (z_\alpha^2 - 1)^{1/2} - z_\alpha \} q_j, \tag{36}$$

where Re stands for the real part and q_j is given by (34). At $x_2 = 0^{\pm}$, we have

$$(z_\alpha^2 - 1)^{1/2} = \begin{cases} \pm i(1 - x_1^2)^{1/2}, & |x_1| < 1, \\ +(x_1^2 - 1)^{1/2}, & x_1 > 1, \\ -(x_1^2 - 1)^{1/2}, & x_1 < -1. \end{cases} \tag{37}$$

The displacement at $x_2 = 0^{\pm}$ is therefore

$$\mathbf{u}(x_1, 0) = \begin{cases} -\mathbf{q}x_1, & \text{for } |x_1| < 1, \\ \{ \pm (x_1^2 - 1)^{1/2} - x_1 \}\mathbf{q}, & \text{for } x_1 > 1 \text{ or } x_1 < -1. \end{cases} \tag{38}$$

Equations $(38)_1$ satisfies the boundary condition (31). It is interesting to see from $(38)_2$ that $\mathbf{u}(x_1, 0)$ for $|x_1| > 1$ is independent of material property.

The stress is obtained from (36) and (7) as

$$\sigma_{i1} = -\mathrm{Re} \sum_{\alpha} \left\{ B_{i\alpha} (\mathbf{A}^{-1})_{\alpha j} p_\alpha \left[\frac{z_\alpha}{(z_\alpha^2 - 1)^{1/2}} - 1 \right] q_j \right\},$$

$$\sigma_{i2} = \mathrm{Re} \sum_{\alpha} \left\{ B_{i\alpha} (\mathbf{A}^{-1})_{\alpha j} \left[\frac{z_\alpha}{(z_\alpha^2 - 1)^{1/2}} - 1 \right] q_j \right\}.$$

Since $(z_\alpha^2 - 1)^{1/2} \to z_\alpha$ as $|\mathbf{x}| \to \infty$, the boundary condition (33) is satisfied. Using (37) the stress at $x_2 = 0^{\pm}$ has the expression

$$\boldsymbol{\sigma}_1 = -\mathrm{Re}\{\mathbf{BPA}^{-1}\mathbf{q}(\Psi(x_1) - 1)\},$$
$$\boldsymbol{\sigma}_2 = \mathrm{Re}\{\mathbf{BA}^{-1}\mathbf{q}(\Psi(x_1) - 1)\},$$

in which

$$(\boldsymbol{\sigma}_1)_i = \sigma_{i1}, \quad (\boldsymbol{\sigma}_2)_i = \sigma_{i2}, \tag{39}$$

and the diagonal matrix $\mathbf{P}$ and $\Psi(x_1)$ are given by

$$\mathbf{P} = \mathrm{diag}\,(p_1, p_2, p_3),$$

$$\Psi(x_1) = \begin{cases} \dfrac{x_1}{\pm i(1 - x_1^2)^{1/2}}, & \text{for } |x_1| < 1, \ x_2 = 0^{\pm}, \\ \pm \dfrac{x_1}{(x_1^2 - 1)^{1/2}}, & \text{for } x_1 > 1 \text{ or } < -1, \ x_2 = 0. \end{cases} \tag{40}$$

It is shown in (Ting, 1988) that

$$\mathbf{BPA}^{-1} = \mathbf{N}_3 + \mathbf{N}_1^T(\mathbf{S}^T + i\mathbf{I})\mathbf{H}^{-1},$$
$$\mathbf{N}_1 = -\mathbf{T}^{-1}\mathbf{R}^T, \quad \mathbf{N}_3 = \mathbf{R}\mathbf{T}^{-1}\mathbf{R}^T - \mathbf{Q}. \tag{41}$$

Using (12), (40), and (41), the stress at $x_2 = 0$ has the real expressions

$$\boldsymbol{\sigma}_1 = (\mathbf{N}_3\mathbf{H} + \mathbf{N}_1^T\mathbf{S}^T)\mathbf{H}^{-1}\mathbf{q} \mp \frac{x_1}{(1 - x_1^2)^{1/2}} \mathbf{N}_1^T\mathbf{H}^{-1}\mathbf{q},$$

$$\boldsymbol{\sigma}_2 = -\mathbf{S}^T\mathbf{H}^{-1}\mathbf{q} \pm \frac{x_1}{(1 - x_1^2)^{1/2}} \mathbf{H}^{-1}\mathbf{q}, \tag{42}$$

for $|x_1| < 1, x_2 = 0^{\pm}$ and

$$\boldsymbol{\sigma}_1 = -\left(\pm \frac{x_1}{(x_1^2 - 1)^{1/2}} - 1 \right) (\mathbf{N}_3\mathbf{H} + \mathbf{N}_1^T\mathbf{S}^T)\mathbf{H}^{-1}\mathbf{q},$$

$$\boldsymbol{\sigma}_2 = \left(\pm \frac{x_1}{(x_1^2 - 1)^{1/2}} - 1 \right) \mathbf{S}^T\mathbf{H}^{-1}\mathbf{q}, \tag{43}$$

for $x_1 > 1$ or $x_1 < -1, x_2 = 0$. The upper sign is for $x_1 > 1$ and the lower sign for $x_1 < -1$. We see from (42) that to satisfy the boundary condition (32) we must have

$$(\mathbf{H}^{-1}\mathbf{q})_2 = 0.$$

This leads to, letting $\mathbf{h} = \mathbf{H}^{-1}$,

$$\omega = \frac{h_{21}\epsilon_1^{\infty} + h_{23}\epsilon_5^{\infty}}{h_{22}},$$

which provides the angular rotation of the rigid inclusion. The fact that $\mathbf{H}$ is positive definite assures us that $h_{22} > 0$. If $h_{21} = h_{23} = 0$, which is the case for isotropic materials, $\omega = 0$ and there is no rotation of the line inclusion due to the uniform loading at infinity.

Equations (42) and (43) show that there is a square root singularity in stress at $x_1 = \pm 1, x_2 = 0$ unless $\mathbf{H}^{-1}\mathbf{q} = 0$. But $\mathbf{H}^{-1}\mathbf{q} = 0$ means $\mathbf{q} = \mathbf{0}$ because $\mathbf{H}^{-1}$ is positive definite. Therefore, the disturbed solution always generates a stress singularity at the tips of the inclusion. If $r > 0$, we see from (43) that

$$\lim_{r \to o} (2r)^{1/2} \boldsymbol{\sigma}_1(1 + r, 0) = -(\mathbf{N}_3\mathbf{H} + \mathbf{N}_1^T\mathbf{S}^T)\mathbf{H}^{-1}\mathbf{q},$$
$$\lim_{r \to o} (2r)^{1/2} \boldsymbol{\sigma}_2(1 + r, 0) = \mathbf{S}^T\mathbf{H}^{-1}\mathbf{q}.$$

This provides the stress singularity coefficients (Wang et al., 1985) for the rigid line inclusion in anisotropic elastic materials. We also have, on the inclusion

$$\lim_{r \to o} (2r)^{1/2} \boldsymbol{\sigma}_1(1 - r, 0^+) = -\mathbf{N}_1^T\mathbf{H}^{-1}\mathbf{q},$$
$$\lim_{r \to o} (2r)^{1/2} \boldsymbol{\sigma}_2(1 - r, 0^+) = \mathbf{H}^{-1}\mathbf{q}.$$

For isotropic materials

$$\mathbf{N}_1 = -\begin{bmatrix} 0 & 1 & 0 \\ \dfrac{\nu}{1 - \nu} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{N}_3 = -\mu \begin{bmatrix} \dfrac{2}{1 - \nu} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and using (13) and (15), the only nonzero limits are

$$\lim_{r \to o} (2r)^{1/2} \sigma_{22}(1 + r, 0) = -\mu \frac{2(1 - 2\nu)}{3 - 4\nu} \epsilon_1^{\infty},$$

$$\lim_{r \to o} (2r)^{1/2} \sigma_{13}(1 + r, 0) = \mu\epsilon_5^{\infty},$$

$$\lim_{r \to o} (2r)^{1/2} \sigma_{11}(1 + r, 0) = 2\mu \frac{3 - 2\nu}{3 - 4\nu} \epsilon_1^{\infty},$$

$$\lim_{r\to o}\ (2r)^{1/2}\sigma_{12}(1-r,\ 0^+)=\mu\,\frac{4(1-\nu)}{3-4\nu}\,\epsilon_1^\infty,$$

$$\lim_{r\to o}\ (2r)^{1/2}\sigma_{23}(1-r,\ 0^+)=\mu\epsilon_5^\infty.$$

In closing this section, we point out that Yang and Chou (1982) studied the elliptic inclusion in an anisotropic solid of cubic symmetry. They obtained explicit solutions for stresses around the inclusion which include the degenerate case of a rigid line inclusion.

5 An Elastic Line Inclusion

The problem becomes more complicated when the line inclusion is elastic and deformable. In this case the normal stresses on two surfaces of the inclusion are identical if we neglect the bending rigidity, but the shear stresses will have a discontinuity which acts as an external force on the elastic inclusion. We assume that the inclusion is also anisotropic with the elastic compliance $W_{ij}^{(i)}$.

For the state disturbed from the homogeneous deformation, we seek a solution which satisfies zero stress at infinity while on $x_2 = 0$, $|x_1| < 1$,

$$\partial\mathbf{u}(x_1, 0)/\partial x_1 = -\mathbf{q}(x_1), \tag{44}$$

where $\mathbf{q}(x_1)$ is an unknown function of x_1 to be determined. We see that if $\tilde{\epsilon}_k^{(i)}$ is the average strains across the thickness of the inclusion,

$$\begin{aligned} q_1(x_1)&=\epsilon_1^\infty-\tilde{\epsilon}_1^{(i)}(x_1),\\ q_3(x_1)&=\epsilon_5^\infty-\tilde{\epsilon}_5^{(i)}(x_1),\end{aligned} \tag{45}$$

while $-q_2(x_1)$ represents the rotation of the line inclusion. If $q_2(x_1)$ is independent of x_1, the line inclusion remains a straight line after the deformation. Otherwise, it deforms into a curved line. Following Stroh (1958), the solution which satisfies zero stress at infinity and prescribed displacement gradient $-\mathbf{q}(x_1)$ at $x_2 = 0$, $|x_1| < 1$ can be written as (Li, 1988):

$$\begin{aligned} u_k=&\frac{1}{\pi}\int_o^1 d\mu\int_{-1}^1 q_k(\mu\xi)\xi(1-\xi^2)^{-1/2}d\xi\\ &-\frac{1}{\pi}\,\mathrm{Re}\Big\{\sum_\alpha A_{k\alpha}A_{\alpha j}^{-1}\int_o^1 d\mu(z_\alpha^2-\mu^2)^{-1/2}\\ &\int_{-1}^1 q_j(\mu\xi)(\mu+z_\alpha\xi)(1-\xi^2)^{-1/2}d\xi\Big\},\end{aligned} \tag{46}$$

$$\begin{aligned} \Phi_k=&\frac{1}{\pi}\,\mathrm{Re}\Big\{B_{k\alpha}A_{\alpha j}^{-1}\int_o^1 d\mu\int_{-1}^1 q_j(\mu\xi)\xi(1-\xi^2)^{-1/2}d\xi\Big\}\\ &-\frac{1}{\pi}\,\mathrm{Re}\Big\{\sum_\alpha B_{k\alpha}A_{\alpha j}^{-1}\int_o^1 d\mu(z_\alpha^2-\mu^2)^{-1/2}\\ &\int_{-1}^1 q_j(\mu\xi)(\mu+z_\alpha\xi)(1-\xi^2)^{-1/2}d\xi\Big\}.\end{aligned} \tag{47}$$

We will now proceed to determine $\mathbf{q}(x_1)$ so that the displacements at $x_2 = 0$, $|x_1| < 1$ are compatible with the deformation of the line inclusion. We first observe that the first terms on the right of (46) and (47) are constants. We next reduce the double integrals in the second terms to a single integral. This is accomplished by replacing the variable ξ by $\eta = \mu\xi$ and changing the order of integration (Stroh, 1958). Finally, we differentiate and employ (7) to obtain

$$u_{k,1}=\frac{1}{\pi}\,\mathrm{Re}\sum_\alpha\{A_{k\alpha}A_{\alpha j}^{-1}(z_\alpha^2-1)^{-1/2}\gamma_j(z_\alpha)\}, \tag{48}$$

$$\sigma_{k1}=-\frac{1}{\pi}\,\mathrm{Re}\sum_\alpha\{B_{k\alpha}A_{\alpha j}^{-1}p_\alpha(z_\alpha^2-1)^{-1/2}\gamma_j(z_\alpha)\},$$

$$\sigma_{k2}=\frac{1}{\pi}\,\mathrm{Re}\sum_\alpha\{B_{k\alpha}A_{\alpha j}^{-1}(z_\alpha^2-1)^{-1/2}\gamma_j(z_\alpha)\}, \tag{49}$$

$$\gamma_j(z_\alpha)=\int_{-1}^1 q_j(\xi)\,\frac{(1-\xi^2)^{1/2}}{z_\alpha-\xi}\,d\xi. \tag{50}$$

The integral in (50) has a singularity at $\xi = z_\alpha = x_1 + p_\alpha x_2$. At $x_2 = 0$ and $|x_1| < 1$, the singularity is on the path of integration. Thus, the stress σ_{k1}, σ_{k2} at the surfaces of inclusion cannot be obtained by simply replacing z_α by x_1 in the integrand. There is an extra contribution due to the integration along the half-circle of very small radius around the singular point $\xi = x_1$. This extra contribution is

$$\mp i\pi q_j(x_1)(1-x_1^2)^{1/2},$$

where $\mp$ is for $x_2 = 0^\pm$. Therefore, if we let

$$y_j(x_1)=\frac{1}{\pi}\int_{-1}^1 q_j(\xi)\,\frac{(1-\xi^2)^{1/2}}{x_1-\xi}\,d\xi, \tag{51}$$

we have

$$\gamma_j(x_1)=\begin{cases}\pi y_j(x_1), & \text{if } |x_1|>1\\ \pi\{y_j(x_1)\mp i(1-x_1^2)^{1/2}q_j(x_1)\}, & \text{if } |x_1|<1.\end{cases} \tag{52}$$

Using (12), (41), and (52), the displacement gradient and stress along $x_2 = 0$ has the following expressions:

$$\begin{aligned} \mathbf{u}_{,1}&=-\mathbf{q},\\ \boldsymbol{\sigma}_1&=(\mathbf{N}_3\mathbf{H}+\mathbf{N}_1^T\mathbf{S}^T)\mathbf{H}^{-1}\mathbf{q}(x_1)\mp(1-x_1^2)^{-1/2}\mathbf{N}_1^T\mathbf{H}^{-1}\mathbf{y}(x_1),\\ \boldsymbol{\sigma}_2&=-\mathbf{S}^T\mathbf{H}^{-1}\mathbf{q}(x_1)\pm(1-x_1^2)^{-1/2}\mathbf{H}^{-1}\mathbf{y}(x_1),\end{aligned} \tag{53}$$

for $|x_1| < 1, x_2 = 0^\pm$, and

$$\mathbf{u}_{,1}=\pm(x_1^2-1)^{-1/2}\mathbf{y}(x_1), \tag{54}$$

$$\begin{aligned} \boldsymbol{\sigma}_1&=\mp(x_1^2-1)^{-1/2}(\mathbf{N}_3\mathbf{H}+\mathbf{N}_1^T\mathbf{S}^T)\mathbf{H}^{-1}\mathbf{y}(x_1),\\ \boldsymbol{\sigma}_2&=\pm(x_1^2-1)^{-1/2}\mathbf{S}^T\mathbf{H}^{-1}\mathbf{y}(x_1),\end{aligned} \tag{55}$$

for $x_1 > 1$ or $< -1, x_2 = 0$. It should be pointed out that if q_j is a constant, the integral in (51) can be integrated to give

$$\frac{1}{\pi}\int_{-1}^1\frac{(1-\xi^2)^{1/2}}{x_1-\xi}\,d\xi$$

$$=\begin{cases}x_1, & \text{if } |x_1|<1\\ (x_1\mp\sqrt{x_1^2-1}), & \text{if } x_1>1, \text{ or } <-1.\end{cases} \tag{56}$$

Equations (53), (54), and (55) then reduce to (42), $(38)_2$, and (43), respectively.

Let

$$2\tau(x_1)=\sigma_2(x_1, 0^+)-\sigma_2(x_1, 0^-),\quad |x_1|<1,$$

be the difference in the surface tractions on both surfaces of the inclusion. Substituting in the above equation from $(53)_2$ we have

$$\mathbf{H}\tau(x_1)=\frac{1}{\pi(1-x_1^2)^{1/2}}\int_{-1}^1\mathbf{q}(\xi)\,\frac{(1-\xi^2)^{1/2}}{x_1-\xi}\,d\xi. \tag{57}$$

Equation (57) provides the relation between $-\mathbf{q}(x_1)$, the disturbed displacement gradient on the inclusion, and $2\tau(x_1)$, the difference in the surface tractions on the inclusion. From the assumption that the inclusion has no bending ridigity, σ_{22} must be continuous which means

$$\tau_2(x_1)=0.$$

From (57) we see that this can be satisfied if we let

$$\{\mathbf{H}^{-1}\mathbf{q}(x_1)\}_2 = 0,$$

or letting $\mathbf{h} = \mathbf{H}^{-1}$,

$$q_2(x_1) = \frac{h_{21}q_1(x_1) + h_{23}q_3(x_1)}{h_{22}}. \tag{58}$$

With $\tau_2(x_1) = 0$, (57) can be written as

$$\hat{\mathbf{H}}\hat{\tau}(x_1) = \frac{1}{\pi(1-x_1^2)^{1/2}} \int_{-1}^{1} \hat{\mathbf{q}}(\xi)\, \frac{(1-\xi^2)^{1/2}}{x_1 - \xi}\, d\xi, \tag{59}$$

$$\hat{\mathbf{H}} = \begin{bmatrix} H_{11} & H_{13} \\ H_{31} & H_{33} \end{bmatrix}, \quad \hat{\tau} = \begin{bmatrix} \tau_1 \\ \tau_3 \end{bmatrix}, \quad \hat{\mathbf{q}} = \begin{bmatrix} q_1 \\ q_3 \end{bmatrix}. \tag{60}$$

We will derive an integral equation for $\hat{\tau}(x_1)$. We see that $\hat{\mathbf{H}}$ is positive definite. We also see from (59) that if $\hat{\tau}(x_1)$ is an odd function in x_1, which is the case as we will show later, $\hat{\mathbf{q}}(x_1)$ and $q_2(x_1)$ of (58) are even functions of x_1.

We consider next stresses and strains at $x_2 = 0$ and, for simplicity, will drop the subscript 1 from x_1. By considering the equilibrium of the segment $(x, 1)$ of the inclusion and using (53) we have

$$\begin{aligned}
\tilde{\sigma}_1^{(i)} &= \sigma_1^{\infty} + \frac{1}{d}\int_x^1 \tau_1(t)dt, \\
\tilde{\sigma}_5^{(i)} &= \sigma_5^{\infty} + \frac{1}{d}\int_x^1 \tau_3(t)dt, \\
\tilde{\sigma}_2^{(i)} &= \sigma_2^{\infty} - S_{j2}(\mathbf{H}^{-1}\mathbf{q})_j, \\
\tilde{\sigma}_4^{(i)} &= \sigma_4^{\infty} - S_{j3}(\mathbf{H}^{-1}\mathbf{q})_j, \\
\tilde{\sigma}_6^{(i)} &= \sigma_6^{\infty} - S_{j1}(\mathbf{H}^{-1}\mathbf{q})_j,
\end{aligned} \tag{61}$$

where $\tilde{\sigma}_i^{(i)}$ is the average stress over the thickness of the inclusion and $2d$ is the nondimensionalized thickness of the inclusion. Using (17) and the relation

$$\tilde{\epsilon}_k^{(i)} = \hat{W}_{kj}^{(i)}\tilde{\sigma}_j^{(i)},$$

(45) can be written as

$$\hat{\mathbf{q}}(x) = \mathbf{g} - \pi\mathbf{M}\int_x^1 \hat{\tau}(t)dt + \mathbf{\Gamma}\mathbf{H}^{-1}\mathbf{q}(x) \tag{62}$$

where $\hat{\mathbf{q}}$ and $\hat{\tau}$ are defined in (60) and

$$\mathbf{g} = \begin{bmatrix} [\hat{W}_{1j}]\sigma_j^{\infty} \\ [\hat{W}_{5j}]\sigma_j^{\infty} \end{bmatrix},$$

$$\mathbf{M} = \frac{1}{\pi d}\begin{bmatrix} \hat{W}_{11}^{(i)} & \hat{W}_{15}^{(i)} \\ \hat{W}_{51}^{(i)} & \hat{W}_{55}^{(i)} \end{bmatrix},$$

$$\mathbf{\Gamma} = \begin{bmatrix} \hat{W}_{16}^{(i)}S_{j1} + \hat{W}_{12}^{(i)}S_{j2} + \hat{W}_{14}^{(i)}S_{j3} \\ \hat{W}_{56}^{(i)}S_{j1} + \hat{W}_{52}^{(i)}S_{j2} + \hat{W}_{54}^{(i)}S_{j3} \end{bmatrix}, \quad j = 1, 2, 3.$$

In the above equation, Γ_{kj} is a 2×3 constant matrix and $\hat{\mathbf{M}}$ is positive definite. If we substitute (62) into (59), and making use of (56) and (57), we have

$$\hat{\mathbf{H}}\hat{\tau}(x) = \frac{1}{(1-x^2)^{1/2}}\left\{x\mathbf{g} - \mathbf{M}\int_{-1}^{1} \hat{\tau}(t)k(x,t)dt\right\} + \hat{\mathbf{\Gamma}}\hat{\tau}(x), \tag{63}$$

in which $\hat{\mathbf{\Gamma}}$ is a 2×2 matrix which is obtained from $\mathbf{\Gamma}$ by deleting the second column of $\mathbf{\Gamma}$. The kernel $k(x, t)$ is given by

$$\begin{aligned}
k(x,t) &= \int_{-1}^{t} \frac{(1-\xi^2)^{1/2}}{x-\xi}\, d\xi, \\
&= (1-x^2)^{1/2}\ln\left|\frac{1-xt+(1-t^2)^{1/2}(1-x^2)^{1/2}}{x-t}\right| \\
&\quad + x(\pi - \cos^{-1}t) - (1-t^2)^{1/2}, \quad |x| < 1.
\end{aligned} \tag{64}$$

Equation $(56)_1$ is recovered by setting $t = 1$ in (64). Noting that $k(-x, -t) = k(x, t) - \pi x$, it is not difficult to show from (63) that $\hat{\tau}(x)$ is an odd function of x.

Equation (63) represents two coupled integral equations for $\tau_1(x)$, and $\tau_3(x)$. Before we show that these two equations can be decoupled under certain conditions, we will look at two special cases. When $\mathbf{g} = \mathbf{0}$, which is (30), $\hat{\tau} = \mathbf{0}$ and, hence, $\tau = \mathbf{0}$ is the solution which means $\mathbf{q} = \mathbf{0}$ by (57). The disturbed solution then vanishes and the homogeneous solution is the solution for the elastic inclusion. This agrees with the result stated in Section 3. Next, consider the special case in which the inclusion is rigid. We then have $\hat{W}_{ij}^{(i)} = 0$ and (63) reduces to

$$\hat{\mathbf{H}}\hat{\tau}(x) = \frac{x}{(1-x^2)^{1/2}}\,\mathbf{g}, \qquad \mathbf{g} = \begin{bmatrix} \hat{W}_{1j}\sigma_j^{\infty} \\ \hat{W}_{5j}\sigma_j^{\infty} \end{bmatrix}. \tag{65}$$

We see from $(56)_1$ that (59) can be reduced to $(65)_1$ if $\hat{\mathbf{q}}$ is a constant given by $\mathbf{g}$ of $(65)_2$. With (58) we conclude that q_2 is also a constant and we have recovered the results obtained in Section 4.

Consider the eigenvalue problem

$$\mathbf{M}\mathbf{e} = \lambda(\hat{\mathbf{H}} - \hat{\mathbf{\Gamma}})\mathbf{e}. \tag{66}$$

Let λ_1, λ_2 be the eigenvalues and $\mathbf{e}_1$, $\mathbf{e}_2$ be the associated eigenvectors. If the inclusion is much "harder" than the matrix, $\hat{\mathbf{H}} - \hat{\mathbf{\Gamma}}$ is dominated by $\hat{\mathbf{H}}$ which is positive definite. If we delete $\hat{\mathbf{\Gamma}}$ from (66), it can be shown (Hildebrand, 1952) that λ_1, λ_2 are real and positive and $\mathbf{e}_1$, $\mathbf{e}_2$ are independent vectors. Introducing the 2×2 matrices $\mathbf{\Omega} = [\mathbf{e}_1, \mathbf{e}_2]$, $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2)$, equation (66) for $\lambda = \lambda_1$, λ_2 can be written in one equation as $\mathbf{M}\mathbf{\Omega} = (\hat{\mathbf{H}} - \hat{\mathbf{\Gamma}})\mathbf{Q}\mathbf{\Lambda}$. Assuming that $\mathbf{\Omega}$ and $\hat{\mathbf{H}} - \hat{\mathbf{\Gamma}}$ are nonsingular, (63) reduces to

$$\hat{\tau}^*(x) = \frac{1}{(1-x^2)^{1/2}}\left\{x\mathbf{g}^* - \mathbf{\Lambda}\int_{-1}^{1} \hat{\tau}^*(t)k(x,t)dt\right\}, \tag{67}$$

$$\hat{\tau}^* = \mathbf{\Omega}^{-1}\hat{\tau}, \quad \mathbf{g}^* = \mathbf{\Omega}^{-1}(\hat{\mathbf{H}} - \hat{\mathbf{\Gamma}})^{-1}\mathbf{g}.$$

Equation (67) is uncoupled into two equations. If we divide the first equation by g_1^* and the second by g_2^*, both equations have the form

$$\tau(x) = \frac{1}{(1-x^2)^{1/2}}\left\{x - \lambda\int_{-1}^{1} \tau(t)k(x,t)dt\right\}, \tag{68}$$

which is a Fredholm integral equation of the second kind.

It should be pointed out that both $\mathbf{M}$ and $\mathbf{\Gamma}$ in (66) vanish for a rigid inclusion and $\lambda = 0$. Hence, λ is small when the inclusion is much harder than the matrix.

6 Solutions of the Integral Equations

Because of the complexity of the kernel there is no analytical solution available for (68). However, under the condition that λ is very small, i.e., the inclusion is much "harder" than the matrix material, (68) can be solved approximately by an asymptotic expansion. More general cases are solved by a numerical discretization.

The solution to (68) is a function of both the space variable x and the eigenvalue λ, i.e., $\tau = \tau(x, \lambda)$. It can be easily seen that for $\lambda = 0$, the solution is simply given by

$$\tau(x,0) = \tau^{(o)}(x) = x(1-x^2)^{-1/2}. \tag{69}$$

For small λ, $\tau(x, \lambda)$ can be expanded into a power series of λ as

$$\tau(x,\lambda) = \sum_{n=0}^{\infty} \lambda^n \tau^{(n)}(x). \tag{70}$$

It can be shown that (70) converges to the true solution for small λ (Tricomi, 1985). By substituting (70) into (68), and equating the like powers of λ, we have

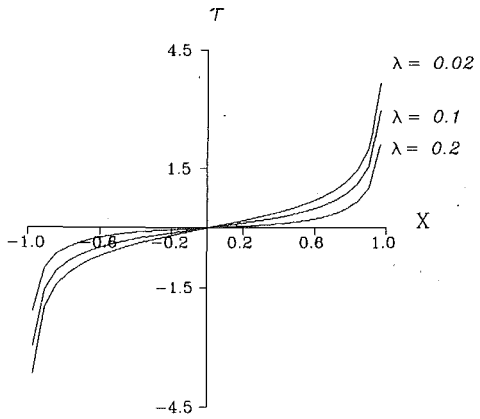


Fig. 2 Asymptotic solutions for τ for various values of λ

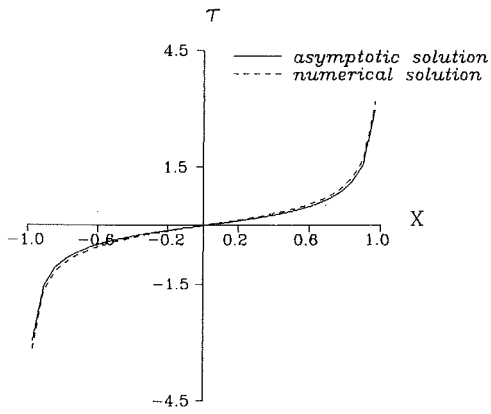


Fig. 3 Asymptotic and numerical solutions for $\lambda = 0.1$

$$\tau^{(n)}(x) = -\frac{1}{(1-x^2)^{1/2}}\int_{-1}^{1} \tau^{(n-1)}(t)k(x,t)dt,$$

$$n = 1, 2, 3 \ldots,$$

in which $\tau^{(o)}(x)$ is given in (69). For $n = 1$ we have

$$\tau^{(1)}(x) = -2\tau^{(o)}(x) - (1-x^2)^{1/2}\ln|(1+x)/(1-x)|. \quad (71)$$

Equation (69) corresponds to the solution for an inextensible rigid inclusion (Erdogan and Gupta, 1972). The solution to (68) for nonzero but small λ can be approximated by taking the first two terms in the expansion

$$\tau(x,\lambda) = \frac{1}{(1-x^2)^{1/2}}\,[x - 2x\lambda - \lambda(1-x^2)\ln|(1+x)/(1-x)|].$$

This is plotted in Fig. 2 for various values of λ.

For the numerical solution, we rewrite (68) as

$$y(x) + \lambda\int_{-1}^{1} y(t)k_1(x,t)dt = x,$$

where

$$y(t) = (1-t^2)^{1/2}\tau(t),$$

$$k_1(x,t) = k(x,t)/(1-t^2)^{1/2}.$$

The results of numerical solutions for $\lambda = 0.1$ and 0.2 are shown in Figs. 3 and 4. Comparisons with the asymptotic solutions shown in the figures indicated that the agreement is excellent for $\lambda < 0.1$.

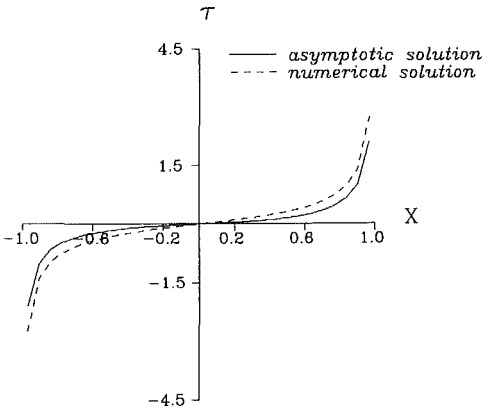


Fig. 4 Asymptotic and numerical solutions for $\lambda = 0.2$

7 Concluding Remarks

The Stroh formalism is employed to analyze the problem of a line inclusion in the general anisotropic elastic solid. For the rigid inclusion, a real and closed-form solution is obtained for the displacement and stress at the inclusion as well as along the extended line of the inclusion. The rotation of the rigid line inclusion is also obtained explicitly. For the elastic inclusion, a pair of Fredhold integral equations for the difference in the surface traction on both sides of the inclusion is derived. The pair can be decoupled and asymptotic and numerical solutions of the integral equations are presented. Regardless of whether the inclusion is rigid or elastic, there is a square root singularity in stress at the tip of the inclusion for general loading at infinity. For certain special loadings, the presence of the inclusion is irrelevant and the homogeneous solution is the solution.

The problem considered here could also be formulated using Eshelby's framework for an ellipsoidal inhomogeneity and taking the limit when the ellipsoid degenerates into a line inclusion. One still obtains a Fredholm integral equation for the unknown strains inside the line inclusion. However, Eshelby's formulation is for three-dimensional problems. In specializing to two-dimensional deformations considered here, many identities in the Stroh formalism, which relate complex expressions to real expressions, cannot be used and the solution remains in a complex form. Another problem one encounters is that for the ellipsoid or elliptic cylinder inclusion, the stress and strain are uniform in the inclusion. In the degenerate case of an elastic line inclusion, the stress and strain are not uniform in the line inclusion. Therefore, while it is possible to obtain the solution for a crack or a rigid line inclusion from that for an elliptic hole or a rigid elliptic inclusion, it is not possible to do so for an elastic line inclusion (Hwu and Ting, 1989).

Acknowledgments

This work is supported in part by the U.S. Army Research Office through grant DAAL 03-88-K-0079 with the University of Illinois at Chicago.

References

Atkinson, C., 1973, "Some Ribbon-Like Inclusion Problems," *Int. J. Engng. Sci.* Vol. 11, pp. 243-265.

Ballarini, R., 1987, "An Integral Equation Approach for Rigid Line Inhomogeneity Problems," *Int. J. Fracture*, Vol. 33, pp. R23-R26.

Barnett, D. M., and Lothe, J., 1973, "Synthesis of the Sextic and the Integral Formalism for Dislocation, Green's Function and Surface Waves in Anisotropic Elastic Solids," *Phys. Norv.*, Vol. 7, pp. 13-19.

Chadwick, P., and Smith, G. D., 1977, "Foundations of the Theory of Surface Waves in Anisotropic Elastic Materials," *Adv. Appl. Mech.*, Vol. 17, pp. 303-376.

Chou, Y. T., and Wang, Z. Y., 1983, "Stress Singularity at the Tip of a Rigid Flat Inclusion," *Recent Developments in Applied Mathematics*, F. F. Ling and I. G. Tadjbakhsh, eds., Rensselaer Press, New York, pp. 21-30.

Erdogan, F., and Gupta, G. D., 1972, "Stress Near a Flat Inclusion in Bonded Dissimilar Materials," *Int. J. Solids Structures*, Vol. 8, pp. 533-547.

Eshelby, J. D., Read, W. T., and Shockley, W., 1953, "Anisotropic Elasticity

With Applications to Dislocation Theory," *ACTA Metallurgica*, Vol. 1, pp. 251–259.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proc. Royal Soc., A.*, Vol. 241, pp. 376–396.

Eshelby, J. D., 1959, "The Elastic Field Outside an Ellipsoidal Inclusion," *Proc. Royal Soc., A.*, Vol. 252, pp. 561–569.

Hildebrand, F. B., 1952, *Methods of Applied Mathematics*, Prentice-Hall, Inc., New York.

Hwu, Chyanbin, and Ting, T. C. T., 1989, "Two-Dimensional Problems of the Anisotropic Elastic Solid With an Elliptic Inclusion," *Q. J. Mech. Appl. Math*, in press.

Li, Qianqian, 1988, "Line Inclusions in Anisotropic Elastic Solids," Ph.D. Thesis, University of Illinois at Chicago.

Mura, T., 1987, *Micromechanics of Defects in Solids*, 2nd ed., Martinus Nijhoff Pub., Boston, Mass., pp. 129–176.

Mura, T., 1988, "Inclusion Problems," *App. Mech. Rev.*, Vol. 41, ASME New York, pp. 15–20.

Muskheleshvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noorhoff Ltd., Groningen.

Selvadurai, A. S. P., 1980, "The Displacements of a Flexible Inhomogeneity Embedded in a Transversely Isotropic Elastic Medium," *Fiber Science and Technology*, Vol. 14, pp. 251–259.

Sendeckyj, G. P., 1970, "Elastic Inclusion Problems in Plane Elastostatics," *Int. J. Solids Structures*, Vol. 6, pp. 1535–1543.

Stroh, A. N., 1958, "Dislocations and Cracks in Anisotropic Elasticity," *Phil. Mag.*, Vol. 3, pp. 625–646.

Stroh, A. N., 1962, "Steady-State Problems in Anisotropic Elasticity," *J. Math. Phys.*, Vol. 41, pp. 77–103.

Ting, T. C. T., 1982, "Effects of Change of Reference Coordinates on the Stress Analyses of Anisotropic Elastic Materials," *Int. J. Solids Structures*, Vol. 18, pp. 139–152.

Ting, T. C. T., 1986, "Explicit Solution and Invariance of the Singularities at an Interface Crack in Anisotropic Composites," *Int. J. Solids Structures*, Vol. 9, pp. 965–983.

Ting, T. C. T., and Hwu, C., 1988, "Sextic Formalism in Anisotropic Elasticity for Almost Non-Semisimple Matrix N," *Int. J. Solids Structures*, Vol. 24, pp. 65–76.

Ting, T. C. T., 1988, "Some Identities and the Structure of Ni in the Stroh Formalism of Anisotropic Elasticity," *Q. Appl. Math.*, Vol. 46, pp. 109–120.

Tricomi, F. G., 1985, *Integral Equations*, Dover Pub., Inc., New York.

Wang, Z. Y., Zhang, H. T., and Chou, Y. T., 1985, "Characteristics of the Elastic Field of a Rigid Line Inhomogeneity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, pp. 818–822.

Wang, Z. Y., Zhang, H. T., and Chou, Y. T., 1986, "Stress Singularity at the Tip of a Rigid Line Inhomogeneity Under Antiplane Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 459–462.

Yang, H. C., and Chou, Y. T., 1982, "The ⟨111⟩ Elliptic Inclusion in an Anisotropic Solid of Cubic Symmetry," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 353–360.

Surface Displacements and Stress Field Generated by a Semi-Ellipsoidal Surface Inclusion

Brian N. Cox
Science Center,
Rockwell International Corp.,
Thousand Oaks, Calif. 91360

This paper presents calculations of the displacement and stress fields generated by semi-ellipsoidal surface inclusions containing uniform transformation strains or eigenstrains. The inclusion is assumed to have the same elastic constants as the rest of the material. This is a reasonable assumption for modeling transformed zones in transformation toughened ceramics and localized plasticity in individual surface grains in alloys. Analytical results are obtained for special cases and numerical results for general cases. The approach is particularly useful for accurately calculating the anomalous fields at the intersection of the boundary of the inclusion and the free surface. It is shown that, in many physically important cases, all components of the stress tensor are zero or constant on the free surface within the inclusion. For shallow inclusions or inclusions of general geometry suffering volume conserving transformation strains, the stress fields are also approximately uniform throughout the inclusion. This result greatly simplifies modeling of localized deformation in certain materials under complex external loads.

1 Introduction

Understanding the fundamental mechanisms of the degradation and failure of composite materials requires experimental and theoretical analysis of events occurring on the scale of the microstructure. For fiber-reinforced composites, transformation toughened ceramics, and polycrystalline alloys, the relevant scale falls between tenths and tens of microns. The only experimental techniques with sufficient spatial resolution to yield data over such gauge lengths are various methods of measuring surface displacements. Outstanding amongst these are stereoscopy (e.g., Williams et al., 1980; Cox et al., 1986) and digital image analysis (James et al., 1988), which can measure differential surface strains over submicron gauge lengths from pairs of SEM micrographs (e.g., Morris et al., 1988). Techniques of measuring surface displacements with inferior but still useful spatial resolution include moire interferometry and the tracking of very fine grids of holes cut in surface overlayers (Bradley, 1987). Techniques for measuring bulk or subsurface strains, such as X-ray diffraction, Raman spectroscopy, and optical birefringence, are restricted in the materials to which they can be applied and fail to provide the required spatial resolution.

The deduction of information concerning surface and subsurface deformation from measurements of surface displacements alone poses special problems. Given no other information, the surface displacements are not necessarily sufficient to determine a unique solution. Although solutions have been found in special cases (Mura et al., 1986), the problem is, in general, ill-posed. However, unique solutions may often be found if other restrictions can be placed on them. Gao and Mura (1988) have pursued a mathematical approach showing that the requirement that the L_2 norm of the deformation be a minimum is sufficient to render the problem well posed. A physically-based approach can be equally successful, since the cumulative insight of experiments and theoretical models usually specifies strict limits on the admissible forms of the deformation. In this case, one proceeds by postulating parametric models of the deformation and optimizing the parameters by comparison of calculated surface displacements with data. To construct manageable parametric models of the deformation and to perform optimization of the parameters, it is very helpful to carry out as much of the modeling as possible analytically and to have efficient numerical procedures in the absence of analytical results.

This paper presents analytical and numerical results for one such model system, which has proven very useful in analyzing surface displacement data in individual grains of aluminum alloys (Morris et al., 1987; Cox et al., 1987) and transformed zones around cracks in magnesia partially stabilized zirconia (Cox et al., 1988). The case considered is that of a semi-ellipsoidal inclusion that suffers a stress-free transformation strain (Robinson, 1951; Eshelby, 1957) or eigenstrain (Mura, 1982), ϵ^t_{ij}, defining the change in shape the inclusion would suffer in the absence of the constraining matrix. The inclusion has semi-axes a and b in the plane (x_1, x_2) of the free surface, and c in the normal direction x_3 (Fig. 1). Both the inclusion and the matrix are assumed to be homogeneous and isotropic, and have the same elastic constants. The last assumption is

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

Discussion on this paper should be addressed to the 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 14, 1988; final revision, September 12, 1988.

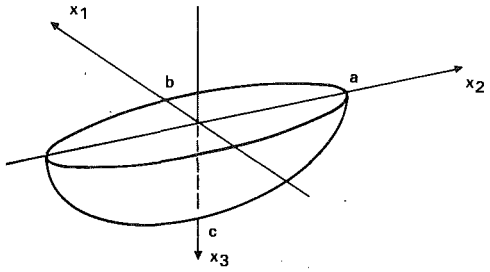


Fig. 1 The semi-ellipsoidal transformed zone, showing its relationship to the coordinate system

reasonable for studying transformed zones in transformation toughened ceramics and plastic deformation in metal alloys. It also simplifies the calculation considerably, allowing more general choices of geometry and eigenstrain. Theories of elastically inhomogeneous inclusions have been restricted to semi-spheroidal inclusions undergoing axially symmetric eigenstrains (Kouris and Mura, 1988a and 1988b).

The computational procedure comprises the following steps:

(1) Evaluate Eshelby's analytic solutions for the stresses $\sigma_{ij}^{(1)}$ and strains $\epsilon_{ij}^{(1)}$ inside (Eshelby, 1957) and outside (Eshelby, 1959) an elliposidal inclusion in an infinite medium.

(2) Introduce a free surface on the plane $x_3 = 0$ by applying a continuum of normal and in-plane point forces, t_i, to cancel exactly the normal stress $\sigma_{33}^{(1)}$ and shear stresses $\sigma_{31}^{(1)}$ and and $\sigma_{23}^{(1)}$ calculated to exist there from Step 1. Find the corrections to the displacements, stresses, and strains calculated in Step 1 associated with the introduction of the free surface by integrating the products of t_i and Mindlin's (1936) Green's functions for a half space over the plane $x_3 = 0$.

These steps can be followed to calculate the displacements, strains, and stresses anywhere in the half space with equal facility. In the work described below, complete stress fields and surface displacements will be reported for several cases. The chosen approach is useful because it is very efficient computationally and it allows certain results to be found analytically. A similar approach has been followed previously to calculate stresses around hemispherical indentations (Chiang et al., 1982).

2 Numerical Methods

The evaluation of Eshelby's analytic solutions for an ellipsoidal inclusion (Step 1) requires lengthy but straightforward programming and follows the original papers. All stresses $\sigma_{ij}^{(1)}$ and strains $\epsilon_{ij}^{(1)}$ are found point by point by evaluating explicit expressions involving certain elliptic integrals. Programs have been written for which the input is (1) the semi-axes a, b, and c; (2) the coordinates $\mathbf{x}$ of a given point of interest; and (3) Poisson's ratio. The output is a matrix S of rank six defined so that

$$\epsilon_\alpha^{(1)}(\mathbf{x}) = \sum_\beta \epsilon_\beta^t S_{\beta\alpha}^c(\mathbf{x}), \quad \alpha = 1, \ldots, 6, \tag{1}$$

with the notation for the row vectors $\epsilon_\alpha^{(1)}$ and ϵ_α^t that their components refer in order to the components ϵ_{11}, ϵ_{22}, ϵ_{33}, ϵ_{12}, ϵ_{23}, and ϵ_{31} of the corresponding strain tensors $\epsilon_{ij}^{(1)}$ and ϵ_{ij}^t. Corresponding displacements $u_i^{(1)}$ can be found by integration of $\epsilon_{ij}^{(1)}$ and stresses $\sigma_{ij}^{(1)}$ directly from the constitutive relations.

When two or more of the semi-axes a, b, and c are equal, Eshelby's expressions become indeterminate. On a 32-bit machine with approximately seven-figure precision, one finds, in practice, that errors of less than one part in 10^4 are suffered if the semi-axes differ by at least 1 percent. For other cases, i.e., the near-spheroid or near-sphere, acceptably accurate solutions may be found by interpolating between solutions for nonspheroidal cases. For example, the solution when $a = (1+\delta)b$, $|\delta| < 0.01$, may be found by interpolating between the solutions for $a = 0.99b$ and $a = 1.01b$. (Note that all aspects of the solutions are smooth functions of a/b at $a/b = 1$.)

Step 2 requires evaluation of the integrals

$$u_i^{(2)}(\mathbf{x}) = \int\int G_{ij}(\mathbf{x}-\mathbf{x}')t_j(\mathbf{x}')dx_1'dx_2' \tag{2a}$$

and

$$\sigma_{ij}^{(2)}(\mathbf{x}) = \int\int H_{ijk}(\mathbf{x}-\mathbf{x}')t_k(\mathbf{x}')dx_1'dx_2', \tag{2b}$$

where $G_{ij}(\mathbf{x}-\mathbf{x}')$ and $H_{ijk}(\mathbf{x}-\mathbf{x}')$ are Mindlin's (1936) Green's functions for the displacements and stresses generated in a half space $x_3 \geq 0$ by point forces acting at $\mathbf{x}'$ on $x_3 = 0$. All such integrals are conveniently divided into two contributions, one, $u_i^{\text{in}}(\mathbf{x})$ or $\sigma_{ij}^{\text{in}}(\mathbf{x})$, arising from integration over the domain of the inclusion only, and the other, $u_i^{\text{out}}(\mathbf{x})$ or $\sigma_{ij}^{\text{out}}(\mathbf{x})$, arising from integration over the remainder of the plane $x_3 = 0$. Note especially that the contributions u_i^{in} and σ_{ij}^{in} are dominant when the inclusion is deep (large c) and either $\epsilon_{ij}^t = \delta_{ij}$ (hydrostatic expansion) or $\epsilon_{ij}^t = \delta_{ij}\delta_{i3}$ (expansion in the normal direction only).

The total displacement and stress fields are given by

$$u_i(\mathbf{x}) = u_i^{(1)}(\mathbf{x}) + u_i^{(2)}(\mathbf{x}) = u_i^{(1)}(\mathbf{x}) + u_i^{\text{in}}(\mathbf{x}) + u_i^{\text{out}}(\mathbf{x}) \tag{3a}$$

and

$$\sigma_{ij}(\mathbf{x}) = \sigma_{ij}^{(1)}(\mathbf{x}) + \sigma_{ij}^{(2)}(\mathbf{x}) = \sigma_{ij}^{(1)}(\mathbf{x}) + \sigma_{ij}^{\text{in}}(\mathbf{x}) + \sigma_{ij}^{\text{out}}(\mathbf{x}). \tag{3b}$$

From Eshelby (1957, 1959), it is known that the stress $\sigma_{23}^{(1)}$ or $\sigma_{31}^{(1)}$ to be cancelled in Step 2 is nonzero on $x_3 = 0$ if ϵ_{23}^t or ϵ_{31}^t is nonzero. Thus, for many important cases, including purely dilatational transformations and plastic deformation in surface grains in alloys, when the only nonzero shear strain is ϵ_{12}, only normal point forces $t_3(\mathbf{x}')$ need appear in equations (2).

2.1 Surface Displacements. For surface displacements $u_i(x_1, x_2, 0)$, the Green's function in equation (2a) has one of the forms $1/r$, $\cos\theta/r$, or $\sin\theta/r$, where $r = |\mathbf{x}-\mathbf{x}'|$ and θ is the angle between $\mathbf{x}-\mathbf{x}'$ and the x_1-axis. The contribution $u_i^{\text{in}}(\mathbf{x})$ can therefore be reduced easily to a one-dimensional integral over a finite domain, which can be calculated accurately in negligible time. The contribution $u_i^{\text{out}}(\mathbf{x})$ is found as follows. The domain outside the inclusion on (x_1, x_2) is divided into a triangular mesh bounded on the outside by an ellipse whose semi-axes are a multiple R of a and b. In practice, it is always adequate to set $R = 5$. The tractions t_i are evaluated in step 1 at each vertex of the grid and at the midpoint of each side of each triangular element. Each of the three components t_i is then approximated in each element by a quadratic function of x_1 and x_2, with the coefficients found by interpolating over the values at the vertices and midpoints. The contribution to $u_i^{(2)}(\mathbf{x})$ from each element then comprises integrals of functions of the form x_1/r, $x_1\cos\theta/r$, x_1^2/r, etc., which can be found analytically, so that the $1/r$ singularity in Mindlin's Green's function presents no numerical difficulty. For small c/a or c/b, the traction t_3 becomes concentrated outside the inclusion to the neighborhood of the interface, and a modest increase becomes necessary in the density of the grid there.

The method of solution just outlined is particularly efficient if surface displacements are the only information required, as is often the case in interpreting experimental measurements. The problem has been reduced to evaluating a numerically simple, two-dimensional integral for each displacement component at each point. For example, a scan of the net displacements $u_i(\mathbf{x})$, $i = 2$ or 3, at 30 points along the x_2-axis

takes approximately 20 seconds of CPU time on a VAX 11/780 (a 32-bit minicomputer). To obtain the same information from a three-dimensional finite element program would take much longer. Furthermore, the present approach has no difficulty in treating the anomalous displacement and stress fields found at the intersection of the free surface and the interface of the inclusion and the matrix.

2.2 Subsurface Stresses. While surface displacements often have special importance, being the only experimentally accessible quantity over microscopic gauge lengths, the same approach can of course be used to calculate subsurface displacements or stresses (equation (2*b*)). Concentrating on the stresses, from which the displacements can always be deduced, one again finds that the contributions $\sigma_{ij}^{(\mathrm{in})}$ can, in every case, be reduced to a numerically simple, one-dimensional integral. This is particularly useful for evaluating the anomalous stresses near the surface on the boundary of the inclusion, where finite element methods are hard pressed.

The contributions $\sigma_{ij}^{(\mathrm{out})}$ are conveniently evaluated using a triangular mesh similar to that used for calculating the surface displacements. When $x_3 \neq 0$ in equations (2), the Green's functions H_{ijk} and G_{ij} are more complicated algebraically, but they are never singular. Therefore, the integrals in equation (2*b*) are evaluated in a slightly different way. The entire integrand, rather than just t_i, is approximated by a quadratic function in x_1' and x_2'. The resulting integrals of functions of the form x_1', $x_1'^2$, $x_1'x_2'$, etc., are evaluated analytically in each triangle. Stresses on the surface $x_3 = 0$, where H_{ijk} can be singular, are easily found by extrapolating from values at subsurface points.

3 Results

For brevity in what follows, most of the results presented are for eigenstrains in which $\epsilon_{23}^t = \epsilon_{31}^t = 0$, so that only the normal tractions $t_3(\mathbf{x}')$ are nonzero in equation (2.2). (In-plane tractions t_1 and t_2 are required outside the inclusion when $\epsilon_{23}^t = \epsilon_{31}^t = 0$ only if the free surface to be created in Step 2 is other than an equatorial plane of the whole ellipsoid.) Some analytical results are presented for surface displacements generated by tractions t_1 and t_2 in the plane strain limit. In general cases, the numerical treatment of in-plane tractions would be completely analogous to that of normal tractions. Note that the condition $\epsilon_{23}^t = \epsilon_{31}^t = 0$ is consistent at a free surface in the Levy-Mises description of plasticity (e.g., Hill, 1950).

3.1 Conditions for u_i^{out} and $\sigma_{ij}^{\mathrm{out}}$ to be Negligible. A survey of the stresses $\sigma_{33}^{(1)}$ calculated on the plane $x_3 = 0$ in Step 1 (Eshelby's problem) can be summarized as follows. When $\epsilon_{ij}^t = \delta_{ij}$ (hydrostatic expansion) or $\epsilon_{ij}^t = \delta_{ij}\delta_{i3}$ (expansion in the normal direction only), $\sigma_{33}^{(1)}$ becomes small and diffuse outside the inclusion when $c \gg a$ and $c \gg b$. In such cases, u_i^{out} and $\sigma_{ij}^{\mathrm{out}}$ become negligible, and the net displacements and stresses are simply the sum of the solution to Eshelby's problem and u_i^{in} or $\sigma_{ij}^{\mathrm{in}}$. Since the problem is entirely linear, it follows that the same simplification applies for any combination $\epsilon_{ij}^t = \alpha\delta_{ij} + \beta\delta_{ij}\delta_{i3}$. For other ϵ_{ij}^t, $\sigma_{33}^{(1)}$ retains comparable magnitudes inside and outside the inclusion for large c, and u_i^{out} and $\sigma_{ij}^{\mathrm{out}}$ must be taken into account.

3.2 Surface Displacements. The surface displacements can sometimes be found analytically. Such cases illustrate qualitative characteristics of more general cases, which are useful in interpreting experimental surface displacement data. These analytical results are presented next, along with representative numerical calculations. The various possibilities are conveniently divided according to symmetry.

Axially Symmetric Cases. For a spheroidal inclusion with $a = b$, the contribution $u_i^{\mathrm{in}}(0, x_2, 0)$ can be found analytically:

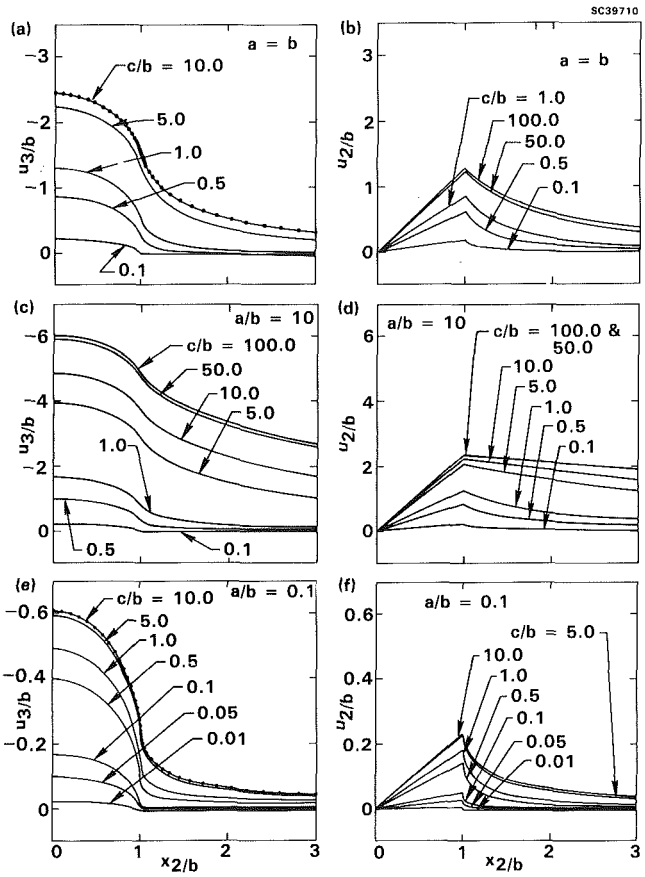


Fig. 2 The normal displacement $u_3(x)$ ((*a*), (*c*), and (*e*)) and the in-plane displacement $u_2(x)$ ((*b*), (*d*), and (*f*)) when $\epsilon_{ij}^t = \delta_{ij}$ for the cases $a = 1$ and $b = 1$ ((*a*) and (*b*)); $a = 10$ and $b = 1$ ((*c*) and (*d*)); and $a = 1$ and $b = 10$ ((*e*) and (*f*)). The semi-axis c has the values marked, and Poisson's ratio was taken to be 0.3. The dots in (*a*) and (*e*) indicate the values of x_2/b at which u_3 was calculated, with values elsewhere found by Lagrangian interpolation.

$$u_3^{\mathrm{in}}(0,x_2,0) = t_3(\mathbf{0}) \cdot \frac{1-\nu}{\pi\mu} \left\{ (a-x_2)K\left[\frac{2\sqrt{ax_2}}{a+x_2}\right] + (a+x_2)E\left[\frac{2\sqrt{ax_2}}{a+x_2}\right] \right\}, \tag{4}$$

where ν is Poisson's ratio, μ the shear modulus, and K and E are complete elliptic integrals of the first and second kinds; and

$$u_2^{\mathrm{in}}(0,x_2,0) = \begin{cases} -t_3(\mathbf{0})\,\dfrac{1-2\nu}{4\pi\mu}\,x_2 & |x_2| < a \\ -t_3(\mathbf{0})\,\dfrac{1-2\nu}{4\pi\mu}\,\dfrac{a}{x_2} & |x_2| > a. \end{cases} \tag{5}$$

Thus, the partial derivative $\partial u_3/\partial x_2$ possesses a logarithmic singularity at the interface $x_2 = a$ on the free surface, since, as $k \to 1$, $K(k) \to 1/2 \ln[16/(1-k)]$ (Byrd and Friedman, 1971). Note, however, that the shear strain ϵ_{23} vanishes on the free surface $x_3 = 0$, so that $\partial u_2/\partial x_3 = -\partial u_3/\partial x_2$. These displacement gradients correspond to a rotation through $\pi/2$ of the element at the intersection of the interface and the free surface about an axis tangential to that intersection. The strain ϵ_{22}, on

the other hand, remains constant within the inclusion, since $u_2^{\text{out}}(0,x_2,0)=0$ for $x_2<a$, a result which follows, as does equation (5), from the fact that for a surface point $\mathbf{x}$

$$\oint_o^{2\pi} G_{23}(\mathbf{x}-\mathbf{x}')d\phi = -\frac{1-2\nu}{4\pi\mu}\oint_o^{2\pi}\frac{x_2-x_2'}{(x_1-x_1')^2+(x_2-x_2')^2}\,d\phi$$

$$= \begin{cases} 0 & \text{if } x_1'^2+x_2'^2<\rho^2 \\ -\dfrac{1-2\nu}{2\mu}\dfrac{x_2'}{x_1'^2+x_2'^2} & \text{if } x_1'^2+x_2'^2>\rho^2\ , \end{cases} \tag{6}$$

where the line integral is evaluated around a circle of radius ρ centered on the origin. This conclusion holds for any axially symmetric, uniform, stress-free transformation strain ϵ_{ij}^t. Since the x_2-axis can be chosen arbitrarily when $a=b$ and ϵ_{ij}^t is axially symmetric, it follows that all components of the stress tensor are uniform on $x_3=0$ within the transformed zone. It will be seen below that numerical results show that this conclusion holds far more generally.

Numerical calculations of the net surface displacements $u_i(\mathbf{x})$, including the contributions u_i^{out}, are shown in Figs. 2(*a*) and 2(*b*) for the case of hydrostatic expansion $\epsilon_{ij}=\delta_{ij}$. As the ratio of depth to surface radius (i.e., c/a) increases, the surface displacements tend asympotically to limiting values. For $c/a>10$, they exhibit no further significant change with increasing depth. However, it should be remembered that one usually measures not absolute displacements but rather strains or relative displacements. For surface strains or the relative displacements $u_i(\mathbf{x})-u_i(\mathbf{0})$ measured over gauge lengths comparable to a and b, the calculated results are already within ~ 1 percent of their asymptotic values in the limit $c\to\infty$ when $c\simeq 2a$. This is found to be the case for all transformations ϵ'.

Plane Strain Cases. In the limit $a/b\to\infty$, the contribution u_3^{in} is unbounded, but the partial derivative $\partial u_3^{\text{in}}/\partial x_2$ remains finite. For normal surface tractions t_3, it is given along the x_2-axis by

$$\frac{\partial u_3^{\text{in}}}{\partial x_2} = \frac{1-\nu}{2\pi\mu}\,t_3(\mathbf{0})\int_{-b}^{b}\int_{-\infty}^{\infty}\frac{x_2'-x_2}{[x_1'^2+(x_2'-x_2)^2]^{3/2}}\,dx_1'dx_2'$$

$$= \frac{1-\nu}{2\pi\mu}\,t_3(\mathbf{0})\,\ln\left|\frac{b-x_2}{b+x_2}\right|, \tag{7}$$

which expression again exhibits a logarithmic singularity. Hence,

$$u_3^{\text{in}}(0,x_2,0)-u_3^{\text{in}}(\mathbf{0}) = \frac{1-\nu}{\pi\mu}\,(-t_3(\mathbf{0}))[(x_2-b)\ln|1-x_2/b| - (x_2+b)\ln|1+x_2/b|]. \tag{8}$$

For the in-plane displacements, one finds $u_1^{\text{in}}=0$, and

$$u_2^{\text{in}}(0,x_2,0) = \begin{cases} -t_3(\mathbf{0})\cdot\dfrac{1-2\nu}{2\mu}\cdot x_2 & |x_2|\le b \\ -t_3(\mathbf{0})\cdot\dfrac{1-2\nu}{2\mu}\cdot b\cdot\dfrac{x_2}{|x_2|} & |x_2|\ge b. \end{cases} \tag{9}$$

For in-plane tractions t_2, the results for u_2^{in} and u_3^{in} are reversed:

$$\frac{\partial u_2^{\text{in}}(0,x_2,0)}{\partial x_2} = \frac{t_2(\mathbf{0})}{\pi\mu}\cdot\ln\left|\frac{b-x_2}{b+x_2}\right| \tag{10}$$

and

$$u_3^{\text{in}}(0,x_2,0) = \begin{cases} \dfrac{1-\nu}{2\mu}\cdot t_2(\mathbf{0})\cdot x_2 & |x_2|\le b \\ \dfrac{1-2\nu}{2\mu}\cdot t_2(\mathbf{0})\cdot b\cdot\dfrac{x_2}{|x_2|} & |x_2|\ge b. \end{cases} \tag{11}$$

For in-plane tractions, t_1, one has $u_2^{\text{in}}=u_3^{\text{in}}=0$ and

$$\frac{\partial u_1^{\text{in}}(0,x_2,0)}{\partial x_2} = \frac{t_1(\mathbf{0})}{\pi\mu}\cdot\ln\left|\frac{b-x_2}{b+x_2}\right|. \tag{12}$$

General Geometries. The in-plane components u_1^{in} and u_2^{in} can still be obtained analytically within the inclusion when $a\neq b$ and $\epsilon_{23}^t=\epsilon_{31}^t=0$. For u_2^{in} one finds

$$u_2(x_1,x_2,0) = -t_3(\mathbf{0})\cdot\frac{1-2\nu}{4\pi\mu}\int_\epsilon\int\frac{x_2-x_2'}{(x_1-x_1')^2+(x_2-x_2')^2}\,dx_1'dx_2'$$

$$= -t_3(\mathbf{0})\,\frac{1-2\nu}{2\mu}\cdot\left(\frac{b}{a+b}\,x_2\right)\quad (x_1/a)^2+(x_2/b)^2<1, \tag{13}$$

where ϵ is the elliptical intersection of the inclusion and the free surface. This expression is independent of x_1 and linear in x_2, showing that this contribution to the strain $\partial u_2/\partial x_2$ inside the inclusion is always uniform.

When c is not large or $\epsilon_{ij}^t\neq\alpha\delta_{ij}+\beta\delta_{ij}\delta_{i3}$, the numerically evaluated contributions u_i^{out} become significant. The net surface displacements u_2 and u_3, comprising the superposition of Steps 1 and 2, are presented for the case of hydrostatic expansion ($\epsilon_{ij}^t=\delta_{ij}$) in Figs. 2(*c*)–2(*f*). The displacements are shown along the x_2-axis, which is the minor semi-axis in Figs. 2(*c*) and 2(*d*) and the major semi-axis in Figs. 2(*e*) and 2(*f*). Along the major semi-axis, the displacements are similar to those for the spheroidal inclusion (Figs. 2(*a*) and 2(*b*)), except that u_3 falls away much more rapidly at and beyond the interface. Along the minor semi-axis, u_3 exhibits only a mild drop from the center of the inclusion to the interface, and falls away gradually far out into the matrix.

A series of numerical calculations showed that, even when u_i^{out} was not negligible, the net in-plane surface displacements u_1 and u_2 were always linear and the surface stresses σ_{11} and σ_{22} therefore uniform within the inclusion when $\epsilon_{23}^t=\epsilon_{31}^t=0$. This was shown to be true within the four-figure accuracy of the calculations over many randomly chosen intervals of the surface inside the inclusion. The three transformation strains $\epsilon_{ij}^t=\delta_{ij}$, $\epsilon_{ij}^t=\delta_{ij}\delta_{i3}$, and $\epsilon_2^t=1$, $\epsilon_3^t=-1$, all other components zero were considered. Since $\sigma_{33}^{(1)}=0$ on $x_3=0$ for the pure shear transformation ϵ_{12}^t, all transformation strains for which $\epsilon_{23}^t=\epsilon_{31}^t=0$ can be constructed from these cases. The in-plane surface displacements were found to be linear in the inclusion for all choices of a, b, and c.

The importance of assuming that the inclusion has the same elastic constants as the matrix can be assessed by comparison with the work of Kouris and Mura (1987) for inhomogeneous semispheroidal inclusions. The case of a hemispherical inclusion suffering the eigenstrain $\epsilon_{ij}^t=\delta_{ij}\delta_{i3}$ (e.g., differential thermal expansion in the normal direction x_3) is presented in Fig. 3. The calculation of Kouris and Mura (dashed curve) is for the case of equal Poisson's ratio in the inclusion and the matrix and for $\Gamma=2.0$, where Γ is the ratio of the shear modulus or Young's modulus in the inclusion to that in the matrix. For the displacements u_2 and u_3, there is strikingly little difference between the two calculations. For the interpretation of experimental measurements of either in-plane or normal surface displacements, it is clear that the assumption of equal elastic constants will be a very reasonable one in many cases.

3.3 Stress Fields. The stress fields at the intersection of an interface and a free surface are well known to be anomalous (Bogy, 1975). Once again it is very helpful in studying this region to separate the contributions σ_{ij}^{in} and σ_{ij}^{out}, and to obtain analytical results in limiting cases. In the following, Young's modulus is denoted E.

3.3.1 The Contribution σ_{ij}^{in}.

Plane Strain. When $a \to \infty$, σ_{ij} can be evaluated analytically from equation (2b). One has

$$\sigma_{22}^{in}(x_2,x_3) = \frac{x_3 t_3(0)}{\pi}\left[\frac{x_2+b}{R_2^2} - \frac{x_2-b}{R_1^2}\right] + \frac{t_3(0)}{\pi}\left[\tan^{-1}\left(\frac{x_2-b}{x_3}\right) - \tan^{-1}\left(\frac{x_2+b}{x_3}\right)\right] \quad (14a)$$

$$\sigma_{33}^{in}(x_3,x_3) = \frac{x_3 t_3(0)}{\pi}\left[\frac{x_2-b}{R_1^2} - \frac{x_2+b}{R_2^2}\right] + \frac{t_3(0)}{\pi}\left[\tan^{-1}\left(\frac{x_2-b}{x_3}\right) - \tan^{-1}\left(\frac{x_2+b}{x_3}\right)\right] \quad (14b)$$

$$\sigma_{23}^{in}(x_2,x_3) = \frac{t_3(0)}{\pi}\left[\frac{x_3^2}{R_2^2} - \frac{x_3^2}{R_1^2}\right] \quad (14c)$$

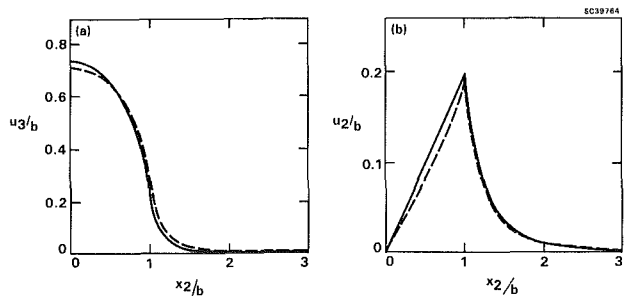


Fig. 3 Comparison of calculations for a hemispherical inclusion undergoing the transformation $\epsilon_{ij}^t = \delta_{ij}\delta_{i3}$. Solid curve: present calculations. Dashed curve: Kouris and Mura (1987), with $\Gamma = 2.0$, where Γ is the ratio of the Young's moduli in the inclusion and the matrix; for Poisson's ratio equal to 0.3.

where $R_1^2 = (x_2-b)^2 + x_3^2$ and $R_2^2 = (x_2+b)^2 + x_3^2$. Of course, these results may be found alternatively from Flamant's (1892) solution for normal line forces applied to a half space. Note that, on the free surface $x_3 = 0$, σ_{22}^{in} is uniform when $|x_2| < b$ and zero when $|x_2| > b$, in concurrence with equation (9). When $|x_2| = b$, the maximum magnitude of σ_{23}^{in} occurs on the free surface $x_3 = 0$. When $|x_2| \neq b$, $\sigma_{23}^{in} = 0$ when $x_3 = 0$, and the maximum magnitude of σ_{23}^{in} is found below the surface. These characteristics of σ_{23}^{in} are also found when $a \neq b$, and persist in the net stress σ_{23}.

General Geometries. For cases of finite a and b, $\sigma_{ij}^{in}(x_1,x_2,x_3)$ is conveniently evaluated in polar coordinates (r,θ) in the plane $x_3 = 0$, with the origin at (x_1,x_2). The integral over r can be performed analytically for all cases, and the ensuring integral over θ is easily calculated using cubic splines. All stresses σ_{ij}^{in} can thus be calculated quickly and accurately, even near the intersection of the interface and the free surface. Complete results are presented for the case $a = b$ in Fig. 4. They are qualitatively representative of the general case $a \neq b$ and similar to the analytical expressions equations (14) for plane strain. Each curve shows a scan parallel to the x_3-axis when $x_1 = 0$ and x_2/b has the value marked. Note that all stresses vary very rapidly near $(0,b,0)$. The stress σ_{22}^{in} is discontinuous at $(0,b,0)$, but only at that point since, along the x_2-axis, it is the component of stress normal to the interface. The stresses σ_{11}^{in} and σ_{22}^{in} are constant everywhere on the surface of the inclusion, in accordance with the surface displacements of equation (13).

As for surface displacements, the sum of σ_{ij}^{in} and the stresses calculated in Step 1 is asymptotically equal to the net stress when c is large and $\epsilon_{ij}^t = \alpha\delta_{ij} + \beta\delta_{ij}\delta_{i3}$.

3.3.2 Total Stress Fields for Some Important Cases. For cases where c is not large or $\epsilon_{ij}^t \neq \alpha\delta_{ij} + \beta\delta_{ij}\delta_{i3}$, the contributions σ_{ij}^{out} are significant. The numerical solutions presented next were obtained according to the procedures described in Section 2.

Hydrostatic Dilation. The case of hydrostatic expansion, $\epsilon_{ij}^t = \delta_{ij}$, for a spheroidal inclusion $(a = b)$ is shown is Figs. 5, 6, and 7. This case may arise, for example, from thermal

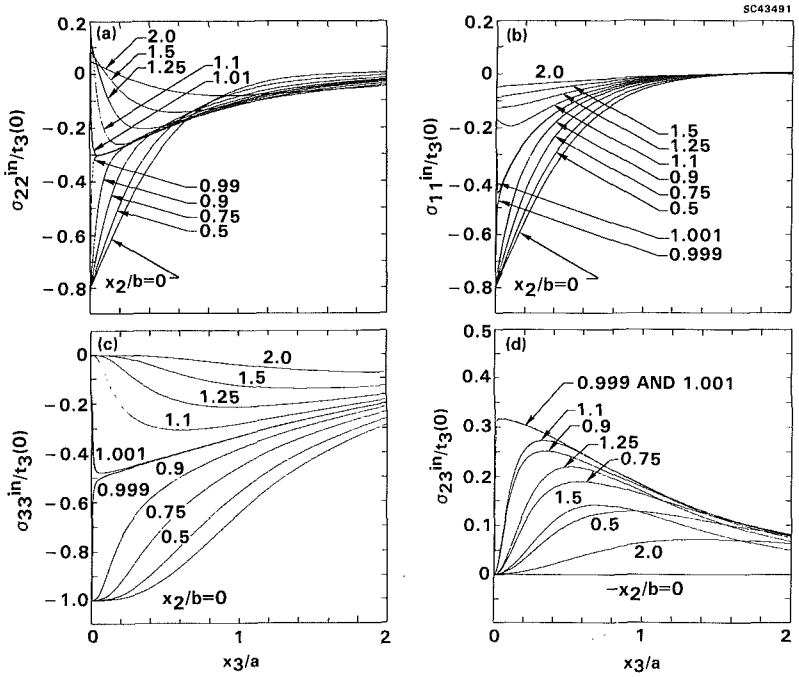


Fig. 4 The contribution σ_{ij}^{in} as a function of x_3 for $x_2 = 0$ and the marked values of x_1 when only the normal tractions t_3 are nonzero for the case $a = b$. Positive t_3 corresponds to tractions directed into the half space $x_3 \geq 0$.

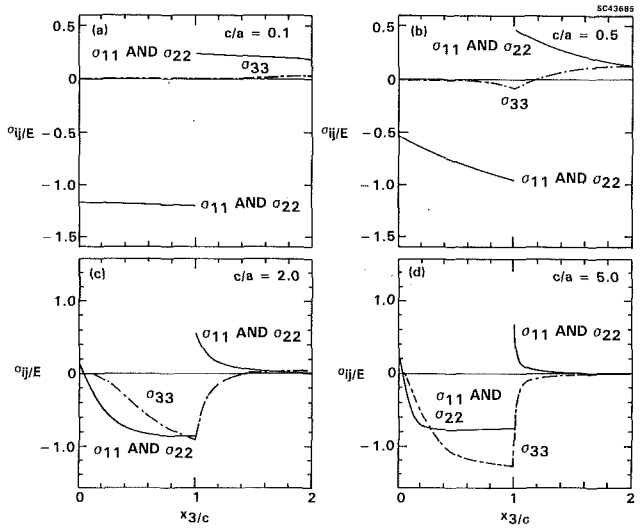


Fig. 5 The net stress σ_{ij} down the x_3-axis for the marked ratios c/a when $a=b$ and $\epsilon^t_{ij}=\delta_{ij}$

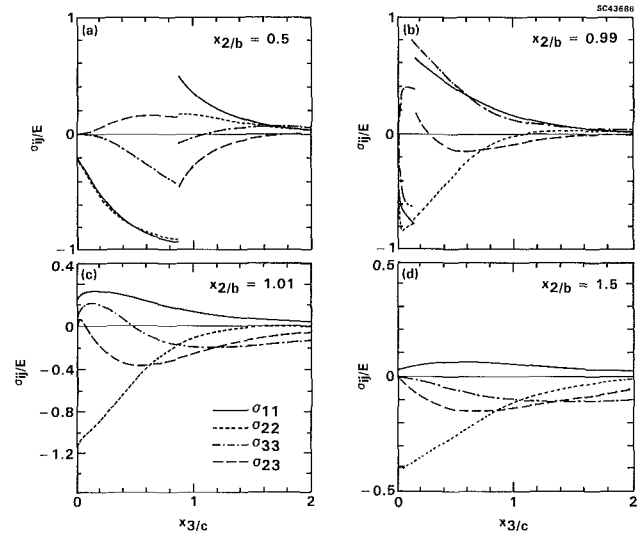


Fig. 6 The net stress σ_{ij} as a function of x_3 when $x_1=0$ and x_2 has the values marked, for $a=b=c$ and $\epsilon^t_{ij}=\delta_{ij}$

mismatch between inclusion and matrix, inhomogeneous densification during sintering, and phase transformations, e.g., in phase transformation toughened alloys and ceramics. The stresses are shown down the x_3-axis in Fig. 5 for various ratios c/a. For shallow inclusions (Fig. 5(a)), the stresses are very nearly uniform throughout the depth of the inclusion. (This has been found to be generally the case for any transformations for which $\epsilon^t_{23}=\epsilon^t_{31}=0$). The stress σ_{33} is tensile beneath the inclusion. As c/a increases, the stresses all become nonuniform within the inclusion. For $c/a \gtrsim 1.5$, the stresses σ_{11} and σ_{22} switch from being compressive on the surface of the inclusion to being tensile. This result is significant in the study of processing inhomogeneities in sintered ceramics. Microcracking is sometimes found to occur at the apex of surface breaking spheroids of unusually dense material (Lange and Metcalf, 1983; Lange, 1988). The same phenomenon is found for inclusions for which $a \neq b$. As a/b increases, σ_{22} is tensile at the surface for lower and lower values of c/b, until, for $a/b \to \infty$, the critical value of c/b is $\simeq 0.75$.

The stresses throughout and around the inclusion are illustrated in Fig. 6 for the hemispherical case $a=b=c$. Each curve in Fig. 6 shows a component of stress as a function of x_3 for $x_1=0$ and the marked values of x_2. Near the interface (Figs. 6(b) and 6(c)), the stresses vary very rapidly within the

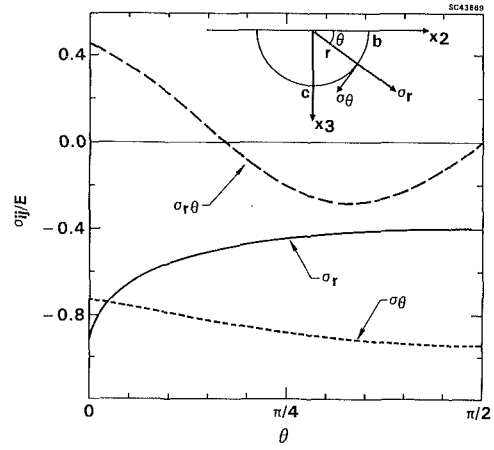


Fig. 7 Interface stress components σ_r, σ_θ, and $\sigma_{r\theta}$ resolved into the coordinates of the interface (see inset) as functions of the angle θ, for the case $a=b=c$ and $\epsilon^t_{ij}=\delta_{ij}$

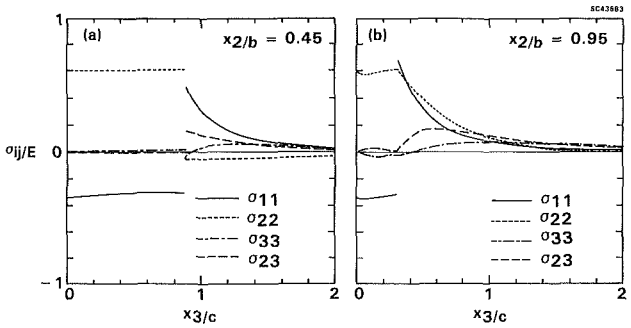


Fig. 8 The net stress σ_{ij} as a function of x_3 for $x_1=0$ and the marked values of x_2 when $a=c=2b$, $\epsilon^t_1=1$, $\epsilon^t_2=-1$, and all other ϵ^t_{ij} are zero

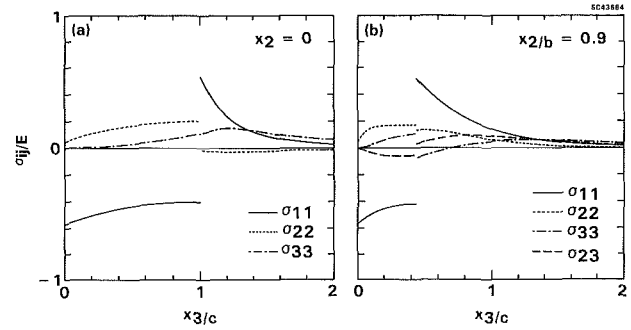


Fig. 9 The net stress σ_{ij} as a function of x_3 for $x_1=0$ and the marked values of x_2 when $a=b=c$, $\epsilon^t_1=1$, $\epsilon^t_2=\epsilon^t_3=-1/2$, and all other ϵ^t_{ij} are zero

inclusion. In Fig. 7, stress components resolved in the coordinates of the interface (inset) are plotted around the interface in the plane $x_1=0$. The angle θ in Fig. 7 is defined by $x_2=b\cos\theta$ and $x_3=b\sin\theta$, where (x_2,x_3) lies on the interface. The stresses shown are those within the inclusion at the interface. Note especially that the shear stress $\sigma_{r\theta}$ is nonzero on the free surface at the interface, but zero elsewhere on the free surface, as required by the boundary conditions. Since $\sigma_{r\theta}$ would drive interfacial failure when the normal stress σ_r across the interface is compressive, failure in such cases would initiate at the free surface.

Shear Transformations. Plastic deformation in crystals is usually volume conserving, which is to say $\epsilon^t_{mm}=0$ (with the summation convention for repeated indices). From many numerical calculations, one induces the important result for such cases that the stress fields are approximately uniform throughout the inclusion for all a, b, and c. For $\epsilon^t_1=1$, $\epsilon^t_2=-1$, and all other components of ϵ^t equal to zero, typical

stress variations are shown in Fig. 8 for the case $b/a = 2$ and $c/a = 1$. Right out to the edges of the inclusion, all stress components are approximately constant. As c/a decreases, the constancy is even more perfect.

For the transformation strain $\epsilon_1^t = 1$, $\epsilon_2^t = -1/2$, $\epsilon_3^t = -1/2$, all other components zero, typical stress variations are shown in Fig. 9. The variations are larger than in Fig. 8, because the tractions t_3 are much larger for a transformation strain for which $\epsilon_3^t \neq 0$. But it is still a useful approximation to consider the stress field uniform in cases where it is already an approximation to consider the inclusion ellipsoidal.

4 Conclusions

Various analytical and numerical calculations of the surface displacements and stress field generated by a semi-ellipsoidal surface inclusion have been presented. For transformation strains for which ϵ_{31}^t and ϵ_{23}^t are zero, general characteristics (for any a, b, and c) include the following:

(1) All components of the stress tensor are zero or constant on the free surface within the inclusion.

(2) For transformations that are also volume conserving ($\epsilon_{mm}^t = 0$), and for shallow inclusions, the stress fields are approximately uniform throughout the volume of the inclusion.

(3) Surface strains or differential displacements are only weakly dependent on the depth of the inclusion over gauge lengths comparable to the width of the inclusion when the depth exceeds the lesser of $2a$ and $2b$.

(4) The shear strain along the interface, when resolved in the plane of the interface, has it maximum at the free surface.

Analytical results and accurate numerical calculations have been presented to reveal the nature of the anomalous stress fields near the intersection of the interface and the free surface. The numerical methods presented are especially efficient for calculating experimentally observable surface displacements.

Acknowledgments

The author is indebted to D. B. Marshall for useful comments, and to D. Kouris and T. Mura for providing the unpublished results reproduced in Fig. 3. The work was supported by Rockwell International Independent Research and Development.

References

Bogy, D. B., 1975, "The Plane Solution for Joined Dissimilar Elastic Semistrips Under Tension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 42, pp. 93–98.

Bradley, W., 1988, private communication.

Byrd, P. F., and Friedman, M. D., 1971, *Handbook of Elliptic Integrals for Engineers and Scientists*, Springer-Verlag, New York.

Chiang, S. S., Marshall, D. B., and Evans, A. G., 1982, "The Response of Solids to Elastic/Plastic Indentation. I. Stresses and Residual Stresses," *Journal of Applied Physics*, Vol. 53, pp. 298–311.

Cox, B. N., Morris, W. L., and James, M. R., 1986, "High Sensitivity, High Spatial Resolution Strain Measurements in Alloys and Composites," *Proc. Conf. on Nondestructive Testing and Evaluation of Advanced Materials and Composites*, Colorado Springs, Colo., pp. 25–39.

Cox, B. N., Marshall, D. B., Kouris, D., and Mura, T., 1988, "Surface Displacement Analysis of the Transformed Zone in Magnesia Partially Stabilized Zirconia," ASME *Journal of Engineering Materials and Technology*, Vol. 110, pp. 105–109.

Cox, B. N., Morris, W. L., and James, M. R., 1987, "Two-Stage Microplastic Surface Deformation in Al 2219-T851," *Acta Metallurgica*, Vol. 35, pp. 1289–1299.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," *Proceedings of the Royal Society London*, Vol. 241A, pp. 376–396.

Eshelby, J. D., 1959, "The Elastic Field Outside an Ellipsoidal Inclusion," *Proceedings of the Royal Society London*, Vol. 252A, pp. 561–569.

Flamant, A., 1892, "Sur la Répartition des Pressions dans un Solide Rectangulaire Chargé Transversalement," *Comptes Rendus*, Vol. 114, pp. 1465–1468.

Gao, Z., and Mura, T., 1988, "A Nondestructive Measurement of Nonelastic Deformation by the Regularization Method," ASME JOURNAL OF APPLIED MECHANCIS, in press.

Hill, R., 1950, *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford, p. 38.

James, M. R., Morris, W. L., and Cox, B. N,. 1988, "A High Accuracy Automated Strain Field Mapper," submitted to *Experimental Mechanics*.

Kouris, D., Tsuchida, E., and Mura, T., 1988a, "The Hemispheroidal Inhomogeneity at the Free Surface of an Elastic Half-Space," ASME JOURNAL OF APPLIED MECHANICS, in press.

Kouris, D., and Mura, T., 1988b, "The Elastic Field of a Hemispherical Inhomogeneity at the Free Surface of an Elastic Half Space," *Journal of Mechanics and Physics of Solids*, in press.

Kouris, D., and Mura, T., 1987, private communication.

Lange, F. F., 1988, private communication.

Lange, F. F., and Metcalf, M., 1983, "Processing-Related Fracture Origins: II, Agglomerate Motion and Cracklike Internal Surface Caused by Differential Sintering," *Journal of American Ceramic Society*, Vol. 66, pp. 398–406.

Mindlin, R. D., 1936, "Force at a Point in the Interior of a Semi-Infinite Solid," *Physics*, Vol. 7, pp. 195–202.

Morris, W. L., Cox, B. N., and James, M. R., 1987, "Microplastic Surface Deformation of Al 2219-T851," *Acta Metallurgica*, Vol. 35, pp. 1055–1065.

Morris, W. L., Inman, R. V., and Cox, B. N., 1988, "Microscopic Deformation in a Heated Unidirectional Graphite/Epoxy Composite," *Journal of Materials Science*, in press.

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

Mura, T., Cox, B., and Gao, Z., 1986, "Computer-Aided Nondestructive Measurements of Plastic Strains from Surface Displacements," *Proceedings of International Conference on Computational Mechanics*, S. N. Atluri, ed., Springer-Verlag, Tokyo, pp. II 43–48.

Robinson, K., 1951, "Elastic Energy of an Ellipsoidal Inclusion in an Infinite Solid," *Journal of Applied Physics*, Vol. 22, pp. 1045–1054.

Williams, D. R., Davidson, D. L., and Lankford, J., 1980, "Fatigue Crack Tip Strains by Stereoimaging," *Experimental Mechanics*, Vol. 20, pp. 134–149.

An Exact Transient Study of Dislocation Emission and its Effects on Dynamic Fracture Initiation

L. M. Brock
Mem. ASME

J.-S. Wu

Department of Engineering Mechanics,
University of Kentucky,
Lexington, KY 40506

Closed-form transient solutions for the micromechanical process of screw dislocation emission from a stationary crack which is subjected to SH-wave diffraction are presented. The dislocations are allowed to leave the crack edge in arbitrary directions, either singly or in pairs. Imposition of an emission criterion that is both based on the dislocation force concept and is similar to criteria applied to quasi-static emission studies allows expressions for the instants of dislocation emission and arrest and the distance traveled by the dislocation to be obtained. These expressions are studied for their dependence on parameters such as emission direction and speed, and several distinctive dynamic effects are observed. A standard fracture criterion is then imposed, and conditions for determining whether fracture will precede or follow emission are established in terms of real time. Finally, some estimates for the orders of magnitudes of the parameters involved in this micromechanical process are given.

Introduction

Rice and Thomson (1974), Li (1981), Ohr (1985), and Lin and Thomson (1985) have discussed fracture and its brittleness or ductility in terms of the emission of dislocations from the crack edge. The discussions are generally nontransient, perhaps reflecting the view (Hirth and Lothe, 1982) that dislocation motion is sluggish enough to minimize or preclude dynamic effects. However, Achenbach and Brock (1973) have shown that important details of fracture under dynamic loading may not be discernable in nontransient analyses. More to the point, Brock (1989) has shown that fracture initiation can be sensitive to the timing of dislocation emission events.

Brock's work, however, treated only a single dislocation emitted in the same plane as the crack. The present work, therefore, extends the transient study of dislocation emission and its role in fracture under dynamic loading by considering both dislocations which leave the crack edge in various directions, and dislocations which are emitted in pairs. This latter consideration will allow insight into multiple emission processes.

As in Brock's (1989) work, screw dislocations and semi-infinite stationary Mode III cracks in unbounded, linearly elastic, isotropic, homogeneous solids are treated. There is little inherent additional difficulty in treating edge dislocation emission from Mode I and Mode II cracks. However, the associated mathematical solutions for the screw dislocation/Mode III case can be obtained in closed form, which is a distinct advantage if general physical insight is a major goal.

In the next section, the solution for the basic problem of a screw dislocation leaving a crack which is subjected to dynamic loading is discussed.

Basic Problem and its Formal Solution

Consider the semi-infinite crack $y=0$, $x<0$ in Fig. 1(*a*). The unbounded material containing the crack is at rest except for a step-stress pulse traveling as a horizontally polarized shear (SH) wave at right angles to the crack plane. The wave

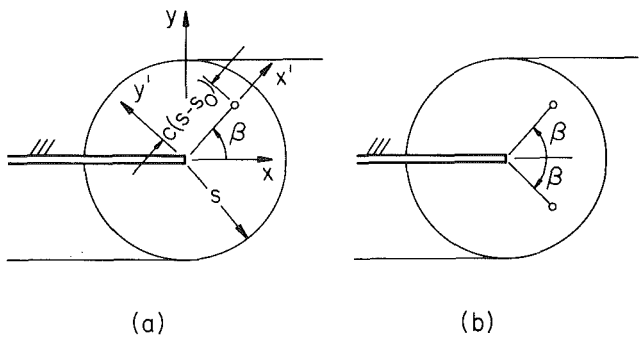


Fig. 1(*a*) Wave pattern generated by single dislocation emission and 1(*b*) wave pattern generated by emission of dislocation pair

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10-15, 1989.

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 on the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 2, 1988; final revision, December 1, 1988.

Paper No. 89-WA/APM-5.

reaches the crack plane at $s=0$, where $s=$(shear wave speed)$\times$(time). At $s=s_o>0$, a single right-handed screw dislocation is emitted from the crack edge, and travels rectilinearly away at an angle $\beta(0\le\beta<\pi)$ with respect to the crack plane, and with a constant, subcritical speed. The wave motion engendered by this process is also indicated in Fig. 1(*a*).

The process itself can be treated as one of antiplane strain, so that only the out-of-plane displacement $w(x,y,s)$ exists. By linear superposition, this can be written as

$$w=w_o+W_o+w_b+W_b \tag{1}$$

where $(w_o,\ w_b)$ are the displacements which would be generated by the SH-wave and emitted dislocation if no crack existed, and $(W_o,\ W_b)$ are the displacements generated in order to cancel the $(w_o,\ w_b)$-induced stresses from the crack surfaces.

The problem involving $(w_o,\ W_o)$ has been treated by Achenbach (1970), and expressions for w_o and the stress generated by W_o directly ahead of the crack derived. Brock (1989) has extended this work and obtained complete expressions for W_o itself everywhere. Thus, we have

$$\mu w_o=\tau(y-s) \tag{2}$$

for $s>y$,

$$\mu W_o=\frac{\sqrt{2}}{\pi}\tau\sqrt{(r-x)}\sqrt{(s-r)}$$

$$+\frac{\tau}{2}s\left(1+\frac{2}{\pi}\tan^{-1}\frac{r-s-x}{\sqrt{2}\sqrt{(r-x)}\sqrt{(s-r)}}\right)$$

$$-\frac{\tau}{2}|y|\left(1+\frac{2}{\pi}\tan^{-1}\frac{s-x-r}{\sqrt{2}\sqrt{(r+x)}\sqrt{(s-r)}}\right) \tag{3}$$

for $s\ge r$, and

$$\pm\mu W_o=\tau(s-|y|)H(s-|y|) \tag{4}$$

for $\pm y>0, r>s$, where

$$r=\sqrt{(x^2+y^2)} \tag{5}$$

and μ is the shear modulus, while $\tau>0$ is the magnitude of the step-stress.

The expression for w_b can be obtained either by generalizing results due to Nabarro (1951) or by specializing a general three-dimensional dislocation loop solution by Brock (1986):

$$\frac{2\pi}{b}w_b=\tan^{-1}\frac{cy'\sqrt{(T^2-r^2)}}{x'(x'-cT)+(y')^2},\quad T=s-s_o\ge r. \tag{6a,b}$$

In Fig. 1(*a*) it can be seen that

$$x'=x\cos\beta+y\sin\beta,\quad y'=y\cos\beta-x\sin\beta \tag{7a,b}$$

are coordinates aligned with the dislocation path. In (6*a*,*b*), c is the dislocation speed nondimensionalized with respect to the shear wave speed, i.e, $0<c<1$, while b is the Burgers vector magnitude.

Finally, by following Achenbach (1970), it is easily shown that for $\pm y<0$ the formal solution

$$\pm W_b=\frac{1}{\pi}\iint\frac{\partial W_b}{\partial y}(p,0,q)\frac{dpdq}{\sqrt{[(s-q)^2-(x-p)^2-y^2]}} \tag{8}$$

exists for W_b, where the variables $(p,0,q)$ correspond to (x,y,s) and the region of integration is delineated by positive values of the radical argument, and by the location of nonzero values of the derivative along $y=0$. These nonzero values are

$$\frac{\partial W_b}{\partial y}=-\frac{\partial w_b}{\partial y} \tag{9}$$

for $y=0,\ -s<x<0$ and

$$\frac{\partial W_b}{\partial y}=\frac{1}{\pi\sqrt{(\eta-\xi)}}\int_{\eta_o}^{\xi}\frac{\partial w_b}{\partial y}\left(\frac{v-\xi}{\sqrt{2}},\ 0,\ \frac{v+\xi}{\sqrt{2}}\right)\frac{\sqrt{(\xi-v)}}{\eta-v}dv,$$

$$\eta_o=\frac{1}{\sqrt{2}}s_o \tag{10a,b}$$

for $y=0,\ 0<x<s$, where $\sqrt{2}\eta=s+x$ and $\sqrt{2}\xi=s-x$ are characteristic variables which arise naturally in the solution process (Achenbach, 1970). The integrations associated with (8) and (10) are nontrivial but, as indicated by the work of Brock (1989) can, in fact, be carried out. This is done in the next section.

Integrations for W_b

From (6*a*) it is easily shown that

$$\frac{2\pi}{cb}\frac{\partial w_b}{\partial y}=\frac{\sqrt{(T^2-x^2)}}{x}\frac{x\cos\beta-cT}{(x\cos\beta-cT)^2+(1-c^2)x^2\sin^2\beta} \tag{11}$$

for all $y=0,\ T>|x|$. Substitution of this result into (10) yields, upon elimination of $(\xi,\ \eta)$ in terms of $(x,\ s)$ and introduction of the integration variable change $\sqrt{2}v=\sqrt{2}\xi(1-u)+s_o u$,

$$\frac{\pi^2}{cb}\frac{\partial W_b}{\partial y}=\frac{\sqrt{2}}{\sqrt{x}\sqrt{(T-x)}}\int_0^1\sqrt{\left(\frac{1-u}{u}\right)}$$

$$\frac{u(\cos\beta-c)+2c}{[u(\cos\beta-c)+2c]^2+(1-c^2)u^2\sin^2\beta}\frac{du}{u+\dfrac{2x}{T-x}}. \tag{12}$$

The integration in (12) is over the branch cut $0<u<1$ of the integrand, which itself exhibits simple poles at

$$u=\frac{-2x}{T-x},\quad u=\frac{2c}{z-c},\quad z=\cos\beta\pm i\sqrt{(1-c^2)}\sin\beta \tag{13a-c}$$

and vanishes as $0(u^{-2})$ when $|u|\to\infty$. These poles lie off the branch cut, so that Cauchy residue theory can readily be applied to give

$$\frac{\partial W_b}{\partial y}=-\frac{\partial w_b}{\partial y}+\frac{b}{2\pi}\sqrt{[c(1+c)]}\sqrt{\left(\frac{T-x}{x}\right)}$$

$$\frac{x(1-c-c\cos\beta)-cT}{(x\cos\beta-cT)^2+(1-c^2)x^2\sin^2\beta}\cos\frac{\beta}{2} \tag{14}$$

for $y=0, 0<x<s$. Substitution of (9), (11), and (14) into (8), along with the introduction of $(\xi,\ \eta)$ and their integration variable counterparts $\sqrt{2}u=p+q$ and $\sqrt{2}v=p-q$, gives the more explicit result

$$\pm\frac{\pi^2}{b}W_b=-c\int_{\eta_o}^{v_o}dv\sqrt{\left(\frac{v-\eta_o}{\xi-v}\right)}\int_{\eta_o}^{M}$$

$$du\sqrt{\left(\frac{u-\eta_o}{M-u}\right)}\frac{(u-v)\cos\beta-c(u+v-2\eta_o)}{D(u-v)}$$

$$+\sqrt{[2c(1+c)]}\cos\frac{1}{2}\beta\int_{\eta_o}^{v^*}dv\sqrt{\left(\frac{v-\eta_o}{\xi-v}\right)} \tag{15}$$

$$\int_v^M du\frac{(u-v)(1-c-c\cos\beta)-c(u+v-2\eta_o)}{D\sqrt{(u-v)}\sqrt{(M-u)}}$$

where

$$D=[(u-v)\cos\beta-c(u+v-2\eta_o)]^2+(1-c^2)(u-v)^2\sin^2\beta \tag{16}$$

$$\sqrt{2}v_o = \sqrt{2}\xi - \frac{y^2}{\sqrt{2}\eta - s_o}, \quad \sqrt{2}M = \sqrt{2}\eta - \frac{y^2}{\sqrt{2}(\xi - v)}, \quad \sqrt{2}v^* = s - r. \tag{17a-c}$$

The first term in (15) involves a u-integration over the branch cut $\eta_o < u < M$ of an integrand which exhibits simple poles at

$$u = v, \quad u = v + 2c\,\frac{v-\eta_o}{z-c} \tag{18a,b}$$

and behaves as $0(u^{-2})$ when $|u| \to \infty$. The complex conjugate poles (18*b*) lie off of the branch cut for all $\eta_o < v < v_o$, but the pole (18*a*) lies on the cut unless $v^* < v < v_o$. For the second term in (15), the u-integration is over the branch cut $v < u < M$ of an integrand which also behaves as $0(u^{-2})$ when $|u| \to \infty$, but which exhibits poles (18*b*) off the branch cut for all $\eta_o < v < v^*$. Cauchy residue theory can then be applied, and (15) becomes

$$\pm \frac{2\pi}{b} W_b = \mathrm{Re}\int_{v^*}^{v_o} \frac{\sqrt{2}\sqrt{(z+c)}\,dv}{\sqrt{[(z-c)(r^2-s^2) + 2\sqrt{2}v(zs + c(s_o - x)) + 2cs_o(x-s) - 2w^2(z+c)]}} + \int_{v^*}^{v_o} \frac{\sqrt{2}dv}{\sqrt{[r^2 - (s - v\sqrt{2}v)^2]}} \tag{19}$$

for $\pm y < 0$, where (ξ, η) have once again been eliminated in terms of (x, s), so that now

$$\sqrt{2}v_o = \frac{s^2 - r^2 - s_o(s-x)}{T+x}. \tag{20}$$

The v-integrations in (19) can be performed by use of standard integral tables, with the results that, for all $s \geq r$,

$$\frac{2\pi}{b} W_b = \frac{\pi}{2} + \mathrm{Re}\left(\sin^{-1}\frac{zx - cT + \left(\frac{z+c}{T+x}\right)y^2}{\sqrt{[(zx-cT)^2 + (z^2-c^2)y^2]}}\right) - \sin^{-1}\left[\frac{Tx + r^2}{r(T+x)}\right] - \mathrm{Re}\left(\sin^{-1}\frac{zx - cT + (z+c)(r-x)}{\sqrt{[(zx-cT)^2 + (z^2-c^2)y^2]}}\right). \tag{21}$$

With (21) in hand, the complete solution (1) for the basic problem is now ideal for purposes of both computation and analysis. In the next section, this solution is used to examine a criterion for dislocation emission.

Criterion for Dislocation Emission

Following the precedent set in quasi-static emission analyses, e.g., Ohr (1985), we adopt the postulate (Bilby, Cottrell, Swinden, 1963; Shaw, 1984) that dislocation motion by glide cannot occur unless the glide plane force can overcome the lattice friction, which can be identified with the yield stress.

To obtain the glide plane force, we adopt the Peach-Koehler scheme (Hirth and Lothe, 1982), whereby the force can be calculated in terms of stress components evaluated at the dislocation itself. More specifically, the glide plane force here involves the glide plane stress

$$\tau_{y'z'} = \mu\frac{\partial w}{\partial y'} = \mu\left(\cos\beta\,\frac{\partial w}{\partial y} - \sin\beta\,\frac{\partial w}{\partial x}\right) \tag{22}$$

which is evaluated by first choosing $x' = cT$, and then setting $y' = 0$ (cf., Dundurs, 1968). Performing this operation in view of (1) and (6) yields the following expression for the glide plane force per unit length of dislocation edge:

$$\frac{1}{b}F_{x'} = 2\frac{\tau}{\pi}\sqrt{\left(\frac{s-cT}{cT}\right)}\cos\frac{\beta}{2} + \frac{2}{\pi}\tau\tan^{-1}\left[\frac{\sqrt{(s^2 - c^2T^2\sin^2\beta)} - s + cT\cos\frac{\beta}{2}}{2\sqrt{(cT)}\sqrt{(s-cT)}\cos\frac{\beta}{2}}\right] - \frac{\mu b}{4\pi}\frac{B(\beta, c)}{cT} \tag{23a}$$

$$B(\beta, c) = \frac{1}{\sqrt{(1-c^2)}}\left[1 + \frac{c(1+c)}{1+c\cos\beta}\cos^2\frac{\beta}{2}\right]. \tag{23b}$$

For $\beta = 0$, (23) reduces to the expression obtained by Brock (1989) for emission in the crack plane. When c is set to zero while cT is allowed to remain finite, the last term in (23*a*) reduces to the quasi-static result for dislocation emission in the crack plane (Majumdar and Burns, 1981). A twofold dynamic effect is thereby made apparent: First, the dislocation speed, i.e., c, appears explicitly in the term that represents the dislocation contributions to its own glide plane force. Then it is the speed which couples the emission angle β into the term. The lack of angle dependence in the quasi-static result is well known (e.g., Brock and Wu (1988a)).

Brock (1988) argued that $cT/s_o \ll 1$ during the emission process. That is, the distance cT traveled by an emitted dislocation is small compared to the distance traveled by a shear wave during the interval prior to emission. Adopting the same argument here in view of (6*b*) yields the following approximation for the right-hand side of (23*a*):

$$\frac{2}{\pi}\tau\sqrt{\left(\frac{s_o}{cT}\right)}\cos\frac{\beta}{2} - \frac{\mu b}{4\pi}\frac{B(\beta, c)}{cT}. \tag{24}$$

Satisfaction of the emission criterion requires that (24) exceed the value σ, where σ is the yield stress. Equating (24) with σ, therefore, gives a quadratic equation for the values of $\sqrt{(cT)}$ at which the force drops below the critical level. Because cT is the distance of the dislocation from the crack edge, this is analogous to quasi-static results for screw dislocation emission in the crack plane. Here, of course, as already noted, the equation includes an emission angle, as well as two essentially dynamic parameters (c, s_o).

The quasi-static results, e.g., Ohr (1985), interpret the smaller distance obtained from the quadratic as the point where the dislocation force overcomes resistance to glide, while the larger distance locates where the emitted dislocation arrests. Furthermore, the smaller distance must be a fraction of the dislocation core radius (Hirth and Lothe, 1982), while the larger value must exceed the core radius. Following Brock (1989), we also adopt this interpretation and, as a result, are able to find not only the distance d^* from the crack edge at which arrest occurs, but also the instants (s_o, s^*) at which, respectively, emission and arrest occur:

$$\lambda\frac{d^*}{h} = (\omega B)^2, \quad s_o = \lambda\left(\frac{\pi\sigma}{2\tau}\sec\frac{\beta}{2}\right)^2\left[1 + \sqrt{\left(\frac{d^*}{\lambda h}\right)}\right]^2 h, \quad s^* = \frac{d^*}{c} + s_o. \tag{25a-c}$$

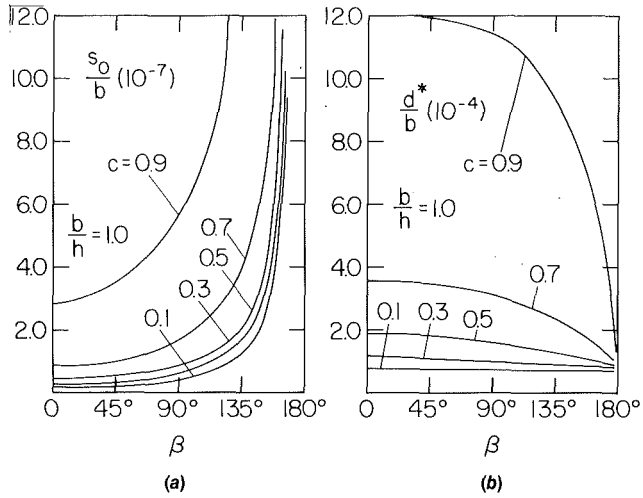


Fig. 2(*a*) Emission instant versus angle at various speeds: single dislocation; 2(*b*) arrest distance versus emission angle at various speeds: single dislocation

Here, h is the core radius, while

$$\left(\frac{\tau}{\pi\sigma}\cos\frac{\beta}{2}\right)^2 \geq \sqrt{\left(\frac{d^*}{\lambda h}\right)},\ \lambda\frac{d^*}{h}\geq 1,\ \omega=\frac{1}{4\pi}\frac{\mu b}{\sigma h} \tag{26a-c}$$

and the dimensionless parameter λ is the aforementioned fraction of the core radius, i.e., $0<\lambda\leq 1$. Substitution of (25a,b) reduces (26a) to

$$(\lambda-\omega B)^2\geq 0 \tag{27}$$

which is always satisfied. It should be noted that for $\lambda=1$ the strict equality in (27) implies that $d^*=h$, i.e., the dislocation in essence does not leave the crack edge. For many engineering materials, $\mu/\sigma\sim 0(10^3)$ and $b/h\sim 0(1)$, while $\tau/\sigma\sim 0(10)^{-1}$ for a high but noncritical stress level. Equations (25a,b) then show that $d^*/h>1$, while simultaneously,

$$\frac{d^*}{s_o}=\left(\frac{2\tau}{\pi\sigma}\cos\frac{\beta}{2}\right)^2\left[1+\sqrt{\left(\frac{\lambda h}{d^*}\right)}\right]^{-2}\ll 1,\ \frac{h}{s_o}\ll 1. \tag{28a,b}$$

The first inequality confirms the assumption made in employing the approximation (24), while (28b) implies that the dislocation emission and arrest times are not necessarily insignificant.

Several other observations can also be made: First, (25) shows that if λ is treated as a specified constant, the arrest distance is independent of the SH-wave. However, the instants of emission and arrest vary inversely with the square of τ, which means that, while the arrest distance would remain finite in the limit as the SH-wave disappears, the emission and arrest process would never actually occur. Then, the form of B shows, as previously implied, a dynamic overshoot effect through its explicit dependence on c. This effect is especially pronounced at high ($c\to 1$) values. It should be noted that this independence of arrest distance and applied loading can also occur in quasi-static analyses of screw dislocation emission in the crack plane, as pointed out by Brock (1989).

Numerical Results

To illustrate the behavior predicted by the emission criterion, we plot in Figs. 2(a,b) the dimensionless ratios s_o/b and d^*/b, respectively, versus β for various values of c. For these plots we choose the dimensionless parameters

$$\frac{b}{h}=1.0,\ \frac{\tau}{\mu}=0.0001,\ \frac{\sigma}{\mu}=0.001,\ \lambda=1.0. \tag{29a-d}$$

That is, the core radius and Burgers vector magnitudes are equal, the SH-wave step-stress is high but noncritical, and the smaller value of cT corresponds to the core radius itself.

Figure 2(a) shows that the time interval between diffraction and emission increases rapidly both with emission angle and dislocation speed. In particular, this time is seen to approach infinity for dislocations which attempt to leave the crack edge by paths very near the crack surface. Figure 2(b) shows that the distance from the crack edge at which an emitted dislocation arrests also increases with dislocation speed, but varies inversely with the emission angle. Thus, emission on the plane directly ahead of the crack will occur earlier and the emitted dislocation will travel farther than on other planes. The variation of d^* with c is clearly a dynamic overshoot phenomenon, while the variation of s_o with c is noteworthy for another reason: It suggests that screw dislocation emission will occur first for any β when $c\to 0$. That is, the emission process prefers in this respect to take place, in effect, quasi-statically.

The magnitudes of the ratios in Figs. 2(a, b) are also noteworthy: Because typical Burgers vector magnitudes are $0(10^{-10})$ m (Hirth and Lothe, 1982), Fig. 2(b) shows that arrest distances are $0(10^{-6})$m. Thus, this result of the essentially micromechanical emission process could possibly be measured, albeit with difficulty. In regard to Fig. 2(a), it is known (Achenbach 1973) that shear wave speeds for metals can be $0(10^3)$ m/sec. Therefore, the times to emission shown in Fig. 2(a) are $0(10^{-6})$sec, which is, again, barely in a measurable range. Returning to Fig. 2(b) in view of (25c) and (28a), however, the results are not as encouraging: The d^*/b magnitudes suggest that the difference between the instants of emission and arrest can definitely be less then $0(10^{-6})$ sec. Therefore, it may not be possible to distinguish between the two instants.

Emission of a Dislocation Pair

To gain insight into multiple dislocation emission, we now consider the situation shown in Fig. 1(b): When $s=s_o$, two right-hand screw dislocations leave the crack edge at angles $\pm\beta$, $\beta>0$ with respect to the crack plane. The solution (1) is again valid, as are (2)–(4). Simple symmetry arguments show that the expression (21) for W_b need be modified only by affixing the factor 2, while (6a) for w_b becomes

$$\frac{2\pi}{b}w_b=\tan^{-1}\frac{cy'\sqrt{(T^2-r^2)}}{x'(x'-cT)+(y')^2}+\tan^{-1}\frac{cy''\sqrt{(T^2-r^2)}}{x''(x''-cT)+(y'')^2},\ T\geq r \tag{30a,b}$$

$$x''=x\cos\beta-y\sin\beta,\ y''=y\cos\beta+x\sin\beta. \tag{31a,b}$$

The procedure just described for obtaining the glide plane force on the dislocation can again be followed and, for the dislocation which is emitted at angle β, it is easily shown that (23)–(28) are again valid, with the exception that the definition (23b) must be replaced by

$$B(\beta,c)=\frac{\sqrt{(1+c)}}{1+c\cos\beta}\left[\frac{\sqrt{(1-c)}}{1-c\cos\beta}+\frac{2c}{\sqrt{(1-c)}}\cos^2\frac{\beta}{2}\right]. \tag{32}$$

The general observations made earlier concerning (25)–(28) are also still valid, so we now present in Figs. 3(a, b) plots analogous to those in Figs. 2(a, b): Again, the values (29a–d) are chosen. Figure 3(b) is similar to Fig. 2(b), in terms of (β, c)-dependence and order of magnitude, although the arrest distance variation with β is more pronounced. In Fig. 3(a), the emission time dependence on c noted in Fig.

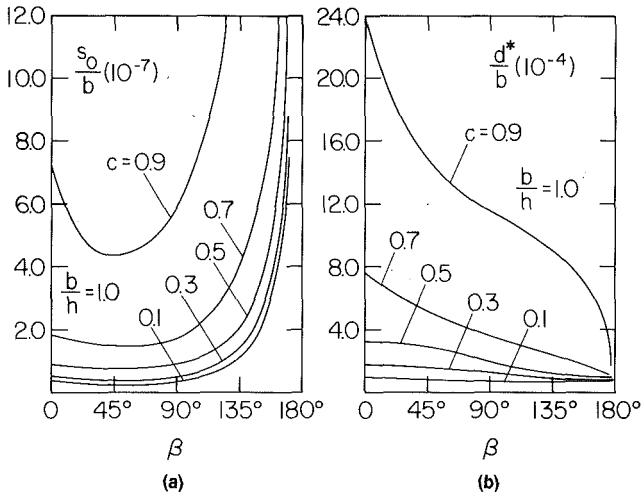


Fig. 3(a) Emission instant versus angle at various speeds: dislocation pair; 3(b) arrest distance versus emission angle at various speeds: dislocation pair

$2(a)$ is preserved, as is the order of magnitude. The β-dependence, however, is somewhat different: Specifically, the emission time exhibits a nonzero β minimum, so that there exists for every c a specific value $\beta>0$ for which emission of dislocation pairs will occur first. Differentiation of $(25b)$ with respect to β leads to the equation

$$\cos\frac{1}{2}\beta-\frac{\lambda+\omega B(\beta,\,c)}{\lambda+\omega B(0,\,c)}=0 \tag{33}$$

for this value, where B is given by (32) and $(\lambda,b/h)=1$ for the cases illustrated in Fig. $3(a)$. Equation (33) gives real values for β only if $B(0,c)\geq B(\beta,c)$, but it can be shown that this is indeed the case for either $(23b)$ or (32) for all

$$0<c<1,\qquad 0<\beta<\pi.$$

A direct comparison of Figs. 2 and 3 also shows that a pair of dislocations might well be emitted before a single dislocation which moves at an angle to the crack plane. Moreover, it is possible for the dislocation pair to travel farther than such a single dislocation would prior to arrest. It should also be noted that the conclusion drawn from Fig. 2 is possible in Fig. 3, too: The emission process prefers, in the sense of minimum s_o, to occur quasi-statically.

The Role of Emission in Fracture

We now apply our emission studies to the question of fracture initiation at the crack edge and its characterization: For purposes of illustration, the stress intensity factor criterion is adopted. For Mode III fracture initiation, the relevant dynamic stress intensity factor K_3 can be obtained from the definition

$$K_3=\lim_{x\to 0^+}\sqrt{(2\pi x)}\tau_{yz}(x,\,0,\,s),\quad \tau_{yz}=\mu\frac{\partial w}{\partial y}. \tag{34a,b}$$

If fracture begins prior to emission, then from the previous analysis it can be shown that the fracture criterion has the form

$$2\tau\sqrt{\left(\frac{2}{\pi}s\right)}=K_3^c,\quad s<s_o \tag{35}$$

where K_3^c is the critical value of K_3. Solution of (35) for the instant of fracture initiation s_c gives

$$s_c=\frac{\pi}{2}\left(\frac{\mu}{2\tau}\right)^2 k_c^2 h<s_o,\quad k_c=\frac{1}{\mu\sqrt{h}}K_3^c \tag{36a,b}$$

where k_c is a dimensionless constant. If the constraint in $(36a)$ is not satisfied, then dislocation emission occurs before fracture initiation, and the results of the previous analysis can be used to show that the fracture criterion becomes

$$2\tau\sqrt{\left(\frac{2}{\pi}T\right)}-\frac{\mu b}{\sqrt{(2\pi)}}\sqrt{\left(1+\frac{1}{c}\right)}\frac{1}{\sqrt{T}}\cos\frac{\beta}{2}=K_3^c \tag{37}$$

for $s<s^*+d^*$, where $T/s_o\gg 1$. The inequality constraint simply acknowledges that the previous analysis is not valid once the signal of dislocation arrest reaches an observation point. The second follows from the fact that, as can be gleaned from the sign difference between the two terms on the left-hand side of (37), the emitted dislocation relaxes the crack edge stress field. Such dislocation shielding of the crack edge is a common (Majumdar and Burns, 1981; Thomson and Sinclair, 1982) effect, and the consequence here is that it will be some time after $s=s_o$ before fracture can begin. Equation (37) is quadratic in $\sqrt{T}$, and it is easily shown that two real roots always exist, but that one is negative. The positive real root gives, therefore, the fracture initiation instant, s_c, as

$$s_o=s_o+\frac{\pi}{2}\left(\frac{\mu}{2\tau}\right)^2\left[\frac{1}{2}k_c+\sqrt{\left[\left(\frac{1}{2}k_c\right)^2+\frac{2\tau b}{\pi\mu h}\sqrt{\left(1+\frac{1}{c}\right)}\right]}\right]^2 h<s^*+d^*. \tag{38}$$

For the general case $\mu/\sigma\sim 0(10^3)$, $b/h\sim 0(1)$, $\tau/\sigma\sim 0(10^{-3})$ it can be shown that, indeed, $T/s_o\gg 1$.

To further examine the constraint in $(36a)$, we substitute $(25b)$ and find that the relation

$$B-4\pi\lambda\frac{h}{b}\left[\frac{1}{\sqrt{(2\pi)}}k_c\cos\frac{\beta}{2}-\frac{\sigma}{\mu}\right]\geq 0 \tag{39}$$

must hold. Similarly, we find that, for the constraint in (38) to hold, the relation

$$\frac{\tau b}{\sigma h}\sqrt{\left(1+\frac{1}{c}\right)}\frac{B}{2\pi\sqrt{\lambda}}-\frac{1}{2}\sqrt{\left(\frac{\pi}{2}\right)}k_c-\sqrt{\left[\frac{\pi}{2}\left(\frac{1}{2}k_c\right)^2+\frac{\tau b}{\mu h}\sqrt{\left(1+\frac{1}{c}\right)}\cos\frac{\beta}{2}\right]}\geq 0 \tag{40}$$

must be satisfied.

The constraints (39) and (40) give, in effect, conditions for the relative brittleness or ductility of the fracture initiation process: Only those dislocations with $(\beta,\ c)$-values which violate (39) can be emitted prior to fracture. If such dislocations also satisfy (40), then fracture will initiate prior to the instant at which the crack edge is aware that dislocations have arrested. These conditions apply, of course, for either single or paired dislocation emissions, depending on the choice of the function B. In fact, it is easily shown that the two functions are equal only if $\beta=\pi$, and that B defined by (32) exceeds that defined by $(23b)$ for $0\leq\beta<\pi$. This implies appropriately that brittle fracture is more likely to initiate prior to emission of dislocation pairs than prior to single dislocation emission.

More generally, the equations in $(36a)$ and (38) and the constraints (39) and (40) demonstrate again (Brock, 1989) that fracture initiation can be characterized in terms of the timing of various events. Finally, it should be noted that the orders of magnitude obtained for $(s_o,\ s^*)$, i.e., the instants of dislocation emission and arrest, imply that, unless fracture initiates in a time interval of $0(10^{-6})$ sec after diffraction occurs, not only will it then initiate after emission, but after the crack edge is aware of dislocation arrest.

Brief Discussion

This work extended earlier results in the transient analysis of dislocation emission from dynamically loaded cracks. The particular problem of single and paired emissions of screw

dislocations at constant speeds and in arbitrary directions from a crack subjected to SH-wave diffraction was treated, and closed-form solutions given. An emissions criterion based on the dislocation force concept was adopted from quasi-static emission analysis. The transient nature of the analysis increased the robustness of the criterion, however. First of all, the force itself was found now to depend explicitly on dislocation speed and, precisely because of the speed dependence, to also be more sensitive to emission angle values. Secondly, the criterion yielded not only formulas for the distance traveled by the emitted dislocation to arrest, but the instants of emission and arrest as well. These formulas generally depended on dimensionless parameters related to the SH-wave stress, the dislocation speed, the dislocation core radius, the Burgers vector magnitude, and the angle of emission.

Numerical calculations indicated that a single dislocation would most likely be emitted in the crack plane, whereas a pair of dislocations would leave the crack at optimum angles. The calculations also demonstrated that single dislocation emission will generally but not always occur before emission of a dislocation pair. Finally, the calculations showed that the arrest distances and emission times are both small, reflecting the micromechanical nature of the emission process, but not necessarily insignificant. On the other hand, it may be difficult to distinguish between the times of emission and arrest. The arrest distances, it should also be noted, would give insight into the extent of the dislocation-free zone (Thomson and Sinclair, 1982) around the crack edge.

This work also considered the role of dislocation emission in fracture. By using the same analysis and a standard critical stress intensity factor fracture criterion, formulas for the instants of fracture initiation were obtained for both emission-free and postemission situations. These instants were associated with constraints which in effect, provided conditions for assessing the relative brittleness or ductility of the initiation process. Furthermore, the aforementioned magnitudes of the instants of dislocation emission and arrest suggested that, unless fracture initiation occurred less than $0(10^{-6})$ sec after diffraction, it would occur under the influence of both moving and arrested dislocations.

In summary, then, the transient nature of the analysis gave additional insight into the emission process, chiefly by allowing the derivation of actual times of events. Interestingly enough, it was the study of one, the emission instant, which indicated that the emission process prefers to occur quasi-statically, as a series of single events, and in the crack plane. Thus, screw dislocation emission studies such as that by Ohr (1985) may be sufficient for gaining insight into nontemporal parameters. This and the other results obtained here suggest that, in fact, the emission instant might be adopted as part of a more complete emission criterion in future work.

Such future work is now planned to relate emission to crack extension, and some preliminary efforts using approximate transient solutions (Brock and Jolles, 1987) and treating preexisting dislocations (Brock and Wu, 1988a) have already been made. Moreover, the analysis and physical insight devised here is currently being extended to studies of edge dislocation emission from Mode I and Mode II cracks: One preliminary result (Brock and Wu, 1988b) suggests that the quasi-static emission process is not necessarily preferred for the edge dislocation case.

Acknowledgment

This work was partially supported by NSF Grant MEA 8319605.

References

Achenbach, J. D., 1970, "Extension of a Crack by a Shear Wave," *Zeitschrift fur angewandte Mathematik und Physik*, Vol. 21, pp. 887–900.

Achenbach, J. D., 1973, *Wave Propagation in Elastic Solids*, North Holland, Amsterdam.

Achenbach, J. D., and Brock, L. M., 1973, "On Quasi-Static and Dynamic Fracture," *Dynamic Crack Propagation*, G. C. Sih, ed., Noordhoff International, Leyden.

Bilby, B. A., Cottrell, A. H., and Swinden, K. H., 1983, "The Spread of Plastic Yield From a Notch," *Proceedings of the Royal Society of London*, Vol. A272, pp. 304–314.

Brock, L. M., 1986, "A Transient Three-Dimensional Analysis of Non-Uniform Dislocation Distribution Growth by Climb and Glide Over Non-Planar Surfaces," *Proceedings of the Royal Society of London*, Vol. A407, pp. 299–311.

Brock, L. M., 1989, "An Exact Transient Analysis of Dislocation Emission and Fracture," *Journal of the Mechanics and Physics of Solids*, Vol. 37, pp. 47–69.

Brock, L. M., and Jolles, M., 1987, "Dislocation-Crack Edge Interaction in Dynamic Brittle Fracture and Crack Propagation," *International Journal of Solids and Structures*, Vol. 23, pp. 607–619.

Brock, L. M., and Wu, J.-S., 1988a, "Transient Generalized Forces Due to Dislocation Array-Growing Crack Interaction," *International Journal of Solids and Structures*, to appear.

Brock, L. M., and Wu, J.-S., 1988b, "Transient Studies of Screw and Edge Dislocation Generation at Crack Edges," University of Kentucky College of Engineering Technical Report, November.

Dundurs, J., 1968, "Elastic Interaction of Dislocations with Inhomogeneities," *Mathematical Theory of Dislocations*, ASME New York, pp. 70–115.

Hirth, J. P., and Lothe, J., 1982, *Theory of Dislocations*, 2nd ed., Wiley-Interscience, New York.

Li, J. C. M., 1981, "Dislocation Sources" *Dislocation Modelling of Physical Systems*, Pergamon Press, New York, pp. 498–518.

Lin, I.-H., and Thomson, R., 1985, "Cleavage, Dislocation Emission and Shielding for Cracks Under General Loading," *Acta Metallurgica*, Vol. 34, pp. 187–200.

Majumdar, B. S., and Burns, S. J., 1981, "An Elastic Theory of Dislocations, Dislocation Arrays and Inclusions Near a Sharp Crack," *Acta Metallurgica*, Vol. 29, pp. 579–588.

Nabarro, F. R. N., 1951, "The Synthesis of Elastic Dislocation Fields," *Philosophical Magazine*, Vol. 42, pp. 1224–1231.

Ohr, S. M., 1985, "An Electron-Microscope Study of Crack Tip Deformation and its Impact on the Dislocation Theory of Fracture," *Materials Science and Engineering*, Vol. 72, pp. 1–35.

Rice, J. R., and Thomson, R., 1974, "Ductile Versus Brittle Behavior of Crystals," *Philosophical Magazine,* Vol. 29, pp. 73–97.

Shaw, M. C., 1984, "A Critical Review of Mechanical Failure Criteria," ASME *Journal of Engineering Materials and Technology*, Vol. 106, pp. 219–226.

Thomson, R. M., and Sinclair, J. E., 1982, "Mechanics of Cracks Screened by Dislocations," *Acta Metallurgica*, Vol. 30, pp. 1325–1334.

Elastic Yield Zone Around an Interfacial Crack Tip

Edward Zywicz[1]
Mem. ASME

David M. Parks
Mem. ASME

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

A closed-form approximate solution for a small-scale yielding (SSY) *plastic zone around a planar interfacial crack tip, occurring between two dissimilar ideally-bonded elastic half spaces, is obtained by equating the elastically-calculated Mises equivalent stress with the material yield strength,* σ_{ys}. *The dimensionless parameter* $\zeta(\theta)$, *which is defined as* $\zeta(\theta) = \angle \mathbf{K} + \epsilon ln r_p(\theta)$, *where* $\angle \mathbf{K}$ *is the phase angle of the complex stress intensity factor* $\mathbf{K}$, ϵ *is the bimaterial constant, and* $r_p(\theta)$, *is the polar representation of the plastic zone radius, naturally arises. The* SSY *interfacial load angle* (ILPA), *defined as* $\zeta_0 = \angle \mathbf{K} + \epsilon ln(\mathbf{K}\bar{\mathbf{K}}/\sigma^2_{ys}\pi\cosh^2(\pi\epsilon))$, *leads to periodic zone growth. The* ILPA *characterizes the overall applied load phase by combining the oscillatory radial phase shift, attributable to the increase in zone size due to increased loading, with* $\angle \mathbf{K}$. *At a particular angle* θ_0 *from the uncracked interface, the plastic zone radius thus calculated is independent of* $\angle \mathbf{K}$, *proportional to* $\mathbf{K}\bar{\mathbf{K}}$, *and has no oscillatory radial phase dependence. The derived plastic zone expression reproduces the shape characteristics, and it modestly reproduces the zone size when compared with solutions for an elastic/perfectly-plastic solid adjoint to an elastic solid. As the strain-hardening exponent in the plastically deforming medium decreases, agreement between the approximation and various accurate numerical solutions improves. In the limiting case when* $\epsilon = 0$, *the well-known homogeneous elastic solutions for pure Mode I and Mode II are recovered, as well as all possible mixed-mode combinations. Approximate validity conditions for the existence of Williams-type asymptotic fields* (*traction-free crack faces*) *are presented.*

1 Introduction

Much effort has recently been focused on interfaces which exist between dissimilar media, with specific attention being directed toward media separation or fracture events. Publications on the subject, such as Shih and Asaro (1988), Hutchinson et al., (1987), and Rice (1988) clarify several aspects of the oscillatory stress solution originally obtained by Williams (1959) for an interfacial crack, and aim to apply or further extend traditional (homogeneous) fracture mechanics approaches to interface cracking phenomena. Elastic interfacial crack-tip fields between isotropic media are well characterized, although only a limited number of geometries have had their stress fields and stress intensity factors solved exactly. Ting (1986) has presented a rigorous framework for determining the degree of singularity and the asymptotic characteristics for the general interfacial crack between two anisotropic elastic materials. When nonlinear material responses are included, no explicit unifying characterization presently exists to unite the various fracture parameters. However, dimensional analyses by Rice (1988) and by Shih and Asaro (1988) lead to symbolic functional relationships consistent with the present results.

Insight concerning contained crack-tip inelastic deformation zones (in the small-scale yielding, SSY, sense) can be obtained by considering the characteristics contained within the elasticity solution. One approximate method which has been used to determine the plastic zone shape and size around a crack tip in a homogeneous medium is equating the elastically-calculated Mises or Tresca equivalent stress with the yield strength of the material (McClintock and Irwin, 1965; Rooke, 1963). The locus of all points satisfying this condition is considered to be the plastic zone boundary which separates the exterior elastic region from the interior plastically-yielding region. The changes in plastic zone size and shape, with respect to the applied load or stress intensity factor(s), can then be estimated from this expression.

The goal of this work is to present a closed-form approximate plastic-zone solution for an interfacial crack between isotropic linear elastic media, and propose various dimensional and dimensionless quantities, which naturally arise in the derivation, as interfacial fracture parameters that uniquely characterize the interface crack-tip region. Comparisons are made between the approximate solution and various precise numerical solutions to demonstrate its accuracy. Conditions

[1] Currently at Lawrence Livermore National Laboratory, Livermore, Calif. 94550.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

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 on the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 8, 1987; final revision, December 11, 1988.

Paper No. 89-WA/APM-6.

which approximately determine the validity of this expression will be stated.

2 SSY Plastic Zone Approximation

The problem considered is a planar interfacial crack, as shown in Fig. 1, whose constituents have shear moduli μ_j ($j=1, 2$) and Poisson's ratios ν_j. (Subscripts 1 and 2 refer to the upper and lower domains, respectively.) Far field loads produce a local elastic stress field which is well characterized by the complex stress intensity factor $\mathbf{K}$ and associated asymptotic interfacial crack-tip stress fields. Following Hutchinson et al., (1987), the stress intensity factor is defined such that, as $r\to 0$ on $\theta=0$, $\sigma_{yy}+i\sigma_{xy}\to \mathbf{K}r^{i\epsilon}/\sqrt{2\pi r}$. For the interfacial "Griffith" crack configuration of length $2a$, this definition for $\mathbf{K}$ differs from $\mathbf{Q}$, the "stress intensity vector" given by Shih and Asaro (1988), by the complex factor $e^{-i\epsilon\ln 2a}$: $\mathbf{K}=\mathbf{Q}e^{-i\epsilon\ln 2a}$. (See Rice (1988) for calculated examples of $\mathbf{K}$ for various geometries and for the interfacial stress fields.)

The bimaterial constant, ϵ, which modulates the stress and displacement oscillation period, can be defined as

$$\epsilon=\frac{1}{2\pi}\ln\frac{\left(\frac{\kappa_1}{\mu_1}+\frac{1}{\mu_2}\right)}{\left(\frac{\kappa_2}{\mu_2}+\frac{1}{\mu_1}\right)}, \tag{1}$$

where $\kappa_j=3-4\nu_j$ for plane strain and $\kappa_j=(3-\nu_j)/(1+\nu_j)$ for plane stress. We note in passing that Dundurs (1969) has shown that elasticity solutions to problems of this class depend functionally on only two dimensionless functions of ν_1, ν_2, and μ_1/μ_2. The second of these parameters, β_2, can be expressed as

$$\beta_2=\frac{\mu_1(\kappa_2-1)-\mu_2(\kappa_1-1)}{\mu_1(\kappa_2+1)+\mu_2(\kappa_1+1)}.$$

Thus, the bimaterial constant ϵ can also be expressed in terms of β_2; e.g., $2\pi\epsilon=\ln[(1-\beta_2)/(1+\beta_2)]$.

The general stress field for an isotropic elastic solid can be represented by the Muskelishvili potentials (Rice, 1988)

$$\sigma_{xx}+\sigma_{yy} = 2[\phi'+\bar{\phi}'] \tag{2}$$

and

$$\sigma_{yy}-\sigma_{xx}+i2\sigma_{xy} = 2[(\bar{z}-z)\phi''-\phi'+\Omega']. \tag{3}$$

Retaining only the dominant asymptotic term as $r\to 0$, the plane-strain elastic potentials in the upper domain are

$$\phi_1' = a_0 e^{-\pi\epsilon} z^{-1/2-i\epsilon} \tag{4}$$

and

$$\Omega_1' = \bar{a}_0 e^{\pi\epsilon} z^{-1/2+i\epsilon}. \tag{5}$$

Using (2)–(5), an expression for the Mises equivalent stress in region 1 can be obtained. (Appendix A contains the complete general series potential functions and formally calculates the Mises equivalent stress). Equating the Mises equivalent stress, $\bar{\sigma}$, in (A27) with the material yield strength, σ_{ys}, and solving for the radius yields

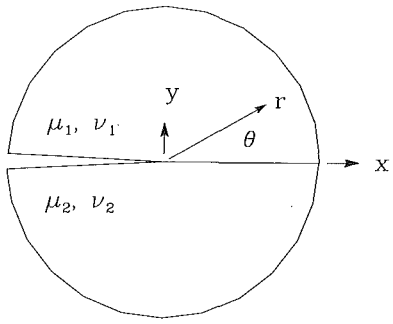


Fig. 1 Schematic interfacial crack tip

where

$$D\equiv\nu_1^2-\nu_1+1, \tag{7}$$

and

$$\zeta(\theta)\equiv\angle\mathbf{K}+\epsilon\ln r_p(\theta). \tag{8}$$

Here θ is the angle measured from the interface, $r_p(\theta)$ is the plastic zone radius from the crack tip, and $\angle\mathbf{K}$ is the phase angle of the (complex) stress intensity factor defined with a branch cut at $\theta=\pi$ such that $\pi>|\angle\mathbf{K}|$. ($\angle\mathbf{K}=\arctan(\Im\mathbf{K}/\Re\mathbf{K})$, which in the homogeneous case, $\epsilon=0$, reduces to $\angle\mathbf{K}=\arctan(K_{II}/K_I)$.) For plane stress conditions, (6) and (8) are still valid; however, (7) is redefined as $D=1$ and the plane stress value for ϵ must be used.

This approximation is valid only when a dominant elastic crack field exists and the maximum extent of the plastic zone is small compared to crack length (L) or other characteristic dimensions (maximum $r_p\ll L$). Further clarification will be stipulated in Section 3.4.

3 Discussion

3.1 Mathematical Considerations. Several interesting mathematical features arise from (6). Foremost, the dimensionless $\zeta(\theta)$ is naturally obtained in the derivation. It removes the dimensional problems associated with assigning length units in $\mathbf{K}$ definitions (Rice, 1988) since $\zeta(\theta)$ is invariant provided $r_p(\theta)$ and $\mathbf{K}$ are evaluated using the same length units. Recall that the generic $\mathbf{K}$ can be expressed as

$$\mathbf{K}=\underline{\sigma}^{\infty}Ce^{-i\epsilon\ln L}\sqrt{\pi L}, \tag{9}$$

where $\underline{\sigma}^{\infty}$ is a complex number (with dimensions of stress) representing the far field load, C is a dimensionless complex geometric constant, and L is the characteristic length dimension. Examination of (9) reveals that when different length units are used to express L, the $\angle\mathbf{K}$ changes. Equation (9) can be rewritten as

$$\mathbf{K}=\|\underline{\sigma}^{\infty}\|\times\|C\|e^{i(\phi-\epsilon\ln L)}\sqrt{\pi L}, \tag{10}$$

where $\|\ \|$ denotes the magnitude of a complex expression,

$$\phi=\angle\underline{\sigma}^{\infty}+\angle C, \tag{11}$$

and

$$\angle\mathbf{K}=\phi-\epsilon\ln L. \tag{12}$$

Substituting (6), (11), and (12) into (8) produces

$$r_p(\theta)=\frac{3\mathbf{K}\bar{\mathbf{K}}}{\sigma_{ys}^2 8\pi\cosh^2(\pi\epsilon)}\times\left\{\begin{array}{l}2\cos(\theta+2\zeta(\theta))\left[\left(\frac{4D}{3}-1\right)e^{2\epsilon(\theta-\pi)}-(2\epsilon\sin\theta+\cos\theta)\right]\\ +e^{2\epsilon(\theta-\pi)}\left[(2\epsilon\sin\theta+\cos\theta)^2+2\left(\frac{4D}{3}-1\right)\right]\\ +e^{2\epsilon(\pi-\theta)}\end{array}\right\}, \tag{6}$$

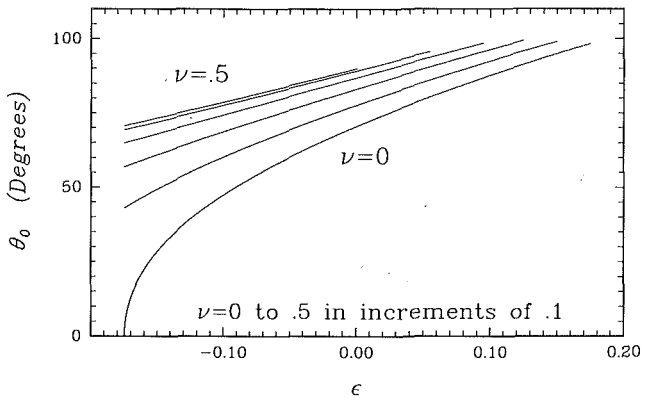


Fig. 2 The angle θ_0 is shown for plane-strain conditions over the complete range of ϵ for various values of ν from 0 to 0.5 in increments of 0.1

$$\zeta(\theta) = \phi - \epsilon \ln L + \epsilon \ln\left\{\frac{\mathbf{K}\bar{\mathbf{K}}}{\sigma_{ys}^2 \pi \cosh^2(\pi\epsilon)} g(\theta, \epsilon, D, \zeta(\theta))\right\}, \quad (13)$$

where $g(\theta, \epsilon, D, \zeta(\theta))$ is a nondimensional function. Using (10), $\mathbf{K}\bar{\mathbf{K}}$ can be expressed as

$$\mathbf{K}\bar{\mathbf{K}} = \|\underline{\sigma}^{\infty}\|^2 \times \|C\|^2 \pi L. \quad (14)$$

Furthermore, (13) can be rearranged and simplified by using (14), reducing to

$$\zeta(\theta) = \phi + 2\epsilon \ln\left\{\frac{\|\sigma^{\infty}\| \times \|C\|}{\sigma_{ys} \cosh(\pi\epsilon)} \sqrt{g(\theta, \epsilon, D, \zeta(\theta))}\right\}. \quad (15)$$

From (15), it is clear that $\zeta(\theta)$ is dimensionless and independent of length units used to express $\mathbf{K}$. This is true as long as a single length measure is assigned to all L used when evaluating $\mathbf{K}$ in, e.g., (10).

For a wide range of engineering interface material properties, an angle θ_0 exists for which the coefficient

$$\left(\frac{4D}{3} - 1\right) e^{2\epsilon(\theta - \pi)} - (2\epsilon \sin\theta + \cos\theta),$$

which multiplies $\cos(\theta + 2\zeta(\theta))$ in (6), is identically zero. Thus, when $\theta = \theta_0$,

$$\left(\frac{4D}{3} - 1\right) e^{2\epsilon(\theta_0 - \pi)} = 2\epsilon \sin\theta_0 + \cos\theta_0. \quad (16)$$

Figure 2 shows the plane-strain θ_0, numerically obtained from (16), for various ν from 0 to 0.5 for the complete range of ϵ, assuming non-negative ν in each material. Note that, θ_0 is generally not the same for plane-strain and plane-stress conditions since, under each condition, ϵ and D have different definitions. The existence of θ_0 indicates that radially, at angle θ_0:

(a) Plastic zone growth is independent of the applied loading phase, $\angle \mathbf{K}$.

(b) The elastically-calculated Mises equivalent stress does not oscillate.

(c) The plastic zone radius is proportional to $\mathbf{K}\bar{\mathbf{K}}$.

Substituting (6) into (8), and defining the SSY interfacial load-phase angle (ILPA), ζ_0, as

$$\zeta_0 \equiv \angle \mathbf{K} + \epsilon \ln\left\{\frac{\mathbf{K}\bar{\mathbf{K}}}{\sigma_{ys}^2 \pi \cosh^2(\pi\epsilon)}\right\}, \quad (17)$$

yields

$$\zeta(\theta) = \zeta_0 + \epsilon \ln\left\{\frac{3}{8}\left[\begin{array}{c} 2\cos(\theta + 2\zeta(\theta))\left[\left(\frac{4D}{3} - 1\right) e^{2\epsilon(\theta - \pi)} - (2\epsilon \sin\theta + \cos\theta)\right] \\ + e^{2\epsilon(\theta - \pi)}\left[(2\epsilon \sin\theta + \cos\theta)^2 + 2\left(\frac{4D}{3} - 1\right)\right] \\ + e^{2\epsilon(\pi - \theta)} \end{array}\right]\right\}. \quad (18)$$

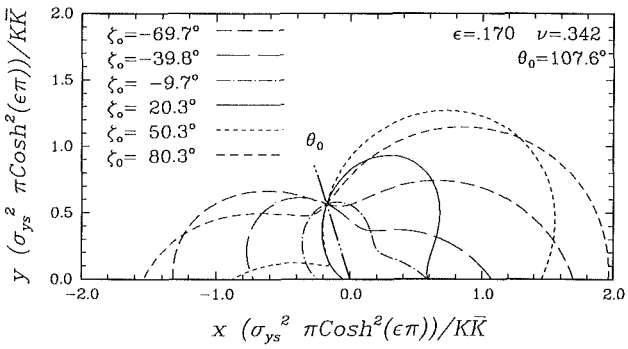


Fig. 3 Approximate plastic zones for various ζ_0 values

Equation (18) reveals that $\zeta(\theta)$ can be additively decoupled into a load-phase dependent quantity, ζ_0, and a transcendental angular dependent function. Alternative definitions of ζ_0, differing trivially by a pure constant, are possible. Such a constant could be chosen, e.g., to approximately normalize the angular function to unity. Equation (17) is a convenient expression for the SSY ILPA since it is an explicit single term representing the total load-phase angle and is common in all $\zeta(\theta)$. *The ILPA totally describes the phase angle of the load by summing the loading phase shift, which is attributable to the change in zone size due to increase in loading, and the load-phase angle* ($\angle \mathbf{K}$).

Shih and Asaro (1988) have independently defined a related load-phase parameter, ξ, for elastic-plastic analysis of interface cracks. Under small-scale yielding conditions, it can be shown that the current load-phase parameter, ζ_0, is related to the parameter ξ of Shih and Asaro by

$$\zeta_0 = \xi - \ln(\pi \cosh^2(\pi\epsilon)). \quad (19)$$

In view of the weak dependence of (19) on ϵ over the practical range of interface elastic constants, ζ_0 and ξ are effectively identical parameterizations of mixity for interface cracks.

Another expression of interest is obtained by evaluating the plastic zone size at $\theta = \theta_0$, in which case, using (16),

$$r_p(\theta_0) = \frac{3\mathbf{K}\bar{\mathbf{K}}}{\sigma_{ys}^2 8\pi \cosh^2(\pi\epsilon)} e^{2\epsilon(\theta_0 - \pi)}\left[\left(\frac{4D}{3} - 1\right) e^{2\epsilon(\theta_0 - \pi)} + e^{-2\epsilon(\theta_0 - \pi)}\right]^2. \quad (20)$$

This entity may prove to be useful in investigating the effects of various material and loading parameters, since it does not contain any radially oscillatory terms and is insensitive to load phase.

3.2 Zone Growth Considerations. From the previous expressions, the overall plastic zone growth characteristics with respect to increasing applied load ($\mathbf{K}$) during SSY can be outlined. After sufficient initial loading has been applied to produce a continuum size plastic zone, the expressions for r_p become valid and applicable. Examination of (6) and (18) shows that zone growth is quasi-proportional to $(\mathbf{K}\bar{\mathbf{K}}/\sigma_{ys}^2)$, and that the zone shape periodically repeats itself with every π increase in ζ_0. For (very) large cracks, it is possible that the plastic zone may repeat its shape during loading. For every

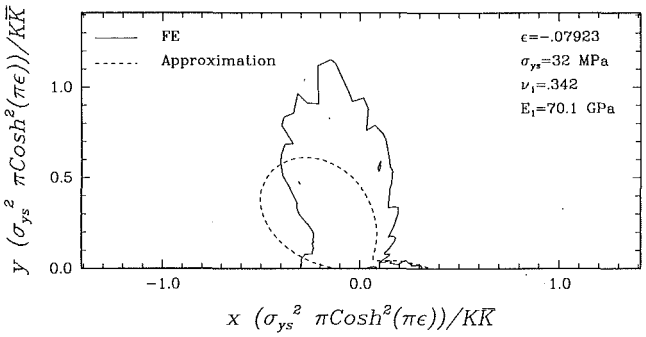


Fig. 4 Plastic zone comparison between the elastic approximation and a finite element solution for an elastic/perfectly-plastic medium adjoint to an elastic medium; $K = 50e^{-0.4636i}$ MPa(m)$^{1/2} + 0.0793i$; $\zeta_0 = -.432$ (-24.7 deg)

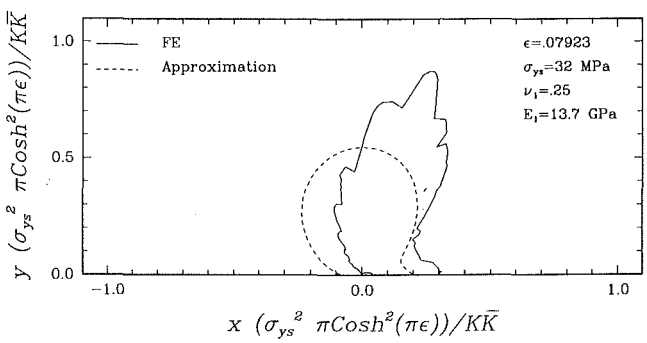


Fig. 6 Plastic zone comparison between the elastic approximation and a finite element solution for an elastic/perfectly-plastic medium adjoint to an elastic medium; $K = 56.7e^{-0.0i}$ MPa(m)$^{1/2} + 0.0793i$; $\zeta_0 = -.007$ ($-.42$ deg)

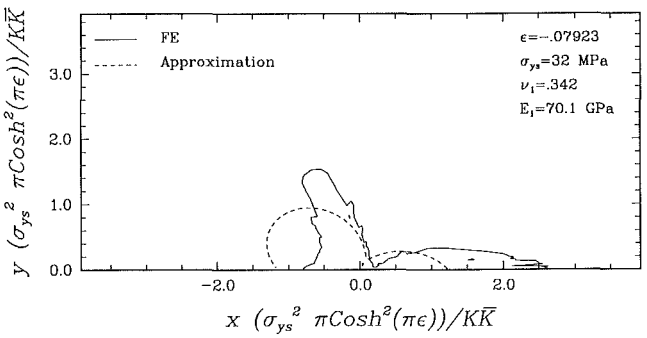


Fig. 5 Plastic zone comparison between the elastic approximation and a finite element solution for an elastic/perfectly-plastic medium adjoint to an elastic medium; $K = 30e^{-0.9272i}$ MPa(m)$^{1/2} + 0.0793i$; $\zeta_0 = -1.004$ (-57.5 deg)

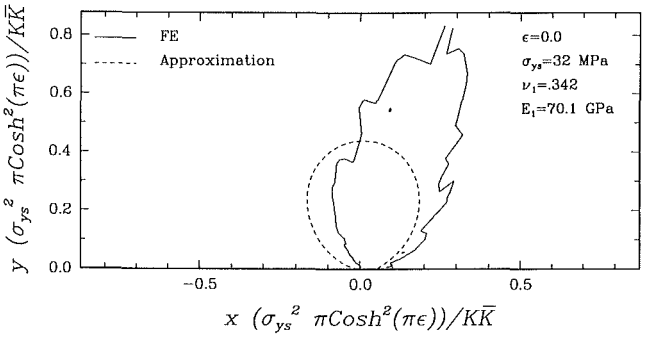


Fig. 7 Plastic zone comparison between the elastic approximation and a finite element solution for an elastic/perfectly-plastic medium adjoint to an elastic medium; $K = 63.4e^{-0.0i}$ MPa(m)$^{1/2} + 0.0i$; $\zeta_0 = 0$

discrete value of ζ_0, a unique zone shape and a unique set of tractions exists along $r_p(\theta)$. Figure 3 shows the plastic zone at various values of ζ_0 for $\epsilon = .170$ and $\nu = .342$. This suggests that ζ_0 uniquely describes the very local crack-tip fields within the zone as long as all previous loading experiences affect the current plastic state in the same manner. For the loading case where cycles of ζ_0 have occurred, this would appear to be true. Since two loadings with unequal tractions can produce identical plastic zones, (e.g., $\zeta_0 = 90$ deg and $\zeta_0 = -90$ deg produce tractions with opposite signs), a full 2π-evaluation of ζ_0 is required to determine all the local fields.

3.3 Comparisons. In the limiting homogeneous elastic case ($\epsilon = 0$), comparison with numerical solutions (Shih, 1974) indicate that the plastic zone shape and size for pure Mode I and Mode II, as well as for various mixed modes, are recovered. Comparing the approximate homogeneous plastic zones with plastic zones numerically obtained for strain-hardening material shows that as the strain-hardening exponent, n, increases (strain α (stress)^n), the elastic approximation overestimates the plastic zone size behind the crack tip and underestimates it ahead of the crack tip. This is accompanied by slight distortional effects which tend to rotate the strain-hardening plastic zone lobes toward the region in front of the crack as compared to the elastic approximation.

Figures 4 to 7 show finite element (FE) calculations of SSY plastic zones for an interfacial crack tip with an elastic/perfectly-plastic medium adjoint to an elastic medium (Zywicz, 1988), and the approximate plastic zones for several values of ϵ, ν_1, μ_1, and $\angle \mathbf{K}$. Although the precise shape is not reproduced, the general size and distribution of the lobe(s), as well as their position(s), are well represented by the simple approximation. An examination of Fig. 5 shows that the size scale is significantly different from that of the other figures, demonstrating the accuracy of the approximation in predicting overall size. Figures 4 to 7 represent the worst case comparisons since perfect plasticity formally represents a strain-hardening exponent of $n = \infty$. The jaggedness of the finite element calculated plastic zones is attributable to extrapolation/approximation errors and mesh discretization. Thus, the jaggedness should only be interpreted as an artifact of the discretization and plotting procedure. Figures 8 and 9 show FE calculations of plastic zones for a deformation theory Ramberg-Osgood strain-hardening material, with strain-hardening exponents $n = 3$ and $n = 10$, respectively, adjoint to a rigid material (Shih and Asaro, 1988), and the approximate plastic zones for several load levels. These FE calculations were performed for a Griffith-type crack, similar to the one shown in Fig. 11, with $\epsilon = .0935$, $L = 2a = 2$ m, and $\nu_1 = .3$, where the stress intensity factor for the geometry and loading is $\mathbf{K} = 1.803\ \sigma^\infty e^{0.1201i}(\mathrm{m})^{1/2 - 0.0935i}$. Here σ^∞ represents the remote stress normal to the crack face (the σ_{yy} stress component, as shown in Fig. 11), and σ_0 is the reference (or yield) stress. The FE plastic zone has been defined as the locus of $\bar{\sigma} = \sigma_0$. The overall sizes and shapes are well characterized by the (asymptotic) approximation. As in the homogeneous case, when the strain-hardening exponent is decreased, the elastic approximation becomes more precise. (Recall, the Ramberg-Osgood material idealization produes a linear response for $n = 1$.) Although the plastic zone radii are not all identically the same at θ_0, the extent of the plastic zone in the vicinity of θ_0 is indeed approximately the same for all loadings (ζ_0).

3.4 Valid Solution Domain. The plastic zone approximation is based upon the assumption that a dominant (Williams type) field exists, as defined in (4)–(5), near the crack tip and transitionally along the plastic zone boundary. This section develops a methodology, based upon exact elasticity solutions for a Griffith crack, for determining approximately when such a Williams-type field exists, and thus defining the valid domain for the characterization of the

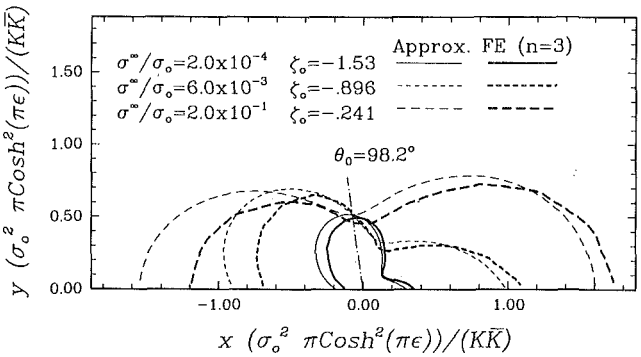


Fig. 8 Plastic zone comparison between the elastic approximation and a finite element solution for a strain-hardening material, $n = 3$, (Shih and Asaro, 1988) for various ζ_0; $\mathbf{K} = \sigma^\infty 1.8025e^{0.1201i}\ (\mathrm{m})^{1/2 - 0.0935i}$, $\epsilon = .0935$, $\nu = .3$

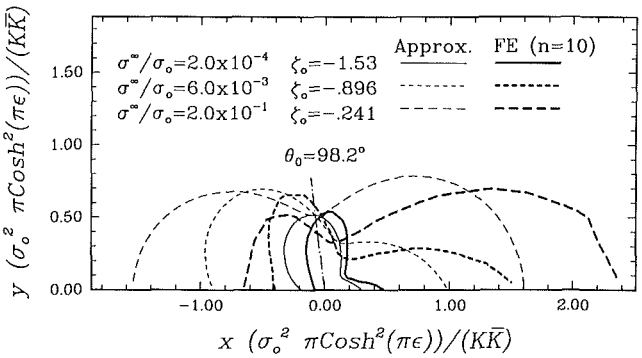


Fig. 9 Plastic zone comparisons between the elastic approximation and a finite element solution for a strain-hardening material, $n = 10$, (Shih and Asaro, 1988) for various ζ_0; $\mathbf{K} = \sigma^\infty 1.8025e^{0.1201i}\ (\mathrm{m})^{1/2 - 0.0935i}$, $\epsilon = .0935$, $\nu = .3$

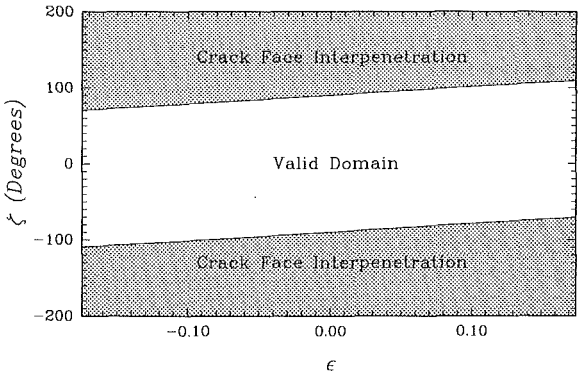


Fig. 10 Crack face interpenetration is shown as a function of ζ and ϵ, for plane-strain conditions

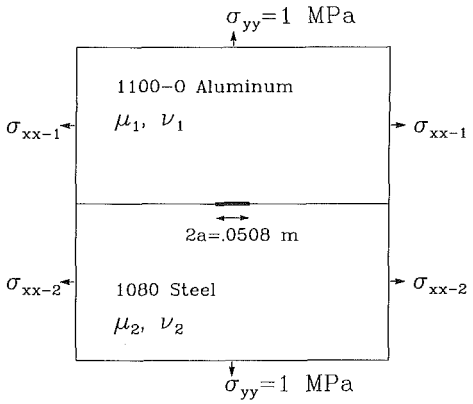


Fig. 11 Geometry for a Griffith-type interfacial crack

plastic zone in terms of ζ_0, $\mathbf{K}$, and material parameters.

In examining the exact elasticity solution for an interfacial crack between two semi-infinite media (Rice and Sih, 1965), the stress potentials can be additively decoupled into singular terms and homogeneous far field terms, and reduced to obtain the dominant asymptotic potentials. Consider Φ, the ϕ_1' stress potential for the Griffith crack, given by Rice and Sih (1965) which is,

$$\Phi = \frac{(z-i2\epsilon a)}{\sqrt{z^2-a^2}}\left(\frac{z+a}{z-a}\right)^{i\epsilon}\frac{\sigma_{yy}^\infty - i\sigma_{xy}^\infty}{1+e^{2\pi\epsilon}} + \frac{\sigma_{xx_1}^\infty + \sigma_{yy}^\infty}{4} - \frac{\sigma_{yy}^\infty}{1+e^{2\pi\epsilon}} + i\left(\frac{\sigma_{xy}^\infty}{1+e^{2\pi\epsilon}} + \frac{2\mu_1\omega_1^\infty}{1+\kappa_1}\right). \tag{21}$$

Here the crack tips are located at $z = \pm a$, and ω_1^∞ is the far-field rotation in Region 1. In the region near the crack tip, the stress potential can be represented by the first line of the right-hand side of (21), namely

$$\Phi \approx \Phi_{\text{near}} = P\left[\frac{(z-i2\epsilon a)}{\sqrt{z^2-a^2}}\left(\frac{z+a}{z-a}\right)^{i\epsilon}\right], \tag{22}$$

where

$$P = \frac{\sigma_{yy}^\infty - i\sigma_{xy}^\infty}{1+e^{2\pi\epsilon}}. \tag{23}$$

To obtain the asymptotic potential, substitute $z = a + x$ in to (22) and assume $\|x\| \ll a$, yielding

$$\Phi_{\text{asymp.}} = P\left[\frac{a(1-i2\epsilon)}{\sqrt{2ax}}\left(\frac{2a}{x}\right)^{i\epsilon}\right] = \bar{\mathbf{K}}\,\frac{x^{-i\epsilon}}{\sqrt{2\pi x}}\,\frac{1}{1+e^{2\pi\epsilon}}. \tag{24}$$

Using (*A*23), (24) can be shown to be identical to (4).

By considering one potential of the exact elasticity solution for a Griffith-type crack, (21), an error parameter can be constructed which represents the discrepancy between the exact solution and the (Williams-type) dominant asymptotic solution, (4)-(5). Normalizing (24) by (22), which is the singular portion of (21), yields the asymptotic norm

$$N = \frac{\Phi_{\text{asymp.}}}{\Phi_{\text{near}}} = \left(1 + \frac{\hat{x}}{0.5 - i\epsilon}\right)^{-1}(1+\hat{x})^{1/2 - i\epsilon}. \tag{25}$$

Here $\hat{x} = x/2a$ is the normalized distance with respect to crack length, and N represents the portion of the singular potential term represented by the asymptotic potential, given by (24). Evaluating (25) along the interface at $\hat{x} = 0.1$ yields $N = 0.874$ for $\epsilon = 0$, while the extreme values $\epsilon = \pm 0.1748$ (positive ν), $N = 0.889e^{\pm .0694i}$. For all ϵ, as $\hat{x} \to 0$, $N \to 1$. This, in conjunction with the previous observations, indicates that the asymptotic expression reproduces the singular term reasonably well over the entire domain where the singular potential term dominates. (From (21), it can be shown that at $\hat{x} \approx 0.13$, the singular term contributes to the total stress potential an amount, equal in magnitude, to that of the homogeneous term. For $|\hat{x}| < 0.13$, the singular portion dominates.)

Based upon the previous discussion, the asymptotic representation, (4)-(5) or (24), is representative in the crack-tip region where

$$\frac{L}{10} \geq r \geq 0. \tag{26}$$

Here, L is the characteristic dimension. (Note, a slight modification has been made for convenience, and that is to limit the domain to $L/10$ instead of $L/8$). Such a conclusion is also typical of homogeneous crack solutions.

A second condition must also be satisfied if (4)-(5) are to

represent the actual dominant asymptotic behavior; namely, that any perturbations within the dominant asymptotic solution domain must be small compared to that domain and occur near the crack tip. Using a St. Venant's-type argument, this can be expressed mathematically as

$$r_{\text{perturbation}_{\max}} \leq \frac{3L}{100}. \tag{27}$$

Such perturbations could include plastic zones and crack face contact and interpenetration, if present. (Note, Williams-type fields, (4)–(5), are based upon the condition that the crack faces are traction-free.) Equation (27) represents a very conservative restriction and, depending upon the actual conditions, it may be appropriate to relax it somewhat.

The asymptotic relative crack-face displacement (CFD), $\Delta\mathbf{u}$, as a function of r (Hutchinson et al., 1987) is

$$\Delta\mathbf{u}(r) \equiv \mathbf{u}(r,\theta=\pi) - \mathbf{u}(r,\theta=-\pi) = \frac{(C_1+C_2)\mathbf{K}r^{i\epsilon}\sqrt{r}}{2\sqrt{2\pi}(1+i2\epsilon)\cosh(\pi\epsilon)}, \tag{28}$$

where

$$\mathbf{u}(r) = u_y(r) + iu_x(r), \tag{29}$$

and C_j are defined according to (A7). Following (A24), ζ is introduced and is defined as

$$\zeta = \angle\mathbf{K} + \epsilon\ln r. \tag{30}$$

Substituting (30) into (28) yields

$$\Delta\mathbf{u}(r) = \frac{(C_1+C_2)\|\mathbf{K}\|\sqrt{r}e^{i\zeta}}{2\sqrt{2\pi}\cosh(\pi\epsilon)(1+i2\epsilon)}. \tag{31}$$

Crack-face interpenetration occurs when $\Delta u_y < 0 \Rightarrow \Re\Delta\mathbf{u} < 0$, or when

$$\cos\zeta + 2\epsilon\sin\zeta < 0. \tag{32}$$

The critical values ζ_l, the beginning and ending points of interpenetration, occur when

$$\cos\zeta_l + 2\epsilon\sin\zeta_l = 0 \tag{33}$$

or,

$$\tan\zeta_l = -\frac{1}{2\epsilon}. \tag{34}$$

Note that for the homogeneous case, the condition represented by (32) occurs any time a negative K_I is applied.

The previous conditions on r, (26) and (27), coupled with the oscillatory crack-face behavior, can be restated as valid solution domain conditions in terms of ζ (via (30)), $\mathbf{K}$ and material parameters. Thus, Williams-type fields, (4) and (5), will exist transitionally along the plastic zone boundary if and only if

$$\cos\zeta + 2\epsilon\sin\zeta > 0 \begin{cases} \epsilon > 0 & \zeta_0 \leq \zeta \leq \zeta_{\max} \\ \epsilon < 0 & \zeta_{\max} \leq \zeta \leq \zeta_0 \end{cases} \tag{35}$$

and

$$\frac{\mathbf{K}\bar{\mathbf{K}}}{\sigma_{ys}^2\pi\cosh^2(\pi\epsilon)} \leq .03L, \tag{36}$$

where

$$\zeta_{\max} = \angle\mathbf{K} + \epsilon\ln(L/10). \tag{37}$$

The condition described by (35) requires that no crack-face contact or interpenetration occurs between the plastic zone boundary and the maximum valid extent of the dominant asymptotic field. It also assumes that the size of the plastic zone along the crack face can be approximated by the characteristic length, $\mathbf{K}\bar{\mathbf{K}}/\sigma_{ys}^2\pi\cosh^2(\pi\epsilon)$. Figure 10 shows for plane strain the values of ζ, as a function of ϵ, which will not produce crack-face interpenetration. In order for (35) to be true, both ζ_0 and $\zeta_{\max}$, as well as the entire path which connects them, must be in the unshaded region of Fig. 10. Note that for $\epsilon = 0$ (homogeneous case) the admissible range is $|\zeta_0| < \pi/2$, corresponding to $K_I > 0$.

Expressions (35)–(37) are necessary, but not sufficient, conditions for a Williams-type field to exist. Crack closure beyond $L/10$ is possible and must be ruled by other considerations, such as global geometrical and loading factors, or by other solutions. However, for a (remotely-loaded) Griffith crack, Comninou and Schmueser (1979) showed that crack closure is continuous from the crack tip outwards; thus if closure exists beyond $L/10$, it will also occur within $L/10$ (with respect to one crack tip). Henceforth, (35)–(37) are also sufficiently validity conditions for a Griffith-type crack.

4 Conclusion

An approximate expression for the plastic zone around an interfacial crack tip in small-scale yielding has been presented. It modestly reproduced the characteristic size and shape, as compared to various precise numerical solutions, with increasing accuracy as the strain-hardening exponent approached unity. The overall crack-tip plastic zone size was found to be quasi-proportional to $(\mathbf{K}\bar{\mathbf{K}}/\sigma_{ys}^2)$. Plastic zones were found to change shape with applied load in a periodic manner dependent upon an interfacial load-phase angle (ILPA), ζ_0. The IPLA was identified as a comprehensive single load-phase angle which determines the zone shape and tractions along the zone boundary, and which may serve as a parameter uniquely characterizing the fields within the zone. Approximate conditions for determining the applicability of this expression were stated in terms of ζ_0 and $\zeta_{\max}$, where $\zeta_{\max}$ is dependent upon the characteristic length in the problem.

From the previous derivations it appears that the ILPA (ζ_0), ϵ, (possibly ν), and the magnitude of $\mathbf{K}$ (expressed as $\mathbf{K}\bar{\mathbf{K}}$ or J, where J is the J-Integral), are the local interfacial fracture mechanics variables needed to describe interfacial SSY behavior. Following homogeneous fracture mechanics, it seems natural to construct interfacial fields analogous to HRR or slip-line fields, utilizing the same material idealizations and similar framework, but with the degree of local (plastic) mode mixity being now dependent upon ζ_0. Using ζ_0 and J as loading conditions describing the SSY plastic zone boundary, the characteristics deep within the zone should be identifiable.

In order to familiarize readers with the application of these concepts to interfacial fracture mechanics, a hypothetical example is included in Appendix B. It demonstrates how to determine various local crack-tip quantities.

Acknowledgments

Special gratitude is expressed to Prof. J. R. Rice for useful discussions.

This work was supported by ONR under grant N00014-84-K-0495.

References

Brown, S. B., Kim, H. K., and Anand, L., 1989, "An Internal Variable Constitutive Model for Hot Working Metals," *International Journal of Plasticity*, in the press.

Comninou, M., and Schmueser, D., 1979, "The Interface Crack in a Combined Tension-Compression and Shear Field," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 345–348.

Dundurs, J., 1969, "Edge-bonded Dissimilar Orthogonal Elastic Wedges Under Normal and Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 650–652.

Hertzberg, R. W., 1976, *Deformation and Fracture Mechanics of Engineering Materials*, John Wiley and Sons, New York, p. 20.

Hutchinson, J. W., Mear, M., and Rice, J. R., 1987, "Crack Paralleling an Interface Between Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 828–832.

McClintock, F. A., and Irwin, G. R., 1965, "Plasticity Aspects of Fracture Mechanics," *Fracture Toughness Testing and Its Applications*, ASTM STP 381, ASTM, Philadelphia, pp. 84–114.

Rice, J. R., 1988, "Elastic Fracture Mechanics Concepts for Interfacial Cracks," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 98–103.

Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," ASME JOURNAL OF APPLIED MECHANICS, Vol. 32, pp. 418–423.

Rooke, D. P., 1963, "Elastic Yield Zone Round a Crack Tip," Royal Aircraft Establishment, Farnborough, Tech. Note CPM 29.

Shih, C. F., 1974, "Small-Scale Yielding Analysis of Mixed Mode Plane Strain Crack Problems," *Fracture Analysis*, ASTM STP 560, ASTM, Philadelphia, pp. 187–210.

Shih, C. F., and Asaro, R. J., 1988, "Elastic-Plastic Analysis of Cracks on Bi-material Interfaces; Part 1: Small Scale Yielding," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 299–316.

Ting, T. C., 1986, "Explicit Solution and Invariance of the Singularities at an Interface Crack in Anisotropic Composites," *International Journal of Solids and Structures*, Vol. 22, No. 9, pp. 965–983.

Williams, M. L., 1959, "The Stresses Around a Fault or Crack in Dissimilar Media," *Bulletin of the Seismological Society of America*, Vol. 49, pp. 199–204.

Zywicz, E., 1988, "On Elastic-Plastic Cracks Between Dissimilar Media," Ph.D. Thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass.

APPENDIX A

Mises Equivalent Stress Derivation

An asymptotic expression for the Mises equivalent stress around an interfacial plane-strain crack tip, as a function of r and θ, is derived.

The general series potential functions for an interface crack, as expressed by Rice (1988), are

$$\phi_1' = e^{-\pi\epsilon} z^{-1/2-i\epsilon} f(z) + 2C_2 g(z)/(C_1+C_2), \tag{A1}$$

$$\phi_2' = e^{\pi\epsilon} z^{-1/2-i\epsilon} f(z) + 2C_1 g(z)/(C_1+C_2), \tag{A2}$$

$$\Omega_1' = e^{\pi\epsilon} z^{-1/2+i\epsilon} \bar{f}(z) - 2C_2 \bar{g}(z)/(C_1+C_2), \tag{A3}$$

and

$$\Omega_2' = e^{-\pi\epsilon} z^{-1/2+i\epsilon} \bar{f}(z) - 2C_1 \bar{g}(z)/(C_1+C_2), \tag{A4}$$

with

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \tag{A5}$$

$$g(z) = \sum_{n=0}^{\infty} b_n z^n \tag{A6}$$

and

$$C_j = (1+\kappa_j)/\mu_j. \tag{A7}$$

Here μ_j are the shear moduli, $\kappa_j = 3-4\nu_j$ for plane strain and $\kappa_j = (3-\nu_j)/(1+\nu_j)$ for plane stress, ν_j are the Poisson's ratios, and the subscripts 1 and 2 refer to the domains above and below the interface, respectively.

From (2) and (3), the individual stress components can be expressed as

$$\sigma_{xx} = \frac{1}{2}(B+\bar{B}), \tag{A8}$$

$$\sigma_{yy} = \frac{1}{2}(A+\bar{A}), \tag{A9}$$

and

$$\sigma_{xy} = \frac{-i}{2}(A-\bar{A}), \tag{A10}$$

where

$$A = (\bar{z}-z)\phi'' + \bar{\phi}' + \Omega', \tag{A11}$$

and

$$B = 2\phi' + \bar{\phi}' - \Omega' - (\bar{z}-z)\phi''. \tag{A12}$$

Here $i=\sqrt{-1}$ and a bar denotes the complex conjugate. For plane-strain isotropic elastic solids, the Mises equivalent stress is

$$\bar{\sigma}^2 = (\sigma_{xx}^2 + \sigma_{yy}^2)D + (\sigma_{xx}\sigma_{yy})F + 3\sigma_{xy}^2, \tag{A13}$$

with

$$D = \nu^2 - \nu + 1 \tag{A14}$$

and

$$F = 2\nu^2 - 2\nu - 1, \tag{A15}$$

where ν is the Poisson's ratio of the solid. For plane-stress isotropic elastic solids (A13) is still valid, but (A14) and (A15) are redefined as $D=1$ and $F=-1$, respectively. After substituting (A8)–(A10) into (A13) and doing some *complex* algebra, (A13) is written as

$$\bar{\sigma}^2 = \frac{1}{2}\Re\{(D-3)AA + (D+3)A\bar{A} + DB\bar{B} + DBB + FAB + FA\bar{B}\}. \tag{A16}$$

Further simplification is obtained by using (A11) and (A12), so that (A16) becomes

$$\bar{\sigma}^2 = \Re\{3(\bar{z}-z)(z-\bar{z})\phi''\bar{\phi}'' - 6(\bar{z}-z)\phi''\bar{\phi}' + 6(\bar{z}-z)\phi''\bar{\Omega}' + (8D-6)\phi'\phi' + (8D-3)\phi'\bar{\phi}' + 3\Omega'\bar{\Omega}' - 6\Omega'\bar{\phi}'\}. \tag{A17}$$

The asymptotic potential functions for the upper domain, (4) and (5), are obtained by considering the dominant term in (A1) and (A3) as $r\to 0$. At this point attention shall be focused upon the upper domain since the lower domain solution is obtainable by substituting $-\epsilon$ for ϵ. Differentiating (4), using $z=re^{i\theta}$, expanding out (4) and (5), and defining

$$\mathcal{J} = a_o r^{-i\epsilon}, \tag{A18}$$

we obtain

$$\phi_1' = \mathcal{J} e^{\epsilon(\theta-\pi)} e^{-i\theta/2} r^{-1/2}, \tag{A19}$$

$$\phi_1'' = \mathcal{J} e^{\epsilon(\theta-\pi)} e^{-i3\theta/2} r^{-3/2}\left(-\frac{1}{2} - i\epsilon\right), \tag{A20}$$

and

$$\Omega_1' = \bar{\mathcal{J}} e^{\epsilon(\pi-\theta)} e^{-i\theta/2} r^{-1/2}. \tag{A21}$$

Substituting (A19)–(A21) into (A17) yields

$$\bar{\sigma}^2 = \frac{1}{r}\Re\{\mathcal{J}\mathcal{J}[(-3-i6\epsilon)(e^{-i2\theta}-1) - 6 + e^{2\epsilon(\theta-\pi)}(8D-6)(\cos\theta - i\sin\theta)] + \mathcal{J}\bar{\mathcal{J}}[e^{2\epsilon(\theta-\pi)}(1-\cos 2\theta)\left(\frac{3}{2}+6\epsilon^2\right) - e^{2\epsilon(\theta-\pi)}(-3-i6\epsilon)(e^{-i2\theta}-1) + e^{2\epsilon(\theta-\pi)}(8D-3) + 3e^{2\epsilon(\pi-\theta)}]\}. \tag{A22}$$

The constant a_0 is related to the complex stress intensity factor $\mathbf{K}$ (Rice, 1988) via

$$a_0 = \frac{\bar{\mathbf{K}}}{2\sqrt{2\pi}\cosh(\pi\epsilon)}. \tag{A23}$$

Defining ζ as

$$\zeta = \angle \mathbf{K} + \epsilon \ln r, \tag{A24}$$

and using (A18), we find

$$\mathcal{J}\mathcal{J} = \frac{\mathbf{K}\bar{\mathbf{K}}}{8\pi\cosh^2(\pi\epsilon)}(\cos 2\zeta - \sin 2\zeta), \tag{A25}$$

and

$$\mathcal{J}\bar{\mathcal{J}} = \frac{\mathbf{K}\bar{\mathbf{K}}}{8\pi\cosh^2(\pi\epsilon)}. \tag{A26}$$

The complete expression for the Mises equivalent stress is obtained by substituting (A25) and (A26) into (A22), and can be expressed as

$$\bar{\sigma}^2 = \frac{3\mathbf{K}\bar{\mathbf{K}}}{r 8\pi\cosh^2(\pi\epsilon)} \times \left\{ \begin{array}{l} 2\cos(\theta + 2\zeta(\theta))\left[\left(\frac{4D}{3} - 1\right)e^{2\epsilon(\theta-\pi)} - (2\epsilon\sin\theta + \cos\theta)\right] \\ \quad + e^{2\epsilon(\theta-\pi)}\left[(2\epsilon\sin\theta + \cos\theta)^2 + 2\left(\frac{4D}{3} - 1\right)\right] \\ \quad + e^{2\epsilon(\pi-\theta)} \end{array} \right\}. \tag{A27}$$

APPENDIX B

Interfacial Crack Example

A detailed hypothetical example demonstrating the procedures to characterize a plane-strain interfacial Griffith-type crack between 1100-O Aluminum and 1080 Steel is presented. The geometry considered is shown in Fig. 11, and the material properties are listed in Table 1. From (1), $\epsilon = .03373$. For this geometry, with the appropriate σ_{xx_j} imposed such that the interface remains straight, the stress intensity factor for the right-hand crack tip in terms of the far-field stresses is (Rice, 1988)

$$\mathbf{K} = (\sigma_{yy} + i\sigma_{xy})(1 + i2\epsilon)(2a)^{-i\epsilon}\sqrt{\pi a}.$$

The stress intensity factor for the left-hand crack tip is the same as for the right-hand crack tip because the applied load is symmetric ($\sigma_{xy} = 0$). Substituting in for the numerical values $\sigma_{yy} = 1$ MPa, $\epsilon = 0.03373$, and $2a = 0.0508$ m yields,

$$\mathbf{K} = .2831 e^{-0.03315i}\ \text{MPa(m)}^{1/2 - 0.03373i}.$$

Using (17), the ILPA is $\zeta_0 = -.33982$ radians (-19.47 deg). The characteristic plastic zone length $\mathbf{K}\bar{\mathbf{K}}/\sigma_{ys}^2 \pi\cosh^2(\pi\epsilon) = 1.577 \times 10^{-5}$ m. Evaluating (36) indicates that the characteristic plane zone length is sufficiently small compared to crack length. (Alternatively, from (6), (17), and (18) the maximum size of the plastic zone is 8.88×10^{-6}m and occurs at $\theta = 122$ deg. Comparing r_p to the crack length gives, $r_p/2a = 1.748 \times 10^{-4}$.) From (37), $\zeta_{\max} = -.2113$ radians (-12.11 deg). Checking (35) indicates that no crack face interpenetration is anticipated. Thus, at this loading all the SSY conditions and the assumption of no crack face interpenetration are satisfied.

Table 1 Material properties for 1100-O Aluminum and 1080 Steel (Hertzberg, 1976)

Material	μ(GPa)	ν	σ_{ys}(MPa)
1100-O Al	26.1	.342	42.[1]
1080 Steel	80.7	.300	585.

[1] Brown et al., 1989.

Thermal Stresses at the Edge of a Bimetallic Thermostat

An-Yu Kuo
Consultant,
Structural Integrity Associates,
San Jose, Calif. 95118
Assoc. Mem., ASME

The plane stress problem of a semi-infinite, bimetallic thermostat subjected to uniform heating or cooling is treated with the theory of elasticity. Solutions to this problem are expressed as the sum of a basic solution for a bimetallic strip of infinite length and a series of complementary solutions. Interlayer peeling stresses at the free edge of the bimetallic thermostat are shown to be singular or nonsingular (but still higher than the nominal values) depending upon whether the combination of the two Dundurs' bimaterial constants, $\alpha(\alpha-2\beta)$, are greater or less than zero. In an example problem, current solutions agree well with finite element results while results predicted by a modified beam theory show a large deviation from the other two solutions near the free edge. Boundary layer effects near the free edges of a bimetallic thermostat are also discussed.

1 Introduction

Thermal stress in bimetallic thermostats subjected to uniform heating or cooling has been of interest for many years. Timoshenko (1925) and later Boley and Weiner (1960) solved the problem by the classical beam theory for an infinitely long, bimaterial elastic strip. Chen and Nelson (1979) used the concept of force equilibrium in calculating thermal stresses of bonded joints in electronic devices. Recently, Suhir (1986) presented a solution for a finite length, bimetallic thermostat with free edges at both ends. Adding to the normal stresses predicted by the strength of material approach (Timoshenko, 1925, and Boley and Weiner, 1960), Suhir calculated interlayer peeling and shear stresses at the free edges by the use of a simplified interface compliance. With this simplified interface compliance, Suhir was able to show high stress concentration near the free edges. However, as later pointed out by Razaqpar and Suhir (1987), the simplified approach used by Suhir (1986) has its inherent shortcomings in predicting accurate interlayer peeling and shear stresses near the free edges.

It has been shown by Bogy (1968, 1970), Dundurs (1969), Hein and Erdogan (1971), and many others that, under certain combinations of material properties, stresses at a bimaterial wedge, such as the edge of a bimetallic thermostat, may behave singularly. This paper presents an analytical solution to the thermal stresses at the free edge of a semi-infinite, bimetallic strip subjected to uniform heating or cooling.

2 Formulation

As illustrated in Fig. 1, this paper considers a semi-infinite, bimetallic thermostat subjected to an uniform temperature change of ΔT. Plane stress condition is assumed in this paper. The plane strain solutions can be easily obtained by extending the plane stress solutions through the use of an equivalent Young's modulus and Poisson's ratio. It is also assumed that both materials are homogeneous, isotropic, and elastic with Young's moduli E'' and E', Poisson's ratios ν'' and ν', and coefficients of thermal expansion δ'' and δ', respectively, for the upper and lower layers. Thicknesses of the two layers are h'' for the upper strip and h' for the lower strip. Throughout this paper, the superscript prime is designated to variables or functions in the lower strip and the superscript double-prime is devoted to variables or functions in the upper strip.

As the first step, we will find a basic solution to the problem of an infinitely long, bimetallic thermostat under uniform temperature change ΔT (see Fig. 2). For semi-infinite thermostats, a complementary solution will be needed later to account for the free edge. It has been shown by Boley and Weiner (1960) that, for an infinitely long rectangular beam under uniform temperature change, the classical beam theory can provide exact solutions to the problem. Thus, as shown in Fig. 2, moment equilibrium leads to the following condition:

$$M' + M'' = P\ (h' + h'')/2. \tag{1}$$

Continuities of curvature and x-displacement at the interface lead to another two conditions:

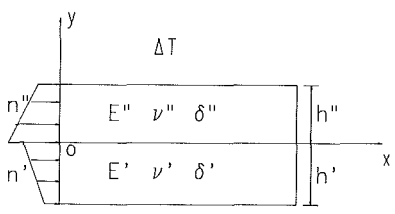


Fig. 1 A bimetallic thermostat

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10-15, 1989.

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 on the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, February 12, 1988; final revision, November 8, 1988.

Paper No. 89-WA/APM-7.

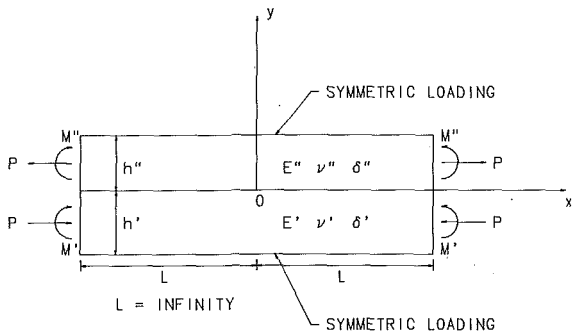


Fig. 2 An infinitely long, bimetallic strip

$$\frac{M'}{E'I'}=\frac{M''}{E''I''} \tag{2}$$

$$\delta'\Delta T+\frac{P}{E'A'}+\frac{M'h'}{2E'I'}=\delta''\Delta T+\frac{P}{E''A''}+\frac{M''h''}{2E''I''} \tag{3}$$

where A and I are cross-sectional area and bending rigidity, respectively, and, as shown in Fig. 2, P and M are force and moments, respectively, acting at the center line of the two strips. With equations (1)–(3), the three unknowns, P, M', and M'', can be solved explicitly, and thereafter, the normal stresses σ'_{xx} and σ''_{xx} can be estimated by

$$\sigma'_{xx}(x,y)=\sigma'_{xx}(y)=-\frac{P}{A'}-\frac{M'}{I'}\left(y+\frac{h'}{2}\right) \tag{4}$$

$$\sigma''_{xx}(x,y)=\sigma''_{xx}(y)=\frac{P}{A''}-\frac{M''}{I''}\left(y-\frac{h''}{2}\right). \tag{5}$$

Note that, in the classical beam theory, all the other stress components, σ'_{xy}, σ''_{xy}, σ'_{yy}, and σ''_{yy}, vanish. It is obvious that equations (4) and (5) are valid only for infinitely long thermostats. For a semi-infinite thermostat with a free edge at the left end ($x=0$), the aforementioned solutions must be superimposed with a complementary solution to account for the stress-free boundary conditions at the edge. As illustrated in Fig. 1, the complementary solution is the solution of a semi-infinite, bimetallic thermostat ($\Delta T=0$) subjected to edge loads n' and n'', which are the negative of the normal stresses predicted by equations (4) and (5).

Numerically, the complementary solution can be obtained by the alternating method. That is, by solving and superimposing a series of two fundamental problems: (*a*) two joined quarter-spaces loaded at the boundary (see Fig. 3), and (*b*) an infinitely long, bimetallic strip loaded symmetrically (with respect to $x=0$) at both the upper ($y=h'$) and lower ($y=h''$) surfaces (see Fig. 2).

3 Surface Loading in Two Joined Quarter-Spaces

As shown in Fig. 3, the first fundamental problem is a composite half space with prescribed surface normal loading $n''(y)$ and $n'(y)$ on the boundary surface ($h''<y<-h'$). This problem has been solved by Bogy (1968, 1970) through the use of two Airy stress functions and the Mellin transform. It was later shown by Dundurs (1969) and Bogy (1970) that, as the distance r from the free edge approaches zero, stresses near the free edge have an asymptotic term of $r^{-\lambda}$, Ln(r), or r^0 ($r\to 0$) depending upon whether $\alpha(\alpha-2\beta)$ is greater, equal, or less than zero, where $\lambda=\lambda(\alpha,\beta)$ is a positive number between 0 and 1, and α and β are the two composite material parameters defined by Dundurs (1969). Definition of λ will be given in the latter part of this section. Bogy (1970) has also pointed out that the logarithmic singularity, Ln(r) as $r\to 0$, exists only when there is a finite discontinuity in the shear load on the boundary of the composite half plane, i.e., only when

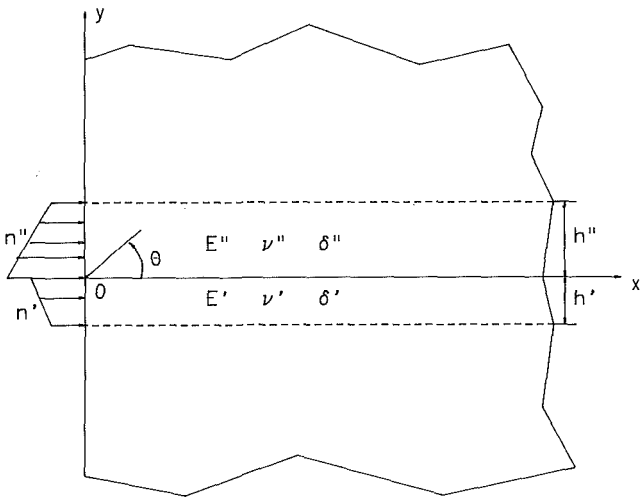


Fig. 3 Two joined quarter-space loaded at the boundary

$t'(0)\neq t''(0)$ where t' and t'' are applied shear loads on the boundary surface. Since neither equations (4) and (5), nor solutions to the second fundamental problem, which will be discussed in the next section, would generate any shear loads on the boundary surface ($x=0$), the logarithmic singularity does not exist in the first fundamental problem. That implies that thermal stresses resulting from an uniform temperature change at the edge of a semi-infinite bimetallic thermostat will behave asymptotically like either $r^{-\lambda}$ or r^0, depending on whether $\alpha(\alpha-2\beta)$ is greater or less than zero.

Taken from Bogy's paper (1970), stress solutions to the first fundamental problem are

$$\sigma'_{rr}(r,\theta)=\frac{1}{\pi r}\int_0^\infty \mathrm{Re}\hat{\sigma}'_{rr}[(-1-i\eta,\theta)\,r^{i\eta}]\,d\eta$$

$$-\frac{1}{2}\lim_{s\to -1}\,[(s+1)\,\hat{\sigma}'_{rr}(s,\theta)\,r^{-1}] \tag{6}$$

$$\sigma'_{\theta\theta}(r,\theta)=\frac{1}{\pi r}\int_0^\infty \mathrm{Re}[\hat{\sigma}'_{\theta\theta}(-1-i\eta,\theta)\,r^{i\eta}]\,d\eta \tag{7}$$

$$\sigma'_{r\theta}(r,\theta)=\frac{1}{\pi r}\int_0^\infty \mathrm{Re}[\hat{\sigma}'_{r\theta}(-1-i\eta,\theta)\,r^{i\eta}]\,d\eta \tag{8}$$

and similar equations for $\sigma''_{jk}(r,\theta)$ by replacing $\hat{\sigma}'_{jk}$ with $\hat{\sigma}''_{jk}$. In equations (6)–(8), $\hat{\sigma}'_{jk}$ are defined as

$$\hat{\sigma}'_{rr}(s,\theta)=\frac{s+1}{16D}\left(s\sum_{k=1}^{4}+4\sum_{k=3}^{4}\right)\{C'^{(k)}\,[M_1{}^{(k)}L_1+M_3{}^{(k)}L_3]\} \tag{9}$$

$$\hat{\sigma}'_{\theta\theta}(s,\theta)=-\frac{s(s+1)}{16D}\sum_{k=1}^{4}\{C'^{(k)}\,[M_1{}^{(k)}L_1+M_3{}^{(k)}L_3]\} \tag{10}$$

$$\hat{\sigma}'_{r\theta}(s,\theta)=-\frac{s+1}{16D}\left(s\sum_{k=1}^{4}+2\sum_{k=3}^{4}\right)\{(-1)^kS'^{(k)}[M_1{}^{(k)}L_1+M_3{}^{(k)}L_3]\} \tag{11}$$

where

$$L_1(s)=-\frac{1}{s(s+1)}\int_0^{h'}n'(r)\,r^{s+1}\,dr \tag{12}$$

$$L_3(s)=-\frac{1}{s(s+1)}\int_0^{h''}n''(r)\,r^{s+1}dr \tag{13}$$

$$D = D(\alpha,\beta,\lambda) = [\beta \cos^2\left(\frac{\lambda\pi}{2}\right) + (\alpha-\beta)(s+1)^2]^2$$

$$+ \cos^2\left(\frac{s\pi}{2}\right)\sin^2\left(\frac{s\pi}{2}\right) - \alpha^2 (s+1)^2 \tag{14}$$

$$M_j^{(k)}(s) = P_j^{(k)}(s) + Q_j^{(k)}(s) \cos^2\left(\frac{s\pi}{2}\right) \tag{15}$$

$$\bar{M}_{j+2}^{(k)}(s) = \bar{P}_{j+2}^{(k)}(s) + \bar{Q}_{j+2}^{(k)}(s) \cos^2\left(\frac{s\pi}{2}\right) \tag{16}$$

$$\alpha = \frac{E'-E''}{E'+E''},\ \beta = \frac{2[E'(1-\nu'')-E''(1-\nu')]}{4(E'+E'')}, \tag{17}$$

and λ is the root between 0 and 1 of the equation $D(\alpha,\beta,\lambda-2)=0$. Similar definitions for $\hat{\sigma}_{jk}''(s,\theta)$ can be written by replacing $M_j^{(k)}$ with $\bar{M}_{j+2}^{(k)}$ in equations (9)–(11). In the aforesaid equations, functions $C'^{(k)}$, $C''^{(k)}$, $S'^{(k)}$, $S''^{(k)}$, $P_j^{(k)}$, $\bar{P}_j^{(k)}$, $Q_j^{(k)}$, and $\bar{Q}_j^{(k)}$ are defined in pages 1293 and 1294 of Bogy's paper (1970) and thus are not repeated in this paper.

Thus, to get the solution for the first fundamental problem, we can convert the two surface loading functions, $n'(r)$ and $n''(r)$, into the frequency domain by equations (12) and (13), and thereafter, calculate stresses by equations (6)–(8).

4 Surface Loading in an Infinite Bimetallic Strip

In the second fundamental problem we need to solve, as shown in Fig. 2, an infinitely long, bimetallic strip loaded symmetrically with respect to the y-axis ($x=0$) at both top ($y=h''$) and bottom ($y=-h''$) surfaces. Like the first fundamental problem, solutions to the second fundamental problem can also be expressed in terms of two Airy stress functions, $\phi'(x,y)$ and $\phi''(x,y)$ for the upper and lower layers, respectively. To solve the second fundamental problem, we first convert the boundary loading at top and bottom surfaces into an integral form as follows:

$$\sigma_{yy}''(x,h'') = \int_0^\infty F_1(s)\cos(sx)\,ds \tag{18}$$

$$\sigma_{xy}''(x,h'') = \int_0^\infty F_2(s)\sin(sx)\,ds \tag{19}$$

$$\sigma_{yy}'(x,-h') = \int_0^\infty G_1(s)\cos(sx)\,ds \tag{20}$$

$$\sigma_{xy}'(x,-h') = \int_0^\infty G_2(s)\sin(sx)\,ds \tag{21}$$

where F_1, F_2, G_1, and G_2 are Fourier transforms of the loading functions applied on the upper and lower surfaces. We can then show that the two stress functions are as follows:

$$\phi''(x,y) = \int_0^\infty [A_1 e^{sy} + A_2 e^{-sy} + A_3 y e^{sy} + A_4 y e^{-sy}]\cos(sx)\,ds \tag{22}$$

$$\phi'(x,y) = \int_0^\infty [A_5 e^{sy} + A_6 e^{-sy} + A_7 y e^{sy} + A_8 y e^{-sy}]\cos(sx)\,ds \tag{23}$$

where the coefficients $A_j = A_j(s)$ are determined by a set of eight simultaneous equations:

$$H_{jk}A_k = R_j \qquad (j,k=1,8). \tag{24}$$

Four out of these eight equations result from the four stress boundary conditions (two each on top and bottom surfaces) and the other four result from the four continuity conditions on the interface (two for stress continuities and two for displacement continuities). Definitions of H_{jk} and R_j are given in the Appendix at the end of this paper.

Thus, for a given set of surface loads, $\sigma_{yy}''(x,h'')$, $\sigma_{xy}''(x,h'')$, $\sigma_{yy}'(x,-h')$, and $\sigma_{xy}'(x,-h')$, solutions to the second fundamental problem are given by equations (22) and (23).

5 Procedure for Total Solution and Numerical Results

The total stress solution to the problem of a semi-infinite, bimetallic thermostat subjected to an uniform temperature change is the superposition of the basic solution derived in Section 2 and a series of the two complementary solutions discussed in Section 3 and 4. The following numerical procedure has been set up to obtain the total solution:

(*a*) Calculate stress $[\sigma_{jk}']_0$ and $[\sigma_{jk}'']_0$ by equations (4) and (5) and set $m=1$.

(*b*) Define n' and n'' as:

$$n'(r) = -[\sigma_{xx}'(0,r)]_{2m-2} \tag{25}$$

$$n''(r) = -[\sigma_{xx}''(0,-r)]_{2m-2}. \tag{26}$$

(*c*) Substitute $n'(r)$ and $n''(r)$ into equations (12) and (13) and calculate stresses $[\sigma_{jk}']_{2m-1}$ by equations (6)–(11). Procedure for calculating $[\sigma_{jk}'']_{2m-1}$ is similar.

(*d*) Calculate $F_1(s)$ by

$$F_1(s) = -\frac{2}{\pi}\int_0^\infty [\sigma_{yy}'(x,h')]_{2m-1}\cos(sx)\,dx \tag{27}$$

and, similarly, for $F_2(s)$, $G_1(s)$, and $G_2(s)$. Then, calculate stresses $[\sigma_{jk}]_{2m}$ according to the Airy stress functions defined in equations (18)–(21).

(*e*) Check convergence. If the solution meets the convergence criterion, go to step (*f*) – otherwise increase m by 1 and go back to step (*b*).

(*f*) Calculate total stresses by

$$\sigma_{jk} = [\sigma_{jk}]_0 + \sum_{m=1}^{M}\{[\sigma_{jk}]_{2m-1} + [\sigma_{jk}]_{2m}\}. \tag{28}$$

There are several ways to set the convergence criterion in step (*e*). In this paper, convergence of the solution is deemed as being achieved when the relative increment of σ_{yy} (0,0) is less than 0.01.

As a numerical example, the same material properties and geometrical dimensions of the problem solved in Suhir's paper (1986) were used in this paper, i.e.,

$E'' = 325000$ MPa, $\nu'' = 0.293$, $\delta'' = 4.9\times10^{-6}$ 1/°C,

$E' = 70380$ MPA, $\nu' = 0.345$, $\delta' = 23.6\times10^{-6}$ 1/°C,

$h' = h'' = 2.5$ mm, $\Delta T = 240$°C.

The two materials in this example problem are molybdenum and aluminum, respectively, for the upper and lower strips. The corresponding bimaterial constants, α and β, and the order of stress singularity, λ, at the free edge are found to be -0.6330, -0.1395, and 0.1485, respectively. Thus, as shown by Dundurs (1969) and Bogy (1970), the interlayer stresses will behave asymptotically like $r^{-0.1485}$ as r approaches zero at the free edge. Numerical quadratures were used to calculate the integrations in the solutions. With only three iterations, the stress solutions converge. Resulting interlayer stress distributions are plotted and compared with the beam theory solutions (Suhir, 1986) and results of a finite element analysis in Figs. 4–7. The stresses shown in Figs. 4–7 have been normalized by a constant, $(E''\delta''\Delta T)$. The normalized asymptotic coefficients of the singular term, which is of order $(r/h'')^{-0.1485}$ as (r/h'') approaches zero, are 0.8038, -0.4148, -1.3733, and 0.3001 for σ_{xx}', σ_{xx}'', σ_{yy}'' $(=\sigma_{yy}')$, and σ_{xy}'' $(=\sigma_{xy}')$, respectively.

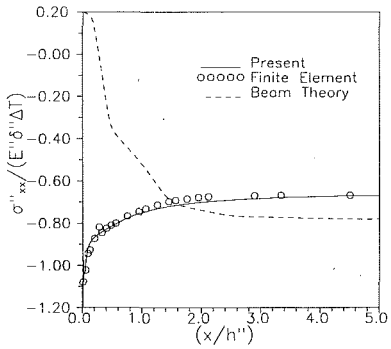


Fig. 4 Comparison of normalized σ''_{xx} (x, 0)

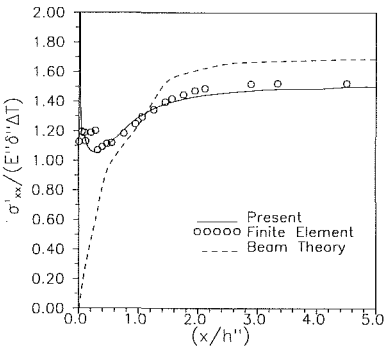


Fig. 5 Comparison of normalized σ'_{xx} (x, 0)

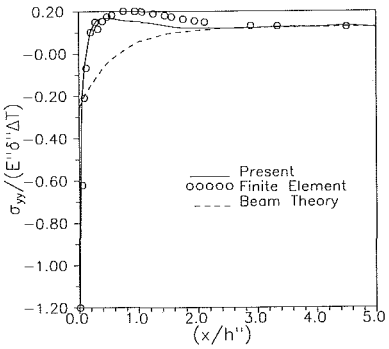


Fig. 6 Comparison of normalized σ_{yy} (x, 0)

The finite element mesh used to obtain the stress results in Figs. 4–7 is shown in Fig. 8. In this finite element model, all the elements are the two-dimensional, eight-node, isoparametric elements. In Figs. 4–7, it is seen that current solutions agree well with finite element results for all stress components. However, results predicted by the modified beam theory (Suhir, 1986) are off by a large margin near the free edge due to the inherent shortcoming of the strength of material approach.

In this example problem, it is observed that the interlayer peeling stress σ_{yy} $(x,0)$ has an asymptotic form of

$$\sigma_{yy}(x,0) = K_{yy}(E''\delta''\Delta T)\left(\frac{x}{h''}\right)^{-0.1485} + \text{higher order terms} \tag{29}$$

near the free edge, where K_{yy} is a constant and is defined as intensity factor of the peeling stress σ_{yy}. To check the convergence of the solution scheme described in Section 5, calculations for the example problem were carried out to twenty iterations and are plotted in Fig. 9. At the end of the first five iterations, the intensity factor of the peeling stress, K_{yy} in equation (29), are -1.3406, -1.3680, -1.3733, -1.3754, and -1.3764, respectively. Although not shown in Fig. 9, a monotoic and quick convergence similar to that for K_{yy} is also

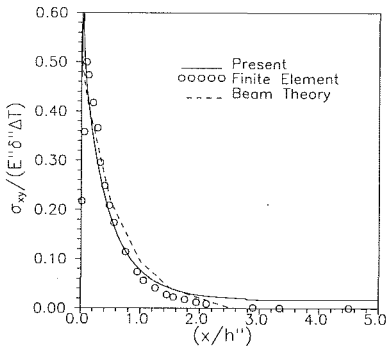


Fig. 7 Comparison of normalized σ_{xy} (x, 0)

Fig. 8 Finite element mesh

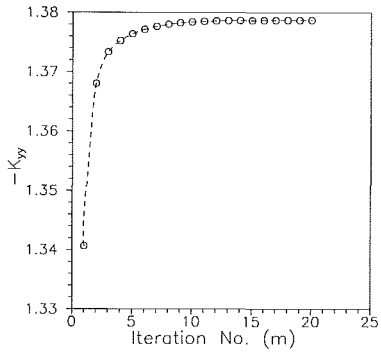


Fig. 9 Convergence study of the solution scheme

observed for K'_{xx}, K''_{xx}, and K_{xy}, which are intensity factors for stress components σ'_{xx}, σ''_{xx}, and σ_{xy}, respectively. Thus, it is concluded that the alternating method and the convergence criteria used in this paper can provide a monotonically convergent solution to the problem.

6 Discussion and Conclusions

For bimaterial thermostats with nonsingular free edges (i.e., $\alpha(\alpha-2\beta)<0$), the solution procedure described in Section 4 is still valid. However, most of the bimetallic structures of practical interest fall in the range of nonzero λ (i.e., $\alpha(\alpha-2\beta)>0$). It is found that, for a singular free edge in a bimetallic thermostat, the stress distributions at a distance of more than three times the thickness of the thinner layer is essentially the same as the nominal stresses predicted by the strength of material approach discussed in Section 2. This observation suggests that, except within a boundary layer range, which is approximately three times the thickness of the thinner layer, the classical beam theory can provide sufficiently accurate results to the problem of a bimetallic thermostat subjected to uniform heating or cooling. Within the boundary layer range, however, interlayer stresses are much higher and an appropriate method such as the alternating method discussed in this paper or finite element methods with a special element at the edge must be used to properly address the free edge effects. Results of this study also suggest that, as long as the total length is greater than six times the thickness of the thinner layer, thermal stresses in a finite length, bimetallic thermostat can be treated as the superposition of two semi-infinite, bimetallic strips. Finally, if the finite element method is used to analyze a bimetallic thermostat with singular edge

$(\alpha(\alpha-2\beta)>0)$, continued refinement of the conventional isoparametric element size at the free edge will not guarantee a convergent peak stress but a special element, which has appropriate interpolation functions built in to account for the singularity at the free edge, is needed to calculate the intensity of the singular term at the edge.

References

Bogy, D. B., 1968, "Edge-Bonded Dissimilar Orthogonal Elastic Wedges Under Normal and Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 460–466.

Bogy, D. B., 1970, "On the Problem of Edge-Bonded Elastic Quarter-Planes Loaded at the Boundary," *International Journal of Solids and Structures*, Vol. 6, pp. 1287–1313.

Boley, B. A., and Weiner, J. H., 1960, *Theory of Thermal Stresses*, John Wiley and Sons, New York.

Chen, W. T., and Nelson, C. W., 1979, "Thermal Stress in Bonded Joints," *IBM J. Res. Develop.*, Vol. 23, No. 2, pp. 179–187.

Dundurs, J., 1969, Discussion on Bogy (1968), ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 650–652.

Hein, V. L., Erdogan, F., 1971, "Stress Singularities in a Two-Material Wedge," *International Journal of Fracture Mechanics*, Vol. 7, pp. 317–330.

Suhir, E., 1986, "Stresses in Bi-Metal Thermostats," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 657–660.

Timoshenko, S. P., 1925, "Analysis of Bi-Metal Thermostats," *Journal of the Optical Society of America*, Vol. 11, pp. 233–255.

Razaqpar, A. G., 1987, Discussion on Suhir's paper (1986), ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, p. 479.

APPENDIX

Definitions of H_{jk} and R_j

$R_1=F_1 \quad R_2=F_2 \quad R_7=G_1 \quad R_8=G_2$

$R_j=0$ for all the others

$H_{11}=-H_{21}=-s^2e^{sh''} \quad H_{12}=H_{22}=-s^2e^{-sh''}$

$H_{13}=h''\ H_{11} \quad H_{14}=h''\ H_{12}$

$H_{23}=s(1+sh'')e^{sh''} \quad H_{24}=s(1-sh'')e^{-sh''}$

$H_{31}=H_{32}=-H_{35}=-H_{36}=H_{63}=H_{64}=-H_{67}=-H_{68}=1.0$

$H_{41}=H_{42}=H_{51}=-H_{52}=(1+\nu'')\delta''/E''$

$H_{43}=-H_{44}=-H_{53}=-H_{54}=2(1-\nu''^2)/E''$

$H_{45}=H_{46}=H_{55}=-H_{56}=-(1+\nu')\delta'/E'$

$-H_{47}=H_{48}=H_{57}=H_{58}=2(1-\nu'^2)/E'$

$H_{61}=-H_{62}=-H_{65}=H_{66}=s$

$H_{75}=-H_{85}=-s^2e^{-sh'} \quad H_{76}=H_{86}=-s^2e^{sh'}$

$H_{77}=-h'H_{74} \quad H_{78}=-h'\ \mathrm{H}_{75}$

$H_{87}=s(1-sh')e^{-sh'} \quad H_{88}=s(1+sh')e^{sh'}$

$H_{jk}=0$ for all the others.

Piero Villaggio
Istituto di Scienza delle Costruzioni,
56126 Pisa, Italy

How to Model a Bonded Joint

The problem is considered of a semi-infinite plane region bonded to a rigid region, with the boundary of contact being in the shape of a cosine curve. It is shown that, when a rigid displacement is applied to the boundary of the elastic region, there is a particular value of the amplitude of the contact curve that minimizes the sum of the strain energy and adhesion energy.

Introduction

In the problem of joining together two separate elastic bodies, the question arises of how to model the profile of the joint in order that it supports a prescribed load. A typical example is shown in Fig. 1, where two bodies, V and V', are bonded along their common surface Σ, while the parts S and S' are subjected to given surface tractions $\mathbf{F}$.

When the bodies are elastic and Σ is known, the stresses throughout the respective volumes may be determined (at least in principle) by solving a contact problem in linear elasticity. In particular, the stresses across Σ can be determined.

For many practical problems, however, although the surfaces S and S' are specified, there is a certain freedom in selecting the shape of Σ, provided, of course, that it is bounded by the curve of intersection between S and S'.

The strength of the connection between V and V' increases with the amount of adhesive material interposed between the two bodies along their common boundary Σ. Thus, it seems that a certain advantage can be achieved by enlarging the area of Σ by giving it a wavy shape instead of flat. But the amplitude of these waves can not be too large since, in this case, a sharp increase of stress around Σ may occur. In practice, a compromise between these two opposite requirements is obtained by joiners when they connect sheets of wood by carving extra interpenetrating teeth, which, on the other hand, are not too sharp nor too deep, since otherwise high concentrations of stresses would be produced.

The designer must therefore decide between two conflicting objectives in modeling joints: to maximize the surface area of Σ by allowing, for instance, Σ to be of oscillatory shape, and, at the same time, to keep Σ reasonably smooth to avoid the introduction of notch stresses.

In this paper a simplified model is studied, which nevertheless provides a rational criterion for the optimal shape of an oscillatory interface between two plane elastic bodies. It is assumed that the two bodies are semi-infinite and joined along a periodic curve of parametric equations $x=u-a_1\sin u$, $y=a_1\cos u$ (Fig. 2) where a_1 is a constant and u ranges on the whole real axis. It is further supposed that one of the two bodies is much more rigid than the other to the extent that it is regarded as practically indeformable. In the present case the lower body is considered as rigid.

The shape of the base of the upper body can be adjusted by varying the amplitude a_1 of the periodic oscillations.

Regarding the load conditions, two cases are considered. In the first, the wavy boundary Σ undergoes a uniform displacement in the direction of the y-axis in order to create a state of pure tension for large values of y; in the second case a uniform displacement is impressed to Σ in the x-direction generating, in this way, a state of pure shear as y becomes large.

The solutions to these two mixed boundary value problems in plane elasticity are achieved by extending a method used by

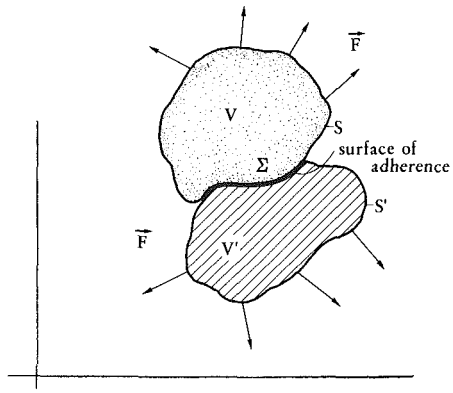


Fig. 1

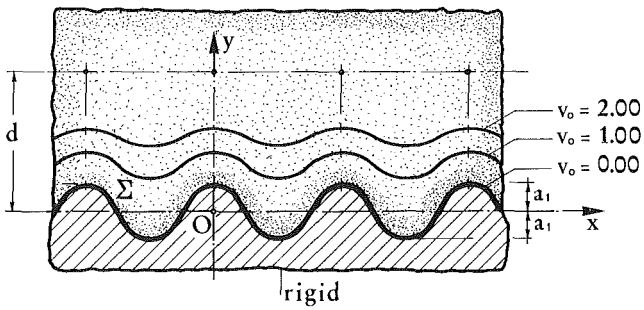


Fig. 2

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

Discussion on this paper should be addressed to the 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 26, 1988; final revision, November 9, 1988.

Weber (1942) for the prestressed half space with a periodically oscillating boundary free from surface tractions.

Under the displacement boundary conditions just described, an increase in a_1 causes a decrease of the strain energy stored within the elastic body. On the other hand, the surface area available for adhesion increases with a_1 and so does the surface energy that can be stored in the bond. It will be shown that there is a single value of a_1 which minimizes the sum of these two energies. This value of a_1 defines the amplitude which must have each wave of the curve Σ of Fig. 2 in order that the joint has the least tendency to rupture. The result may help to explain, in this particular case, empirical formulae used in design (cf., Vinson and Sierakowsky (1986)).

2 Pure Extension

If the upper body is a long cylinder with the generators parallel to the z-axis and the terminal sections remain at right angles to this axis, the state of strain can be regarded as plane, and the displacements ξ, η along the x, y-axes are functions of x, y only and the displacement ζ along z vanishes.

Let the material be elastic, homogeneous, and isotropic, with Young's modulus E and Poisson ratio σ. To simplify the formulae it is sometimes useful to introduce the constant $\alpha = 2(1-\sigma)$.

In absence of body forces, it is known that the displacements may be written in terms of two plane harmonic functions, called the Boussinesq-Papkovič-Neuber functions. More precisely, let $\Phi_0(x, y)$, $\Phi_1(x, y)$ be two harmonic functions and define the plane biharmonic function F to be

$$F = \Phi_0 + x\Phi_1, \tag{1}$$

then, the displacements assume the form (cf., Neuber, 1937, Chapter 4, No. 1),

$$2\mu\xi = -\frac{\partial F}{\partial x} + 2\alpha\Phi_1, \quad 2\mu\eta = -\frac{\partial F}{\partial y}, \tag{2}$$

where μ is the shear modulus related to E and σ by the formula $\mu = E/2(1+\sigma)$.

The stresses associated with these displacements and expressed in terms of the functions F, Φ_0, Φ_1 are given by

$$\sigma_x = -\frac{\partial^2 F}{\partial x^2} + 2\alpha\frac{\partial \Phi_1}{\partial x} + \left(1 - \frac{\alpha}{2}\right)\nabla^2 F,$$

$$\sigma_y = -\frac{\partial^2 F}{\partial y^2} + \left(1 - \frac{\alpha}{2}\right)\nabla^2 F, \quad \tau_{xy} = -\frac{\partial^2 F}{\partial x \partial y} + \alpha\frac{\partial \Phi_1}{\partial y}, \tag{3}$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y$, and the stress-strain relations

$$\sigma_x = 2\mu\left[\frac{\partial \xi}{\partial x} + \frac{\sigma}{(1-2\sigma)}\left(\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y}\right)\right],$$

$$\sigma_y = 2\mu\left[\frac{\partial \eta}{\partial y} + \frac{\sigma}{(1-2\sigma)}\left(\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y}\right)\right],$$

$$\tau_{xy} = \mu\left[\frac{\partial \xi}{\partial y} + \frac{\partial \eta}{\partial x}\right],$$

have been used.

In order to find the functions Φ_0, Φ_1 it is necessary to change independent variables and introduce curvilinear coordinates u, v such that

$$x = u - a_1 e^{-v}\sin u, \quad y = v + a_1 e^{-v}\cos u, \tag{4}$$

where a_1 is a constant strictly less than one. Under this restriction, the mapping (4) is conformal and takes harmonic functions in the u, v-plane into harmonic functions in the x, y-plane. In particular, the images of the lines $v = v_0 =$ constant in the u, v-plane are the lines $x = u - a_1 e^{-v_0}\sin u$, $y = v_0 + a_1 e^{-v_0}\cos u$ in the x, y-plane, and the image of the axis $v = 0$ is the line $x = u - a_1 \sin u$, $y = a_1 \cos u$ (Fig. 2).

The elastic body, occupying the region $v > 0$, is stressed when a rigid translation of the type $\xi = 0$, $\eta = \Delta$, with Δ constant, is applied at the boundary.

The boundary conditions are satisfied when Φ_0 and Φ_1 are given by

$$\Phi_0 = A[(u - a_1 e^{-v}\sin u)^2 - (v + a_1 e^{-v}\cos u)^2] + b_1 e^{-v}\cos u + b_2 e^{-2v}\cos 2u + b_3 v + E(u^2 - v^2),$$

$$\Phi_1 = B(u - a_1 e^{-v}\sin u) + c_1 e^{-v}\sin u, \tag{5}$$

where A, B, b_1, b_2, . . . ,c_1 are constants. It is evident that, with this choice, both Φ_0 and Φ_1 are plane harmonic functions, and the biharmonic function F assumes the form

$$F = (A+B)(u - a_1 e^{-v}\sin u)^2 - A(v + a_1 e^{-v}\cos u)^2 + b_1 e^{-v}\cos u + b_2 e^{-2v}\cos 2u + b_3 v + E(u^2 - v^2) + c_1 e^{-v}\sin u(u - a_1 e^{-v}\sin u). \tag{6}$$

To determine the unknown constants it is necessary to write displacements ξ, η in curvilinear coordinates and observe that the conditions $\xi = 0$, $\eta = \Delta$, are satisfied at $v = 0$. In terms of the new variables u, v, the partial derivative $\partial F/\partial x$ becomes (cf., Neuber, 1973, Chapter 4)

$$\frac{\partial F}{\partial x} = \frac{1}{h^2}\left(\frac{\partial F}{\partial u}\frac{\partial x}{\partial u} + \frac{\partial F}{\partial v}\frac{\partial x}{\partial v}\right), \tag{7}$$

where $h^2 = (\partial x/\partial u)^2 + (\partial x/\partial v)^2 = 1 - 2\ a_1 e^{-v}\cos\ u + a_1^2 e^{-2v}$, and an analogous formula holds for $\partial F/\partial y$, with y instead of x.

Thus, the boundary conditions that $\xi = 0$, $\eta = \Delta$ for $v = 0$, written in terms of F and Φ_1, are equivalent to the following pair of equations

$$2\mu U(u,o) = \frac{1}{h(u,o)}\left[-\frac{\partial F}{\partial u}(u,o) + 2\alpha\Phi_1(u,o)\frac{\partial x}{\partial u}(u,o)\right]$$

$$= \frac{-1}{h(u,o)}\left[2(A+B)(u - a_1\sin u)(1 - a_1\cos u) - 2Aa_1^2\sin u\cos u - b_1\sin u - 2b_2\sin 2u + c_1\cos u(u - a_1\sin u) + c_1\sin u(1 - a_1\cos u) + 2\alpha\Phi_1(u,o)\frac{\partial x}{\partial u}(u,o)\right] = \frac{2\mu\Delta}{h(u,o)}\frac{\partial y}{\partial u}(u,o),$$

$$2\mu V(u,o) = \frac{1}{h(u,o)}\left[-\frac{\partial F}{\partial v}(u,o) + 2\alpha\Phi_1(u,o)\frac{\partial x}{\partial v}(u,o)\right]$$

$$= \frac{-1}{h(u,o)}\left[2(A+B)(u - a_1\sin u)(a_1\sin u) - 2Aa_1\cos u(1 - \cos u) - b_1\cos u - 2b_2\cos 2u + b_3 - c_1\sin u(u - a_1\sin u) + c_1 a_1\sin^2 u + 2\alpha\Phi_1(u,o)\frac{\partial x}{\partial v}(u,o)\right]$$

$$= \frac{2\mu\Delta}{h(u,o)}\frac{\partial y}{\partial v}(u,o).$$

These equations yield seven linear independent equations, from which it is possible to obtain the constants defining the function F in terms of the single constant A. The result of the process of elimination of the constants is

$$B=\frac{(2\alpha-1)A}{2\alpha(\alpha-1)}\,,\quad b_3=-\frac{Aa_1^2}{2\alpha}-2\mu\Delta,$$

$$b_1=-2Aa_1-2\mu\Delta a_1,\quad b_2=Aa_1^2\left(\alpha-\frac{1}{4\alpha}\right),$$

$$E=-\frac{A}{2\alpha}\,,\quad c_1=\frac{Aa_1}{\alpha}\,. \tag{9}$$

From these values of the constants the stress function F and, hence, the full solution is determined to within the constant A. The solution, however, can be further simplified by observing that, in practice, the amplitude a_1 is not only strictly less than one, but even a small fraction of the half-period π. In consideration of this fact the terms containing the powers of second order in a_1 can be disregarded.

The constant A remains indeterminate since the solution to the boundary value problem is not uniquely defined until the displacement is prescribed at least on another point, not lying on the boundary. In the present case it will be assumed that the point with coordinates $u=0$, $v=d$, with $d\gg 1$, remains fixed. Because $u=0$, is on the y-axis, the U-component of displacement is zero by symmetry and, therefore, the only significant condition concerns the vanishing of the V-component of displacement:

$$2\mu V(o,d)=\frac{1}{h(o,d)}\left[-\frac{\partial F}{\partial v}(o,d)+2\alpha\Phi_1(o,d)\frac{\partial x}{\partial v}(o,d)\right]$$

$$=\frac{-1}{h(u,o)}\left[-2A(d+a_1e^{-d}(1-a_1e^{-d})-b_1e^{-d}-2b_2e^{-2d}\right.$$

$$\left.-2Ed+b_3\right]=0. \tag{10}$$

The coefficients b_1, b_2, b_3, E may be expressed in terms of A by means of formulae (9) and a substitution into (10) leads to an equation for A alone. On neglecting higher order terms in a_1 and d^{-1}, this equation reduces to

$$2Ad+\frac{A}{\alpha}\,d-2\mu\Delta=0,$$

which yields for A the surprisingly simple form

$$A=\frac{2\mu\Delta\alpha}{(2\alpha+1)d}\,. \tag{11}$$

Once the functions Φ_0 and Φ_1 are known, the state of stress is also determinable by using the formulae (3). But, in order to calculate the strain energy stored in each strip $0<u<2\pi$, $v>0$, the detailed expression of the single components of stress are not necessary, since it can be directly determined from the resultant of the tractions parallel to the y-axis acting upon the portion of lower boundary of the strip $0<u<2\pi$. If Y is this resultant, the strain energy is simple given by $1/2\ Y\Delta$, as a consequence of Clapeyron's theorem.

The easiest way to find Y is to apply a result of Neuber (1937, Chapter 4, No. 1), which modifies the expression (3) of stress components. In fact, by introducing two new harmonic functions Φ_0', Φ_1' such that

$$\Phi_1=\frac{\partial\Phi_1'}{\partial x}\,,\quad \Phi_0=\alpha\Phi_1'+\Phi_0'\,, \tag{12}$$

and defining

$$F'=\Phi_0'+x\frac{\partial\Phi_1'}{\partial x}\,, \tag{13}$$

the stress assumes the simpler form

$$\sigma_x=\frac{\partial^2F'}{\partial y^2}\,,\quad \sigma_y=\frac{\partial^2F'}{\partial x^2}\,,\quad \tau_{x,y}=-\frac{\partial^2F'}{\partial x\partial y}\,. \tag{14}$$

The construction of Φ_1' is almost immediate if, in the second of (5), the term with c_1 is neglected, and only then with B is taken into account. Under this approximation, Φ_1 becomes Bx and Φ_1' can be written

$$\Phi_1'=\frac{1}{2}B(x^2-y^2)=\frac{1}{2}B\left[(u-a_1e^{-v}\sin u)^2\right.$$

$$\left.+(v+a_1e^{-v}\cos u)^2\right], \tag{15}$$

and, hence, F' is

$$F'=\left(A+\frac{2-\alpha}{2}B\right)(u-a_1e^{-v}\sin u)^2$$

$$-\left(A-\frac{\alpha}{2}B\right)(v+a_1e^{-v}\cos u)^2$$

$$+b_1e^{-v}\cos u+b_2e^{-2v}\cos 2u+b_3v+E(u^2-v^2). \tag{16}$$

Given F', by a known transformation of surface tractions in plane elasticity (cf., Love, 1924, Article 154), the resultant Y is then expressed by the simple formula

$$Y=\frac{\partial F'}{\partial x}(2\pi,o)-\frac{\partial F'}{\partial x}(o,o). \tag{17}$$

Here again, by using (7), Y takes the form

$$Y=4\pi\left(A+\frac{2-\alpha}{2}B\right)+4\pi\frac{E}{(1-\alpha_1)}\,, \tag{18}$$

whence the expression for the strain energy $W=1/2\ Y\Delta$ can be derived explicitly.

3 Simple Shear

The calculation of stress when a rigid translation Δ parallel to the axis of x is applied to the boundary $v=0$ of the upper body is quite similar. The appropriate expressions for the functions Φ_0 and Φ_1 are now

$$\Phi_0=A(u-a_1e^{-v}\sin u)(v+a_1e^{-v}\cos u)$$

$$+b_0u+b_1e^{-v}\sin u+b_2e^{-2v}\sin 2u+Euv,$$

$$\Phi_1=B(v+a_1e^{-v}\cos u)+c_1e^{-v}\cos u, \tag{19}$$

where A, B, E, b_0, . . . , c_1 are new constants. The biharmonic function F is then

$$F=(A+B)(u-a_1e^{-v}\sin u)(v+a_1e^{-v}\cos u)$$

$$+b_0u+b_1e^{-v}\sin u+b_2e^{-2v}\sin 2u$$

$$+Euv+c_1e^{-v}\cos u(u-a_1e^{-v}\sin u), \tag{20}$$

and the conditions that $\xi=\Delta$, $\eta=0$ for $v=0$ yield the two equations of the type

$$2\mu U(u,o)=\frac{1}{h(u,o)}\left[-(A+B)(1-a_1\cos u)\,a_1\cos u\right.$$

$$-(A+B)(u-a_1\sin u)$$

$$\times(-a_1\sin u)-b_0-b_1\cos u-2b_2\cos 2u$$

$$\left.-c_1\cos u(1-a_1\cos u)\right.$$

$$+c_1\sin u(1-a_1\sin u)+2\alpha\Phi_1(u,o)\frac{\partial x}{\partial u}(u,o)\Big]$$

$$=\frac{2\mu\Delta}{h(u,o)}\frac{\partial x}{\partial u}(u,o),$$

$$2\mu V(u,o)=\frac{1}{h(u,o)}\Big[-(A+B)a_1^2\sin u\cos u$$
$$-(A+B)(u-a_1\sin u)(1-a_1\cos u)+b_1\sin u$$
$$+2b_2\sin 2u-Eu-c_1(\cos u)a_1\sin u$$
$$+c_1\cos u(u-a_1\sin u)+2\alpha\Phi_1(u,o)\frac{\partial y}{\partial u}(u,o)\Big]$$

$$=\frac{2\mu\Delta}{h(u,o)}\frac{\partial x}{\partial v}(u,o). \tag{21}$$

These equations are satisfied for each u when the constants obey the following relations

$$c_1=-\frac{2\alpha a_1}{(2\alpha-1)}B,\quad b_2=\frac{\alpha a_1^2}{2(2\alpha-1)}B,$$

$$A=\frac{B}{(2\alpha-1)},\quad b_1=-\frac{2\alpha a_1}{(2\alpha-1)}B+2\mu a_1\Delta,$$

$$E=-\frac{2\alpha}{(2\alpha-1)}B,\quad b_0=\frac{\alpha a_1^2}{(2\alpha-1)}B-2\mu\Delta. \tag{22}$$

The constant B is then determined by requiring that the point (o, d) with $d>>1$, remains fixed. This condition implies that

$$2\mu U(o,d)=\frac{1}{h(o,d)}\Big[-(A+B)(1-a_1e^{-d})(d+a_1e^{-d})$$
$$-b_o-b_1e^{-d}-2b_2e^{-2d}-Ed$$
$$-c_1e^{-d}(1-a_1e^{-d})+2\alpha\Phi_1(o,d)\frac{\partial x}{\partial u}(o,d)\Big]=0 \tag{23}$$

By expressing the other constants in terms of B through (22) and omitting the small terms, it is found that

$$B=-\frac{2\mu\Delta}{2\alpha d}. \tag{24}$$

Once the harmonic functions Φ_0 and Φ_1 are determined, the stress components are given by (3), but a form like (14) is obtainable by introducing the function

$$\Phi_1'=Bxy=B(u-a_1e^{-v}\sin u)(v+a_1e^{-v}\cos u), \tag{25}$$

which satisfies the relation $\Phi_1=\partial\Phi_1'/\partial x$, provided that the coefficient c_1 in the second of (19) is omitted compared with B.

Then, by a formula like (13), the new stress function F' may be written as

$$F'=\Big[A+(\alpha-1)B\Big](u-a_1e^{-v}\sin u)(v+a_1e^{-v}\cos u)$$
$$+b_0u+b_1e^{-v}\sin u+b_2e^{-2v}\sin 2u+Euv. \tag{26}$$

Given F' the resultant of surface tractions, parallel to the x-axis extended along the lower boundary of the strip $0<u<2\pi$, is simply

$$X=-\frac{\partial F'}{\partial y}(2\pi,o)+\frac{\partial F'}{\partial y}(o,o), \tag{27}$$

and consequently, using the expression of F', X becomes

$$X=-2\pi\Big[A-(\alpha-1)B\Big]-2\pi\frac{E}{1-a_1}, \tag{28}$$

and the strain energy is now $W=1/2\ X\Delta$.

4 The Effect of the Adhesive

After having calculated the strain energy stored in the elastic body under pure extension and pure shear, it is now necessary to evaluate the surface energy localized on the welding line Σ, where the adhesive is interposed. The simplest expression for this surface energy derives from the assumption that it is proportional to the length of the line of contact. On denoting by γ the surface energy per unit length, the energy of that part of the boundary $0<u<2\pi$, $v=0$ is given by

$$U=\gamma\int_0^{2\pi}\sqrt{\Big(\frac{\partial x}{\partial u}(u,o)\Big)^2+\Big(\frac{\partial y}{\partial u}(u,o)\Big)^2}\,du. \tag{29}$$

On recalling that the integrand in (29) is just $h(u,0)=\sqrt{1-2\alpha_1\cos u+a_1^2}$, the value of U is explicitly computable (cf., Gradshteyn and Ryzhik, 1965, 2.576)

$$v=4\gamma(1+a_1)E(k) \tag{30}$$

where

$$k=\sqrt{\frac{4a_1}{(1+a_1)^2}}$$

and $E(k)$ is the complete elliptic integral of the second kind. Thus, v increases with a_1 and so a higher amount of surface energy may be obtained by increasing the amplitude of the oscillations. Again, since a_1 is small, the expression of (30) may be simplified by retaining only the first two terms in the power series expansion of $E(k)$ (cf., Jahnke-Emde-Lösch, 1966, V,C)

$$\frac{2}{\pi}E(k)=1-2\frac{k^2}{8}-3\Big(\frac{k^2}{8}\Big)^2-\cdots\cdots$$

To this order of approximation, U assumes the form

$$U\cong 2\pi\gamma\Big[(1+a_1)-\frac{a_1}{(1+a_1)}\Big]=2\pi\gamma\Big(1+\frac{a_1^2}{(1+a_1)}\Big). \tag{31}$$

Once the strain energy W and the surface energy U have been determined as functions of a_1, the optimal shape of the surface is given by the value of a_1 minimizing the sum $(W+U)$.

In the case of pure tension, still assuming a_1 to be small and d large, the optimum value of a_1 is the first positive root of the equation

$$\frac{\partial}{\partial a_1}(W+U)=0 \tag{32}$$

where $W=1/2\ Y\Delta$ and U is given by (31).

After a substitution from (9) and differentiation with respect to a_1, equation (32) is found to be equivalent to the following:

$$2\pi\frac{E\Delta}{(1-a_1)^2}+4\pi\frac{\gamma a_1}{(1+a_1)}-2\pi\frac{\gamma a_1^2}{(1+a_1)^2}=0,$$

which, considering only the terms of first-order in a_1, admits the root

$$a_1=-\frac{E\Delta}{2\gamma}, \tag{33}$$

or better, recalling that $E=-A/2\alpha$ with $A=2\mu\alpha\Delta/(2\alpha+1)d$,

$$a_1=\frac{\mu\Delta^2}{2(\alpha+1)\gamma d}. \tag{34}$$

In the case of pure shear, an equation like (32) must be solved with respect to a_1, with the only difference that now $W=\frac{1}{2}X\Delta$, while U remains unchanged. The resulting equation in a_1 is then

$$\pi\frac{E\Delta}{(1-a_1)^2}+4\pi\frac{\gamma a_1}{(1+a_1)}-2\pi\frac{\gamma a_1^2}{(1+a_1)^2}=0,$$

which, recalling that $E=2\alpha B/(2\alpha-1)$ with $B=-\mu\Delta/\alpha d$, and neglecting the terms of higher order in a_1, yields

$$a_1=\frac{\mu\Delta^2}{2(2\alpha-1)\gamma d}\,. \tag{35}$$

Formulae (34) and (35) show that once the constitutive properties of the upper body are given through μ and γ, and the magnitude of the load is specified by Δ and d, there is a unique value of a_1, minimizing the total energy. The minimum in pure tension is lower than the minimum in simple shear.

References

Gradshteyn, I. S., and Ryzhik, I. M., 1963, *Tables of Integrals Series and Products*, Academic Press, New York.

Jahnke, E., Emde, F., and Lösch, F., 1966, *Tafeln Höherer Funktionen*, Teubner, Stuttgart.

Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., Cambridge University Press.

Neuber, H., 1937, *Kerbspannungslehre*, Springer-Verlag, Berlin.

Vinson, J. R., and Sierakowski, R. L., 1986, *The Behaviour of Structures Composed of Composite Materials*, Martinus Nijhoff, Dordrecht.

Weber, C., 1942, "Halbebene mit periodisch gewelltem Rand," *Z. Angew. Math. Mech.*, Vol. 22, No. 1, pp. 29–33.

Interfacial Stresses in Bimetal Thermostats

E. Suhir
AT&T Bell Laboratories,
Murray Hill, N. J. 07974
Mem. ASME

The magnitude and the distribution of the interfacial stresses in thermostat-like structures are determined on the basis of an elementary beam theory, with consideration of both the longitudinal and the transverse ("through-thickness") interfacial compliances of the thermostat strips. The suggested approach is applicable, generally speaking, to any elongated lap shear assembly, subjected to thermal or external loading.

Introduction

A bimetal thermostat is a useful theoretical model for the analysis of stresses in thermally-mismatched structures. The mechanical behavior of bimetal thermostats was apparently first examined by Timoshenko (1925) on the basis of an elementary beam theory. In Timoshenko's analysis, however, only the normal stresses in the thermostat strips were evaluated using an assumption that these stresses remain unchanged along the strips. As to the interfacial stresses, it was just mentioned that they are of "local" type and concentrate near the strip ends at the distances comparable with the strip thicknesses. At the same time, it should be emphasized that while the normal stresses in the thermostat strips themselves are responsible for the ultimate and fatigue strength of the strips, it is the interfacial shearing and peeling stresses which are responsible for the structural integrity of the thermostat.

Valuable insight into the thermally-induced stresses in heterogeneous structures, including interfacial stresses and the "edge problem," were later provided by Aleck (1949), Bogy (1968, 1970), Hess (1969), Zayfang (1970), Chang (1981), Chen et al. (1982), Wang and Choi (1982), Zwiers et al. (1982), Blech and Kantor (1983), and others on the basis of the theory of elasticity methods. The obtained solutions are, however, so complex that a substantial additional effort seems to be needed to make their utilization convenient in engineering applications.

Therefore, various simplified approaches to the problem in question were suggested during the last decade, mostly in connection with the needs of the microelectronics technology. These approaches, in one way or another, extend the original Timoshenko theory and are based primarily on the elementary methods of strength of materials and structural mechanics, rather than methods of the theory of elasticity. Examples are Grimado (1978), Chen and Nelson (1979), and Suhir (1986). The utilization of the concept of the finite longitudinal interfacial compliance enabled the last author to satisfy the stress-free boundary conditions for the normal stresses in the strip, and to obtain simple formulas for the interfacial stresses.

In the forthcoming analysis we suggest a somewhat more complicated analytical model for evaluation of the interfacial stresses in bimetal thermostats. This model is also based on an elementary beam (or long-and-narrow plate) theory, but considers, in addition to the longitudinal, also the transverse ("through-thickness") interfacial compliance. The obtained solutions satisfy the boundary conditions for the interfacial shearing stresses, as well as for the normal stresses in the thermostat strips themselves.

Note that since the mid-1960s many investigators applied numerical, mainly finite element, methods to analyze bimaterial structures, subjected to thermal loading: Saganuma et al. (1984), Blanchard and Watson (1986), Gerstle and Chambers (1987), and others. Although these methods offer high flexibility in dealing with different geometries and material properties, they could be less appealing than sufficiently simple analytical solutions.

Analysis

Basic Equations. The interfacial shearing $\tau(x)$ and peeling $p(x)$ stresses, which occur in an elongated bimaterial plate experiencing temperature change, are related by the following equilibrium equation (Fig. 1):

$$-\int_{-l}^{x}\int_{-l}^{\xi} p(\xi')d\xi d\xi = D_1 w_1''(x) - \frac{h_1}{2}T(x) = -D_2 w_2''(x) + \frac{h_2}{2}T(x). \quad (1)$$

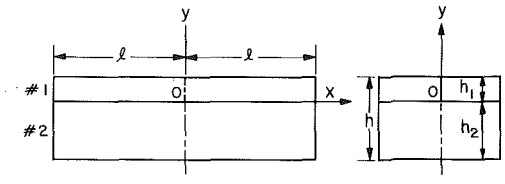


Fig. 1

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, Dallas, Tex., November 25–30, 1990.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 7, 1988; final revision, November 14, 1988.
Paper No 90-WA/APM-1.

Here

$$T(x)=\int_{-l}^{x}\tau(\xi)d\xi \tag{2}$$

is the shearing force caused by the thermal expansion or contraction mismatch of the materials, $w_1(x)$ and $w_2(x)$ are lateral deflections of the strips, which are treated here as elongated rectangular plates,

$$D_1=\frac{E_1h_1^3}{12(1-\nu_1^2)},\quad D_2=\frac{E_2h_2^3}{12(1-\nu_2^2)}$$

are flexural rigidities of these plates, E_1 and E_2 are Young's moduli of the materials, ν_1 and ν_2 are Poisson's ratios, h_1 and h_2 are the strip thicknesses, and l is half the thermostat's length. The origin of the rectangular coordinates x, y is in the middle of the plate on its interface. The relationship (1) simply states that the bending moments

$$M_1(x)=\int_l^x\int_l^\xi p(\xi')d\xi'd\xi'+\frac{h_1}{2}T(x)$$

$$M_2(x)=\int_l^x\int_l^\xi p(\xi')d\xi'd\xi'+\frac{h_2}{2}T(x)$$

acting over the cross-sections of the strips and caused by the stresses $\tau(x)$ and $p(x)$ are equilibrated by the elastic moments $D_1w_1''(x)$ and $D_2w_2''(x)$.

By differentiating the relationship (1) twice, we obtain:

$$p(x)=-D_1w_1^{IV}(x)+\frac{h_1}{2}\tau'(x)=D_2w_2^{IV}(x)-\frac{h_2}{2}\tau'(x). \tag{3}$$

On the other hand, the peeling stress $p(x)$ can be represented as:

$$p(x)=K[w_1(x)-w_2(x)], \tag{4}$$

where K is the through-thickness spring constant. This constant is due to the transverse compliance of the thermostat strips and can be assessed on the basis of the following elementary considerations. The lateral displacements in a long-and-narrow strip due to a transverse load $p(x)$ can be evaluated by the formula $w(x)=(1-\nu^2)h/Ep(x)$, which can be obtained using the Ribière solution for a long-and-narrow strip (see, for instance, Timoshenko and Goodier, 1970). Treating the longitudinal cross-sections of the thermostat plates as long and narrow strips, we obtain the following formula for the total spring constant in the through-thickness direction:

$$K=\left[(1-\nu_1^2)\frac{h_1}{E_1}+(1-\nu_2^2)\frac{h_2}{E_2}\right]^{-1}.$$

From (3) and (4) we have:

$$\left.\begin{aligned} w_1^{IV}(x) &= \frac{D_2}{KD}p^{IV}(x)+\frac{h}{2D}\tau'(x)\\ w_2^{IV}(x) &= -\frac{D_1}{KD}p^{IV}(x)+\frac{h}{2D}\tau'(x)\end{aligned}\right\}, \tag{5}$$

where $D=D_1+D_2$ is the total flexural rigidity and $h=h_1+h_2$ is the total thickness of the thermostat. Then equations (3) result in the following relationship:

$$p^{IV}(x)+4\alpha^4p(x)=4\mu\alpha^4\tau'(x). \tag{6}$$

Here, the parameters μ and α are expressed by the formulas:

$$\mu=\frac{h_1D_2-h_2D_1}{2D},\quad \alpha=\sqrt[4]{\frac{KD}{4D_1D_2}}.$$

Note that the peeling stress $p(x)$ is zero if the combination of the material properties is such that the ratio h_1/h_2 of the strip thicknesses is equal to the ratio D_1/D_2 of the flexural rigidities.

The second equation for the functions $\tau(x)$ and $p(x)$ can be obtained using the condition of compatibility for the longitudinal displacements $u_1(x)$ and $u_2(x)$ of the extreme (interfacial) fibers of the thermostat strips. If the stresses $\tau(x)$ and $p(x)$ were known, then these displacements could be evaluated by the formulas:

$$\left.\begin{aligned} u_1(x) &= \alpha_1\Delta tx-\lambda_1\int_{-l}^{x}T(\xi)d\xi+\kappa_1\tau(x)-\frac{h_1}{2}w_1'(x)\\ u_2(x) &= \alpha_2\Delta tx+\lambda_2\int_{-l}^{x}T(\xi)d\xi-\kappa_2\tau(x)+\frac{h_2}{2}w_2'(x)\end{aligned}\right\}, \tag{7}$$

where α_1 and α_2 are coefficients of thermal expansion of the materials, $\lambda_1=(1-\nu_1^2)/E_1h_1$, and $\lambda_2=(1-\nu_2^2)/E_2h_2$ are in-plane compliances of the strips, $\kappa_1=2(1+\nu_1)h_1/3E_1$ and $\kappa_2=2(1+\nu_2)h_2/3E_2$ are their interfacial compliances (Suhir, 1986), and Δt is the temperature change. The first terms in (7) are unrestricted thermal expansions. The second terms are due to the thermal mismatch forces $T(x)$ and reflect an assumption that these forces are uniformly distributed over the strip thicknesses. The third terms account for the nonuniform distribution of the aforementioned forces, and are based on an assumption that the corresponding corrections can be evaluated by taking into account only the shearing stresses in the given cross-section. The last terms are due to bending.

The compatibility condition $u_1(x)=u_2(x)$ results in the equation:

$$\kappa\tau(x)-\lambda_{12}\int_0^xT(\xi)d\xi=\Delta\alpha\Delta tx+\frac{h_1}{2}w_1'(x)+\frac{h_2}{2}w_2'(x), \tag{8}$$

where $\lambda_{12}=\lambda_1+\lambda_2$ and $\kappa=\kappa_1+\kappa_2$ are the total in-plane compliance and the total interfacial compliance, respectively, and $\Delta\alpha=\alpha_2-\alpha_1$. By differentiating (8), with consideration of (2), we find:

$$\kappa\tau'(x)-\lambda_{12}T(x)=\Delta\alpha\Delta t+\frac{h_1}{2}w_1''(x)+\frac{h_2}{2}w_2''(x), \tag{9}$$

$$\kappa\tau''(x)-\lambda_{12}\tau(x)=\frac{h_1}{2}w_1'''(x)+w_2'''(x). \tag{10}$$

Using the formulas (5), equation (10) can be presented in the form:

$$\tau''(x)-k^2\tau(x)=mp'''(x), \tag{11}$$

where the following notation is used:

$$k=\sqrt{\frac{\lambda}{\kappa}},\quad \lambda=\lambda_{12}+\frac{h^2}{4D}=\lambda_1+\lambda_2+\frac{h^2}{4D},\quad m=\frac{\mu}{K\kappa}.$$

Equations (6) and (11) form a system from which the stress functions $\tau(x)$ and $p(x)$ can be determined. In order to separate these functions, we rewrite (11) in the form

$$p^{IV}(x)=\frac{1}{m}[\tau'''(x)-k^2\tau'(x)]$$

and substitute this equation into (6). Then we have:

$$p(x)=\frac{1}{4\alpha^4m}[(1+\epsilon)k^2\tau'(x)-\tau'''(x)]. \tag{12}$$

Here

$$\epsilon=\frac{D}{D_1D_2}\frac{\mu^2}{\lambda}=\frac{3}{1+\left(\dfrac{2h\sqrt{D_1D_2}}{h_1D_2-h_2D_1}\right)^2}$$

is the dimensionless parameter of the peeling stress. This parameter changes from $\epsilon=0$, when $h_1/h_2=D_1/D_2$, to $\epsilon=3$, when $D_2\gg D_1$ or $D_1\gg D_2$. Now substituting (12) into (11), we obtain the following homogeneous differential equation for shearing stress function $\tau(x)$:

$$\tau^{VI}(x)-(1+\epsilon)k^2\tau^{IV}(x)+4\alpha^4\tau''(x)-4\alpha^4k^2\tau(x)=0. \quad (13)$$

By introducing the expression for the derivative $\tau'(x)$ from (6) into (12), one can obtain an identical equation for the peeling stress function $p(x)$:

$$p^{VI}(x)-(1+\epsilon)k^2p^{IV}(x)+4\alpha^4p''(x)-4\alpha^4k^2p(x)=0. \quad (14)$$

Thus, both the shearing and the peeling stresses can be determined on the basis of the same ordinary homogeneous differential equation of the sixth order.

The equation for the shearing force $T(x)$ can be obtained, using (13) and (2), in the form:

$$T^{VII}(x)-(1+\epsilon)k^2T^{V}(x)+4\alpha^4T'''(x)-4\alpha^4k^2T'(x)=0. \quad (15)$$

Boundary Conditions. The function $\tau(x)$ must be antisymmetric with respect to the origin. In our analysis we assume that this function satisfies the following boundary conditions at the edge $x=l$:

$$\tau(l)=0, \quad \tau'(l)=\frac{\Delta\alpha\Delta t}{\kappa}, \quad \tau''(l)=0. \quad (16)$$

The first of the conditions (16) follows from the equilibrium equation (1), which could be, by differentiation, presented in the form:

$$\int_{-l}^{x}p(\xi)d\xi=D_1w_1'''(x)-\frac{h_1}{2}\tau(x)=-D_2w_2'''(x)+\frac{h_2}{2}\tau(x).$$

The peeling stress $p(x)$ must be self-equilibrated, and therefore the left part of this equation becomes zero for $x=l$. The elastic terms are also zero at $x=l$, since there are no concentrated lateral forces at this cross-section and, therefore, $w_2'''(l)=w_2'''(l)=0$. This leads to the condition $\tau(l)=0$. The second condition in (16) can be obtained from (9) using the facts that there are no external forces at the free edges, i.e., $T(l)=0$, and that there are no bending moments in the cross-section $x=l$, so that $w_1''(l)=w_2''(l)$. The third condition follows from (10). Indeed, since $\tau(l)=0$ and, in addition, $w_1'''(l)=w_2'''(l)=0$, then the second derivative of the shearing stress function must also be zero at this cross-section.

It is important to mention, however, that elastic analyses (see, for instance, Bogy, 1968, 1970; Hein and Erdogan, 1971) indicate that stress singularities generally occur at the corners of geometric boundaries joining dissimilar materials, and that the interfacial shearing stress may become unbounded at the corner, depending on the particular material combination and remote loading. The assumption $\tau(l)=0$ seems, nonetheless, consistent with the elementary approach taken in the present study.

The peeling stress $p(x)$ must be symmetric with respect to the origin and should satisfy the equilibrium conditions for the lateral forces and bending moments:

$$\int_{-l}^{l}p(x)dx=0, \quad \int_{-l}^{l}\int_{-l}^{x}p(\xi)d\xi dx=0.$$

Obviously, these conditions are fulfilled automatically as long as the conditions $T(l)=0$ and $\tau(l)=0$ are satisfied.

The function $T(x)$ must also be symmetric with respect to the origin and should satisfy the conditions:

$$T(l)=0, \quad T'(l)=0, \quad T''(l)=\frac{\Delta\alpha\Delta t}{\kappa}, \quad T'''(l)=0.$$

The first condition reflects the fact that no external forces act on the free edge. The remaining three conditions follow from the conditions (16) and the relationship (2).

Solutions. In order to obtain solutions to the differential equations (13), (14), and (15), we form a characteristic equation:

$$\beta^6-(1+\epsilon)k^2\beta^4+4\alpha^4\beta^2-4\alpha^4k^2=0. \quad (17)$$

By introducing a new unknown

$$\zeta=\beta^2-\frac{1+\epsilon}{3}k^2,$$

this equation can be reduced to a simple cubic equation:

$$\zeta^3+3q_1\zeta+2q_0=0. \quad (18)$$

Here

$$q_0=-\frac{2}{3}(2-\epsilon)\alpha^4k^2-\left(\frac{1+\epsilon}{3}k^2\right)^3,$$

$$q_1=\frac{4}{3}\alpha^4+\left(\frac{1+\epsilon}{3}k^2\right)^2.$$

The analysis of the discriminant $d=q_0^2+q_1^3$ of the equation (18) indicates that the d value is always positive, and therefore this equation has one real and two conjugate complex roots:

$$\zeta_1=a+b, \quad \zeta_{2,3}=-\frac{a+b}{2}\pm i\sqrt{3}\frac{a-b}{2}.$$

Here $i=\sqrt{-1}$, and

$$a=\sqrt[3]{-q_0+\sqrt{d}}, \quad b=\sqrt[3]{-q_0-\sqrt{d}}.$$

Accordingly, the characteristic equation (17) has the following roots:

$$\left.\begin{aligned}\beta_{1,2}&=\pm\sqrt{\zeta_1+\frac{1+\epsilon}{3}k^2}\\ \beta_{3,4}&=\pm\sqrt{\zeta_2+\frac{1+\epsilon}{3}k^2}=\pm\gamma_1\pm i\gamma_2\\ \beta_{5,6}&=\pm\sqrt{\zeta_3+\frac{1+\epsilon}{3}k^2}=\pm\gamma_1\mp i\gamma_2\end{aligned}\right\},$$

where

$$\gamma_1=\sqrt{\frac{1}{2}\left(\sqrt{\mu_1^2+\mu_2^2}+\mu_1\right)}, \quad \gamma_2=\sqrt{\frac{1}{2}\left(\sqrt{\mu_1^2+\mu_2^2}-\mu_1\right)},$$

and

$$\mu_1=-\frac{a+b}{2}+\frac{1+\epsilon}{3}k^2, \quad \mu_2=\sqrt{3}\frac{a-b}{2}.$$

Thus, the solutions to the equations (13), (14), and (15) can be presented in the form:

$$\tau(x)=C_1\sinh\beta_1x+C_3\cosh\gamma_1x\sin\gamma_2x+C_5\sinh\gamma_1x\cos\gamma_2x, \quad (19)$$

$$p(x)=C_2\cosh\beta_1x+C_4\cosh\gamma_1x\cos\gamma_2x+C_6\sinh\gamma_1x\sin\gamma_2x, \quad (20)$$

$$T(x)=A_2\cosh\beta_1x+A_4\cosh\gamma_1x\cos\gamma_2x+A_6\sinh\gamma_1x\sin\gamma_2x+C_0, \quad (21)$$

where the antisymmetry of the function $\tau(x)$ and the symmetry of the functions $p(x)$ and $T(x)$ have been taken into consideration.

Constants of Integration. The relationship between the constants with the odd indices in (19), on the one hand, and the constants with the even indices in (20), on the other hand, can be obtained by substituting (19) and (20) into (6):

$$C_2=\frac{4\mu\alpha^4\beta_1}{\beta_1^4+4\alpha^4}\,C_1,\quad C_4=4\mu\alpha^4(\delta_1C_3+\delta_2C_5),$$
$$C_6=4\mu\alpha^4(\delta_2C_3-\delta_1C_5). \tag{22}$$

Here

$$\delta_1=\frac{f_1\gamma_2-f_2\gamma_1}{f_1^2+f_2^2},\quad \delta_2=\frac{f_1\gamma_1+f_2\gamma_2}{f_1^2+f_2^2},$$
$$f_1=\mu_1^2-\mu_2^2+4\alpha^4,\quad f_2=2\mu_1\mu_2.$$

The relationships between the constants C_1, C_3, and C_5 and the constants A_2, A_4, and A_6 can be found on the basis of the equation $T'(x)=\tau(x)$, which follows from (2). After substituting (19) and (21) in this equation we have:

$$A_2=\frac{C_1}{\beta_1},\quad A_4=\frac{\gamma_1C_5-\gamma_2C_3}{\gamma_1^2+\gamma_2^2},\quad A_6=\frac{\gamma_1C_3+\gamma_2C_5}{\gamma_1^2+\gamma_2^2}. \tag{23}$$

The constant C_0 can be determined using the condition $T(l)=0$, which results in the formula:

$$C_0=-A_2\cosh u-A_4\cosh v_1\cos v_2-A_6\sinh v_1\sin v_2. \tag{24}$$

The constants C_1, C_3, and C_5 in (19) can be found using the boundary conditions (16). These result in the following system of linear algebraic equations:

$$\left.\begin{aligned}
&C_1\sinh u+C_3\cosh v_1\sin v_2+C_5\sinh v_1\cos v_2=0\\
&C_1\beta_1\cosh u+C_3(\gamma_1\sinh v_1\sin v_2+\gamma_2\cosh v_1\cos v_2)+\\
&+C_5(\gamma_1\cosh v_1\cos v_2-\gamma_2\sinh v_1\sin v_2)=\frac{\Delta\alpha\Delta t}{\kappa}\\
&C_1\beta_1^2\sinh u+C_3(\mu_1\cosh v_1\sin v_2+\mu_2\sinh v_1\cos v_2)+\\
&+C_5(\mu_1\sinh v_1\cos v_2-\mu_2\cosh v_1\sin v_2)=0
\end{aligned}\right\}. \tag{25}$$

Here, $u=\beta_1 l$, $v_1=\gamma_1 l$, $v_2=\gamma_2 l$. Then we have:

$$\left.\begin{aligned}
C_1&=-\frac{\Delta\alpha\Delta t}{\kappa D_0}\left(\operatorname{cotanh}2v_1-\frac{\cos 2v_2}{\sinh 2v_1}\right)\frac{1}{\cosh u}\\
C_3&=\frac{\Delta\alpha\Delta t}{\kappa D_0}\left(\frac{\sin v_2}{\sinh v_1}+\frac{3}{2}\,\frac{\zeta_1}{\mu_2}\,\frac{\cos v_2}{\cosh v_1}\right)\tanh u\\
C_5&=\frac{\Delta\alpha\Delta t}{\kappa D_0}\left(\frac{\cos v_2}{\cosh v_1}-\frac{3}{2}\,\frac{\zeta_1}{\mu_2}\,\frac{\sin v_2}{\sinh v_1}\right)\tanh u
\end{aligned}\right\}, \tag{26}$$

where

$$D_0=\left[\gamma_1+\frac{3}{2}\,\frac{\zeta_1}{\mu_2}\,\gamma_2+\left(\gamma_2-\frac{3}{2}\,\frac{\zeta_1}{\mu_2}\,\gamma_1\right)\frac{\sin 2v_2}{\sinh 2v_1}\right]\tanh u$$
$$-\beta_1\left(\operatorname{cotanh}2v_1-\frac{\cos 2v_2}{\sinh 2v_1}\right)$$

is the determinant of the system (25). This concludes the solution of the problem.

Special Cases

In the case of zero through-thickness stiffness ($K=0$), the peeling stress is also zero and (17) yields:

$$\beta=\sqrt{1+\epsilon}\,k.$$

Then we have:

$$\zeta=\frac{2(1+\epsilon)}{3}\,k^2,\quad q_0=q_1=-\left(\frac{1+\epsilon}{3}\,k^2\right)^2,\quad \alpha=0,$$

$$a=b=\frac{1+\epsilon}{3}\,k^2,\quad \mu_1=\mu_2=0,\quad \gamma_1=\gamma_2=0,$$

and the equation (19) for the shearing stress reduces to the following simple formula:

$$\tau(x)=C_1\sinh\beta_1 x. \tag{27}$$

The constant integration of C_1 in this formula can be obtained from the nonzero boundary condition in (16):

$$C_1=\frac{\Delta\alpha\Delta t}{\kappa\beta_1\cosh\beta_1 l}.$$

The maximum shearing stress occurs at the end cross-section $x=l$:

$$\tau_{\max}=\tau(l)=\frac{\Delta\alpha\Delta t}{\kappa\beta_1}\tanh\beta_1 l.$$

For sufficiently large $\beta_1 l$ values, $\tanh\beta_1 l$ is approximately equal to unity, and the maximum shearing stress becomes independent of the plate size.

In the case of infinitely large through-thickness stiffness ($K\to\infty$), the basic equations (6) and (11) yield:

$$p(x)=\mu\tau'(x),\quad \tau''(x)-k^2\tau(x)=0. \tag{28}$$

Note that in this extreme special case the equation for the peeling stress $p(x)$ can be still presented in the same form as the equation for the shearing stress $\tau(x)$. Indeed, from the first equation in (28) we have:

$$\tau''(x)=\frac{1}{\mu}\,p'(x).$$

Introduction of this relationship into the second equation in (28) yields:

$$\tau(x)=\frac{1}{\mu k^2}\,p'(x).$$

Substitution of this formula into the first formula in (28) results in the equation:

$$p''(x)-k^2p(x)=0, \tag{29}$$

which is identical to the second equation in (28) for the shearing stress.

Equations (28) and (29) have the following simple solutions:

$$\tau(x)=C_1\sinh kx,\quad p(x)=C_2\cosh kx, \tag{30}$$

where the antisymmetry of the function $\tau(x)$ and the symmetry of the function $p(x)$ are taken into account.

Introducing (30) into the first equation in (28), we conclude that the constants C_1 and C_2 are related as follows:

$$C_2=k\mu C_1.$$

Evidently, since there is only one independent constant of integration in the solutions (30), all the three boundary conditions in (16) cannot be satisfied. Using the nonzero condition, we have:

$$C_1=\frac{\Delta\alpha\Delta t}{\kappa k\cosh kl},\quad C_2=\mu\frac{\Delta\alpha\Delta t}{\kappa\cosh kl}.$$

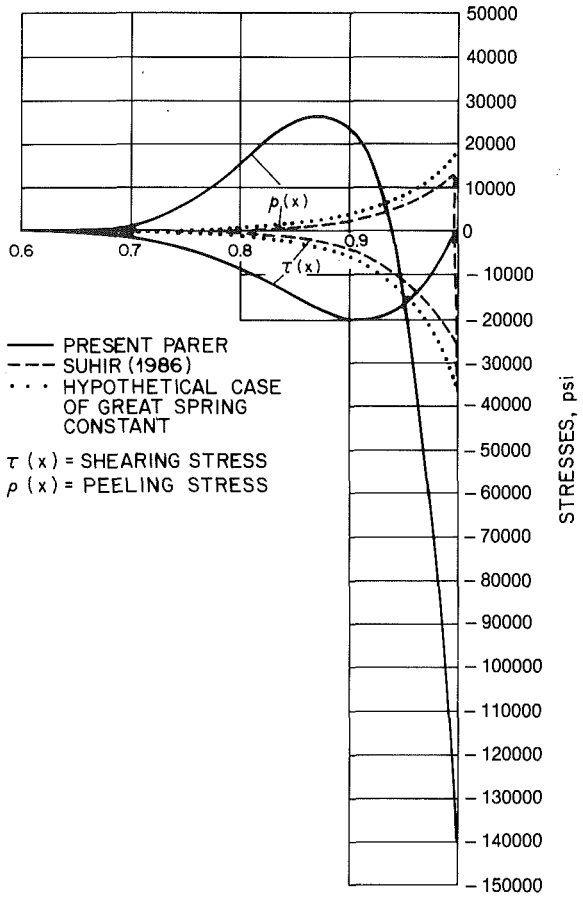


Fig. 2

The maximum interfacial stresses occur at the end cross-sections:

$$\tau_{\max} = \frac{\Delta\alpha\Delta t}{\kappa k} \tanh kl, \quad p_{\max} = \mu \frac{\Delta\alpha\Delta t}{\kappa}.$$

Note that the distributed peeling stresses determined by the second formula in (30) do not satisfy the conditions of self-equilibrium. Therefore, concentrated forces

$$N_0 = \frac{1}{2}\int_0^l p(x)dx = \mu \frac{\Delta\alpha\Delta t}{k\kappa} \tanh kl = \mu\tau_{\max} \qquad (31)$$

should be introduced at the end cross-sections to satisfy these conditions (Suhir, 1986).

Comparing the formulas for the maximum shearing stress in the cases $K=0$ and $K\to\infty$, we conclude that when the spring constant K increases from zero to infinity, the maximum shearing stress increases by a factor of

$$\eta = \frac{\beta_1}{k} = \sqrt{1+\epsilon}.$$

Since the ϵ value changes from $\epsilon=0$ (in the cases of $h_1/h_2 = D_1/D_2$) to $\epsilon=3$ (in the cases of $D_2 \gg D_1$ or $D_1 \gg D_2$), the factor η changes from $\eta=1$ to $\eta=2$. Therefore, for thermostats characterized by small ϵ values (and small peeling stresses), the through-thickness sifffness has a small effect on the maximum shearing stress. This effect increases with an increase in the ϵ value, but even for a significant difference in the flexural rigidities and thicknesses of the thermostats strips (when the parameter ϵ is close to 3) the factor η does not exceed two. This justifies the fact that in an approximate analysis the maximum shearing stress can be determined without considering the effect of the finite through-thickness stiffness at all. This is, of course, a conservative approach, which results in a reasonable overestimation of the actual stress.

Numerical Example

The numerical example is carried out for a molybdenum/aluminum thermostat. The following input data is used: $E_1 = 3.247\times10^5$MPa, $\nu_1 = 0.293$, $\alpha_1 = 4.9\times10^{-6}$ 1/°C, $h_1 = 2.54$ mm, $E_2 = 7.033\times10^4$MPa, $\nu_2 = 0.345$, $\alpha_2 = 23.6\times10^{-6}$ 1/°C, $h_2 = 2.54$ mm, $l = 25.4$ mm, $\Delta t = 240$°C.

The calculated stresses $\tau(x)$ and $p(x)$ are plotted in Fig. 2 as solid lines. The maximum shearing stress occurs in the vicinity of the cross-section $x \cong 0.9$ and is about 138MPa. The maximum value of the peeling stress takes place at the end cross-section and is 920MPa. The calculated spring constant value is about $K = 2.7\times10^{13}$N/m^3. The stresses predicted by the simplified formulas, assuming $K\to\infty$, are shown as dotted lines. As evident from the obtained data, these formulas conservatively predict the maximum shearing stress, but underestimate the extremes of the peeling stress. (Note that since the formula (31) determines the maximum force at the end cross-section, the corresponding stress shown in Fig. 2 was evaluated, assuming that this stress is distributed over the same length as in the nonsimplified analysis.) Stresses calculated on the basis of the present theory for a hypothetical case of a very large spring constant ($K = 10^{12}$lb/in^3) are shown as broken lines. These curves are rather close to the curves, obtained on the basis of the approach, assuming infinitely large through-thickness stiffness.

Conclusion

An engineering theory of interfacial stresses in thermostat-like structures is developed. In this theory the boundary conditions at the short edges are satisfied not only for the normal stress in the thermostat strips, but also for the interfacial stresses. A numerical example showed that the formulas, obtained on the basis of a simplified approach assuming infinitely large through-thickness stiffness, satisfactorily predict the maximum shearing stresses, but may underestimate the peeling stresses. Therefore, it is recommended that the latter be evaluated on the basis of the more accurate analysis presented in this paper.

Acknowledgment

This author acknowledges with thanks the help of R. Thompson, who developed a computer program for stress calculation.

References

Aleck, B. J., 1949, "Thermal Stresses in a Rectangular Plate Clamped Along an Edge," ASME JOURNAL OF APPLIED MECHANICS., Vol. 16, pp. 118–122.

Blanchard, J. P., and Watson, R. D., 1986, "Residual Stresses in Bonded Armor Tiles for On-Vessel Fusion Components," *Nuclear Engng. Design/Fusion*, Vol. 4, pp. 61–66.

Blech, J. J., and Kantor, Y., 1984, "An Edge Problem Having No Singularity at the Corner," *Computers and Structures*, Vol. 18, No. 4, pp. 609–617.

Bogy, D. B., 1968, "Edge-bonded Dissimilar Orthogonal Elastic Wedges Under Normal and Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 460-473.

Bogy, D. B., 1970, "On the Problem of Edge-Bonded Elastic Quarter-Planes Loaded at the Boundary," *Int. J. Solid Structures*, Pergamon Press, Vol. 6, pp. 1287–1313.

Chang, Fo-Van, 1983, "Thermal Contact Stresses of Bi-Metal Strip Thermostat," *Appl. Math. J. Mech.*, Vol. 4, No. 3, Tsing-hua Univ., Beijing, China, pp. 363–376.

Chen, W. T., and Nelson, C. W., 1979, "Thermal Stress in Bonded Joints," *IBM J. Res. Devel.*, Vol. 23, No. 2, pp. 178–188.

Grimado, P. B., 1978, "Interlaminar Thermoelastic Stresses in Layered Beams," *J. Thermal Stresses*, No. 1, pp. 75–86.

Gerstle, Jr., F. P., and Chambers, R. S., 1987, "Analysis of End Stresses in Glass-Metal Bi-Material Strips," ASME WAM, Boston, Mass., Dec. 1987.

Hein, V. L., and Erdogan, F., 1971, "Stress Singularities in a Two-Material Wedge," *Int. J. of Fractions Mechanics*, Vol. 7, pp. 317–330.

Hess, M. S., 1969, "The End Problem for a Laminated Elastic Strip-II. Differential Expansion Stresses," *J. Composite Materials*, Vol. 3, pp. 630–641.

Saganuma, K., Okamoto, T., and Koizumi, M., 1984, "Effect of Interlayers in Ceramic-Metal Joints with Thermal Expansion Mismatches," *J. Am. Cer. Soc.*, Vol. 67, No. 12, pp. C256–C257.

Suhir, E., 1986. "Stresses in Bi-Metal Thermostats," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 657–660.

Timoshenko, S. P., 1925, "Analysis of Bi-Metal Thermostats," *J. Opt. Soc. of America*, Vol. 11, pp. 233–255.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Electricity*, 3rd ed. McGraw-Hill, New York.

Wang, S. S., and Choi, I., 1982, "Boundary-Layer Effects in Composite Dominates: Part 1: Free-Edge Stress Singularities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 541–548; Part 2: Free-Edge Stress Solutions and Basic Characteristics, Ibid, pp. 549–560.

Zayfang, R., 1971, "Stress and Strains in a Plate Bonded to a Substrate," *Solid State Elect.*, Vol. 14, pp. 1035–1039.

Zwiers, R. I., Ting, T. C. T., and Spilker, R. L., 1982, "On the Logarithmic Singularity of Free-Edge Stress in Dominated Composite Under Uniform Extension," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 561–569.

Yuan Ruo Wang
Research Assistant.

Tsu-Wei Chou
Professor.
Mem. ASME

Center for Composite Materials and
Department of Mechanical Engineering,
University of Delaware,
Newark, DE 19716

Three-Dimensional Transient Interlaminar Thermal Stresses in Angle-Ply Composites

This paper studies the three-dimensional transient interlaminar thermal stresses in elastic, angle-ply laminated composites due to sudden changes in the thermal boundary conditions. The transient temperature field and transient interlaminar thermal stresses of the laminate are obtained by solving the heat conduction equation and by a zeroth-order perturbation analysis of the equilibrium equations, respectively. Numerical results for a four-layer angle-ply laminate have shown that the interlaminar normal stress near the free edge is significantly higher than that in the interior region and it increases rapidly with the fiber volume fraction.

Introduction

With the increasing applications of advanced fiber composites under severe environment, the thermomechanical behavior of such materials, especially metal and ceramic matrix composites, has received considerable attention. This is mostly because that the temperature at which the metal and ceramic composites could be utilized is much higher than that for polymer-based composites.

Among the tremendous research interests in the analysis and design of advanced fiber composites, the problem of "free-edge effect" of laminated composites has attracted considerable attention. Both experimental studies and approximate analytical solutions have indicated that there exists highly localized regions of stress concentration near laminate free edges due to the geometrical as well as material discontinuities. The highly localized, boundary layer stress coupled with the relatively low interlaminar strength is often detrimental to the durability of laminated composites. Thus, there is the need to establish adequate analytical and experimental techniques to investigate the interlaminar stress behavior near free edges of laminates due to thermal and mechanical loadings.

Considerable effort has been made to investigate the behavior of edge stresses. Pipes and Pagano (1970) first employed finite difference method to study the nature of interlaminar stresses in symmetric composite laminates due to mechanical loadings. Wang and Crossman (1977) used the finite element method to investigate the edge effect in symmetric composite laminates subjected to uniaxial tension and a uniform temperature change. Wang and Dickson (1978) extended the Galerkin procedure to reveal the singularity behavior of interlaminar stress in composite laminates. Wang and Yuan (1983) presented a hybrid finite element method for analyzing the composite laminate elasticity problems with singularities.

The interlaminar edge stress problem also has been examined by analytical techniques. For example, Pipes and Pagano (1974) developed an approximate elasticity solution for the response of a finite width, angle-ply composite laminate under uniform axial strain and yielded the sinusoidal-hyperbolic series solution form of the interlaminar stresses. Hsu and Herakovich (1976, 1977) used the perturbation method to obtain a zeroth-order solution for edge effects in angle-ply composite laminates subjected to a uniform strain. Wang and Choi (1979, 1982) employed the Lekhnitskii's complex variable potential approach to investigate the singularity of boundary layer stresses in composite laminates subjected to a uniform extension and a uniform temperature change. Recently, Kassapoglou and Lagace (1987) obtained the closed-form solutions to the problem of interlaminar stresses at a straight-free edge of cross-ply and angle-ply laminates using the force balance method and the minimum complementary energy principle. Comparison of the existing results of studies of free-edge effect has shown good agreements in the far field and appreciable difference near the laminate boundary.

To the authors' knowledge, there is a lack of fundamental understanding of the interlaminar stress concentration induced by transient temperature field. Wang and Chou (1987, 1988) initiated the study of this problem. Transient thermal stress analyses in unidirectional fiber composites have been reported by Wang and Chou (1985, 1986) and Wang, Pipes, and Chou (1986). With the development of laminated metal and ceramic matrix composites for elevated temperature applications, there is certainly the need for better understanding of the transient interlaminar stresses.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10-15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 23, 1987; final revision, December 12, 1988.

Paper No. 89-WA/APM-14.

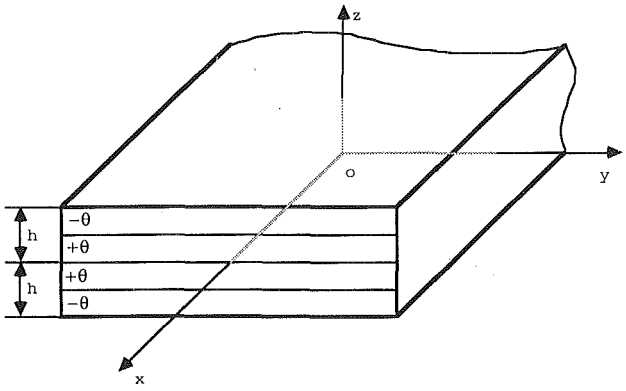


Fig. 1 Geometry of the angle-ply laminate for analytical modeling

This paper analyzes the transient heat transfer and thermoelastic problem of a balanced, symmetric angle-ply laminate by dividing the plate into the interior and boundary layer regions. In the interior region of the laminate, the stress and displacement are studied by the classical laminate theory, while in the boundary layer region, a perturbation method is applied to the elastic governing equations. This approach is adopted because the transient interlaminar thermal stress in the boundary layer region can not be assessed accurately by the classical laminate theory due to the existence of high gradient displacement and stress fields, and the three-dimensional nature of the boundary layer effect. The Prandtl's matching principle of perturbation theory is imposed to match the solutions of these two regions. A four-layer $(-\theta/\theta)_s$ laminate is presented as a numerical example. The boundary layer stress singularities, and the fiber volume fraction and fiber orientation effects have been studied. The laminate thickness-to-width ratio influence and stress solution sensitivity to the composite elastic and thermal properties have also been assessed. Finally, the transient thermal effect on boundary layer stress is compared with that induced by the application of uniaxial tension.

Thermal Stress Field

1 Basic Equations. A 4-ply symmetric $(-\theta/\theta)_s$ composite laminate is considered in this study. The laminate is of thickness $2h$ and width $2b$; it is infinite in extent along the x-direction (Fig. 1). Since the thermal boundary conditions are uniform along the surfaces $y = \pm b$ (Fig. 2), the displacements are independent of the x-axis and expressed as:

$$\begin{aligned} u &= u(y,z,t) \\ v &= v(y,z,t) \\ w &= w(y,z,t) \end{aligned} \tag{1}$$

The time variable t will not be written out for convenience in the following discussions. The equilibrium equations are

$$\begin{aligned} &\partial\sigma_x/\partial x + \partial\tau_{xy}/\partial y + \partial\tau_{xz}/\partial z = 0 \\ &\partial\tau_{xy}/\partial x + \partial\sigma_y/\partial y + \partial\tau_{yz}/\partial z = 0 \\ &\partial\tau_{xz}/\partial x + \partial\tau_{yz}/\partial y + \partial\sigma_z/\partial z = 0. \end{aligned} \tag{2}$$

Here, σ and τ denote normal and shear stress, respectively.

The stress-strain relations for orthotropic materials are (Vinson and Chou, 1975; Vinson and Sierakowski, 1986):

$$\begin{aligned} \sigma_x &= C_{11}\epsilon_x + C_{12}\epsilon_y + C_{13}\epsilon_z + C_{16}\gamma_{xy} - \beta_1 T \\ \sigma_y &= C_{12}\epsilon_x + C_{22}\epsilon_y + C_{23}\epsilon_z + C_{26}\gamma_{xy} - \beta_2 T \\ \sigma_z &= C_{13}\epsilon_x + C_{23}\epsilon_y + C_{33}\epsilon_z + C_{36}\gamma_{xy} - \beta_3 T \\ \tau_{yz} &= C_{44}\gamma_{yz} \\ \tau_{xz} &= C_{55}\gamma_{xz} \\ \tau_{xy} &= C_{16}\epsilon_x + C_{26}\epsilon_y + C_{36}\epsilon_z + C_{66}\gamma_{xy} - \beta_6 T \end{aligned} \tag{3}$$

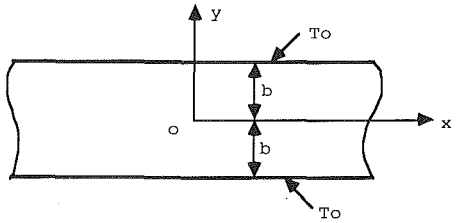


Fig. 2 Thermal boundary conditions

where

$$\begin{aligned} \beta_1 &= \alpha_x C_{11} + \alpha_y C_{12} + \alpha_z C_{13} + \alpha_{xy} C_{16} \\ \beta_2 &= \alpha_x C_{12} + \alpha_y C_{22} + \alpha_z C_{23} + \alpha_{xy} C_{26} \\ \beta_3 &= \alpha_x C_{13} + \alpha_y C_{23} + \alpha_z C_{33} + \alpha_{xy} C_{36} \\ \beta_6 &= \alpha_x C_{16} + \alpha_y C_{26} + \alpha_z C_{36} + \alpha_{xy} C_{66}. \end{aligned}$$

In equation (3), C_{ij} are elastic stiffness constants; α_x, α_y, α_z are coefficients of thermal expansion, and $T = T(y, t)$ denotes the transient temperature field (Appendix A). α_{xy} is given in Appendix C.

Using the stress-strain equation (3) and strain-displacement relationship (Vinson and Chou, 1975; Vinson and Sierakowski; 1986), the equilibrium equations are written in terms of displacements.

$$\begin{aligned} &C_{66}\partial^2 u/\partial y^2 + C_{55}\partial^2 u/\partial z^2 + C_{26}\partial^2 v/\partial y^2 \\ &\qquad + C_{36}\partial^2 w/\partial y\partial z = \beta_6 \partial T/\partial y \\ &C_{26}\partial^2 u/\partial y^2 + C_{22}\partial^2 v/\partial y^2 + C_{44}\partial^2 v/\partial z^2 \\ &\qquad + (C_{23} + C_{44})\partial^2 w/\partial y\partial z = \beta_2 \partial T/\partial y \\ &C_{36}\partial^2 u/\partial y\partial z + (C_{44} + C_{23})\partial^2 v/\partial y\partial z + C_{44}\partial^2 w/\partial y^2 \\ &\qquad + C_{33}\partial^2 w/\partial z^2 = 0. \end{aligned} \tag{4}$$

The following nondimensional variables and constants are introduced: $U = u/h$, $V = v/h$, $W = w/h$, $Y = y/b$, $Z = z/h$ and $Q_{ij} = C_{ij}/C_{\max}$, where $C_{\max}$ is the largest value among all the C_{ij}'s. Thus, the equilibrium equations could be written as follows after eliminating (h/b) and higher-order terms (Hsu and Herakovich, 1976, 1977).

$$\begin{aligned} Q_{55}\partial^2 U/\partial Z^2 &= 0 \\ Q_{44}\partial^2 V/\partial Z^2 &= 0 \\ Q_{33}\partial^2 W/\partial Z^2 &= 0. \end{aligned} \tag{5}$$

The solution of equation (5) assumes the following general form in terms of functions A, B, C, D, E, and F of Y.

$$\begin{aligned} U &= A(Y)Z + B(Y) \\ V &= C(Y)Z + D(Y) \\ W &= E(Y)Z + F(Y). \end{aligned} \tag{6}$$

With the following conditions of symmetry with respect to the laminate midplane,

$$\begin{aligned} \partial U(Y,0)/\partial Z &= 0 \\ \partial V(Y,0)/\partial Z &= 0 \\ W(Y,0) &= 0, \end{aligned} \tag{7}$$

the solution of equation (6) becomes

$$\begin{aligned} U &= B(Y,t) \\ V &= D(Y,t) \\ W &= E(Y,t)Z. \end{aligned} \tag{8}$$

2 Solution for the Interior Region. The classical plate theory is assumed to hold in the interior region, such that $\sigma_z = \tau_{xz} = \tau_{yz} = 0$. Utilizing the condition of $\sigma_z = 0$, and equations (3) and (8), we obtain

$$E(Y,t) = (\beta_3/C_{33})T(Y,t). \tag{9}$$

It is assumed here that h/b is sufficiently small (< 10 percent) and can be neglected. Obviously, the conditions of $\tau_{xz} = \tau_{yz} = 0$ are satisfied with the solution form of equation (8).

The symmetry conditions at the laminate central plane are

$$\begin{aligned} U(0,Z) &= 0 \\ V(0,Z) &= 0 \\ \partial W(0,Z)/\partial Y &= 0. \end{aligned} \tag{10}$$

Equations (10) lead to

$$B(Y) = -B(-Y) \qquad D(Y) = -D(-Y).$$

Furthermore, the requirement of continuity of displacements in U and V results in

$$B^{(1)} = B^{(2)} = \ldots = B(Y) \qquad D^{(1)} = D^{(2)} = \ldots = D(Y).$$

Here, the superscripts denote the laminae in the laminate. Displacement continuity in W is not satisfied in the present zeroth-order approximation. Continuity in W could be achieved only when higher-order terms of (h/b) are included.

To obtain explicit expressions of $B(Y,t)$ and $D(Y,t)$, it is necessary to consider the force equilibrium.

$$\begin{aligned} \Sigma F_y &= \Sigma \sigma_y{}^{(k)} t_k = 0 \\ \Sigma F_x &= \Sigma \tau_{xy}{}^{(k)} t_k = 0 \end{aligned} \tag{11}$$

where $t_k = h_k/h$, h_k = thickness of the kth lamina and $h = \Sigma h_k$. So, $B(Y,t)$ and $D(Y,t)$ for each layer are given by

$$\begin{aligned} (h/b)B(Y,t) &= [q_3 Q_1(Y,t) - q_2 Q_4(Y,t)]/(q_2 q_5 - q_3{}^2) \\ (h/b)D(Y,t) &= [q_5 Q_1(Y,t) - q_3 Q_4(Y,t)]/(q_3{}^2 - q_2 q_5) \end{aligned} \tag{12}$$

where

$$\begin{aligned} &Q_1(Y,t) = \Sigma[(C_{23}/C_{33})\beta_3 - \beta_2]^{(k)} \underline{T}^{(k)}(Y,t) h_k \\ &q_2 = \Sigma C_{22}{}^{(k)} h_k \quad q_3 = \Sigma C_{26}{}^{(k)} h_k \quad q_5 = \Sigma C_{66}{}^{(k)} h_k \\ &Q_4(Y,t) = \Sigma[(C_{36}/C_{33})\beta_3 - \beta_6]^{(k)} \underline{T}^{(k)}(Y,t) h_k \\ &\underline{T}^{(k)}(Y,t) = \int_o^Y T^{(k)}(Y,t)\, dY. \end{aligned}$$

The displacement field in the interior region is as equation (8), and $B(Y,t)$, $D(Y,t)$, and $E(Y,t)$ are given by equations (9) and (12).

3 Solution for the Boundary Layer Region. Following Hsu and Herakovich (1976, 1977), a stretching transformation parameter is introduced.

$$\eta = (1 - Y)/(h/b) \tag{13}$$

Then the equilibrium equation (4) becomes

$$\begin{aligned} &Q_{66}\partial^2 U/\partial\eta^2 + Q_{55}\partial^2 U/\partial Z^2 + Q_{26}\partial^2 V/\partial\eta^2 \\ &\quad - Q_{36}\partial^2 W/\partial\eta\partial Z = (h/b)(\beta_6/C_{\max})\partial T/\partial Y \\ &Q_{26}\partial^2 U/\partial\eta^2 + Q_{22}\partial^2 V/\partial\eta^2 + Q_{44}\partial^2 V/\partial Z^2 \\ &\quad - Q_{23} + Q_{44})\partial^2 W/\partial\eta\partial Z = (h/b)(\beta_2/C_{\max})\partial T/\partial Y \\ &\quad - Q_{36}\partial^2 U/\partial\eta\partial Z - (Q_{23} + Q_{44})\partial^2 V/\partial\eta\partial Z \\ &\quad + Q_{44}\partial^2 W/\partial\eta^2 + Q_{33}\partial^2 W/\partial Z^2 = 0. \end{aligned} \tag{14}$$

To satisfy the Prandtl's matching principle, the following expressions of the displacement field are assumed.

$$\begin{aligned} U^{(k)} &= B(Y,t) + Pe^{\lambda\eta}\cos(\alpha Z) \\ V^{(k)} &= D(Y,t) + Re^{\lambda\eta}\cos(\alpha Z) \\ W^{(k)} &= E(Y,t)Z + Se^{\lambda\eta}\sin(\alpha Z) \end{aligned} \tag{15}$$

Here, $B(Y,t)$, $D(Y,t)$, and $E(Y,t)$ are the interior region solutions of equations (9) and (12); P, R, and S are undetermined coefficients; α is an undetermined positive quantity; λ is the characteristic.

Substituting the $U^{(k)}$, $V^{(k)}$ and $W^{(k)}$ expressions into the equilibrium equations (14), we obtain

$$\begin{aligned} &(Q_{66}\lambda^2 - Q_{55}\alpha^2)P + Q_{26}\lambda^2 R - Q_{36}\lambda\alpha S = 0 \\ &Q_{26}\lambda^2 P + (Q_{22}\lambda^2 - Q_{44}\alpha^2)R - (Q_{23} + Q_{44})\lambda\alpha S = 0 \\ &Q_{36}\lambda\alpha P + (Q_{23} + Q_{44})\lambda\alpha R + (Q_{44}\lambda^2 - Q_{33}\alpha^2)S = 0. \end{aligned} \tag{16}$$

For nontrivial solutions of P, R, and S, the determinants of these algebraic equations must vanish and, thus, the characteristic equation is

$$\begin{vmatrix} Q_{66}\lambda^2 - Q_{55}\alpha^2 & Q_{26}\lambda^2 & -Q_{36}\lambda\alpha \\ Q_{26}\lambda^2 & Q_{22}\lambda^2 - Q_{44}\alpha^2 & -(Q_{23} + Q_{44})\lambda\alpha \\ Q_{36}\lambda\alpha & (Q_{23} + Q_{44})\lambda\alpha & Q_{44}\lambda^2 - Q_{33}\alpha^2 \end{vmatrix} = 0. \tag{17}$$

The six roots of λ have been found in the following forms (Appendix B):

$$\begin{aligned} \lambda_{1,2} &= \pm a^{(k)}\alpha \\ \lambda_{3,4} &= \pm b^{(k)}\alpha \\ \lambda_{5,6} &= \pm c^{(k)}\alpha \end{aligned} \tag{18}$$

where $a^{(k)}$, $b^{(k)}$ and $c^{(k)}$ are three positive constants. The positive roots of λ are dropped to avoid divergence in the displacement field.

Thus, the displacements in the boundary layer region can be written as follows:

$$\begin{aligned} U^{(k)} &= B(y,t) + (P_1 e^{-a\alpha\eta} + P_2 e^{-b\alpha\eta} + P_3 e^{-c\alpha\eta})\cos(\alpha Z) \\ V^{(k)} &= D(Y,t) + (R_1 e^{-a\alpha\eta} + R_2 e^{-b\alpha\eta} + R_3 e^{-c\alpha\eta})\cos(\alpha Z) \\ W^{(k)} &= E(Y,t)Z + (S_1 e^{-a\alpha\eta} + S_2 e^{-b\alpha\eta} + S_3 e^{-c\alpha\eta})\sin(\alpha Z). \end{aligned} \tag{19}$$

There are ten unknowns for the kth layer solution (P_1, P_2, P_3, R_1, R_2, R_3, S_1, S_2, S_3, and α).

The available equations are: (*i*) three stress boundary conditions, $\sigma_y(b, z) = 0$, $\tau_{xy}(b, z) = 0$ and $\tau_{yz}(b, z) = 0$, (*ii*) six equilibrium equations (16), and (*iii*) the integrated equilibrium condition

$$\int_o^{1/2} \tau_{xy}(0,Z)\, h dZ = \int_o^1 \tau_{xz}(Y, 1/2)\, b dY. \tag{20}$$

The nine equations of (*i*) and (*ii*) are summarized next:

$$\begin{aligned} &Q_{22}aR_1 + Q_{22}bR_2 + Q_{22}cR_3 + Q_{23}S_1 + Q_{23}S_2 \\ &\quad + Q_{23}S_3 + Q_{26}aP_1 + Q_{26}bP_2 + Q_{26}cP_3 = F_1/\alpha\cos(\alpha Z) \\ &Q_{26}aR_1 + Q_{26}bR_2 + Q_{26}cR_3 + Q_{36}S_1 + Q_{36}S_2 \\ &\quad + Q_{36}S_3 + Q_{66}aP_1 + Q_{66}bP_2 + Q_{66}cP_3 = F_2/\alpha\cos(\alpha Z) \\ &Q_{44}R_1 + Q_{44}R_2 + Q_{44}R_3 - Q_{44}aS_1 - Q_{44}bS_2 - Q_{44}cS_3 = 0 \\ &B_aR_1 + C_aS_1 + A_aP_1 = 0 \\ &D_aR_1 + E_aS_1 + B_aP_1 = 0 \\ &B_bR_2 + C_bS_2 + A_bP_2 = 0 \\ &D_bR_2 + E_bS_2 + B_bP_2 = 0 \\ &B_cR_3 + C_cS_3 + A_cP_3 = 0 \\ &D_cR_3 + E_cS_3 + B_cP_3 = 0 \end{aligned} \tag{21}$$

where

$$\begin{aligned} F_1 &= -Q_{22}DD - Q_{23}E_1 - Q_{26}DB + \beta_2 T_o/C_{\max} \\ F_2 &= -Q_{26}DD - Q_{36}E_1 - Q_{66}DB + \beta_6 T_o/C_{\max} \end{aligned}$$

$$\begin{aligned} A_a &= Q_{66}a^2 - Q_{55} & A_b &= Q_{66}b^2 - Q_{55} & A_c &= Q_{66}c^2 - Q_{55} \\ B_a &= Q_{26}a^2 & B_b &= Q_{26}b^2 & B_c &= Q_{26}c^2 \\ C_a &= Q_{36}a & C_b &= Q_{36}b & C_c &= Q_{36}c \\ D_a &= Q_{22}a^2 - Q_{44} & D_b &= Q_{22}b^2 - Q_{44} & D_c &= Q_{22}c^2 - Q_{44} \\ E_a &= (Q_{23} + Q_{44})a & E_b &= (Q_{23} + Q_{44})b & E_c &= (Q_{23} + Q_{44})c \end{aligned}$$

$$\begin{aligned} DQ_1 &= \Sigma[(C_{23}/C_{33})\beta_3 - \beta_2]^{(k)} T_o h_k \\ DQ_4 &= \Sigma[(C_{36}/C_{33})\beta_3 - \beta_6]^{(k)} T_o h_k \end{aligned}$$

Table 1 Fiber thermoelastic properties

	ρ (g/cm^3)	E1 (Gpa)	E2	G12	G23	υ_{12}	υ_{23}	α_1(1/°C) x10^6	α_2	K1 (W/m,K)	K2	Cp (J/kg,K)
SiC	3.2	406	406	169	169	0.2	0.2	5.2	5.2	86	86	670
Al_2O_3	3.9	385	385	154	154	0.26	0.26	8.5	8.5	30	30	800
T300	1.77	221	14	9	4.8	0.2	0.25	1	10	84	8.4	920

ρ: density

E1,E2,G12,G23: axial,transverse and shear Young's modulus

υ_{12}, υ_{23}: Poisson's ratio

α_1, α_2: thermal expansion coefficient

K1,K2: thermal conductivity

Cp: specific heat

Table 2 Matrix thermoelastic properties

	ρ (g/cm^3)	E (Gpa)	G	υ	α (1/°C) x10^6	K (W/m,K)	Cp (J/kg,K)
BG	2.2	63.7	28	0.21	3.25	1.09	711
LAS	2.42	85	35	0.22	1	1.5	800
Al	2.7	69	26	0.33	23.6	300	481
Mg	1.7	45.5	17	0.33	26	169	67

BG: borosilicate glass
LAS: lithium aluminosilicate
Al: aluminum
Mg: magnesium

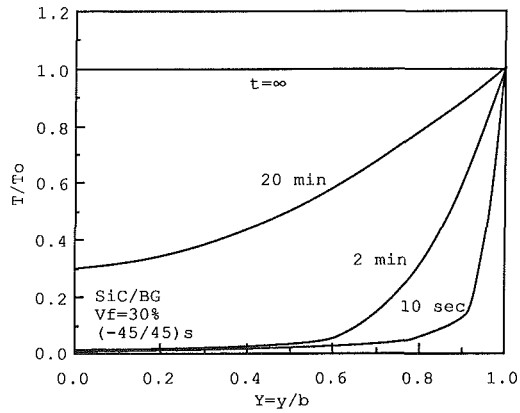


Fig. 3 Transient temperature distributions

Table 3 Identifications of composite systems

fiber / matrix	SiC	Al_2O_3	T300
BG	(11)	(12)	(13)
LAS	(21)	(22)	(23)
Al	(31)	(32)	(33)
Mg	(41)	(42)	(43)

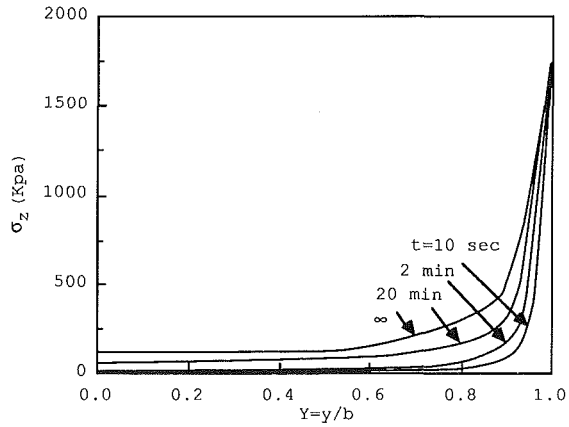


Fig. 4 Transient interlaminar thermal stress of a SiC/BG (−45 deg/45 deg)$_s$ laminate for V_f = 30 percent and T_o = 1°C

$$DB = (q_3 DQ_1 - q_2 DQ_4)/(q_2 q_5 - q_3^2)$$

$$DD = (q_5 DQ_1 - q_3 DQ_4)/(q_3^2 - q_2 q_5)$$

$$E_1 = \beta_3 T_o / C_{33}.$$

Therefore, the P, R, and S could be solved in terms of α. To determine α, the force equilibrium equation (20) must be considered.

Numerical Results

A four-layer $(-\theta/\theta)_s$ laminated plate is taken as a numerical example. Each layer is 5 mm in thickness (h_k), 200 mm in width (b). The fiber and matrix thermoelastic properties come from Chou and Yang (1986), Chamis (1984), Brennan and Prewo (1982), and are listed in Tables 1 and 2. SiC/borosilicate glass (BG) laminate is used as a baseline composite system for demonstration of the results, and numerical computations are also performed for eleven other composite systems (Table 3).

From Appendixes B and C, the eigenvalues of the characteristic equation (17) are obtained for (−45 deg/45 deg)$_s$ SiC/BG laminate.

$$\lambda_1 = -\alpha\sqrt{\omega_1} = -1.1018\alpha$$

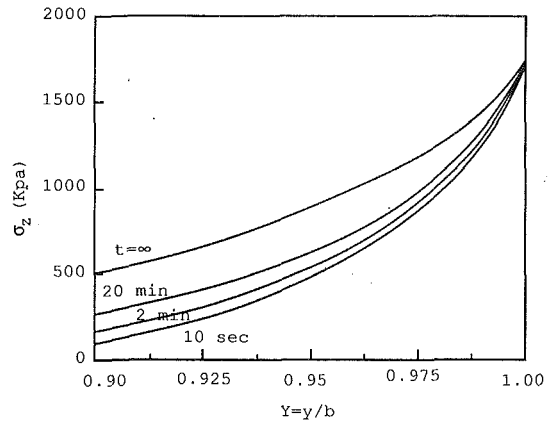


Fig. 5 Detailed edge stress distribution of Fig. 4

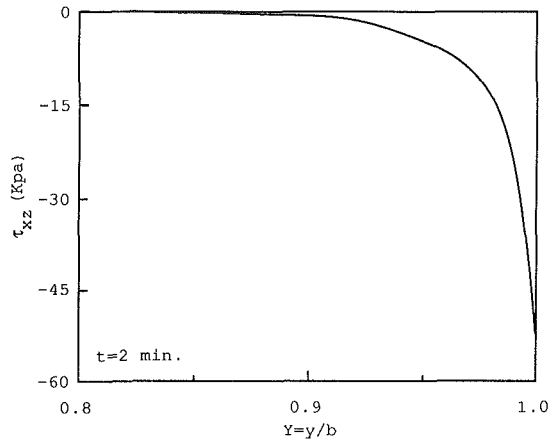


Fig. 6 Transient interlaminar thermal stress of a $(-45\ \text{deg}/45\ \text{deg})_S$ SiC/BG laminate for $V_f = 30$ percent and $T_o = 1°C$

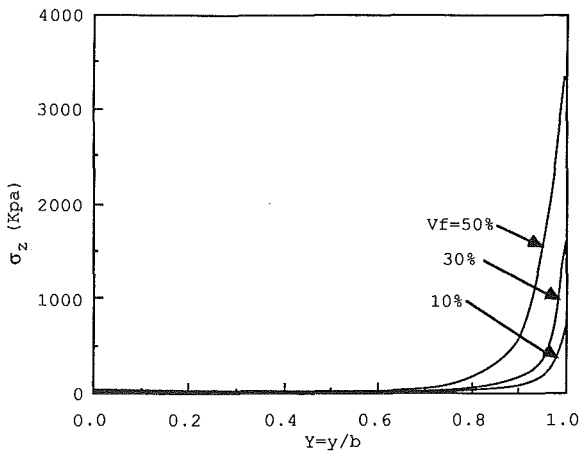


Fig. 7 Fiber volume fraction effect on interlaminar normal stress for a SiC/BG $(-45\ \text{deg}/45\ \text{deg})_S$ laminate at $T_o = 1°C$ and $t = 2$ min

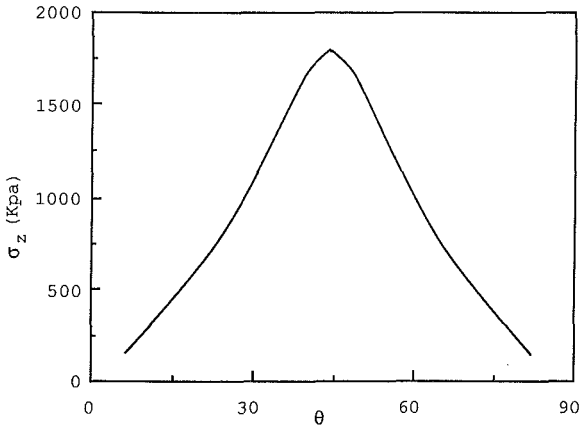


Fig. 8 Fiber orientation effect on interlaminar normal stress for a SiC/BG $(-\theta/\theta)_S$ laminate at $T_o = 1°C$, $t = 2$ min, $V_f = 30$ percent, and Y = 1.0

$$\lambda_2 = -\alpha\sqrt{\omega_2} = -0.7917\alpha$$
$$\lambda_3 = -\alpha\sqrt{\omega_3} = -1.0\alpha$$

The solution of equation (21) is

$$R_1 = 0.5555\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$R_2 = 0.3129\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$R_3 = -0.9247\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$S_1 = -0.7380\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$S_2 = -0.2120\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$S_3 = 0.9247\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$P_1 = -0.2576\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$P_2 = 0.1451\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}$$
$$P_3 = 0.9247\times10^{-6}\{T_o/[\alpha\cos(\alpha Z)]\}.$$

Equation (20) leads to a transcendental equation for determining α at $Z = 1/2$ and different time t.

$t = 10$ s	2 min.	20 min.	∞
$\alpha = 6.2831$	6.2833	8.0614	8.9868

Then, the displacement field is readily obtained from equation (19), and the transient thermal interlaminar stresses are derived from the displacement field.

The transient temperature profile is depicted in Fig. 3 for a $(-45\ \text{deg}/45\ \text{deg})_s$ SiC/BG laminate of $V_f = 30$ percent. There exists a sharp variation of the temperature field along the width of the plate at the beginning of the sudden heating, and it is obvious that the temperature distribution becomes uniform as time tends to infinite.

The transient interlaminar normal stress distribution of the $(-45\ \text{deg}/45\ \text{deg})_s$ SiC/BG laminate which is subjected to a sudden edge heating of the magnitude $T_o = 1°C$ at $t = 0^+$ is demonstrated in Fig. 4. No stress singularity is found as a consequence of the assumed displacement field, but it is apparent that the interlaminar normal stress increases very significantly as approaching to the boundary ($Y = 1$). The detailed local interlaminar normal stress distribution ($Y = 0.9 \sim 1.0$) is shown in Fig. 5. The stress at $Y = 1.0$ is about $3 \sim 20$ times higher than that at $Y = 0.99$ for different transient times. This indicates the existence of high stress concentration. As the heating proceeds, the overall interlaminar normal stress increases smoothly, but the stress which is very close to the boundary remains almost constant. On the other hand, the interlaminar normal stress tends to zero away from the free edge of the laminate due to the adoption of the classical laminate theory in the interior region.

Unlike the interlaminar shear stresses induced by axial tension, which are in the same order of magnitude as the interlaminar normal stress (Hsu and Herakovich, 1977), the transient thermal interlaminar shear stress τ_{xz} in Fig. 6 is very small compared to the interlaminar normal stress.

The fiber volume fraction effect is studied in Fig. 7. The interlaminar stress increases significantly with the fiber volume fraction due to the composite stiffness increase. For $(-\theta/\theta)_s$

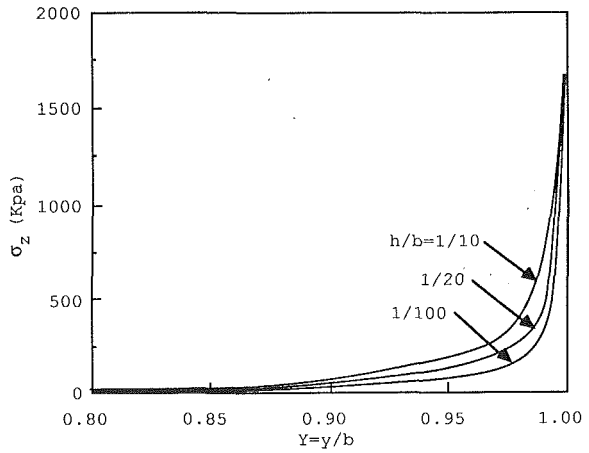


Fig. 9 *h/b* effect on interlaminar stress for a SiC/BG $(-45\ \text{deg}/45\ \text{deg})_s$ laminate at $V_f = 30$ percent, $T_o = 1°C$, and $t = 2$ min

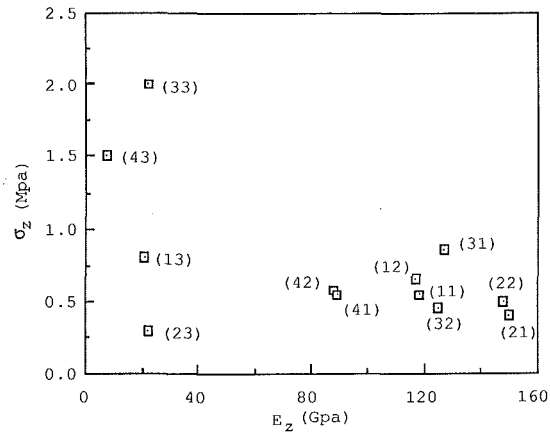


Fig. 11 Interlaminar stress versus modulus of twelve composite systems for $(-45\ \text{deg}/45\ \text{deg})_s$ laminates at $V_f = 30$ percent, $Y = 0.99$, and $t = 2$ min

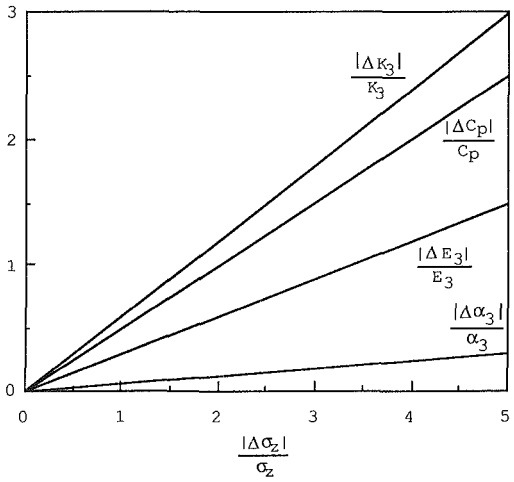


Fig. 10 Parametric studies of stress solution sensitivity to composite elastic and thermal properties, SiC/BG $(-45\ \text{deg}/45\ \text{deg})_s$ laminate is used as the baseline materials. $t = 2$ min and $Y = 0.99$.

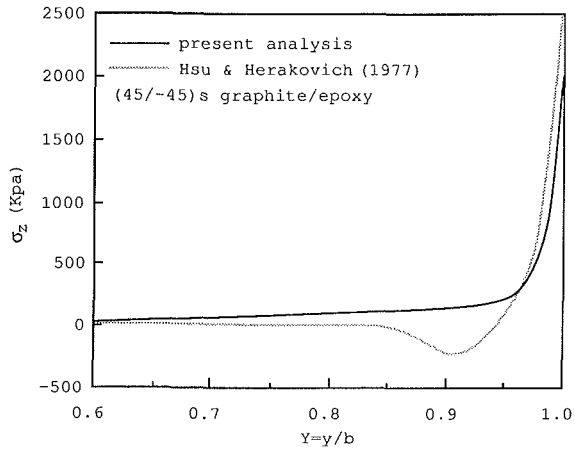


Fig. 12 Comparison of transient thermal effect and axial tension

angle-ply laminates, the interlaminar stress reaches its peak value when $\theta = 45$ deg (Fig. 8).

The laminate thickness-to-width ratio effect is investigated in Fig. 9. The interlaminar stress increases with the h/b value, and the stresses (at $Y \approx 1.0$) are the same for different h/b values. The present analysis is based upon the thin laminated plate theory, and neglects h/b and higher-order terms. Therefore, the theoretical prediction will be more accurate for smaller h/b values.

Figure 10 provides the parametric study of the stress solution sensitivity to the composite elastic and thermal properties. The Young's modulus (E_3) and thermal expansion coefficient (α_3) along the plate thickness direction have a more significant effect on the stress solution than the thermal conductivity (K_3) and specific heat (C_p). This is so because the linear stress-strain relationship is adopted and terms containing K_3 and C_p are in negative exponential form in the present analysis.

Figure 11 presents the interlaminar thermal stresses of twelve composite systems (Table 3) versus their moduli. The data are generated for $(-45\ \text{deg}/45\ \text{deg})_s$ laminates at $V_f = 30$ percent, $Y = 0.99$ and $t = 2$ min.

Finally, the transient thermal effect in the boundary layer identified from the present analysis is compared with that induced by uniaxial tension for a $(45\ \text{deg}/-45\ \text{deg})_s$ graphite/epoxy laminate (Hsu and Herakovich, 1977). The uniaxial tensile strain ϵ_x is assumed to be 10^{-3} for mechanical loading, and the boundary layer thermal condition is $T_o = 10°C$ for computing the transient thermal effect. It can be seen from Fig. 12 that the interlaminar normal stress tends to zero away from the boundary layer ($Y < 0.8$), and increases sharply as one approaches to the free edge for both cases.

Conclusions

(1) Transient interlaminar thermal stresses within the boundary layer region of an angle-ply composite laminate are three-dimensional in nature. They can not be determined by the classical laminate theory, but they have been examined explicitly by the present method.

(2) The interlaminar stress σ_z is very significant close to the free edge of the laminate and increases with fiber volume fraction due to the composite stiffness increase.

(3) The interlaminar normal stress reaches its maximum at $\theta = 45$ deg for $(-\theta/\theta)_s$ laminates, and a minimum for unidirectional composites ($\theta = 0$ deg or 90 deg).

(4) The interlaminar thermal shear stress is small compared to interlaminar normal stress.

(5) The overall interlaminar normal stress increases with the plate thickness-to-width ratio (h/b), but the peak value of interlaminar stress (at $Y \approx 1.0$) is independent of (h/b). The present analysis is suitable for small (h/b) values.

(6) The sensitivity of the stress solution to composite thermal and elastic properties is of practical interest. The results of these parametric studies indicate that the elastic stiffness and thermal expansion coefficient have a more severe influence on

the stress solution than the thermal conductivity and the specific heat (see Fig. 10). This is due to the assumption of the linear stress-strain relation and because the terms containing specific heat and thermal conductivity are in negative exponential form.

Discussions

(1) There is the need to define the boundary layer width on a rational basis. Pipes and Pagano (1970) defined the boundary layer width as the distance from the boundary at which the interlaminar shear stress is about three percent of the value calculated at the free edge. Since the interlaminar shear stress is relatively small in the present case, the applicability of this definition is questionable. Wang and Choi (1982) used an alternative definition of boundary layer width based on strain energy density consideration and defined the boundary layer width as the distance from the free edge where the strain energy density is three percent higher than that obtained in the far field.

In the present studies of transient thermal stress, the boundary layer width based on the definition just discussed changes with time. There is the need of further studies of boundary layer width for the transient case.

(2) There is still the uncertainty about the existence of the singular property of free-edge stresses from both mathematical and physical viewpoints. No mathematical stress singularity is found based upon the present method of analysis and assumed displacement field.

(3) Delamination of composites could occur due to the coupling of low interlaminar strength and relatively large interlaminar stresses induced by transient thermal effects.

Acknowledgment

This work is partially supported by the Office of Naval Research. Dr. L. H. Peebles, Jr. is the technical monitor.

References

Brennan, J. J., and Prewo, K. M., 1982, "Silicon Carbide Fiber Reinforced Glass-Ceramic Matrix Composites Exhibiting High Strength and Toughness," *J. of Mat. Sci.*, Vol. 17, pp. 2371–2383.

Chamis, C. C., 1984, "Simplified Composite Micromechanical Equations for Hygral, Thermal, and Mechanical Properties," *SAMPE Quarterly*, April.

Chou, T. W., and Yang, J. M., 1986, "Structure-Performance Maps of Polymeric, Metal, and Ceramic Matrix Composites," *Metallurgical Transaction*, Vol. 17A, pp. 1547–1559.

Hsu, P. W., and Herakovich, C. T., 1976, "A Perturbation Solution for Interlaminar Stresses in Bidirectional Laminates," *Comp. Mat. Testing and Design, 4th Conference*, pp. 296–316.

Hsu, P. W., and Herakovich, C. T., 1977, "Edge Effects in Angle-Ply Composite Laminates," *J. of Comp. Mat.*, Vol. 11, pp. 422–428.

Kassapoglou, C., and Lagace, P. A., 1987, "Closed Form Solution for the Interlaminar Stress Field in Angle-Ply and Cross-Ply Laminates," *J. of Comp. Mat.*, Vol. 21, pp. 292–308.

Ozisik, M. N., 1980, *Heat Conduction*, John Wiley and Sons, Inc., New York.

Pipes, R. B., and Pagano, N. J., 1970, "Interlaminar Stresses in Composite Laminates under Uniform Axial Extension," *J. of Comp. Mat.*, Vol. 4, pp. 538–548.

Pipes, R. B., and Pagano, N. J., 1974, "Interlaminar Stresses in Composite Laminates – An Approximate Elasticity Solution," ASME JOURNAL OF APPLIED MECHANICS, Vol. 41, pp. 668–672.

Vinson, J. R., and Chou, T. W., 1975, *Composite Materials and Their Use in Structures*, Applied Science Publishers Ltd., London.

Vinson, J. R., and Sierakowski, R. L., 1986, *The Behavior of Structures Composed of Composite Materials*, Martinus Nijhoff Publishers.

Wang, S. S., and Choi, I., 1979, "Boundary Layer Thermal Stresses in Angle-Ply Composite Laminates," *Modern Development in Composite Materials and Structures*, J. R. Vinson, ed., pp. 315–341.

Wang, S. S., and Choi, I., 1982, "Boundary-Layer Effects in Composite Laminates: Part 1 – Free-Edge Stress Singularities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 541–548.

Wang, S. S., and Choi, I., 1982, "Boundary-Layer Effects in Composite Laminates: Part 2 – Free-Edge Stress Solutions and Basic Characteristics," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 549–560.

Wang, H. S., and Chou, T. W, 1985, "Transient Thermal Stress Analysis of a Rectangular Orthotropic Slab," *J. of Comp. Mat.*, Vol. 19, pp. 424–442.

Wang, H. S., and Chou, T. W., 1986, "Transient Thermal Behavior of a Thermally and Elastically Orthotropic Medium," *AIAA Journal*, Vol. 24, pp. 664–672.

Wang, H. S., Pipes, R. B., and Chou, T. W., 1986, "Thermal Transient Stresses due to Rapid Cooling in a Thermally and Elastically Orthotropic Medium," *Metallurgial Transactions*, Vol. 17A, pp. 1051–1055.

Wang, Y. R., and Chou, T. W., 1987, "3-D Analysis of Transient Interlaminar Thermal Stress of Cross-Ply Composites," *ICCM-6*, Vol. 4, pp. 383–393.

Wang, Y. R., and Chou, T. W., 1988, "3-D Analysis of Transient Interlaminar Thermal Stress of Laminated Composites," *Symposium on Mechanics of Composite Materials*, ASME AMD Vol. 92, pp. 185–192.

Wang, A. S. D., and Crossman, F. W., 1977, "Some New Results on Edge Effect in Symmetric Composite Laminates," *J. of Comp. Mat.*, Vol. 11, pp. 92–106.

Wang, A. S. D., and Crossman, F. W., 1977, "Edge Effects on Thermally Induced Stresses in Composite Laminates," *J. of Comp. Mat.*, Vol. 11, pp. 300–312.

Wang, J. T. S., and Dickson, J. N., 1978, "Interlaminar Stresses in Symmetric Composite Laminates," *J. of Comp. Mat.*, Vol. 12, pp. 390–402.

Wang, S. S., and Yuan, F. G., 1983, "A Hybrid Finite Element Approach to Composite Laminate Elasticity Problems with Singularities," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 835–844.

APPENDIX A

Temperature Field

The heat conduction equation for general anisotropic solid of constant conductivity coefficients without internal heat generation in rectangular coordinate system is (Ozisik, 1980):

$$K_{11}\partial^2 T/\partial x^2 + K_{22}\partial^2 T/\partial y^2 + K_{33}\partial^2 T/\partial z^2 + 2K_{12}\partial^2 T/\partial x\partial y + 2K_{13}\partial^2 T/\partial x\partial z + 2K_{23}\partial^2 T/\partial y\partial z = \rho C_p \partial T/\partial t. \quad (A1)$$

Here, K_{ij} are the coefficients of heat conduction.

We consider an x-direction infinite plate subjected to a temperature field $T = T_o$ on two edges ($y = \pm b$) at time $t = 0^+$ (Figs. 1 and 2) and assume that temperature field in each layer is independent of the thickness direction, i.e., $T = T(y, t)$ only.

The heat conduction equation for each layer is

$$\partial^2 T/\partial y^2 = (1/g^2)\partial T/\partial t \quad (A2)$$

where $g^2 = K_y/\rho c_p$, K_y, ρ and c_p are the coefficient of heat conduction in y-direction, mass density, and specific heat, respectively.

The boundary and initial conditions are

$$T(\pm b, t) = T_o \quad (A3)$$

$$T(y, 0) = 0. \quad (A4)$$

The solution of governing equation (A2) by the method of separation of variables is

$$T = T_o\,(1 + \Sigma B_n \cos\omega_n Y e^{-D_n t}) \quad (A5)$$

where

$$B_n = (-1)^n 4/(2n-1)\pi$$
$$D_n = [(n-1/2)\pi g/b]^2$$
$$\omega_n = (n-1/2)\pi$$
$$Y = y/b.$$

APPENDIX B

Solution of Characteristic Equation

The expansion of the characteristic equation (17) is

$$C_1\lambda^6 + C_2\alpha^2\lambda^4 + C_3\alpha^4\lambda^2 + C_4\alpha^6 = 0 \quad (B1)$$

where

$$C_1 = (Q_{22}Q_{66} - Q_{26}{}^2)Q_{44}$$
$$C_2 = Q_{22}Q_{36}{}^2 + Q_{33}Q_{26}{}^2 + Q_{66}Q_{23}{}^2 + 2Q_{23}Q_{44}Q_{66} - Q_{22}Q_{33}Q_{66} - Q_{22}Q_{44}Q_{55} - 2Q_{23}Q_{26}Q_{36} - 2Q_{44}Q_{26}Q_{36}$$

$$C_3 = Q_{33}Q_{44}Q_{66} + Q_{22}Q_{33}Q_{55} - Q_{44}Q_{36}^2 - Q_{55}Q_{23}^2 - 2Q_{23}Q_{44}Q_{55}$$
$$C_4 = -Q_{33}Q_{44}Q_{55}.$$

Let $\lambda = +\alpha\sqrt{\omega}$, then equation (*B*2) becomes

$$C_1\omega^3 + C_2\omega^2 + C_3\omega + C_4 = 0. \qquad (B2)$$

Let $\omega = \gamma - C_2/(3C_1)$, where $C_1 \neq 0$, then equation (*B*2) becomes

$$\gamma^3 + p\gamma + q = 0 \qquad (B3)$$

where

$$p = C_3/C_1 - C_2^2/(3C_1^2)$$
$$q = (2C_2^3/27C_1^3) - (C_2C_3)/(3C_1^2) + C_4/C_1.$$

Some properties of equation (*B*3) are discussed in the following based upon the discriminant $\Delta = (q/2)^2 + (p/3)^3$:

Case 1: $\Delta > 0$, There exists one real root and two complex roots.

Case 2: $\Delta = 0$, There exist three real roots, two of them are equal.

Case 3: $\Delta < 0$, There exist three unequal real roots.

Let $\gamma = \phi + \psi$, then the characteristic equation (*B*3) could be separated into two equations

$$\phi^3 + \psi^3 = -q$$
$$\phi\psi = -p/3 \qquad (B4)$$

which give the following solutions

$$\phi^3 = [-q + (\pm\sqrt{q^2 + 4p^3/27})]/2$$
$$\psi^3 = [-q - (\pm\sqrt{q^2 + 4p^3/27})]/2. \qquad (B5)$$

APPENDIX C

Rule of Mixtures

The rule-of-mixtures which is used to predict the composite elastic and thermal properties from the fiber and matrix properties is as follows (Chamis, 1984):

$$E_1 = E_{f1}V_f + E_mV_m \quad E_2 = E_3 = E_m/[1 - \sqrt{V_f}(1 - E_m/E_{f2})]$$
$$G_{12} = G_{13} = G_m/[1 - \sqrt{V_f}(1 - G_m/G_{f12})]$$
$$G_{23} = G_m/[1 - \sqrt{V_f}(1 - G_m/G_{f23})]$$
$$\upsilon_{12} = \upsilon_{13} = \upsilon_{f12}V_f + \upsilon_mV_m \quad \upsilon_{23} = \upsilon_{32} = E_2/(2G_{23}) - 1$$
$$\upsilon_{31} = \upsilon_{21} = E_2\upsilon_{12}/E_1 \quad \alpha_1 = (\alpha_{f1}E_{f1}V_f + \alpha_mE_mV_m)/E_1$$
$$\alpha_2 = \alpha_3 = \alpha_{f2}\sqrt{V_f} + (1 - \sqrt{V_f})(1 + V_f\upsilon_mE_{f1}/E_1)\alpha_m$$
$$K_1 = K_{f1}V_f + K_mV_m$$
$$K_2 = K_3 = (1 - \sqrt{V_f})K_m + K_m\sqrt{V_f}/[(1 - \sqrt{V_f})(1 + K_m/K_{f2})]. \qquad (C1)$$

From the basic thermoelastic properties of the composite, equation (*C*1), the elements in the stiffness matrix of stress-strain relationship in the direction of material principal axis are (Vinson and Chou, 1975; Vinson and Sierakowski, 1986)

$$C_{\underline{11}} = (1 - \upsilon_{23}\upsilon_{32})E_1/\Delta \qquad C_{\underline{12}} = (\upsilon_{12} + \upsilon_{32}\upsilon_{13})E_2/\Delta$$
$$C_{\underline{13}} = (\upsilon_{13} + \upsilon_{12}\upsilon_{23})E_3/\Delta \qquad C_{\underline{22}} = (1 - \upsilon_{13}\upsilon_{31})E_2/\Delta$$
$$C_{\underline{23}} = (\upsilon_{23} + \upsilon_{21}\upsilon_{13})E_3/\Delta \qquad C_{\underline{33}} = (1 - \upsilon_{12}\upsilon_{21})E_3/\Delta$$
$$C_{\underline{44}} = G_{23} \qquad C_{\underline{55}} = G_{31} \qquad C_{\underline{66}} = G_{12}. \qquad (C2)$$

Here $\Delta = 1 - \upsilon_{12}\upsilon_{21} - \upsilon_{23}\upsilon_{32} - \upsilon_{13}\upsilon_{31} - 2\upsilon_{21}\upsilon_{32}\upsilon_{13}$ and the underlined subscripts 1, 2, and 3 denote the principal material axis. The angle between the fiber direction and the reference axis is defined as θ. The elastic constants and thermal expansion coefficients in terms of the reference axis are (Vinson and Chou, 1975; Vinson and Sierakowski, 1986)

$$C_{11} = C_{\underline{11}}\cos^4\theta + 2(C_{\underline{12}} + 2C_{\underline{66}})\cos^2\theta\sin^2\theta + C_{\underline{22}}\sin^4\theta$$
$$C_{12} = (C_{\underline{11}} + C_{\underline{22}} - 4C_{\underline{66}})\cos^2\theta\sin^2\theta + C_{\underline{12}}(\cos^4\theta + \sin^4\theta)$$
$$C_{13} = C_{\underline{13}}\cos^2\theta + C_{\underline{23}}\sin^2\theta$$
$$C_{16} = -C_{\underline{22}}\cos\theta\sin^3\theta + C_{\underline{11}}\cos^3\theta\sin\theta - (C_{\underline{12}} + 2C_{\underline{66}})\cos\theta\sin\theta(\cos^2\theta - \sin^2\theta)$$
$$C_{22} = C_{\underline{11}}\sin^4\theta + 2(C_{\underline{12}} + 2C_{\underline{66}})\cos^2\theta\sin^2\theta + C_{\underline{22}}\cos^4\theta$$
$$C_{23} = C_{\underline{13}}\sin^2\theta + C_{\underline{23}}\cos^2\theta$$
$$C_{26} = -C_{\underline{22}}\cos^3\theta\sin\theta + C_{\underline{11}}\cos\theta\sin^3\theta - (C_{\underline{12}} + 2C_{\underline{66}})\cos\theta\sin\theta(\cos^2\theta - \sin^2\theta)$$
$$C_{33} = C_{\underline{33}}$$
$$C_{36} = (C_{\underline{13}} - C_{\underline{23}})\cos\theta\sin\theta$$
$$C_{44} = C_{\underline{44}}\cos^2\theta + C_{\underline{55}}\sin^2\theta$$
$$C_{45} = (C_{\underline{55}} - C_{\underline{44}})\cos\theta\sin\theta$$
$$C_{55} = C_{\underline{55}}\cos^2\theta + C_{\underline{44}}\sin^2\theta$$
$$C_{66} = (C_{\underline{11}} + C_{\underline{22}} - 2C_{\underline{12}})\cos^2\theta\sin^2\theta + C_{\underline{66}}(\cos^2\theta - \sin^2\theta)^2$$
$$\alpha_x = \alpha_1\cos^2\theta + \alpha_2\sin^2\theta$$
$$\alpha_y = \alpha_1\sin^2\theta + \alpha_2\cos^2\theta$$
$$\alpha_z = \alpha_3$$
$$\alpha_{xy} = (\alpha_1 - \alpha_2)\cos\theta\sin\theta. \qquad (C3)$$

The Extended Free Formulation of Finite Elements in Linear Elasticity

Carlos A. Felippa
Department of Aerospace Engineering Sciences
and Center for Space Structures and Controls,
University of Colorado,
Boulder, Colo. 80309-0429

The free formulation of Bergan and Nygård (1984) has been successfully used in the construction of high-performance finite elements for linear and nonlinear structural analysis. In its original form the formulation combines nonconforming internal displacement assumptions with a specialized version of the patch test. The original formulation is limited, however, by strict invertibility conditions linking the assumed displacement field to the nodal displacements. The present paper lifts those restrictions by recasting the free formulation within the framework of a mixed-hybrid functional that allows internal stresses, internal displacements, and boundary displacements to vary independently. This functional contains a free parameter and includes the potential energy and the Hellinger-Reissner principles as special cases. The parameter appears in the higher-order stiffness of the element equations.

1 Introduction

Bergan and Nygård (1984) have developed the so-called free formulation (FF) for the construction of displacement-based incompatible finite elements. This work consolidated a decade of research of Bergan and co-workers at Trondheim, milestones of which may be found in Bergan and Hanssen (1976), Hanssen et al. (1979), and Bergan (1980). The products of this research have been finite elements of high performance, especially for plates and shells. Linear applications are reported in the aforementioned papers as well as in Bergan and Wang (1984), Bergan and Felippa (1985), and Felippa and Bergan (1987); whereas nonlinear applications are presented in Bergan and Nygård (1985) and Nygård (1986). By "high performance" it is meant that solution of engineering accuracy can be obtained with coarse meshes of simple elements, and that those elements exhibit low distortion sensitivity.

The original FF was based on nonconforming displacement assumptions, the principle of virtual work and a specialized form of Irons' patch test that Bergan and Hanssen (1976) called the individual element test. A key ingredient of the FF is the separation of the element stiffness matrix into the sum of two parts, called *basic* and *higher-order* stiffness, respectively. The basic part is constructed for convergence and the higher order part for numerical stability and (in recent work) accuracy.

An intriguing question has been: Does the FF fit in a variational framework? This was partly answered by Bergan and Felippa (1985), who showed that the basic stiffness part was equivalent to a constant-stress hybrid element. But persistent efforts by the present author to encompass the higher-order stiffness within a hybrid variational principle were unsuccessful until the development of parametrized mixed-hybrid functionals in Felippa (1989a, 1989b). With the help of these more general functionals it is possible to show that the FF is a very special type of mixed-hybrid element which does not fit within the classical Hellinger-Reissner principle. In retrospect, the classification of FF elements as hybrids is not surprising. Under mild conditions studied in the Appendix, hybrid elements satisfy Irons' patch test *a priori*, and the FF development has been founded on that premise.

To encompass the FF within the hybrid framework, the following assumptions must be invoked.

(1) A specific hybrid functional, identified as Π_γ^d in the sequel, is constructed. This functional depends linearly on a parameter γ.

(2) Three fields are assumed over each element:
   (*a*) a constant stress field,
   (*b*) an internal displacement field $\mathbf{u}$ defined by n_q generalized coordinates collected in vector $\mathbf{q}$, and
   (*c*) a boundary displacement field $\mathbf{d}$ defined by n_v nodal displacements collected in vector $\mathbf{v}$. Both $\mathbf{d}$ and $\mathbf{u}$ must represent rigid body motions and constant strain states exactly.

(3) The number of generalized coordinates, n_q, equals the number of nodal displacements, n_v, and the square transformation matrix $\mathbf{G}$ relating $\mathbf{v} = \mathbf{Gq}$ is nonsingular.

In Felippa (1989b) it is shown that substituting the finite element expansions into Π_γ^d, rendering the functional stationary with respect to the degrees-of-freedom, and eliminating both internal fields by a combination of static condensation and kinematic constraints, leads to the FF stiffness equations in terms of the nodal displacements $\mathbf{v}$. The parameter γ appears as a coefficient of the higher-order stiffness. These stiffness

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10-15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 6, 1988; final revision, December 29, 1988.
Paper No. 89-WA/APM-12.

equations can be readily implemented into any displacement-based finite element code.

This variational pathway to FF is of interest for two reasons. First, it explains the behavior of FF elements as regards convergence, stability, and accuracy. Second, it opens up the door to extensions that are not obvious from a physical standpoint. Two such extensions involve: retaining higher-order stress fields, and allowing more internal displacement modes than nodal displacements, that is, $n_q > n_v$. The main purpose of this paper is to study these two extensions, which are shown to be closely related. The resulting framework for deriving finite elements in elasticity is called the extended free formulation (EFF).

2 Governing Equations

Consider a *linearly elastic body* under static loading that occupies the volume V. The body is bounded by the surface S, which is decomposed into S: $S_d \cup S_t$. Displacements are prescribed on S_d whereas surface tractions are prescribed on S_t. The outward unit normal on S is denoted by $\mathbf{n} \equiv n_i$.

The three unknown volume fields are displacements $\mathbf{u} \neq u_i$, infinitesimal strains $\mathbf{e} \equiv e_{ij}$, and stresses $\boldsymbol{\sigma} \equiv \sigma_{ij}$. The problem data include: the body force field $\mathbf{b} \equiv b_i$ in V, prescribed displacements $\hat{\mathbf{d}}$ on S_d, and prescribed surface tractions $\hat{\mathbf{t}} \equiv \hat{t}_i$ on S_t.

The relations between the volume fields are the strain-displacement equations

$$\mathbf{e} = \tfrac{1}{2}(\nabla \mathbf{u} + \nabla^T \mathbf{u}) = \mathbf{D}\mathbf{u} \text{ or } e_j \tfrac{1}{2}(u_{i,j} + u_{j,i}) \text{ in } V, \tag{1}$$

the constitutive equations

$$\boldsymbol{\sigma} = \mathbf{E}\mathbf{e} \text{ or } \sigma_{ij} = E_{ijkl} e_{kl} \text{ in } V, \tag{2}$$

and the equilibrium (balance) equations

$$-\mathbf{div}\ \boldsymbol{\sigma} = \mathbf{D}^* \boldsymbol{\sigma} = \mathbf{b} \text{ or } \sigma_{ij,j} + b_i = 0 \text{ in } V, \tag{3}$$

in which $\mathbf{D}^* = -\mathbf{div}$ denotes the adjoint operator of $\mathbf{D} = 1/2\ (\nabla + \nabla^T)$.

The stress vector with respect to a direction defined by the unit vector $\mathbf{v}$ is denoted as $\boldsymbol{\sigma}_v = \boldsymbol{\sigma}.\mathbf{v}$, or $\sigma_{vi} = \sigma_{ij} v_j$. On S the surface-traction stress vector is defined as

$$\boldsymbol{\sigma}_n = \boldsymbol{\sigma}.\mathbf{n}, \text{ or } \sigma_{ni} = \sigma_{ij} n_j. \tag{4}$$

With this definition the traction boundary conditions may be stated as

$$\boldsymbol{\sigma}_n = \hat{\mathbf{t}} \text{ or } \sigma_{ij} n_j = \hat{t}_i \text{ on } S_t, \tag{5}$$

and the displacement boundary conditions as

$$\mathbf{u} = \hat{\mathbf{d}} \text{ or } u_i = \hat{d}_i \text{ on } S_d. \tag{6}$$

3 Notation

Field Dependency. In variational methods of approximation we do not work, of course, with the exact fields that satisfy the governing equations (1)–(3), (5)–(6), but with *independent* (primary) fields, which are subject to variations, and *dependent* (secondary, associated, derived) fields, which are not. The approximation is determined by taking variations with respect to the independent fields.

An *independently varied* field will be identified by a superposed tilde, for example, $\tilde{\mathbf{u}}$. A dependent field is identified by writing the independent field symbol as superscript. For example, if the displacements are independently varied, the derived strain and stress fields are

$$\mathbf{e}^u \tfrac{1}{2}(\nabla + \nabla^T)\tilde{\mathbf{u}} = \mathbf{D}\tilde{\mathbf{u}},\ \boldsymbol{\sigma}^u = \mathbf{E}\mathbf{e}^u = \mathbf{E}\mathbf{D}\tilde{\mathbf{u}}. \tag{7}$$

An advantage of this convention is that $\mathbf{u}$, $\mathbf{e}$, and $\boldsymbol{\sigma}$ may be reserved for the *exact* fields.

Integral Abbreviations. Volume and surface integrals will be abbreviated by placing domain-subscripted parentheses and square brackets, respectively, around the integrand. For example:

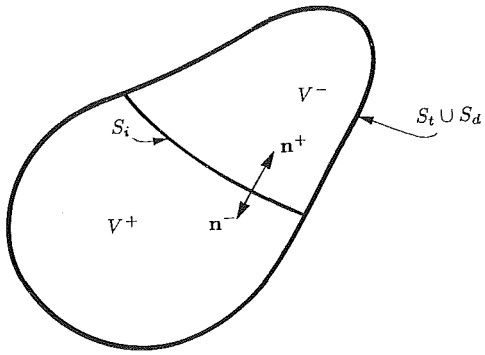


Fig. 1 Internal interface example

$$(f)_V \overset{\text{def}}{=} \int_V f\, dV,\ [f]_S \overset{\text{def}}{=} \int_S f\, dS,\ [f]_{S_d} \overset{\text{def}}{=}$$

$$\int_{S^{\hat{d}}} f\, dS,\ [f]_{S_t} \overset{\text{def}}{=} \int_{S_t} f\, dS. \tag{8}$$

If $\mathbf{f}$ and $\mathbf{g}$ are vector functions, and $\mathbf{p}$ and $\mathbf{q}$ tensor functions, their inner product over V is denoted in the usual manner

$$(\mathbf{f},\mathbf{g})_V \overset{\text{def}}{=} \int_V \mathbf{f}.\mathbf{g} dV = \int_V f_i g_i dV,$$

$$(\mathbf{p},\mathbf{q})_V \overset{\text{def}}{=} \int_V \mathbf{p}.\mathbf{q} dV = \int_V p_{ij} q_{ij} dV, \tag{9}$$

and, similarly, for surface integrals, in which case square brackets are used.

Domain Assertions. The notation

$$(a = b)_V,\ [a = b]_S,\ [a = b]_{S_d},\ [a = b]_{S_t}, \tag{10}$$

is used to assert that the relation $a = b$ is valid at each point of V, S, S_d, and S_t, respectively.

Internal Interfaces. In the following subsections a variational principle is constructed, in which boundary displacements $\mathbf{d}$ can be varied independently from the internal displacements $\mathbf{u}$. These displacements play the role of Lagrange multipliers that relax internal displacement continuity. Variational principles of this form will be called *displacement-generalized*, or *d*-generalized for short.

The choice of $\mathbf{d}$ as independent field is *not* variationally admissible on S_d or S_t. We must therefore extend the definition of boundary to include *internal interfaces* collectively designated as S_i. Thus,

$$S : S_d \cup S_t \cup S_i. \tag{11}$$

On S_i neither displacements nor tractions are prescribed. A simple case is illustrated in Fig. 1, in which the interface S_i divides V into two subvolumes: V^+ *and* V^-. An interface such as S_i on Fig. 1 has two "sides" called S_i^+ and S_i^-, which identify S_i viewed as boundary of V^+ and V^-, respectively. At smooth points of S_i, the unit normals $\mathbf{n}^+$ and $\mathbf{n}^-$ point in opposite directions.

The integral abbreviations (8)–(9) generalize as follows, using Fig. 1 for definiteness. A volume integral is the sum of integrals over the subvolumes:

$$(f)_V \overset{\text{def}}{=} \int_{V^+} f\, dV + \int_{V^-} f\, dV. \tag{12}$$

An integral over S_i includes two contributions:

$$[g]_{S_i} \overset{\text{def}}{=} \int_{S_i^+} g^+ \, dS + \int_{S_i^-} g^- \, dS, \tag{13}$$

where g^+ and g^- denotes the value of the integrand g on S_i^+ and S_i^-, respectively. These two values may be different if g is discontinuous or involves a projection on the normals.

4 The Hu-Washizu Principle

There are several essentially equivalent statements of the Hu-Washizu functional of linear elasticity. The starting form used here is the four-field functional

$$\Pi_W^d(\tilde{\mathbf{u}},\tilde{\mathbf{e}},\tilde{\sigma},\tilde{\mathbf{d}}) = \tfrac{1}{2}(\sigma^e,\tilde{\mathbf{e}})_V + (\tilde{\sigma},\mathbf{e}^u - \tilde{\mathbf{e}})_V - P^d, \tag{14}$$

where P^d is the "forcing" potential

$$P^d(\tilde{\mathbf{u}},\tilde{\sigma},\tilde{\mathbf{d}}) = (\mathbf{b},\tilde{\mathbf{u}})_V + [\tilde{\sigma}_n,\tilde{\mathbf{u}} - \hat{\mathbf{d}}]_{S_d} + [\hat{\mathbf{t}},\tilde{\mathbf{u}}]_{S_t} + [\tilde{\sigma}_n,\tilde{\mathbf{u}} - \tilde{\mathbf{d}}]_{S_i}. \tag{15}$$

The function Π_W^t is called *d*-generalized in the sense that the volume fields $\tilde{\mathbf{u}}$, $\tilde{\mathbf{e}}$, $\tilde{\sigma}$, *and* the surface displacement field $\tilde{\mathbf{d}}$ are subject to independent variations, whereas in the conventional form of the principle the relation $[\mathbf{d} = \tilde{\mathbf{u}}]_{S_d \cup S_i}$ is enforced *a priori*. The superscript d is used to distinguish it from the *t*-generalized variant

$$\Pi_W^t(\tilde{\mathbf{u}},\tilde{\mathbf{e}},\tilde{\sigma},\tilde{\mathbf{t}}) = \tfrac{1}{2}(\sigma^e,\tilde{\mathbf{e}})_V + (\tilde{\sigma},\mathbf{e}^u - \tilde{\mathbf{e}})_V - P^t, \tag{16}$$

in which the surface tractions $\tilde{\mathbf{t}}$ are varied independently from the internal stress field $\tilde{\sigma}$. This is the starting form in the classical textbook of Washizu (1968). Parametrized versions of (16) are studied in further detail in Felippa (1989a).

Functionals that are not d or t-generalized will be called *conventional*. The three versions differ only in the forcing potential term.

5 Parametrization

Constraining the Hu-Washizu functional (14) by selectively enforcing field equations and boundary conditions *a priori* yields six functionals listed (in their conventional form) in Chapter 4 of the monograph of Oden and Reddy (1983). Of particular interest for the present study are the *d*-generalized Hellinger-Reissner functional

$$\Pi_R^d(\tilde{\mathbf{u}},\tilde{\sigma},\tilde{\mathbf{d}}) = -\tfrac{1}{2}(\tilde{\sigma},\mathbf{e}^\sigma)_V + (\tilde{\sigma},\mathbf{e}^u)_V - P^d, \tag{17}$$

as well as the *d*-generalized potential energy functional

$$\Pi_P^d(\tilde{\mathbf{u}},\tilde{\mathbf{d}}) = \tfrac{1}{2}(\sigma^u,\mathbf{e}^u)_V - P^d. \tag{18}$$

These two functionals are special cases of the following parametrized form

$$\Pi_\gamma^d(\tilde{\mathbf{u}},\tilde{\sigma},\tilde{\mathbf{d}}) = \tfrac{1}{2}(1-\gamma)(\sigma^u,\mathbf{e}^u)_V - \tfrac{1}{2}\gamma(\tilde{\sigma},\mathbf{e}^\sigma)_V + \gamma(\tilde{\sigma},\mathbf{e}^u)_V - P^d, \tag{19}$$

where γ is a scalar. If $\gamma = 1$ and 0 we obtain the functionals Π_R^d and Π_P^d, respectively. Parametrized forms, such as (19), of the elasticity variational principles were studied by Chien (1983).

First Variation. Defining the γ-weighted stresses

$$\sigma^\gamma \overset{\text{def}}{=} \gamma\tilde{\sigma} + (1-\gamma)\sigma^u \text{ in } V, \quad \sigma_n^\gamma \overset{\text{def}}{=} \gamma\tilde{\sigma}_n + (1-\gamma)\sigma_n^u \text{ on } S \tag{20}$$

the first variation of (19) can be written

$$\delta\Pi_\gamma^d = \gamma(\mathbf{e}^u - \mathbf{e}^\sigma,\delta\tilde{\sigma})_V - (\mathbf{div}\ \sigma^\gamma + \mathbf{b},\ \delta\tilde{\mathbf{u}})_V - [\hat{\mathbf{t}} - \sigma_n^\gamma,\delta\tilde{\mathbf{u}}]_{S_t} - [\tilde{\sigma}_n - \tilde{\sigma}_n^\gamma,\delta\tilde{\mathbf{u}}]_{S_d} - [\mathbf{u} - \hat{\mathbf{d}},\delta\tilde{\sigma}_n]_{S_d} - [\tilde{\sigma}_n - \tilde{\sigma}_n^\gamma,\delta\tilde{\mathbf{u}}]_{S_i} - [\tilde{\mathbf{u}} - \tilde{\mathbf{d}},\delta\tilde{\sigma}_n]_{S_i} - [\tilde{\sigma}_n,\delta\tilde{\mathbf{d}}]_{S_i}. \tag{21}$$

Since $\mathbf{d}$ is unique on S_i whereas $\tilde{\mathbf{u}}$ and $\tilde{\sigma}$ are generally discontinuous on it, the interface integrals in (21) split as follows:

$$[\tilde{\sigma}_n - \tilde{\sigma}_n^\gamma,\delta\tilde{\mathbf{u}}]_{S_i} = [\tilde{\sigma}_n^+ - \tilde{\sigma}_n^{\gamma+},\delta\tilde{\mathbf{u}}^+]_{S_i+} + [\tilde{\sigma}_n^- - \tilde{\sigma}_n^{\gamma-},\delta\tilde{\mathbf{u}}^-]_{S_i^-},$$
$$[\tilde{\mathbf{u}} - \tilde{\mathbf{d}},\delta\tilde{\sigma}]_{S_i} = [\tilde{\mathbf{u}}^+ - \tilde{\mathbf{d}},\delta\tilde{\sigma}_n^+]_{S_i^+} + [\tilde{\mathbf{u}}^- - \tilde{\mathbf{d}},\delta\tilde{\sigma}_n^-]_{S_i^-}, \tag{22}$$
$$[\tilde{\sigma}_n,\delta\tilde{\mathbf{d}}]_{S_i} = [\tilde{\sigma}_n^+,\delta\tilde{\mathbf{d}}]_{S_i^+} + [\tilde{\sigma}_n^-,\delta\tilde{\mathbf{d}}]_{S_i^-} = [\tilde{\sigma}_n^+ - \tilde{\sigma}_n^-,\delta\tilde{\mathbf{d}}]_{S_i}.$$

Setting the first variation to zero and taking (22) into account, the Euler equations and natural boundary conditions for $\gamma \neq 0$ are found to be

$$(\mathbf{e}^u = \mathbf{e}^\sigma)_V,\ (\mathbf{div}\ \sigma^\gamma + \mathbf{b} = 0)_V,\ [\sigma_n^\gamma = \hat{\mathbf{t}}]_{S_t},$$
$$[\sigma_n = \sigma_n^\gamma]_{S_d},\ [\mathbf{u} = \hat{\mathbf{d}}]_{S_d},\ [\sigma_n^{\gamma+} - \sigma_n^+ = \mathbf{0}]_{S_i}, \tag{23}$$
$$[\sigma_n^{\gamma-} - \sigma_n^- = \mathbf{0}]_{S_i},\ [\mathbf{u}^+ = \mathbf{u}^- = \mathbf{d}]_{S_i},\ [\sigma_n^+ - \sigma_n^- = \mathbf{0}]_{S_i}.$$

The constitutive equations do not appear since they are enforced *a priori* in Π_γ^d. If $\gamma = 0$, the first equation $(\mathbf{e}^u = \mathbf{e}^\sigma)_V$, drops out.

Modified Forcing Potential. Substituting $\mathbf{d}$ in lieu of $\mathbf{u}$ in the forcing potential (15)

$$P^d(\tilde{\mathbf{u}},\tilde{\sigma},\tilde{\mathbf{d}}) = (\mathbf{b},\tilde{\mathbf{u}})_V + [\tilde{\sigma}_n,\tilde{\mathbf{d}} - \hat{\mathbf{d}}]_{S_d} + [\hat{\mathbf{t}},\tilde{\mathbf{d}}]_{S_t} + [\tilde{\sigma}_n,\tilde{\mathbf{u}} - \tilde{\mathbf{d}}]_{S_i} \tag{24}$$

is *not* variationally admissible because incorrect Euler equations result. A correct potential that resembles (24) can be obtained in two stages. First, surface terms $[\tilde{\sigma}_n,\ \tilde{\mathbf{u}} - \tilde{\mathbf{d}}]_{S_t}$ and $[\tilde{\sigma}_n,\ \tilde{\mathbf{u}} - \tilde{\mathbf{d}}]_{S_t}$ are added and subtracted to produce

$$P^d(\tilde{\mathbf{u}},\tilde{\sigma},\tilde{\mathbf{d}}) = (\mathbf{b},\tilde{\mathbf{u}})_V + [\tilde{\sigma}_n,\tilde{\mathbf{d}} - \hat{\mathbf{d}}]_{S_d} - [\tilde{\sigma}_n - \hat{\mathbf{t}},\tilde{\mathbf{u}}]_{S_t} + [\tilde{\sigma}_n,\tilde{\mathbf{u}} - \tilde{\mathbf{d}}]_S. \tag{25}$$

Second, $\hat{\mathbf{t}}$ is assumed to be in the range of $\tilde{\sigma}_n$ and the condition $[\tilde{\sigma}_n = \hat{\mathbf{t}}]_{S_t}$ satisfied *a priori*, reducing (25) to

$$P^d(\tilde{\mathbf{u}},\tilde{\sigma},\tilde{\mathbf{d}}) = (\mathbf{b},\ \tilde{\mathbf{u}})_V + [\tilde{\sigma}_n,\tilde{\mathbf{d}} - \hat{\mathbf{d}}]_{S_d} + [\hat{\mathbf{t}},\tilde{\mathbf{d}}]_{S_t} + [\tilde{\sigma}_n,\tilde{\mathbf{u}} - \tilde{\mathbf{d}}]_S. \tag{26}$$

This expression differs from (24) in that the all-important surface dislocation integral is taken over S rather than S_i. Further simplification results if the displacement boundary conditions $[\tilde{\mathbf{d}} = \hat{\mathbf{d}}]_{S_d}$ are exactly satisfied:

$$P^d(\tilde{\mathbf{u}},\tilde{\sigma},\tilde{\mathbf{d}}) = (\mathbf{b},\tilde{\mathbf{u}})_V + [\hat{\mathbf{t}},\tilde{\mathbf{d}}]_{S_t} + [\tilde{\sigma}_n,\tilde{\mathbf{u}} - \tilde{\mathbf{d}}]_S. \tag{27}$$

This expression of P^d is used in the sequel, as modifications required to account for the case $[\tilde{\mathbf{d}} \neq \hat{\mathbf{d}}]_{S_d}$ are of minor importance.

6 Energy Balancing

Distances. Let $U(\epsilon) = 1/2\,(\mathbf{E}\epsilon,\epsilon)_V$ denote the strain energy associated with field ϵ. We may rewrite (19) as a potential-energy deviator

$$\Pi_\gamma^d = \Pi_P^d - \gamma U(\mathbf{e}^u - \mathbf{e}^\sigma), \tag{28}$$

because

$$\frac{\Pi_\gamma^d - \Pi_P^d}{\gamma/2} = (\tilde{\sigma},\mathbf{e}^\sigma - \mathbf{e}^u)_V - (\tilde{\sigma} - \sigma^u,\mathbf{e}^u)_V$$
$$= (\sigma^u - \tilde{\sigma},\mathbf{e}^u - \mathbf{e}^\sigma)_V = (\mathbf{E}\mathbf{e}^u - \mathbf{E}\mathbf{e}^\sigma,\mathbf{e}^u - \mathbf{e}^\sigma)_V. \tag{29}$$

If $\mathbf{E}$ is positive definite, $U(\mathbf{e}^u - \mathbf{e}^\sigma) \geq 0$ and, consequently,

$$\Pi_\gamma^d \leq \Pi_P^d \text{ if } \gamma > 0. \tag{30}$$

If $\tilde{\mathbf{u}}$ is kinematically admissible, Π_P^d exceeds the exact potential energy as will be shown. It follows that to improve solutions in

energy, we expect to take $\gamma \geq 0$. Thus, principles associated with $\gamma < 0$ have limited practical interest.

Let $\Pi(\mathbf{u})$ denote the exact potential energy

$$\Pi(\mathbf{u}) = \tfrac{1}{2}(\sigma,\mathbf{e})_V - (\mathbf{b},\mathbf{u})_V - [\hat{\mathbf{t}},\mathbf{u}]_{S_t}, \tag{31}$$

where σ and $\mathbf{e}$ denotes the exact stress and strain field, respectively. If $\tilde{\mathbf{u}}$ is kinematically admissible and thus satisfies $[\tilde{\mathbf{u}} = \hat{\mathbf{d}}]_{S_d}$, then the energy distance from $\Pi_P^d(\tilde{\mathbf{u}})$ to the exact functional (31) is (see, e.g., Section 34 of Gurtin (1972))

$$\Pi_P^d - \Pi = \tfrac{1}{2}(\sigma^u - \sigma, \mathbf{e}^u - \mathbf{e})_V = U(\mathbf{e}^u - \mathbf{e}). \tag{32}$$

Adjusting γ. To derive an "energy balanced" approximation we impose the condition $\Pi_\gamma^d = \Pi$, which yields

$$\gamma_b = \frac{U(\mathbf{e}^u - \mathbf{e})}{U(\mathbf{e}^u - \mathbf{e}^\sigma)} = \frac{(\sigma^u - \sigma, \mathbf{e}^u - \mathbf{e})}{(\sigma^u - \tilde{\sigma}, \mathbf{e}^u - \mathbf{e}^\sigma)}. \tag{33}$$

For example, if we assume that the exact stresses and strains lie halfway between the approximate fields,

$$\sigma = \tfrac{1}{2}(\sigma^u + \tilde{\sigma}), \quad \mathbf{e} = \tfrac{1}{2}(\mathbf{e}^\sigma + \mathbf{e}^u), \tag{34}$$

then $\gamma_b = 1/4$. But, as the exact stresses and strains for the elasticity problem are not generally known in advance, the practical determination of γ_b has been based on application of (33) to element "patches" under simple load systems, as discussed in Bergan and Felippa (1985) and Felippa and Bergan (1987).

Error Estimates. The strain difference $\mathbf{e}^u - \mathbf{e}^\sigma$ may be used as a pointwise measure of solution accuracy, and the associated "dislocation work" $U(\mathbf{e}^u - \mathbf{e}^\sigma)$ as an energy error measure for applications such as adaptive mesh refinement.

7 Finite Element Discretization

In this section the finite element discretization of Π_γ^d is studied. Following usual practice in finite element work, the components of stresses and strains are arranged as one-dimensional arrays whereas the elastic moduli in $\mathbf{E}$ are arranged as a square symmetric matrix. The FE assumption is globally written

$$(\tilde{\mathbf{u}} = \mathbf{N}\mathbf{q})_V, \quad (\tilde{\sigma} = \mathbf{A}\mathbf{a})_V, \quad [\tilde{\mathbf{d}} = \mathbf{V}\mathbf{v}]_S. \tag{35}$$

Here, matrices $\mathbf{N}$, $\mathbf{A}$, and $\mathbf{V}$ collect generalized-displacement shape functions, internal stress modes, and interface displacement modes, respectively, whereas column vectors $\mathbf{q}$, $\mathbf{a}$, and $\mathbf{v}$ collect generalized internal displacements, stress mode amplitudes, and generalized interface displacements, respectively. The assumed volume fields $\tilde{\sigma}$ and $\tilde{\mathbf{u}}$ need not be continuous across S_i. The derived fields are

$$(\mathbf{e}^u = \mathbf{D}\mathbf{N}\mathbf{q} = \mathbf{B}\mathbf{q})_V, \quad (\sigma^u = \mathbf{E}\mathbf{B}\mathbf{q})_V, \quad (\mathbf{e}^\sigma = \mathbf{E}^{-1}\tilde{\sigma} = \mathbf{E}^{-1}\mathbf{A}\mathbf{a})_V. \tag{36}$$

Inserting these expressions into Π_γ^d with the forcing potential (27), we obtain the algebraic form

$$\Pi_\gamma^d(\mathbf{a},\mathbf{q},\mathbf{s}) = \tfrac{1}{2}(1-\gamma)\mathbf{q}^T\mathbf{K}_u\mathbf{q} - \tfrac{1}{2}\gamma\mathbf{a}^T\mathbf{C}\mathbf{a} + \gamma\mathbf{q}^T\mathbf{Q}\mathbf{a} - \mathbf{q}^T\mathbf{P}\mathbf{a} + \mathbf{v}^T\mathbf{L}\mathbf{a} - \mathbf{q}^T\mathbf{f}_q - \mathbf{v}^T\mathbf{f}_v \tag{37}$$

where

$$\mathbf{K}_u = (\mathbf{B}^T\mathbf{E}\mathbf{B})_V = \mathbf{K}_u^T, \ \mathbf{C} = (\mathbf{A}^t\mathbf{E}^{-1}\mathbf{A})_V = \mathbf{C}^T, \ \mathbf{Q} = (\mathbf{B}^T\mathbf{A})_V, \quad \mathbf{L} = [\mathbf{V}^T\mathbf{A}_n]_S, \ \mathbf{P} = [\mathbf{N}^T\mathbf{A}_n]_S, \ \mathbf{f}_q = (\mathbf{N}^T\mathbf{b})_V, \ \mathbf{f}_v = [\mathbf{N}^T\hat{\mathbf{t}}]_{S_t}. \tag{38}$$

The matrices $\mathbf{K}_u$, $\mathbf{C}$, $\mathbf{Q}$, $\mathbf{L}$, and $\mathbf{P}$ are called internal-displacement stiffness, compliance, leverage, nodal-force lumping, and boundary dislocation matrices, respectively. Making (37) stationary yields the linear system

$$\begin{bmatrix} -\gamma\mathbf{C} & \gamma\mathbf{Q}^T - \mathbf{P}^T & \mathbf{L}^T \\ \gamma\mathbf{Q} - \mathbf{P} & (1-\gamma)\mathbf{K}_u & 0 \\ \mathbf{L} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{a} \\ \mathbf{q} \\ \mathbf{v} \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{f}_q \\ \mathbf{f}_v \end{Bmatrix}. \tag{39}$$

The first matrix equation is the discrete analog of the first, fifth, and eighth relations in (24), and expresses internal and boundary compatibility. The third matrix equation is the discrete analog of the last relation, and expresses equilibrium across S_i. The second matrix equation is the discrete analog of the remaining relations, and expresses internal and boundary equilibrium.

It is shown later (in Section 9) that if the assumed stress modes in $\mathbf{A}$ are *divergence free* (self-equilibrating), then $\mathbf{P} \equiv \mathbf{Q}$, and (39) simplifies to

$$\begin{bmatrix} -\gamma\mathbf{C} & -(1-\gamma)\mathbf{Q}^T & \mathbf{L}^T \\ -(1-\gamma)\mathbf{Q} & (1-\gamma)\mathbf{K}_u & \mathbf{0} \\ \mathbf{L} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{a} \\ \mathbf{q} \\ \mathbf{v} \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{f}_q \\ \mathbf{f}_v \end{Bmatrix}. \tag{40}$$

These results are now reinterpreted in terms of hybrid elements.

8 Hybrid Elements

Approach. The preceding treatment is relevant to the construction of *displacement-connected hybrid elements*. Hybrid elements based on more restricted assumptions were originally constructed by Pian and co-workers (see Pian, 1964; Pian and Tong, 1969; Pian, 1973). From current perspective, the principal features of the hybrid formulation are:

(*A*) The domain is subdivided into elements *before* the variational principle is established.

(*B*) Continuity requirements across element boundaries are relaxed by introducing boundary tractions or boundary displacements as Lagrange multiplier fields.

(*C*) All stress and internal-displacement degrees-of-freedom are eliminated (by either static condensation or kinematic constraints) at the *element level.*

(*A*) says that hybrid functionals are effectively *mesh-dependent*, because the domain subdivision process introduces element boundaries which must be treated as *internal interfaces*, and therefore become part of S_i. Previous developments remain valid if one reinterprets "body" as "individual element," "volume" as "element volume," and "surface" as "interelement boundary."

Continuity and Connectors. The internal fields $\tilde{\sigma}$ and $\tilde{\mathbf{u}}$ may be discontinuous across elements. The boundary displacement field $\tilde{\mathbf{d}}$, however, must be continuous on S_i, i.e., it must have the same value on adjacent elements. This condition may be satisfied if $\tilde{\mathbf{d}}$ on an interface separating two elements *is uniquely interpolated by nodal values on that interface*. It is natural to take such nodal values as entries of $\mathbf{v}$, which automatically becomes the vector of *connected node displacements* or *connectors*.

9 Kinematic Relations

In this and subsequent sections we work with an *individual element* unless otherwise noted. The element volume is V and the element surface is S: $S_d \cup S_t \cup S_i$. The $\mathbf{v}$ subvector contains n_v element-connector degrees-of-freedom, whereas $\mathbf{q}$ and $\mathbf{a}$ contain n_q and n_a internal freedoms, respectively. We shall assume that $n_q \geq n_v$.

The first matrix equation (the discrete compatibility equation) in (39) can be interpreted as the dislocation-energy balance statement

$$\frac{1}{2}\gamma(\tilde{\sigma},\mathbf{e}^u-\mathbf{e}^\sigma)_V-\mathbf{a}^T(\mathbf{P}^T\mathbf{q}-\mathbf{L}^T\mathbf{v})_V=0. \tag{41}$$

Setting $\gamma=0$ and observing that $\mathbf{a}$ is arbitrary, (41) forces the kinematic constraint

$$\mathbf{P}^T\mathbf{q}=\mathbf{L}^T\mathbf{v} \tag{42}$$

to be satisfied. The same relation emerges if $\gamma\neq 0$ but the element displacements are forced to obey

$$(\tilde{\sigma},\mathbf{e}^u-\mathbf{e}^\sigma)_V=0 \tag{43}$$

as an *optimality* condition which says that the work of the strain error over the assumed stress field vanishes for arbitrary element motions. The constraint (42) plays a key role in subsequent derivations. An immediate consequence is that the first matrix equation in (39) reduces to the equivalent of (43), namely $\gamma\mathbf{a}^T(-\mathbf{C}\mathbf{a}+\mathbf{Q}^T\mathbf{q})=\mathbf{0}$, thus, if $\gamma\neq 0$,

$$\mathbf{a}=\mathbf{C}^{-1}\mathbf{Q}^T\mathbf{q},\ \text{or}\ \mathbf{a}=\mathbf{C}^{-1}\mathbf{L}^T\mathbf{v}\ \text{if}\ \mathbf{P}\equiv\mathbf{Q}. \tag{44}$$

Next, suppose that $\mathbf{q}$ and $\mathbf{v}$ are connected by the linear algebraic relations

$$\mathbf{v}=\mathbf{G}\mathbf{q}, \tag{45}$$

$$\mathbf{q}=\mathbf{H}\mathbf{v}, \tag{46}$$

where $\mathbf{G}$ is a $\mathbf{n_v}\times\mathbf{n_q}$ transformation matrix and $\mathbf{H}$ is a $n_q\times n_v$ transformation matrix. The determination of these matrices and their connecting relationships is discussed later. Using (45)–(46) the constraint (42) may be stated in two ways:

$$\mathbf{P}^T=\mathbf{L}^T\mathbf{G},\ \mathbf{P}^t\mathbf{H}=\mathbf{L}^T. \tag{47}$$

Elimination of $\mathbf{a}$ and $\mathbf{q}$ in (39) through (44)–(46), with account taken of the second of (47), yields the external stiffness equations

$$\mathbf{K}\mathbf{v}=\mathbf{f}, \tag{48}$$

in which

$$\mathbf{K}=\gamma[\mathbf{L}\mathbf{C}^{-1}\mathbf{Q}^T\mathbf{H}+\mathbf{H}^T\mathbf{Q}\mathbf{C}^{-1}\mathbf{L}^T-\mathbf{L}\mathbf{C}^{-1}\mathbf{L}^T]+(1-\gamma)\mathbf{H}^T\mathbf{K}_u\mathbf{H},\ \mathbf{f}=\mathbf{f}_v+\mathbf{H}^T\mathbf{f}_q. \tag{49}$$

If $\mathbf{P}\equiv\mathbf{Q}$, system (40) reduces to (48) but with

$$\mathbf{K}=\gamma\mathbf{L}\mathbf{C}^{-1}\mathbf{L}^T+(1-\gamma)\mathbf{H}^T\mathbf{K}_u\mathbf{H}. \tag{50}$$

10 Internal Field Decomposition

To gain further insight into the structure of the element stiffness equations (48) and eventually link up with the free formulation, we proceed to decompose both internal element fields as follows.

Stress Decomposition. The assumed stress field, $\tilde{\sigma}$, is decomposed into a mean value, $\bar{\sigma}$, and a deviator:

$$\tilde{\sigma}=\bar{\sigma}+\tilde{\sigma}_h=\bar{\sigma}+\mathbf{A}_h\mathbf{a}_h, \tag{51}$$

in which

$$\bar{\sigma}=(\tilde{\sigma})_V/v,\ (\mathbf{A}_h)_V=\mathbf{0}, \tag{52}$$

where $v=(1)_V$ denotes the element volume measure. The second relation in (52) is obtained by integrating (51) over V and noting that $\mathbf{a}_h$ is arbitrary.

Internal Displacement Decomposition. Next, the $\tilde{\mathbf{u}}$ assumption is decomposed into rigid body, constant strain, and higher-order displacements:

$$\tilde{\mathbf{u}}=\mathbf{N}_r\mathbf{q}_r+\mathbf{N}_c\mathbf{q}_c+\mathbf{N}_h\mathbf{q}_h. \tag{53}$$

Applying the strain operator $\mathbf{D}=1/2\,(\nabla+\nabla^T)$ to $\tilde{\mathbf{u}}$ we get the associated strain field:

$$\mathbf{e}^u=\mathbf{D}\mathbf{N}_r\mathbf{q}_r+\mathbf{D}\mathbf{N}_c\mathbf{q}_c+\mathbf{D}\mathbf{N}_h\mathbf{q}_h=\mathbf{B}_r\mathbf{q}_r+\mathbf{B}_c\mathbf{q}_c+\mathbf{B}_h\mathbf{q}_h. \tag{54}$$

But $\mathbf{B}_r=\mathbf{D}\mathbf{N}_r$ vanishes because $\mathbf{N}_r$ contains only rigid body modes. We are also free to select $\mathbf{B}_c=\mathbf{D}\mathbf{N}_c$ to be the identity matrix $\mathbf{I}$ if the generalized coordinates $\mathbf{q}_c$ are identified with the mean (volume-averaged) strain values $\bar{\mathbf{e}}^u$. Consequently, (54) simplifies to

$$\mathbf{e}^u=\bar{\mathbf{e}}^u+\mathbf{e}_h^u=\bar{\mathbf{e}}^u+\mathbf{B}_h\mathbf{q}_h, \tag{55}$$

in which

$$\mathbf{q}_c=\bar{\mathbf{e}}^u=(\mathbf{e}^u)_V/v,\qquad (\mathbf{B}_h)_V=\mathbf{0}. \tag{56}$$

Equation Partitioning. Assume that all elastic moduli in $\mathbf{E}$ are *constant* over the element. The degree-of-freedom partition

$$\mathbf{a}=\begin{Bmatrix}\bar{\sigma}\\ \\ \mathbf{a}_h\end{Bmatrix},\quad \mathbf{q}=\begin{Bmatrix}\mathbf{q}_r\\ \bar{\mathbf{e}}^u\\ \mathbf{q}_h\end{Bmatrix}, \tag{57}$$

induces the following partition of the general element equations (39)

$$\begin{bmatrix}
-\gamma v\mathbf{E}^{-1} & \mathbf{0} & -\bar{\mathbf{P}}_r^T & \gamma v\mathbf{I}-\bar{\mathbf{P}}_c^T & -\bar{\mathbf{P}}_h^T & \bar{\mathbf{L}}^T\\
\mathbf{0} & -\gamma\mathbf{C}_h & -\mathbf{P}_{hr}^T & -\mathbf{P}_{hc}^T & \gamma\mathbf{Q}_h^T-\mathbf{P}_{hh}^T & \mathbf{L}_h^T\\
-\bar{\mathbf{P}}_r & -\mathbf{P}_{hr} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\
\gamma v\mathbf{I}-\bar{\mathbf{P}}_c & -\mathbf{P}_{hc} & \mathbf{0} & (1-\gamma)v\mathbf{E} & \mathbf{0} & \mathbf{0}\\
-\bar{\mathbf{P}}_h & \gamma\mathbf{Q}_h-\mathbf{P}_{hh} & \mathbf{0} & \mathbf{0} & (1-\gamma)\mathbf{K}_{qh} & \mathbf{0}\\
\bar{\mathbf{L}} & \mathbf{L}_h & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}
\end{bmatrix}
\begin{Bmatrix}\bar{\sigma}\\ \mathbf{a}_h\\ \mathbf{q}_r\\ \bar{\mathbf{e}}^u\\ \mathbf{q}_h\\ \mathbf{v}\end{Bmatrix}=
\begin{Bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{f}_{qr}\\ \mathbf{f}_{qc}\\ \mathbf{f}_{qh}\\ \mathbf{f}_v\end{Bmatrix}$$

where

$$\mathbf{C}_h=(\mathbf{A}_h^T\mathbf{E}^{-1}\mathbf{A}_h)_V,\ \mathbf{Q}_h=(\mathbf{B}_h^T\mathbf{A}_h)_V,\ \mathbf{K}_{qh}=(\mathbf{B}_h^T\mathbf{E}\mathbf{B}_h)_V,$$
$$\bar{\mathbf{P}}_x=[\mathbf{N}_{xn}^T]_S,\ x=r,c,h,\ \mathbf{P}_{hx}=[\mathbf{N}_x^T\mathbf{A}_{hn}]_S,\ x=r,c,h,$$
$$\bar{\mathbf{L}}=[\mathbf{V}_n^T]_S,\ \mathbf{L}_h=[\mathbf{V}^T\mathbf{A}_{hn}]_S,\ \mathbf{f}_{qx}=(\mathbf{N}_x^T\mathbf{b})_V,\ x=r,c,h. \tag{59}$$

Integral Transformations. Application of the divergence theorem to the work of the mean stress on $\mathbf{e}^u$ yields

$$(\bar{\sigma},\mathbf{e}^u)_V=(\bar{\sigma},\bar{\mathbf{e}}^u+\mathbf{B}_h\mathbf{q}_h)_V=v\bar{\sigma}^T\bar{\mathbf{e}}^u+\bar{\sigma}^T(\mathbf{B}_h)_V\mathbf{q}_h=v\bar{\sigma}^T\bar{\mathbf{e}}^u$$
$$=[\bar{\sigma}_n,\tilde{\mathbf{u}}]_S=[\bar{\sigma}_n,\mathbf{N}_r\mathbf{q}_r+\mathbf{N}_c\bar{\mathbf{e}}^u+\mathbf{N}_h\mathbf{q}_h]_S$$
$$=\bar{\sigma}^T(\bar{\mathbf{P}}_r\mathbf{q}_r+\bar{\mathbf{P}}_c\bar{\mathbf{e}}^u+\bar{\mathbf{P}}_h\mathbf{q}_h). \tag{60}$$

Hence,

$$\bar{\mathbf{P}}_r=\mathbf{0},\ \bar{\mathbf{P}}_c=v\mathbf{I},\ \bar{\mathbf{P}}_h=\mathbf{0}. \tag{61}$$

A similar analysis of the stress-deviator work $(\tilde{\sigma}_h,\mathbf{e}^u)_V$ does not yield simple forms for the $\mathbf{P}_{hx}$ matrices unless $\tilde{\sigma}_h$ is *divergence-free*, in which case

$$\mathbf{P}_{hr}=\mathbf{0},\ \mathbf{P}_{hc}=\mathbf{0},\ \mathbf{P}_{hh}=\mathbf{Q}_h. \tag{62}$$

Hence, $\mathbf{P}\equiv\mathbf{Q}$ as claimed in Section 7. Inserting (61)–(62) into (58) yields the partitioned form of (40):

$$\begin{bmatrix} -\gamma v\mathbf{E}^{-1} & \mathbf{0} & \mathbf{0} & -(1-\gamma)v\mathbf{I} & \mathbf{0} & \bar{\mathbf{L}}^T \\ \mathbf{0} & -\gamma\mathbf{C}_h & \mathbf{0} & \mathbf{0} & -(1-\gamma)\mathbf{Q}_h^T & \mathbf{L}_h^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -(1-\gamma)v\mathbf{I} & \mathbf{0} & \mathbf{0} & (1-\gamma)v\mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -(1-\gamma)\mathbf{Q}_h & \mathbf{0} & \mathbf{0} & (1-\gamma)\mathbf{K}_{qh} & \mathbf{0} \\ \bar{\mathbf{L}} & \mathbf{L}_h & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \bar{\sigma} \\ \mathbf{a}_h \\ \mathbf{q}_r \\ \bar{\mathbf{e}}^u \\ \mathbf{q}_h \\ \mathbf{v} \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{f}_{qr} \\ \mathbf{f}_{qc} \\ \mathbf{f}_{qh} \\ \mathbf{f}_v \end{Bmatrix} \tag{63}$$

Orthogonality Conditions. If the higher-order stresses are divergence free so that $\mathbf{P} \equiv \mathbf{Q}$, the relations (47) partition as

$$\begin{bmatrix} \mathbf{0} & v\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_h^T \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{L}}^T \\ \mathbf{L}_h^T \end{bmatrix} [\mathbf{G}_r\ \mathbf{G}_c\ \mathbf{G}_h],$$

$$\begin{bmatrix} \mathbf{0} & v\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_h^T \end{bmatrix} \begin{bmatrix} \mathbf{H}_r \\ \mathbf{H}_c \\ \mathbf{H}_h \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{L}}^T \\ \mathbf{L}_h^T \end{bmatrix}, \tag{64}$$

whence the relations

$$\begin{aligned} &\bar{\mathbf{L}}^T\mathbf{G}_r = \mathbf{0},\ \bar{\mathbf{L}}^T\mathbf{G}_c = v\mathbf{I},\ \bar{\mathbf{L}}^T\mathbf{G}_h = \mathbf{0},\ \bar{\mathbf{L}}^T = v\mathbf{H}_c, \\ &\mathbf{L}_h^T\mathbf{G}_r = \mathbf{0},\ \mathbf{L}_h^T\mathbf{G}_c = \mathbf{0},\ \mathbf{L}_h^T\mathbf{G}_h = \mathbf{Q}_h^T,\ \mathbf{L}_h^T = \mathbf{Q}_h^T\mathbf{H}_h. \end{aligned} \tag{65}$$

The first four were obtained through other means by Bergan (1980) and Bergan and Nygård (1984), who called them the *force orthogonality* conditions on account of the physical interpretation of $\bar{\mathbf{L}}$ as a "boundary nodal force lumping" matrix in the free formulation studied next.

If the higher-order stresses are not divergence-free, the last four of (65) are replaced by

$$\begin{aligned} \mathbf{L}_h^t\mathbf{G}_r = \mathbf{P}_{hr}^T,\quad \mathbf{L}_h^T\mathbf{G}_c = \mathbf{P}_{hc}^T,\quad \mathbf{L}_h^T\mathbf{G}_h = \mathbf{P}_{hh}^T, \\ \mathbf{L}_h^T = \mathbf{P}_{hr}^T,\quad \mathbf{H}_r + \mathbf{P}_{hc}^T\mathbf{H}_c + \mathbf{P}_{hh}^T\mathbf{H}_h. \end{aligned} \tag{66}$$

11 The Free Formulation

The free formulation of Bergan and Nygård (1984) was originally conceived as an incompatible finite element displacement model that passes a cancelling-tractions version of the patch test which Bergan and Hanssen (1975) called the individual patch test. Here the formulation is reinterpreted in the context of the hybrid principle (19). The assumptions that lead to the FF are listed in the Introduction and will be studied in further detail.

Constant Internal Stress. The internal stress field is *constant*. Consequently, there are no $\mathbf{a}_h$ parameters, reducing (63) to

$$\begin{bmatrix} -\gamma v\mathbf{E}^{-1} & \mathbf{0} & -(1-\gamma)v\mathbf{I} & \mathbf{0} & \bar{\mathbf{L}}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -(1-\gamma)v\mathbf{I} & \mathbf{0} & (1-\gamma)v\mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & (1-\gamma)\mathbf{K}_{qh} & \mathbf{0} \\ \bar{\mathbf{L}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \bar{\sigma} \\ \mathbf{q}_r \\ \bar{\mathbf{e}}^u \\ \mathbf{q}_h \\ \mathbf{v} \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{f}_{qr} \\ \mathbf{f}_{qc} \\ \mathbf{f}_{qh} \\ \mathbf{f}_v \end{Bmatrix}. \tag{67}$$

Invertible G. Matrix $\mathbf{G}$ in (45) is constructed by *nodal collocation*, that is, by evaluating the expansion $\bar{\mathbf{u}} = \mathbf{N}\mathbf{q}$ at the element boundary nodes. This establishes the transformation

$$\mathbf{v} = \mathbf{G}\mathbf{q} = [\mathbf{G}_r\ \mathbf{G}_c\ \mathbf{G}_h] \begin{Bmatrix} \mathbf{q}_r \\ \bar{\mathbf{e}}^u \\ \mathbf{q}_h \end{Bmatrix}. \tag{68}$$

According to the assumptions listed in the Introduction, matrix $\mathbf{G}$ is square and nonsingular so inverting (68) we get $\mathbf{q} = \mathbf{G}^{-1} = \mathbf{H}\mathbf{v}$ or

$$\mathbf{q} = \begin{Bmatrix} \mathbf{q}_r \\ \bar{\mathbf{e}}^u \\ \mathbf{q}_h \end{Bmatrix} = \begin{bmatrix} \mathbf{H}_r \\ \mathbf{H}_c \\ \mathbf{H}_h \end{bmatrix} \mathbf{v} = \begin{bmatrix} \mathbf{H}_r \\ v^{-1}\bar{\mathbf{L}}^T \\ \mathbf{H}_h^T \end{bmatrix} \mathbf{v}. \tag{69}$$

The FF Stiffness Equations. Eliminating $\bar{\sigma}$ and $\mathbf{q}$ from (67) yields the FF stiffness equations

$$\mathbf{K}\mathbf{v} = [\mathbf{K}_b + (1-\gamma)\mathbf{K}_h]\mathbf{v} = \mathbf{f}, \tag{70}$$

where

$$\mathbf{K}_b = v^{-1}\bar{\mathbf{L}}\mathbf{E}^{-1}\bar{\mathbf{L}}^T,\quad \mathbf{K}_h = \mathbf{H}_h^T\mathbf{K}_{qh}\mathbf{H}_h,$$

$$\mathbf{f} = \mathbf{f}_v + \mathbf{H}_r^T\mathbf{f}_{qr} + v^{-1}\bar{\mathbf{L}}\mathbf{f}_{qc} + \mathbf{H}_h\mathbf{f}_{qh}. \tag{71}$$

In the free formulation, $\mathbf{K}_b$ and $\mathbf{K}_h$ receive the name *basic* and *higher-order* stiffness matrices, respectively. A 1/2 scaling of $\mathbf{K}_h$ derived from energy-balancing studies was recommended by Bergan and Felippa (1985) for a plane-stress element. This corresponds to taking $\gamma = 1/2$. But in general the value of γ can be expected to be dependent on the type and geometry of the element.

As $\mathbf{K}_b$ is rank-deficient (except for the simplex elements) choosing $\gamma = 1$, which corresponds to the d-generalized Hellinger-Reissner functional (17), is *not* admissible.

12 The Extended Free Formulation

In the extended free formation (EFF) the number of internal displacement freedoms, $n_q = \dim(\mathbf{q})$, is allowed to exceed the number of nodal displacement connectors $n_v = \dim(\mathbf{v})$. We can establish the relation (68) as before, but matrix $\mathbf{G}$ will now be rectangular and cannot be directly inverted. One way of circumventing this difficulty is to retain $n_q - n_v = \dim(\mathbf{a}_h)$ higher-order stress modes; an alternative procedure is discussed in Section 13. The stress modes are assumed to be divergence-free so (62) holds. The available relations are

$$\mathbf{v} = \mathbf{G}\mathbf{q},\quad \mathbf{C}_h\mathbf{a}_h = \mathbf{L}_h^T\mathbf{v} = \mathbf{Q}_h^T\mathbf{q}_h, \tag{72}$$

which can be combined to form the matrix system

$$\begin{Bmatrix} \mathbf{v} \\ \mathbf{a}_h \end{Bmatrix} = \begin{bmatrix} \mathbf{G}_r & \mathbf{G}_c & \mathbf{G}_h \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_h^{-1}\mathbf{Q}_h^T \end{bmatrix} \begin{Bmatrix} \mathbf{q}_r \\ \bar{\mathbf{e}}^u \\ \mathbf{q}_h \end{Bmatrix}. \tag{73}$$

The matrix on the right side is square, and invertible if $\mathbf{G}$, $\mathbf{C}_h$, and $\mathbf{Q}_h$ have full rank. Solving for $\mathbf{q}$ and eliminating $\mathbf{a}_h$ one obtains

$$\mathbf{q}=\begin{Bmatrix}\mathbf{q}_r\\ \mathbf{q}_c\\ \mathbf{q}_h\end{Bmatrix}=\begin{bmatrix}\mathbf{H}_r & \mathbf{0}\\ \mathbf{H}_c & \mathbf{0}\\ \mathbf{H}_h' & \mathbf{J}_h\end{bmatrix}\begin{Bmatrix}\mathbf{v}\\ \mathbf{a}_h\end{Bmatrix}=$$

$$\begin{bmatrix}\mathbf{H}_r\\ \mathbf{H}_c\\ \mathbf{H}_h'+\mathbf{J}_h\mathbf{C}_h^{-1}\mathbf{Q}_h^T\end{bmatrix}\mathbf{v}=\begin{bmatrix}\mathbf{H}_r\\ \mathbf{H}_c\\ \mathbf{H}_h\end{bmatrix}\mathbf{v}, \tag{74}$$

where $\mathbf{H}_h'$ and $\mathbf{J}_h$ result from the inversion process. Since $\mathbf{H}_h'$ $\mathbf{G}_h + \mathbf{J}_h\,\mathbf{C}_h^{-1}\,\mathbf{Q}_h^T = \mathbf{I}$, we can express $\mathbf{H}_h$ as

$$\mathbf{H}_h=\mathbf{H}_h'+\mathbf{I}-\mathbf{H}_h'\mathbf{G}_h. \tag{75}$$

Having $\mathbf{H}$ available, replacing into (48)–(50) we obtain the EFF stiffness equations

$$\mathbf{K}\mathbf{v}=[\mathbf{K}_b+\mathbf{K}_{bh}+(1-\gamma)\mathbf{K}_h]\mathbf{v}=\mathbf{f}, \tag{76}$$

where $\mathbf{K}_b$, $\mathbf{K}_h$, and $\mathbf{f}$ are the same as in (71), and

$$\mathbf{K}_{bh}=\mathbf{L}_h\mathbf{C}_h^{-1}\mathbf{L}_h^T. \tag{77}$$

Is $\gamma = 1$ now admissible? If $\mathbf{K}_b + \mathbf{K}_{bh}$ has correct rank, yes! Curiously enough, if the body force field $\mathbf{b}$ vanishes and $\gamma = 1$, (76) are precisely the stiffness equations for the original equilibrium-stress-assumed hybrid elements of Pian (1964), which can, of course, be constructed without any internal displacement assumptions.

13 Hierarchical Connector Augmentation

An alternative approach to building an invertible transformation such as (73) consists of augmenting $\mathbf{v}$ with $n_q - n_v$ connector degrees-of-freedom collected in subvector $\mathbf{v}_x$. These must be selected to give a square transformation of the form

$$\begin{Bmatrix}\mathbf{v}\\ \mathbf{v}_x\end{Bmatrix}=\begin{bmatrix}\mathbf{G}_r & \mathbf{G}_c & \mathbf{G}_h\\ \mathbf{0} & \mathbf{0} & \mathbf{G}_x\end{bmatrix}\begin{Bmatrix}\mathbf{q}_r\\ \bar{\mathbf{e}}^u\\ \mathbf{q}_h\end{Bmatrix}. \tag{78}$$

If this approach is followed, it is important to choose $\mathbf{v}_x$ in *hierarchical* fashion so that the expanded $\mathbf{G}$ has the structure just shown. In other words, $\mathbf{v}_x$ *must not be "excited" by rigid body or constant strain motions*. Otherwise the interelement compatibility of boundary displacements is generally violated for such motions, and the patch test discussed in the Appendix fails.

Inversion of (78) provides the $\mathbf{H}$ matrix. The FF stiffness equations (70) can be constructed with the strain-energy contribution from $\mathbf{v}_x$ flowing to the higher-order stiffness $\mathbf{K}_h$. Finally, the $\mathbf{v}_x$ freedoms can be statically condensed.

Which EFF approach is better? The decision seems to be element-dependent. The choice primarily hinges on whether it is easier to choose divergence-free stress modes than hierarchical connectors while preserving element invariance. If both approaches appear equally feasible, there is not presently enough experience to decide which one is preferable.

14 Concluding Remarks

The qualifier *free* in "free formulation" was meant to emphasize "freedom from conformity requirements" that are a pervasive part of the conventional displacement formulation, and the possibility of constructing the basic and higher-order stiffness contributions through largely independent assumptions. But when the FF is studied from a variational standpoint, several constraints become immediately apparent. The extended FF releases the most troublesome one at the cost of buying more complicated stress assumptions, or additional hierarchical connectors. So it is fair to state that the admirable goal of absolute freedom has not yet been attained.

The development of the EFF as reported here was motivated by difficulties encountered in the construction of the following elements:

3-Node Plane Stress Triangle with Nodal Rotations. Similar to the element constructed by Bergan and Felippa (1985), but with a *fully quadratic* internal displacement field. Thus, $n_v = 9$, $n_q = 12$ and three additional self-equilibrating stress fields are needed.

4-Node Tetrahedron with Nodal Rotations. The extension of the previous element to three dimensions has $n_v = 12$, $n_q = 18$ and six additional stress fields are needed.

Assuming fully-quadratic internal displacement fields eliminates the higher-order mode selection difficulties discussed by Bergan and Felippa (1985). Progress in the derivation of these elements will be reported in subsequent papers.

Acknowledgments

The preparation of this paper was jointly supported by the Office of Naval Research under Contract N0001486-C-0082, and by the Naval Research Laboratory under Grant N00014-87-K-2018. The author is indebted to the referees for many constructing suggestions, including the reference to the work of Chien (1983).

References

Bergan, P. G., and Hanssen, L., 1976, "A New Approach for Deriving 'Good' Finite Elements," *The Mathematics of Finite Elements and Applications*, Vol. II, MAFELAP II Conference, Brunel University, 1975, by J. R. Whiteman, ed., Academic Press, London.

Bergan, P. G., 1980, "Finite Elements Based on Energy Orthogonal Functions," *Int. J. Num. Meth. Engrg.*, Vol. 15, pp. 1141–1555.

Bergan, P. G., and Nygård, M. K., 1984, "Finite Elements with Increased Freedom in Choosing Shape Functions," *Int. J. Num. Meth. Engrg.*, Vol. 20, pp. 643–664.

Bergan, P. G., and Wang, X., 1984, 'Quadrilateral Plate Bending Elements With Shear Deformations," *Computer and Structures*, Vol. 19, pp. 25–34.

Bergan, P. G., and Felippa, C. A., 1985, "A Triangular Membrane Element with Rotational Degrees of Freedom," *Computer Methods in Applied Mechanics and Engineering*, Vol. 50, pp. 25–69.

Bergan, P. G., and Nygård, M. K., 1985, "Nonlinear Shell Analysis Using Free Formulation Finite Elements," *Proc. Europe-US Symposium on Finite Element Methods for Nonlinear Problems*, Trondheim, Norway, August 1985, Springer-Verlag, Berlin.

Chien, W. Z., 1983, "Method of High-Order Lagrange Multiplier and Generalized Variational Principles of Elasticity with More Forms of Functionals," *Applied Mathematics and Mechanics*, Vol. 4, pp. 143–159.

Felippa, C. A., and Bergan, P. G., 1987, "A Triangular Plate Bending Element Based on an Energy-Orthogonal Free Formulation," *Computer Methods in Applied Mechanics and Engineering*, Vol. 61, pp. 129–160.

Felippa, C. A., 1989a, "Parametrized Multifield Variational Principles in Elasticity: I. Mixed Functionals," *Communications in Applied Numerical Methods*, Vol. 5, pp. 79-88.

Felippa, C. A., 1989b, "Parametrized Multifield Variational Principles in Elasticity: II. Hybrid Functionals and the Free Formulation," *Communications in Applied Numerical Methods*, Vol. 5, pp. 89–90.

Fraeijs de Veubeke, B. M., 1974, "Variational Formulation and the Patch Test," *Int. J. Num. Meth. Engrg.*, Vol. 8, pp. 783–801.

Gurtin, M., 1972, "The Linear Theory of Elasticity," *Encyclopedia of Physics*, Vol. VIa/2, by C. Truesdell, ed., Springer-Verlag, Berlin.

Hanssen, L., Syvertsen, T. G., and Bergan, P. G., 1979, "Stiffness Derivation Based on Element Convergence Requirements," *The Mathematics of Finite Elements and Applications*, Vol. III, MAFELAP III Conference, Brunel University, 1978, by J. R. Whiteman, ed., Academic Press, London.

Nygård, M. K., 1986, "The Free Formulation for Nonlinear Finite Element Analysis with Applications to Shells," Dr. Ing. Dissertation, Division of Structural Mechanics, The Norwegian Institute of Technology, Trondheim, Norway.

Oden, J. T., and Reddy, J. N., 1983, *Variational Methods in Theoretical Mechanics*, 2nd ed., Springer-Verlag, Berlin.

Pian, T. H. H., 1964, "Derivation of Element Stiffness Matrices by Assumed Stress Distributions," *AIAA Journal*, Vol. 2, pp. 1333–1336.

Pian, T. H. H., and Tong, P., 1969, "Basis of Finite Element Methods for Solid Continua," *Int. J. Numer. Meth. Engrg.*, Vol. 1, pp. 3–29.

Pian, T. H. H., and Tong, P., 1969, "Basis of Finite Element Methods for Solid Continua," *Int. J. Numer. Meth. Engrg.*, Vol. 1, pp. 3-29.

Pian, T. H. H., 1973, "Finite Element Methods by Variational Principles with Relaxed Continuity Requirements," *Variational Methods in Engineering*, Vol. 1, by C. A. Brebbia and H. Tottenham, eds., Southampton University Press, Southampton, U.K.

Washizu, K., 1968, *Variational Methods in Elasticity and Plasticity*, Pergamon Press, Oxford.

APPENDIX A

The Cancelling-Traction Patch Test

It is instructive to study whether this element class passes the patch test for an arbitrary γ. To investigate this question we use the sketch of Fig. 1 and view the subvolumes V^+ and V^- as *two elements* connected along S_i with an external traction boundary S_t. Both elements are in a state of *constant stress* σ_0. The prescribed surface tractions are $[\hat{\mathbf{t}} = \sigma_{0n}]_{S_t}$ and the body forces $\mathbf{b}$ vanish.

First, take (63) to be the governing discrete equations for the two-element assembly. The only nonzero forces are $\mathbf{f}_v = [\mathbf{V}^T\hat{\mathbf{t}}]_{S_t}$. The key observation is that

$$\bar{\mathbf{L}} = [\mathbf{V}_n^T]_S = [\mathbf{V}_n^T]_{S_t}, \tag{79}$$

because the integral over S_i vanishes as $(\mathbf{V}_+ = \mathbf{V}_-)_{S_i}$ on account of the interface compatibility conditions stated in Section 8, and $\mathbf{n}^+ = -\mathbf{n}^-$. Now, for any γ it can be verified that the solution of (63) is that demanded by the patch test, namely

$$\bar{\sigma} = \sigma_0 = \tilde{\sigma}^u,\ \mathbf{a}_h = \mathbf{0},\ \mathbf{q}_r = \text{arbitrary},$$
$$\tilde{\mathbf{e}}^u = \mathbf{E}^{-1}\bar{\sigma}_0,\ \mathbf{q}_h = \mathbf{0},\ \mathbf{v} = \mathbf{L}^T\bar{\sigma}_0 + \mathbf{G}_r\mathbf{q}_r. \tag{80}$$

In checking this assertion one finds that the following relations, listed in (65), must be satisfied:

$$\bar{\mathbf{L}}^t\mathbf{G}_r = \mathbf{0},\ \bar{\mathbf{L}}^T\mathbf{G}_c = v\mathbf{I},\ \mathbf{L}_h^T\mathbf{G}_c = \mathbf{0},\ \mathbf{L}_h\mathbf{G}_r = \mathbf{0}. \tag{81}$$

If instead we take the more general equations (59), verification of the solution (81) demands that

$$\bar{\mathbf{P}}_r = \mathbf{0},\ \bar{\mathbf{P}}_c = v\mathbf{I},\ \bar{\mathbf{P}}_h = \mathbf{0},\ \mathbf{P}_{hr}^T = \mathbf{L}_h^T\mathbf{G}_r,$$
$$\mathbf{P}_{hc}^T = \mathbf{L}_h^T\mathbf{G}_c,\ \mathbf{P}_{hh}^T = \mathbf{L}_h^T\mathbf{G}_h. \tag{82}$$

The first three follow from the divergence theorem as shown in (60). But the last three, listed in (66), are a consequence of the kinematic constraint (43), which is thus directly correlated to satisfaction of the patch test.

As noted by Fraeijs de Veubeke (1973), the physical meaning of this form of the patch test is that the *interface virtual work is zero* when the element patch is in a constant stress state.

A Boundary Integral Equation Formulation in Derivative Unknowns for Two-Dimensional Potential Problems

Joo Ho Choi
Research Associate,
Center for Computer-Aided Design.

Byung Man Kwak[1]
Visiting Professor,
Department of Biomechanical Engineering.
Mem. ASME

The University of Iowa,
Iowa City, Iowa 52242

A boundary integral equation called Derivative BIE is developed for two-dimensional potential problems in terms of tangential and normal derivatives of the potential on the boundary, by integrating by parts the Cauchy formula. The potential values on the boundary can be calculated by integration after the solution is obtained. The primary unknowns in this formulation can be of direct interest in a shape design sensitivity analysis where the tangential derivatives of the potential are also required. The method is applied to several test problems, and the results show better accuracy than those by the conventional boundary element method, not only for the derivatives of the potential but also for the potential itself.

Introduction

One of the most popular and earlier applications of the boundary element method (BEM) is found in the potential problems governed by the Laplace equation such as potential flow and heat conduction. Generally the unknown variables in the boundary integral equation (BIE) consist of two kinds. One is the potential on the Neumann part, and the other is its normal derivative on the Dirichlet part. In this conventional formulation, both the potential and its normal derivative are approximated by the same interpolation function, despite they are different in smoothness. The solution thus obtained on the boundary, however, is known to be usually more accurate than those by a classical finite element method (FEM) with similar meshes. Specifically, as for the flux on the boundary the BEM solves it directly, while in the FEM it is calculated by a numerical differentiation of the potential and an extrapolation to the boundary, which might be a source of poor accuracy.

There is a class of problems in which the tangential derivative information of the potential on the boundary is as equally important as its normal derivative, such as in the free boundary value problems (Liggett and Liu, 1982) and shape design sensitivity analyses (Kwak and Choi, 1987). This is more obvious if we look into the plane elasticity problems where the critical stress occurs usually on the boundary. It is often expressed in terms of the von Mises stress, which depends not only on the tractions, but also on the tangential derivative of the displacements (Banerjee and Butterfield, 1981).

In the conventional BEM, tangential derivative of the potential on the boundary is calculated by a numerical differentiation after the primary solution is obtained. This process, however, degrades the accuracy by an order as compared to that of the original potential. One obvious way to alleviate this problem is to employ higher-order interpolation functions for the potential. Another method is to use a new formulation that was developed in this paper. In this approach, the tangential derivative of the potential is taken as an unknown variable instead of the potential itself, such that both the tangential and normal derivatives are the primary unknowns on the boundary. Once we get the boundary solutions, the potential value can be obtained by an integration of its tangential derivative along the boundary, which is numerically more advantageous than a differentiation.

There are a few papers which consider the tangential derivatives, although they are somewhat different from ours, either in the method of derivation or in the usage.

Athanasiadis (1985) formulated a BIE for the derivatives of the potential on the boundary, in an attempt to derive many different kinds of integral equations in the heat conduction problem. However, he ended up by adding a stronger singular kernel in his equation than that of the conventional one, which was not desirable. Katz (1982) used a tangential derivative term, but it was introduced only as a means for a better calculation of interior values, especially for points close to the boundary. The boundary solution from the conventional BEM was still used. Recently, Ghosh et al. (1986) included the tangential derivatives of displacement as unknowns in their elasticity formulation, which was brought to our attention only in the last stage of this writing. The resulting formula connotes an idea fairly similar to ours, even though we deal with

[1] On leave from the Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Seoul, Korea.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 23, 1988; final revision, December 15, 1988.

Paper No. 89-WA/APM-13.

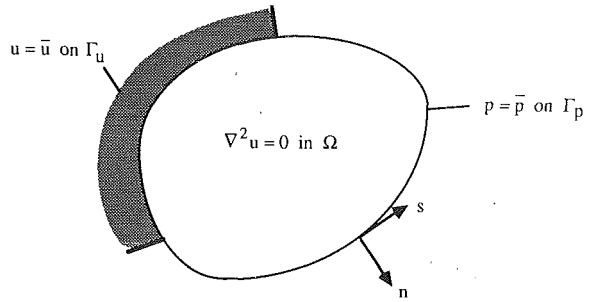


Fig. 1 Definition of potential problem

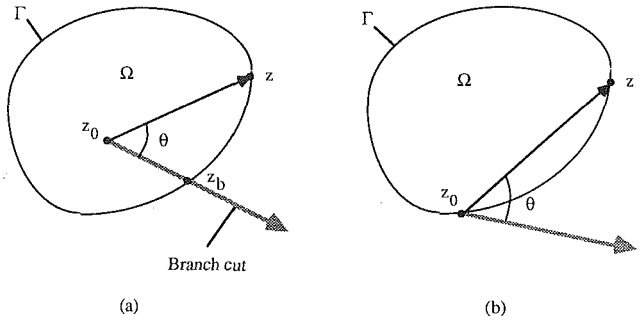


Fig. 2 Multivaluedness of complex logarithm function; (*a*) when the point z_0 is within the domain, (*b*) when the point z_0 is on the boundary

two-dimensional potential problems and the method of derivation is different.

In this paper, a systematic development of the new BIE, called derivative BIE, is presented for two-dimensional potential problems. The derivation is based on integration by parts of the Cauchy formula in the complex variable theory. In the following, the conventional BIE is first reviewed in complex variable theory, before our new BIE is formulated. The advantage over the conventional boundary elements will then be presented illustrating some numerical examples.

Conventional BIE From Complex Variable Theory

The most usual way for the formulation of the classical BIE has been to apply Green's formula for the potential, introducing a suitable fundamental solution. Of special importance, however, is the fact that the well-known Cauchy integral formula for complex analytic function is a generalized expression of the conventional BIE for the real potential in the complex domain (Carrier et al., 1966).

Let $\phi(z)$ be an analytic function in a simply-connected complex domain Ω with a simply-closed boundary Γ. Then $\phi(z)$ can be written in terms of two real variable functions such that

$$\phi = u(x,y) + iw(x,y), \quad z \in \Omega \tag{1}$$

where $z = x + iy$, $i = \sqrt{(-1)}$, u is a potential function to be sought, and w is a conjugate function which satisfies the Cauchy-Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial w}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial w}{\partial x}. \tag{2}$$

It is noted that u and w are harmonic and satisfy the Laplace equation. Then the Cauchy integral formula states that (Carrier et al., 1966)

$$\frac{1}{2\pi i}\int_\Gamma \frac{\phi(z)}{z - z_0}\,dz = \begin{cases} \phi(z_0), & z_0 \in \Omega \\ \alpha\phi(z_0), & z_0 \in \Gamma \\ 0, & z_0 \in \text{ outside } \Gamma \end{cases} \tag{3}$$

where α denotes the interior angle at z_0 on the boundary divided by 2π, which will be one half if the boundary is smooth.

The complex expressions of equation (3) can now be written in terms of their real and imaginary parts to obtain a pair of integral equations for u and w, which are similar to the conventional BIE:

$$\alpha u(\mathbf{x}_0) = \frac{1}{2\pi}\int_\Gamma \left\{ w\frac{\partial}{\partial s}\ln r + u\frac{\partial}{\partial n}\ln r \right\} ds \tag{4}$$

$$\alpha w(\mathbf{x}_0) = \frac{1}{2\pi}\int_\Gamma \left\{ w\frac{\partial}{\partial n}\ln r - u\frac{\partial}{\partial s}\ln r \right\} ds \tag{5}$$

where $\mathbf{x}_0$ is on Γ, and $r = |\mathbf{x} - \mathbf{x}_0|$. These equations have coupled expressions with respect to the potential u and its conjugate w, hence, two equations may be solved simultaneously for both u and w. However, what we need now is the BIE for the potential u only, which can be done by integrating by parts the first one in the right-hand side of equation (4) and utilizing the Cauchy-Riemann relations for the (n, s) coordinates. Then the desired BIE is obtained as,

$$\alpha u(\mathbf{x}_0) = \int_\Gamma \{p(\mathbf{x})G(\mathbf{x},\mathbf{x}_0) - u(\mathbf{x})F(\mathbf{x},\mathbf{x}_0)\}\,ds, \quad \mathbf{x}_0 \in \Gamma \tag{6}$$

where G and F denote the fundamental solution and its derivative, respectively, given by

$$G(\mathbf{x},\mathbf{x}_0) = -\frac{1}{2\pi}\ln|\mathbf{x} - \mathbf{x}_0|,$$

$$F(\mathbf{x},\mathbf{x}_0) = \frac{\partial}{\partial n}G(\mathbf{x}, \mathbf{x}_0) \tag{7}$$

and $p \equiv \partial u/\partial n$ denotes the flux on the boundary.

In a well-posed problem, only the value of u or p is prescribed at each boundary point as shown in Fig. 1, i.e.,

$$\left.\begin{aligned} u &= \bar{u} \quad \text{on } \Gamma_u \\ p &= \bar{p} \quad \text{on } \Gamma_p \end{aligned}\right\} \tag{8}$$

where Γ_u and Γ_p denote the Dirichlet and the Nemann boundary, respectively, and $\Gamma_u \cup \Gamma_p \equiv \Gamma$. If we discretize equation (6) approximating the boundary and the variables u and p with suitable shape functions and apply the resulting equation at each collocation point, a system of equations is obtained, which determines the remaining unknowns u and p on their respective boundary. Although the potential u and its flux p have a different smoothness requirements, the same shape function is taken for both u and p in most boundary elements. If the values for the tangential derivative of u on the boundary are needed, they are calculated by a numerical differentiation after the solution is obtained, which yields a result one order less smoother than u or p.

Formulation of Derivative BIE

Consider first a simply-connected domain. The derivative BIE is then developed starting from an integration by parts of Cauchy's formula (3), to obtain

$$\frac{1}{2\pi i}\int_\Gamma \ln(z - z_0)\phi'(z)\,dz = \begin{cases} \phi(z_b) - \phi(z_0), & z_0 \in \Omega \\ 0, & z_0 \in \Gamma \text{ and outside } \Gamma \end{cases} \tag{9}$$

where $\phi'(z) \equiv d\phi/dz$ is the derivative of ϕ in complex sense, and z_b is a point of intersection on Γ, by a branch cut originating at z_0 in Ω and passing to infinity, as shown in Fig. 2(*a*). This is due to the multivaluedness of complex logarithmic function, which has appeared as a result of the integration by parts. Especially when z_0 is on Γ, care should be taken to the direction of the branch cut for equation (9) to hold at z_0 on Γ, such that the cut should not cross any point on the boundary except the point itself as shown in Fig. 2(*b*).

To now derive the derivative BIE, the complex expressions are rewritten in terms of the real and imaginary parts, utilizing the following relations

$$\ln(z-z_0)=\ln|\mathbf{x}-\mathbf{x}_0|+i\theta(\mathbf{x}-\mathbf{x}_0) \tag{10}$$

$$\begin{aligned}\phi'(z)dz=\frac{\partial\phi}{\partial s}\,ds&=\left(\frac{\partial u}{\partial s}+i\,\frac{\partial w}{\partial s}\right)ds\\&=\left(\frac{\partial u}{\partial s}+i\,\frac{\partial u}{\partial n}\right)ds\\&=(q+ip)ds\end{aligned} \tag{11}$$

where the symbol θ denotes the angle between the vector $\mathbf{x}-\mathbf{x}_0$ and a reference direction which corresponds to that of the branch cut. In equation (11), $q\equiv\partial u/\partial s$ denotes the tangential derivative of u on the boundary. Substitute equations (10) and (11) into equation (9) when z_0 is on Γ, and take the real and imaginary parts, respectively, from the resulting equation. Then a pair of boundary integral equations are obtained for p and q:

$$\int_\Gamma \{G(\mathbf{x},\mathbf{x}_0)p(\mathbf{x})+H(\mathbf{x},\mathbf{x}_0)q(\mathbf{x})\}ds=0,\quad \mathbf{x}_0\in\Gamma \tag{12}$$

$$\int_\Gamma \{H(\mathbf{x},\mathbf{x}_0)p(\mathbf{x})-G(\mathbf{x},\mathbf{x}_0)q(\mathbf{x})\}ds=0,\quad \mathbf{x}_0\in\Gamma \tag{13}$$

where H denotes the conjugate of the fundamental solution G, given by

$$H(\mathbf{x},\mathbf{x}_0)=-\frac{1}{2\pi}\,\theta(\mathbf{x}-\mathbf{x}_0). \tag{14}$$

Note here that if the integration by parts on the boundary were applied directly to the conventional BIE of the potential problem, only equation (12) could be obtained.

It is interesting to note that unlike equations (4) and (5), the two equations just derived are expressed in terms of the potential u, or actually its derivatives only. That is, there appears no conjugate function w in equations (12) and (13). This implies that the two real integral equations hold simultaneously for the derivatives of u at each point of the boundary. Therefore, one can use either equation (12) or (13) at his convenience at any collocation point $\mathbf{x}_0$, provided that certain continuity of the potential and single valuedness are satisfied, as will be discussed later.

It is further observed that the kernel H has no singular behavior which stabilizes the integration process and removes the necessity for special treatment. However, when the boundary element equations are made from equation (12) or (13), the elements of H are found to be all of similar order in magnitude, possibly leading to a poor behavior of matrix, as experienced in a test when only equation (12) was used. This problem, however, can be solved by suitably selecting equations out of (12) and (13) such that G-term—which shows diagonally dominant behavior—takes place in the diagonal of the main matrix. This is possible by applying (12) on the Dirichlet part, and by applying (13) on the Neumann part, respectively. Thus, one finally obtains the derivative BIE for q and p suitable for a numerical calculation as

$$\left.\begin{aligned}&\int_\Gamma G(\mathbf{x},\mathbf{x}_0)p(\mathbf{x})ds=-\int_\Gamma H(\mathbf{x},\mathbf{x}_0)q(\mathbf{x})ds,\ \mathbf{x}_0\in\Gamma_q\\&\int_\Gamma G(\mathbf{x},\mathbf{x}_0)q(\mathbf{x})ds=\int_\Gamma H(\mathbf{x},\mathbf{x}_0)p(\mathbf{x})ds,\ \mathbf{x}_0\in\Gamma_p\end{aligned}\right\} \tag{15}$$

where Γ_q corresponds to Γ_u, hence, $\Gamma_q\cup\Gamma_p\equiv\Gamma$.

Now the boundary conditions (8) can be rewritten as

$$\left.\begin{aligned}q&=\bar{q}\quad\text{on }\Gamma_q\\p&=\bar{p}\quad\text{on }\Gamma_p\end{aligned}\right\} \tag{16}$$

where $\bar{q}$ is obtained by differentiating u along Γ_u. A distinguishing feature of the derivative BIE is that homogeneous boundary conditions are encountered, since all the variables consist of the derivatives of u. A typical example is a heat conduction problem with zero flux and constant temperature in each boundary, as shown in the examples in this paper. For these problems, the generated equations will obviously be homogeneous, yielding a trivial solution for p and q unless the coefficient of unknowns is a singular operator. Hence, to ensure a nontrivial unique solution some auxiliary conditions relating given information on the potential u need be imposed. Assuming that the potential is continuous throughout the boundary, the relation between u and q is utilized for this purpose as follows:

$$\int_{\Gamma_p} q\,ds=\Delta u=u_2-u_1 \tag{17}$$

where u_1 and u_2 denote the prescribed values of the potential at the starting and ending points of the Neumann boundary Γ_p, where q is unknown. This condition is to be imposed for every segment of Γ_p.

Now, equation (15) should be solved under constraint (17). The classical Lagrange multiplier method (Hildebrand, 1965) can be efficiently used. As in the conventional BIE, equation (15) is discretized and reordered for the unknowns to obtain

$$\begin{bmatrix}\mathbf{G}_{qq} & \mathbf{H}_{qp}\\ -\mathbf{H}_{pq} & \mathbf{G}_{pp}\end{bmatrix}\begin{Bmatrix}p_q\\ q_p\end{Bmatrix}=\begin{bmatrix}-\mathbf{H}_{qq} & -\mathbf{G}_{qp}\\ -\mathbf{G}_{pq} & \mathbf{H}_{pp}\end{bmatrix}\begin{Bmatrix}\bar{q}\\ \bar{p}\end{Bmatrix}\equiv\begin{Bmatrix}b\end{Bmatrix} \tag{18}$$

where p_q and q_p are the remaining unknowns on their respective boundary, and the vector b simply represents the resulting values of the second matrix operation. Note here that while in the conventional BEM only the columns between F and G are exchanged (Banerjee and Butterfield, 1981), both the row and column are exchanged in (18), so that the diagonal terms are always G. Next, the constraint (17) becomes, after discretization,

$$c^Tq_p=d. \tag{19}$$

Introducing the Lagrange multiplier λ for constraint (19), the complete system of equations is obtained as

$$\begin{bmatrix}\mathbf{G}_{qq} & \mathbf{H}_{qp} & 0\\ -\mathbf{H}_{pq} & \mathbf{G}_{pp} & c\\ 0 & c^T & 0\end{bmatrix}\begin{Bmatrix}p_q\\ q_p\\ \lambda_\alpha\end{Bmatrix}=\begin{Bmatrix}b\\ d\end{Bmatrix}. \tag{20}$$

Once the solutions for p and q are obtained from this equation, the potential u will be calculated, if necessary, by integrating q along the boundary.

Conjugate Fundamental Solution

As was discussed in the previous section, one should evaluate carefully the integration involving the conjugate fundamental solution H because of its multivaluedness. It has the following form

$$I(\mathbf{x}_0)=\int_\Gamma \theta(\mathbf{x}-\mathbf{x}_0)\sigma(\mathbf{x})ds,\quad \mathbf{x}_0\in\Gamma \tag{21}$$

where σ denotes either p or q as shown in equation (15), and the integration over $\mathbf{x}$ is done along the contour Γ counterclockwise starting from $\mathbf{x}_0$.

The angle $\theta(\mathbf{x}-\mathbf{x}_0)$ can be measured from any reference direction, but the change in the reference direction brings into an additional term

$$I_0 = \theta_0 \int_\Gamma \sigma \, ds \tag{22}$$

where θ_0 is the magnitude of the angular change of the reference direction at $\mathbf{x}_0$. Unless the integral in equation (22) vanishes, θ_0 should be carried on as additional unknowns in the formulated BIE (15). The developed BIE, however, holds independently of θ_0, if there is no net flux or circulation, which means that the integral in equation (22) vanishes identically. When σ is p, the integral denotes the total flux over the whole boundary. When σ is q it is the circulation, and no circulation means that the potential u should be single-valued and continuous over the whole boundary. In the BIE developed here, this condition can be satisfied by the condition (17), which is imposed on q, assuming u is continuous. However, for problems with nonzero circulation, the formulation should be appropriately modified.

In the following no circulation and no net flux will be assumed, so θ_0 does not appear. The reference direction can then be chosen arbitrarily, say, in the outward normal at $\mathbf{x}_0$ on the boundary. The angle θ is measured counterclockwise with respect to this direction, and does not undergo jump as $\mathbf{x}$ traverses the boundary. Then, for the geometries shown in Fig. 3, the angle is added by 2π on the thicker curve part of the boundary, which allows θ to go beyond the range between 0 and 2π, varying continuously throughout the boundary except at $\mathbf{x}_0$ itself (Jaswon and Symm, 1977). Hence, the whole domain can be considered as a single-valued branch. This can be thought of as a kind of analytic continuation applied to θ.

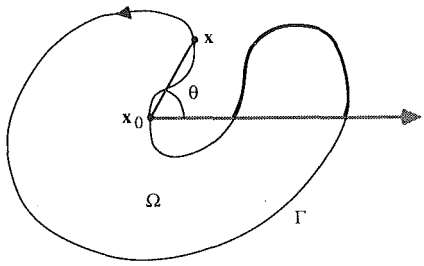


Fig. 3 Measurement of the angle θ in conjugate fundamental solution for a simply-closed boundary

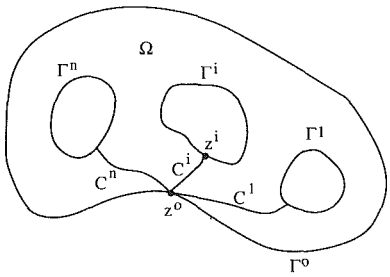


Fig. 4 Mutiply-connected domain

Multiply-Connected Domain

Unlike the conventional BIE, the derivative BIE just described needs a slight modification for a multiply-connected domain because of the multivaluedness of conjugate fundamental solution H.

Consider a multiply-connected domain as shown in Fig. 4, where Γ^i, $i=1, 2, \ldots n$ denote inner holes and Γ^o denotes the outer boundary. Let $\Gamma \equiv \sum_{i}^{n} \Gamma^i \cup \Gamma^o$. Following the usual treatment, the problem can be considered as simply-connected by introducing an arbitrary cut C^i from each inner hole to a common point on the outer boundary. Then, the integration by parts of Cauchy's formula (3) gives

$$\frac{1}{2\pi i}\int_\Gamma \ln(z-z_0)\phi'(z)dz = \begin{cases} 0, & z_0 \in \Gamma^o \\ \phi(z^o)-\phi(z^i), & z_0 \in \Gamma^i,\ i=1, 2, \ldots, n \end{cases} \tag{23}$$

where z^i and z^o are the end points of the cut C^i on Γ^i and Γ^o, respectively. Note that the integration on the inner boundary Γ^i is done in the clockwise direction. When z_0 is on Γ^o the equation (23) is identical to (9) of the simply-connected one, which bears no further difficulty. On the other hand, when z_0 is on Γ^i, the difference of complex potential between z^i and z^o appears on the right-hand side of equation (23) because of multivaluedness of logarithmic function. Therefore, the equations when $\mathbf{x}_0$ is on the inner boundary Γ^i should be modified accordingly as follows:

$$\left.\begin{aligned} &\int_\Gamma G(\mathbf{x},\mathbf{x}_0)p(\mathbf{x})ds + \int_\Gamma H(\mathbf{x},\mathbf{x}_0)q(\mathbf{x})ds = u_d^i,\ \mathbf{x}_0 \in \Gamma_q^i \\ &\int_\Gamma G(\mathbf{x},\mathbf{x}_0)q(\mathbf{x})ds - \int_\Gamma H(\mathbf{x},\mathbf{x}_0)p(\mathbf{x})ds = -w_d^i,\ \mathbf{x}_0 \in \Gamma_p^i \end{aligned}\right\} \tag{24}$$

where Γ_q^i and Γ_p^i denote the Dirichlet and the Neumann subboundaries of inner boundary Γ^i, with $\Gamma_q^i \cup \Gamma_p^i \equiv \Gamma^i$, and u_d^i and w_d^i denote changes of u and w from those at $\mathbf{x}^i$ to $\mathbf{x}^o$ on the cut C^i, respectively.

Since the cut is arbitrary, one can choose the end points of the cut to belong to a segment of the Dirichlet boundary where the potential is prescribed; otherwise, u_d^i will not appear because $\Gamma_q^i = \varnothing$ in equation (24). Then u_d^i becomes known; hence, there is actually one more unknown w_d^i, requiring an additional equation for each inner boundary. As in the simply-connected case, the following relation between q and u must be imposed for every segment of the Neumann boundary:

$$\int_{\Gamma_{p_j}} q \, ds = \Delta u^j \tag{25}$$

where Δu^j denotes the prescribed difference of u between the starting and the ending point of the jth Neumann boundary patch Γ_{p_j}. These equations satisfy the condition for uniqueness and also determine w_d^i. Therefore, the complete solution is obtained by solving equations (24) and (25) simultaneously. It is noted that the Γ-integrals in equation (24) are performed over a single-valued branch, while there is a jump in argument across the cut and at $\mathbf{x}_0$.

Numerical Implementation

Numerical implementation of the present formulation is essentially the same as the conventional one—by introducing suitable boundary elements and integration. However, because of the argument measurement and the end points of the cuts introduced, additional bookkeeping is necessary. Furthermore, the resulting matrix has a somewhat different shape from that in the usual BEM. The block matrix in bold letters in equation (20) can be singular, although the whole matrix is not. This may add some numerical difficulty, which has not, as yet, been looked at in this paper.

As in the conventional BEM, the numerical treatment of the corner points should be made with care, and needs more study. In the conventional boundary elements, three variables—the potential and its two derivatives in each normal direction on the corner—are defined on the corner node, where two out of three are prescribed, making the solution possible for the remaining one (Banerjee and Butterfield, 1981). However, in the present case four variables—two

Table 1 Results for Problem 2 at selected points

	Potential	Tangential Derivative q along AB			Flux p along BC		Total
	u at E	A	E	B	B	F	Flux
Exact Solution	.5850	1.443	.9618	.7214	.7214	.7214	.000
Conv. QUA	.5854	1.408	.9414	.7094	.7156	.7200	.76(−2)
Deriv. LIN	.5875	1.463	.9556	.6854	.7410	.7168	.28(−1)
Deriv. QUA	.5849	1.437	.9623	.7209	.7222	.7215	.63(−3)

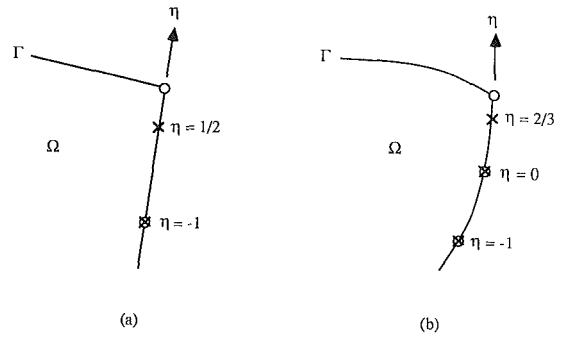


Fig. 5 Partially discontinuous elements; (a) linear element, (b) quadratic element. Symbol " o " denotes the geometric node and symbol " × " denotes the collocation point.

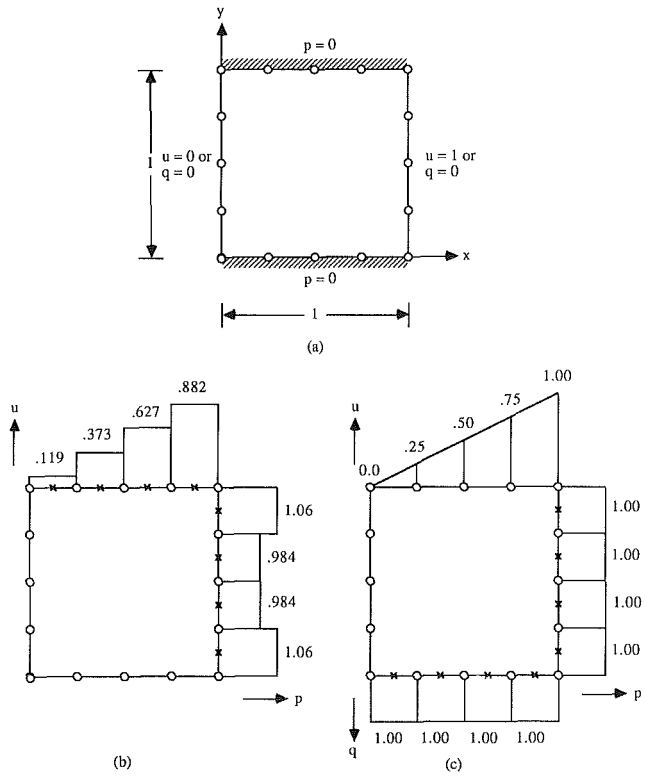


Fig. 6 Heat conduction in a rectangle; (a) problem definition, (b) solution by conventional BEM, and (c) solution by derivative BEM

derivatives in each tangential and normal direction—are defined, and still only two are known on the corner, which makes the equation indeterminate. To avoid this problem, partially discontinuous elements (Patterson and Sheikh, 1984) can be introduced, as shown in Fig. 5, which is to locate a collocation point near the corner instead of on the corner. Another approach, which has been suggested by one of the referees, is to enforce the continuity of ∇u at a corner and to use the equation

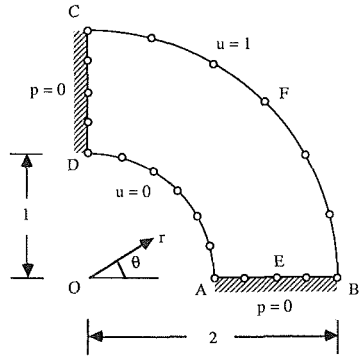


Fig. 7 Heat conduction in an annular cylinder

$$u_{,i} = \frac{\partial u}{\partial n} n_i + \frac{\partial u}{\partial s} s_i.$$

Example Applications

Four examples are presented to illustrate the use of the derivative BEM, which consider only the simply-connected domain. For each problem, constant, linear, and quadratic elements are tested, and the results are compared to analytic solutions when possible and those by the conventional boundary elements otherwise. Also, the total flux, which should be theoretically zero, is calculated for each problem to check the solution quality.

Throughout the following results, potential values by the present method are one-order smoother than the solutions p and q, and the corner values of p and q are the extrapolated values obtained from the partially-discontinuous elements.

Problem 1: Heat Conduction in a Rectangle. The boundary of this problem is discretized using 16 nodes with boundary conditions, as shown in Fig. 6(a), which has a linearly-varying solution for the temperature in the x direction. Computed results using constant elements are given in Fig. 6(b) and 6(c) for the conventional and present method, respectively. In the conventional methods, q can not be calculated, since u is constant on each element. Figure 6(c) shows that the results match the exact solution, and u varies linearly, while p and q are constant on each element. From these observations, some nature of the present method can be clearly understood.

Problem 2: Heat Conduction in an Annular Cylinder. An annular cylinder is subjected to different constant temperatures on the inner and outer boundaries. Because of its symmetry only one quarter is considered, with the geometry and the boundary conditions as given in Fig. 7. The results at some selected points as well as the total flux are given in Table 1, which shows that the present method gives better accuracy than the conventional one for both p and q values. Note in this table that the tangential derivatives by quadratic elements of conventional BEM are averaged values of numerical differentiation of the solution for u between each adjacent elements.

Table 2 Results for Problem 3 at selected points

	Potential	Tangential Derivative q along DE			Flux p along AE			Total
	u at F	E	F	D	A	G	E	Flux
Conv. QUA (refined)	.5874	2.210	1.762	1.303	.0013	1.553	2.216	.42(−6)
Conv. QUA	.5875	2.180	1.740	1.292	.0135	1.549	2.216	.16(−3)
Deriv. LIN	.5915	2.255	1.758	1.296	−.051	1.565	2.216	.51(−2)
Deriv. QUA	.5874	2.210	1.764	1.303	−.007	1.547	2.212	.14(−3)

Table 3 Results for Problem 4 at selected points

	Potential			Flux p		Total
	B	C	H	G	I	Flux
Conv. QUA (refined)	.6667	.9008	.5496	.2948	−.5787	.12(−4)
Jaswon and Symm	.6667	.9009	.5495	–	–	–
Conv. LIN	.6672	.9025	.5513	.2937	−.5810	.11(−1)
Conv. QUA	.6665	.9003	.5496	.2966	−.5783	.93(−3)
Deriv. LIN	.6665	.9003	.5492	.2949	−.5772	.37(−3)
Deriv. QUA	.6665	.9009	.5499	.2949	−.5774	.37(−3)

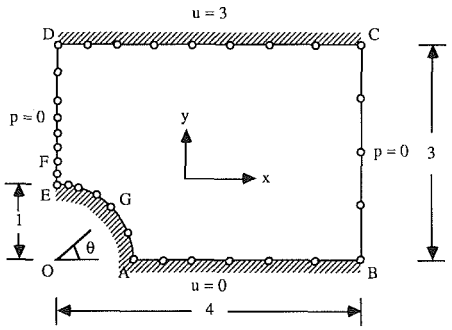


Fig. 8 Potential flow past a cylinder in a channel

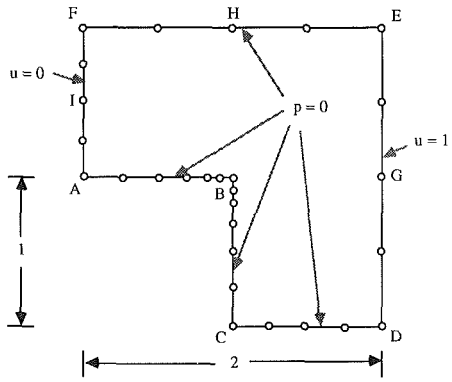


Fig. 9 Potential problem with reentrant corner

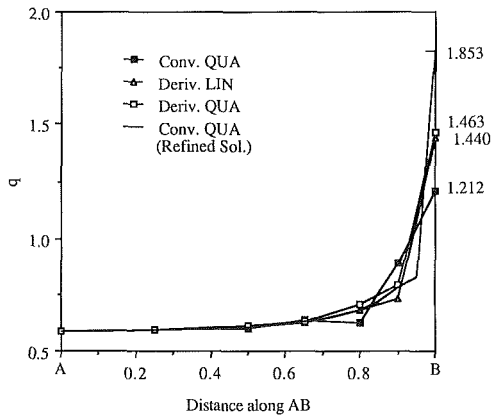


Fig. 10 Tangential derivative of the potential along AB in Fig. 9

Problem 3: Potential Flow Past a Cylinder in a Channel. Consider a flow past a cylinder located between two flat plates. The potential u in this problem actually denotes the stream function. With properly specified boundary conditions, only one quarter of the domain is considered, and 32 nodes are used for this geometry as shown in Fig. 8. As this problem has no analytical solution, the accuracies of the derivative BEM are compared with a solution obtained from a conventional quadratic element model which is four times denser than the test model here. The results at some selected points are given in Table 2. While the values for potential and flux show similar accuracy in both methods, the tangential derivatives show excellent accuracy compared to the conventional ones.

Problem 4: Potential Problem With Reentrant Corner. The geometry and the boundary conditions treated are shown in Fig. 9, where 28 nodes are used for discretization. This problem has a reentrant corner, which gives rise to a singularity for the derivatives at point B. As in the previous problem, the results from a four times more refined model by the conventional BEM are used as a reference of comparison. The results are also compared with those obtained by Jaswon and Symm (1977), where this problem is solved by a conventional method taking into special account the singularity at B. The values of the potential and the flux at some selected points are given in Table 3, in which better accuracy by the present method is observed for the potential, and also seen is the good agreement with those given by Jaswon and Symm (1977). The tangential derivatives are shown in Fig. 10. The singular behavior at B is better approximated by the present method with both linear and quadratic elements than the conventional quadratic element model.

Summary and Discussions

A new boundary integral equation, called derivative BIE, is developed for two-dimensional potential problems by integrating by parts the Cauchy formula and employing tangential and normal derivatives of the potential on the boundary as the primary unknown variables. This formulation has several advantages over the conventional method in that the tangential derivative along the boundary is obtained directly from the generated equations, and the potential value is calculated afterwards by integrating the solution obtained, which can improve the accuracy of the derivatives as well as the potential itself. In the developed BIE, a new kernel called conjugate fundamental solution is introduced, which is regular, but a careful treatment is needed because of its multivaluedness. To ensure uniqueness of the solution an auxiliary condition is necessary, relating the unknown derivatives to the prescribed potential information. The expression is simple when the potential is assumed continuous throughout the boundary. General problems with multivalued or discontinuous potential, such as the flow with nonzero circulation, can also be treated with some modifications. Multiply-connected domains can be treated as well but with additional terms in the formulation. Extra bookkeeping efforts for the computational implementation may be necessary as compared to the conventional method. Although not studied yet, the method, being general in nature, can be applied advantageously to many important applications, especially when accuracy of the tangential derivatives is crucial such as in a shape design sensitivity analysis.

Four examples are presented to illustrate the behavior of the solutions. Comparisons with the conventional method show that the present method has provided better accuracy not only for the potential but also for the flux values. In particular, for the problem with a reentrant corner where a singularity exists, the present method shows significantly better distribution of the derivative, even with the rather coarse linear element model.

Acknowledgment

While the essential idea was established when the authors were at Korea Advanced Institute of Science and Technology the first author, in particular, would like to thank Prof. Edward J. Haug, Director of Center for Computer Aided Design, for allowing a good environment conducive to the successful finishing of the present manuscript, and he wishes to acknowledge the support of the Korea Science and Engineering Foundation for his stay.

References

Athanasiadis, G., 1985, "Direct and Indirect Boundary Element Methods for Solving the Heat Conduction Problem," *Computational Methods in Applied Mechanics and Engineering*, Vol. 49, pp. 37–54.

Banerjee, P. K., and Butterfield, R., 1981, *Boundary Element Methods in Engineering Science*, McGraw-Hill, New York.

Carrier, G. F., Krook, M., and Pearson, C. E., 1966, *Functions of Complex Variable*, McGraw-Hill, New York.

Ghosh, N., Rajiyah, H., Ghosh, S., and Mukherjee, S., 1986, "A New Boundary Element Method Formulation for Linear Elasticity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 69–76.

Hildebrand, F. B., 1965, *Methods of Applied Mathematics*, Prentice-Hall, Englewood Cliffs, N.J.

Jaswon, M. A., and Symm, G. T., 1977, *Integral Equation Methods in Potential Theory and Elastostatics*, Academic Press.

Katz, C., 1982, "Some Improvements in 2D Boundary Elements Using Integration by Parts," *Proceedings of 4th International Seminar on Boundary Elements*, C. A. Brebbia, ed., Springer-Verlag, pp. 72–85.

Kwak, B. M., and Choi, J. H., 1987, "Shape Design Sensitivity Analysis Using Boundary Integral Equation for Potential Problems," *Computer Aided Optimal Design: Structural and Mechanical Systems*, C. A. Mota Soares, ed., Springer-Verlag, pp. 633–642.

Liggett, J. A., and Liu, L-F., 1982, *The Boundary Integral Equation Method Applied to Flow in Porous Media*, Allen and Unwin.

Patterson, C., and Sheikh, M. A., 1984, "Interelement Continuity in the Boundary Element Method," *Topics in Boundary Element Research, Vol. 1: Basic Principles and Applications*, Springer-Verlag, pp. 123–141.

D. J. Steigmann[1]

A. C. Pipkin

Division of Applied Mathematics,
Brown University,
Providence, R.I. 02912

Wrinkling of Pressurized Membranes

1 Introduction

Real membranes often exhibit wrinkled regions, which are not predicted by membrane theory in its usual form. The configuration of the wrinkled region is controlled by the small bending stiffness of the material. In membrane *theory*, as distinct from the theory of plates and shells, this bending stiffness is treated as zero, so the theory cannot give the details of the deformation in a wrinkled region.

The onset of wrinkling is associated with the appearance of compressive stresses in membrane solutions. Since states with compressive stress are unstable (Steigmann, 1986), such solutions are not physically meaningful. To obtain a solution with no compressive stress without resort to plate theory, tension field theory can be used (Wagner, 1929; Reissner, 1938; Kondo et al., 1955). However, particularly when a membrane is only partly wrinkled, it is usually not simple to decide which theory to use in a given part of the membrane.

It has recently been shown (Pipkin, 1986) that tension field theory can be incorporated into ordinary membrane theory simply by replacing the strain energy function by a suitable *relaxed* energy density, and then proceeding as usual with the equations for finite elastic deformations. In wrinkled regions, the relaxed energy density represents the average energy per unit initial area over a region containing many wrinkles. The stress-strain relation obtained from a relaxed energy function gives stresses that are never compressive, but in states of strain for which wrinkling is indicated, one principal stress component is zero and the theory reproduces all of the main assumptions of tension field theory.

A solution within this theory gives only the average deformation in a wrinkled region, with nothing to indicate the detailed structure of the wrinkles. The designation of a particular region as wrinkled can be deferred until after the solution has been obtained, but this final step is necessary for proper physical interpretation of the mathematical result.

In the present paper we outline the theory of finite elastic deformations of isotropic membranes, subject to edge loads and normal pressure. Here we consider only membranes that are planar in the stress-free reference state, or that are initially developable surfaces, such as cylinders and cones. Kinematics are discussed in Section 2, and the stress-strain relations and equilibrium equations are outlined in Section 3. In Section 4 we describe the main features of a relaxed energy density, and in particular we give the kinematic criterion by which a given state of deformation can be classified as wrinkled or not.

Within the theory, a wrinkled region is simply defined as a region in which the stress is uniaxial at each point. In such regions the deformation has some special features, which we describe in Section 5. In particular, the two fundamental equations for the tension direction and the corresponding stretch, which we have derived elsewhere (Steigmann and Pipkin) for the special case of no normal pressure, remain valid for inflated membranes.

The special features of the deformation in a wrinkled region can be used to construct solutions when the approach by ordinary elasticity theory might be difficult, as we have shown elsewhere (Steigmann and Pipkin). In the present paper, however, we make no use of these special features. We solve two problems without any reference to tension field theory, and then identify certain regions of the deformed membrane as wrinkled.

The problem in Section 6 concerns an infinite strip that is sheared and subjected to a uniform pressure. The deformation is controllable in this case. That is, the deformation can be specified in advance, and the resulting stress field is in equilibrium regardless of the form of the strain energy function. For materials that are incompressible in bulk, we show that the deformation represents a wrinkled state whenever the amount of shear exceeds a certain definite function of the curvature. This explicit criterion is valid for all isotropic, incompressible materials.

In Section 7 we consider the inflation of a semi-infinite tube that is tied off at one end. The solution is easily obtained by Kydoniefs' method (Kydoniefs, 1969). We then show that near the tied-off end, the solution represents a wrinkled state. The shape of the membrane in the wrinkled region is described by a specific function whose form is independent of material properties; such properties enter only through a certain scale factor.

2 Kinematics and Notation

We consider deformations of a membrane that occupies

[1]Presently at the Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta T6G 2G8, Canada.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 19, 1988; final revision, November 2, 1988.

Paper No. 89-WA/APM-10.

some region of the x_1, x_2 plane in its reference state. Let $\mathbf{e}_i$ $(i=1, 2, 3)$ be unit vectors in the coordinate directions. Then the reference position of a particle is $\mathbf{x}=x_a\mathbf{e}_a$, where summation over the range $a=1, 2$ is implied. In a deformation of the membrane, the particle initially at $\mathbf{x}$ goes to the place $\mathbf{r}=r_i\mathbf{e}_i$, where summation over the range $i=1, 2, 3$ is implied. The gradient ∇ is the two-dimensional operator

$$\nabla=\mathbf{e}_a\partial/\partial x_a. \tag{1}$$

The deformation gradient $\mathbf{F}$, defined by $d\mathbf{r}=\mathbf{F}\cdot d\mathbf{x}$, is defined equally well in terms of its transpose $\mathbf{F}^t$ by

$$\mathbf{F}^t=\nabla\mathbf{r}=\mathbf{e}_a(\partial/\partial x_a)r_i\otimes\mathbf{e}_i. \tag{2}$$

We use dyadic notation, in which $\mathbf{a}\otimes\mathbf{b}$ is the tensor whose ij-component is $(\mathbf{a}\otimes\mathbf{b})_{ij}=a_ib_j$.

By an extension of the polar decomposition theorem (Pipkin, 1986), the deformation gradient can also be expressed in terms of principal directions and principal stretches as

$$\mathbf{F}^t=\nabla\mathbf{r}=\lambda\mathbf{v}\otimes\mathbf{u}+\mu\mathbf{v}^*\otimes\mathbf{u}^*. \tag{3}$$

Here $\mathbf{v}$ and $\mathbf{v}^*$ are orthogonal unit vectors in the x_1, x_2 plane, while $\mathbf{u}$ and $\mathbf{u}^*$ are orthogonal unit vectors tangential to the deformed surface. The principal stretches λ and μ are nonnegative by definition.

We use the strain tensor $\mathbf{g}$ defined by

$$\begin{aligned}\mathbf{g}&=\mathbf{F}\cdot\mathbf{F}^t=(\nabla\mathbf{r})^t\cdot(\nabla\mathbf{r})\\&=\lambda^2\mathbf{u}\otimes\mathbf{u}+\mu^2\mathbf{u}^*\otimes\mathbf{u}^*.\end{aligned} \tag{4}$$

We also use the unit tensor δ on the deformed surface, which can be defined in terms of $\mathbf{u}$ and $\mathbf{u}^*$ by

$$\delta=\mathbf{u}\otimes\mathbf{u}+\mathbf{u}^*\otimes\mathbf{u}^*. \tag{5}$$

Let I and J be the isotropic strain invariants defined by

$$J=\lambda\mu=(\det\mathbf{g})^{1/2} \tag{6}$$

and

$$I=\lambda+\mu=(\text{tr }\mathbf{g}+2J)^{1/2}. \tag{7}$$

Finally, let $\mathbf{n}$ be a unit vector normal to the deformed surface,

$$\mathbf{n}=\mathbf{u}\times\mathbf{u}^*. \tag{8}$$

3 Stress and Equilibrium

We consider isotropic elastic membranes, for which there is a strain energy W per unit initial area that can be expressed as a symmetric function of the stretches λ and μ. W can equally well be expressed as a function of the symmetric invariants I and J. Let f_1 and f_2 be the principal forces,

$$f_1=\partial W/\partial\lambda,\ f_2=\partial W/\partial\mu. \tag{9}$$

These are the normal forces that must be applied to a unit square of material to stretch it into a rectangle of dimensions λ and μ. The stresses σ_1 and σ_2, which are the forces per unit current length in the deformed state, are

$$\sigma_1=f_1/\mu,\ \sigma_2=f_2/\lambda. \tag{10}$$

Let $\mathbf{T}$ be the engineering stress, given by

$$\mathbf{T}=f_1\mathbf{u}\otimes\mathbf{v}+f_2\mathbf{u}^*\otimes\mathbf{v}^*. \tag{11}$$

Then in the deformed state, the force across an element that had length ds and rightward normal ν in the undeformed state is

$$\mathbf{T}\cdot\nu ds=f_1\mathbf{u}(\mathbf{v}\cdot\nu ds)+f_2\mathbf{u}^*(\mathbf{v}^*\cdot\nu ds). \tag{12}$$

Let σ be the ordinary stress, which measures force per unit current length in the deformed configuration:

$$\sigma=(f_1/\mu)\mathbf{u}\otimes\mathbf{u}+(f_2/\lambda)\mathbf{u}^*\otimes\mathbf{u}^*. \tag{13}$$

This is related to $\mathbf{T}$ by

$$\sigma=J^{-1}\mathbf{T}\cdot\mathbf{F}^t, \tag{14}$$

where J and $\mathbf{F}^t$ are defined by (6) and (3). By using (4) and (5), it can be shown that

$$\sigma=(W_I/IJ)\mathbf{g}+(W_I/I+W_J)\delta, \tag{15}$$

where

$$W_I=\partial W/\partial I \text{ and } W_J=\partial W/\partial J. \tag{16}$$

We consider the equilibrium of a membrane loaded by a pressure $p\mathbf{n}$ per unit of deformed area. We note that J in (6) is the deformed area per unit initial area. Then equilibrium of an arbitrary part of the membrane requires that

$$\oint\mathbf{T}\cdot\nu ds+\int\int p\mathbf{n}Jdx_1dx_2=\mathbf{0}, \tag{17}$$

in which the integrals are taken over an arbitrary part of the initial domain and over its perimeter. By using the divergence theorem and the arbitrariness of the region, we obtain the pointwise equilibrium equation

$$\nabla\cdot\mathbf{T}^t+pJ\mathbf{n}=\mathbf{0}. \tag{18}$$

4 The Relaxed Energy Density

Real membranes have some small bending stiffness, which is neglected when membrane theory is used. The deformation of a real membrane may exhibit a wrinkled region, in which the size and configuration of the wrinkles is determined by the bending stiffness. Membrane theory, in which the stiffness is zero, cannot predict the details of such a deformation. However, the mean deformation and stress in a wrinkled region can be predicted from ordinary membrane theory if the strain energy function is replaced by a suitable *relaxed* energy density (Pipkin, 1986). The stress derived from a relaxed energy density is never compressive, and the kind of instability that would be produced by compressive stress is taken into account automatically.

For $\lambda>1$, let $w(\lambda)$ be the smallest value of the stretch μ such that $f_2>0$ when $\mu>w(\lambda)$. We call $w(\lambda)$ the *natural width* in simple tension. For $\mu=w(\lambda)$, the material is in a state of simple tension with

$$f_2=0,\ f_1=f_1[\lambda,\ w(\lambda)]=f(\lambda)\text{(say)}. \tag{19}$$

For example, for any material that is incompressible in bulk, the stretch of the membrane in the thickness direction is $1/\lambda\mu$, and in simple tension this is equal to μ, so

$$w(\lambda)=\lambda^{-1/2}. \tag{20}$$

The force $f(\lambda)$ is proportional to the force that would be required to stretch a string of the material to λ times its initial length.

For $\mu<w(\lambda)$ (with $\lambda>1$), where the given strain energy function might yield a compressive stress $f_2<0$, the relaxed energy density is defined to be equal to its value at the natural width,

$$W(\lambda,\mu)=W[\lambda,w(\lambda)](\lambda>1,\ \mu\leqq w(\lambda)). \tag{21}$$

Because of the symmetry of W with respect to λ and μ, similarly

$$W(\lambda,\mu)=W[w(\mu),\mu]\ (\mu>1,\ \lambda\leqq w(\mu)). \tag{22}$$

When both λ and μ are less than unity, the given strain energy function might well predict both $f_1<0$ and $f_2<0$. In such cases the relaxed energy density is defined to be equal to $W(1,1)$, which we take to be zero.

$$W(\lambda,\mu)=0(\lambda\leqq 1,\ \mu\leqq 1). \tag{23}$$

From these properties of the relaxed energy density it follows that the principal forces satisfy

$$f_1>0, f_2>0 \text{ if } \mu>w(\lambda) \text{ and } \lambda>w(\mu), \tag{24}$$

$$f_1=f(\lambda), f_2=0 \text{ if } \lambda>1 \text{ and } \mu\leqq w(\lambda) \tag{25}$$

$$f_1 = 0,\ f_2 = f(\mu) \text{ if } \mu > 1 \text{ and } \lambda \leqq w(\lambda) \tag{26}$$

$$f_1 = f_2 = 0 \text{ if } \lambda \leqq 1 \text{ and } \mu \leqq 1. \tag{27}$$

5 Tension Field Theory

When a relaxed energy density is used, ordinary membrane theory incorporates tension field theory automatically. The stress and deformation in a tension field have special features that may facilitate the process of solution in some problems, although these features need not be used explicitly.

We say that a region of a deformed membrane in which $\mu < w(\lambda)$ is *wrinkled*, even though the deformation predicted by the theory will generally be perfectly smooth. In such a region the stress takes the form

$$\mathbf{T} = f(\lambda)\mathbf{u} \otimes \mathbf{v}, \tag{28}$$

and the force across every arc is parallel to $\mathbf{u}$:

$$\mathbf{T} \cdot \nu ds = f(\lambda)\mathbf{u}(\mathbf{v} \cdot \nu ds). \tag{29}$$

For arcs parallel to $\mathbf{v}$ in the initial domain, $\mathbf{v} \cdot \nu = 0$, and no force is exerted across such an arc. The curves defined by the direction field $\mathbf{v}$ are the *stress trajectories*. Since $f > 0$ (when $\lambda > 1$), the force is purely tensile. The deformation carries the stress trajectories onto the curves with tangent $\mathbf{u}$ on the deformed surface. These curves are the *tension lines*.

In a wrinkled region, the equilibrium equation (18) reduces to the form

$$\nabla \cdot (f\mathbf{v})\mathbf{u} + f(\mathbf{v} \cdot \nabla)\mathbf{u} + pJ\mathbf{n} = \mathbf{0}. \tag{30}$$

Since both $\mathbf{n}$ and $\mathbf{v} \cdot \nabla \mathbf{u}$ are orthogonal to $\mathbf{u}$, this implies that

$$\nabla \cdot (f\mathbf{v}) = 0 \tag{31}$$

$$f(\mathbf{v} \cdot \nabla)\mathbf{u} = -pJ\mathbf{n}. \tag{32}$$

The former result expresses the fact that the amount of force channeled between two stress trajectories is constant along their length.

Let $\partial/\partial s$ be the derivative with respective to arc length along a tension line. In terms of this arc length the stretch λ is $\mathbf{v} \cdot \nabla s$, and since $J = \lambda\mu$, (32) can be written as

$$(f/\mu)\partial \mathbf{u}/\partial s = -p\mathbf{n}. \tag{33}$$

Here f/μ is the tension (per unit deformed length) in a tension line.

Because the change in $\mathbf{u}$ along a tension line is purely in the normal direction, each tension line is a geodesic on the deformed surface (Zak, 1982). It is useful to parameterize the deformed surface by a system of geodesic coordinates (Struik, 1950). Let ψ be a parameter that numbers the tension lines. On the deformed surface, select some curve that is orthogonal to the tension line at each point. Let ϕ be the distance of an arbitrary point from this base curve, measured along the geodesic that passes through that point. Then the curves $\phi = \text{constant}$ are *geodesic parallels*, and each of these curves is orthogonal to the geodesics (Struik, 1950). With position on the deformed surface given as a function of ϕ and ψ, we have

$$\partial \mathbf{r}/\partial \phi = \mathbf{u},\ \partial \mathbf{r}/\partial \psi = m\mathbf{u}^* (\text{say}), \tag{34}$$

where $\mathbf{u}^*$ is the second principal direction, orthogonal to $\mathbf{u}$.

The deformation is specified by giving ϕ and ψ as functions of $\mathbf{x}$. The deformation gradient is then

$$\nabla \mathbf{r} = \nabla \phi \otimes \partial \mathbf{r}/\partial \phi + \nabla \psi \otimes \partial \mathbf{r}/\partial \psi. \tag{35}$$

With (34), comparison with (3) shows that

$$\lambda \mathbf{v} = \nabla \phi \text{ and } \mu \mathbf{v}^* = m \nabla \psi, \tag{36}$$

where m is the magnitude of $\partial \mathbf{r}/\partial \psi$.

From (31) and (36) we obtain two fundamental equations for λ and $\mathbf{v}$,

$$\nabla \cdot [f(\lambda)\mathbf{v}] = 0,\ \nabla \times (\lambda \mathbf{v}) = \mathbf{0}. \tag{37}$$

These were derived for the special case $p = 0$ elsewhere (Steigmann and Pipkin). They can be reduced to a single second-order equation for ϕ by using (36). Alternatively, ψ can be chosen so as to represent a stress potential,

$$f(\lambda)\mathbf{v} = \nabla \psi \times \mathbf{e}_3, \tag{38}$$

and (37) then gives a second-order equation for ψ. The curves $\phi = \textit{constant}$ and $\psi = \textit{constant}$ are orthogonal in the initial domain as well as on the deformed surface, when they are treated as material lines.

The preceding results are of use in solving some problems. For the examples in Sections 6 and 7, however, we do not use any of these results explicitly. We merely solve the given problem and then, after the fact, identify certain states as wrinkled.

6 Example: Sheared and Pressurized Strip

As an example, consider a strip that initially occupies the region $-\infty < x < \infty$, $-H < y < H$ (here $x = x_1$ and $y = x_2$). We consider a deformation in which the strip is sheared parallel to its length and subjected to a uniform pressure on one side. The deformation has the form

$$\mathbf{r}(x,y) = (x + \kappa y)\mathbf{e}_1 + R\mathbf{i}(\theta) - \mathbf{e}_3 R\cos\theta_o, \tag{39}$$

where

$$\mathbf{i}(\theta) = \mathbf{e}_3\cos\theta + \mathbf{e}_2\sin\theta,\ \theta = \omega y,\ \theta_o = \omega H. \tag{40}$$

Here κ is the amount of shear. The deformed membrane is cylindrical, with radius of curvature R. The parameter ω is determined by the requirement that the edges $y = \pm H$ have no displacement in the y or z directions:

$$\omega H = \arcsin(H/R). \tag{41}$$

The deformation gradient is

$$\nabla \mathbf{r} = (\mathbf{e}_1 + \kappa \mathbf{e}_2) \otimes \mathbf{e}_1 + R\omega \mathbf{e}_2 \otimes \mathbf{j}(\theta), \tag{42}$$

where $\mathbf{j}(\theta) = \mathbf{i}'(\theta)$:

$$\mathbf{j}(\theta) = -\mathbf{e}_3\sin\theta + \mathbf{e}_2\cos\theta. \tag{43}$$

The vector $\mathbf{j}(\theta)$ is tangential to a parallel of latitude on the cylinder. The unit tensor on the deformed surface can be represented as

$$\delta = \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{j}(\theta) \otimes \mathbf{j}(\theta). \tag{44}$$

From (4), with (42), the strain $\mathbf{g}$ is

$$\mathbf{g} = (1 + \kappa^2)\mathbf{e}_1 \otimes \mathbf{e}_1 + \kappa R\omega(\mathbf{e}_1 \otimes \mathbf{j} + \mathbf{j} \otimes \mathbf{e}_1) + (R\omega)^2 \mathbf{j} \otimes \mathbf{j}. \tag{45}$$

Then, from (6) and (7), the invariants I and J are

$$I = [\kappa^2 + (1 + R\omega)^2]^{1/2},\ J = R\omega. \tag{46}$$

From the stress-strain relation (15), with (44) to (46), the shearing stress is

$$\mathbf{e}_1 \cdot \sigma \cdot \mathbf{j} = (W_I/I)\kappa, \tag{47}$$

and the x-component of the equilibrium equation is satisfied because this is independent of y. The hoop stress $\mathbf{j} \cdot \sigma \cdot \mathbf{j}$ is also constant, so the $\mathbf{j}$ component of the equilibrium equation is satisfied. The normal component of the equilibrium equation gives the pressure required to support the deformation:

$$pR = \mathbf{j} \cdot \sigma \cdot \mathbf{j} = (W_I/I)(R\omega + 1) + W_J. \tag{48}$$

We have verified that the specified deformation is an equilibrium configuration. However, for an arbitrary form of W, it might well be the case that one of the principal stress components is negative, so that the given state of deformation is unstable. When W is a relaxed energy density, compressive stress cannot occur, but one of the principal stresses may be zero for some deformations.

To determine whether or not the given deformation represents a wrinkled state, it is necessary to calculate the larger and smaller principal stretches λ and μ. Then, if

$\mu < w(\lambda)$, the given deformation represents a wrinkled state. Let us use the form $w(\lambda) = \lambda^{-1/2}$ that applies to all incompressible materials. Then the wrinkling criterion is $\mu^2\lambda < 1$, or $\mu J < 1$, since $J = \lambda\mu$. The stretches λ and μ are the larger and smaller roots of the equation

$$x^2 - Ix + J = 0. \tag{49}$$

From this, the stretch μ is easily determined. Then a rather lengthy algebraic manipulation brings the condition $\mu J < 1$ into the form

$$\kappa^2 > (1 + J^{-2})(J^2 - 1)^2. \tag{50}$$

Here, $J = R\omega$, with ω defined in terms of R by (41). The family of deformations is parameterized by the amount of shear κ and the curvature $1/R$. The wrinkling criterion (50), which is valid for all isotropic incompressible materials, compares the amount of shear κ to a certain function of the curvature. We note that $J \geqq 1$ in all cases, with $J = 1$ only when there is no pressure.

When the wrinkling criterion (50) is satisfied, we do not expect to observe a deformation of the form (39) experimentally. Instead, we expect to see a deformation like (39) with a periodic wave superposed on it. The principal direction $\mathbf{u}$ defines the direction parallel to the wave crests. Because $\mathbf{u}$ is an eigenvector of $\mathbf{g}$, it is easily determined from (45). With

$$\mathbf{u} = \mathbf{e}_1 \cos\alpha + \mathbf{j}(\theta)\sin\alpha, \tag{51}$$

the angle α is given by

$$\tan\alpha = (Q + J^2 - 1 - \kappa^2)/2J\kappa, \tag{52}$$

where Q is the positive root of

$$Q^2 = (1 + \kappa^2 + J^2)^2 - 4J^2. \tag{53}$$

The distance between crests of the wave should depend on the bending stiffness of the material, and membrane theory gives no information about this.

7 Example: Pressurized Tube

Now consider a membrane that initially has the form of a circular cylinder of radius R and length L, with L large in comparison to R. The location of a particle on the undeformed cylinder is specified by its cylindrical coordinates (R,θ,z). Because the cylinder is developable into the plane, the general remarks in Sections 2 to 5 are all applicable, with $x_1 = R\theta$ and $x_2 = z$.

We consider certain axisymmetric deformations without twist, in which the particle initially at (θ,z) goes to the place

$$\mathbf{r} = r(z)\mathbf{i}(\theta) + \zeta(z)\mathbf{k}, \tag{54}$$

where $\mathbf{i}(\theta)$ and $\mathbf{k}$ are unit vectors in the radial and axial directions. The element of arc length along the deformed meridian is defined by

$$ds^2 = dr^2 + d\zeta^2. \tag{55}$$

The principal stretches are

$$\lambda = ds/dz,\ \mu = r/R. \tag{56}$$

The principal directions corresponding to the stretch λ are $\mathbf{v} = \mathbf{k}$ and

$$\mathbf{u} = \mathbf{k}\cos\gamma + \mathbf{i}(\theta)\sin\gamma, \tag{57}$$

where

$$\sin\gamma = dr/ds. \tag{58}$$

We treat the tube as semi-infinite, in the region $0 \leq z < \infty$. The end $z = 0$ is contracted to zero radius, $r(0) = 0$, and sealed so that the tube can contain an internal pressure. If the inflation pressure is only moderate, for large z the tube can approach a uniform cylindrical state with constant stretches λ_1 and μ_1. The radius of the deformed cylinder in this limiting state is $r_1 = R\mu_1$. Equilibrium requires pr_1 to be equal to the hoop stress f_2/λ, so

$$pR = f_2(\lambda_1,\mu_1)/\lambda_1\mu_1. \tag{59}$$

For equilibrium under no externally-imposed force, the total force across a plane $\zeta = constant$ must be zero, so

$$2\pi R f_1 \cos\gamma - p\pi r^2 = 0. \tag{60}$$

Thus,

$$\cos\gamma = pR\mu^2/2f_1, \tag{61}$$

where we have replaced r by the dimensionless radius $\mu = r/R$. For $z \to \infty$, γ approaches zero and thus

$$pR = 2f_1(\lambda_1,\mu_1)/\mu_1^2. \tag{62}$$

We regard μ_1 as given. Then λ_1 is determined from the equation obtained by eliminating p from (59) and (62), which is equivalent to $2\sigma_1 = \sigma_2$. Then the boundary conditions to be used are

$$r(0) = \zeta(0) = 0,\ r \to R\mu_1 \text{ and } \zeta \sim \lambda_1 z \text{ as } z \to \infty. \tag{63}$$

Let $H(\lambda,\mu)$ be defined by

$$H(\lambda,\mu) = \lambda f_1(\lambda,\mu) - W(\lambda,\mu). \tag{64}$$

Equilibrium requires H to be constant (Pipkin, 1968). Thus,

$$H(\lambda,\mu) = H(\lambda_1,\mu_1). \tag{65}$$

We are following the procedure explained by Kydoniefs (1969). From (65) we obtain λ as a function of μ. (It is easy to show that $d\lambda/d\mu$ is finite along a curve $H = \text{constant}$ if $\partial f_1/\partial\lambda > 0$, so a single-valued locus $\lambda = \lambda(\mu)$ is obtained.) The function $\lambda(\mu)$ is substituted into (61), with p given by (59) or (62). Then (61) gives $\cos\gamma$ as a function of μ. With (55), (56), and (58), the various unknowns can then be found by integration. In particular, the shape of the deformed tube is given by

$$\zeta/R = \int_o^\mu \operatorname{ctn}\gamma(\mu')\,d\mu'. \tag{66}$$

We remark that this procedure fails if the expression (61) for $\cos\gamma$ is greater than unity at stretches near the state λ_1, μ_1, and we state without proof that this is the case when the uniform state λ_1, μ_1 is unstable, and only then. Stability has been discussed by Corneliussen and Shield (1961) and by Haughton and Ogden (1979). When λ_1 and μ_1 are only moderately greater than unity, the uniform cylindrical state is stable, and Kydoniefs' method is valid.

Now, let us consider the possibility that $\mu < w(\lambda)$ in some region, so that the deformation is wrinkled there. Because $\mu = 0$ at $z = 0$, the membrane is certainly wrinkled near the end $z = 0$. By setting $\mu = 0$ in (65) we obtain an equation for the stretch λ_o in the wrinkled region,

$$H(\lambda_o,0) = H(\lambda_1,\mu_1). \tag{67}$$

Because H is independent of μ for $\mu < w(\lambda)$, the stretch λ is constant at the value λ_o throughout the wrinkled region. In fact, for axisymmetric deformations of cylinders, λ is always constant in wrinkled regions (Steigmann and Pipkin).

With λ_o known, we evaluate the corresponding natural width $w(\lambda_o)$. Then the deformation is wrinkled in that part of the end region for which the dimensionless radius μ is less than $w(\lambda_o)$.

Because λ is constant and f_1 is independent of μ in a wrinkled region, then f_1 is constant at the value $f(\lambda_o)$. Let P be the constant defined by

$$P^2 = pR/2f(\lambda_o). \tag{68}$$

Then from (61),

$$\operatorname{ctn}\gamma = (P\mu)^2/[1 - (P\mu)^4]^{1/2}. \tag{69}$$

It follows that the shape (66) is given by

$$P\zeta/R = F(Pr/R), \tag{70}$$

where $F(x)$ is the function defined by the integral

$$F(x) = \int_o^x t^2(1-t^4)^{-1/2}dt. \tag{71}$$

For small x, $F(x)$ is given by

$$F(x) = (1/3)x^3 + (1/14)x^7 + (3/88)x^{11} + \ldots . \tag{72}$$

It is interesting that the end shape (70) depends on p and the properties of the material only through the scale factor P/R. Assuming that the interior of a sausage is approximately a fluid under uniform pressure, the function F defines the shape of the end of a sausage. This is the shape in the region where the sausage skin is wrinkled, near the place where it is tied off.

Acknowledgment

The work described in this paper was carried out under a grant NSF-8702866 from the National Science Foundation. We gratefully acknowledge this support.

References

Corneliussen, A. H., and Shield, R. T., 1961, "Finite Deformations of Elastic Membranes with Application to the Stability of an Inflated and Extended Tube," *ARMA*, Vol. 7, pp. 273–304.

Haughton, D. M., and Ogden, R. W., 1979, "Bifurcation of Inflated Circular Cylinders of Elastic Material under Axial Loading. I. Membrane Theory for Thin-Walled Tubes," *JMPS*, Vol. 27, pp. 179–212.

Kondo, K., Iai, T., Moriguti, S., and Murasaki, T., 1955, "Tension Field Theory," *Memoirs of the Unifying Study of the Basic Problems in Engineering Sciences by Means of Geometry*, Vol. 1, C-V, Gakujutsu Bunken Fukyu-Kai, Tokyo, pp. 61–85.

Kydoniefs, A. D., 1969, "Finite Axisymmetric Deformations of an Initially Cylindrical Elastic Membrane Enclosing a Rigid Body," *QJMAM*, Vol. 22, pp. 319–331.

Pipkin, A. C., 1968, "Integration of an Equation in Membrane Theory," *ZAMP*, Vol. 19, pp. 818–819.

Pipkin, A. C., 1986, "The Relaxed Energy Density for Isotropic Elastic Membranes," *IMA J. Appl. Math*, Vol. 36, pp. 85–99.

Reissner, E., 1938, "On Tension Field Theory," *Proc. 5th Int. Cong. Appl. Mech.*, pp. 88–92.

Steigmann, D. J., 1986, "Proof of a Conjecture in Elastic Membrane Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 955–956.

Steigmann, D. J., and Pipkin, A. C., "Finite Deformations of Wrinkled Membranes," *QJMAM*, forthcoming.

Struik, D. J., 1950, *Lectures on Classical Differential Geometry*, Addison-Wesley, Cambridge.

Wagner, H., 1929, "Ebene Blechwandtrager mit sehr Dunnem Stegblech," *Z. Flugtechnik u. Motorluftschiffahrt*, Vol. 20, pp. 200–207, 227–233, 256–262, 279–284, 306–314.

Zak, M., 1982, "Statics of Wrinkling Films," *J. Elast*, Vol. 12, pp. 51–63.

Simultaneous Optimization of Beams and Their Elastic Foundations for Minimum Compliance

R. H. Plaut
Professor,
Charles E. Via, Jr. Department of Civil Engineering,
Virginia Polytechnic Institute and State University,
Blacksburg, VA. 24061
Mem. ASME

Nonuniform beams on nonuniform elastic foundations are considered. The beams have sandwich cross-sections with cores of negligible stiffness and are subjected to a uniformly distributed load. The total volume of the beam and the total stiffness of the foundation (which may include elastic supports) are specified. Both the cross-sectional area of the beam and the stiffness distribution of the foundation are design functions. They are chosen to minimize the compliance (or, equivalently, the area displaced by the beam deflection function). The calculus of variations is used to derive optimality conditions, and results are obtained for cantilevers and pinned-pinned beams. Several types of solutions are found, involving a single elastic support or a region of uniform foundation bordered at internal locations by elastic supports. In comparison to a reference uniform beam with uniform foundation, the decrease in compliance is significant.

Introduction

Optimization of beams attached to elastic foundations has received some attention in the past. Typically, a uniform foundation is specified and the variation of the beam cross-section is optimized (e.g., Plaut, Johnson, and Olhoff, 1986). In six recent papers, however, the foundation has been optimized, rather than the beam. Szelag and Mróz (1979) minimized the total foundation stiffness for a specified fundamental frequency of free vibrations. In Taylor and Bendsøe (1984), a beam was displaced downwards and the foundation stiffness distribution was chosen to minimize the maximum pressure.

Dems, Plaut, Banach, and Johnson (1987) considered a foundation with piecewise-constant stiffness and minimized a measure of the beam deflection. In Plaut (1987), the compliance (i.e., the work done by the load) was minimized, and the optimal solutions involved elastic supports and regions of uniform foundation. In addition to beams, circular plates also were treated. Finally, optimal elastic foundations for maximum buckling load were determined in Shin, Haftka, and Plaut (1988) and Shin, Haftka, Watson, and Plaut (1988).

In the present paper, both the distribution of beam material and the distribution of foundation stiffness are optimized. A cantilever and a pinned-pinned beam with sandwich cross-sections are considered. The total volume of the beam and the total stiffness of the foundation are specified. A uniformly distributed load is applied, and the compliance is minimized. (For this loading, the compliance is proportional to the area displaced by the beam in the vertical plane.) Optimality conditions are derived using the calculus of variations, and optimal solutions are obtained for several values of minimum cross-sectional area and a range of values of total foundation stiffness.

Formulation

Consider an elastic beam of length L which has a sandwich cross-section. The face sheets have varying thickness and their cross-sectional area is denoted $A_f(X)$ where $0<X<L$. The mass and stiffness of the core are neglected. Let A_u and EI_u represent the face-sheet area and the bending stiffness of a reference uniform beam which has the same total volume. The beam is attached to an elastic foundation of the Winkler type with varying stiffness coefficient $K(X)$. Elastic supports with stiffnesses C_i at locations $X=L_i$ may be included in the foundation. A downward, uniformly distributed load q is applied to the beam, and the resulting downward deflection is denoted $W(X)$.

The analysis is carried out in terms of the nondimensional quantities

$$x=X/L, \quad \alpha(x)=A_f(xL)/A_u, \quad k(x)=L^4K(xL)/(EI_u),$$

$$c_i=L^3C_i/(EI_u), \quad a_i=L_i/L, \quad w(x)=EI_uW(xL)/(L^4q) \quad (1)$$

where $0<x<1$ (see Fig. 1). Then the equilibrium equation is given by

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, August 9, 1988; final revision, January 10, 1989.

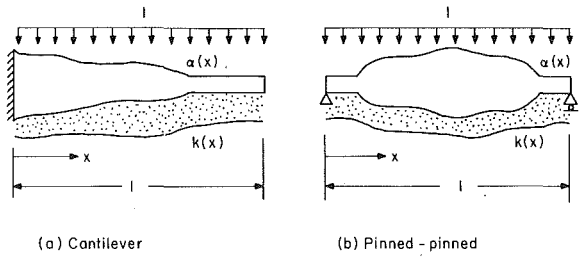


Fig. 1 Geometry of beams in nondimensional terms

$$[\alpha(x)w''(x)]'' + k(x)w(x) = 1 \tag{2}$$

(Plaut, Johnson, and Olhoff, 1986; Plaut, 1987). At $x = a_i$, if $0 < a_i < 1$, the functions w, w', and w'' are continuous, while

$$w'''(a_i^-) - w'''(a_i^+) = c_i w(a_i). \tag{3}$$

The constraint of given total volume is

$$\int_0^1 \alpha(x)dx = 1. \tag{4}$$

The specified (nondimensional) total foundation stiffness is denoted K_T, so that

$$\int_0^1 k(x)dx + \sum_{i=1}^{n} c_i = K_T \tag{5}$$

where n is the number of elastic supports. Additional constraints are given by

$$\alpha(x) \geq \alpha_m, \quad k(x) \geq 0, \quad c_i \geq 0, \quad a_i < a_{i+1} \tag{6}$$

where α_m is a minimum value for the (nondimensional) area. Also, $k(x)$ turns out to be continuous except at $x = a_i$.

The design variables are $\alpha(x)$, $k(x)$, c_i, and a_i, and the objective function to be minimized is the compliance G defined by

$$G = \int_0^1 w(x)dx \tag{7}$$

where $w(x) \geq 0$ for the beams to be analyzed. The following augmented functional G^* is constructed:

$$\begin{aligned} G^* = &\int_0^1 w dx + \int_0^1 \lambda[-(\alpha w'')'' - kw + 1]dx \\ &+ \mu\left(\int_0^1 k dx + \sum_{i=1}^{n} c_i - K_T\right) + \int_0^1 \beta(-k + \theta^2)dx \\ &+ \sum_{i=1}^{n} \gamma_i(-c_i + \phi_i^2) + \Lambda\left(\int_0^1 \alpha dx - 1\right) \\ &+ \int_0^1 \Gamma(\alpha - \alpha_m - \psi^2)dx + \sum_{i=1}^{n-1} \nu_i(a_i - a_{i+1} + \psi_i^2). \end{aligned} \tag{8}$$

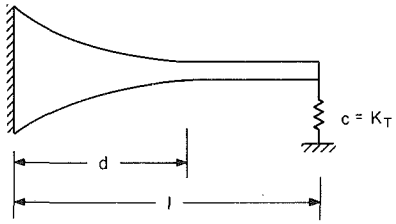
In equation (8), $\lambda(x)$, μ, $\beta(x)$, γ_i, Λ, $\Gamma(x)$, and ν_i are Lagrange multipliers, while $\theta(x)$, ϕ_i, $\psi(x)$, and ψ_i are slack variables. G^* is made stationary with respect to the design variables, Lagrange multipliers, and slack variables (Dems et al., 1987).

For the beams to be treated here, the resulting equation and boundary conditions for $\lambda(x)$ are the same as those for $w(x)$, so that $\lambda(x) = w(x)$. (This only occurs because the load is uniform and the objective function is given by equation (7).) Then the remaining stationary equations lead to the following optimality conditions, where Λ and μ are constants:

$$[w''(x)]^2 = \Lambda \text{ if } \alpha(x) > \alpha_m; \tag{9}$$

$$w^2(x) = \mu \text{ if } k(x) > 0; \tag{10}$$

$$w^2(a_i) = \mu, \quad w(a_i)w'(a_i) = 0 \text{ if } c_i > 0. \tag{11}$$



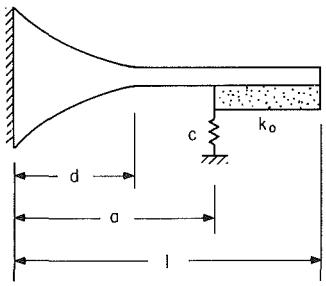


Fig. 2 Types of optimal solutions for cantilever

Optimal solutions are obtained with the use of equations (2)–(5), (9)–(11), and boundary conditions on $w(x)$. In the examples, Λ and μ are positive and $w'(a_i) = 0$. When $k(x) > 0$, the deflection is constant (equation (10)), so that $w''(x) = 0$, and then equation (9) requires that $\alpha(x) = \alpha_m$. It follows from equation (2) that $k(x)$ will be a constant. When $\alpha(x) > \alpha_m$, $w'' \neq 0$ (equation (9)) and $w(x)$ is not constant. Then equation (10) requires that $k(x) = 0$, and it follows from equations (2) and (9) that $\alpha(x)$ will be a quadratic function. (If $\alpha(x)$ were not proportional to the cross-sectional moment of inertia, equation (9) would involve $\alpha(x)$ and the optimal variation of the area would not be quadratic.)

Cantilever

Consider the cantilever shown in Fig. 1(*a*). The boundary conditions are $w = w' = 0$ at $x = 0$ and $\alpha w'' = (\alpha w'')' = 0$ at $x = 1$. If there is an elastic support with stiffness c at $x = 1$, the last of these conditions is replaced by $(\alpha w'')' = cw$ at $x = 1^-$.

Two types of optimal solutions are possible, as illustrated in Fig. 2. For sufficiently small values of the total foundation stiffness K_T, Type I occurs, in which the area α is nonuniform (decreasing) in a region $0 < x < d$, $\alpha = \alpha_m$ for $d \leq x \leq 1$, and the foundation consists of an elastic support at $x = 1$. As K_T is increased, $w'(1)$ becomes zero and then Type II governs, in which α is nonuniform (decreasing) in a region $0 < x < d$, $\alpha = \alpha_m$ for $d \leq x \leq 1$, an elastic support of stiffness c exists at some location $x = a$ $(d < a < 1)$, and there is a uniform foundation with stiffness k_0 for $a < x < 1$. The governing equations can be solved analytically for $w(x)$ and $\alpha(x)$, leading to a sixth-order polynomial equation to be solved for d in the Type I solution, and four coupled polynomial equations involving d, a, c, and k_0 in the Type II solution.

Numerical results are presented in Fig. 3 and Table 1 for $\alpha_m = 0.8$, 0.6, 0.4, and 0.2. Points A, B, and C correspond to $K_T = 0$, the transition from Type I optimal solutions to Type II solutions (marked by circles in Fig. 3), and $K_T = 50$, respectively. The maximum value of α, denoted $\alpha_{\max}$, occurs at $x = 0$ and is listed in Table 1. The last column in Table 1 compares the compliance associated with the optimal solution to that of a reference beam (Heténi, 1946) which has uniform cross-

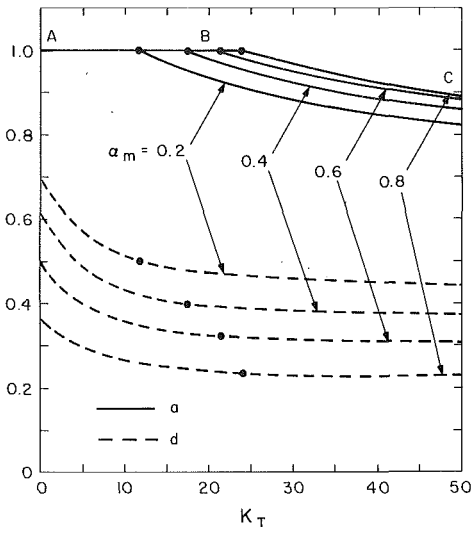


Fig. 3 Optimal values of d and a for cantilever, as a function of K_T; $\alpha_m = 0.8$, 0.6, 0.4, and 0.2

section (with the same total volume) and a uniform foundation (with the same total foundation stiffness).

When $K_T = 0$ (point A), only $\alpha(x)$ is a design variable. The value of d varies from 0.365 for $\alpha_m = 0.8$ to 0.695 for $\alpha_m = 0.2$, and the reduction in compliance ranges from 24.6 percent to 40.5 percent. Point B occurs in the range $11.8 \le K_T \le 24.2$ for the results presented, with the corresponding value of d varying from 0.236 for $\alpha_m = 0.8$ to 0.500 for $\alpha_m = 0.2$. With further increase in K_T, d decreases further and a also decreases. For these Type II solutions between B and C, the reduction in compliance is in the range 55.3–64.1 percent.

In the Type II solutions, the optimal ratio c/K_T is given by

$$c/K_T = (a-d)(a+2d)/[3(a+d) - 2(a^2+ad+d^2)]. \quad (12)$$

At point C ($K_T = 50$), this ratio has the values 0.707, 0.668, 0.600, and 0.490 for $\alpha_m = 0.8$, 0.6, 0.4, and 0.2, respectively.

Pinned-Pinned Beam

In this section, the pinned-pinned beam of Fig. 1(b) is considered. The boundary conditions are $w = \alpha w'' = 0$ at $x = 0, 1$. Figure 4 depicts the types of optimal solutions which occur, with Type I governing for small values of K_T, Type II for an intermediate range of K_T, and Type III for large values of K_T. The solutions are symmetric about $x = 1/2$. The area is at its minimum allowable value near the pinned ends and also, for Type II and Type III solutions, in a central region of the beam. In the first two types of solutions, the optimal foundation is a central elastic support, while in Type III it is a uniform foundation about the center of the beam, bordered by elastic supports.

As for the cantilever, one can obtain an analytical solution for the deflection $w(x)$ and area $\alpha(x)$, with polynomial equations to be solved for the design variables: b in the Type I optimal solution, b and d in Type II, and b, d, a, c, and k_0 in Type III. Optimal values of b, d, and a are plotted as a function of K_T in Fig. 5 for the case $\alpha_m = 0.8$. Similar curves are found for other values of α_m. Numerical values are listed in Table 2 for $\alpha_m = 0.8$, 0.6, 0.4, and 0.2.

When $K_T = 0$ (point A), the optimal value of b in Fig. 4(a) varies from 0.192 for $\alpha_m = 0.8$ to 0.035 for $\alpha_m = 0.2$, the value of $\alpha(1/2)$ varies from 1.29 to 1.49, and the reduction in compliance is in the range 8.7–16.3 percent. As K_T increases, b decreases and $\alpha(1/2)$ decreases. The transition from a Type I to a Type II optimal solution occurs at point B, when $\alpha(1/2)$ has the value α_m. The beam then has a form as in Fig. 4 (b) but

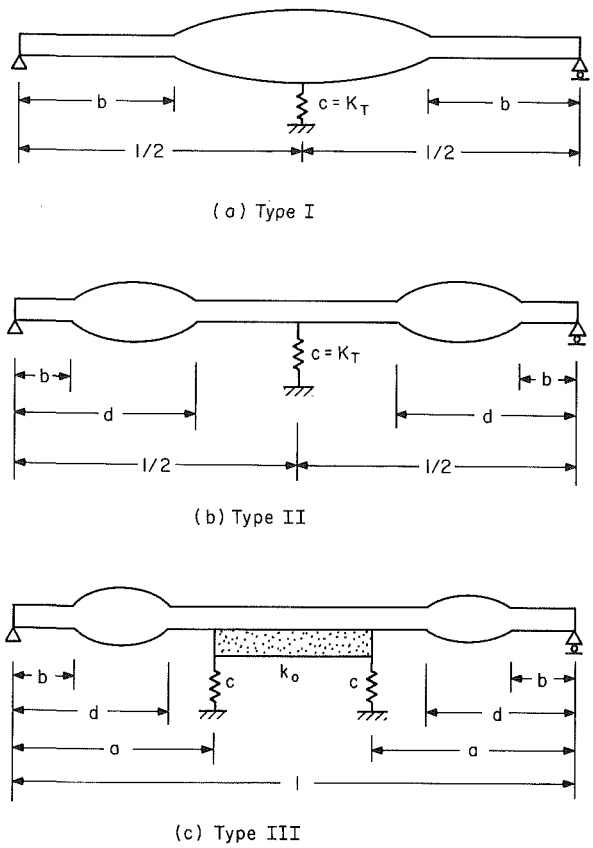


Fig. 4 Types of optimal solutions for pinned-pinned beam

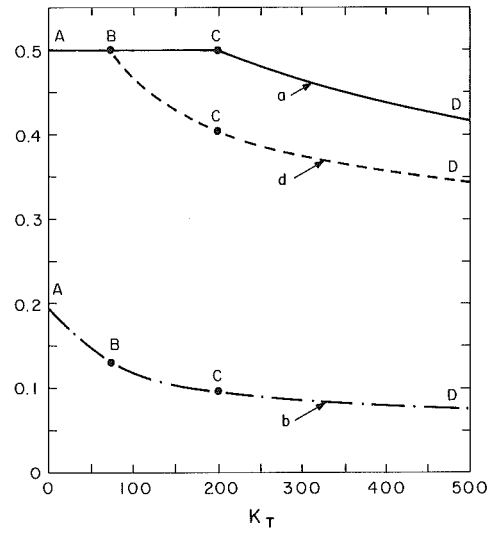


Fig. 5 Optimal values of b, d, and a for pinned-pinned beam, as a function of K_T; $\alpha_m = 0.8$

with $d = 1/2$, i.e., no central uniform region. At point B, the optimal value of b is a root of the equation

$$8\alpha_m b^3 - 12\alpha_m b^2 + 6(2-\alpha_m)b - \alpha_m = 0. \quad (13)$$

The corresponding total foundation stiffness varies from $K_T = 73$ for $\alpha_m = 0.8$ to $K_T = 166$ for $\alpha_m = 0.2$. It is noted that these values of K_T are beyond the K_T range shown in Fig. 3 for the cantilever.

From point B to point C, where the Type II solution of Fig. 4 (b) governs, the values of b and d decrease as K_T increases. The area function $\alpha(x)$ for $b \le x \le d$ is quadratic and sym-

metric about its center, $x=(b+d)/2$. Maximum values, α_{max}, are listed in Table 2.

As K_T increases beyond point C in Table 2 (and in Fig. 5 for $\alpha_m=0.8$), the Type III solution of Fig. 4(*c*) is optimal. The lengths b, d, and a all decrease as K_T increases. The area is symmetric in the region $0\le x\le a$, i.e., $a-d=b$ and α_{max} occurs at $x=a/2$.

The stiffness of the elastic supports in the Type III solutions satisfies the equation

$$c/K_T=a/[2(1-a]. \tag{14}$$

At point D, where $K_T=500$, the ratio c/K_T is 0.358, 0.364, 0.367, and 0.370 for $\alpha_m=0.8$, 0.6, 0.4, and 0.2, respectively. For the Type II and Type III solutions given in Table 2, the reduction in compliance is in the range 29.9–39.4 percent, in comparison with the reference beam.

Concluding Remarks

Beams attached to elastic foundations were considered in this paper. The distributions of beam material and foundation stiffness were optimized simultaneously, with total beam material and foundation stiffness specified. Cantilevers and pinned-pinned beams were treated. The beams had sandwich cross-sections with cores of negligible stiffness, and were subjected to a uniformly distributed load. The compliance was minimized.

Several types of optimal solutions were obtained, as illustrated in Figs. 2 and 4. The governing type depends on the minimum allowable cross-sectional area of the face sheets and the total foundation stiffness (relative to the beam stiffness). If the total foundation stiffness is sufficiently small, the optimal solution is a single elastic support; if it is sufficiently large, the optimal solution contains a region of uniform foundation, bordered at internal locations by elastic supports. Where the foundation exists, the beam is uniform (with its minimum area). For the pinned-pinned beam, the optimal area distribution satisfies special symmetry properties.

Even though the formulation allows the foundation to be nonuniform, the optimal solution does not involve a continuously-varying function $k(x)$. Either k is zero, a positive constant k_0, or a delta function (corresponding to an elastic support). This property is due to two factors, the uniform load and the objective function used in this study. The quadratic variation of the area when it is nonuniform is caused by these factors plus the assumption of a sandwich cross-section.

In comparison to a reference beam with uniform cross-section and uniform elastic foundation, the optimal solutions decrease the compliance significantly. Typical results are presented in the last column of Tables 1 and 2. The reduction in compliance for these cases is as high as 64.1 percent for cantilevers and 39.4 percent for pinned-pinned beams.

Table 1 Optimal values for cantilever

α_m	Point	d	a	K_T	α_{max}	Percent decrease
0.8	A	0.365	1	0	1.98	24.6
0.8	B	0.236	1	24.2	2.62	58.8
0.8	C	0.227	0.890	50.0	2.94	55.3
0.6	A	0.500	1	0	2.40	34.0
0.6	B	0.322	1	21.5	3.35	61.7
0.6	C	0.305	0.880	50.0	3.93	58.1
0.4	A	0.615	1	0	2.69	39.6
0.4	B	0.401	1	17.5	3.81	63.4
0.4	C	0.371	0.859	50.0	4.75	59.9
0.2	A	0.695	1	0	2.89	40.5
0.2	B	0.500	1	11.8	4.00	64.1
0.2	C	0.442	0.822	50.0	5.51	60.8

References

Dems, K., Plaut, R. H., Banach, A. S., and Johnson, L. W., 1987, "Optimization of Elastic Foundation for Minimum Beam Deflection," *International Journal of Solids and Structures*, Vol. 23, pp. 1551–1562.

Hetényi, M., 1946, *Beams on Elastic Foundation*, The University of Michigan Press, Ann Arbor, Mich.

Plaut, R. H., 1987, "Optimal Beam and Plate Foundations for Minimum Compliance," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, pp. 255–257.

Plaut, R. H., Johnson, L. W., and Olhoff, N., 1986, "Bimodal Optimization of Compressed Columns on Elastic Foundations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 130–134.

Shin, Y. S., Haftka, R. T., and Plaut, R. H., 1988, "Simultaneous Analysis and Design for Eigenvalue Maximization," *AIAA Journal*, Vol. 26, pp. 738–744.

Shin, Y. S., Haftka, R. T., Watson, L. T., and Plaut, R. H., 1988, "Tracing Structural Optima as a Function of Available Resources by a Homotopy Method," *Computer Methods in Applied Mechanics and Engineering*, Vol. 70, pp. 151–164.

Szelag, D., and Mróz, Z., 1979, "Optimal Design of Vibrating Beams with Unspecified Support Reactions," *Computer Methods in Applied Mechanics and Engineering*, Vol. 19, pp. 333–349.

Taylor, J. E., and Bendsøe, M. P., 1984, "An Interpretation for Min-max Structural Design Problems Including a Method for Relaxing Constraints," *International Journal of Solids and Structures*, Vol. 20, pp. 301–314.

Table 2 Optimal values for pinned-pinned beam

α_m	Point	b	d	a	K_T	α_{max}	Percent decrease
0.8	A	0.192	—	—	0	1.29	8.7
0.8	B	0.132	0.500	0.500	73	1.21	29.9
0.8	C	0.096	0.404	0.500	200	1.29	37.0
0.8	D	0.075	0.343	0.417	500	1.36	37.0
0.6	A	0.123	—	—	0	1.40	12.7
0.6	B	0.076	0.500	0.500	108	1.31	34.0
0.6	C	0.061	0.439	0.500	202	1.39	37.8
0.6	D	0.047	0.374	0.421	500	1.51	38.0
0.4	A	0.074	—	—	0	1.47	15.0
0.4	B	0.043	0.500	0.500	138	1.38	36.1
0.4	C	0.037	0.463	0.500	201	1.45	38.2
0.4	D	0.028	0.387	0.423	500	1.62	38.8
0.2	A	0.035	—	—	0	1.49	16.3
0.2	B	0.019	0.500	0.500	166	1.45	37.4
0.2	C	0.017	0.483	0.500	197	1.49	38.7
0.2	D	0.013	0.412	0.425	500	1.73	39.4

The Buckling of Thin-Walled Open-Profile Bars

Morris Ojalvo
Columbus, Ohio 43221

The theory of buckling for thin-walled open-profile bars is criticized. Its several derivations are faulted for violating statics, using a variational theorem approximately, using an incorrect variational statement, and/or using an inconsistent filament representation of the bar. Significantly, the theory yields buckling loads that contradict engineering expectations. A theory to replace it with general equations for computing buckling loads is presented. A problem solved under the old and new theories shows how torsional buckling is viewed under the new.

Introduction

H. Wagner (1929) laid down several concepts from which a general theory for thin-walled open-profile bars has grown. Some are known to be incorrect, others endure. A Wagner concept that endures, though it is not correct, is the notion that longitudinal stresses can induce torsional moment on normal sections of a deformed bar. The reasoning which attempts to justify what is called the Wagner effect and which is based on equilibrium considerations views the bar as if it were a bundle of longitudinal elements acting, to an extent, independently of each other. Such a conceptualization is inconsistent with the model used to define the bar's other characteristics. Concerning this, Lenz and Vielsack (Lenz, 1980) conclude that ". . . the assumptions and results of the theory of torsional buckling based on a filament model cannot be brought into conformity with the assumptions and results which a theory based on the principles of modern continuum theory suggests." While this conclusion is reached through an examination of a thin-walled tube, the reasoning applies as well to other sections. In general, equilibrium method derivations in which Wagner effect terms appear are deficient for not identifying the free body with which torsional equilibrium is expressed and/or violating statical principles (Ojalvo, 1987).

Variational method derivations in which Wagner effect terms appear fare no better. Some fail because the underlying theorem on which the derivation must be based (the theorem of stationary potential energy) is misused: Components of a finite strain tensor are used in the strain energy expression when infinitesimal strain expressions are clearly called for by the theorem (Ojalvo, 1982, 1987). Kappus (1937) and F. Bleich (1952) use the inconsistent multifilament model to determine the potential of external longitudinal tractions that act at the ends of a bar. For their analysis the ends of a filament move towards one another because of the sag resulting from its transverse displacements. The potential of the end tractions is the integral over an end face of the product of the relative movement for the ends of a filament and the end traction. At best the procedure is approximate. It yields different potentials for different distributions of the end tractions even when their statical resultants (be they centroidal forces, couples, or both) are the same (H. Bleich, 1956; Ojalvo, 1981). It will be shown that potentials do not depend on a particular distribution of end tractions when a single model for the bar is adhered to for all phases of the derivation.

Other derivations which have produced the conventional results for buckling equations (Goto et al., 1985; Hasegawa et al., 1985, Nishino et al., 1973) must be questioned for their use of Washizu's virtual work principle (theorem) and/or his principle (theorem) of stationary potential energy for finite displacements (Washizu, 1968, equations (3.49) and (3.68)). The proofs offered for these theorems depend on equilibrium equations (ibid. equations (3.22), (3.23)) that have no basis in a conventional understanding of statics: Quantities called pseudo stresses are determined for faces of elementary parallelepipeds of the material in the deformed state with a Lagrangian representation. That a deformed body's geometry can be expressed with coordinates of the undeformed state is not in dispute. What is objected to is a failure to use actual stresses acting on a defined parallelepiped for the equilibrium equations.

Cywinski and others (1986, 1982, 1971) have determined torsional buckling loads for tapered *I*-shaped columns with the established theory which accepts the Wagner hypothesis. The constant depth columns achieve taper with a gradual decrease of flange widths from a midlength maximum to a minimum at the ends. Torsional buckling loads were compared with those of columns that were similar except that their flange widths were uniform for the column's entire length. The natural expectation is that the uniform cross-section column yields the higher buckling load when cross-sections are the same at midlength. Surprisingly, the opposite was found. Lind (1973) performed additional computations to resolve the paradox but only succeeded in confirming that the conventional theory predicts substantially higher buckling loads for the tapered columns. Thus, once again the conventional theory is found wanting.

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

Discussion on this paper should be addressed to the 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 26, 1988; final revision, October 25, 1988.

A new theory with equations for buckling is presented next. It does not accept the validity of the Wagner effect and differs in one other important way from the conventional theory. Its derivation follows a discussion of the theorem of stationary potential energy.

The Theorem and Its Application to Buckling Theory

The following with caveats serves us as a statement of the theorem of stationary potential energy: The first variation of the potential is zero for a body in static equilibrium (Fung, 1965, pp. 284–288). The statement contemplates a domain in space occupied by a body. The variation of potential is in consequence of an arbitrary continuous displacement field that satisfies the body's physical constraints, both external and internal. It is required that suitable potential functions exist for external loads, body forces, and internal stresses. The last often limits the use of the theorem to bodies that have not previously suffered large displacements.

Buckling is said to occur when distortions of a new type appear in a bar that has already been deformed by loads or support movements. Thus a bar that is subject to only longitudinal contraction before buckling acquires flexural and, possibly, torsional distortion when buckling begins.

The analysis for buckling considers the bar in its buckled configuration and, because the buckled bar is in equilibrium, the theorem of stationary potential energy applies. Variations of the potential, however, can be the result of only variations of the buckling distortions because buckling occurs without a change in the pre-buckling distortions. The foregoing allows a bar's potential to be determined from the unstressed state as a datum or from the loaded but as yet unbuckled state. We follow Bleich (1952) and make the latter state the datum.

The Bar

The particular longitudinal line with which a bar is modeled is crucial to the analysis for buckling. Each point of the line represents a transverse cross-section (profile) of the bar. The average lateral movement of a profile of the undeformed bar is the lateral movement of the corresponding point of the line. Similarly, the profile's average longitudinal movement is the longitudinal movement of the corresponding point of the line. The extentional strains for the line are the average extentional strains for the corresponding profiles. Planes intersecting the line at right angles are called normal planes. The direction of the line is what the theory assumes for the longitudinal direction of the deflected bar at the corresponding profile.

The line assumes importance when the theorem of stationary potential energy is used for the derivation of buckling equations because, as will become obvious, different lines result in different expressions for the potential of external and body forces. For equilibrium derivations the line defines the orientation of normal planes. These are bounding surfaces for bar lengths that are considered as free bodies for the analysis. The torsion equilibrium condition is expressed for an axis that is normal to a normal plane and it is thus that the choice of line impacts on the derived buckling equations for equilibrium method derivations (Ojalvo, 1981).

We use the line of centroids to model a bar. Adopting the line of shear centers without, at the same time, adopting the dubious Wagner hypothesis leads to at least one anomaly: Namely, that a monosymmetric I-beam with uniform moment buckles at the same value of the moment irrespective of whether the larger or the smaller flange is in compression (Ojalvo, 1987). All derivations that reach the conventional equations for buckling use the sheer center line for the modeling of the bar and either explicitly assume the validity of the Wagner hypothesis or do something of a questionable nature to insure the appearance of Wagner effect terms in the final results.

Uniform Compression

Point c (Fig. 1) locates the centroid of a thin-walled bar's profile. Principal centroidal axes x and y serve as a global coordinate system to which transverse displacements and the location of points on a profile are referred. $S(x_o, y_o)$ locates the shear center, I_x and I_y are principal centroidal moments of inertia, and I_ω is the warping torsion constant of a profile. All points of a profile may move laterally when a bar deforms (Fig. 2). Because the displacements of a buckled bar are assumed small and profiles are assumed not to distort in their own planes, one may represent lateral movement for the points of a profile as though the profile is embedded in a rigid normal plane that is translated parallel to the x and y axes and then

Nomenclature

A = profile area
$\bar{a}$ = distance on y axis to point of transverse load application
c = point locating centroid of a profile
E = Youngs modulus
I_x, I_y = profile area moments of inertia about x and y axes
G = shear modulus of elasticity
I_ω = warping torsion constant for a profile
J = St. Venant torsion constant for a profile
L = length of a bar
P = compressive end load acting on a bar
p = transverse distributed load
$\bar{P}$ = total potential
r_o = distance between c and S on a profile
S = a point locating the shear center of a profile
U = strain energy
u, v = shear center displacements in the x and y directions
u_c, v_c = centroid displacements in the x and y directions
V, V_1, V_2 = external and body force load potentials
V_y = bar shear in the y direction
w = displacement in the z direction
w_c = average profile displacement in the z direction
x, y = centroidal principal axes of a profile
x_o, y_o = x and y coordinates of S
z = longitudinal axis and coordinate
α = a deflection angle, see Fig. 6(*b*)
θ = a small rotation of the profile about a longitudinal axis
ξ,η = principal centroidal axes of a profile in a normal plane of the deflected bar
$(')$ = designates differentiation with respect to z
a = a quantity equal to $2d - y_o$
I_o = polar moment of inertia of a profile about its shear center
P_θ = torsional buckling load
t = profile wall thickness
σ = a traction in the z direction acting on an end face of a bar
M_x = internal bar moment about the x axis

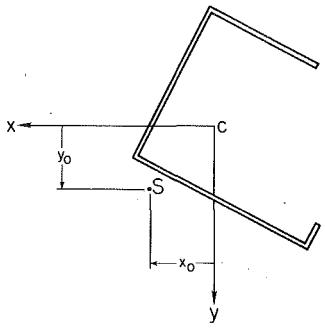


Fig. 1 Profile

rotated through a small angle about an axis that is perpendicular to it. Let $\theta = \theta(z)$ be the rotation where z is a coordinate of the line of centroids of the undeformed bar. The displacements of c parallel to the x and y axes are u_c and v_c. Those of S are u and v. The principal axes of the profile for a normal plane of the deformed bar are ξ and η (Fig. 2). Displacements u, v, u_c, and v_c are related according to

$$u_c = u + y_o\theta \tag{1a}$$

and

$$v_c = v - x_o\,\theta. \tag{1b}$$

Let the load consist of a distribution of tractions σ in the z direction acting on an end face of the bar at $z = 0$ and let the bar be restrained longitudinally at its opposite end $z = L$. The resultant of the load is a compressive force P acting through the centroid of the end face.

$$P = -\int_A \sigma\, dA \tag{2}$$

where the integral is over the profile area A.

A nonzero bimoment on an end face signifies a boundary condition in θ that is not homogeneous. The bimoment is $\int \sigma\,\omega\, dA$ where ω is the principal double-sectorial area coordinate of a location on the profile. Since boundary condition equations in θ must be homogeneous for a buckling problem, $\int \sigma\,\omega\, dA = 0$ is a restriction on permissible distributions of σ. Additional restrictions $\int \sigma\, x\, dA = \int \sigma\, y\, dA = 0$ apply because P intersects c both before and after buckling.

The load potential measured from the loaded but not yet buckled bar is

$$V = \int_A \sigma\, w\, dA \tag{3}$$

where w expresses longitudinal displacements of the end profile for a small buckling deformation of the bar and $\sigma = \sigma(x,y)$. Displacement w is (Vlasov, 1961)

$$w = w_c - u'x - v'y - \theta'\omega \tag{4}$$

where w_c is the average longitudinal displacement of a profile and a prime, ($'$), indicates a differentiation with respect to z. The potential of loads σ is seen to be

$$V = -P\, w_c \tag{5}$$

after substitution from equation (4) is made in equation (3) and cognizance is taken of equation (2) and the above three restrictions on the distribution of σ. w_c as used in equation (5) is also the distance by which end profiles approach each other along the z axis when a bar buckles. It is also the distance by which the ends of the centroid line approach each other. The relative movement is due entirely to sag displacements u_c and v_c. Extentional strain of the centroid line plays no part because that type of distortion is of a type that occurs in the bar prior to buckling.

It can be shown that w_c is given by

$$w_c = \frac{1}{2}\int_0^L [(u_c')^2 + (v_c')^2]dz \tag{6}$$

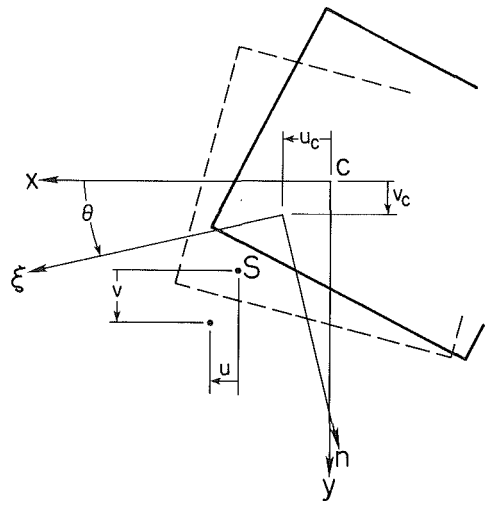


Fig. 2 Deflected profile

provided terms smaller by two or more degrees of magnitude than $(u_c')^2$ or $(v_c')^2$ are discarded. After u_c' and v_c' are expressed with u', v', and θ' in accordance with equations (1), one has for w_c

$$w_c = \frac{1}{2}\int_0^L [(u')^2 + (v')^2 + (r_o)^2(\theta')^2 + 2\theta'(y_o u' - x_o v')]\, dz \tag{7}$$

where $r_o^2 = x_o^2 + y_o^2$ is the square of the distance between the centroid and shear center of a profile. Using displacements of the shear center in place of displacements of the centroid in the expression for w_c is a matter of convenience. The use of u and v for dependent variables leads to simpler expression for the strain energy. u_c and v_c continue to define the transverse displacements of the bar.

The strain energy U for the distortions of the displacement field associated with buckling is (Bleich, 1952)

$$U = \frac{1}{2}\int_0^L [EI_y(u'')^2 + EI_x(v'')^2 + GJ(\theta')^2 + EI_\omega(\theta'')^2]\, dz \tag{8}$$

where E is Young's modulus, G is the modulus of elasticity in shear, and J is the St. Venant torsion constant for the profile.

The total potential $\bar{P}$ is the sum of U and V and obtained with equations (5), (7), and (8). The first variation of the functional $\bar{P}$, $\delta\bar{P}$, is equal to zero in accordance with the theorem of stationary potential energy. It leads to three Eulerian differential equations

$$EI_y u'''' + P\,u'' + P\,y_o\theta'' = 0 \tag{9a}$$

$$EI_x v'''' + Pv'' - Px_o\,\theta'' = 0 \tag{9b}$$

$$EI_\omega\theta'''' + [P(r_o)^2 - GJ]\theta'' + Py_o u'' - Px_o v'' = 0 \tag{9c}$$

with which the buckling of uniformly compressed bars is studied.

Equations (9) are not new, having been previously derived with an equilibrium method derivation (Ojalvo, 1981). The present demonstrates that equilibrium and variational method derivations yield identical results if the idealization for the bar is the same in both. It can not be otherwise for both rest on the same principles of statics, geometric continuity, and material behavior.

The old theory which uses the Wagner hypothesis and defines normal planes with the shear center line yields equations that differ in only one respect from equations 9: Where r_o^2 occurs in equations (9c) one finds in the old theory a term which is the square of the profile area's radius of gyration about the

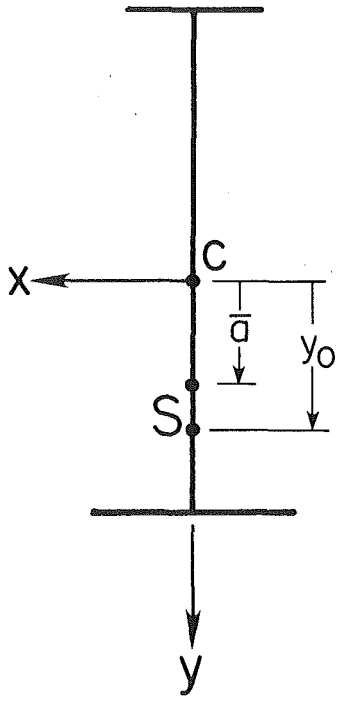


Fig. 3 Monosymmetric profile

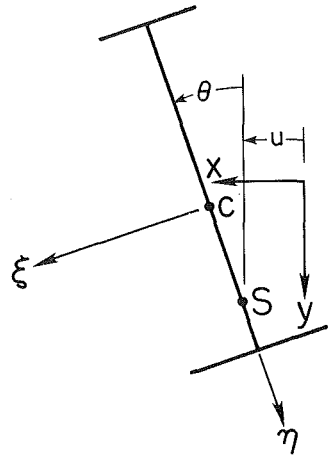


Fig. 4 Deflected monosymmetric profile

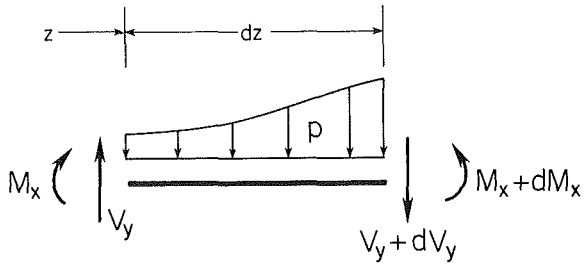


Fig. 5 Element *dz*

shear center (Timoshenko, 1961). Solutions based on the equations are profoundly influenced by the difference.

Bars With a Longitudinal Plane of Symmetry

Transverse loads are applied to a bar in its y–z plane of symmetry (Fig. 3). End loads which may also act on the bar have resultants which, at least until buckling occurs, are in the y–z plane. The bar's flexural stiffness about its x axis is large compared to its flexural stiffness about its y axis so that prebuckling curvature about the x axis may be ignored, and buckling without curvature about that axis may be presumed. As a consequence, the strain energy of the virtual buckling distortions is (Bleich, 1952)

$$U = \frac{1}{2}\int_0^L [EI_y(u'')^2 + GJ(\theta')^2 + EI_\omega(\theta'')^2]dz. \quad (10)$$

The potential of a centroidal longitudinal load is determined separately from the potential of end moments about the x axis and transverse loads:

(*a*) Potential of Centroidal Longitudinal Load. The shear center line for the buckled bar has a curvature component v'' in the y–z plane even though there is no flexural distortion about the ξ axis (the displaced x axis) of profiles. v'' is due entirely to twist distortion and flexure about the η axes. One may substitute u'' for curvature about the η axis because θ is small (Fig. 4) and then it is seen that v'' is very nearly $\theta \cdot u''$. Evidently, v'' is of a higher order (smaller by a degree of magnitude) than u'' and, because of it, v' is of a higher order than u'. Thus, the term including $(v')^2$ in equation (7) may be eliminated. Taking this and the fact that $x_o = 0$ into consideration causes the potential V_1 derived from equations (5) and (7) to be

$$V_1 = -\frac{1}{2}P\int_0^L [(u')^2 + (y_o)^2(\theta')^2 + 2y_o\theta' u']dz. \quad (11)$$

(*b*) End Moments and Transverse Loads Potential. Transverse loads p are at first assumed to act at the centroid line. A modification introduced later accounts for a positioning of p above or below the centroid. The virtual work theorem is used to facilitate the determination of the potential of these loads. See Ojalvo (1961) for a statement of the theorem in the form in which it is used. The loaded but as yet not buckled bar is the traction field contemplated by the theorem and the displacement field contemplated is that which characterizes buckling.

An element of the unbuckled bar is represented by a length dz of the centroid line (Fig. 5). It is in equilibrium under moments M_x, shears V_y, and loads $p = p(z)$. M_x and V_y are components of the internal stress resultant on normal planes as determined from a linear analysis which assumes an undeflected bar in its establishment of the equilibrium conditions.

The displacement field characterizing buckling produces rigid body motion for the element dz, curvature $u'' + y_o\theta''$ of the centroid line about the η axis, and twist θ' about a longitudinal axis through S. By the virtual work theorem, the work of the forces and moments of Fig. 5, for all elements and in consequence of the buckling displacements, is equal to the work of the end moments and transverse loads on the bar for the same displacements. V_2, the potential of these moments and loads, is the negative of this work. Work as used here is a scalar product which should not be confused with energy stored within an elastic body as the result of gradually applied loads.

The traction field of Fig. 5 does no work for the rigid body motion of dz because the forces and moments are in static equilibrium. Forces V_y do no work because transverse shear distortions do not occur in the displacement field. Only moments M_x do work, and this in consequence of the rotations of the end faces of element dz about axes parallel to the global x axis.

Curvature $u'' + y_o\theta''$ causes a relative rotation $(u'' + y_o\theta'')dz$ of the planes bounding element dz. The relative rotation about the η axis (Fig. 4) has a component $-\theta(u'' + y_o\theta'')dz$ about the x axis. Twist θ' about a longitudinal axis through S produces additional curvature of the line of centroids in the y–z plane (Figs. 6(*a*), (*b*)). The increment of longitudinal rotation for the bounding planes of the element produces a rel-

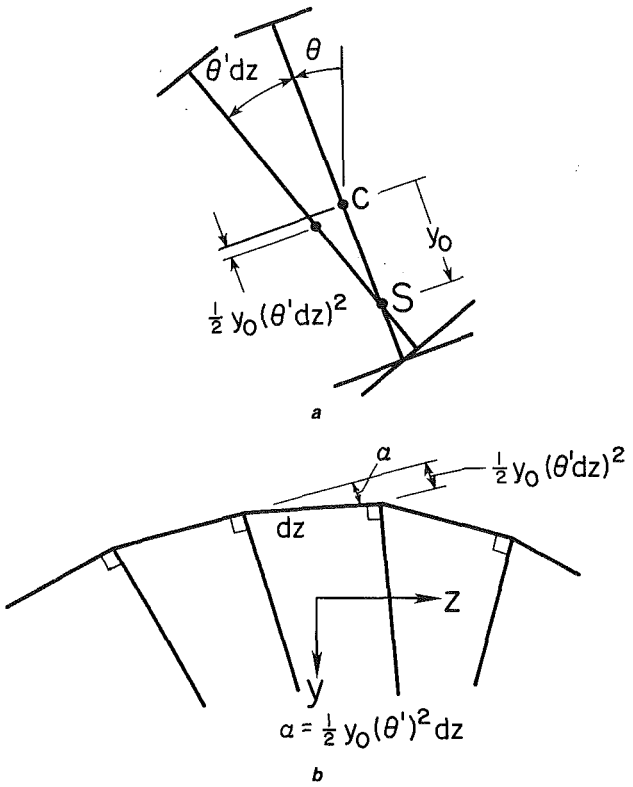


Fig. 6 Curvature induced by twist distortion

ative displacement of the centroids in the η direction of $\frac{1}{2}y_o(\theta' dz)^2$. The line of centroids is depicted (Fig. 6(*b*)) as a string of elements dz with slope discontinuities α where they join. The representation of the centroid line in this manner facilitates understanding how twist affects curvature in the y–z plane without introducing further approximation to the the derivation. The normal to a segment is represented by a perpendicular in the y–z plane at its right end. Relative displacement $\frac{1}{2}y_o(\theta' dz)^2$ between the ends is shown as a distance parallel to the normal for the preceding segment. Deflection angles α are $\frac{1}{2}y_o(\theta')^2 dz$ so that curvature of the centroid line about the x axis resulting from twist is $-\frac{1}{2}y_o(\theta')^2$.

The curvature of the centroid line about the x axis from both distortions is $-[\theta(u'' + y_o\theta'') + \frac{1}{2}y_o(\theta')^2]$. By the virtual work theorem, the work of the loads is

$$-\int_0^L M_x[\theta(u'' + y_o(\theta'') + \tfrac{1}{2}y_o(\theta')^2]\, dz.$$

Before writing the expression for V_2 we note that the potential is increased by $\int_0^L \frac{1}{2}p\,\bar{a}\,\theta^2\, dz$ when p is distributed on a line with y coordinate $\bar{a}$ (Fig. 3). With this,

$$V_2 = \int_0^L \{M_x[\theta(u'' + y_o\,\theta'') + \tfrac{1}{2}\,y_o\,(\theta')^2] + \tfrac{1}{2}\,p\bar{a}\theta^2\}\, dz. \quad (12)$$

The total potential found by adding U (equation (10)), V_1, and V_2 is

$$\bar{P} = \int_0^L \{\tfrac{1}{2}EI_y(u'')^2 + \tfrac{1}{2}EI_\omega(\theta'')^2 - \tfrac{1}{2}P(u')^2 + \tfrac{1}{2}[GJ - P(y_o)^2 + M_x\,y_o](\theta')^2 + M_x\,\theta\,u'' - Py_o\,\theta'u' + M_xy_o\theta\,\theta'' + \tfrac{1}{2}p\bar{a}\,\theta^2\}\, dz. \quad (13)$$

The Eulerian equations used for the determination of buckling loads and mode shapes for bars with a plane of symmetry are obtained from the condition $\delta\bar{P} = 0$. These are, after V_y is substituted for M_x' and $-p$ for M_x'',

$$EI_yu'''' + Pu'' + (M_x + Py_o)\theta'' + 2V_y\theta' - p\theta = 0 \quad (14a)$$

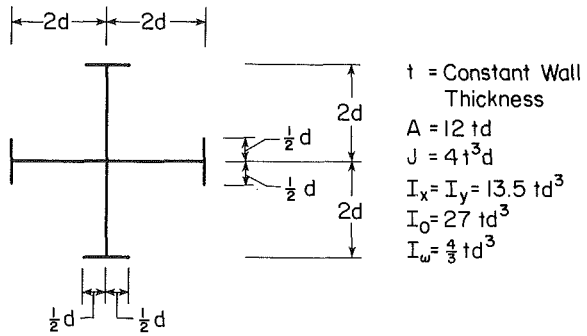


Fig. 7 Cruciform

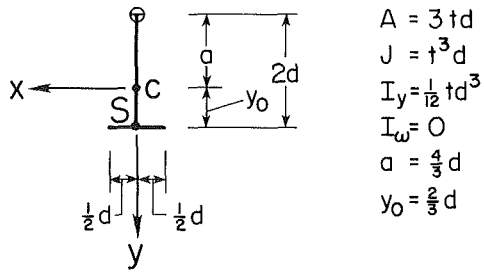


Fig. 8 *T* section

and

$$EI_\omega\,\theta'''' + [P(y_o)^2 - GJ + M_x\,y_o]\theta'' + V_y\,y_o\,\theta' - p(y_o - \bar{a})\theta + [P\,y_o + M_x]\,u'' = 0. \quad (14b)$$

Equation (14*b*) differs considerably from what would be obtained with the old theory.

Example

The torsional buckling load of a column of length L whose profile is a flanged cruciform is required. The profile has a constant thickness t which is sufficiently small compared to its other dimensions so that quantities multiplied by t raised to the power three may be ignored in the computation of section properties I_x, I_y, and I_ω (Fig. 7).

The torsional buckling equation under the theory using the Wagner hypothesis is (Bleich, 1952)

$$EI_\omega\,\theta'''' + \left[P\frac{I_o}{A} - GJ\right]\theta'' = 0 \quad (15)$$

where I_o is the shear center polar moment of inertia of the profile area. For boundary conditions $\theta = \theta'' = 0$ at z equal to zero and L the buckled form is defined by a half sine wave variation of θ and the buckling load P_θ is

$$P_\theta = \frac{A}{I_o}\left(GJ + EI_\omega\frac{\pi^2}{L^2}\right) = \frac{16}{9}G\frac{t^3}{d} + \frac{16}{27}E\frac{\pi^2}{L^2}td^3. \quad (16)$$

Buckling of a purely torsional nature is never indicated by the new theory. If, however, the cruciform column is perceived as four T columns and the Ts were to buckle simultaneously in a lateral-torsional mode with a common longitudinal axis of enforced rotation, the effect would be the same as a pure torsional buckling of the cruciform. The new theory permits a determination of the buckling load for one such T (Fig. 8). The enforced axis of rotation intersects its profile at the toe of the stem.

Equations (14) (and equations (9), for the matter) may not be used when there is an enforced axis of rotation. Such axes impose conditions which must be reflected in the total potential before the Eulerian equations are determined. Their effect is to reduce the number of dependent variables. For the T bar

considered (Fig. 8), the imposed conditions $u = -2\,d\,\theta$, $u' = -2\,d\,\theta'$, and $u'' = -2\,d\,\theta''$ are used to eliminate u from equation (13). For uniform compression p and M_x are absent so that $\bar{P}$ for the bar considered reduces to

$$\bar{P} = \frac{1}{2}\int_0^L [(EI_\omega + EI_y\,\{2d\}^2)\,(\theta'')^2 + (GJ - P\,a^2)\,(\theta')^2]\,dz \qquad (17)$$

where $a = 2d - y_o$. The buckling equation for the T bar obtained from $\delta\bar{P} = 0$ is

$$(EI_\omega + EI_y\,\{2d\}^2)\,\theta'''' + (P\,a^2 - GJ)\,\theta'' = 0. \qquad (18)$$

The buckling load obtained with equation (18) is $\frac{1}{4}P_\theta$ where P_θ is the buckling load for the cruciform. With boundary conditions $\theta = \theta'' = 0$ at z equal to zero and L and with the section properties as tabulated in Fig. 8, four times the buckling load for the T is found to be

$$P_\theta = \frac{4}{a^2}\left[GJ + (EI_\omega + EI_y\,\{2d\}^2)\,\frac{\pi^2}{L^2}\right] = \frac{9}{4}\,G\,\frac{t^3}{d} + \frac{3}{4}\,\frac{\pi^2}{L^2}\,E\,t\,d^3. \qquad (19)$$

Buckling loads for the example under the old and the proposed theories are in the ratio of 64 to 81. Such discrepancy indicates a need for experimental verification. Before proceeding too far in this direction, however, it is essential that thought be given to the nature of buckling and failure.

A buckling load comes from the solution of a mathematical eigenvalue problem. The equations of the problem are obtained with an idealization of the actual bar. Thus, the idealization assumes infinitesimal displacements, a profile which retains its shape, the absence of shear strains in the bar's middle surface, and that St. Venant torsion theory applies even when the torsion is not uniform. It is therefore not remarkable that tests do not yield what can be clearly identified as buckling loads. At best they indicate a load range where small load increments begin to produce large displacement increments.

Failure load, on the other hand, is clearly identified in a test. It is a load which is of paramount interest to designers of machines and structures.

The study of buckling is justified by the hope that correlation exists between failure load and buckling load. Such correlation is generally acknowledged for columns that fail by excessive bending. Correlation is less certain for columns which fail with substantial torsional deformation and for beam columns and beams loaded in a plane of symmetry. This is largely due to the lack of a great deal of experimental information in this area. Nevertheless, experimental failure loads for columns and one beam column have been assembled and comparisons made with buckling loads (Ojalvo, 1983). These indicate that a better correlation is obtained with the proposed theory of buckling. Under the old theory some loads are greater while others are smaller than the failure load. Buckling loads computed with the proposed theory were always higher than the corresponding experimental failure loads. Discrepancies with the proposed theory averaged 6 percent and were as much as 14 percent. With the old theory failure loads were, on the average, 7 percent higher than the buckling load, and in one instance the failure load was 20 percent higher.

Conclusions

A bar buckling theory should, whenever possible, show consistency in its idealization of the bar. The multifilament model used for the Wagner effect is inconsistent with the single filament model used elsewhere by the theory.

It seems evident that buckling must be defined before it may be examined mathematically. Our definition is couched in terms of the onset of new types of distortions.

The virtual work theorem can be useful in the sometimes difficult task of evaluating the potential of the loads for a variational derivation of the buckling equations.

References

Bleich, F., 1952, *Buckling Strength of Metal Structures*, McGraw-Hill, New York.

Bleich, H., 1956, "Refinement of the Theory of Torsional Buckling of Thin-Walled Columns," *Proceedings, First Midwestern Congress, Solid Mechanics*, Urbana, Ill.

Cywinski, Z., and Kollbrunner, C. F., 1971, "Drilknicken dunnwandiger I-Stabe mit verandelichen, doppelt-symmetrischen Querschnitten," *Institute for Engineering Research*, Zurich, Vol. 18, pp. 1–35.

Cywinski, Z., and Kollbrunner, C. F., 1982, "Neues zu einem Paradoxon des Drill-knickens," *Institute for Engineering Research*, Zurich, Vol. 50, pp. 1–32.

Cwyinski, Z., 1986, "On a Certain Paradox of Torsional Buckling: State of the Art," *Colloquium on Stability of Steel Structures, Proceedings*, Hungary.

Fung, Y. C., 1965, *Foundations of Solid Mechanics*, Prentice-Hall, Englewood Cliffs, N.J., pp. 284–288.

Goto, Y., Matsuura, S., Hasegawa, A., and Nishino, F., 1985, "A New Formulation of Finite Displacement Theory of Curved and Twisted Bars," *Proceedings of JSCE, Structural Engineering/Earthquake Engineering*, Vol. 2, No. 2.

Hasegawa, A., Liyanage, K., Ikeda, T., and Nishino, F., 1985, "A Concise and Explicit Formulation of Out-of-plane Instability of Thin-walled Members," *Proceedings of JSCE, Structural Engineering/Earthquake Engineering*, Vol. 2, No. 1.

Kappus, R., 1937, "Drillknicken zentrisch gedruckter Stabe mit offenen Profil im elastischen Bereich," *Luftfahrt-Forschung*, also, NACA Tech. Mem. 851.

Lenz, J., and Vielsack, P., 1980, "Eine kritsche Bemerkung zur Theorie des Drill-knickens," *Der Stahlbau*, Vol. 49, No. 8, p. 245.

Lind, N. C., 1973, Review of Cywinski et al. paper, *Applied Mechanics Reviews*, Vol. 26, p. 574.

Nishino, F., Kasemset, C., and Lee, S. L., 1973, "Variational Formulation of Stability Problems for Thin-walled Members," *Ingenieur-Archiv*, Vol. 43, pp. 58–68.

Ojalvo, M., 1961, Discussion of "Principle of Virtual Work in Structural Analysis," *Journal of the Structural Division, ASCE*, Vol. 87, No. 2, pp. 61–64.

Ojalvo, M., 1981, "(The)Wagner Hypothesis in Beam and Column Theory," *Journal of the Engineering Mechanics Division, ASCE*, Vol. 107, No. 4, pp. 669–677.

Ojalvo, M., 1982, Discussion of "Lateral-Torsional Buckling of Tapered I-Beams," *Journal of the Structural Division, ASCE*, Vol. 108, No. 2, pp. 503–504.

Ojalvo, M., 1983, Closure to "(The)Wagner Hypothesis in Beam and Column Theory," *Journal of the Engineering Mechanics Division, ASCE*, Vol. 109, No. 3, pp. 924–932.

Ojalvo, M., 1987, "Discussion of "Buckling of Monosymmetric I-Beams Under Moment Gradient," *Journal of Structural Engineering, ASCE*, Vol. 113, No. 6, pp. 1387–1391.

Timoshenko, S. P., with Gere, J. M., 1961, *Theory of Elastic Stability*, 2nd. ed., McGraw-Hill, New York.

Vlasov, V. Z., 1961, *Thin-Walled Elastic Beams*, 2nd. ed., Israel Program for Scientific Translations, PST Cat. 428.

Wagner, H., 1929, "Verdrehung und Knickung von offenen Profilen," *25th Anniversary Publication, Technische Hochschule, Danzig*, pp. 329–343, NACA translation, TM 807.

Washizu, K., 1968, *Variational Methods in Elasticity and Plasticity*, Pergamon Press, Oxford.

Cantilever Rod in Cross Wind

C. Y. Wang
Professor,
Michigan State University,
East Lansing, Mich. 48824
Mem. ASME

A thin elastic rod is held at one end in a strong cross wind. The nonlinear large deformation equations are formulated and solved by perturbation and numerical integration. The problem is governed by a nondimensional parameter K *representing the relative importance of aerodynamic drag to flexural rigidity. For large* K, *phenomena such as nonuniqueness, instability, and hysteresis may occur.*

Introduction

The study of the behavior of thin elastic rods in cross wind is important in the design of antennas on roof tops and moving vehicles. Due to the interaction of fluid mechanics and elasticity, literature on the flexible rod has been scarce. The drag of inclined rigid circular cylinders was experimentally documented by several researchers (Hoerner, 1958). The drag of a long curved cylinder may be estimated by integrating the drag coefficients of element inclined straight cylinders. The method has been applied successfully on ocean cables (McCormick, 1973), where weight is important but the flexural rigidity can be ignored. For antennas in a strong cross wind studied in this paper, the problem is substantially different since flexural rigidity and aerodynamic drag dominate while the effect of self-weight can be ignored.

Formulation

Consider a thin cantilever in cross wind shown in Fig. 1(*a*). The cantilever is relatively inextensible, has uniform properties, and one end is fixed at an angle α with the uniform flow. Let s' be the arc length from origin, L be the rod length and θ be the local angle of inclination. For an elemental length ds' (Fig. 1(*b*)), the free stream velocity gives rise to a normal drag per length q_n and a tangential drag per length q_t. The horizontal force X' acting on the elemental length is then

$$X' = \int_{s'}^{L} (-q_n \sin\theta - q_t \cos\theta) ds'. \tag{1}$$

Similarly, the vertical force is

$$Y' = \int_{s'}^{L} (q_n \cos\theta - q_t \sin\theta) ds'. \tag{2}$$

A balance of local moment m on ds' gives

$$dm = X' \sin\theta ds' - Y' \cos\theta ds'. \tag{3}$$

If the rod is thin enough, the local moment is proportional to the local curvature (elastica model, see Frisch-Fay, 1962)

$$m = EI \frac{d\theta}{ds'} \tag{4}$$

where EI is the flexural rigidity.

Now, q_n and q_t are complicated functions of cylinder cross-section, free stream velocity V, angle θ, and Reynolds number. Experiments on rigid inclined circular cylinders (Hoerner, 1958) support the cross-flow principle, that for high, subcritical Reynolds numbers the net steady resistance is the sum of the resistance due to the normal component of the free stream, and that in turn is due to the tangential component. Thus

$$q_n = c_n \frac{1}{2} \rho U_n |U_n| D \tag{5}$$

$$q_t = c_t \frac{1}{2} \rho U_t |U_t| D \tag{6}$$

where ρ is the fluid density, D is the diameter, C_n and C_t are drag coefficients in directions normal and tangential to the cylinder, and $U_n = U \sin\theta$, $U_t = U \cos\theta$. Experiments also show C_n and C_t are approximately constant at high subcritical Reynolds numbers (McCormick, 1973; Schlichting, 1979).

Normalize the arc length by L, the forces by EI/L^2, and drop the primes. Equations (1)–(4) yield

$$X = -K \int_s^1 (\sin^2\theta |\sin\theta| + \lambda \cos^2\theta |\cos\theta|) ds \tag{7}$$

$$Y = K \int_s^1 \sin\theta \cos\theta (|\sin\theta| - \lambda |\cos\theta|) ds \tag{8}$$

$$\frac{d^2\theta}{ds^2} = X \sin\theta - Y \cos\theta. \tag{9}$$

Here $\lambda \equiv C_t/C_n$ is usually small and $K = C_n \rho D L^3 U^2/2EI$ is an

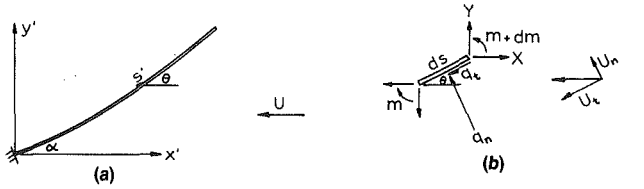


Fig. 1 *(a)* **The coordinate system and** *(b)* **forces on an elemental length**

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10-15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y., 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 23, 1988; final revision, January 8, 1989.
Paper No. 89-WA/APM-8.

important nondimensional parameter signifying the relative importance of drag due to velocity and resistance due to rigidity. Equations (7) and (8) in differential form are

$$\frac{dX}{ds} = K(\sin^2\theta\,|\sin\theta| + \lambda\cos^2\theta\,|\cos\theta|) \tag{10}$$

$$\frac{dY}{ds} = K\sin\theta\cos\theta(\lambda\,|\cos\theta| - |\sin\theta|). \tag{11}$$

The boundary conditions are

$$X(1) = Y(1) = \frac{d\theta}{ds}(1) = 0, \quad \theta(0) = \alpha. \tag{12}$$

After $\theta(s)$ is found, the configuration of the rod (x, y) is obtained by integrating

$$\frac{dx}{ds} = \cos\theta, \quad \frac{dy}{ds} = \sin\theta, \quad x(0) = y(0) = 0. \tag{13}$$

Perturbation Solution for Small K

Small K implies relatively high flexural rigidity. We expect the rod to be almost straight. Let $K = \epsilon \ll 1$ and we expand

$$\theta = \alpha + \epsilon\theta_1 + \epsilon^2\theta_2 + 0(\epsilon^3) \tag{14}$$

$$X = \epsilon X_1 + \epsilon^2 X_2 + 0(\epsilon^3) \tag{15}$$

$$Y = \epsilon Y_1 + \epsilon^2 Y_2 + 0(\epsilon^3). \tag{16}$$

The first-order terms of equations (9)–(12) are

$$\frac{d^2\theta_1}{ds^2} = X_1\sin\alpha - Y_1\cos\alpha \tag{17}$$

$$\frac{dX_1}{ds} = \sin^2\alpha\,|\sin\alpha| + \lambda\cos^2\alpha\,|\cos\alpha| \equiv k_1 \tag{18}$$

$$\frac{dY_1}{ds} = \sin\alpha\cos\alpha(\lambda\,|\cos\alpha| - |\sin\alpha|) \equiv k_2. \tag{19}$$

The solutions subjected to

$$X_1(1) = Y_1(1) = \frac{d\theta_1}{ds}(1) = \theta_1(0) = 0 \tag{20}$$

are

$$X_1 = k_1(s-1), \quad Y_1 = k_2(s-1) \tag{21}$$

$$\theta_1 = \left(\frac{s^3}{6} - \frac{s^3}{2} + \frac{s}{2}\right)\sin\alpha\,|\sin\alpha|. \tag{22}$$

The proper expansion for the absolute values of $\sin\theta$ and $\cos\theta$ is

$$|\sin\theta| = |\sin\alpha| + \epsilon\theta_1\cos\alpha\,\mathrm{sgn}(\sin\alpha) + \ldots \tag{23}$$

$$|\cos\theta| = |\cos\alpha| - \epsilon\theta_1\sin\alpha\,\mathrm{sgn}(\cos\alpha) + \ldots \tag{24}$$

where sgn denotes the sign of the argument. The second-order equations are

$$\frac{d^2\theta_2}{ds^2} = X_2\sin\alpha - Y_2\cos\alpha + \theta_1(X_1\cos\alpha + Y_1\sin\alpha) \tag{25}$$

$$\frac{dX_2}{ds} = \theta_1 k_3, \quad \frac{dY_2}{ds} = \theta_1 k_4 \tag{26}$$

where

$$k_3 \equiv \sin\alpha\cos\alpha[\sin\alpha\,\mathrm{sgn}(\sin\alpha) - \lambda\cos\alpha\,\mathrm{sgn}(\cos\alpha) + 2|\sin\alpha| - 2\lambda|\cos\alpha|] \tag{27}$$

$$k_4 \equiv \cos 2\alpha[\lambda|\cos\alpha| - |\sin\alpha|] - \sin\alpha\cos\alpha[\lambda\sin\alpha\,\mathrm{sgn}(\cos\alpha) + \cos\alpha\,\mathrm{sgn}(\sin\alpha)]. \tag{28}$$

The solution is

$$X_2 = \left(\frac{s^4}{24} - \frac{s^3}{6} + \frac{s^2}{4} - \frac{1}{8}\right)\sin\alpha\,|\sin\alpha|k_3 \tag{29}$$

$$Y_2 = \left(\frac{s^4}{24} - \frac{s^3}{6} + \frac{s^2}{4} - \frac{1}{8}\right)\sin\alpha\,|\sin\alpha|k_4 \tag{30}$$

$$\theta_2 = \left(\frac{s^6}{720} - \frac{s^5}{120} + \frac{s^4}{48} - \frac{s^2}{16} + \frac{3}{40}s\right)\sin\alpha \times |\sin\alpha|\cos\alpha[\sin\alpha\,\mathrm{sgn}(\sin\alpha) + |\sin\alpha| - \lambda|\cos\alpha|] + \left(\frac{s^6}{180} - \frac{s^5}{30} + \frac{s^4}{12} - \frac{s^3}{12} + \frac{s}{20}\right) \times\lambda\sin\alpha\cos\alpha|\sin\alpha\cos\alpha|. \tag{31}$$

The force experienced at the origin is

$$-X = \epsilon k_1 + \epsilon^2\sin\alpha|\sin\alpha|\frac{k_3}{8} + 0(\epsilon^3) \tag{32}$$

$$-Y = \epsilon k_2 + \epsilon^2\sin\alpha|\sin\alpha|\frac{k_4}{8} + 0(\epsilon^3). \tag{33}$$

The moment at the origin, normalized by EI/L, is

$$M = \frac{d\theta}{ds}(0) = \sin\alpha|\sin\alpha|\left\{\frac{\epsilon}{2} + \epsilon^2\frac{\cos\alpha}{40}[3\sin\alpha\,\mathrm{sgn}(\sin\alpha) + 3|\sin\alpha| - \lambda|\cos\alpha|] + 0(\epsilon^3)\right\}. \tag{34}$$

The tip angle at $s = 1$ is

$$\theta(1) = \alpha + \sin\alpha|\sin\alpha|\left\{\frac{\epsilon}{6} + \epsilon^3\cos\alpha\frac{19}{720}[\sin\alpha\,\mathrm{sgn}(\sin\alpha) + |\sin\alpha| - \frac{3}{19}\lambda|\cos\alpha|] + 0(\epsilon^3)\right\}. \tag{35}$$

The configuration of the rod is described by the Cartesian coordinates

$$x = \int_0^s\cos\theta\,ds = s\cos\alpha - \epsilon\sin^2\alpha|\sin\alpha|\left(\frac{s^4}{24} - \frac{s^3}{6} + \frac{s^2}{4}\right)0(\epsilon^2) \tag{36}$$

$$y = \int_0^s\sin\theta\,ds = s\sin\alpha + \epsilon\cos\alpha\sin\alpha|\sin\alpha|\left(\frac{s^4}{24} - \frac{s^2}{6} + \frac{s^2}{4}\right) + 0(\epsilon^2). \tag{37}$$

Numerical Integration of the $\alpha = 0$ Case

For large K the nonlinear equations (9) – (13) can only be integrated numerically. Owing to the boundary conditions, it is more convenient to change the independent variable to $r \equiv 1 - s$. Thus, the equations are

$$\frac{d^2\theta}{dr^2} = X\sin\theta - Y\cos\theta \tag{38}$$

$$\frac{dX}{dr} = -K(\sin^2\theta\,|\sin\theta| + \lambda\cos^2\theta\,|\cos\theta|) \tag{39}$$

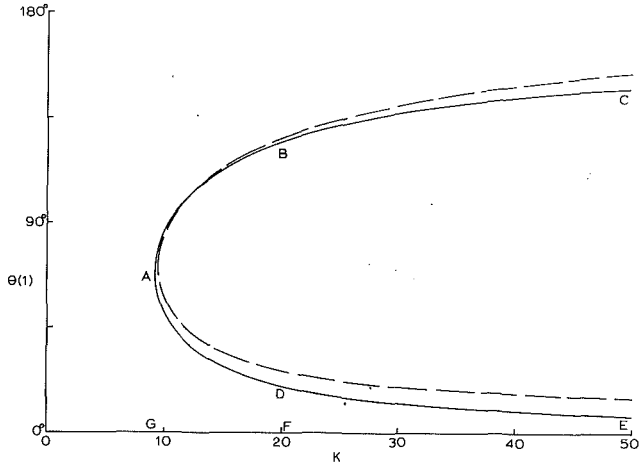


Fig. 2 End angle as a function of K for $\alpha = 0$; ——— $\lambda = 0$; ---- $\lambda = 0.1$

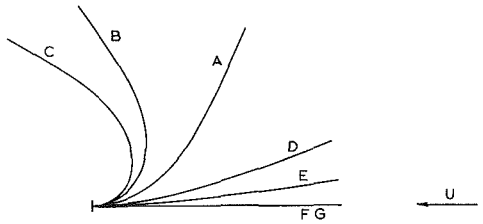


Fig. 3 Configurations for $\alpha = 0$, $\lambda = 0.014$. A: $K = 9.2$; B: $K = 20$, C: $K = 50$, D: $K = 20$, E: $K = 50$. The states correspond to those shown in Fig. 2. D and E are unstable.

$$\frac{dY}{dr} = -K \sin\theta \cos\theta(\lambda|\cos\theta| - |\sin\theta) \tag{40}$$

$$\text{at } r = 0 \quad X = Y = \frac{d\theta}{dr} = 0. \tag{41}$$

For given K, λ we guess θ at $r = 0$ and integrate equations (38)–(41) as an initial value problem by the Runge-Kutta-Fehlberg algorithm. The integration terminates at $r = 1$ where we check whether $\theta = \alpha$. If not, the initial guess is adjusted.

Figure 2 shows the end angle $\theta(1)$ plotted against K for $\alpha = 0$ or when the cantilever rod is pointing towards the wind. The solutions are the trivial (straight) solution $\theta(1) = 0$ and the curve EDABC. The value of λ for circular rods is less than 0.02 and the solution curve is indistinguishable from the $\lambda = 0$ curve. For $0 < K < 9.2$ there is only one solution, the trivial one. When K is large, say $K = 20$, three solutions at states F,D,B are possible. Note that the nontrivial solutions do not bifurcate from the trivial solution. The analytic proof of nonbifurcation is given in the Appendix. The graph of the moment at the base, $\theta'(0)$, show similar characteristics. Thus, small deformation theory and stability theory would not be able to predict such nonuniqueness. The phenomena is very different from a tip-loaded Euler column which has a pitchfork bifurcation.

Now let us look at the stability of the nonunique solutions at large K. Since the trivial solution does not bifurcate, it is stable to infinitesimal perturbations. The branch ADE has negative slope, i.e., the higher the velocity the lower the deformation and strain energy. We conclude this branch is unstable and can not be realized in practice. The branch ABC has positive slope and is thus stable.

What would happen if the speed or K is gradually increased for a straight rod at $\alpha = 0$? It will remain straight unless a large enough finite perturbation causes it to jump to a nontrivial branch. For example, suppose at $K = 20$ a perturbation causes the rod to jump from State F to State D. Since State D is unstable, the rod instantly snaps through State A and settles at State B. The branch BC will be followed as K is further increased. As K is decreased from Stable B, the rod follows to the upper branch to State A $K = 9.2$. Then there is a sudden snap back to straight configuration at State G and remain straight when K falls below 9.2. The hystersis loop would happen only if there is an energy input from State F to State D. At higher K, the required perturbation energy becomes smaller. Figure 3 shows the shape of the rod at various states.

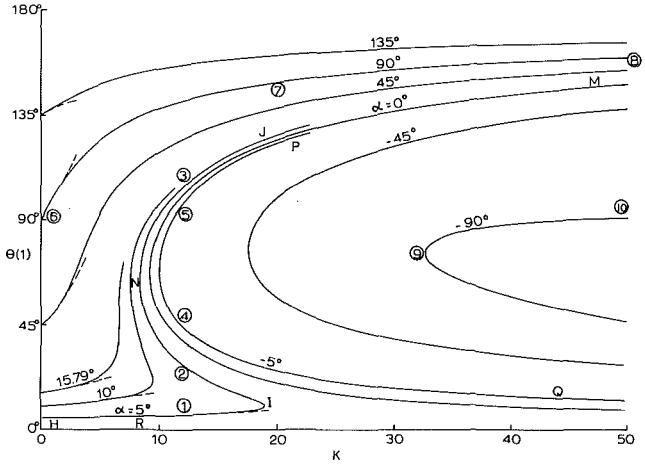


Fig. 4 End angle as a function of K; $\lambda = 0.014$ and various α. Dashed lines are from equation (35).

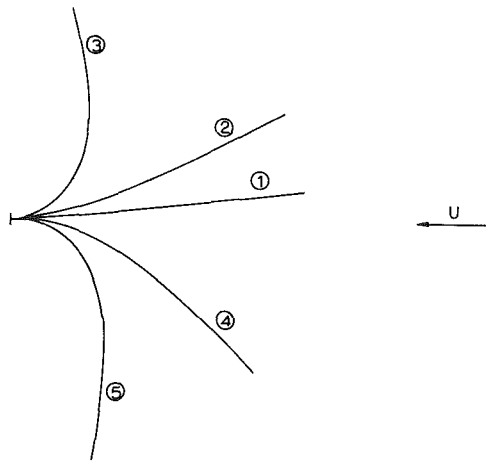


Fig. 5 Nonunique configurations at $K = 12$ for $\alpha = 5$ deg. ①, ③, ⑤ are stables ②, ④ are unstable. The states corresponds to those indicated in Fig. 4.

The Inclined Cantilever

When $\alpha \neq 0$ or 180 deg, the undeformed rod is pointing at an angle to the free stream. For nonzero K, the rod will always be bent. Figure 4 shows the end angle $\theta(1)$ for $\lambda = 0.014$ and various α. Changing the sign of α is equivalent to changing the sign of $\theta(1)$. Thus, negative α represents the case when the rod is bent with opposite curvature, i.e., to the other side. Let us study the case when $\alpha = \pm 5$ deg. The solution is unique when $0 < K < 8.8$, three solutions (one unstable) when $8.8 < K < 10$, five solutions (two unstable) for $10 < K < 18.4$, and three solutions (one unstable) for $k > 18.4$. The behavior of the slightly inclined rod ($\alpha = 5$ deg) differs from the parallel rod with $\alpha = 0$. As K is increased from zero, the rod starts to bend following the curve HI. Then, even without an outside perturbation, it snaps to State J and follows JM as the cross flow is further increased. When K decreases the path MJN snap RH is followed. Thus the rod has an natural hystersis loop. The states on the curve PQ (bent to the other side) does not happen

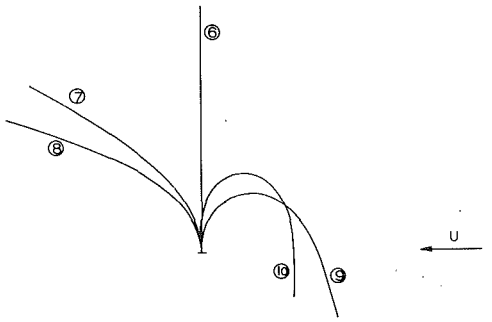


Fig. 6 Configurations for $\alpha = 90$ deg, $\lambda = 0.014$; ⑥: $K = 0$, ⑦: $K = 20$, ⑧: $K = 50$, ⑨ $K = 33.6$, ⑩ $K = 50$

unless there is a large enough appropriate perturbation. The five equilibrium configurations for the same $K = 12$ is shown in Fig. 5. The natural hysteresis loop and jumps occur only for values of α between 0 deg and 15.79 deg. For larger angles the rod would bend smoothly. The configurations for a cantilever rod held perpendicular to the free stream are shown in Fig. 6. In general, the higher the angle α, the higher the minimum value of K for the rod to be able to bend to the other side. Our approximate solutions compare well with exact numerical results for low K.

Our study, however, considers only in-phase deformations, out-of-plane perturbations, such as those due to vortex shedding, would certainly cause the α negative states (or those bent to the other side) to be swept by the cross wind into the stable α positive states. Figure 7 shows the maximum normalized moment which occurs at the base of the cantilever rod. For clarity, only the most stable states are shown. Figure 8 shows the horizontal and vertical forces experienced at the base. Although the drag $(-X)$ is always positive, the transverse force $(-Y)$ may be negative at low K and low α.

Discussions and Conclusion

For rough circular cylinders the value of $\lambda \equiv C_t/C_n$ is less than 0.02 (Hoerner 1958, McCormick 1973). This value is even smaller for smooth cylinders. Our computations show (e.g., Fig. 2) that λ can be set to zero without much error for circular cylinders. However, λ may be nonnegligible for laterally-ridged or finned cylinders where the transverse drag may increase several fold.

Although the present analysis also applies to low K values, we note the Reynolds number should be kept high and subcritical (500 ~ 500,000) such that C_n, C_t are approximately constant. The importance of the nondimensional parameter K can not be over emphasized. For small K ($K < 6.513$), the rod is relatively stiff and the solution is unique. For large K the possible nonlinear phenomena of nonuniqueness, instability, and hysteresis jumps must be considered.

For high, subcritical Reynolds numbers, oscillations due to alternate vortex shedding are always present (Blevins, 1977). Do these oscillations affect our results which are derived from steady, mean deformation equations? The answer is unlikely. Experimental investigations (Keefe, 1961, Schmidt, 1965) on the *magnitude* of the oscillatory forces on a rigid cylinder normal to free stream showed that order of the unsteady lift is of the same order of the steady drag and the order of the unsteady drag is much less than the order of the steady drag. Now, lift is out of the plane of elastic deformation and therefore would not affect our basic equations. Furthermore, this unsteady lift rapidly decrease in magnitude as the local angle θ deviates from 90 deg. On the other hand, the unsteady drag is in the plane of the elastic deformation. However, its magnitude is orders smaller, and therefore would not appreciably affect the deformation. Assuming no resonance

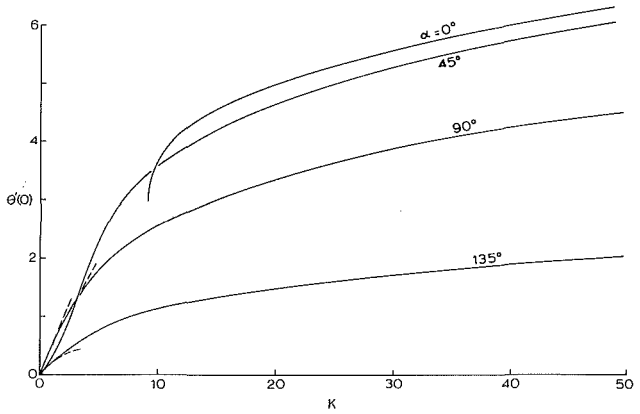


Fig. 7 Maximum moment (at the base) for the most stable states; dashed lines are from equation (34)

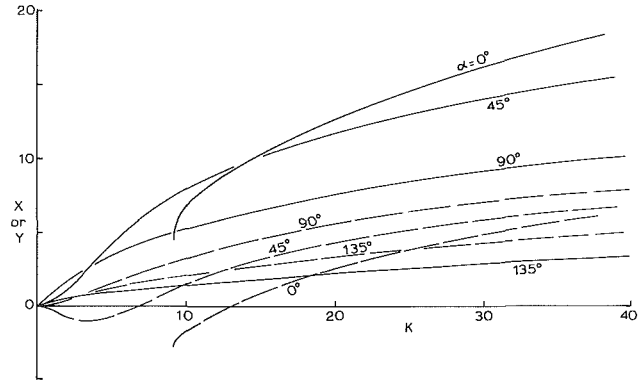


Fig. 8 Forces at the base for the most stable states; ——— horizontal force, ---- vertical force

modes are excited, our results should be accurate. We hope this paper will elicit some experimental research on the interesting flexible cantilever.

References

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

Frisch-Fay, R., 1962, *Flexible Bars*, Butterworths, London.

Hoerner, S. F., 1958, *Fluid Dynamic Drag*, printed by author, New Jersey.

Keefe, R. T., 1961, "An Investigation of the Fluctuating Forces Acting on a Stationary Circular Cylinder in a Subsonic Stream," Univ. Toronto Inst. Aerophys. Rep. No. 76, AFOSR 2147, Quoted in Schnmidt (1965).

McCormick, M. E., 1973, *Ocean Engineering Wave Mechanics*, John Wiley and Sons, New York.

Schlichting, M., 1970, *Boundary Layer Theory*, McGraw-Hill, New York.

Schmidt, L. V., 1965, "Measurments of Fluctuating Air Loads on a Circular Cylinder," *Journal of Aircraft*, Vol. 2, pp. 49–55.

Wang, C. Y., 1984, "On Symmetric Buckling of a Finite Flat-Lying Heavy Sheet," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 278–282.

APPENDIX

Bifurcation Study of the $\alpha = 0$ Case

When $\alpha = 0$, equations (9)–(12) show $\theta = 0$ is a solution. We shall study the possibility of bifurcation from this state. For $\theta << 1$, the equations linearize to

$$\frac{d^2\theta}{ds^2} = X\theta - Y \tag{42}$$

$$\frac{dX}{ds} = K\lambda \tag{43}$$

$$\frac{dY}{ds} = K\lambda\theta. \tag{44}$$

Integration of equations (43) and (44) and using the boundary conditions give

$$X = K\lambda(s-1), \quad Y = K\lambda\int_1^s \theta ds. \tag{45}$$

Substitute into equation (42) and differentiate once yield

$$\frac{d^3\theta}{ds^3} - K\lambda(1-s)\frac{d\theta}{ds} = 0. \tag{46}$$

The general solution is the Airy functions

$$\frac{d\theta}{ds} = C_1 A_i(\mu) + C_2 B_i(\mu) \tag{47}$$

where $\mu \equiv (K\lambda)^{1/3}(1-s)$. But equations (12), (42) show

$$\frac{d\theta}{ds}(1) = \frac{d^2\theta}{ds^2}(1) = 0. \tag{48}$$

Due to uniqueness of an initial value problem, equation (47) gives identically

$$\frac{d\theta}{ds} = 0. \tag{49}$$

Since $\theta(0) = 0$, the only solution is $\theta(s) = 0$, and therefore no bifurcation.

Buckling of beams without the usual pitchfork bifurcation also occur in other instances, for example the heavy horizontal beam (Wang, 1984).

Elastoplastic Buckling of Annular Plates in Pure Shear[1]

E. Ore

D. Durban

Faculty of Aeronautical Engineering,
Technion–I.T.T.,
Haifa 32000, Israel

A linear buckling analysis is presented for annular elastoplastic plates under shear loads. The standard plate buckling equations are used in conjunction with the small strain J_2 flow and deformation theories of plasticity. The main numerical finding is that deformation theory predicts critical loads which are considerably below the predictions obtained with the flow theory. Furthermore, comparison with experimental data for different metals shows a good agreement with the deformation theory results over a wide range of geometries. The limiting buckling problem of a long narrow panel under shear stresses is treated separately. This problem admits an exact solution and it is shown that the critical loads for the panel are approached asymtotically by the annular plate results. Contact is made with earlier studies on the buckling of elastic-orthotropic and elastoplastic shear panels.

1 Introduction

A common method for determining the stress-strain response of metal plates is based on the in-plane torsion test of annular plates. However, the applicability of that test is limited by the phenomenon of out-of-plane buckling when the external torsion moment reaches a critical value.

In this paper we present an elastoplastic buckling analysis of the annular plate in pure shear. The study is within the usual framework of plate buckling theory and small strain plasticity. Material behavior is modeled by the J_2 flow and deformation theories of plasticity with arbitrary hardening characteristics.

The governing equations are given in the next section where we derive the homogeneous fourth-order partial differential equation for the normal velocity at buckling. This equation is supplemented by four homogeneous boundary conditions, for either clamped or simply-supported edges. The resulting eigenvalue problem is examined in Section 3 via a separation of variables solution which leads to an ordinary differential equation for the radial profile of the normal velocity. A finite difference scheme is employed to determine the buckling loads (smallest eigenvalues) for a few representative materials. For thin plates our numerical results agree with the elastic buckling analysis of Dean (1924). In the elastoplastic range there is a difference between the predictions obtained from the two plasticity models. As expected, deformation theory gives critical loads which are considerably below the flow theory results. That difference increases as the plate becomes thicker and can be very wide—up to an order of magitude—in the deep plastic range.

A detailed comparison with recent experimental data obtained by Bauer (1987, 1988) is shown in Section 4. Bauer's tests were made with clamped annular plates over a wide range of geometries and for three different metals. The main finding here is that while deformation theory predictions for the buckling loads are in good agreement with measured results, the flow theory predictions considerably overestimate the values of the experimental critical stresses. That observation highlights the so-called "plastic buckling paradox" (Hutchinson, 1974). It is likely that initial imperfections will lower the maximum load predictions of the flow theory, but engineering-wise, one can use the deformation theory results of the bifurcation loads for all practical purposes.

Finally, in Section 5, we investigate the limiting buckling problem of a narrow shear panel. Here we obtain an exact analytical solution for the bifurcation modes, and the associated eigenvalues follow from simple transcendental equations. Sample calculations for the buckling loads reveal a picture which is very similar to the annular plate results: Flow theory predictions are considerably higher than the deformation theory predictions. We also show that the results for the annular plate approach asymptotically the critical eigenvalues obtained for the narrow panel.

The paper concludes by exploiting a formal correspondence between our elastoplastic analysis and the study by Durban and Stavsky (1982) on the shear buckling of elastic-orthotropic panels. That correspondence leads to useful aysmptotic approximations for the critical loads in the plastic region. A further confirmation of these simplified expressions is obtained from the recipe given by Stowell (1948) for assessing the elastoplastic buckling stress of shear panels.

2 Governing Equations

An annular plate (Fig. 1(*a*)) with inner radius a, outer radius b, and constant thickness h is subjected to uniform shearing

[1]This work was based on part of a thesis to be submitted to the Technion, in partial fulfillment of the requirements for the degree of Doctor of Science.

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 27, 1988; final revision, October 14, 1988.

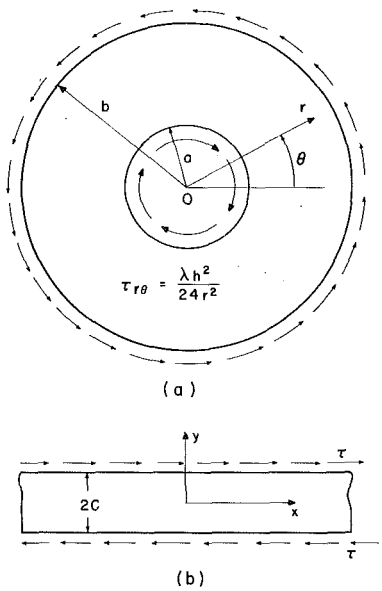


Fig. 1 *(a)* **Notation for the annular plate,** *(b)* **the limiting problem of a long narrow panel. The thickness for both problems is** h.

stresses along the boundaries. The state of stress within the plate is that of pure shear with the statically determined radial profile

$$\tau_{r\theta} = \frac{\lambda h^2}{24 r^2} \tag{1}$$

where λ is a constant, r—the radial coordinate, and the factor of 24 has been introduced for convenience.

The constitutive relations during buckling may be written as

$$\dot{\sigma}_{rr} = E_{rr}\dot{\epsilon}_{rr} + E_{r\theta}\dot{\epsilon}_{\theta\theta} \tag{2}$$

$$\dot{\sigma}_{\theta\theta} = E_{r\theta}\dot{\epsilon}_{rr} + E_{\theta\theta}\dot{\epsilon}_{\theta\theta} \tag{3}$$

$$\dot{\tau}_{r\theta} = 2G_{r\theta}\dot{\epsilon}_{r\theta} \tag{4}$$

where $(\dot{\sigma}_{rr}, \dot{\sigma}_{\theta\theta}, \dot{\tau}_{r\theta})$ are the stress rates, $(\dot{\epsilon}_{rr}, \dot{\epsilon}_{\theta\theta}, \dot{\epsilon}_{r\theta})$ are the strain rates, and $(E_{rr}, E_{r\theta}, E_{\theta\theta}, G_{r\theta})$ are the instantaneous moduli of the material.

Combining (2)–(4) with the standard plate buckling equation, and observing that the only active prebuckling stress component is (1), results in the differential equation

$$\frac{1}{r}\left[r(E_{rr}\kappa_{rr} + E_{r\theta}\kappa_{\theta\theta})\right]_{,rr} - \frac{1}{r}\left(E_{r\theta}\kappa_{rr} + E_{\theta\theta}\kappa_{\theta\theta}\right)_{,r}$$

$$+\frac{4}{r^2}\left(rG_{r\theta}\kappa_{r\theta}\right)_{,r\theta} + \frac{1}{r^2}\left(E_{r\theta}\kappa_{rr} + E_{\theta\theta}\kappa_{\theta\theta}\right)_{,\theta\theta} - \frac{\lambda}{r^2}\kappa_{r\theta} = 0 \tag{5}$$

where the rates of curvature change $(\kappa_{rr}, \kappa_{r\theta}, \kappa_{\theta\theta})$ are given by

$$\kappa_{rr} = w_{,rr} \quad \kappa_{r\theta} = \left(\frac{w_{,\theta}}{r}\right)_{,r} \quad \kappa_{\theta\theta} = \frac{w_{,r}}{r} + \frac{w_{,\theta\theta}}{r^2} \tag{6}$$

with w denoting the out-of-plane velocity at the onset of buckling. Equation (5) is supplemented by boundary conditions which are either those of a clamped edge where

$$w = 0 \qquad w_{,r} = 0, \tag{7}$$

or those of a simply-supported edge where

$$w = 0 \qquad w_{,rr} + \left(\frac{E_{r\theta}}{E_{rr}}\right)\frac{w_{,r}}{r} = 0. \tag{8}$$

The instantaneous moduli in (2)–(4) are determined by the type of constitutive model employed in the analysis. Here we shall use the two small strain versions of the elastoplastic J_2 model. First we have the flow theory with

$$\dot{\sigma}_{ij} = 2G\dot{\epsilon}_{ij} + \lambda_E\delta_{ij}\dot{\epsilon}_{kk} - 3(G - G_T)\frac{S_{ij}S_{kl}\dot{\epsilon}_{kl}}{\sigma_e^2} \tag{9}$$

where G, λ_E are the usual elastic Lamé constants, S_{ij} is the stress deviator, σ_e is the effective stress (in our problem $\sigma_e = \sqrt{3}\,\tau_{r\theta}$), and G_T is the tangent shear modulus defined by

$$\frac{1}{G_T} = \frac{1}{G} + 3\left(\frac{1}{E_T} - \frac{1}{E}\right) \tag{10}$$

where E is the elastic modulus and E_T is the tangent modulus of the uniaxial stress-strain curve (and therefore a known function of the effective stress). The instantaneous moduli associated with (9) for the pure shear field are readily found as

$$E_{rr} = E_{\theta\theta} = \frac{E}{1-\nu^2} \qquad E_{r\theta} = \frac{\nu E}{1-\nu^2} \qquad G_{r\theta} = G_T = G(1 - R_T) \tag{11}$$

where

$$R_T = \frac{1-\eta_T}{1-\left(\frac{1-2\nu}{3}\right)\eta_T} \qquad \eta_T = \frac{E_T}{E} \tag{12}$$

and ν stands for Poisson's ratio. Inserting relation (11)–(12) in (5) and using definitions (6) gives the equation

$$\nabla^4 w - \frac{2(1-\nu)}{r^2}\left[rR_T\left(\frac{w_{,\theta\theta}}{r}\right)_{,r}\right]_{,r} - \frac{S}{r^2}\left(\frac{w_{,\theta}}{r}\right)_{,r} = 0 \tag{13}$$

where

$$S = \frac{1-\nu^2}{E}\lambda. \tag{14}$$

A noteworthy observation here is that the moment-free condition (8) takes the usual elastic form

$$w_{,rr} + \nu\frac{w_{,r}}{r} = 0. \tag{15}$$

The second constitutive model employed in the present study is that of the J_2 deformation theory whose rate form is

$$\dot{\sigma}_{ij} = 2G_s\dot{\epsilon}_{ij} + \lambda_s\delta_{ij}\dot{\epsilon}_{kk} - 3(G_s - G_T)\frac{S_{ij}S_{kl}\dot{\epsilon}_{kl}}{\sigma_e^2} \tag{16}$$

where (G_s, λ_s) are the secant moduli defined by

$$G_s = \frac{E_s}{2(1+\nu_s)} \qquad \lambda_s = \frac{\nu_s E_s}{(1+\nu_s)(1-2\nu_s)} \tag{17}$$

ν_s is the secant Poisson ratio

$$\nu_s = \frac{1}{2} - \left(\frac{1}{2} - \nu\right)\frac{E_s}{E} \tag{18}$$

and E_s is the secant modulus of the uniaxial stress-strain curve. Notice that the secant shear modulus G_s is related to the secant modulus E_s by an expression which resembles (10), namely

$$\frac{1}{G_s} = \frac{1}{G} + 3\left(\frac{1}{E_s} - \frac{1}{E}\right). \tag{19}$$

The instantaneous moduli obtained from the deformation theory follow as

$$E_{rr} = E_{\theta\theta} = \frac{E_s}{1-\nu_s^2} = \bar{E}_s \qquad E_{r\theta} = \nu_s\bar{E}_s \qquad G_{r\theta} = G(1 - R_T) \tag{20}$$

(the shear modulus $G_{r\theta}$ is identical with the one given by the flow theory). Substituting the moduli (20) in (5) and using the kinematical relations (6), we get the equation

$$\frac{1}{r}\left\{r\bar{E}_s\left[w_{,rr}+\nu_s\left(\frac{w_{,r}}{r}+\frac{w_{,\theta\theta}}{r^2}\right)\right]\right\}_{,rr}$$

$$-\frac{1}{r}\left[\bar{E}_s\left(\nu_s w_{,rr}+\frac{w_{,r}}{r}+\frac{w_{,\theta\theta}}{r^2}\right)\right]_{,r}$$

$$+\frac{4}{r^2}G\left[r(1-R_T)\left(\frac{w_{,\theta\theta}}{r}\right)_{,r}\right]_{,r}+\frac{1}{r^2}\bar{E}_s\left(\nu_s w_{,rr}$$

$$+\frac{w_{,r}}{r}+\frac{w_{,\theta\theta}}{r^2}\right)_{,\theta\theta}-\frac{S}{r^2}\left(\frac{E}{1-\nu^2}\right)\left(\frac{w_{,\theta}}{r}\right)_{,r}=0. \tag{21}$$

The moment-free condition from (8) can now be rewritten as

$$w_{,rr}+\nu_s\frac{w_{,r}}{r}=0. \tag{22}$$

Thus, the buckling problem is governed by a partial differential equation, (13) or (21), along the four homogeneous boundary conditions. That system has a nontrivial solution only for certain eigenvalues of the load parameter S (or λ) and we wish to determine the smallest possible $S=S_{cr}$ which is identified with the buckling load. Unlike the analogous elastic buckling problem, equations (13) and (21) have coefficients (R_T, $\bar{E}_s$, ν_s) which vary along the radius in a way that depends on the particular stress-strain characteristics of the material.

3 Solution and Examples

Following the method of Durban and Stavsky (1982), we write the solution for the normal velocity at buckling as

$$w=\mathrm{Re}\{\phi(r)e^{im\theta}\} \qquad m \text{ integer}. \tag{23}$$

Inserting (23) in the flow theory equation (13) results in the ordinary differential equation

$$\phi''''+\frac{2}{r}\phi'''+[2(1-\nu)m^2R_T-1-2m^2]\frac{\phi''}{r^2}$$

$$+[1+2m^2+2(1-\nu)m^2(rR_T'-R_T)-imS]\frac{\phi'}{r^3}$$

$$+[m^4-4m^2-2(1-\nu)m^2(rR_T'-R_T)+imS]\frac{\phi}{r^4}=0 \tag{24}$$

where the prime indicates differentiation with respect to r. Similarly, the deformation theory equation (21) is reduced to

$$\bar{E}_s\phi''''+\frac{2}{r}(r\bar{E}_s'+\bar{E}_s)\phi'''+F_2\frac{\phi''}{r^2}+(F_1-im\lambda)\frac{\phi'}{r^3}$$

$$+(m^2F_o+im\lambda)\frac{\phi}{r^4}=0 \tag{25}$$

where the radial functions F_o, F_1, F_2 are defined by

$$F_o=-r^2\nu_s\bar{E}_s''+r(-2r\nu_s'+2\nu_s+1)\bar{E}_s'$$

$$+(-r^2\nu_s''+2r\nu_s'-2\nu_s-2+m^2)\bar{E}_s-4G(rR_T'+1-R_T) \tag{26}$$

$$F_1=r^2\nu_s\bar{E}_s''+r(2r\nu_s'-2m^2\nu_s-1)\bar{E}_s'$$

$$+(r^2\nu_s''-2m^2r\nu_s'+2m^2\nu_s+1)\bar{E}_s+4m^2G(rR_T'+1-R_T) \tag{27}$$

$$F_2=r^2\bar{E}_s''+(\nu_s+2)r\bar{E}_s'-(2m^2\nu_s+1)\bar{E}_s-4m^2G(1-R_T). \tag{28}$$

The boundary conditions on function ϕ are now

$$\phi=0 \qquad \phi'=0 \tag{29}$$

for a clamped edge and

$$\phi=0 \qquad \phi''+\nu^*\frac{\phi'}{r}=0 \tag{30}$$

for a simply-supported edge, where $\nu^*=\nu$ with the flow theory and $\nu^*=\nu_s$ with the deformation theory.

Equations (24)–(25), along with the proper boundary conditions, can be solved by available numerical schemes based on the finite difference method. This procedure leads to a system of linear homogeneous equations with the requirement for vanishing of the system's determinant Δ at a nontrivial solution. The coefficients of that determinant are complex numbers but the roots (eigenvalues) of the determinant are real. The smallest root of $\Delta=0$ determines the critical eigenvalue S_{cr} at which the plate will buckle.

Sample calculations were performed for three materials represented by the Ramberg-Osgood relation

$$\epsilon=\frac{\sigma_e}{E}+K\left(\frac{\sigma_e}{E}\right)^n \tag{31}$$

where n, K are material constants. The radial profile of the effective stress follows from (1) and (14) as

$$\sigma_e=\frac{\sqrt{3}}{24}\left(\frac{E}{1-\nu^2}\right)\left(\frac{h}{r}\right)^2 S. \tag{32}$$

Accordingly, the tangent and secant moduli become

$$\frac{E_T}{E}=\left\{1+nK\left[\frac{\sqrt{3}\,S}{24(1-\nu^2)}\right]^{n-1}\left(\frac{h}{r}\right)^{2(n-1)}\right\}^{-1} \tag{33}$$

$$\frac{E_s}{E}=\left\{1+K\left[\frac{\sqrt{3}\,S}{24(1-\nu^2)}\right]^{n-1}\left(\frac{h}{r}\right)^{2(n-1)}\right\}^{-1}. \tag{34}$$

All radial coefficients of equations (24)–(25) are therefore explicit functions of the radial coordinate r.

The searching technique which has been used for finding S_{cr} is that of tracing the values of $\mathrm{Re}^2\{\Delta\}+\mathrm{Im}^2\{\Delta\}$ for increasing values of S until the smallest eigenvalue S_{min} is located with sufficient accuracy. This is done for given material constants, fixed geometry (h/a and b/a) and a chosen number m of circumferential waves. The procedure is then repeated with different wave numbers until the smallest (S_{cr}) of the first eigenvalues (S_{min}) is determined.

Figures 2(*a*)–2(*b*) show the variation of the critical eigenvalue S_{cr} with the thickness ratio $h/(b-a)$ for three metals. The uniaxial tension curve is described by (31) with the following material constants:

commercial aluminum:	$E=68700$ MPa $\nu=0.3$	$K=1.27\cdot10^{10}$	$n=3.72$
AL 2014 T6:	$E=69000$ MPa $\nu=0.33$	$K=6.08\cdot10^{31}$	$n=15.62$
ST AISI 4340:	$E=201000$ MPa $\nu=0.28$	$K=7.61\cdot10^{54}$	$n=27.6$

Buckling loads were computed for the radii ratio $b/a=4.18$, with the two J_2 theories, and for clamped and simply-supported plates. Comparison is made also with purely elastic buckling where (24) and (25) coincide with the equation, Dean (1924),

$$\phi''''+\frac{2}{r}\phi'''-(1+2m^2)\frac{\phi''}{r^2}+(1+2m^2-imS)\frac{\phi'}{r^3}$$

$$+(m^4-4m^2+imS)\frac{\phi}{r^4}=0. \tag{35}$$

Initially, for sufficiently thin plates, both theories predict the known linear elastic results: $S_{cr}=88.3$ for the clamped plate (Dean (1924)) and $S_{cr}=52.7$ for the simply-supported plate (Durban and Stavsky (1982)). For thicker plates, however, where buckling occurs in the elastoplastic range, there is a considerable departure of the S_{cr} versus $h/(b-a)$ curves from the purely elastic values. The most striking finding that emerges from the curves displayed in Figs. 2(*a*)–2(*b*) is the increasing

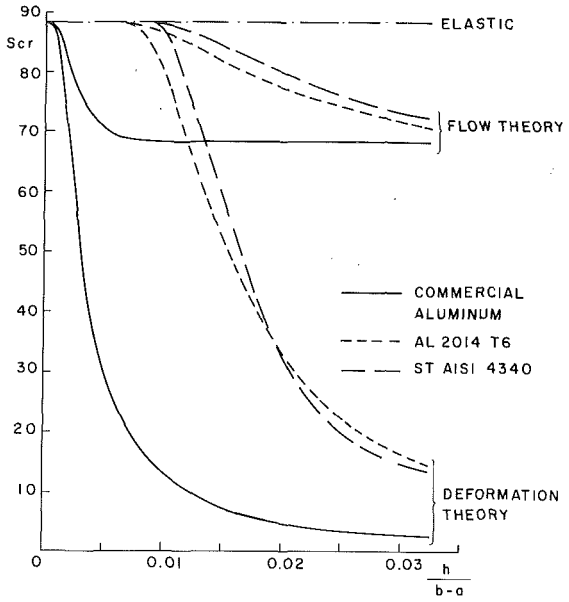


Fig. 2(*a*) Critical eigenvalues for clamped annular plates, $b/a = 4.18$

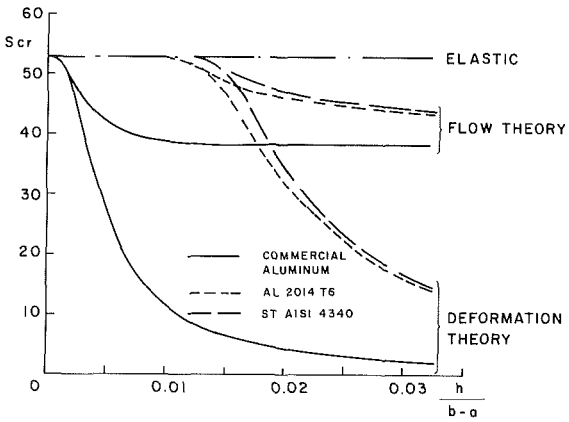


Fig. 2(*b*) Critical eigenvalues for simply-supported plates, $b/a = 4.18$

difference, as the plate is getting thicker, in the prediction of S_{cr} obtained from the two J_2 theories. As expected, the deformation theory gives lower eigenvalues than the flow theory, but the extent of the difference between the corresponding eigenvalues is not common in plastic buckling analysis. That difference appears to increase with the hardening parameter $1/n$. Similar results were obtained for the two cases of mixed boundary conditions with one edge clamped and the other simply-supported. The critical loads for these cases fall between the corresponding results of Figs. 2(*a*) and 2(*b*), and are somewhat higher when the inner boundary is clamped in comparison with a clamped outer boundary. The critical eigenvalues for thin plates, with mixed boundary conditions, were again found to be in agreement with the elastic analysis by Durban and Stavsky (1982). It is worth mentioning in this context that the elastic buckling equation (35) admits an exact solution of the form $\phi \sim r^{\alpha}$ where α stands for the four roots of the characteristic equation associated with (35). Such a simple solution is not possible for the elastoplastic problem, since both (24) and (25) have radially varying coefficients. The only exception to this is the case of a linear-hardening material where E_T is constant; the flow theory equation (24) is then of the Euler type and an exact solution, similar to that of (35), can be found. Note also that in the deep plastic range, where $\eta_T << 1$, we have from (12) that $R_T \approx 1$ and the flow theory equation (24) then describes an elastic-orthotropic solid with nearly constant moduli. This explains the transition of the flow theory curves in Figs. 2(*a*)–2(*b*) to a constant asymptotic value of S_{cr}.

4 Comparison With Experimental Results

We turn now to a comparison of the results of our theoretical analysis with the experimental data reported recently by Bauer (1987, 1988). In these tests several annular plates, with clamped boundaries, were subjected to in-plane torsion. The onset of buckling was determined by tracing the torsion-twist history and observing the formation of a buckling waves pattern. Critical loads were measured for a few metals over a wide range of geometries.

Figures 3(*a*)–3(*c*) display the experimenal values for the critical torsion moment M_{cr} along with the corresponding theoretical predictions obtained from the two J_2 theories. Also shown in Figs. 3(*a*)–3(*c*) is the background curve for purely elastic buckling (35). The critical torsion moment is related to our eigenvalue (14) by the expression

$$M_{cr} = \frac{\pi E h^3}{12(1-\nu^2)} S_{cr}. \tag{36}$$

The stress-strain curves of the tension test for the metals used in the experiments can be described by (31) with the following constants (Lange and Bauer, 1987; Bauer 1987):

CuZn36: $E = 114400$ MPa $\nu = 0.33$ $K = 1.75 \cdot 10^5$ $n = 2.4$
AL 98.7W: $E = 70000$ MPa $\nu = 0.33$ $K = 8.95 \cdot 10^{11}$ $n = 4.46$
ST 1403: $E = 210000$ MPa $\nu = 0.3$ $K = 3.7 \cdot 10^{12}$ $n = 4.9$

It is clearly seen from Figs. 3(*a*)–3(*c*) that the deformation theory predictions are generally in good agreement with the measured values for M_{cr}. Flow theory, by contrast, predicts buckling loads which are considerably above the experimental results. The difference between the critical torsion moments obtained from the two theories is emphasized more in Fig. 4 which shows the ratio M_{cr} (theoretical)/M_{cr} (experimental), for all three metals on a common scale.

It is certainly possible that unavoidable initial imperfections will reduce the maximum load prediction obtained from the flow theory (Hutchinson, 1974). But from a purely practical point of view, the deformation theory analysis of the bifurcation loads appears to be of sufficient reliability. The extent of the difference in the critical eigenvalues predicted by the two theories—for thick plates the ratio between the critical moments can reach an order of magnitude (Fig. 4)—provides a strong example of the "plastic buckling paradox."

The discrepancy between the results for the critical loads, obtained from the flow and deformation theories reflects, of course, the corresponding difference in the magnitude of the instantaneous moduli. Indeed, while moduli (11) of the flow theory essentially retain their elastic values (except $G_{r\theta}$ which is the same in both theories), we find that the deformation theory moduli (20) are considerably below those of the flow theory (again, except $G_{r\theta}$). In the deep plastic range the ratio of moduli E_{ij} obtained from the two theories will be approximately $E_s/E << 1$. A similar observation has been made in the analysis of elastoplastic buckling of the cruciform column (Hutchinson and Budiansksy, 1974).

5 Buckling of a Long Strip in Pure Shear

When the annular plate (Fig. 1(*a*)) becomes very narrow, with $a \to b$, we may expect the critical load to approach that of a long strip (Fig. 1(*b*)) under uniform shear along the boundaries. The solution of the strip problem is fairly simple since, unlike the plate problem, the prebuckling field is homogeneous with $\tau_{xy} \equiv \tau$. The buckling equation is simply

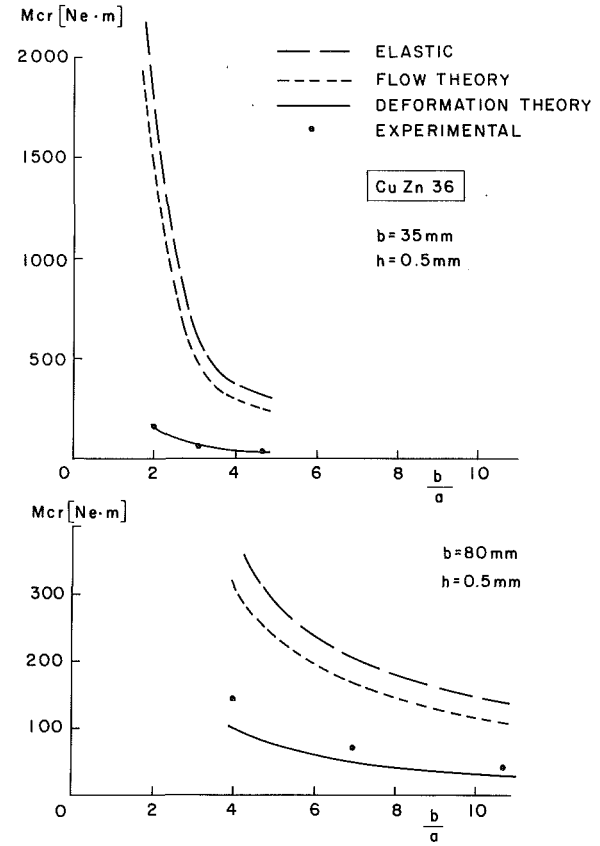


Fig. 3(*a*) Critical torsion moment for CuZn36, clamped boundaries

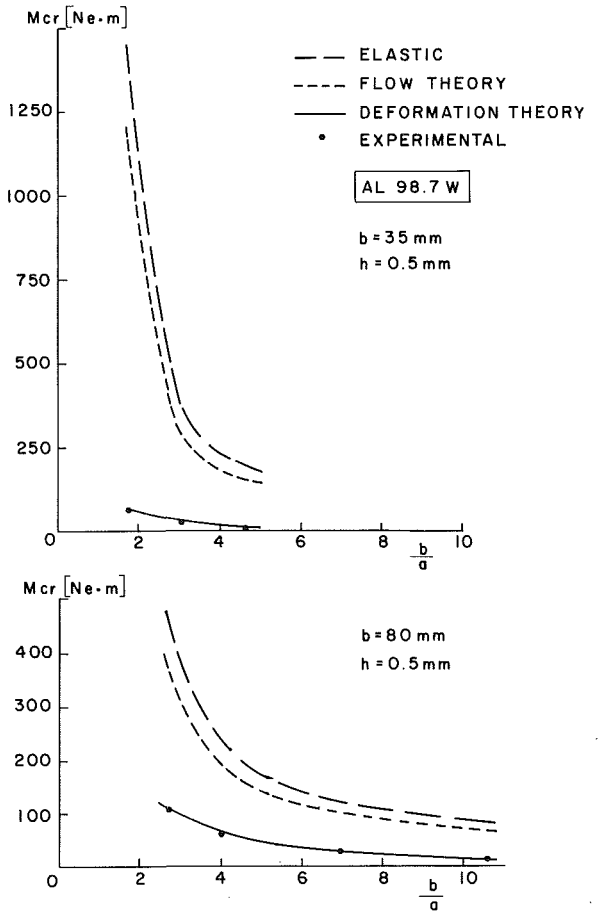


Fig. 3(*b*) Critical torsion moment for Al 98.7 W, clamped boundaries

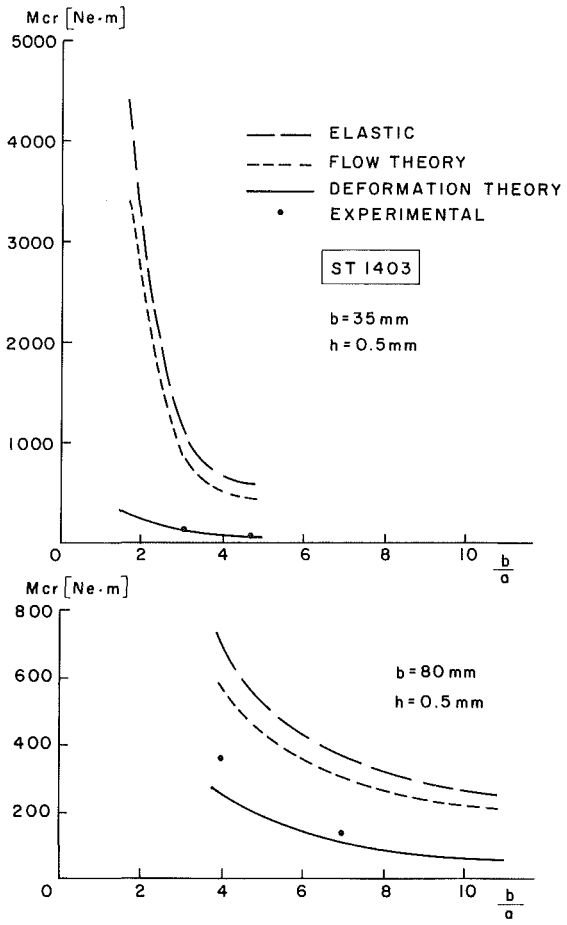


Fig. 3(*c*) Critical torsion moment for ST 1403, clamped boundaries

$$\frac{h^2}{12}\left[E_{xx}w_{,xxxx}+2(E_{xy}+2G_{xy})w_{,xxyy}+E_{yy}w_{,yyyy}\right]$$
$$-2\tau w_{,xy}=0 \tag{37}$$

where $(E_{xx}, E_{xy}, E_{yy}, G_{xy})$ are, with the usual notation, the instantaneous moduli of the material. These moduli are here exactly as in (11) for the flow theory, and as in (20) for the deformation theory, except for the transformation of the (r, θ) directions to the (y, x) directions.

The boundary conditions at the edges $y=\pm c$ are taken as either clamped or simply-supported and we write the solution of (37) in the form

$$w=\mathrm{Re}\{f(y)e^{i\gamma\frac{x}{c}}\} \tag{38}$$

where γ is an unknown parameter. The boundary conditions can now be written as

$$f=0 \quad f'=0 \quad \text{at a clamped edge} \tag{39a}$$

$$f=0 \quad f''=0 \quad \text{at a simply-supported edge} \tag{39b}$$

where here (and throughout this section) the prime denotes differentiation with respect to y. Inserting (38) in (37) we get the ordinary differential equation

$$E_{yy}f''''-2\left(\frac{\gamma}{c}\right)^2(E_{xy}+2G_{xy})f''$$
$$-i\left(\frac{\gamma}{c^3}\right)\lambda^* f'+\left(\frac{\gamma}{c}\right)^4 E_{xx}f=0 \tag{40}$$

where

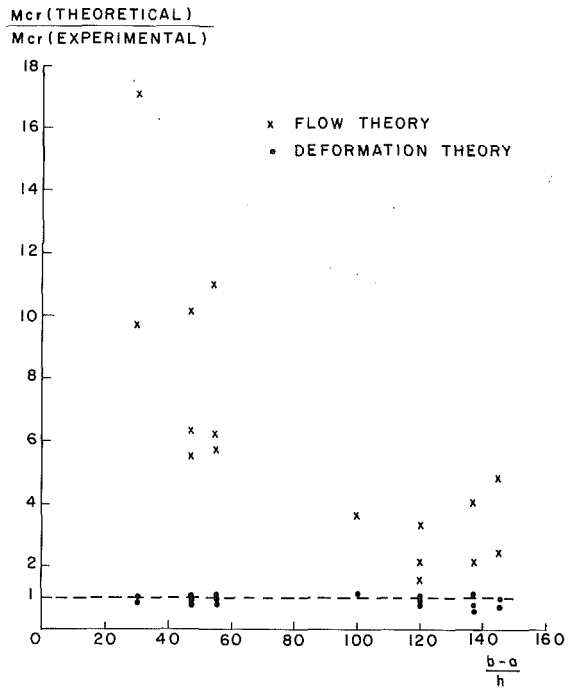


Fig. 4 The ratio M_{cr} (theoretical)/M_{cr} (experimental) for both theories, clamped boundaries

$$\lambda^* = 24\left(\frac{c}{h}\right)^2 \tau. \tag{41}$$

Since the coefficients of (40) are homogeneous we can put the solution for $f(y)$ in the form

$$f = \sum_{p=1}^{4} A_p \exp\left(i\,\frac{\alpha_p}{c}\,y\right) \tag{42}$$

where α_p are the four roots of the characteristic equation

$$E_{yy}\alpha^4 + 2\gamma^2(E_{xy} + 2G_{xy})\alpha^2 + \gamma\lambda^*\alpha + \gamma^4 E_{xx} = 0 \tag{43}$$

and A_p are the four integration constants. Compliance with the boundary conditions (39*a*) or (39*b*) leads to a system of four algebraic equations for constants A_p. The requirement for a nontrivial solution of that system gives the eigenvalue equation for the critical load at which the strip will buckle.

Calculations were made with a few metals (the same as those of Figs. 2(*a*)-2(*b*)) represented by relation (31). Since the effective stress is $\sigma_e = \sqrt{3}\,\tau$ we find that the secant and tangent moduli are here given by

$$\frac{E_T}{E} = \left\{1 + nK\left[\frac{\sqrt{3}\,\lambda^*}{24E}\left(\frac{h}{c}\right)^2\right]^{n-1}\right\}^{-1} \tag{44a}$$

$$\frac{E_s}{E} = \left\{1 + K\left[\frac{\sqrt{3}\,\lambda^*}{24E}\left(\frac{h}{c}\right)^2\right]^{n-1}\right\}^{-1}. \tag{44b}$$

The solution procedure is essentially the same as for the annular plate problem, except that parameter γ of (38) is here continuous. Results for the critical load parameter

$$S^*_{cr} = \frac{1-\nu^2}{2E}\,\lambda^*_{cr} = \frac{3(1-\nu^2)}{E}\left(\frac{2c}{h}\right)^2 \tau_{cr} \tag{45}$$

are shown in Figs. 5(*a*)-5(*b*) for three different metals. The behavior of S^*_{cr} with increasing thickness is similar to what we have seen for the annular plate in Figs. 2(*a*)-2(*b*). In the elastic range, for thin panels, we recover the results of Southwell and Skan (1924): $S^*_{cr} \approx 22.18$ for the clamped panel and $S^*_{cr} \approx 13.21$ for the simply-supported panel. In the plastic range S^*_{cr}

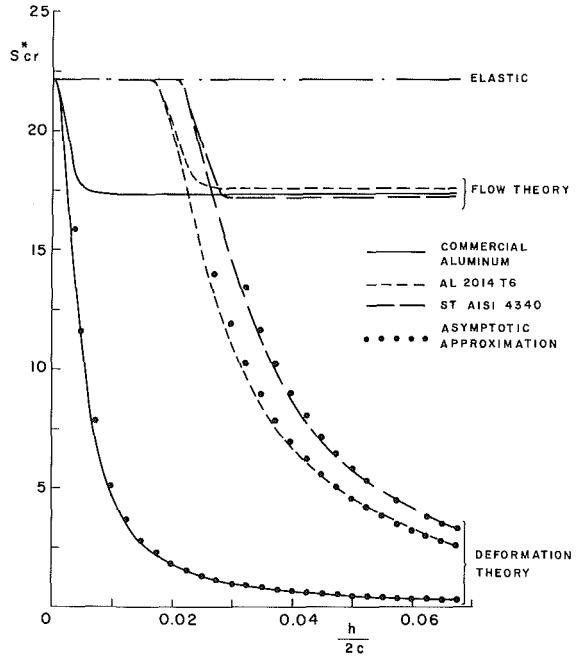


Fig. 5(*a*) Critical eigenvalues for clamped panels

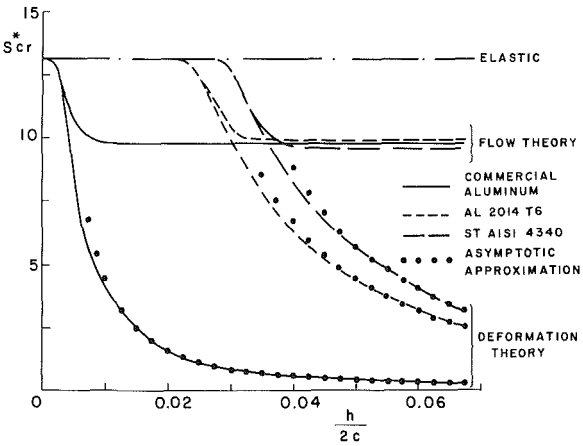


Fig. 5(*b*) Critical eigenvalues for simply-supported panels

decreases with $h/2c$ and there is again a considerable difference in the eigenvalues predicted by the two theories.

Comparison of the critical stresses for the clamped annular plate and the clamped long panel is shown in Fig. 6 for two different materials. The eigenvalues for the annular plate represent in Fig. 6 the critical shear stress at the inner boundary through the parameter

$$T_{cr} = \frac{3(1-\nu^2)}{E}\left(\frac{b-a}{h}\right)^2 \tau_{cr}(r=a). \tag{46}$$

The curves for the shear panel show the variation of S^*_{cr} from (45) with the thickness ratio $h/2c$. In calculating the annular plate curves we have assumed that h/a is maintained constant so that T_{cr} from (46) can be determined for every value of b/a. The variation of T_{cr} with $h/(b-a)$ is shown in Fig. 6 for decreasing values of b/a. It can be clearly seen that T_{cr} approaches asymptotically S^*_{cr} as $b/a \to 1$ (implying that $b-a \to 2c$ for the same thickness). T_{cr} is always higher than S^*_{cr} and the deformation theory curves are closer to the shear panel asymptotes in comparison with the flow theory and purely elastic curves.

Since the instantaneous plastic moduli of the shear panel are homogeneous, it is possible to establish a formal cor-

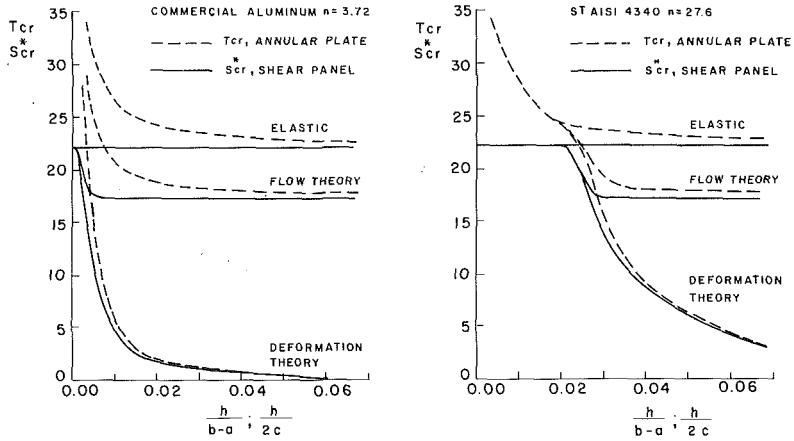


Fig. 6 Comparison of critical eigenvalues for the clamped annular plate (T_{cr}) and the clamped long panel (S^*_{cr}); the curves of the annular plate are computed with $h/a = 0.002$

respondence with the buckling analysis of elastic orthotropic panels. That problem has been studied recently by Durban and Stavsky (1982), where references to earlier work can be found. Critical loads for the elastic orthotropic panel are presented through an eigenvalue k which, in the present notation and with $E_{yy} = E_{xx}$, is defined by

$$k = 3\left(\frac{2c}{h}\right)^2 \frac{\tau_{cr}}{E_{xx}}. \tag{47}$$

The dependence of k on the material parameter

$$\bar{E} = \frac{E_{xx}}{E_{xy} + 2G_{xy}} \tag{48}$$

is shown in Fig. 6 of Durban and Stavsky (1982) for the buckling of elastic orthotropic panels with various boundary conditions. These results can be made to fit to our problem (with $\bar{E} \geq 1$) of plastic buckling, where the instantaneous moduli are load dependent, in the following way: Assuming a value for τ_{cr}, we find from (48) the associated value of $\bar{E}$. The corresponding value of k is then obtained from Fig. 6 in Durban and Stavsky (1982), and (47) will determine the proper thickness ratio $h/2c$ for the assumed critical stress τ_{cr}. Put differently, we have from (45) and (47) the critical load parameter in the form

$$S^*_{cr} = (1-\nu^2)k\frac{E_{xx}}{E}. \tag{49}$$

Sample calculations made by this method have shown a very good agreement with the curves of Figs. 5(*a*)–5(*b*) over the entire elastoplastic range. Furthermore, simple and useful asymptotic approximations for the behavior of S^*_{cr} in the plastic range can be deduced from the elastic-orthotropic curves.

With the flow theory we get from (11)–(12), that in the plastic range, where $R_T \approx 1$, $G_{xy} << E_{xy}$. Hence, parameter $\bar{E}$ from (48) is nearly constant and equals $\bar{E} = 1/\nu$. Denoting the corresponding value of k by $k(1/\nu)$, we find from (49) that

$$S^*_{cr} = k\left(\frac{1}{\nu}\right). \tag{50}$$

Thus, the flow theory results for S^*_{cr} exhibit a transition from the linear elastic values to elastic-orthotropic values which remain constant in the plastic range. That phenomenon is confirmed by the curves of Figs. 5(*a*)–5(*b*) (see also Fig. 6), and the trend can be seen also from the annular plate results in Figs. 2(*a*)–2(*b*). For the three materials considered here we have the values of $\nu = 0.3$ for the commercial aluminum, $\nu = 0.33$ for AL 2014 T6, and $\nu = 0.28$ for ST AISI 4340. The corresponding values of $k(1/\nu)$, as read from Fig. 6 in Durban and Stavsky (1982), are (17.3, 17.5, 17.2)[2] for the clamped panel, and (9.80, 9.93, 9.63) for the simply-supported panel. These values agree with the asymptotic results for S^*_{cr} in Figs. 5(*a*)–5(*b*) to within less than a percent.

With the deformation theory we find from (20), with the aid of (18), that in the plastic range, where $\nu_s \approx 1/2$, $\bar{E} \approx 1/\nu_s \approx 2$, the critical eigenvalue follows from (49) as

$$S^*_{cr} = \frac{4(1-\nu^2)}{3} k(2) \frac{E_s}{E} \tag{51}$$

where $k(2)$ is the elastic orthotropic value of k at $\bar{E} = 2$. Inserting in (51) the power-law approximations of (44*b*) and using (45) gives the asymptotic approximation

$$S^*_{cr} = B\left(\frac{h}{2c}\right)^{-\frac{2(n-1)}{n}}$$

with

$$B = \sqrt{3}(1-\nu^2)\left[\frac{4k(2)}{3\sqrt{3}\,K}\right]^{1/n}. \tag{52}$$

The elastic orthotropic values of $k(2)$ are 18.6 for the clamped panel and 10.9 for the simply-supported panel. The values of coefficient B in the asymptotic approximation (52) for the three metals follow as ($6.2 \cdot 10^{-3}$, $1.69 \cdot 10^{-2}$, $1.81 \cdot 10^{-2}$) for the clamped panel and ($5.37 \cdot 10^{-3}$, $1.63 \cdot 10^{-2}$, $1.77 \cdot 10^{-2}$) for the simply-supported panel. Comparison of (52) with the exact solution shows an excellent agreement (Figs. 5(*a*)–5(*b*)) in the plastic range.

The elastoplastic buckling problem of shear panels has been investigated by Stowell (1948). Using a one term approximation for w, in conjunction with a minimum principle and the deformation theory, the critical load is determined by

$$\tau_{cr} = \frac{1}{\sin 2\phi}\left\{2\sqrt{f_1}\sqrt{1 - \frac{1-c_3}{2}\sin^2 2\phi} + 2f_2[1 + 2\sin^2\phi - (1-c_3)\cos^2\phi]\right\}\frac{\pi^2}{9}\left(\frac{h}{2c}\right)^2 E_s \tag{53}$$

where the wave angle ϕ is found from the solution of the equation

[2] Numbers in brackets are for commercial aluminum, AL 2014 T6, and ST AISI 4340, respectively.

$$\cos 2\phi = \frac{(3-c_3)f_2}{\dfrac{2\sqrt{f_1}}{\sqrt{1-\dfrac{1-c_3}{2}\sin^2 2\phi}} + (3+c_3)f_2}. \tag{54}$$

Here $c_3 = (1 + E_T/E_s)/2$ and (f_1, f_2) are constants that depend on the boundary conditions. For clamped boundaries, $f_1 = 5.14$, $f_2 = 1.24$, while for simply-supported boundaries, $f_1 = f_2 = 1$.

Now, for the Ramberg-Osgood materials (31) we have in the plastic range $c_3 \approx (n+1)/2n$. The solution of (54) is then load-independent, and for the three metals of Figs. 5(*a*)–5(*b*) we find the shear angles ϕ [rad] as (0.628, 0.621, 0.620) for the clamped panel, and (0.576, 0.564, 0.562) for the simply-supported panel. Substituting these values of ϕ in (53), and using the power law approximation for E_s, we recover again the asymptotic expansion (52) with identical power terms and virtually the same values of coefficient B. Stowell (1948) has compared the theoretical predictions of (53)–(54) with the experimental measurements of Gerard (1948) on the buckling of 24S-0 aluminum alloy shear panels. The agreement of the test data with the deformation theory predictions (53)–(54), over the range of thickness ratio $2c/h \approx 45 \div 103$, reveals essentially the same picture as in Fig. 4 of the present paper.

Acknowledgment

We are indebted to Mr. Michael Bauer for his helpful comments and kind assistance in providing the experimental data during the preparation of the manuscript.

References

Bauer, M., 1987, "Faltenbildung beim ebenen torsionsversuch," *Ingenieur-Archiv*, Vol. 57, pp. 39–50.

Bauer, M., 1988, Doctoral thesis, Stuttgart University.

Dean, W. R., 1924, "The Elastic Stability of an Annular Plate," Proc. Roy. Soc. Lond., Vol. A106, pp. 268–284.

Durban, D., and Stavsky, Y., 1982, "Elastic Buckling of Polar Orthotropic Annular Plates in Shear," *Int. J. Solids Structures*, Vol. 18, pp. 51–58.

Gerard, G., 1948, "Critical Shear Stress of Plates Above the Proportional Limit," ASME JOURNAL OF APPLIED MECHANICS, Vol. 15, pp. 7–12.

Hutchinson, J. W., 1974, "Plastic Buckling," *Advances in Applied Mechanics*, Vol. 14, pp. 67–144.

Hutchinson, J. W., and Budiansky, B., 1976, "Analytical and Numerical Study of the Effects of Initial Imperfections on the Inelastic Buckling of a Cruciform Column," *Buckling of Structures*, IUTAM Symposium, Springer-Verlag, pp. 98–105.

Lange, K., and Bauer, M., 1987, "Determining Flow Curves of Thin Sheet Metal by the Plane Torsion Test," *SME Technology Review*, Vol. 2, pp. 346–352.

Southwell, R. V., and Skan, S. W., 1924, "On the Stability Under Shearing Forces of a Flat Elastic Strip," Proc. Roy. Soc. Lond., Vol. A105, pp. 582–607.

Stowell, E. Z., 1948, "Critical Shear Stress of an Infinitely Long Plate in the Plastic Region," NACA TN No. 1681.

Moderately Thick Angle-Ply Cylindrical Shells Under Internal Pressure

Kamal R. Abu-Arja
Postdoctoral Fellow and Instructor.

Reaz A. Chaudhuri[1]
Assistant Professor,
Assoc. Mem. ASME

Department of Civil Engineering,
University of Utah,
Salt Lake City, Utah 84112

Heretofore unavailable closed-form solutions are obtained for unbalanced symmetric as well as balanced unsymmetric angle-ply, moderately thick cylindrical shells subjected to axially varying (axisymmetric) internal pressure loading, under the framework of constant shear-angle theory (CST) or first-order shear deformation theory (FSDT), for any boundary condition. The solutions are obtained for four CST-based kinematic relations, which are extensions of the classical shell theories due to Donnell, Love-Timoshenko, Love-Riessner, and Sanders. The available CLT (classical lamination theory)-based solutions can be obtained from the present solutions in the limiting case of the two transverse shear moduli tending to infinity. Numerical results have been presented for two layer and three layer angle-ply cylindrical shells with simply-supported edges and have been compared with the corresponding CLT-based analytical solutions and also the LCST (layerwise constant shear angle theory)-based finite element solutions.

1 Introduction

Laminated composite structures offer the advantage of high strength-to-weight and stiffness-to-weight ratios, corrosion resistance, fatigue (including sonic fatigue) life, and the possibility of optimum design through the variation of fiber orientation, stacking pattern and choice of fiber, and matrix materials (Calcote, 1969). Due to the low transverse shear moduli relative to the in-plane Young's moduli, the transverse shear deformation effects are more pronounced in laminated fiber-reinforced structures when compared to their isotropic counterparts under same loading. A number of analyses for laminated anisotropic shells, based on the classical lamination theory (CLT), which neglects transverse shear deformation altogether (Love's first approximation theory or Love-Kirchhoff hypothesis) exist in the literature. Surveys of these analyses can be found in the works of Bert (1974a,b), Chaudhuri (1974, 1983) and Abu-Arja (1986). Dong et al. (1962) have formulated a theory of laminated anisotropic shell, which may be considered to be an extension of the work of Reissner and Stavsky (1961) on laminated anisotropic plates to Donnell's shallow shell theory. Cheng and Ho (1963) have presented a CLT-based analysis of laminated anisotropic cylindrical shells using Flugge's kinematic relations. Reuter (1972) has utilized Donnell's shallow shell theory in obtaining CLT-based analytical solutions for angle-ply cylindrical shells under the influence of uniform internal pressure and temperature. Balaraman et al. (1972) and Chaudhuri et al. (1986) have obtained CLT-based closed-form solutions for an arbitrarily laminated cylindrical shell of finite length under uniform internal pressure, utilizing Love-Timoshenko's kinematic relations. The latter work also presents finite element solutions, based on the LCST (layerwise constant shear-angle theory, which assumes layerwise linear variation of the surface-parallel components of displacement), due to Chaudhuri (1983) and Seide and Chaudhuri (1987) for comparison. Bert and Reddy (1982) have presented exact solutions for bending under sinusoidal transverse loading of two-layer thin cylindrical shells of bimodulus material. Reddy (1984) has used the CST (constant shear-angle theory), also known as FSDT (first-order shear deformation theory), based on Mindlin-Reissner hypothesis, to present series solutions for cross-ply open shallow shells of cylindrical as well as doubly-curved geometries utilizing Sanders' kinematic relations. Wilson and Orgill (1986) have studied parametrically the deformation behavior of uniformly-stressed orthotropic cylindrical shells. Hutchinson and El-Azhari (1986) have developed a series solution of the general three-dimensional equations of linear elasticity for studying vibrations of isotropic hollow circular cylindrical body with traction-free surfaces. A brief literature search reveals the nonexistence of exact solutions to the problem of closed lamianted shells of finite length, which incorporate either an approximate shear deformation theory (e.g., CST) or a three-dimensional elasticity theory into the formulation, even for such simple geometry as circular cylinder or such simple loading condition as uniform internal pressure. Recently, Abu-Arja and Chaudhuri (in review) have presented solutions for axisymmetric cross-ply cylindrical shells subjected to uniform internal pressure, under the framework of the CST. The purpose of the present study is to (*a*) obtain exact solutions for balanced unsymmetric as well as unbalanced symmetric angle-ply shear-flexible moderately

[1]To whom correspondence should be sent.

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, N. Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 19, 1987; final revision, August 6, 1987.

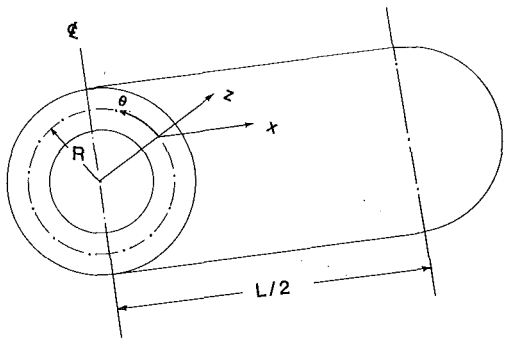


Fig. 1 Geometry of the cylindrical shell

thick cylindrical shells, subjected to axially-varying internal pressure for arbitrary boundary conditions and (*b*) identify the nature of shear deformation behavior of these angle-ply shells by way of comparison with the available CLT-based analytical and LCST-based finite element solutions.

2 Statement of the Problem

The strain-displacement relations of a circular cylindrical shell (Fig. 1), subjected to axisymmetric loading are given by (Reddy, 1984)

$$\epsilon_1 = \epsilon_1^0 + \zeta\kappa_1; \quad \epsilon_2 = \epsilon_2^0 + \zeta\kappa_2; \quad \epsilon_4 = \epsilon_4^0; \quad \epsilon_5 = \epsilon_5^0; \quad \epsilon_6 = \epsilon_6^0 + \zeta\kappa_6 \tag{1}$$

where, using $\partial(\ \ldots\ )/\partial\theta = 0$, because of axisymmetry

$$\epsilon_1^0 = u_{0,x}; \quad \epsilon_2^0 = w/R; \quad \epsilon_4^0 = \phi_\theta - v_0/R; \quad \epsilon_5^0 = w_{,x} + \phi_x; \quad \epsilon_6^0 = v_{0,x}; \quad \kappa_1 = \phi_{x,x}; \quad \kappa_2 = 0; \quad \kappa_6 = \phi_{\theta,x} + c_0 v_{0,x}/R, \tag{2}$$

in which u_0, v_0 are the reference (middle) surface stretching, w is the transverse or radial displacement; ϕ_x, ϕ_θ are the rotations of the reference surface about θ- and x-coordinate axes, respectively. The x and θ-coordinates are equivalent to 1 and 2-coordinates, respectively, and R is the mean radius of the cylinder. c_0 is a constant and assumes the values of -1, 0, 1/2, and 1 for extension of kinematic relations, based on Love's first approximation theory due to Donnell, Reissner, Sanders, and Timoshenko (Kraus, 1967 and Chaudhuri et al., 1986), respectively, to the case of the CST. The equations of equilibrium for a circular, axisymmetric cylindrical shell subjected to axially-varying internal pressure, $p(x)$, are given by (Timoshenko and Woinowsky-Krieger, 1959)

$$N_{x,x} = 0; \quad N_{x\theta,x} + Q_\theta/R = 0; \quad N_\theta/R - Q_{x,x} = p(x); \quad M_{x,x} = Q_x; \quad M_{x\theta,x} = Q_\theta; \quad N_{x\theta} - N_{\theta x} - M_{\theta x}/R = 0; \tag{3}$$

where N_x, N_θ, $N_{\theta x}$, $N_{x\theta}$ are the stress resultants, M_x, M_θ, $M_{x\theta}$, $M_{\theta x}$ are moment resultants, and Q_x, Q_θ are transverse shear forces, all per unit length. The stress resultants, moment resultants (stress couples), and transverse shear forces are expressed in terms of the reference surface strains and changes of curvatures and twist as follows:

$$N_x = A_{11}\epsilon_1^0 + A_{12}\epsilon_2^0 + A_{16}\epsilon_6^0 + B_{11}\kappa_1 + B_{12}\kappa_2 + B_{16}\kappa_6 \tag{4a}$$

$$N_\theta = A_{12}\epsilon_1^0 + A_{22}\epsilon_2^0 + A_{26}\epsilon_6^0 + B_{12}\kappa_1 + B_{22}\kappa_2 + B_{26}\kappa_6 \tag{4b}$$

$$N_{x\theta} = A_{16}\epsilon_1^0 + A_{26}\epsilon_2^0 + A_{66}\epsilon_6^0 + B_{16}\kappa_1 + B_{26}\kappa_2 + B_{66}\kappa_6 + (c_0/R)(B_{16}\epsilon_1^0 + B_{26}\epsilon_2^0 + B_{66}\epsilon_6^0 + D_{16}\kappa_1 + D_{26}\kappa_2 + D_{66}\kappa_6) \tag{4c}$$

$$N_{\theta x} = A_{16}\epsilon_1^0 + A_{26}\epsilon_2^0 + A_{66}\epsilon_6^0 + B_{16}\kappa_1 + B_{26}\kappa_2 + B_{66}\kappa_6 \tag{4d}$$

$$M_x = B_{11}\epsilon_1^0 + B_{12}\epsilon_2^0 + B_{16}\epsilon_6^0 + D_{11}\kappa_1 + D_{12}\kappa_2 + D_{16}\kappa_6 \tag{4e}$$

$$M_\theta = B_{12}\epsilon_1^0 + B_{22}\epsilon_2^0 + B_{26}\epsilon_6^0 + D_{12}\kappa_1 + D_{22}\kappa_2 + D_{26}\kappa_6 \tag{4f}$$

$$M_{\theta x} = M_{x\theta} = B_{16}\epsilon_1^0 + B_{26}\epsilon_2^0 + B_{66}\epsilon_6^0 + D_{16}\kappa_1 + D_{26}\kappa_2 + D_{66}\kappa_6 \tag{4g}$$

$$Q_x = A_{45}\epsilon_4^0 + A_{55}\epsilon_5^0 \tag{4h}$$

$$Q_\theta = \mathrm{K}_{44}\epsilon_4^0 + A_{45}\epsilon_5^0 \tag{4i}$$

where A_{ij}, B_{ij}, and D_{ij} ($i,j = 1, 2, 6$) are extensional, bending-stretching coupling, and bending rigidities, respectively, while A_{ij} ($i,j = 4, 5$) are transverse shear rigidities.

3 Exact Solution of Balanced Unsymmetric Angle-Ply Cylindrical Shell

Exact solution to the problem of a pressurized balanced unsymmetric (for definition, see Reuter (1972)), angle-ply cylindrical shell is derived in this section for arbitrary boundary conditions. For this type of shells,

$$A_{16} = A_{26} = A_{45} = B_{11} = B_{12} = B_{22} = B_{66} = D_{16} = D_{26} = 0. \tag{5}$$

Substitution of equations (2), (4), and (5) into equations (3) will yield

$$A_{11}u_{0,xx} + A_{12}w_{,x}/R + (c_0+1)B_{16}v_{0,xx}/R + B_{16}Q_{\theta,xx}/A_{44} = 0 \tag{6a}$$

$$Q_\theta + c_0 B_{16}u_{0,xx} + c_0 B_{26}w_{,x}/R - RB_{16}w_{,xxx} + RB_{16}Q_{x,xx}/A_{55} + [(c_0+c_0^2)D_{66}/R + RA_{66}]v_{0,xx} + c_0 D_{66}Q_{\theta,xx}/A_{44} = 0 \tag{6b}$$

$$A_{12}u_{0,x}/R + A_{22}w/R^2 + (c_0+1)B_{26}v_{0,x}/R^2 + B_{26}Q_{\theta,x}/(RA_{44}) - Q_{x,x} = p(x) \tag{6c}$$

$$-Q_x + B_{16}v_{0,xx} - D_{11}w_{,xxx} + D_{11}Q_{x,xx}/A_{55} = 0 \tag{6d}$$

$$-Q_\theta + B_{16}u_{0,xx} + B_{26}w_{,x}/R + (c_0+1)D_{66}v_{0,xx}/R + D_{66}Q_{\theta,xx}/A_{44} = 0. \tag{6e}$$

Equation (6*a*), on integration, gives

$$u_{0,x} = -A_{12}w/(A_{11}R) - (c_0+1)B_{16}v_{0,x}/(RA_{11}) - B_{16}Q_{\theta,x}/(A_{11}A_{44}) + C_1 \tag{7}$$

where C_1 is an integration constant. Substitution of $u_{0,x}$, as given by equation (7) into equations (6*b*)–(6*e*), and then successive elimination of v_0, Q_θ, and Q_x from the resulting four coupled ordinary differential equations (O.D.E.), will finally yield a decoupled seventh-order O.D.E in terms of w. This, on integration, will reduce to a sixth-order O.D.E.

$$(A_1D^6 + A_2D^4 + A_3D^2 + A_4)w = (A_5D^4 + A_6D^2 + A_7)p(x) + C_2 \tag{8}$$

where the symbolic operator D^n is defined by

$$D^n = d^n/dx^n \qquad (n = 1, \ldots, 6) \tag{9}$$

and A_i ($i = 1, \ldots, 7$) are given by equations (*A*1), while C_2 is an integration constant. It is noteworthy that in the case of uniform internal pressure, $p(x) = p_0$, the aforementioned seventh-order O.D.E. reduces a homogeneous O.D.E., which implies that the r.h.s. of equation (8) becomes merely an integration constant, C_2^*. It is then evident that the solution, due to the uniform pressure, is a degenerate case of its counterpart due to axially-varying pressure loading, with $C_2^* = A_7p_0 + C_2$. The equation (8), on substitution of $G_{13} = G_{23} \to \infty$, reduces to its CLT-counterpart (Chaudhuri et al., 1986) which, for the case of uniform internal pressure, p_0, is given by

$$A_1^*D^6 + A_2^*D^4 + A_3^*D^2 + A_4^*)w = A_7^*p_0 + C_2 \tag{10}$$

where A_i^* ($i=2, 3, 4, 7$) are as presented in equations (A2). It may also be noted that

$$A_1^* = \lim_{G_{13}=G_{23}\to\infty} A_1 \to 0.$$

w_c, the complementary solution of equation (8), is obtained by assuming $w=e^{\lambda x}$, which yields the characteristic equation

$$\lambda^6 + a_1\lambda^4 + a_2\lambda^2 + a_3 = 0 \tag{11}$$

where

$$a_i = A_{i+1}/A_1 \qquad (i=1, \ldots, 3). \tag{12}$$

Considering $\lambda^2 = m$ reduces equation (11) to a third-degree polynomial which can easily be solved (Korn and Korn, 1968),

$$w_c = \begin{cases} B_1\sinh(\beta_1 x)\sin(\alpha_1 x) + B_2\sinh(\beta_1 x)\cos(\alpha_1 x) + B_3\cosh(\beta_1 x)\cos(\alpha_1 x) + \\ B_4\cosh(\beta_1 x)\sin(\alpha_1 x) + B_5\cosh(\gamma_1 x) + B_6\sinh(\gamma_1 x) & \text{if } F>0 \text{ and } m_1>0 \\ B_1\sinh(\beta_1 x)\sin(\alpha_1 x) + B_2\sinh(\beta_1 x)\cos(\alpha_1 x) + B_3\cosh(\beta_1 x)\cos(\alpha_1 x) + \\ B_4\cosh(\beta_1 x)\sin(\alpha_1 x) + B_5\cos(\gamma_1 x) + B_6\sin(\gamma_1 x) & \text{if } F>0 \text{ and } m_1<0 \\ \sum_{i=1}^{3} f_i(x) & \text{if } F<0 \end{cases} \tag{13}$$

where α_1, β_1, γ_1, m_1, and F are as given by equations (A3). B_i ($i=1, \ldots, 6$) are integration constants and $f_i(x)$ are defined as follows:

$$f_i(x) = \begin{cases} B_i\cosh(\Phi_i x) + B_{i+i}\sinh(\Phi_i x) & \text{if } \mu_i>0 \\ B_i\cos(\Phi_i x) + B_{i+i}\sin(\Phi_i x) & \text{if } \mu_i<0 \end{cases} \tag{14}$$

with μ_i ($i=1, 2, 3$) being given by equations (A4) and $\Phi_i = \sqrt{|\mu_i|}$.

If $p(x)$ is prescribed, w_p, the particular integral of equation (8), can easily be obtained. For example, for the case $p(x)=p_0$,

$$w_p = A_7 p_0 + C_2. \tag{15}$$

Once the complete solution, $w = w_c + w_p$, is known, the remaining quantities, Q_x, Q_θ, u_0, and v_0 can easily be obtained, which are as follows:

$$Q_x = h_5 w^{(5)} + h_6 w_{,xxx} + h_7 w_{,x} \tag{16a}$$

$$Q_\theta = j_1 w^{(5)} + j_2 w_{,xxx} + j_3 w_{,x} \tag{16b}$$

$$v_0 = j_4 w^{(5)} + j_5 w_{,xxx} + j_6 w_{,x} + j_7\int w + (j_8 + j_9 C_1)x + C_3 \tag{16c}$$

$$u_0 = k_1 w^{(5)} + k_2 w_{,xxx} + k_3 w_{,x} + k_4\int w + (k_5 + k_6 C_1)x + C_4 \tag{16d}$$

where C_3 and C_4 are integration constants, while h_i ($i=5, 6, 7$), j_i ($i=1, \ldots, 9$) and k_i ($i=1, \ldots, 6$) are as given by equations (A5).

The ten integration constants, B_i ($i=1, \ldots, 6$) and C_i ($i=1, \ldots, 4$), can be obtained by prescribing five boundary conditions at each of the ends ($x=\pm L/2$) of the cylindrical shell. The five boundary conditions are chosen to be one member from each pair of the following equations

$$(w, Q_x) = (M_x, \phi_x) = (N_x, u_0) = (N_{x\theta}, v_0) = (\phi_\theta, M_{x\theta}) = 0. \tag{17}$$

In the event the loading and the boundary conditions are symmetric, with respect to the central section of the cylinder, $B_2 = B_4 = B_6 = C_3 = C_4 = 0$ and the remaining five constants of integration are obtained from the prescribed boundary condition at one end only.

4 Exact Solution of Unbalanced Symmetric Angle-Ply Cylindrical Shell

Exact solution to the problem of a pressurized unbalanced symmetric (for definition, see Reuter (1972)) angle-ply cylindrical shell is derived in this section for arbitrary boundary conditions. For a symmetric angle-ply shell,

$$A_{45} = 0; \qquad B_{ij} = 0 \qquad (i,j = 1, 2, 6). \tag{18}$$

Substitution of equations (2), (4), and (18) into equations (3) will yield

$$A_{11}u_{0,xx} + A_{12}w_{,x}/R + A_{16}v_{0,xx} = 0 \tag{19a}$$

$$Q_\theta + RA_{16}u_{0,xx} + A_{26}w_{,x} - c_0 D_{16}w_{,xxx} + c_0 D_{16}Q_{x,xx}/A_{55} + [c_0(c_0+1)D_{66}/R + RA_{66}]v_{0,xx} + c_0 D_{66}Q_{\theta,xx}/A_{44} = 0 \tag{19b}$$

$$A_{12}u_{0,x}/R + A_{22}w/R^2 + A_{26}v_{0,x}/R = Q_{x,x} = p(x) \tag{19c}$$

$$-Q_x - D_{11}w_{,xxx} + D_{11}Q_{x,xx}/A_{55} + (c_0+1)D_{16}v_{0,xx}/R + D_{16}Q_{\theta,xx}/A_{44} = 0 \tag{19d}$$

$$-Q_\theta - D_{16}w_{,xxx} + D_{16}Q_{x,xx}/A_{55} + D_{66}Q_{\theta,xx}/A_{44} (c_0+1)D_{66}v_{0,xx}/R = 0. \tag{19e}$$

The system of five coupled O.D.E.'s is solved following the procedure presented in the preceding section. The form of the resulting sixth-order O.D.E., in terms of w and its derivatives, is identical to its counterpart of the unsymmetric case. In the interest of brevity the details of the solution, which are available in Abu-Arja (1986), will not be presented here.

5 Numerical Results

Before obtaining numerical results for moderately thick angle-ply cylindrical shells, thin unidirectional ($R/t=100$) shells of the same ply material and of otherwise the same geometry, were investigated. The three theories—CLT, CST, and LCST—predicted almost identical results. It is therefore presumed that any difference in the predictions of these three theories, in the case of moderately thick angle-ply shells, can be attributed to the effects of thickness and of varying fiber orientation from layer to layer.

The present study investigates, as the first example, a two-layer balanced unsymmetric cylindrical shell with fiber orientations of the layers, $-\theta/\theta$. The layers are of equal thickness. The inner layer makes an angle, $-\theta$, negative sign implying clockwise sense with respect to the positive direction of the generator (x-axis). The length, L, of the cylindrical shell and Young's modulus, E_1, in the direction parallel to the fibers are 5080 mm (200 in.) and 275.8 GPa (40×10^6 psi), respectively.

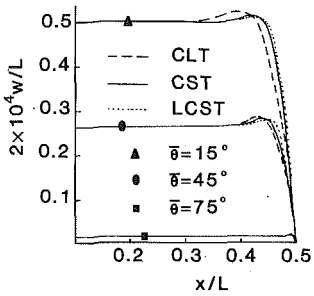


Fig. 2 Variation of displacement, w, along the axial direction of cylindrical shells for -15 deg/15 deg, -45 deg/45 deg, and -75 deg/75 deg laminations

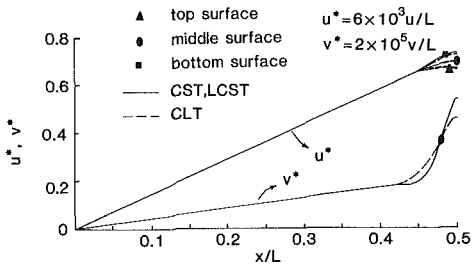


Fig. 3 Variation of u_0 and v_0 displacements along the axial direction of the cylindrical shell for -45 deg/45 deg lamination

The other geometric and material parameters are given in the nondimensional forms as follows:

$$L/R=20;\ R/t=5 \text{ so that } L/(Rt)^{1/2}=44.72$$

$$E_1/E_2=40;\ G_{12}(=G_{13}=G_{23})/E_2=0.5;$$

$$\nu_{12}(=\nu_{13}=\nu_{23})=0.25.$$

In the previous equation t is the wall thickness. E_2 is the surface-parallel Young's modulus transverse to the fiber direction. G_{12}, G_{13}, and G_{23} denote surface-parallel and transverse shear moduli, respectively, while ν_{12} ν_{13}, and ν_{23} are surface-parallel and transverse major Poisson's ratios, respectively. The material properties assumed here are the same as those of Spilker et al. (1976).

Although the procedure is applicable to an arbitrary choice of the admissible boundary conditions, as has been mentioned earlier, space limitation forces the present study to limit itself to only one type. The cylindrical shell is assumed to be simply-supported with SSI type (Chaudhuri, et al., 1986) boundary conditions, which are given by

$$w(\pm L/2)=M_x(\pm L/2)=N_x(\pm L/2)$$
$$=N_{x\theta}(\pm L/2)=\phi_\theta(\pm L/2)=0. \qquad (20)$$

Figure 2 shows the variation of displacement, w, along the axial direction of the cylinder for -15 deg/15 deg, -45 deg/45 deg, and -75 deg/75 deg laminations. These plots compare the present CST solution with the CLT-based analytical and the LCST-based finite element solutions. It may be noted that the three solutions (CLT, CST, LCST) have the same displacement, w, in the central region, because of the predominance of membrane action in this region. However, the plots show disagreement among the three solutions in the edge region. The reason behind this is that the bending action predominates in this region, which brings into action the different transverse shear deformation effects approximated by zero, constant, and layerwise constant shear deformation theories. It may be noted that the difference between the CST and LCST solutions (Fig. 2), for the same thickness, first increases as $\bar{\theta}$ increases and then decreases until the two solutions become identical for $\bar{\theta}=90$ deg (Abu-Arja, 1986). Figure 3 shows the variation of u_0 and v_0 displacements along the axial direction of the cylinder for -45 deg/45 deg lamination. These plots show that the three solutions are the same in the membrane region, while they disagree in the edge region, due to transverse shear effects they approximate. The axial variation of longitudinal stresses, σ_x, for -45 deg/45 deg lamination, is exhibited in Fig. 4. These plots show that σ_x is negligible in the central region, where the three solutions are almost equal. However, σ_x becomes significant in the edge region, where the aforementioned disagreements among the three solutions, are also observed. Figure 5 shows the axial variation of σ_θ for the three solutions with -45 deg/45 deg lamination. It may be noted that the behavior of σ_θ here is different from the behavior of σ_θ for the cross-ply laminates (Abu-Arja and Chaudhuri, in review). While σ_θ behaves like w for cross-ply laminates, σ_θ behaves differently for -45 deg/45 deg because of bending-twisting coupling effect.

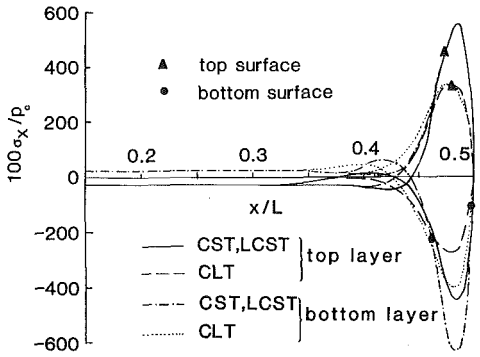


Fig. 4 Axial variation of longitudinal stresses, σ_x, for -45 deg/45 deg lamination

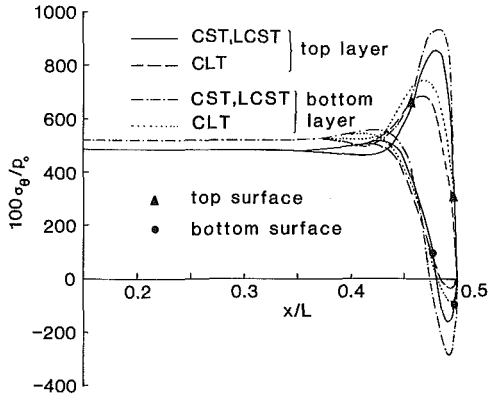


Fig. 5 Axial variation of σ_θ for -45 deg/45 deg lamination

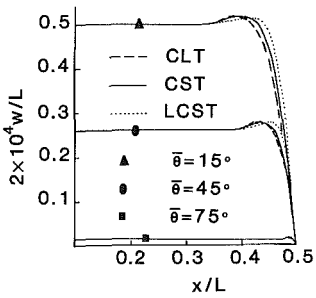


Fig. 6 Axial variation of displacement, w, of three-layered ($\bar{\theta}/-\bar{\theta}/\bar{\theta}$) cylindrical shells for $\bar{\theta}=15$ deg, 45 deg, and 75 deg laminations

The second example is a three-layer, unbalanced symmetric cylindrical shell with fiber orientations of the layers, $\bar{\theta}/-\bar{\theta}/\bar{\theta}$. The layers are of equal thickness. The length, radius, and thickness of the shell, and the material properties of a layer and the boundary conditions are identical to those of the two-

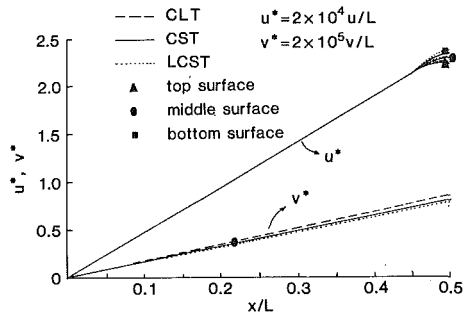


Fig. 7 Variation of u_0 and v_0 along the axial direction of the cylindrical shell for 45 deg/−45 deg/45 deg lamination

layer balanced unsymmetric shell. Figure 6 shows the axial variation of displacement, w, of the cylinder for $\bar{\theta} = 15$ deg, 45 deg, and 75 deg laminations. These plots compare the present CST solution with CLT and LCST solutions. As has been seen earlier, the three solutions have the same displacement, w, in the central region, while they disagree in the boundary region. The reasoning is the same as stated in the preceding paragraph. It is interesting to observe that the CST solution is closer to its CLT counterpart for symmetric lamination (Fig. 6) than for antisymmetric lamination (Fig. 2). The reason behind this is that the shear angle changes drastically from layer to layer and that it changes twice for the symmetric lamination while it changes once for the antisymmetric lamination, for the laminates studied in this investigation. This variation can be accounted for by the LCST, while the CST considers only constant shear deformation through the entire thickness. Figure 7 shows the variation of u_0 and v_0 along the axial direction of the cylinder for the three solutions, with 45 deg/−45 deg/45 deg lamination. These plots are similar to their counterparts for the anitsymmetric shell. Axial variation of the stresses, σ_x and σ_θ, are exhibited in Figs. 8 and 9, respectively. It is interesting to observe that the CST yields stresses, which are identical to those given by the CLT, while considerable disagreement between the CST-based analytical solution and the LCST-based finite element solution is observed. This has been observed in the case of unbalanced symmetric angle-ply plates by Chaudhuri and Seide (1987, to appear). This behavior is unlike the case of balanced unsymmetric shell of identical thickness, where the CST solution is almost identical to its LCST counterpart, while considerable disagreement exists between the CST and the CLT. The reason behind this difference of behavior of the two types of laminations is the aforementioned change of shear angle from layer to layer and insensitivity of the CST to that change.

6 Summary and Conclusions

Heretofore unavailable closed-form solutions are presented for unbalanced symmetric and balanced unsymmetric angle-ply cylindrical shells, subjected to axially-varying (axisymmetric) internal pressure loading, under the framework of FSDT or CST, for arbitrary boundary conditions. A constant term c_0, introduced in this paper, assumes the values of -1, 0, 1/2, and 1 for extension of classical theories, due to Donnell, Love-Reissner, Love-Sanders, and Love-Timoshenko, respectively, to the CST. The currently available CLT-based solutions can be easily obtained as a limiting case of the present solution with $G_{13} = G_{23} \to \infty$. Furthermore, it is interesting to observe that the decoupled O.D.E. in terms of w (resulting from the five coupled O.D.E.'s after successive elimination of u_0, v_0, Q_θ, and Q_x) for the case of uniform internal pressure, p_0, becomes a degenerate case of the same due to axially-varying pressure, $p(x)$.

Numerical results using the CST have been presented for two-layer unbalanced symmetric and three-layer balanced unsymmetric angle-ply cylindrical shells. Comparison of the

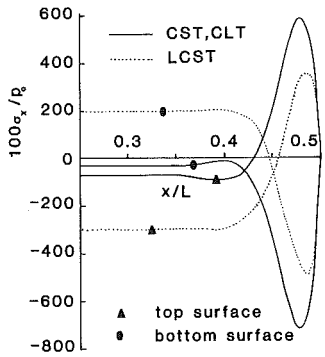


Fig. 8 Axial variation of the stress, σ_x, for 45 deg/−45 deg/45 deg lamination

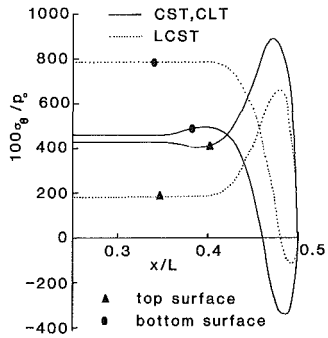


Fig. 9 Axial variation of the stress, σ_θ, for 45 deg/−45 deg/45 deg lamination

results with the CLT-based analytical and LCST-based finite element solutions suggests that, while they agree in the central (membrane) region, as expected, some disagreement among them has been observed in the edge (bending) region, due to the varying degrees of shear deformation effects approximated by the different theories inherent in these solutions. Furthermore, comparison of the stresses for the two types of angle-ply shells of identical thickness indicate different types of shear deformation behavior in these two shells. For the thickness and fiber orientations considered in the present investigation, the present CST-based solution in the case of the balanced unsymmetric shell is almost identical to its LCST counterpart while exhibiting disagreement with the corresponding CLT-based solution, whereas in the case of the unbalanced symmetric shell investigated, the CST and CLT solutions are identical, while considerable disagreement is observed with the corresponding LCST solution. This suggests that while the former case exhibits nearly constant transverse shear deformation through the laminate thickness, significant layerwise variation of this shear deformation occurs in the latter case, to which the CST is totally insensitive and the LCST is one of the viable alternatives to the three-dimensional elasticity theory. Solutions presented herein should serve as bench mark results for future comparisons.

References

Abu-Arja, K. R., 1986, "Static Analysis of Fiber Reinforced Plates and Shells with Shear Deformation, Ph.D. Dissertation, Department of Civil Engineering, University of Utah, Salt Lake City.

Abu-Arja, K. R., and Chaudhuri, R. A., "Influence of Transverse Shear Deformation on the Scaling of Moderately Thick, Cross-Ply Cylindrical Shells," *Journal of Composite Materials*, in press.

Balaraman, K., Kunukkasseril, V. X., and Chaudhuri, R. A., 1972, "Bending of Asymmetrically Laminated Anisotropic Shells Subjected to Internal Pressure," presented at the First Conference of Reinforced Plastics and Their Aerospace Applications, Vikram Sarabhai Space Center, ISRO, Trivandrum, India.

Bert, C. W., 1974, "Analysis of Shells," *Structural Design and Analysis*, Part 1, C. C. Chamis, ed., Vol. 7 of *Composite Materials*, L. J. Broutman and R. H. Krock, eds., John Wiley and Sons, Inc., New York, pp. 207–258.

Bert, C. W., and Francis, P. H., 1974, "Composite Material Mechanics: Structural Mechanics," *AIAA Journal*, Vol. 13, No. 9, pp. 1173–1186.

Bert, C. W., and Reddy, V. S., 1982, "Cylindrical Shells of Bimodulus Material," *Journal of the Engineering Mechanics Division*, ASCE, New York, Vol. 8, No. EM5, pp. 675–688.

Calcote, L. R., 1969, *The Analysis of Laminated Composite Structures*, Van Nostrand Reinhold, New York.

Chaudhuri, R. A., 1974, "Structural Behavior of FRP Rectangular Plates and Cylindrical Shells," M. S. Thesis, Dept. of Aerospace Engineering, I.I.T., Madras, India.

Chaudhuri, R. A., 1983, "Static Analysis of Fiber Reinforced Laminated Plates and Shells with Shear Deformation Using Quadratic Triangular Elements," Ph.D. Dissertation, Department of Civil Engineering, University of Southern California, Los Angeles.

Chaudhuri, R. A., Balaraman, K., and Kunukkasseril, V., 1986, "Arbitrary Laminated Anisotropic Cylindrical Shells Under Internal Pressure," *AIAA Journal*, Vol. 24, No. 11, pp. 1851–1857.

Chaudhuri, R. A., and Seide, P., 1987, "Triangular Finite Element for Analysis of Thick Laminated Plates," *International Journal for Numerical Methods in Engineering*, Vol. 24, No. 6, pp. 1203–1224.

Cheng, S., and Ho, B., 1963, "Stability of Heterogeneous Aelotropic Cylindrical Shell Under Combined Loading," *AIAA Journal*, Vol. 1, No. 4, pp. 892–898.

Dong, S. B., Pister, K. S., and Taylor, R. L., 1962, "On the Theory of Laminated Anisotropic Shells and Plates," *Journal of Aerospace Sciences*, Vol. 29, pp. 969–975.

Hutchinson, J. R., and El-Azhari, S. A., 1986, "Vibrations of Free Hollow Circular Cylinders," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 641–646.

Korn, G. A., and Korn, T. M., 1968, *Mathematical Handbook for Scientists and Engineers*, 2nd ed. McGraw-Hill Book Co., New York.

Kraus, H., 1967, *Thin Elastic Shells*, John Wiley and Sons, Inc., New York.

Reddy, J. N., 1984, "Exact Solutions of Moderately Thick Laminated Shells," *Journal Engineering Mechanics Division*, ASCE, New York, Vol. 108, pp. 794–809.

Reissner, E., and Stavsky, Y., 1961, "Bending and Stretching of Certain Types of Heterogeneous Aelotropic Plates," ASME JOURNAL OF APPLIED MECHANICS, Vol. 28, pp. 402–408.

Reuter, R. C., 1972, "Analysis of Shells Under Internal Pressure," *Journal of Composite Materials*, Vol. 6, pp. 94–113.

Seide, P., and Chaudhuri, R. A., 1987, "Triangular Finite Element for Analysis of Thick Laminated Shells," *International Journal for Numerical Methods in Engineering*, Vol. 24, No. 8, pp. 1563–1579.

Spilker, R. L., et al., 1976, "Use of Hybrid-Stress Finite Element Model for Static and Dynamic Analysis of Multi-Layer Composite Plates and Shells," DEPT AMMRC CTR 76:29, ASRL TR 181-2, M.I.T., Cambridge, Mass.

Timoshenko, S. P., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, 2nd ed., McGraw-Hill Book Co., New York.

Wilson, J. F., and Orgill, G., 1986, "Linear Analysis of Uniformly Stressed, Orthotropic Cylindrical Shells," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, pp. 249–256.

APPENDIX

Definitions of Certain Constants Referred to in the Test:

$$A_1 = e_4e_9e_5 - e_1e_8e_9; \quad A_2 = e_1e_6e_{10} - e_8e_2$$

$$+ e_3e_9e_5 + e_4e_6 - e_2e_5e_{10} - e_1e_7e_9 - e_8e_9;$$

$$A_3 = -e_1e_{10} + e_6e_{10} - e_7e_2 - e_4 + e_3e_6 - e_7e_9;$$

$$A_4 = -e_{10} - e_3;$$

$$A_5 = (e_2e_5c_5 + e_1e_9c_1 - e_1e_6c_5 - e_5e_9b_2)/b_7; \quad A_6 = (e_1c_1$$

$$+ e_2c_1 + e_1e_5 - e_6c_5 - e_6b_2)/b_7; \quad A_7 = (c_5 + b_2)/b_7$$

where

$$e_1 = b_4 - b_2b_8/b_7; \quad e_2 = b_3 + b_2/b_7; \quad e_3 = b_1 - b_2b_6/b_7;$$

$$e_4 = b_5; \quad e_5 = -B_{16}b_8/b_7; \quad e_6 = c_3 + B_{16}/b_7;$$

$$e_7 = -B_{16}b_6/b_7; \quad e_8 = c_2; \quad e_9 = c_5/b_7; \quad e_{10} = c_4 - b_6c_5/b_7$$

with

$$b_1 = -c_0B_{16}A_{12}/(A_{11}R) + c_0B_{20}/R; \quad b_2 = RA_{66}$$

$$- c_0(c_0 + 1)[B_{16}^2/(RA_{11}) - D_{66}/R];$$

$$b_3 = RB_{16}/A_{55}; \quad b_4 = c_0D_{66}/A_{44} - c_0B_{16}^2/(A_{11}A_{44});$$

$$b_5 = -RB_{16}; \quad b_6 = A_{22}/R^2 - A_{12}^2/(R^2A_{11});$$

$$b_7 = (c_0 + 1)[B_{26}/R^2 - A_{12}B_{16}/(R^2A_{11})];$$

$$b_8 = B_{26}/RA_{44} - A_{12}B_{16}/(RA_{11}A_{44});$$

$$b_9 = A_{12}/R; \quad c_1 = B_{16}; \quad c_2 = -D_{11}; \quad c_3 = D_{11}/A_{55};$$

$$c_4 = b_8A_{44};$$

$$c_5 = (c_0 + 1)[D_{66}/R - B_{16}^2/(RA_{11})]; \quad c_6 = b_4/c_0 \tag{A 1}$$

$$A_2^* = D_{11} - [A_{11}B_{16}^2]/[A_{11}(A_{66} + 4D_{66}/R^2) - 4B_{16}^2/R^2];$$

$$A_3^* = [4A_{11}B_{16}B_{26} - 4A_{12}B_{16}^2]/[A_{11}(A_{66}R^2 + 4D_{66})$$

$$- 4B_{16}^2/R^2]$$

$$A_4^* = A_{22}/R^2 - [A_{12}^2(R^2A_{66} + 4D_{66}) - 8A_{12}B_{16}B_{26}$$

$$+ 4A_{11}B_{26}^2]/[A_{11}(A_{66}R^4 + 4D_{66}R^2) - 4B_{16}^2R^2]$$

$$A_7^* = 1.0 \tag{A 2}$$

$$H_1 = -a_1^2/3 + a_2; \quad H_2 = 2[a_1/3]^3 - a_1a_2/3 + a_3;$$

$$F = [H_1/3]^3 + [H_2/2]^2;$$

$$H_3 = [-H_2/2 + \sqrt{F}\,]^{1/3}; \quad H_4 = [-H_2/2 - \sqrt{F}\,]^{1/3};$$

$$m_1 = H_3 + H_4 - a_1/3;$$

$$m_{2,3} = [-(H_3 + H_4)/2 - a_1/3] \pm i[(H_3 - H_4)/2\sqrt{3}] = r \pm is;$$

$$\alpha_1 = [(-r + \sqrt{r^2 + s^2})/2]^{1/2}; \quad \beta_1 = [(r + \sqrt{r^2 + s^2})/2]^{1/2};$$

$$\gamma_1 = \sqrt{|m_1|} \tag{A 3}$$

$$\mu_1 = 2\sqrt{-H_1/3}\,\cos(\alpha_2/3);$$

$$\mu_{2,3} = -2\sqrt{-H_1/3}\,\cos(\alpha_2/3 \pm 60 \text{ deg})$$

where

$$\cos(\alpha_2) = H_2/(2\sqrt{-(H_1/3)^3}) \tag{A 4}$$

$$h_5 = -h_2g_5/h_8; \quad h_6 = -[h_2(g_5h_1 + g_6) + g_7 + h_3g_5]h_8;$$

$$h_7 = -[h_3(g_5h_1 + g_6) + g_8]/h_8;$$

$$h_8 = h_1(g_5h_1 + g_6) - 1; \quad j_1 = e_9h_1h_5; \quad j_2 = e_9h_1h_6 + e_9h_2;$$

$$j_3 = e_9h_1h_7 + e_9h_3 + e_{10};$$

$$j_4 = h_5/b_7 = b_8j_1/b_7; \quad j_5 = h_6/b_7 = b_8j_2/b_7;$$

$$j_6 = h_7/b_7 = b_8j_3/b_7; \quad j_7 = -b_6/b_7$$

$$j_8 = p_0/b_7; \quad j_9 = -b_9/b_7; \quad k_1 = k_7j_4 + k_8j_1;$$

$$k_2 = k_7j_5 + k_8j_2;$$

$$k_3 = k_7j_6 + k_8j_3; \quad k_4 = k_7j_7 - A_{12}/(A_{11}R);$$

$$k_5 = k_7j_8; \quad k_6 = k_7j_9 + 1;$$

where

$$g_1 = e_1e_9; \quad g_2 = e_2 + e_9; \quad g_3 = e_1e_{10} + e_4;$$

$$g_4 = e_3 + e_{10}; \quad g_5 = e_5e_9; \quad g_6 = e_6;$$

$$g_7 = e_5e_{10} + e_8; \quad g_8 = e_7; \quad h_1 = 1/h_4;$$

$$h_2 = (-g_7g_1 + g_5g_3)/(h_4g_1);$$

$$h_3 = (-g_8g_1 + g_5g_4)/(h_4g_1; \quad h_4 = g_6 - g_5g_2/g_1;$$

$$k_7 = -(c_0 + 1)B_{16}/(RA_{11}); \quad k_8 = -B_{16}/(A_{11}A_{44}). \tag{A 5}$$

S. F. Masri
Mem. ASME

R. K. Miller
Mem. ASME

T. J. Dehghanyar

Department of Civil Engineering,
University of Southern California,
Los Angeles, Calif. 90089-0242

T. K. Caughey

Division of Engineering and Applied Science,
California Institute of Technology,
Pasadena, Calif. 91106
Mem. ASME

Active Parameter Control of Nonlinear Vibrating Structures

A simple, yet efficient method is presented for the on-line vibration control of nonlinear, multidegree-of-freedom systems responding to arbitrary dynamic environments. The procedure uses nonlinear auxiliary mass dampers with adjustable motion-limiting stops located at selected positions throughout a given nonlinear system. A mathematical model of the system to be controlled is not needed for implementing the control algorithm. The degree of the primary structure oscillation near each vibration damper determines the damper's actively-controlled gap size and activation time. By using control energy to adjust the damper parameters instead of directly attenuating the motion of the primary system, a significant improvement is achieved in the total amount of energy expended to accomplish a given level of vibration control. In a related paper, the direct method of Lyapunov is used to establish that the response of the controlled nonlinear primary structure is Lagrange stable. Numerical simulation studies of several example systems, as well as an experimental study with a mechanical model, demonstrate the feasibility, reliability, and robustness of the proposed semi-active control method.

1 Introduction

Analytical and experimental studies have shown that a class of nonlinear auxiliary mass dampers, known as impact dampers, may be more efficient than the conventional (linear) dynamic vibration neutralizers in attenuating the response of oscillating structures subjected to nonstationary, wide-band random excitation. The main factor for the effectiveness of properly designed impact dampers in limiting the vibrations of structures emanating from arbitrary dynamic environments is that the relatively small damping forces generated by the impacting (auxiliary) mass introduces chaos in the primary system response by disorganizing the orderly process of amplitude buildup, thus significantly reducing the structural response.

However, as in any passive device, even when the characteristics of a particular damper have been optimized for a given operating condition, its vibration damping efficiency is limited in handling wide-band excitations due to the inability of continuously adapting its governing characteristics to the evolving environment. This limitation of passive dampers is particularly pertinent in applications where not only the rms level of the response but also the peak levels of the primary structure response are of concern, as is the case in most structual applications.

Motivated by the above discussion, the authors have developed and implemented two on-line active control algorithms (henceforth referred to as Methods 1 and 2) that utilize pulse generators to emulate the action of optimally designed impact dampers, to suppress the vibrations of linear as well as nonlinear multidegree-of-freedom (MDOF) flexible structures responding to arbitrary dynamic environments (Masri et al., 1981, 1982).

The aforementioned on-line control procedures have been shown to be quite effective in greatly reducing the rms response as well as the peak response of vibrating structures even when the excitation is nonstationary wide-band random. This significant improvement in efficiency is achieved because the active control algorithms under discussion are designed to maximize the influence of the control actuators either by (1) optimizing the relative magnitude of the control pulses (Method 1) or by (2) choosing the optimum time for applying the control forces (Method 2). Both methods assume the availability of an external energy source to produce the control pulses on demand.

Since in many practical cases the amount of energy available for control purposes is limited, the present study explores an alternate pulse-control strategy that economizes the use of control energy. This is accomplished by devising an on-line control procedure that attempts to optimize the parameters of incorporated impact vibration dampers attached to different locations within the vibrating flexible structure. Instead of using mass-ejection techniques (or equivalent methods) to directly furnish the needed control forces, an internal mechanism of momentum transfer between the primary structure and the auxiliary masses is employed. It is shown that the trade off between vibration damping efficiency and control energy economy does not lead to a major deterioration in the overall vibration reduction of the primary system as compared to what can be achieved with fully active pulse-control methods.

Contributed by the Applied Mechanics Divison of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS 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 26, 1988; final revision, September 21, 1988.

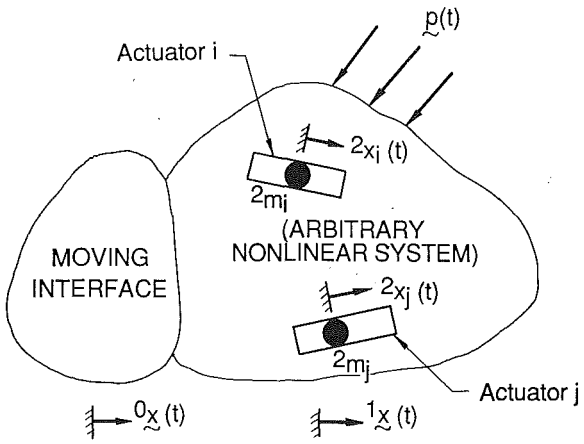


Fig. 1 Model of arbitrary, nonlinear MDOF system under directly applied dynamic loads $p(t)$ and/or interface motions ${}^0x(t)$ that is provided with a number of active vibration controllers. Vectors ${}^1x(t)$ and ${}^2x(t)$ define the absolute displacements of the primary and secondary system, respectively. The mass of actuator number i is designated by ${}^2m_i$.

Section 2 of this paper formulates the problem and presents the semi-active control algorithm, Section 3 presents numerical simulation results to demonstrate the effectiveness of the proposed control strategy under a wide variety of situations, and Section 4 presents some experimental results with a mechanical model to demonstrate the feasibility of the procedure under laboratory conditions. A stability analysis of an idealized version of the semi-active device is available in the work of Karyeaclis and Caughey (1987).

2 Formulation

Consider the arbitrary, nonlinear multidegree-of-freedom system shown in Fig. 1 under directly applied dynamic loads $\mathbf{p}(t) = (p_1, p_2, \ldots, p_{n_1})^T$ and/or support motions ${}^0\mathbf{x}(t) = ({}^0x_1, {}^0x_2, \ldots, {}^0x_{n_0})^T$. The equation of motion for such a system can be expressed as

$$M_{11}{}^1\ddot{\mathbf{x}} + C_{11}{}^1\dot{\mathbf{x}} + K_{11}{}^1\mathbf{x} + M_{10}{}^0\ddot{\mathbf{x}} + C_{10}{}^0\dot{\mathbf{x}} + K_{10}{}^0\mathbf{x} + \mathbf{f}_N(\mathbf{x}, \dot{\mathbf{x}}, t) = \mathbf{p}(t) \quad (1)$$

where

- ${}^1\mathbf{x}(t)$ = $({}^1x_1, {}^1x_2, \ldots, {}^1x_{n_1})^T$ is the system displacement vector,
- ${}^0\mathbf{x}(t)$ = $({}^0x_1, {}^0x_2, \ldots, {}^0x_{n_1})^T$ is the support displacement vector,
- M_{11}, C_{11}, K_{11} = matrices, possibly function of time, each of order $(n_1 \times n_1)$, that charaterize the inertia, damping and stiffness forces associated with the n_1 system degrees-of-freedom,
- M_{10}, C_{10}, K_{10} = matrices, possibly function of time, each of order $(n_1 \times n_0)$, that characterize the inertia, damping, and stiffness forces associated with the interface motions,
- $\mathbf{f}_N(\mathbf{x},\dot{\mathbf{x}},t)$ = an n_1 column vector of nonlinear, nonconservative forces involving ${}^1\mathbf{x}(t)$ as well as ${}^0\mathbf{x}(t)$, and
- $\mathbf{p}(t)$ = an n_1 column vector of directly applied forces.

Assuming, without any loss of generality, that the system mass matrix is diagonal allows equation (1) to be expressed in the form

$${}^1m_i{}^1\ddot{x}_i + {}^1f_i({}^1\mathbf{x}, {}^0\mathbf{x}, {}^1\dot{\mathbf{x}}, {}^0\dot{\mathbf{x}}) = p_i(t); \quad i = 1, 2, \ldots, n_1, \quad (2)$$

where

- ${}^1m_i$ is the mass associated with the system degree-of-freedom i, and
- ${}^1f_i$ is the "restoring force" associated with system DOF i arising from passive interactions.

If the nonlinear system under consideration is now provided with a number (n_2) of active vibration dampers distributed throughout the vibrating structure, then the n_2 equations of motion of the involved DOFs will change from the form of equation (2) to

$${}^1m_j{}^1\ddot{x}_j + {}^1f_j({}^1\mathbf{x}, {}^0\mathbf{x}, {}^1\dot{\mathbf{x}}, {}^0\dot{\mathbf{x}}) - {}^2f_j(z_j, \dot{z}_j) = p_j(t);$$

$$j = 1,2, \ldots, n_2. \quad (3)$$

Additionally, the passive system's n_1 equations of motion will have to be augmented by n_2 equations that govern the motion of the active vibration dampers:

$${}^2m_j{}^2\ddot{x}_j + {}^2f_j(z_j,\dot{z}_j) = p_j(t); \quad j = 1,2, \ldots, n_2, \quad (4)$$

where

- ${}^2\mathbf{x}(t)$ = $({}^2x_1, {}^2x_2, \ldots, {}^2x_{n_1})^T$ = auxiliary masses displacement vector,
- $z_j(t)$ = ${}^2x_j(t) - {}^1x_j(t)$ is the displacement of damper ${}^2m_j$ relative to ${}^1m_j$,
- ${}^2m_j$ = mass of auxiliary damper j, and
- ${}^2f_j$ = interaction forces arising from the presence of damper j.

It is seen from equations (1) through (4) that the following convention is followed in the choice of notation: left superscript (0) pertains to the n_0 interface (support) DOFs, left superscript (1) pertains to the n_1 passive system DOFs, and left superscript (2) pertains to the n_2 auxiliary mass dampers' DOFs.

Consider now the class of nonlinear auxiliary dampers that resemble dynamic vibration neutralizers (DVN) with resilient motion-limiting stops. The performance of such devices under a variety of excitations is available in the work of Masri (1972). The influence of this class of devices on the primary system to which they are attached can be expressed as:

$${}^2f_j(z_j,\dot{z}_j) = {}^2m_j{}^2\ddot{q}_j(z_j,\dot{z}_j,{}^j\theta); \quad j = 1,2, \ldots, n_2, \quad (5)$$

where ${}^2q_j$, the normalized force associated with damper ${}^2m_j$, is given by

$${}^2q_j(z_j,\dot{z}_j,{}^j\theta) = g_j(z_j,d_j) + h_j(z_j,\dot{z}_j,d_j) + r_j(z_j,\dot{z}_j), \quad (6)$$

and the three terms appearing on the right-hand side of equation (6) are:

- $g_j(z_j, d_j)$ = nonlinear conservative force arising from the contact of damper ${}^2m_j$ with its constraining (limiting) stops of characteristic dimension d_j,
- $h_j(z_j,\dot{z}_j,d_j)$ = nonlinear nonconservative force arising from the contact of damper ${}^2m_j$ with its stops, and
- $r_j(z_j,\dot{z}_j)$ = nonlinear nonconservative forces arising from the coupling mechanism between damper ${}^2m_j$ and its attachment location ${}^1m_j$ when the motion-limiting stops are not engaged.

To help interpret the various force terms appearing in the general damper representation of equation (6), consider the following special cases.

2.1 Special Cases.

Case (1); Dynamic Vibration Neutralizer (DVN): This widely used linear damper (also known as the "vibration ab-

sorber'' or ''Frahm damper'') employs a linear elastic element and, quite often, a linear viscous damping element to couple the auxiliary mass to the oscillating structure. Thus, for this case

$$^2q_j(z_j,\dot{z}_j,{}^j\theta) = {}^j\theta_1 z_j + {}^j\theta_2 \dot{z}_j, \tag{7}$$

where

$^j\theta_1$ = stiffness coefficient of the coupling spring, and
$^j\theta_2$ = damping coefficient of the coupling dashpot.

Comparing equation (7) to the general form of equation (6) yields

$$g_j(.) = 0, \tag{8}$$

$$h_j(.) = 0, \text{ and} \tag{9}$$

$$r_j(.) = {}^j\theta_1 z_j + {}^j\theta_2 \dot{z}_j. \tag{10}$$

Case (2); Nonlinear Vibration Neutralizer: In this class of dampers, the coupling element between ^1m_j and ^2m_j has nonlinear characteristics involving the stiffness and/or damping terms. For example, when polynomial-like nonlinearities exist, if the spring has hardening stiffness and the damping forces are a quadratic function of the relative velocity, the three generic components of the damper force appearing in equation (6) become:

$$g_j(.) = 0, \tag{11}$$

$$h_j(.) = 0, \tag{12}$$

$$r_j(.) = {}^j\theta_1 z_j + {}^j\theta_2 \dot{z}_j + {}^j\theta_3 z_j^3 + {}^j\theta_4 \dot{z}_j^2 \tag{13}$$

Case (3); Impact Damper: In an ideal impact damper which is moving freely in a container with a stiff, resilient stops, the components of equation (6) assume the form

$$g_j(.) = {}^j\theta_1[z_j - \text{sgn}(z_j)\,{}^j\theta_3]u(|z_j| - {}^j\theta_3), \tag{14}$$

$$h_j(.) = {}^j\theta_2 \dot{z}_j u(|z_j| - {}^j\theta_3), \tag{15}$$

$$r_j(.) = 0, \tag{16}$$

where

$^j\theta_1$ = stiffness of the slightly resilient damper stops,
$^j\theta_2$ = equivalent viscous damping coefficient involved during impacts,
$^j\theta_3$ = impact damper clearance, equal to one half of the total gap size in the passive damper,
sgn(.) = indicates the algebraic sign of its argument, and
$u(.)$ = unit step function defined by:

$$u(a) = \begin{cases} 1 \text{ if } a > 0, \\ 0 \text{ if } a \le 0. \end{cases}$$

Notice that, in this case, no coupling exists between the colliding masses when the relative displacement of the auxiliary mass is less than the available gap; consequently, the coupling force $r_j(.)$ is zero.

Case (4); Nonlinear Vibration Neutralizer With Motion-Limiting Stops: This device combines features of the conventional DVN and the impact damper. In the terminology of equation (6), it is responsible for the following forces:

$$g_j(.) = {}^j\theta_1[z_j - \text{sgn}(z_j)\,{}^j\theta_3]u(|z_j| - {}^j\theta_3), \tag{17}$$

$$h_j(.) = {}^j\theta_2 \dot{z}_j u(|z_j| - {}^j\theta_3), \tag{18}$$

$$r_j(.) = {}^j\theta_4 z_j + {}^j\theta_5 \dot{z}_j, \tag{19}$$

where it is recognized that forces g_j and h_j are identical to the corresponding terms in Case (3), and force r_j has the same form as in Case (1). Notice that, here, parameters $^j\theta_1$, $^j\theta_2$ and $^j\theta_3$ govern the performance of the damper in its nonlinear range of motion (i.e., when the available gap is exceeded), while parameters $^j\theta_4$ and $^j\theta_5$ determine the behavior of the damper within its linear range.

2.2 Optimization Procedure. Consider again the nonlinear system whose oscillations are to be attenuated:

$$^1m_i\,{}^1\ddot{x}_i + {}^1f_i = p_i(t); \quad i = 1,2, \ldots, n_1 - n_2, \tag{20}$$

$$^1m_j\,{}^1\ddot{x}_j + {}^1\ddot{x}_j + {}^1f_j - {}^2f_j = p_j(t);$$

$$j = n_1 - n_2 + 1, \ldots, n_1, \tag{21}$$

$$^2m_k\,{}^2\ddot{x}_k + {}^2f_k = 0; \quad k = 1,2, \ldots, n_2. \tag{22}$$

Let $\mathbf{y}(t)$, an n_1 column vector, denote a measure of the primary system response of interest. For example, if the structural deformations with respect to a moving base are of concern, $\mathbf{y}(t)$ can be composed of a combination of the primary system relative displacements and velocities. On the other hand, if peak deformations are of interest, the entries in $\mathbf{y}$ can correspond to the maximum deformations of designated locations. Hence, the response of the dynamic system with dampers whose motion is governed by the $(n_1 + n_2)$ equations given in equation (20)–(22) can be expressed as:

$$\mathbf{y}(t) = \mathbf{y}({}^1\theta, {}^2\theta, \ldots, {}^{n_2}\theta). \tag{23}$$

Let the cost function to be minimized be

$$J({}^1\theta, {}^2\theta, \ldots, {}^{n_2}\theta) = \int_{t_0}^{t_0 + T_{\text{opt}}} \mathbf{y}^T(t)\, W \mathbf{y}(t)\, dt, \tag{24}$$

where W is an arbitrary weighting matrix.

In principle, the optimization task is now reduced to seeking the set of damper parameters which will minimize J over the response segment T_{opt}. When this optimization is performed once ''off-line'' for the whole response record, the result is an optimized set of *passive* damper parameters. However, as mentioned in the introduction, passive dampers, even when optimally designed for a particular situation, may have limited effectiveness when operating under wide-band excitations.

On the other hand, the continual optimization and adjustment of the damper parameters (fully active control) requires the ''on-line'' solution of equation (24) and the continuous feedback of the results to the control actuators. This approach, while mathematically appealing, is not feasible for a variety of reasons, the leading one of which is the demanding analytical and computational effort required to determine (let alone adjust) the optimum damper parameters in a small fraction of the structure time constant.

This study presents a compromise solution of the two control options discussed above: (1) passive dampers initially optimized off-line, and (2) fully-active optimized dampers with continuous feedback control. The alternate option of this paper is to trade degraded optimization (i.e., open-loop, suboptimal control) for ease of implementation in real life engineering situations with actual hardware.

The motivation behind the proposed control algorithm is the observed behavior of passive impact dampers configured in the form of dynamic vibration neutralizers with motion-limiting stops. When one such passive damper is attached to an oscillating primary system undergoing transient excitations, the auxiliary mass will sustain repetitive (possibly chaotic) impacts on different sides of its container. The number, location, and intensity of these irregular impacts is a highly nonlinear function of the system characteristics and the nature of the excitation. The ensuing plastic deformations, Coulomb friction, and momentum transfer between the two masses during collisions tend to reduce the vibrations of the primary system.

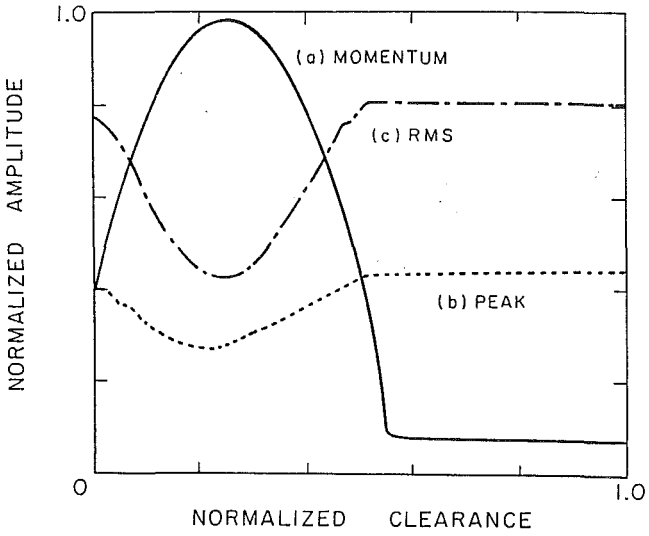


Fig. 2 Transient response of a linear SDOF system provided with a passive impact damper and subjected to swept-sine excitation. A representative segment of the primary system response between two consecutive impacts separated by a time period T_{opt} is considered. The plotted curves show the variation of the indicated quantity with the gap size d, all other parameters remaining the same. The primary system ratio of critical damping is 0.05. The impact damper mass ratio is 0.10 and its coefficient of restitution is $e \approx 0.75$. Notice that the time increment T_{opt} varies with the gap size. (a) Momentum transfer between the colliding masses at the end of the observation time segment from t_o to $t_o + T_{opt}$; (b) peak value of the primary system displacement; and (c) RMS value of the primary system displacement.

If the time of occurrence of one of these impacts is used to define a reference time t_0, then the variation of the peak and rms levels of the primary system response with the gap size that governs the time of occurrence of the succeeding impact will be as indicated in Fig. 2.

Since the predominant mechanism that governs the interaction between ${}^1m_i$ and ${}^2m_i$ is momentum transfer, it is reasonable to expect a strong dependence of the criterion function $J(.)$ on the discontinuity in the velocity of ${}^1\dot{x}_i$ and/or ${}^2\dot{x}_i$ during the impact process. This expectation is borne out by the results depicted in Fig. 2, where the value of the momentum transfer is superimposed on the graph of the constituents of $J(.)$.

It is thus clear that, at least for the example problem shown in Fig. 2, optimizing $J(.)$ is practically identical to seeking an extremum value of the momentum transfer involved in the impact process. For the class of problems under discussion, this condition is equivalent to having an impact occur when the primary system's velocity is at its peak value.

2.3 Semi-Active Control Algorithm. The preceding discussion established the guidelines for a procedure to optimize the operation of semi-active impact dampers configured as mentioned above. To maximize the efficiency of an impact damper between two consecutive impacts, the gap size d should be adjusted so that the following conditions are satisfied:

- For each damper mass ${}^2m_j$, an impact is made to occur when the velocity of the corresponding primary system mass ${}^1m_j$ has reached its peak value. This instant corresponds to the zero crossing of the corresponding primary system displacement.
- The velocities of the various set(s) of two colliding masses must be opposite to each other at the time of impact. This condition insures that the impact process(es) will stabilize the motion of the primary system.

On this basis, the following control algorithm for on-line implementation of the damping device(s) is proposed. The control strategy consists of detecting the displacement from the neutral position (absolute, or relative to a moving support)

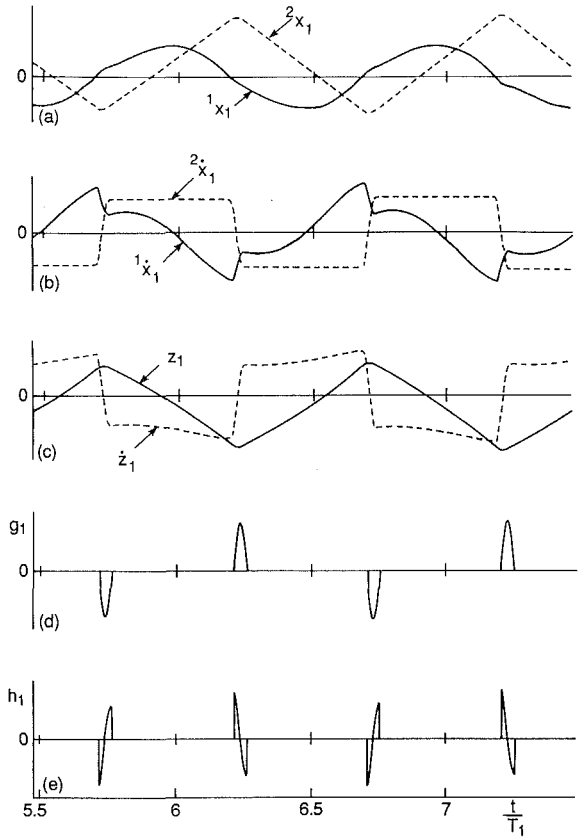


Fig. 3 Time history of a representative segment of the steady-state motion of a SDOF system, that is harmonically excited at resonance and provided with a semi-active impact damper having a mass ratio of 0.1. The time segment shown covers approximately two natural periods T_1 of the primary system. For clarity, the amplitude of all the plotted quantities have been normalized. (a) Absolute displacements of the primary and secondary systems; (b) absolute velocity of the primary and secondary systems; (c) relative displacement and velocity between the primary and secondary systems; (d) nonlinear stiffness force; and (e) nonlinear damping force.

zero crossings of the oscillating structure damper locations, and generating sets of impulsive control forces by inducing a collision between each of the auxiliary masses and their corresponding structure locations. The essential features of this approach can be summarized as follows:

- Virtually no on-line information regarding the global dynamic system characteristics is needed.
- Whether the primary system is linear or nonlinear has no bearing on the algorithm.
- Monitoring of only the system relative displacements at the dampers' locations is required.
- The on-line computation of the optimum clearance distances is reduced to a simple detection process.

To illustrate the application of this approach, a representative segment of the motion of a linear SDOF oscillator being controlled by such a semi-active damper is shown in Fig. 3. The primary system is harmonically excited at resonance. These graphs represent the absolute and relative state variables of the system and the nonlinear conservative and nonconservative control functions, g and h. The amplitude of all time histories in this figure have been normalized to lie between -1.0 and $+1.0$. The length of time segment shown is approximately two natural periods.

In Fig. 3(a), the solid line represents ${}^1x_1(t)$, the absolute displacement of the primary system ${}^1m_1$, while the dashed line represents ${}^2x_1$, the absolute displacement of the secondary system ${}^2m_1$. Similarly, in Fig. 3(b) the solid and dashed lines represent ${}^1\dot{x}_1(t)$, the absolute velocity of ${}^1m_1$ and ${}^2\dot{x}_1(t)$, the absolute velocity of ${}^2m_1$, respectively. The time histories of

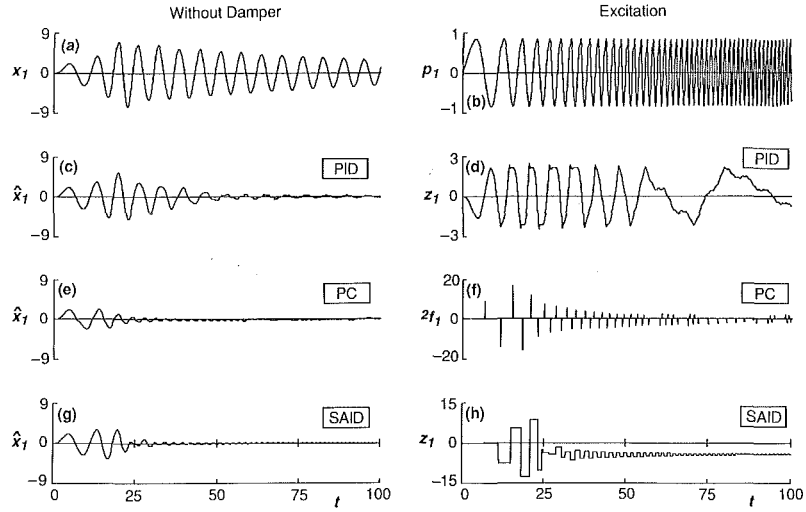


Fig. 4 Swept-sine excitation of a SDOF system with a variety of damping devices. (*a*) primary system response in the absence of any dampers; (*b*) excitation; (*c*) response with a passive impact damper; (*d*) relative displacement between the passive impact damper mass and the primary system; (*e*) response when using on-line pulse control; (*f*) pulse control forces; (*g*) response when using a semi-active impact damper; and (*h*) evolution of the gap size in the semi-active damper.

the relative displacement $z_1 = {}^2x_1 - {}^1x_1$ and relative velocity $\dot{z}_1 = {}^2\dot{x}_1 - {}^1\dot{x}_1$ are represented by the solid and dashed lines, respectively, in Fig. 3(*c*). The time history of the nonlinear stiffness force $g_1(t)$ generated by the contact of the oscillating mass ${}^2m_1$ with its resilient "stops" is shown in Fig. 3(*d*). Similar results for the nonlinear damping force $h_1(t)$ arising during the impact process is shown in Fig. 3(*e*). As seen from these results, suitable control impacts are applied twice every fundamental period of the system. The total control energy exerted on the structure during an impact is the sum of the areas under the *g* and *h* functions.

The only significant disadvantage of this technique is the lack of consideration for possible hardware delays in the activation of the impacting mechanism. Two possible provisions may be adopted to overcome this inadequacy:

- The first obvious choice is to design a high speed activation system with delays that are small when compared to the fundamental period of the structural system.
- The alternative solution is to anticipate the system response and thus activate the impacting mechanism when the displacement of the structure has crossed a certain prescribed threshold level.

3 Stability Analysis

A stability analysis of the device under discussion has been performed and is available in the work of Karyeaclis and Caughey (1987). Using Lyapunov's approach it is shown that, under fairly permissive conditions, the response of a system provided with the SAID under discussion is bounded.

4 Numerical Simulation

The efficiency of the proposed control strategy is demonstrated by presenting numerical simulation results for several SDOF and MDOF models with diverse characteristics, subjected to deterministic and stochastic dynamic environments.

Example (1): SDOF System Under Swept-Sine Excitation. The results shown in Fig. 4 correspond to a linear, viscously damped SDOF system consisting of a mass ${}^1m_1$ having a ratio of critical damping ${}^1\zeta_1 = 0.01$, and initially at rest, that is subjected to swept-sine excitation $F(t)$ given by

$$F(t) = F_0 \sin[\Omega(t)t]. \tag{25}$$

The time variation of the exciting frequency, Ω, is of the form

$$\Omega(t) = at + b. \tag{26}$$

If this linear system is subjected to a swept-sine excitation, shown in Fig. 4(*b*) of amplitude F_0 that varies according to equation (26) between the frequency limits $\Omega(0)/\omega_1 = 0.5$ and $\Omega(T_s)/\omega_1 = 1.5$ in sweep time $T_s/T_1 = 25$, the transient response shown in Fig. 4(*a*) is obtained.

Suppose now that the primary system under consideration is equipped with a conventional impact damper having an auxiliary mass ratio $\mu = 0.1$. Assume that the damper stops are relatively stiff and have impact plastic deformation characteristics equivalent to a coefficient of restitution $e \approx 0.8$ (within the range provided by hardened steel). Let the damper clearance ratio $d^* = [d/{}^1x_{1\max}]$ be optimized in accordance with the response characteristics of such nonlinear devices (Masri and Caughey, 1966) thus yielding an optimum clearance of $d^*_{\text{opt}} \simeq 2.0$.

The normalized response of the primary system with an optimized passive impact damper (PID) will then be as shown in Fig. 4(*c*). Notice that, in this case, the peak amplitude is attenuated by the factor $\simeq 0.7$ relative to the corresponding peak response in Fig. 4(*a*). It is seen in Fig. 4(*d*) that, due to the nature of the passive impact damper, the relative displacement between the colliding masses is constrained to remain within the fixed gap size of $\pm d/2$.

It is clear from the results shown in Fig. 4(*c*) that, while the optimized PID did attenuate the peak response to some extent, its efficiency was limited because it could not adapt to the transient nature of the primary system response. This problem can be easily remedied by using an active on-line pulse control procedure previously developed by the authors. When this control method is applied to the primary system under discussion, it results in the response shown in Fig. 4(*e*). The control forces that are used here are governed by the following rule ("active" viscous damping):

$$F_c(t) = \begin{cases} -{}^2c_1\,{}^1\dot{x}_1(t), & t_0 \le t \le (t_0 + T_d) \\ 0 & \text{otherwise} \end{cases}$$

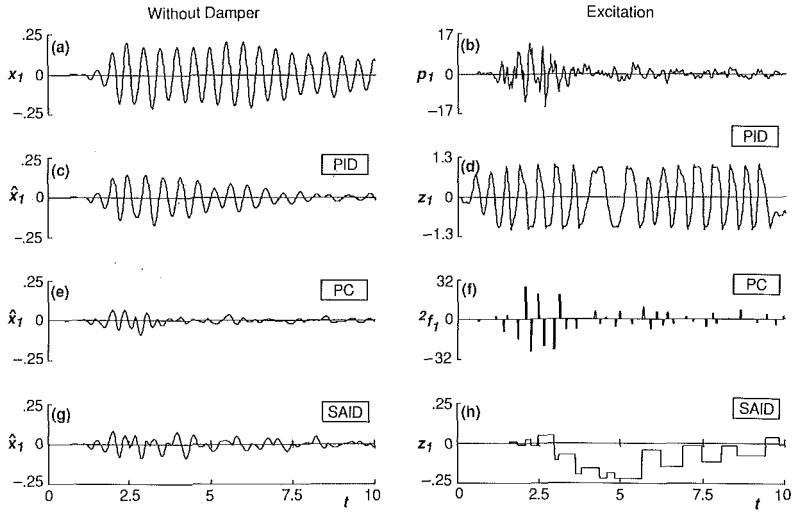


Fig. 5 Nonstationary random excitation of a SDOF system with the same variety of damping devices used in Fig. (4).

where t_0 is the pulse initiation time and T_d is the pulse duration.

The time history of the control forces generated by the actuators that are using an external energy source are shown in Fig. 4(*f*). Notice that, due to the nature of this control algorithm, the actuation time of the pulse control forces coincides with the primary system displacement zero crossings (equivalent to velocity peaks). In addition, the magnitude of the control force is changed once every one-half system period to maintain a value which is a constant factor of the primary system velocity.

The attenuation in peak amplitude with the on-line pulse control is $[^1\hat{x}_{1\max}/^1x_{1\max}] \simeq 0.3$, which is substantially better than what was achieved with the optimum passive dampers discussed above. Obviously, the cost of this added efficiency is the need to furnish an external energy source for the expenditure of the control energy.

If a semi-active impact damper is now attached to the primary system under discussion, the response shown in Fig. 4(*g*) is obtained. The auxiliary mass has the same ratio ($\mu = 0.10$) used by the passive impact damper discussed previously, and the damper stops have the same coefficient of restitution as for the PID. The evolution of the adjustable stops is shown in Fig. 4(*h*).

As might be expected, the efficiency of the SAID (a peak reduction factor in the ratio of $\simeq 0.4$) is better than what was achieved by the optimized passive impact damper, but not as good as the active pulse-control procedure. A clear visual explanation for the improved damping efficiency of the SAID is furnished by Fig. 4(*h*) where it is seen that the envelope of the optimum gap size, was being adapted to closely match that of the primary system response. Furthermore, the time-varying gap size caused the collisions between the oscillating masses to occur at a time when the interaction force $^2f_1(.)$ components had the most beneficial effect (as regarding motion attenuation) on the primary system. Thus, $^2f_1$, the combined force due to g_1 and h_1, is seen to play the same role, and to have the same qualitative features, as the active control force $F_c(t)$ shown in Fig. 4(*f*).

Example (2): SDOF Under Nonstationary Random Excitation. This case is similar to the one in Example (1) except that the disturbance is a wide-band nonstationary random excitation. The identical damper parameters of Example (1) are used again. The performance characteristics of the various dampers are shown in Fig. 5. The relative efficiency of various damping devices matches the results under swept-sine excitation shown in Fig. 4.

The time history of the adaptive gap size is shown in Fig. 5(*h*). The lack of any discernible pattern in the evolution of the optimum gap size reflects the nature of the random disturbance. The complex changes of $z_1(t)$ between impacts clearly illustrate the handicaps passive dampers have to cope with, since their initial (fixed) gap size cannot change in time to accommodate quiescent or active episodes of the random response.

Example (3): Linear MDOF System Under Nonstationary Random Excitation. Consider a MDOF linear frame structure that is subjected to wide-band random interface motion and without any directly applied loads. The chain-like nature of this example is in no way a requirement of the control algorithm under discussion; it is merely a convenient choice so as to make the system resemble, for example, a building-like structure undergoing earthquake ground motion.

The response of this structure under a simulated earthquake, operating without any auxiliary mass dampers, is shown in the LHS column of plots in Fig. 6. If a SAID of mass ratio $\mu_i = 0.05$, $i = 1, 2, 3$ is attached to each of the three "stories," then the controlled response would be as shown in the middle column of plots in Fig. 6, and the corresponding variable gap sizes are shown in the RHS column of plots.

Variable $y_i(t)$ represents the displacement of the ith level in the structure with respect to the oscillating base. $y_1(t)$ is closest to the base and $y_3(t)$ is the farthest away. For clarity, different scales are used for the ordinates of the plots corresponding to y_1, y_2, and y_3 in Figs. 6(*a*), (*b*), and (*c*). However, the middle column of plots uses the same amplitude scales as the corresponding uncontrolled cases. Comparison of the controlled and uncontrolled responses of various locations indicates that nearly the same percentage vibration attenuation is achieved at each of the controlled locations.

The influence of the SAID location on the efficiency of the device is demonstrated in Fig. 7, where a single SAID is attached to different locations in the MDOF system under discussion. The LHS column of plots in Fig. 7 shows schematic diagrams of the attachment points of the SAID, the middle column of plots presents the controlled response of the top mass m_3 (not necessarily the location of the SAID), and the RHS column of plots gives the evolution if the optimum gap size for each of the three tests.

The plots in Fig. 7 show the effects of the SAID locations on the attenuation of the relative displacement of the top floor. Notice that evolution of the damper clearance is clearly dependent on the local oscillations in the vicinity of the damper. The same relative reduction in the response is attained for the loca-

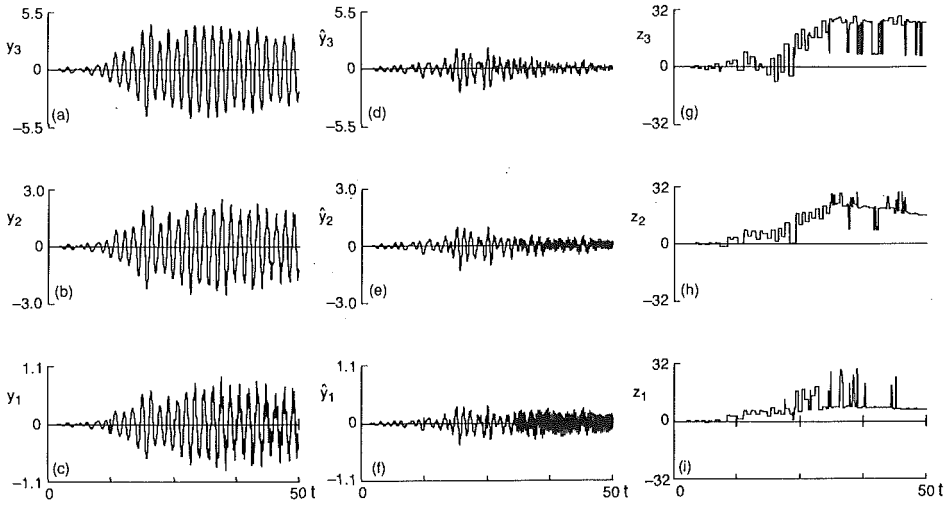


Fig. 6 Response of a linear frame structure, resembling a 3-story building, under nonstationary base excitation. The left-hand side, column of plots represent the transient response without augmented damping, the middle column shows the corresponding response (plotted to the same scale) when a separate SAID is attached to each level in the structure, and the right-hand column of plots shows the evolution of the dampers stops. Variable $y_i(t)$ represents the displacement of the ith level in the structure with respect to the oscillating base. The mass of each damper is 5 percent of the corresponding location mass.

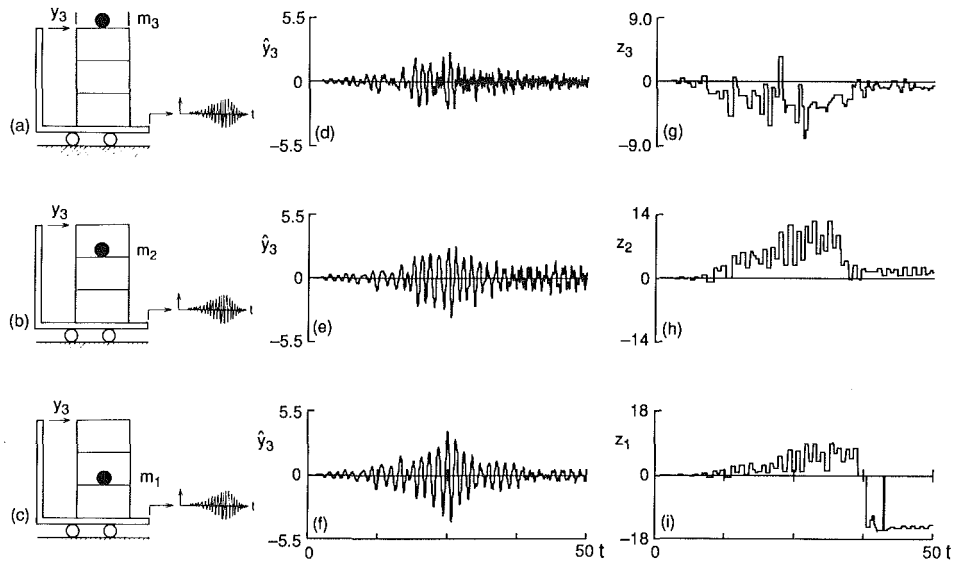


Fig. 7 Influence of the controller location on the response of the structure in Fig. 6. The top row of plots indicate in part (*a*) the location of a single SAID, whose mass equals 5 percent the total primary structure mass, attached to the top "floor" m_3 of the structure, the controlled response of m_3 relative to the moving base is shown in (*d*), and the evolution of the damper gap is shown in (*g*). Similar results are shown in the middle row of plots for the case where the location of the damper is moved from m_3 to m_2, and in the bottom row for the case where the damper location is moved to m_1.

tions that are not shown in the figure. Everything else being the same, it is clear that the top floor is the best location to use if a single SAID is to be employed.

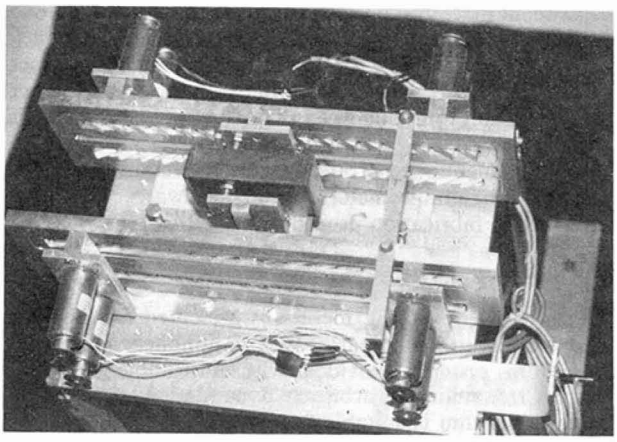
The applied excitation is identical to that used in conjunction with the system of Fig. 6. The mass ratio of the single SAID used in each of the three cases illustrated in Fig. 7 was equivalent to that total mass ratio incorporated in the three SAID used simultaneously in Fig. 6.

Information about the effect of the placement of active control devices on vibrating structures is available in the works of Lindberg and Longman (1984) and Chassiakos et al. (1988).

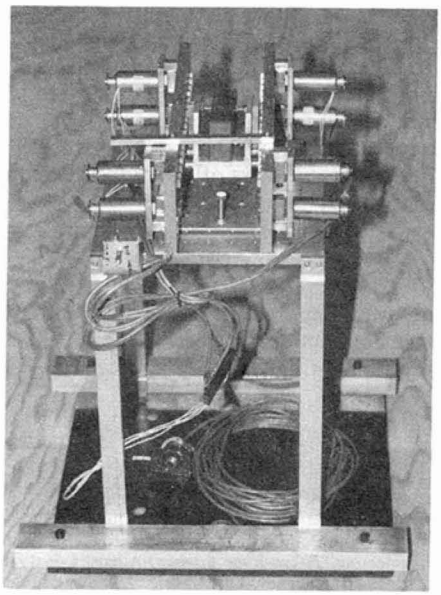
5 Mechanical Model

5.1 Apparatus. A mechanical model resembling a SDOF frame structure was designed and fabricated to investigate the SAID under realistic laboratory conditions. This rudimentary system, shown in Fig. 8, consisted of the following major components:

- a rectangle-shaped container (approximately 35×10 cm in plan) used to constrain the motion of the auxiliary mass,
- a bearing-mounted auxiliary mass which was allowed to



(a)



(b)

Fig. 8 Mechanical model used for implementing the SAID algorithm: (a) top view and (b) front view

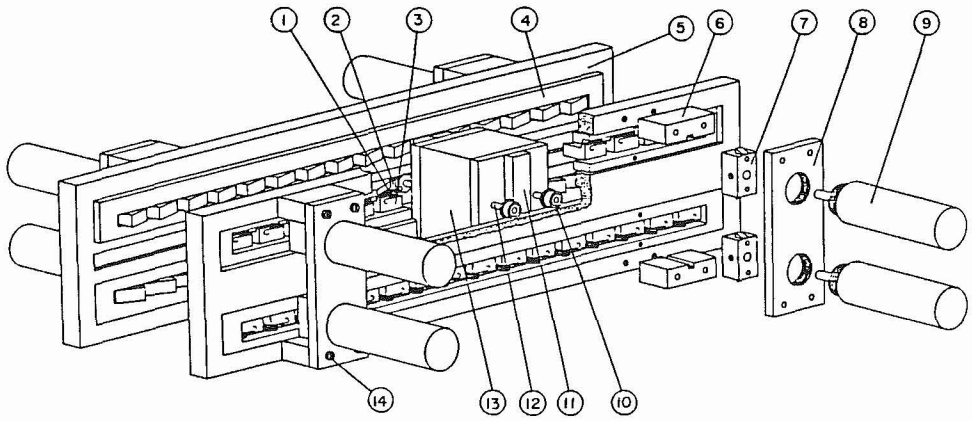
move, with slight friction, in grooves at the center of the container face plates which were rigidly connected to the primary structure,

- four movable panels that were used to position the stoppers and cause an impact between the auxiliary mass and the primary structure. Each of the panels could be moved forward and backward, relative to the centerline of the panel, and
- a set of 16 spring-loaded "stoppers" mounted on each of the panels. These stoppers are hinged wedges, approximately 2-cm apart. They are connected to the moving panel with pins designed to allow the auxiliary mass to move freely in only one direction. The top set of panels allowed unimpeded motion in one direction along the longitudinal axis of the container while the lower set did the same for the opposite direction of motion. Consequently, the panels provided electromechanically-controlled ratchet action.

The principal goals of this design (schematic shown in Fig. 9) is to eliminate all sensors required to monitor the state of the auxiliary mass and to simplify the control to an on/off-type algorithm (i.e., there is no need to compute and supply a value for the gap size). The latter feature minimizes the computation time involved in the decision making process. This shortens the delay, thus allowing more time for hardware activation.

The control logic was implemented on a Z-80 microprocessor using the FORTH language. Upon detection of the displacement zero-crossing by the microprocessor, an "*impact*" command is issued. This event switches a relay circuit, activating magnetic solenoids and moving the impact barriers. As a result, an impact in the desired direction will occur. The solenoids used were capable of moving the panels into impact position in 4 milliseconds (about 1/20th of the fundamental period of the structure).

5.2 Experimental Studies. The response of the structure photographed in Fig. 8 with and without a SAID, under swept-sine and wide-band random excitation is shown in the upper and lower parts of Fig. 10, respectively. The damper mass ratio μ was $\simeq 0.10$. It is seen that under swept-sine excitation the peak displacement response is $\simeq 0.45$ of the cor-



1. RETURN
2. PIN
3. STOPPER
4. MOVING PANEL
5. FACE PLATE
6. SPACER
7. SOLENOID BRACKET
8. SOLENOID MOUNTING PLATE
9. MAGNETIC SOLENOID
10. BEARING
11. MASS SUPPORT
12. AXIAL
13. AUXILIARY MASS
14. SCREW

Fig. 9 Schematic SAID control apparatus

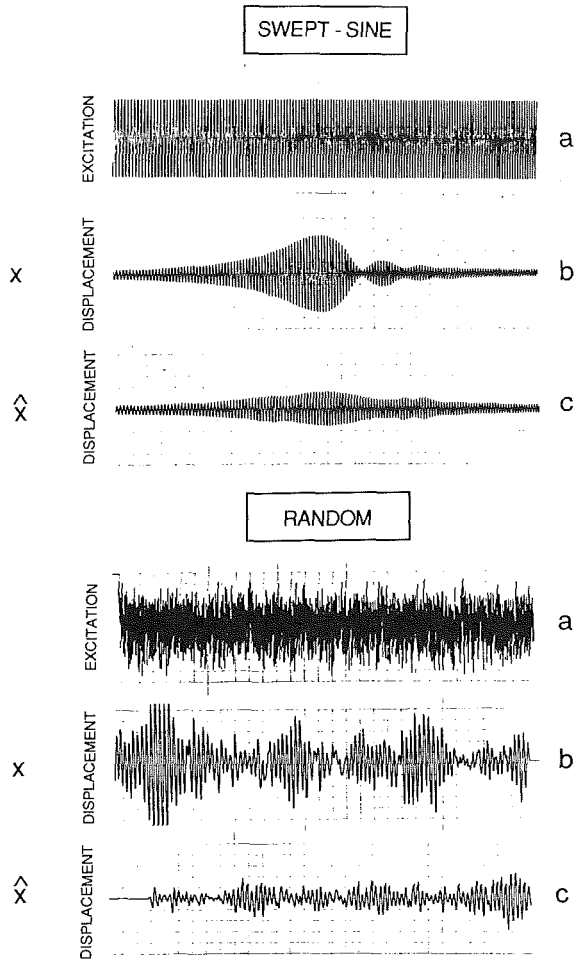


Fig. 10 Transient responses of the mechanical model with and without a SAID. The top and bottom groups of plots correspond to the swept-sine and random excitations, respectively.

responding value in the absence of the damper, while in the case of random excitation this quantity ${}^1\hat{y}_1/{}^1y_1$ is $\simeq 0.42$.

Among the factors contributing to the lowered efficiency of the test apparatus are:

- relatively low coefficient of restitution due to the use of aluminum in constructing the controlling fixture. Analytical studies indicate a much superior performance for the SAID with relatively high values of e, which reduce the loss of impact (control) energy,
- spacing between adjacent stoppers (resolution) was not fine enough to ensure optimum impacts at all times, and
- fabrication inaccuracies (contributed by inexperienced student machinists) in the controller assembly introduced a significant amount of backlash (dead-space nonlinearity) thus increasing the influence of mechanical energy dissipation (at the expense of momentum transfer) on the interaction forces between the structure and the auxiliary mass.

The aforementioned problems can be circumvented by the use of more suitable materials and hardware coupled with more precise fabrication procedures.

6 Summary and Conclusions

A simple, yet efficient, method is presented for the on-line *parameter* control of linear as well as nonlinear multidegree-of-freedom systems provided with adjustable-gap impact dampers responding to arbitrary dynamic loads. The on-line control algorithm is suitable for situations in which detailed knowledge of the system structure is not available; only *local* measurements in the vicinity of each of the attached impact dampers are needed with this adaptive control method to determine the evolution of each impact damper clearance so as to optimize the vibration attenuation efficiency of the individual dampers.

A stability analysis, simulation studies, and experimental tests with a mechanical model have demonstrated the feasibility, reliability, and robustness of the proposed semi-active on-line control method.

Acknowledgments

This study was supported in part by a grant from the National Science Foundation. The assistance of V. Chavakula in the numerical simulation studies is appreciated.

References

Chassiakos, A. G., Masri, S. F., Bekey, G. A., and Miller, R. K., 1988, "Optimum Controller Location for Mitigating Earthquake Induced Response of Structures Provided With Point Actuators," *Proc. Ninth World Conference on Earthquake Engineering, Tokyo-Kyoto, Japan*, August 2–9, 1988.

Karyeaclis, M. P., and Caughey, T. K., 1987, "Stability of a Semi-Active Impact Damper," submitted to ASME JOURNAL OF APPLIED MECHANICS.

Lindberg, Jr., R. E., and Longman, X., 1984, "On the Number and Placement of Actuators for Independent Modal Space Control," *Journal of Guidance*, Vol. 7, No. 2, pp. 215–221.

Masri, S. F., 1972, "Theory of the Dynamic Vibration Neutralizer With Motion-Limiting Stops," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 563–568.

Masri, S. F., Bekey, G. A., and Caughey, T. K., 1981, "Optimium Pulse Control of Flexible Structures," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, pp. 619–626.

Masri, S. F., Bekey, G. A., and Caughey, T. K., 1982, "On-Line Control of Nonlinear Flexible Structures," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, No. 4, pp. 877–884.

Masri, S. F., and Caughey, T. K., 1966, "On the Stability of the Impact Damper," ASME JOURNAL OF APPLIED MECHANICS, Vol. 33, pp. 586–592.

S. L. Lau
Senior Lecturer,
Department of Civil and
Structural Engineering,
Hong Kong Polytechnic, Hong Kong
Mem. ASME

Y. K. Cheung
Pro-vice Chancellor,
Professor and Head.

Shuhui Chen
Research Student.

Department of Civil and
Structural Engineering,
University of Hong Kong, Hong Kong

An Alternative Perturbation Procedure of Multiple Scales for Nonlinear Dynamics Systems

An alternative perturbation procedure of multiple scales is presented in this paper which is capable of treating various periodic and almost periodic steady-state vibrations including combination resonance of nonlinear systems with multiple degrees-of-freedom. This procedure is a generalization of the Lindstedt-Poincaré method. To show its essential features a typical example of cubic nonlinear systems, the clamped-hinged beam, is analyzed. The numerical results for the almost periodic-free vibration are surprisingly close to that obtained by the incremental harmonic balance (IHB) method, and the analytical formulae for steady-state solution are, in fact, identical with that of conventional method of multiple time scales. Moreover, detail calculations of this example revealed some interesting behavior of nonlinear responses, which is of significance for general cubic systems.

I Introduction

It is well known that the perturbation method is one of the commonly used quantitative methods for analyzing nonlinear problems. Nayfeh (1973, 1981) has presented an account of various perturbation techniques, pointing out their similarities, differences, and advantages, as well as their limitations. The most representative perturbation methods used in nonlinear structural vibrations are the Lindstedt-Poincaré method, the method of multiple time scales, and the KBM method. The first method employed by earlier astronomers expands the dependent variable and frequency in power series of small parameter, resulting in a set of linear ordinary differential equations which can be solved successively. With this method, one directly determines the periodic motions. In comparison with the Lindstedt-Poincaré method, the method of multiple time scales appears more involved, but it can provide a more general solution which is able to treat various resonance phenomena and therefore has been widely applied to nonlinear vibration problems in recent years (e.g., Nayfeh, 1983, 1984; Sridharet et al., 1975, 1978; Mook et al., 1985, 1986).

In this paper, a different perturbation procedure of multiple scales is presented. This method is, in fact, a generalization of Lindstedt-Poincaré method and is capable of treating various complicated resonances of multiple degrees-of-freedom systems. In this procedure, the multiple time variables $\tau_n = \omega_n t$ (but not the time scales $T_n = \epsilon^n t$ as in the standard procedure of multiple scales) are employed, so that it may be called the method of multiple dimensions. The dependent variables, as well as the frequencies ω_n, are expanded into power series of small parameter, and the original nonlinear equations are then separated into a series of linear partial differential equations, which can be solved in a stepwise manner. The solution of each approximation expressed in the form of multiple Fourier series is generally an almost periodic steady-state vibration. Only when the nonlinear frequencies are commensurable with each other will the periodic steady-state solution be reduced. The governing equations for the amplitudes and frequencies are algebraic equations, which can usually be solved by routine methods.

The main advantages of the present approach are its intuition in idea and versatility in application. Moreover, it leads directly to the almost periodic or periodic steady-state solutions, which are probably the most attempted in practice.

For convenience of presentation of the general procedure, only cubic nonlinear system is treated in this paper. Obviously, the same procedure can be applied to the systems with different kind of nonlinearities.

To demonstrate the application of the present method, the nonlinear vibrations of a clamped-hinged beam, which is a typical example of systems with cubic nonlinearity possessing internal resonances, are analyzed. The numerical results of almost periodic-free vibrations are quite close to that obtained by IHB method (Lau, Cheung, and Wu, 1983). The internal resonance of the beam has a similar characteristic with that of elastic thin plates (Lau, Cheung, and Wu, 1984). However, it is found by this example that the nonlinear response characteristic is usually excitation-level dependent. It is interesting that this fact provides an explanation of the discrepancy in behavior between different computed results (see Iu and Lau et al., 1983).

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

Discussion on this paper should be addressed to the 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 1987; final revision, August 1988.

II General Procedure

Consider the oscillations of a system having cubic nonlinearities

$$\frac{d^2u_n}{dt^2}\,\Omega_n^2 u_n+\epsilon\,\mu_n\,\frac{du_n}{dt}+\epsilon\sum_{m=1}^{N}\sum_{p=1}^{N}\sum_{q=1}^{N}\Gamma_{nmpq}u_m\,u_p\,u_q=f_n\cos(\Omega t)$$

$$n=1,\,2,\,\ldots \tag{1}$$

where u_n are generalized coordinates of the linear normal modes, Ω_n linear natural frequencies, Γ_{nmpq} coefficients of the nonlinear terms, μ_n coefficients of modal viscous damping, f_n excitation amplitudes, Ω the exciting frequency, and ϵ a small constant parameter.

Following the same consideration of Lau, Cheung, and Wu (1983), one can first introduce multiple time variables defined as

$$\tau_n=\eta\omega_n t \tag{2}$$

in which ω_n are the nonlinear frequenices of responses (in free or forced vibration) generally incommensurable with one another, and η is a rational number depending on the resonance to be sought. The introduction of multiple times τ_n enables the procedure to treat general multiple degrees-of-freedom systems, especially under almost periodic vibrations. For a more detailed explanation of this consideration, please refer to Lau et al., (1983). The generalized coordinates u_n $(n=1,\,2,\,\ldots)$ are then regarded as functions of independent variables τ_n. Let u_n and ω_n be expanded in power series of ϵ similar to that of Lindstedt-Poincaré method:

$$u_n(\tau_1,\,\tau_2,\,\ldots\,\tau_N)=\sum_{k=0}^{\infty}u_{nk}(\tau_1,\,\tau_2,\,\ldots\,\tau_N)\epsilon^k \tag{3}$$

$$\omega_n=\sum_{k=0}^{\infty}\omega_{nk}\epsilon^k \tag{4}$$

and assume that

$$f_n=\sum_{k=0}^{\infty}f_{nk}\,\epsilon^k \tag{5}$$

whereupon

$$\frac{du_n}{dt}=\sum_{i=1}^{N}\eta\,\omega_i\,\frac{\partial u_n}{\partial\tau_i}=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\epsilon^{k+1}D_k u_{nl} \tag{6}$$

$$\frac{d^2u_n}{dt^2}=\sum_{i=1}^{N}\sum_{j=1}^{N}\eta^2\,\omega_i\,\omega_j\,\frac{\partial^2u_n}{\partial\tau_i\partial\tau_j}=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\epsilon^{k+l+m}\,D_{kl}^2u_{nm} \tag{7}$$

where the operators D_k and D_{kl}^2 are introduced for conciseness:

$$D_k\equiv\sum_{i=1}^{N}\eta\,\omega_{ik}\,\frac{\partial}{\partial\tau_i} \tag{8}$$

$$D_{kl}^2\equiv D_k\,D_l\equiv\sum_{i=1}^{N}\sum_{j=1}^{N}\eta^2\,\omega_{ik}\,\omega_{jl}\,\frac{\partial^2}{\partial\tau_i\partial\tau_j}. \tag{9}$$

Substituting equations (3), (5), (6), and (7) into equation (1) and then equating the coefficients of powers of ϵ, we obtain

$$D^2{}_{00}u_{n0}+\Omega_n^2u_{n0}=f_{n0}\cos T \tag{10}$$

$$D_{00}^2u_{n1}+\Omega_n^2n_{n1}=-2D_{01}^2u_{n0}-\mu_nD_0u_{n0}-\sum_{m=1}^{N}\sum_{p=1}^{N}\sum_{q=1}^{N}\Gamma_{nmpq}u_{m0}u_{p0}u_{q0}+f_{n1}\cos T \tag{11}$$

$$D_{00}^2u_{n2}+\Omega_n^2n_{n2}=-2D_{01}^2u_{n1}-D_{11}^2u_{n0}-2D_{02}^2u_{n0}-\mu_n(D_0u_{n1}+D_1u_{n0})-\sum_{m=1}^{N}\sum_{p=1}^{N}\sum_{q=1}^{N}\alpha_{nmpq}\,u_{m0}\,u_{p0}\,u_{q1}+f_{n2}\cos T \tag{12}$$

. . . . . .

where

$$\alpha_{nmpq}=\Gamma_{nmpq}+\Gamma_{npqm}+\Gamma_{nqmp}$$

$$T=\Omega\,t.$$

In the case of resonance, it can be expressed for generality as

$$T=\sum_{i=1}^{N}a_i\tau_i \tag{13}$$

in which a_i are rational constants. The solution of equation (10) can be expressed as

$$u_{n0}=A_{n0}\cos(\tau_n+\phi_{n0})+F_{n0}\cos T \tag{14}$$

in which A_{n0} and ϕ_{n0} are integration constants, and

$$F_{n0}=f_{n0}/\left(\Omega_n^2-\sum_{i=1}^{N}\sum_{j=1}^{N}\eta^2\omega_{i0}\omega_{j0}a_ia_j\right) \tag{15}$$

$$\omega_{n0}=\Omega_n/\eta. \tag{16}$$

While the term $A_{n0}\cos(\tau_n+\phi_{n0})$ is a solution of the corresponding homogeneous equation,

$$D_{00}^2u_{n0}+\Omega_n^2u_{n0}=0. \tag{17}$$

Note that there are many other solutions of equation (17), such as

$$\cos\left(\frac{\Omega_n}{\omega_{i0}}\tau_i\right),\ \cos\left(\frac{2\Omega_n}{\omega_{i0}}\tau_i-\frac{\Omega_n}{\omega_{n0}}\tau_n\right),$$

$$\cos\left(\sum_{i=1}^{N}\frac{2\Omega_n}{(N-1)\omega_{i0}}\tau_i-\frac{\Omega_n}{\omega_{n0}}\tau_n\right)\quad i\neq n$$

etc., but these are precluded in the present solution. To explain this let us be reminded of the fact that for Lindstedt's method, the solution is expanded into a single Fourier series, i.e.,

$$\sum_n[A_n\cos(n\tau)+B_n\sin(n\tau)]$$

similarly, for the present case with multiple time variables, the generalized steady-state solution should be in the form of multiple Fourier series (Lau et al., 1983).

$$\sum_{j_i}\sum_{j_2}\cdots\sum_{j_m}\left[A_{j_1j_2\ldots j_m}\cos\left(\sum_k j_k\tau_k\right)+B_{j_1j_2\ldots j_m}\sin\left(\sum_k j_k\tau_k\right)\right],$$

where j_k $(k=1,\,2,\,\ldots)$ are integers. Therefore, within the frame of this solution form, the solution for the homogeneous equation (17) should be taken as

$$u_{n0}=C_{n0}\cos\tau_n+E_{n0}\sin\tau_n=A_{n0}\cos(\tau_n+\phi_{n0}). \tag{18}$$

Inserting (14) into (11), we obtain

$$D_{00}^2 u_{n1} + \Omega_n^2 u_{n1} = 2\eta^2 \omega_{n0}\omega_{n1} A_{n0} \cos(\tau_n + \phi_{n0})$$
$$+ 2\sum_{i=1}^{N}\sum_{j=1}^{N} \eta^2 \omega_{i0}\omega_{j1} a_i a_j F_{n0} \cos T$$
$$+ \mu_n \eta \omega_{n0} A_{n0} \sin(\tau_n + \phi_{n0}) + \mu_n \sum_{i=1}^{N} \eta\, \omega_{i0} a_i F_{n0} \sin T$$
$$- \sum_{m=1}^{N}\sum_{p=1}^{N}\sum_{q=1}^{N} \left\{ \frac{1}{4} \Gamma_{nmpq} A_{m0} A_{p0} A_{q0} \right.$$
$$\times \cos(\tau_m + \tau_p + \tau_q + \phi_{m0} + \phi_{p0} + \phi_{q0})$$
$$+ \frac{1}{4} \alpha_{nmpq} [A_{m0} A_{p0} A_{q0} \cos(\tau_m + \tau_p - \tau_q + \phi_{m0} + \phi_{p0} - \phi_{q0})$$
$$+ A_{m0} A_{p0} F_{q0} [\cos(\tau_m + \tau_p + T + \phi_{m0} + \phi_{p0})$$
$$+ \cos(\tau_m + \tau_p - T + \phi_{m0} + \phi_{p0})$$
$$+ \cos(\tau_m - \tau_p + T + \phi_{m0} - \phi_{p0}) + \cos(\tau_m - \tau_p - T + \phi_{m0} - \phi_{p0})]$$
$$+ A_{m0} F_{p0} F_{q0} [\cos(\tau_m + 2T + \phi_{m0}) + \cos(\tau_m - 2T + \phi_{m0})$$
$$+ 2\cos(\tau_m + \phi_{m0})]]$$
$$\left. + \frac{1}{4} \Gamma_{nmpq} F_{m0} F_{p0} F_{q0} [3\cos T + \cos 3T] \right\}$$
$$+ f_{n1} \cos T. \tag{19}$$

The terms containing $\cos(\tau_n + \phi_{n0})$ or $\sin(\tau_n + \phi_{n0})$ on the right-hand side of equation (19) will produce the secular terms, which should not be parts of a uniformly valid expansion. To eliminate these secular terms, the coefficients of $\cos(\tau_n + \phi_{n0})$ and $\sin(\tau_n + \phi_{n0})$ must be zero. This leads to a set of algebraic equations governing the relationships between amplitudes A_{n0} and frequencies ω_{n1}. Thus, various resonances corrected up to the first order can be easily obtained by solving these equations.

Having known A_{n0} and ω_{n1}, we can determine u_{n1} from (19). Then, substitute them into equation (12) and continue this procedure to determine $u_{n2}, u_{n3} \ldots$, and so on.

The procedure is obviously the generalization of the Lindstedt-Poincaré method and will be demonstrated in the next section in which the details of treating the specific resonances of beam problem are considered.

III Application to Nonlinear Vibrations of Clamped-Hinged Beam

This problem is a typical example of cubic nonlinearity system possessing the phenomenon of internal resonance, and was first analyzed by Nayfeh, Mook, and Sridhar (1974, 1975) using the standard method of multiple scales. Lau, Cheung, and Wu (1983) analyzed the almost periodic-free vibration of this problem numerically by the use of the IHB method with multiple time variables. In this section we consider both the free vibrations and the forced vibrations including fundamental resonance and combination resonance under the influence of internal resonance using the method developed in Section II.

1 Almost Periodic-Free Vibration. The governing equations describing the transverse-free vibration of a undamped clamped-hinged beam using a two-mode shape approximation finally can be written as (Lau et al., 1983):

$$\frac{d^2 u_1}{dt^2} + \Omega_1^2 u_1 + \alpha_{11} u_1^3 + \alpha_{12} u_1^2 u_2 + \alpha_{13} u_1 u_2^2 + \alpha_{14} u_2^3 = 0$$
$$\frac{d^2 u_2}{dt^2} + \Omega_2^2 u_2 + \alpha_{21} u_2^3 + \alpha_{22} u_2^2 u_1 + \alpha_{23} u_2 u_1^2 + \alpha_{24} u_1^3 = 0. \tag{20}$$

Where u_1 and u_2 are normal mode coordinates, t is the normalized time, $\Omega_1 = 1$, $\Omega_2 = 3.24064$ are the first and the second normalized linear frequencies, respectively, coefficients α_{ij} ($i = 1, 2$; $j = 1, 2, 3, 4$) are constants (see the Appendix). Obviously, equation (20) is a special case of equation (1) with $\mu_n = 0$, $f_n = 0$. $N = 2$ and $\epsilon = 1$.

The following solution clearly illustrates that the introduction of multiple time variables are necessitated for obtaining the almost-periodic vibration solutions.

Following the procedure developed in Section II, we can obtain each successive approximation solution as follows:

The first approximation solutions are

$$\omega_{10} = \Omega_1 \qquad \omega_{20} = \Omega_5 \tag{21}$$
$$u_{10} = A_{10} \cos\tau_1 \qquad u_{20} = A_{20} \cos\tau_2. \tag{22}$$

The second approximation solutions are

$$\omega_{11} = \frac{1}{2\Omega_1}\left(\frac{3}{4}\alpha_{11} A_{10}^2 + \frac{1}{2}\alpha_{13} A_{20}^2\right)$$
$$\omega_{21} = \frac{1}{2\Omega_2}\left(\frac{3}{4}\alpha_{21} A_{20}^2 + \frac{1}{2}\alpha_{23} A_{10}^2\right) \tag{23}$$

$$u_{11} = C_1^{(1)} \cos\tau_1 + C_{12}^{(1)} \cos 3\tau_1 + C_{13}^{(1)} \cos\tau_2 + C_{14}^{(1)} \cos 3\tau_2$$
$$+ C_{15}^{(1)} \cos(2\tau_1 + \tau_2) + C_{16}^{(1)} \cos(2\tau_1 - \tau_2) + C_{17}^{(1)} \cos(2\tau_2 + \tau_1)$$
$$+ C_{18}^{(1)} \cos(2\tau_2 - \tau_1)$$
$$u_{21} = C_{21}^{(1)} \cos\tau_2 + C_{22}^{(1)} \cos 3\tau_2 + C_{23}^{(1)} \cos\tau_1 + C_{24}^{(1)} \cos 3\tau_1$$
$$+ C_{25}^{(1)} \cos(2\tau_2 + \tau_1) + C_{26}^{(1)} \cos(2\tau_2 - \tau_1) + C_{27}^{(1)} \cos(2\tau_1 + \tau_2)$$
$$+ C_{28}^{(1)} \cos(2\tau_1 - \tau_2). \tag{24}$$

The third approximation solutions are

$$\omega_{12} = \frac{1}{2\omega_{10} A_{10}}\left[-\omega_{11}^2 A_{10} + \frac{3}{4}\alpha_{11} A_{10}^2 (3C_{11}^{(1)} + C_{12}^{(1)})\right.$$
$$+ \frac{1}{4}\alpha_{12} A_{10}^2 (3C_{23}^{(1)} + C_{24}^{(1)})$$
$$+ \frac{1}{2}\alpha_{12} A_{10} A_{20} (2C_{13}^{(1)} + C_{15}^{(1)} + C_{16}^{(1)})$$
$$+ \frac{1}{2}\alpha_{13} A_{10} A_{20} (2C_{21}^{(1)} + C_{27}^{(1)} + C_{28}^{(1)})$$
$$+ \frac{1}{4}\alpha_{13} A_{20}^2 (2C_{11}^{(1)} + C_{17}^{(1)} + C_{18}^{(1)})$$
$$\left. + \frac{3}{4}\alpha_{14} A_{20}^2 (2C_{23}^{(1)} + C_{25}^{(1)} + C_{26}^{(1)})\right]$$
$$\omega_{22} = \frac{1}{2\omega_{20} A_{20}}\left[-\omega_{21}^2 A_{20} + \frac{3}{4}\alpha_{21} A_{20}^2 (3C_{21}^{(1)} + C_{22}^{(1)})\right.$$
$$+ \frac{1}{4}\alpha_{22} A_{20}^2 (3C_{13}^{(1)} + C_{14}^{(1)})$$
$$+ \frac{1}{2}\alpha_{22} A_{20} A_{10} (2C_{23}^{(1)} + C_{25}^{(1)} + C_{26}^{(1)})$$
$$+ \frac{1}{2}\alpha_{23} A_{20} A_{10} (2C_{11}^{(1)} + C_{17}^{(1)} + C_{18}^{(1)})$$
$$+ \frac{1}{4}\alpha_{23} A_{10}^2 (2C_{21}^{(1)} + C_{27}^{(1)} + C_{28}^{(1)})$$
$$\left. + \frac{3}{4}\alpha_{24} A_{10}^2 (2C_{13}^{(1)} + C_{15}^{(1)} + C_{16}^{(1)})\right] \tag{25}$$

Table 1 The values of C_{ij} in equations (27)

Method	IHB Method	Present Method		IHB Method	Present Method	
		2nd Approximation	3rd Approximation		2nd Approximation	3rd Approximation
ω_1	1.01556	1.01568	1.01555	1.0427	1.0436	1.0426
ω_2	3.25841	3.25845	3.25841	3.2898	3.2902	3.2898
C_{11}	0.300_{E0}	0.300_{E0}	0.300_{E0}	0.500_{E0}	0.500_{E0}	0.500_{E0}
C_{12}	0.228_{E-3}	0.235_{E-3}	0.227_{E-3}	0.997_{E-3}	0.109_{E-2}	0.998_{E-3}
C_{13}	-0.323_{E-3}	-0.324_{E-3}	-0.323_{E-3}	-0.151_{E-2}	-0.150_{E-2}	-0.149_{E-2}
C_{14}	-0.353_{E-5}	-0.349_{E-5}	-0.353_{E-5}	-0.167_{E-4}	-0.161_{E-4}	-0.168_{E-4}
C_{15}	-0.389_{E-4}	-0.397_{E-4}	-0.389_{E-4}	-0.174_{E-3}	-0.184_{E-3}	-0.173_{E-3}
C_{16}	-0.224_{E-2}	-0.195_{E-2}	-0.220_{E-2}	-0.143_{E-1}	-0.901_{E-2}	-0.123_{E-1}
C_{17}	0.340_{E-4}	0.343_{E-4}	0.340_{E-4}	0.155_{E-3}	0.159_{E-3}	0.155_{E-3}
C_{18}	0.647_{E-4}	0.648_{E-4}	0.647_{E-4}	0.301_{E-3}	0.300_{E-3}	0.300_{E-3}
C_{21}	0.150_{E0}	0.150_{E0}	0.150_{E0}	0.250_{E0}	0.250_{E0}	0.250_{E0}
C_{22}	0.386_{E-4}	0.389_{E-4}	0.386_{E-4}	0.176_{E-3}	0.180_{E-3}	0.176_{E-3}
C_{23}	0.628_{E-3}	0.633_{E-3}	0.627_{E-3}	0.289_{E-2}	0.293_{E-2}	0.286_{E-2}
C_{24}	0.530_{E-3}	0.466_{E-3}	0.521_{E-3}	0.333_{E-2}	0.216_{E-2}	0.286_{E-2}
C_{25}	-0.420_{E-4}	-0.430_{E-4}	-0.420_{E-4}	-0.187_{E-3}	-0.199_{E-3}	-0.186_{E-3}
C_{26}	-0.101_{E-3}	-0.100_{E-3}	-0.100_{E-3}	-0.486_{E-3}	-0.463_{E-3}	-0.467_{E-3}
C_{27}	0.216_{E-3}	0.222_{E-3}	0.216_{E-3}	0.952_{E-3}	0.103_{E-3}	0.947_{E-3}
C_{28}	-0.417_{E-3}	-0.420_{E-3}	-0.416_{E-3}	-0.192_{E-2}	-0.195_{E-2}	-0.189_{E-2}

$$u_{12} = C_{11}^{(2)} \cos\tau_1 + C_{12}^{(2)} \cos 3\tau_1 + C_{13}^{(2)} \cos\tau_2 + C_{14}^{(2)} \cos 3\tau_2 + C_{15}^{(2)} \cos(2\tau_1 + \tau_2) + C_{16}^{(2)} \cos(2\tau_1 - \tau_2) + C_{17}^{(2)} \cos(2\tau_2 + \tau_1) + C_{18}^{(2)} \cos(2\tau_2 - \tau_1) + \ldots$$

$$u_{22} = C_{21}^{(2)} \cos\tau_2 + C_{22}^{(2)} \cos 3\tau_2 + C_{23}^{(2)} \cos\tau_1 + C_{24}^{(2)} \cos 3\tau_1 + C_{25}^{(2)} \cos(2\tau_2 + \tau_1) + C_{26}^{(2)} \cos(2\tau_2 - \tau_1) + C_{27}^{(2)} \cos(2\tau_1 + \tau_2) + C_{28}^{(2)} \cos(2\tau_1 - \tau_2) + \ldots, \quad (26)$$

where coefficients $C_{ij}^{(1)}$ and $C_{ij}^{(2)}$ ($i=1, 2$; $j=2, \ldots, 8$) are given in the Appendix. Therefore, the almost-periodic steady-state solutions to the third-order approximation are given by

$$\omega_n = \omega_{n0} + \omega_{n1} + \omega_{n2} \quad (n=1, 2)$$

$$u_1 = C_{11} \cos\tau_1 + C_{12} \cos 3\tau_1 + C_{13} \cos\tau_2 + C_{14} \cos 3\tau_2 + C_{15} \cos(2\tau_1 + \tau_2) + C_{16} \cos(2\tau_1 - \tau_2) + C_{17} \cos(2\tau_2 + \tau_1) + C_{18} \cos(2\tau_2 - \tau_1) + \ldots$$

$$u_2 = C_{21} \cos\tau_2 + C_{22} \cos 3\tau_2 + C_{23} \cos\tau_1 + C_{24} \cos 3\tau_1 + C_{25} \cos(2\tau_2 + \tau_1) + C_{26} \cos(2\tau_2 - \tau_1) + C_{27} \cos(2\tau_1 + \tau_2) + C_{28} \cos(2\tau_1 - \tau_2) + \ldots \quad (27)$$

where

$$\tau_1 = \omega_1 t, \qquad \tau_2 = \omega_2 t \qquad (\eta = 1)$$

$$C_{i1} = A_{i0} + C_{i1}^{(1)} + C_{i1}^{(2)}$$

$$C_{ij} = C_{ij}^{(1)} + C_{ij}^{(2)} \quad i=1, 2; j=2, 3, \ldots 8. \quad (28)$$

In order to make comparisons the coefficients C_{ij}, calculated by the present method and by IHB method (Lau et al., 1983), are listed in Table 1. It can be seen that the discrepancies of the results between the two methods are quite small. In fact, the results of the second approximation have already been accurate enough for the case of moderately large amplitudes.

2 Forced Vibration With Internal Resonance. For the forced vibration of the beam with two-mode approximation, the governing equations read:

$$\frac{d^2u_1}{dt^2} + \Omega_1^2 u_1 + \mu_1 \frac{du_1}{dt} + \alpha_{11} u_1^3 + \alpha_{12} u_1^2 u_2 + \alpha_{13} u_1 u_2^2 + \alpha_{14} u_2^3 = f_1 \cos T$$

$$\frac{d^2u_2}{dt^2} + \Omega_2^2 u_2 + \mu_2 \frac{du_2}{dt} + \alpha_{21} u_2^3 + \alpha_{22} u_2^2 u_1 + \alpha_{23} u_2 u_1^2 + \alpha_{24} u_1^3 = f_2 \cos T \quad (29)$$

where μ_1 and μ_2 are viscous damping coefficients, f_1 and f_2 are forcing coefficients, and $T=\Omega t$, Ω is exciting frequency.

2.1 Fundamental Resonance, Ω Near Ω_1. In this case, we should take $\Omega=\omega_1$, i.e.,

$$T = \tau_1. \quad (30)$$

It should be stressed that ω_1, as well as ω_2, are nonlinear response frequencies in the present formulation.

The second linear frequency, Ω_2, is nearly three times the fundamental frequency Ω_1, and therefore, the internal resonance is likely to occur. For a periodic solution we should let $\omega_2 = 3\omega_1$, i.e.,

$$\tau_2 = 3\tau_1. \quad (31)$$

Since $T=\tau_1$, f_{10} must be zero, otherwise it will produce a secular term in u_{10}. Hence we have $F_{10}=0$. Similarly, to eliminate secular terms, the coefficients of $\cos(\tau_1+\phi_{10})$ and $\cos(\tau_2+\phi_{20})$ in equation (19) for $n=1$ and $n=2$ must be zero, respectively. Thus we obtain the solvability conditions relating amplitudes A_{n0}, phase angles ϕ_{n0}, and frequencies ω_{n1}.

$$\mu_1\eta\omega_{10}A_{10} + \frac{1}{4}\alpha_{12}A_{10}^2A_{20}\sin(\phi_2 - 3\phi_{10}) - \frac{1}{4}\alpha_{12}A_{10}^2F_{20}\sin\phi_{10} + \frac{1}{4}\alpha_{13}A_{10}A_{20}F_{20}\sin(\phi_{20} - 2\phi_{10}) - \frac{3}{2}\alpha_{14}A_{20}^2F_{20}\sin\phi_{10} - \frac{1}{4}\alpha_{13}A_{10}F_{20}^2\sin 2\phi_{10} + \frac{3}{4}\alpha_{14}A_{20}F_{20}^2\sin(\phi_{20} - \phi_{10}) - \frac{3}{4}\alpha_{14}F_{20}^3\sin\phi_{10} + f_{11}\sin\phi_{10} = 0$$

$$2\eta^2\omega_{10}\omega_{11}A_{10} - \frac{3}{4}\alpha_{11}A_{10}^3 - \frac{1}{2}\alpha_{13}A_{10}A_{20}^2 - \frac{1}{4}\alpha_{12}A_{10}^2A_{20}\cos(\phi_{20} - 3\phi_{10}) - \frac{3}{4}\alpha_{12}A_{10}^2F_{20}\cos\phi_{10} - \frac{1}{4}\alpha_{13}A_{10}A_{20}F_{20}\cos(\phi_{20} - 2\phi_{10}) - \frac{3}{2}\alpha_{14}A_{20}^2F_{20}\cos\phi_{10} - \frac{1}{4}\alpha_{13}A_{10}F_{20}^2(\cos 2\phi_{10} + 2) - \frac{3}{4}\alpha_{14}A_{20}F_{20}^2\cos(\phi_{20} - \phi_{10}) - \frac{3}{4}\alpha_{14}F_{20}^3\cos\phi_{10} + f_{11}\cos\phi_{10} = 0$$

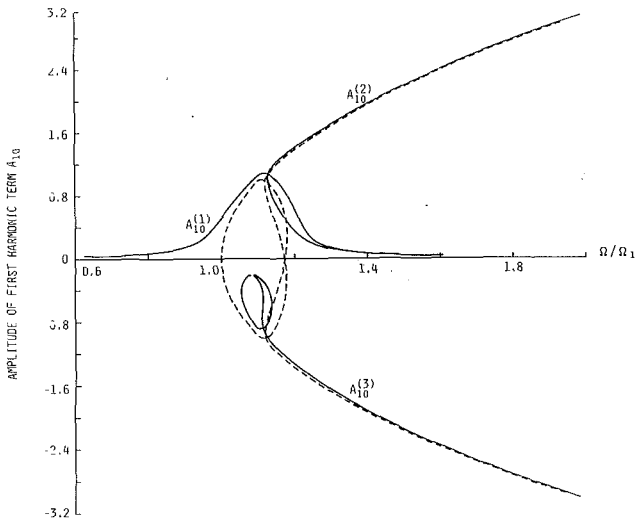


Fig. 1 Forced frequency response as $f_{11}=0.03$, $T=\tau_1$ ——— forced frequency response, - - - - - ω_1 backbone curve, A_{10} = amplitude of first harmonic term

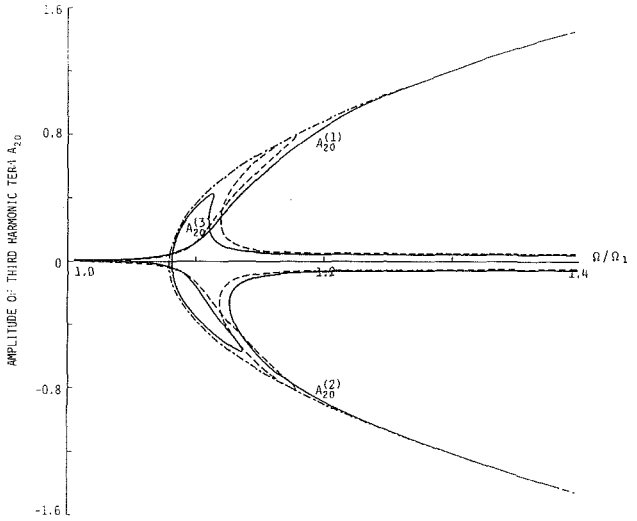


Fig. 2 Forced frequency response as $f_{11}=0.03$, $T=\tau_1$ ——— forced frequency response, - - - - - ω_1 backbone curve, —·— ω_2 backbone curve. A_{20} = amplitude of third harmonic term.

$$\mu_2\eta\omega_{20}A_{20}-\frac{1}{4}\alpha_{24}A_{10}^3\sin(\phi_{20}-3\phi_{10})$$

$$-\frac{1}{4}\alpha_{23}A_{10}^2F_{20}\sin(\phi_{20}-2\phi_{10})$$

$$-\frac{1}{4}\alpha_{22}A_{10}F_{20}^2\sin(\phi_{20}-\phi_{10})-\frac{1}{4}\alpha_{21}F_{20}^3\sin\phi_{20}=0$$

$$2\eta^2\omega_{20}\omega_{21}A_{20}-\frac{3}{4}\alpha_{21}A_{20}^3-\frac{1}{2}\alpha_{23}A_{20}A_{10}^2$$

$$-\frac{1}{4}\alpha_{24}A_{10}^3\cos(\phi_{20}-3\phi_{10})$$

$$-\frac{3}{2}\alpha_{21}A_{20}F_{20}^2-\frac{1}{4}\alpha_{23}A_{10}^2F_{20}\cos(\phi_{20}-2\phi_{10})$$

$$-\alpha_{22}A_{20}A_{10}F_{20}\cos\phi_{10}-\frac{1}{4}\alpha_{22}A_{10}F_{20}^2\cos(\phi_{20}-\phi_{10})$$

$$-\frac{1}{4}\alpha_{21}F_{20}^3\cos\phi_{20}=0. \tag{32}$$

Nayfeh and Mook (1979) analyzed equations (29) with the method of standard multiple scales in which they considered $\mu_n=0$ and $f_{n0}=0$, $n=1, 2$. If we take $\eta=1$, $\mu_1=\mu_2=0$, and $F_{20}=0$ in (32), then obviously $\phi_{10}=\phi_{20}=0$. Hence, equations of (32) are reduced to

$$2\omega_{10}\omega_{11}A_{10}-\frac{3}{4}\alpha_{11}A_{10}^3-\frac{1}{2}\alpha_{13}A_{10}A_{20}^2$$

$$-\frac{1}{4}\alpha_{12}A_{10}^2A_{20}+f_{11}=0$$

$$2\omega_{20}\omega_{21}A_{20}-\frac{3}{4}\alpha_{21}A_{20}^3-\frac{1}{2}\alpha_{23}A_{20}A_{10}^2$$

$$-\frac{1}{4}\alpha_{24}A_{10}^3=0. \tag{33}$$

It can be found that equations (33) are, in fact, the same as those obtained by Nayfeh and Mook.

Figures 1 and 2 show the frequency response curves $\Omega-A_{10}$ and $\Omega-A_{20}$ for undamped forced vibration. The undamped, free vibration backbone curves for A_{10} and A_{20} are also plotted with dotted line and dash-dotted line, respectively, to facilitate the understanding for the relation between forced and free vibration. It is apparent that there are two separate branches of solution, i.e., the "in-phase" and the "out-of-phase" resonances. The in-phase resonance is shown by curves $A_{10}^{(1)}$, $A_{10}^{(2)}$ and $A_{20}^{(1)}$, $A_{20}^{(2)}$, while the out-of-phase resonance is shown by $A_{10}^{(3)}$ and $A_{20}^{(3)}$. It can be seen from Figs. 1 and 2 that the superharmonic resonances caused by internal resonance exist in both in-phase and out-of-phase responses. However, detail calculation reveals that the out-of-phase response is excitation-level dependent. If the excitation increases beyond a certain critical level, the out-of-phase superharmonic resonance disappears. An example of response at critical excitation is shown in Figs. 3 and 4. This phenomenon can be explained in that the out-of-phase response curve will shift towards far right to the backbone curve as the applied forces increase beyond the critical excitation. Therefore, no superharmonic resonance can be excited.

It is worth pointing out that the forced vibration responses of beam with internal resonance exhibit the same character as those of thin plates and sandwich plates computed with the IHB method by Lau, Cheung, and Wu (1984), and by Iu (1985), respectively, as they are all of cubic nonlinearity with similar frequency distribution. The particularly interesting point is that the out-of-phase superharmonic resonance, as shown in Figs. 1 and 2, which occurs exactly in the same manner in the case of thin plate, does not appear in the computed response of sandwich plate (i.e., its responses are similar to those shown in Figs. 3 and 4). This discrepancy may now be explained as such that the former results correspond to an undercritical excitation, while the latter results an overcritical excitation.

2.2 Combination Resonance, Ω Near $1/2(\Omega_1+\Omega_2)$. For treating this case, we should take $\Omega=1/2(\omega_1+\omega_2)$, i.e.,

$$T=\frac{1}{2}(\tau_1+\tau_2) \tag{34}$$

and let

$$\tau_2=3\ \tau_1 \tag{35}$$

in line with the periodic solution. Thus, T can also be rewritten in terms of τ_1 alone as

$$T=2\ \tau_1. \tag{36}$$

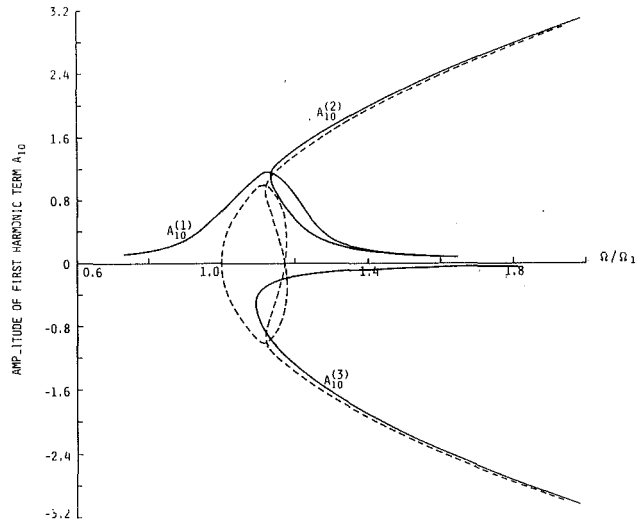


Fig. 3 Forced frequency response as $f_{11} = 0.064$, $T = \tau_1$, A_{10} = amplitude of first harmonic term

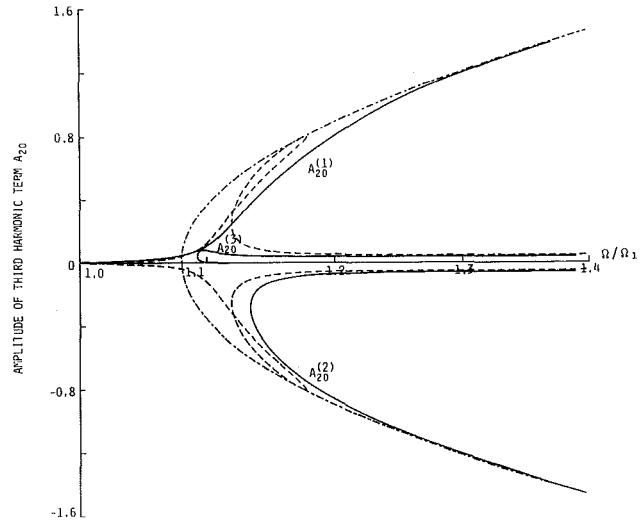


Fig. 4 Forced frequency response as $f_{11} = 0.064$, $T = \tau_1$, A_{20} = amplitude of third harmonic term

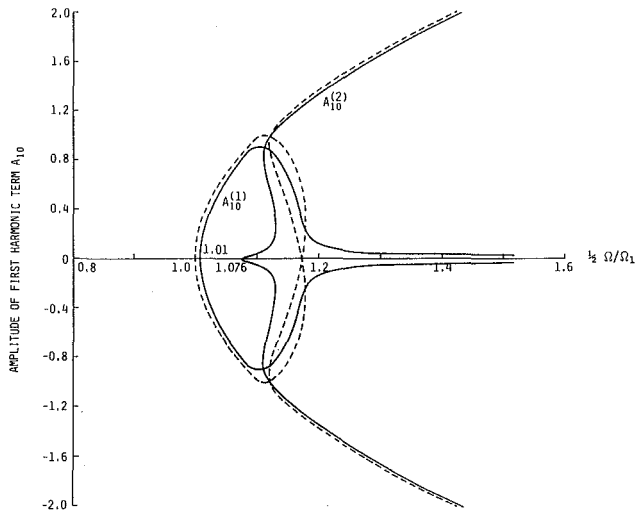


Fig. 5 Combination response with internal resonance $T = 1/2\ (\tau_1 + \tau_2)$, $\tau_2 = 3\tau_1$, $f_{10} = 0.5$, $f_{20} = 0.5$, A_{10} = amplitude of first harmonic term

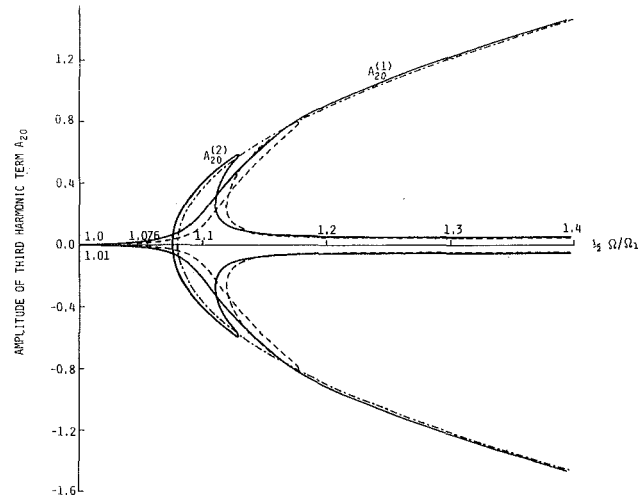


Fig. 6 Combination response with internal resonance $T = 1/2\ (\tau_1 + \tau_2)$, $\tau_2 = 3\tau_1$, $f_{10} = 0.5$, $f_{20} = 0.5$, A_{10} = amplitude of third harmonic term

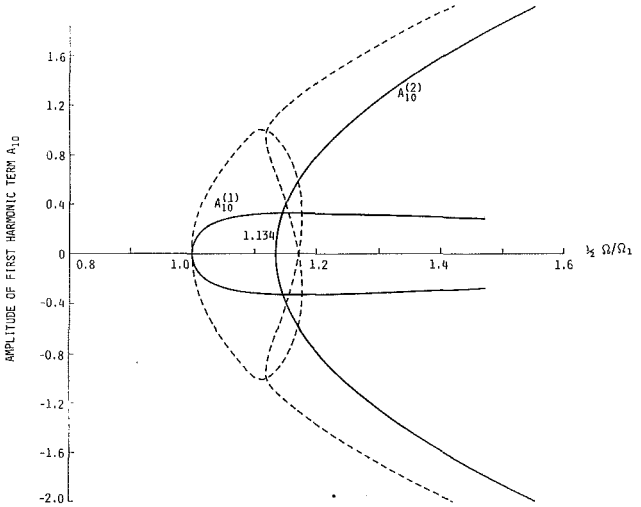


Fig. 7 Combination response with internal resonance $T = 1/2\ (\tau_1 + \tau_2)$, $\tau_2 = 3\tau_1$, $f_{10} = 2$, $f_{20} = 2$, A_{10} = amplitude of first harmonic term

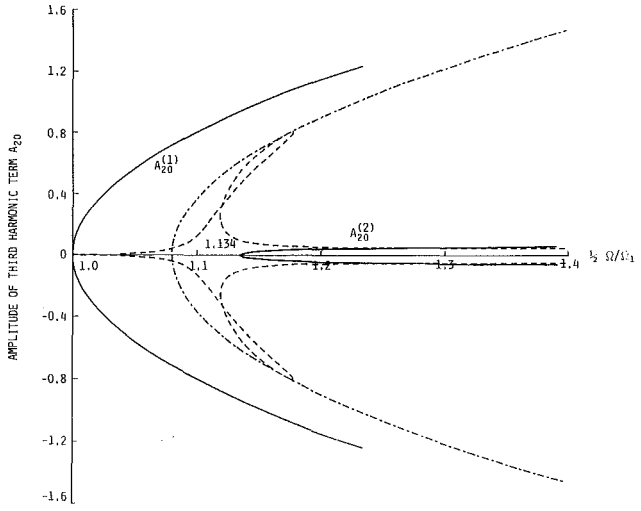


Fig. 8 Combination response with internal resonance $T = 1/2\ (\tau_1 + \tau_2)$, $\tau_2 = 3\tau_1$, $f_{10} = 2$, $f_{20} = 2$, A_{10} = amplitude of third harmonic term

The solvability conditions are obtained from equation (19) by letting the coefficients of $\cos(\tau_1+\phi_{10})$ and $\cos(\tau_2+\phi_{20})$ for $n=1$ and $n=2$ be zero, respectively.

$$\mu_1\eta\,\omega_{10}A_{10}+\frac{1}{4}\alpha_{12}A_{10}^2A_{20}\sin(\phi_{20}-3\phi_{10})$$

$$-\left(\frac{1}{4}\alpha_{12}F_{10}^2+\frac{1}{2}\alpha_{13}F_{10}F_{20}\right.$$

$$\left.+\frac{3}{4}\alpha_{14}F_{20}^2\right)A_{20}\sin(\phi_{20}+\phi_{10})=0$$

$$2\eta^2\omega_{10}\omega_{11}A_{10}-\frac{3}{4}\alpha_{11}A_{10}^3-\frac{1}{2}\alpha_{13}A_{10}A_{20}^2$$

$$-\frac{1}{4}\alpha_{12}A_{10}^2A_{20}\cos(\phi_{20}-3\phi_{10})$$

$$-\left(\frac{3}{2}\alpha_{11}F_{10}^2+\alpha_{12}F_{10}F_{20}+\frac{1}{2}\alpha_{13}F_{20}^2\right)A_{10}$$

$$-\left(\frac{1}{4}\alpha_{12}F_{10}^2+\frac{1}{2}\alpha_{13}F_{10}F_{20}\right.$$

$$\left.+\frac{3}{4}\alpha_{14}F_{20}^2\right)A_{20}\cos(\phi_{20}+\phi_{10})=0$$

$$\mu_2\eta\,\omega_{20}A_{20}-\frac{1}{4}\alpha_{24}A_{10}^3\sin(\phi_{20}-3\phi_{10})$$

$$-\left(\frac{3}{4}\alpha_{24}F_{10}^2+\frac{1}{2}\alpha_{23}F_{10}F_{20}\right.$$

$$\left.+\frac{1}{4}\alpha_{22}F_{20}^2\right)A_{10}\sin(\phi_{20}+\phi_{10})=0$$

$$2\eta^2\omega_{20}\omega_{21}A_{20}-\frac{3}{4}\alpha_{21}A_{20}^3-\frac{1}{2}\alpha_{23}A_{10}^2A_{20}$$

$$-\frac{1}{4}\alpha_{24}A_{10}^3\cos(\phi_{20}-3\phi_{10})$$

$$-\left(\frac{1}{2}\alpha_{23}F_{10}^2+\alpha_{22}F_{10}F_{20}+\frac{3}{2}\alpha_{21}F_{20}^2\right)A_{20}$$

$$-\left(\frac{3}{4}\alpha_{24}F_{10}^2+\frac{1}{2}\alpha_{23}F_{10}F_{20}\right.$$

$$\left.+\frac{1}{4}\alpha_{22}F_{20}^2\right)A_{10}\cos(\phi_{20}+\phi_{10}). \tag{37}$$

Figures 5–8 show the undamped combination resonance curves plotted against $\Omega/2\Omega_1$ from the solutions of equations (37), with $\mu_1=\mu_2=0$ and $f_{10}=0.5, 2, f_{20}=0.5, 2$, respectively. Obviously, the curves are symmetrical with the frequency axis.

From Figs. 5–8, some interesting phenomena can be listed as follows:

(1) The combination resonance only occurs when $\Omega>2\Omega_1$, i.e., the exciting frequency is greater than twice the linear fundamental frequency Ω_1.

(2) There are two branches of solution which indicate that two different responses may exist beyond certain exciting frequency.

(3) The response is also excitation-level dependent. For higher level excitation, A_{10} and A_{20} are dominant in different response curves, respectively. However, for lower-level excitation, the response curves become more involved. The dominance of modes may exchange in a single response curve as energy transfers from mode to mode occurring at a certain frequency range.

(4) Since the first mode vibrates at a frequency of $\Omega/2$ and the second mode at $3\Omega/2$, the first-mode dominant response can be regarded as a subharmonic resonance of order 1/2, while the second-mode dominant response has a superharmonic resonance of order 3/2.

Figures 9 and 10 show the combination resonance response curves with damping ratios $\mu_1=\mu_2=0.001$ and 0.003, respectively. They are all plotted against $\Omega/2\Omega_1$ from the solution of

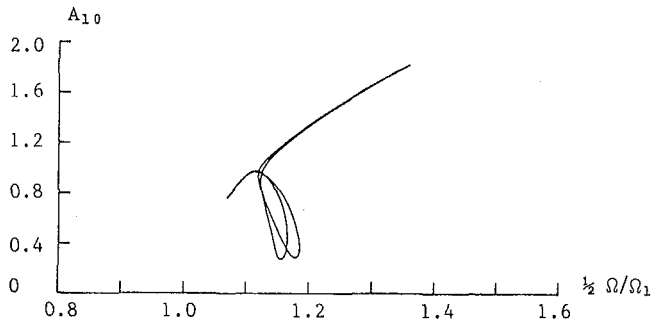


Fig. 9(a)

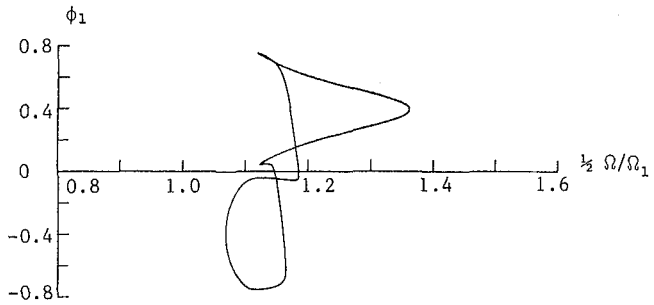


Fig. 9(b)

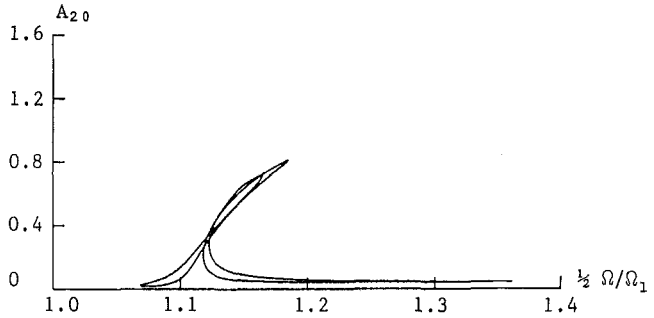


Fig. 9(c)

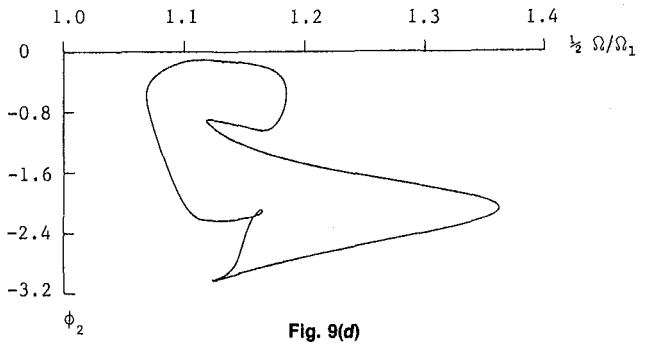


Fig. 9(d)

Fig. 9 Damped combination response with internal resonance $T=1/2(\tau_1+\tau_2)$, $\tau_2=3\tau_1$, $f_{10}=f_{20}=0.5$, $\mu_1=\mu_2=0.001$; (a) Frequency-amplitude curve of first harmonic term, (b) Frequency-phase curve of first harmonic term, (c) Frequency-amplitude curve of third harmonic term, and (d) Frequency-phase curve of third harmonic term

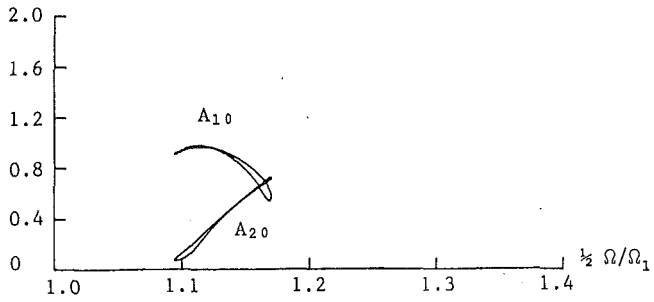


Fig. 10 Damped combination response with internal resonance $T = 1/2$ $(\tau_1 + \tau_2)$, $\tau_2 = 3\tau_1$, $f_{10} = f_{20} = = 0.5$, $\mu_1 = \mu_2 = 0.003$; A_{10} = amplitude of first harmonic term; A_{20} = amplitude of third harmonic term

equation (37) with the forcing term $f_{10} = f_{20} = 0.5$. It is interesting to note that the originally separate response curves in undamped systems become connected together to form a complex loop due to the influence of damping. Figure 10 indicates that the response loops shrink rather rapidly when increasing the damping ratio. This fact implies that there must be a critical damping ratio where the combination resonance may be completely suppressed.

IV Concluding Remarks

(1) An alternative perturbation procedure of multiple scales for nonlinear dynamic systems is presented. It is capable of treating periodic and almost-periodic steady-state vibrations for multiple DOF systems, with various resonances including the combination resonance, which the conventional Lindstedt-Poincaré method cannot apply. Obviously, this method can be further generalized.

(2) The aim of this paper is to introduce the essence of the method, so only solutions of nonlinear responses are considered. However, to fully understand the complete picture of nonlinear vibration behavior of the system, the inclusion of stability analysis is definitely necessary. In fact, for a given periodic vibration, the stability analysis can be carried out by considering the corresponding variational differential equations and applying the Floquet theorem. Hsu (1972, 1973, 1974) has developed an efficient method for approximating the transition matrix of the variational equations, during one period, by a series of step functions. Hsu's is a very convenient method to be implemented on a computer.

References

Hsu, C. S., 1972, "Impulsive Parametric Excitation: Theory," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 551–558.

Hsu, C. S., 1974, "On Approximating a General Linear Periodic System," *J. Math Analysis and Applic.*, Vol. 45, pp. 234–251.

Hsu, C. S., and Cheng, W. H., 1973, "Applications of the Theory of Impulsive Parametric Excitation and New Treatments of General Parametric Excitation Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, pp. 78–86.

Iu, V. P., 1985, "Nonlinear Vibration Analysis of Multilayer Sandwich Structure by Incremental Finite Elements," Ph.D. Thesis, Univ. of Hong Kong.

Lau, S. L., Cheung, Y. K., and Wu, S. Y., 1983, "Incremental Harmonic Balance Method With Multiple Time Scales for Aperiodic Vibration of Nonlinear Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 50, pp. 871–876.

Lau, S. L., Cheung, Y. K., and Wu, S. Y., 1984, "Nonlinear Vibration of Thin Elastic Plates, Part 2: Internal Resonance by Amplitude—Incremental Finite Element," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 845–851.

Mook, D. T., Plaut, R. H., and Haquang, N., 1985, "The Influence of an Internal Resonance on Non-Linear Structural Vibrations Under Subharmonic Resonance Conditions," *J. of Sound and Vibration*, Vol. 102, pp. 473–492.

Mook, D. T., Plaut, R. H., and Haquang, N., 1986, "The Influence of an Internal Resonance on Non-Linear Structural Vibrations Under Combination Resonance Conditions," *J. of Sound and Vibration*, Vol. 104, pp. 229–241.

Nayfeh, A. H., 1973, *Perturbation Methods*, John Wiley and Sons, New York.

Nayfeh, A. H., 1981, *Introduction to Perturbation Techniques*, John Wiley and Sons, New York.

Nayfeh, A. H., 1983, "Response of Two-Degree-of-Freedom System to Multifrequency Parametric Excitations," *J. of Sound and Vibration*, Vol. 88, pp. 1–10.

Nayfeh, A. H., 1984, "Combination Tones in the Response of Single-Degree-of-Freedom Systems With Quadratic and Cubic Non-Linearities," *J. of Sound and Vibration*, Vol. 92, pp. 379–386.

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

Nayfeh, A. H., Mook, D. T., and Sridhar, S., 1974, "Nonlinear Analysis of the Forced Response of Structural Elements," *J. Acoust. Soc. Am.*, Vol. 55, pp. 281–291.

Sridhar, S., Mook, D. T., and Nayfeh, A. H., 1975, "Nonlinear Resonances in the Forced Response of Plates, Part 1: Symmetric Response of Circular Plates," *J. of Sound and Vibration*, Vol. 41, pp. 359–373.

Sridhar, S., Mook, D. T., and Nayfeh, A. H., 1978, "Nonlinear Resonances in the Forced Response of Plates, Part 2: Asymmetric Responses of Circular Plates," *J. of Sound and Vibration*, Vol. 59, pp. 159–170.

Sridhar, S., Nayfeh, A. H., and Mook, D. T., 1975, "Nonlinear Resonances in a Class of Multi-Degree-of-Freedom Systems," *J. Acoust Soc. Am.*, Vol. 58, pp. 113–123.

APPENDIX

$$\Omega_1 = 1, \qquad \Omega_2 = 3.2406$$

$$\alpha_{11} = 0.278769, \qquad \alpha_{12} = -0.311074,$$

$$\alpha_{13} = 1.11585, \qquad \alpha_{14} = -0.386361,$$

$$\alpha_{21} = 3.87030, \qquad \alpha_{22} = 3\,\alpha_{14},$$

$$\alpha_{23} = \alpha_{13}, \qquad \alpha_{24} = \frac{1}{3}\,\alpha_{12},$$

$$C_{12}^{(1)} = \frac{\alpha_{11} A_{10}^3}{32\omega_{10}^2},$$

$$C_{13}^{(1)} = \frac{3\alpha_{14} A_{20}^3 + 2\alpha_{12} A_{10}^2 A_{20}}{4(\omega_{20}^2 - \omega_{10}^2)},$$

$$C_{14}^{(1)} = \frac{\alpha_{14} A_{20}^3}{4(9\omega_{20}^2 - \omega_{10}^2)},$$

$$C_{15}^{(1)} = \frac{\alpha_{12} A_{10}^2 A_{20}}{4(3\omega_{10}^2 + 4\omega_{10}\omega_{20} + \omega_{20}^2)},$$

$$C_{16}^{(1)} = \frac{\alpha_{12} A_{10}^2 A_{20}}{4(3\omega_{10}^2 - 4\omega_{10}\omega_{20} + \omega_{20}^2)}, \qquad C_{17}^{(1)} = \frac{\alpha_{13} A_{10} A_{20}^2}{16(\omega_{20}^2 + \omega_{10}\omega_{20})},$$

$$C_{18}^{(1)} = \frac{\alpha_{13} A_{10} A_{20}^2}{16(\omega_{20}^2 - \omega_{10}\omega_{20})},$$

$C_{22}^{(1)}$ to $C_{28}^{(1)}$ are in the same form as $C_{12}^{(1)}$ to $C_{18}^{(1)}$ except α_{1i}, A_{10}, A_{20}, ω_{10} and ω_{20} are replaced by α_{2i}, A_{20}, A_{10}, ω_{20} and ω_{10}, respectively.

$$C_{12}^{(2)} = \frac{-d_{12}}{8\omega_{10}^2}, \qquad C_{13}^{(2)} = \frac{d_{13}}{\omega_{10}^2 - \omega_{20}^2}, \qquad C_{14}^{(2)} = \frac{d_{14}}{\omega_{10}^2 - 9\omega_{20}^2},$$

$$C_{15}^{(2)} = -\frac{d_{15}}{3\omega_{10}^2 + 4\omega_{10}\omega_{20} + \omega_{20}^2},$$

$$C_{16}^{(2)} = -\frac{d_{16}}{3\omega_{10}^2 - 4\omega_{10}\omega_{20} + \omega_{20}^2},$$

$$C_{17}^{(2)} = -\frac{d_{17}}{4\omega_{10}\omega_{20} + 4\omega_{20}^2}, \qquad C_{18}^{(2)} = \frac{d_{18}}{4\omega_{10}\omega_{20} - 4\omega_{20}^2},$$

where

$$d_{12} = 18\omega_{10}\omega_{11} C_{12}^{(1)} - \frac{3}{4}\,\alpha_{11} A_{10}^2 (C_{11}^{(1)} + 2C_{12}^{(11)})$$

$$- \frac{1}{4}\,\alpha_{12} A_{10}^2 (C_{23}^{(1)} + 2C_{24}^{(1)}) - \frac{1}{2}\,\alpha_{12} A_{10} A_{20} (C_{15}^{(1)} + C_{16}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{10}A_{20}(C_{27}^{(1)} + C_{28}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{20}^2 C_{12}^{(1)} - \frac{3}{2}\alpha_{14}A_{20}^2 C_{24}^{(1)}$$

$$d_{13} = 2\omega_{20}\omega_{21}C_{13}^{(1)} - \frac{3}{4}\alpha_{11}A_{10}^2(2C_{13}^{(1)} + C_{15}^{(1)} + C_{16}^{(1)})$$

$$- \frac{1}{4}\alpha_{12}A_{10}^2(2C_{21}^{(1)} + C_{27}^{(1)} + C_{28}^{(1)})$$

$$- \frac{1}{2}\alpha_{12}A_{10}A_{20}(2C_{11}^{(1)} + C_{17}^{(1)} + C_{18}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{10}A_{20}(2C_{23}^{(1)} + C_{25}^{(1)} + C_{26}^{(1)})$$

$$- \frac{1}{4}\alpha_{13}A_{20}^2(3C_{13}^{(1)} + C_{14}^{(1)}) - \frac{3}{4}\alpha_{14}A_{20}^2(3C_{21}^{(1)} + C_{22}^{(1)})$$

$$d_{14} = 18\omega_{20}\omega_{21}C_{14}^{(1)} - \frac{3}{2}\alpha_{11}A_{10}^2 C_{14}^{(1)} \frac{1}{2}\alpha_{12}A_{10}^2 C_{22}^{(1)}$$

$$- \frac{1}{2}\alpha_{12}A_{10}A_{20}(C_{17}^{(1)} + C_{18}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{10}A_{20}(C_{25}^{(1)} + C_{26}^{(1)})$$

$$- \frac{1}{4}\alpha_{13}A_{20}^2(C_{13}^{(1)} + 2C_{14}^{(1)}) - \frac{3}{4}\alpha_{14}A_{20}^2(C_{21}^{(1)} + 2C_{22}^{(1)})$$

$$d_{15} = (8\omega_{10}\omega_{11} + 4\omega_{10}\omega_{21} + 4\omega_{20}\omega_{11} + 2\omega_{20}\omega_{21})C_{15}^{(1)}$$

$$- \frac{3}{4}\alpha_{11}A_{10}^2(C_{13}^{(1)} + 2C_{15}^{(1)}) - \frac{1}{4}\alpha_{12}A_{10}^2(C_{21}^{(1)} + 2C_{27}^{(1)})$$

$$- \frac{1}{2}\alpha_{12}A_{10}A_{20}(C_{11}^{(1)} + C_{12}^{(1)} + C_{17}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{10}A_{20}(C_{23}^{(1)} + C_{24}^{(1)} + C_{25}^{(1)})$$

$$- \frac{1}{4}\alpha_{13}A_{20}^2(2C_{15}^{(1)} + C_{16}^{(1)}) - \frac{3}{4}\alpha_{14}A_{20}^2(2C_{27}^{(1)} + C_{28}^{(1)})$$

$$d_{16} = (8\omega_{10}\omega_{11} - 4\omega_{10}\omega_{21} - 4\omega_{20}\omega_{11} + 2\omega_{20}\omega_{21})C_{16}^{(1)}$$

$$- \frac{3}{4}\alpha_{11}A_{10}^2(C_{13}^{(1)} + 2C_{16}^{(1)})$$

$$- \frac{1}{4}\alpha_{12}A_{10}^2(C_{21}^{(1)} + 2C_{28}^{(1)}) - \frac{1}{2}\alpha_{12}A_{10}A_{20}(C_{12}^{(1)} + C_{18}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{10}A_{20}(C_{23}^{(1)} + C_{24}^{(1)} + C_{26}^{(1)})$$

$$- \frac{1}{4}\alpha_{13}A_{20}^2(C_{15}^{(1)} + 2C_{16}^{(1)}) - \frac{3}{4}\alpha_{14}A_{20}^2(C_{27}^{(1)} + 2C_{28}^{(1)})$$

$$d_{17} = (2\omega_{10}\omega_{11} + 4\omega_{10}\omega_{21} + 4\omega_{20}\omega_{11} + 8\omega_{20}\omega_{21})C_{17}^{(1)}$$

$$- \frac{3}{4}\alpha_{11}A_{10}^2(2C_{17}^{(1)} + C_{18}^{(1)}) - \frac{1}{4}\alpha_{12}A_{10}^2(2C_{25}^{(1)} + C_{26}^{(1)})$$

$$- \frac{1}{2}\alpha_{12}A_{10}A_{20}(C_{13}^{(1)} + C_{14}^{(1)} + C_{15}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{10}A_{20}(C_{21}^{(1)} + C_{22}^{(1)} + C_{27}^{(1)})$$

$$- \frac{1}{4}\alpha_{13}A_{20}^2(C_{11}^{(1)} + 2C_{17}^{(1)}) - \frac{3}{4}\alpha_{14}A_{20}^2(C_{23}^{(1)} + 2C_{25}^{(1)})$$

$$d_{18} = (2\omega_{10}\omega_{11} - 4\omega_{10}\omega_{21} - 4\omega_{20}\omega_{11} + 8\omega_{20}\omega_{21})C_{18}^{(1)}$$

$$- \frac{3}{4}\alpha_{11}A_{10}^2(C_{17}^{(11)} + 2C_{18}^{(1)}) - \frac{1}{4}\alpha_{12}A_{10}^2(C_{25}^{(1)} + 2C_{26}^{(1)})$$

$$- \frac{1}{2}\alpha_{12}A_{10}A_{20}(C_{13}^{(1)} + C_{14}^{(1)} + C_{16}^{(1)})$$

$$- \frac{1}{2}\alpha_{13}A_{10}A_{20}(C_{22}^{(1)} + C_{28}^{(1)})$$

$$- \frac{1}{4}\alpha_{13}A_{20}^2(C_{11}^{(1)} + 2C_{18}^{(1)}) - \frac{3}{4}\alpha_{14}A_{20}^2(C_{23}^{(1)} + 2C_{26}^{(1)}).$$

$C_{22}^{(2)}$ to $C_{28}^{(2)}$ are in the same forms as $C_{12}^{(2)}$ to $C_{18}^{(2)}$ except ω_{10}, ω_{20}, and d_{1i} are replaced by ω_{20}, ω_{10}, and d_{2i} ($i = 2, \ldots 8$), respectively. Similarly, d_{22} to d_{28} are in the same form as d_{12} to d_{18}, except A_{10}, A_{20}, ω_{10}, ω_{20}, $C_{1i}^{(1)}$, $C_{2i}^{(1)}$, and α_{1i} are replaced by A_{20}, A_{10}, ω_{20}, ω_{10}, C_{2i}, C_{1i}, and α_{2i}, respectively.

Jan Parczewski

Wojciech Blajer

Department of Mechanics,
Technical University of Radom,
26-600 Radom, Poland

On Realization of Program Constraints: Part I—Theory

The problem of realization of program constraints is considered. The classical theory approach based on replacing the constriant reactions by adequate control forces has been generalized to the case when the control forces are not collinear with program constraint vectors or, in the extreme, when the control forces do not project in these directions at all (control forces are tangent to constraint manifolds). A classification of possible ways of program constraint realization is proposed and a general solution of the problem is presented.

1 Introduction

In the classical theory of mechanical systems an idea of reactions of ideal constraints is introduced. In principle, these reactions are postulated to be collinear with so-called "constraint vectors" or, in other words, orthogonal to constraint manifolds (see Arnold (1978), Gutowski (1971), Kamman and Huston (1984), Kane (1968), Nejmark and Fufajev (1972), Wang and Huston (1987), and Wittenburg (1977)). According to these works, the constraint reactions can be written in the matrix notation as

$$R = B\,\lambda, \tag{1}$$

where B is an $n \times m$ full-rank matrix of constraint vectors and λ is an m-dimensional vector of Lagrange multipliers associated with m constraints imposed on the system.

Equations of motion with the constraining forces (equation (1)) coupled with the equations of corresponding constraints then become the Lagrange's equations of first order. In a general case, when the dynamic equations in quasi-coordinates are considered, the full set of governing equations may be expressed as follows (refer also to Blajer (1988a,b), Hemami et al. (1979,1981), Lötstedt (1982), Nikravesh (1984,1985), and Wittenburg (1977)):

$$M(q)\dot{\omega} = f(\omega,q,t) + B\,\lambda, \tag{2a}$$

$$\dot{q} = g(\omega,q,t), \tag{2b}$$

$$B^T\,\dot{\omega} + b(\omega,q,t) = 0, \tag{2c}$$

where M is an $n \times n$ symmetric positive-definite matrix, $\omega = [\omega_1, \ldots, \omega_n]^T$ and $q = [q_1, \ldots, q_n]^T$ are vectors of quasi-velocities and generalized coordinates, respectively, t is time, f and g are n-dimensional vectors, and b is an m-dimensional vector. If the motion in generalized coordinates is considered, ω becomes the vector of generalized velocities and (2b) simplifies to $\dot{q} = \omega$.

The equation (2c) represents a set of m-constraint equations in the second-order kinematic form. They are linear in $\dot{\omega}$, i.e., B can be a function of ω, q, and t. Usually, however, constraints imposed on the system are of the form of geometric and/or first-order kinematic constraints, $u(q,t) = 0$ or $v(\omega,q,t) = 0$, respectively. Thus, in order to transform them to the form (2c), they must be differentiated with respect to time twice or once, respectively. As a result, according to the constraint form, the rows of B^T are

$$B^T = \begin{cases} (g_\omega u_q)^T = u_q^T g_\omega^T & \text{for } u(q,t) = 0, \quad (3a) \\ v_\omega^T & \text{for } v(\omega,q,t) = 0. \quad (3b) \end{cases}$$

The constraint vectors, defined by (2c) or (3), projected in the directions of quasi-coordinates $\pi = [\pi_1, \ldots, \pi_n]^T$, $\dot{\pi} = \omega$, are contained in B as columns. Evidently, if geometric or first-order kinematic constraints are imposed, their transformation to the form (2c) yields appropriate conditions imposed on the initial value problem of (2).

In applications, see Gear et al. (1985), Lötstedt et al. (1982,1986), Hemami et al. (1979,1981); the Lagrange multipliers are often eliminated from the equations (2), i.e.,

$$\lambda = -(B^T M^{-1} B)^{-1}(B^T M^{-1} f + b). \tag{4}$$

Then, the equation (2a) becomes

$$M\,\dot{\omega} = f + B(B^T M^{-1} B)^{-1}(B^T M^{-1} f + b) \tag{5}$$

and, including (2b), a set of $2n$ ordinary differential equations (ODEs) in $[\omega,q]^T$ is obtained.

The aforementioned approach and other techniques for solving (2), as well as other types of equations of motion used in analytical mechanics (Lagrange's equation of second order, Hamilton's canonical equations, Gibbs' equations, Kane's equations, . . .), all of them originate more or less directly from the fundamental postulate of orthogonality of constraint reactions (1) to the corresponding constraint manifolds. (See Arnold (1978), Hemami et al. (1979,1981), Kamman and Huston (1984), Nejmark and Fufajew (1972), and Wang and Huston (1987)). The constraints imposed on the system are called ideal constraints and usually refer to material con-

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 30, 1988; final revision, November 1, 1988.

straints, since the constraining forces represent the reactions of environment to the system. The situation may change, however, when a system with program constraints is considered.

Program constraints are meant here as requirements imposed on the system motion (see Blajer (1988a,b), Gutowski (1972), and Walker et al. (1984)). The motion consistent with these constraints (program motion) must be ensured by an exactly adopted model of control (program control). Control forces, however, being an inner feature of the system, may have arbitrary directions in relation to the constraint manifolds and may not be able to replace the constraint reactions in the sense of classical mechanics. As a consequence, the approach defined by (4) and (5) may fail because of the noninvertibility of a matrix corresponding to $B^TM^{-1}B$ (λ now denotes the vector of control forces).

The objective of this paper is to present possible ways of realization of program constraints and to formulate a general mathematical model for determination of control ensuring exact execution of assumed program. It will be shown that the program constraint realization is possible when control forces do not satisfy the condition of orthogonality to constraint manifolds and, in the extreme, when they are tangent to these manifolds. The results obtained generalize, to some extent, the classical theory of constrained dynamical systems.

2 Formulation of the Problem

An n-degree-of-freedom controlled system with m program constraints is considered. According to the formulation (2) of motion equations for a system with "material" constraints, the governing equations for a controlled mechanical system subjected to program constraints have been taken in the following form (refer also to Blajer (1988a,b) and Hemami et al. (1979,1981)):

$$M\,\dot{\omega} = f + A\,C, \tag{6a}$$

$$\dot{q} = g, \tag{6b}$$

$$B^T\,\dot{\omega} + b = 0, \tag{6c}$$

where $C=[C_1, \ldots, C_m]^T$ is a vector of control forces and A is a full-rank $n\times m$ matrix of control force representation in the n-dimensional space of π. Similarly, as in (2c), B is an $n\times m$ full-rank matrix of program constraint vectors, $m\le n$.

It is worth noting that the solution of (6) and the determination of control reactions ensuring the realization of program constraints by using the classical theory approach described in Section 1 will be possible as long as the matrix $B^TM^{-1}A$ is invertible. In many works such a condition is postulated *a priori* (or, simply, $A=B$) (see Do Sanh (1984), Gutowski and Radziszewski (1969), and Hemami et al. (1979,1981)), which is equivalent to the demand that the control reactions replace the reactions of program constraints treated as ideal "material" ones. In a general case, however, the directions of control forces, being an individual characteristic of the system, may have nothing to do with the directions of program constraint vectors contained in B as columns. Hence, the matrix $B^TM^{-1}A$ may be singular and the classical theory approach may be not valid. In this case the solution of the problem described needs a modified procedure.

Let us factorize the matrix B as follows:

$$B = A\,P + A^*Q = (A, A^*)\begin{bmatrix} P \\ Q \end{bmatrix}, \tag{7}$$

where the full-rank $n\times k$ ($k=n-m$) matrix A^* is an orthogonal complement of $M^{-1}A$ such that

$$(A^*)^T\,M^{-1}\,A = 0. \tag{8}$$

Table 1

1	$\operatorname{rank}(P)=m$ $\operatorname{rank}(Q)=0$	OR (orthogonal realization)
2	$\operatorname{rank}(P)=m$ $0<\operatorname{rank}(Q)=q\le m$	NOR (nonideal orthogonal realization)
3	$0<\operatorname{rank}(P)=p<m$ $0<\operatorname{rank}(Q)=q\le m$ $p+q\ge m$	MR (mixed realization)
4	$\operatorname{rank}(P)=0$ $\operatorname{rank}(Q)=m$ $2m\le n$	TR (tangent realization)

The $m\times m$ matrix P and $k\times m$ matrix Q can be found from the following relation:

$$\begin{bmatrix} P \\ Q \end{bmatrix} = [A, A^*]^{-1}B. \tag{9}$$

The factorization (7) consists in the determination of tangent and orthogonal components of constraint vectors B in relation to the directions of control forces C in the n-dimensional space. According to the ranks of P and Q, the following classification of program constraint realization is proposed (see Table 1).

Substituting $\dot{\omega}$ from (6a) into (6c), and considering (7) and (8), it can be found that

$$(P^TA^T + Q^T(A^*)^T)\times M^{-1}f + P^TA^TM^{-1}A\;C + b = 0. \tag{10}$$

Analyzing (10), the possibility of OR and NOR can be easily deduced since P and $A^TM^{-1}A$ are invertible and (10) can be solved univocally for C, i.e.,

$$C = -(A^TM^{-1}A)^{-1}(P^T)^{-1}((P^TA^T + Q^T(A^*)^T)M^{-1}f + b). \tag{11}$$

In the classical case of OR ($Q=0$ and $B=A$), the relation (11) transforms to (4) and C plays the role of the Lagrange multiplier vector.

In the cases of MR and TR, equation (10) cannot be solved for C—the matrix P is not invertible. In these cases, a special approach to the problem must be undertaken. However, prior to the presentation of the solution, another form of the governing equations of motion will be introduced. It will be of some use in further considerations.

Premultiplying (6a) by the full-rank matrix $(A, A^*)^T$ or, in other words, projecting the equations (6a) in the directions of vectors contained as columns in (A, A^*) and considering (8), it follows that

$$A^T\dot{\omega} = A^TM^{-1}f + A^TM^{-1}A\;C, \tag{12a}$$

$$(A^*)^T\dot{\omega} = (A^*)^T\,M^{-1}f. \tag{12b}$$

Now, the equations (12b), (6b), and (6c) form a new set of equations of program motion, the dimension of which is reduced to $2n$ (note that the dimension of (6) is $2n+m$). Since the matrix $A^TM^{-1}A$ is invertible, (12a) will serve only for determination of demanded control ensuring the exact realization of program constraints. The problem of control determination then becomes a secondary one and follows from the transient dynamic solution of program motion equations. Obviously, in the cases of OR and NOR, the program control will be found rather from (4) or (11) than from (12a).

The aforementioned approach is used in classical mechanics for reducing the equations of motion with constraints (see Amirouche et al. (1987,1988), Hemami et al. (1979,1981), Kamman and Huston (1984), and Wang and Huston (1987)),

and is known as the Maggi transformation (refer to Nejmark and Fufajew (1972)).

The set of program motion equations (12*b*), (6*b*), and (6*c*) can be converted to a standard ODE system in $[\omega,q]^T$ if

$$\operatorname{rank}\left(\begin{bmatrix}(A^*)^T\\ B^T\end{bmatrix}\right)=\max=n. \tag{13}$$

Considering that

$$\begin{bmatrix}(A^*)^T\\ B^T\end{bmatrix}=\begin{bmatrix}(A^*)^T\\ P^TA^T+Q^T(A^*)^T\end{bmatrix}=\begin{bmatrix}I & 0\\ 0 & P^T\end{bmatrix}\begin{bmatrix}(A^*)^T\\ A^T\end{bmatrix}+\begin{bmatrix}0\\ Q^T(A^*)^T\end{bmatrix}, \tag{14}$$

the existence of condition (13) for the cases of OR and NOR is evident. However, for the cases of MR and TR,

$$\operatorname{rank}\left(\begin{bmatrix}(A^*)^T\\ B^T\end{bmatrix}\right)<\max, \tag{15}$$

and the problem cannot be solved using the classical theory.

3 Solution of the Problem

As it was shown in Sections 1 and 2, for the cases of OR and NOR, the problem of control ensuring the realization of program of motion can be easily solved using the classical theory of constrained systems or a slightly modified (generalized) version of this theory. For the case of OR, the demanded control can be found from (4), where $C=\lambda$, and the governing equations of program motion are formed by (5) and (2*b*) (or (12*b*) and (6*c*), alternatively). For the case of NOR, the program control can be determined from (11), whereas the equations of program motion can be composed either of (6*a*) (after substituting C from (11)) and (6*b*) or of (12*b*) and (6*c*).

Let us concentrate now on the solution of the problem when the cases of MR and TR are faced. First, we will show that the case of MR can be transformed to a form equivalent to TR, and then the solution for both the cases will be presented.

For the purpose of the analysis, let us rewrite the reduced dynamic equation (12*b*) and the constraint equation (6*c*) as follows:

$$(A^*)^T\dot{\omega}=(A^*)^TM^{-1}f, \tag{16a}$$

$$(P^TA^T+Q^T(A^*)^T)\dot{\omega}+b=0. \tag{16b}$$

Denote now that for the case of MR, $\operatorname{rank}(P)=p=m-l$ (see Table 1). In this case, l rows of P are linearly dependent, i.e.,

$$U^TP^T=0, \tag{17}$$

where U is an $m\times l$ full-rank matrix. Let U^* be a complement of U in the m-dimensional space. Thus, $\operatorname{rank}((U^*)^TP^T)=\max=p$. Premultiplying now the equation (16*b*) by the matrix $(U^*, U)^T$, and considering (16*a*), the equations (16) can be manipulated to the following form:

$$(A^*)^T\dot{\omega}=(A^*)^TM^{-1}f, \tag{18a}$$

$$(U^*)^TP^TA^T\dot{\omega}=-(U^*)^T(b+Q^TA^*M^{-1}f), \tag{18b}$$

$$U^TQ^T(A^*)^T\dot{\omega}=-U^Tb. \tag{18c}$$

Since $\operatorname{rank}([(A^*)^T, (U^*)^TP^TA^T]^T)=\max=n-l$, the form (18) is structurally similar to (16) and the equations (16*b*) and (18*c*) represent the program constraints which are realized by "tangent" control forces. Substituting $\dot{\omega}$ from (6*a*), these equations can be rewritten in the generalized form as follows:

$$w(\omega,q,t)=\begin{cases}Q^T(A^*)^TM^{-1}f+b & (19a)\\ U^TQ^T(A^*)^TM^{-1}f+U^Tb. & (19b)\end{cases}$$

Note that the dimension of (19*b*), referring to the MR case, is $l=m-p$, whereas, the dimension of (19*a*) is m.

Taking into account the previous considerations, the schemes for solution of the problem for MR and TR cases are similar. In the following we will deal with the case of TR, where $P=0$, $\operatorname{rank}(Q)=m$, and $w=B^TM^{-1}f+b$. Now, the governing equations of program motion can be reformulated as a set of differential/algebraic equations (DAEs), i.e.,

$$(A^*)^T\dot{\omega}-(A^*)^TM^{-1}f=0, \tag{20a}$$

$$\dot{q}-g(\omega,q,t)=0, \tag{20b}$$

$$w(\omega,q,t)=0. \tag{20c}$$

According to the theory of DAEs, a range of ODE methods can be used to solve the problem stated in equations (20). The idea of using ODE methods for solving DAE systems directly was introduced by Gear (1971) and is based on the backward Euler method. The original algorithm analysis was performed under the assumption that the index of system is equal to one. For our case it means that the matrix $((A^*)^T, w_\omega^T)^T$ is nonsingular, where w_ω denotes the n×m Jacobian matrix. With some care, techniques based on this method can be constructed for solving DAE systems even if the index exceeds one (for details refer to Brenan (1983)). For the purpose of this paper, however, a reduction technique to rewrite the system in the form with lower index will be applied to get a set of DAEs with index equal to one.

The technqiue is based on the algorithms proposed by Gear (1984), Gear and Petzold (1984), Lötstedt and Petzold (1986), and, apart from reducing the index of systems, it is also useful for determining their index value. The applied algorithm can be stated as follows:

Algorithm 1. (1) Differentiate with respect to time the equation (20*c*) to get

$$w_\omega^T\dot{\omega}+w_q^T\dot{q}+w_t=w_\omega^T\dot{\omega}+h(\omega,q,t)=0. \tag{21}$$

(2) If the matrix $((A^*)^T, w_\omega^T)^T$ is nonsingular, then we are done (the equations (20*a*), (20*b*), and (21) can be transformed to a set of ODEs).

(3) Otherwise, premultiply the set of equations (20*a*) and (21) by a nonsingular n×n matrix R to zero out a maximal number of rows of $((A^*)^T, w_\omega^T)^T$ and permute the zero rows to the bottom to obtain

$$\begin{bmatrix}A_1\\ 0\end{bmatrix}\dot{\omega}+\begin{bmatrix}f_1(\omega,q,t)\\ w_1(\omega,q,t)\end{bmatrix}=0. \tag{22}$$

Now, apply the process to this new system, which is in the form of the equations (20*a*) and (20*c*).

Of course, by differentiating the algebraic equations a number of integration constants are introduced, which means that we must determine the correct initial conditions. This can be done by satisfying $w(\omega_o,q_o,0)=0$, $w_1(\omega_o,q_o,0)=0$ or other algebraic equations which may appear in the next steps of Algorithm 1.

4 General Remarks

A more detailed discussion of this formulation is provided at the end of Part II of this paper, where practical implications of the general theory are demonstrated. In this section we would like to emphasize only some interesting conclusions resulting from the analysis investigated in the previous section.

One of the most valuable results seems to be the generalization of the concept of constraint reactions. The mathematical

formulation presented in the paper can make it possible to analyze mechanical systems with so-called nonideal constraints (constraints with friction, for example). Moreover, a new type of constraint realization has been defined. It has been proved that the realization is also possible by constraining forces which are tangent to the constraint manifolds. Obviously, this refers mainly to a subclass of constraints—the program constraints, the realization of which is ensured by control reactions.

One may face difficulties in the determination of orthogonal complements of the matrix $M^{-1}A$ and U, and in the determination of the matrix R introduced in Algorithm 1. As it will be shown in Part II of the paper, the task seems quite simple for small systems. For large systems, the methods suggested by Kamman and Huston (1984) or by Wang and Huston (1987) may be valuable. The problem has not been considered in this paper.

In Section 3 the governing equations of program motion have been introduced in the form of the DAE system—the equations (16) and (18). One of the important characteristics that determines the behavior of DAEs is the index of the system. Following the definitions suggested by Gear (1984), Gear and Petzold (1984), and Lötstedt and Petzold (1986), the index of (16) equals zero for the cases of OR and NOR. In these cases the condition (13) is fulfilled and (16) can be transformed to a standard ODE form. For the cases of MR and TR, the corresponding index exceeds one and the proposed algorithm must be used. It might be worth noting that if the program of motion consists of any geometric constraints, they must be first differentiated twice to be transformed to the form (2*c*). Considering this the global index of the problem for the cases of MR and TR is at least four. Examples of such systems will be presented in Part II.

References

Amirouche, F. M. L., and Ider, S. K., 1988, "Determination of Constraint Forces in Multibody Systems Dynamics Using Kane's Equations," *Journal de Mécanique Theorique et Appliquée*, Vol. 7, No. 1, pp. 3–20.

Amirouche, F. M. L., and Jia, T., 1987, "Automatic Elimination of the Undetermined Multipliers in Kane's Equations Using a Pseudo Uptriangular Decomposition (PUTD) Method," *Computers and Structures*, Vol. 27, No. 2, pp. 203–210.

Arnold, V. I., 1978, *Mathematical Methods in Classical Mechanics*, Springer-Verlag, Berlin, Heidelberg, and New York.

Blajer, W., 1988a, "Modelling and Control of Systems with Program Constraints," *Advances in Modelling and Simulation*, AMSE Press, New York, Vol. 12, No. 2, pp. 53–63.

Blajer, W., 1988b, "Modelling of Aircraft Program Motion with Application to Circular Loop Simulation," *The Aeronautical Journal*, Vol. 92, August-September, pp. 289–296.

Brenan, K. E., 1983, "Stability and Convergence of Differentiation Approximations for Higher Index Differential-Algebraic Systems with Applications in Trajectory Control," Ph.D. Thesis, University of California at Los Angeles.

Do Sanh, 1984, "On the Motion of Controlled Mechanical Systems," *Advances in Mechanics*, Vol. 2, No. 7, pp. 3–23.

Gear, C. W., 1971, "The Simultaneous Numerical Solution of Differential-Algebraic Equations," *IEEE Transactions on Circuit Theory*, Vol. CT-18, pp. 89–95.

Gear, C. W., 1984, "Differential-Algebraic Equations," NATO ASI Series, Vol. F9, Springer-Verlag, Berlin and Heidelberg, pp. 323–334.

Gear, C. W., and Petzold, L. R., 1984, "ODE Methods for the Solution of Differential/Algebraic Systems," *SIAM Journal of Numerical Analysis*, Vol. 21, No. 4, pp. 716–728.

Gear, C. W., Leimkuhler, B., and Gupta, G. K., 1985, "Automatic Integration of Euler-Lagrange Equations with Constraints," *Journal of Computational and Applied Mathematics*, Vol. 12–13, pp. 77–90.

Gutowski, R., 1971, *Analytical Mechanics*, (in Polish), PWN, Warsaw.

Gutowski, R., and Radziszewski, B., 1969, "On the Exact Control Problems of Motion of Mechanical Systems with Program Constraints," *Nonlinear Vibration Problems*, No. 10, pp. 3011–3016.

Hemami, H., and Weimer, F. C., 1981, "Modelling of Nonholonomic Dynamic Systems with Applications," ASME JOURNAL OF APPLIED MECHANICS, Vol. 48, No. 1, pp. 177–182.

Hemami, H., and Wyman, B. F., 1979, "Indirect Control of the Forces of Constraint in Dynamic Systems," ASME *Journal of Dynamic Systems, Measurement, and Control*, Vol. 101, pp. 355–360.

Kamman, J. W., and Huston, R. L., 1984, "Dynamics of Constrained Multibody Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, No. 4, pp. 899–903.

Kane, T. R., 1968, *Dynamics*, Holt, Rinehart, and Winston, New York.

Lötstedt, P., 1982, "Mechanical Systems of Rigid Bodies Subjected to Unilateral Constraints," *SIAM Journal of Applied Mathematics*, Vol. 42, No. 2, pp. 281–296.

Lötstedt, P., and Petzold, L. R., 1986, "Numerical Solution of Nonlinear Differential Equations with Algebraic Constraints I: Convergence Results for Backward Differentiation Formulas," *Mathematics of Computations*, Vol. 46, No. 174, pp. 491–516.

Nejmark, J. I., and Fufajev, N. A., 1972, "Dynamics of Nonholonomic Systems," (in Russian), Mir, Moscow.

Nikravesh, P. E., 1984, "Some Methods for Dynamics of Constrained Mechanical Systems: a Survey," NATO ASI Series, Vol. F9, Springer-Verlag, Berlin and Heidelberg, pp. 351–368.

Nikravesh, P. E., Wehage, R. A., and Known, O. K., 1985, "Euler Parameters in Computational Kinematics and Dynamics. Part 1," ASME *Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 107, pp. 358–365.

Walker, R. A., Gupta, N. K., Duke, E. L., and Patterson, B., 1984, "Developments in Flight Test Trajectory Control," AIAA Paper No. 84-0240.

Wang, J. T., and Huston, R. L., 1987, "Kane's Equations with Undetermined Multipliers—Application to Constrained Multibody Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, No. 2, pp. 424–429.

Wittenburg, J., 1980, "Dynamics of Systems of Rigid Bodies," Teubner, Stuttgart.

Wojciech Blajer

Jan Parczewski

Department of Mechanics,
Technical University of Radom,
26-600 Radom, Poland

On Realization of Program Constraints: Part II—Practical Implications

Practical implications of a general mathematical model of realization of program constraints are investigated. Illustrative examples of different types of this realization are demonstrated and discussed. Some general conclusions concerning the problem of program constraint realization are drawn.

1 Introduction

A general methodology for analyzing the dynamics of controlled mechanical systems subjected to program constraints was presented in Part I of this paper. Different types of program constraint realization were classified there. A general mathematical model for determination of control forces ensuring the exact realization of program constraints and for formulation of governing equations of program motion were also contained there.

In Part II of this paper, practical implications of the general formulation are demonstrated and discussed. The examples of particle and aircraft trajectory motion are used to present different types of possible ways of program constraint realization by adequate control forces. Through these examples, the general formulation of Part I becomes clearer. The practical implications enabled us also to draw some general conclusions concerning the mathematical model of Part I. They are discussed at the end of this paper.

Many of the equations and definitions used in Part II refer closely to Part I of the paper. Reading both parts as a whole is suggested.

2 Particle Trajectory Motion

Let us consider a particle of mass m and charge q moving in the gravitational, electric, and magnetic fields. Let us limit ourselves to the case of planar motion and assume that the vector of gravity acceleration lies along the z-axis (downwards), the vector of electric force is parallel to the x-axis, and the vector of magnetic induction lies along the y-axis (see Fig. 1).

The governing equations of the particle motion expressed in the path axes are, as follows:

$$m\dot{V} = -aV^2 - mg\,\sin\gamma + qE\,\cos\gamma, \tag{1a}$$

$$mV\dot{\gamma} = -mg\,\cos\gamma - qE\,\sin\gamma + qVB_M, \tag{1b}$$

where V is the velocity, a is a constant value (aV^2 denotes the drag force), g is the acceleration due to gravity, E is the electric field intensity, B_M is the value of magnetic induction, and γ is the inclination angle of $\bar{V}$. The path axes have been chosen since, in these axes, it is convenient to formulate the orthogonal and tangent directions to the predetermined path (constraint).

The equations (1) must be completed with the kinematic equations:

$$\dot{x} = V\,\cos\gamma, \tag{2a}$$

$$\dot{z} = V\,\sin\gamma. \tag{2b}$$

Let us assume now that the particle is postulated to move along a prescribed trajectory (program constraint), the equation of which is

$$f(x,z) = 0. \tag{3}$$

Assuming that $f(x_o,z_o) = 0$ and $tg\gamma_o = -f_x(x_o,z_o)/f_z(x_o,z_o)$, the geometric constraint (3) can be transformed (after two differentiations) to the second-order kinematic form, i.e.,

$$V\dot{\gamma} - V^2\kappa = 0, \tag{4}$$

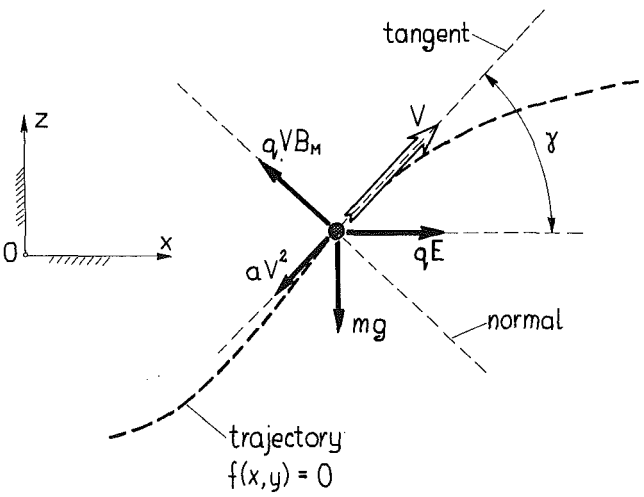


Fig. 1 Particle trajectory motion

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 30, 1988; final revision, November 1, 1988.

where $\kappa=(f_{xx}f_z^2+f_{zz}f_x^2)/(f_x^2+f_z^2)^{3/2}$ is the curvature of (3). Since, in the sense of equations (1) the quantity $V\dot{\gamma}$ should be interpreted as a quasi-acceleration (the corresponding quasi-velocity has no physical interpretation), the matrix M and the constraint vector B (see Part I) can be formulated as follows:

$$M=\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}, \quad B=\begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{5}$$

Example 1 (OR). Let us postulate that the particle motion along the trajectory (3) is ensured by changes in the Lorentz force qVB_M (control force) or, in other words, by changes in the value of magnetic induction B_M (control parameter). In accordance with the mathematical formulation of Part I, we can find that the matrix A of control force representation and the matrix A^* (orthogonal complement of $M^{-1}A$) are

$$A=\begin{bmatrix} 0 \\ 1 \end{bmatrix}; \quad A^*=\begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{6}$$

It is evident that $B=A$, $P=I$ (1×1 identity matrix), $Q=0$ and, consequently, we have a case of orthogonal realization (OR). Then, the actual value B_M ensuring the exact realization of the program (3) can be found from the relation following from (1*b*) and (4), i.e.,

$$B_M=\frac{mV^2\kappa+mg\cos\gamma+qE\sin\gamma}{qV}. \tag{7}$$

Taking into account equation (7), the dynamic equations (1) transform to

$$m\dot{V}=-aV^2-mg\sin\gamma+qE\sin\gamma, \tag{8a}$$

$$mV\dot{\gamma}=mV^2\kappa. \tag{8b}$$

Example 2 (NOR). Let us assume now that B_M is a suitably differentiable arbitrary function and that the particle trajectory motion is controlled by the electric force $q\,E$ (E is a control parameter). Then, the matrices A and A^* are

$$A=\begin{bmatrix} \cos\gamma \\ -\sin\gamma \end{bmatrix}; \quad A^*=\begin{bmatrix} \sin\gamma \\ \cos\gamma \end{bmatrix}. \tag{9}$$

Following the mathematical model proposed in Part I, we can find that

$$\begin{bmatrix} P \\ Q \end{bmatrix} = (A,A^*)^{-1}B=\begin{bmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{bmatrix}\times\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$=\begin{bmatrix} -\sin\gamma \\ \cos\gamma \end{bmatrix}. \tag{10}$$

Since, for $\sin\gamma\neq 0$, rank $(P)=1=\max$, and rank $(Q)>0$, a case of nonideal orthogonal realization (NOR) has been obtained. The actual value of control parameter E ensuring the particle trajectory motion can be found from the equation (10) of Part I, i.e.,

$$E=-\frac{mg\cos\gamma-qB_MV+mV^2\kappa}{q\sin\gamma}. \tag{11}$$

Note that for $\sin\gamma=0$, the realization of the control defined by (10) is impossible, and that for $\cos\gamma=0$, the problem becomes an OR case (rank $(Q)=0$).

Considering (11), the dynamic equations (1) become

$$m\dot{V}=-aV^2-mg\sin\gamma-(mg\cos\gamma-qB_MV+mV^2\kappa)ctg\gamma, \tag{12a}$$

$$mV\dot{\gamma}=mV^2\kappa. \tag{12b}$$

Comparing the equations (8) and (12), it is worth noting that in both cases (OR and NOR) the dynamic equation (1*b*) has been replaced by the same equation $mV\dot{\gamma}=mV^2\kappa$, which is equivalent to (4). This equation expresses the actual value of external force projections in the normal direction demanded for trajectory motion. From this condition the control forces qVB_N in Example 1 and qE in Example 2 are derived. Since the control reaction qVB_M is orthogonal to the trajectory, it can be interpreted as an ideal constraint reaction. The control reaction qE, however, is not "ideal" and gives projections in both normal and tangent directions. The tangent projection is represented in (12*a*) by the last factor.

Example 3 (OR). This example will also be classified as an OR case. The analysis will be of some use in further discussion.

In addition to the trajectory constraint (3), an additional constraint on particle velocity is postulated;

$$V=\phi(x,z,t). \tag{13}$$

If $V_o=\phi(x_o,z_o,0)$, the condition (13) can be replaced by its differentiated form

$$\dot{V}-b=0, \tag{14}$$

where $b=V\phi_x\cos\gamma+V\phi_z\sin\gamma+\phi_t$.

In the example considered now, the number of constraints equals the number of degrees-of-freedom. As a consequence, both of the dynamic equations (1) will be replaced by the constraint equations in the second-order kinematic form (4) and (14). Moreover, both control forces, qVB_M and qE, must be applied to ensure the realization of the program.

The matrix of constraint vectors can be written as

$$B=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{15}$$

and the matrix of control force representation

$$A=\begin{bmatrix} 0 & \cos\gamma \\ 1 & -\sin\gamma \end{bmatrix}, \tag{16}$$

where the vector of control forces is meant as $C=[qVB_M,qE]^T$. Note that the matrix A^* does not exist, so the matrix Q does not exist either.

It is easy to deduce that

$$P=A^{-1}B=\begin{bmatrix} tg\gamma & 1 \\ \cos^{-1}\gamma & 0 \end{bmatrix}\times\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$=\begin{bmatrix} 1 & tg\gamma \\ 0 & \cos^{-1}\gamma \end{bmatrix}, \tag{17}$$

and that rank $(P)=\max=2$, if only $\cos\gamma\neq 0$. Then, in accordance to the classification provided in Table 1 of Part I, a case of OR has been obtained.

The governing equations of the program motion can be composed now of the constraint conditions (4) and (14) and the kinematic equations (2). The actual values of control forces qVB_M and qE (or control parameters B_M and E) follow from the equation (11) of Part I, i.e.,

$$C=\begin{bmatrix} qVB_M \\ qE \end{bmatrix}$$

$$=\begin{bmatrix} mg\cos^{-1}\gamma + aV^2 tg\gamma + mV^2\kappa + mbtg\gamma \\ aV^2\cos^{-1}\gamma + mgtg\gamma + mb\cos^{-1}\gamma \end{bmatrix}, \quad (18)$$

where b is defined in (14).

Note that in this example, classified as an OR case, the realization of the program is not "ideal" in the sense of classical theory. Since $A \neq B$, the control forces qVB_M and qE cannot be, in principle, identified with ideal constraint reactions. It refers to the qE control force giving projections in directions of both constraint vectors defined in (15). Of course, for $\gamma=0$ or $\gamma=\pi$, the realization becomes "ideal", and for $\gamma=\mp\pi/2$, the realization defined by (18) is impossible.

3 Aircraft Trajectory Motion

An aircraft program flight along a prescribed trajectory may be used to introduce examples of the other types of program constraint realization classified in Part I. For a more detailed discussion of the problem of aircraft trajectory motion, the reader is referred also to Blajer et al. (1987, 1988a,b). Here, we will quote the governing equations of motion and discuss a simple case of planar motion.

Using similar notation as in Section 2, the governing equations are (see Fig. 2):

$$m\dot{V} = -\frac{1}{2}\rho\, S\, V^2\, c_D(\alpha) - mg\sin\gamma + T\cos\alpha, \quad (19a)$$

$$mV\dot{\gamma} = \frac{1}{2}\rho\, S\, V^2\, c_L(\alpha) - mg\cos\gamma + T\sin\alpha, \quad (19b)$$

$$J\dot{Q} = \frac{1}{2}\rho\, S\, V^2\, c_a\left(c_{mo}(\alpha,Q) + \frac{\partial c_m}{\partial \delta_H}\delta_H\right), \quad (19c)$$

where ρ is the air density (for simplicity $\rho=$const), S is the wing area, c_D and c_L are the drag and lift force coefficients, respectively, c_a is the mean chord value, c_m is the pitching moment coefficient, J is the aircraft moment of inertia, α is the value of attack, Q is the aircraft angular velocity, T is the thrust force, and δ_H is the elevator deflection.

The modeled aircraft can be controlled by changes in T and δ_H values. Assuming that c_m depends linearly on δ_H (see (19c)), the control forces are T and $1/2(\rho\, S\, V^2\, c_\alpha)\partial c_m/\partial\delta_H(\delta_H)$ (T and δ_H are control parameters).

The dynamic equations (19) must be completed with the kinematic differentional equations

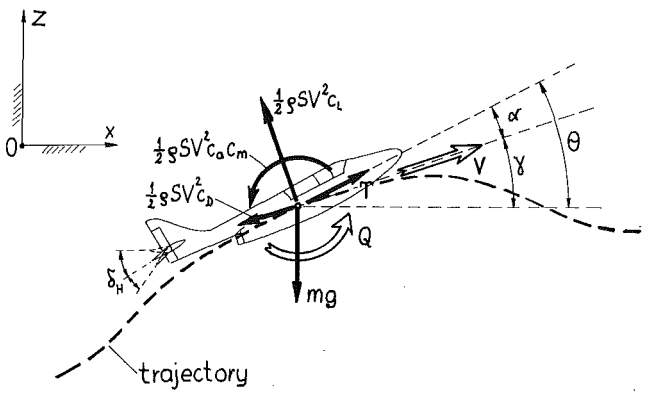


Fig. 2 Aircraft trajectory motion

$$\dot{x} = V\cos\gamma, \quad (20a)$$

$$\dot{z} = V\sin\gamma, \quad (20b)$$

$$\dot{\theta} = Q, \quad (20c)$$

and the geometric relation following from Fig. 2

$$\theta = \alpha + \gamma. \quad (21)$$

Example 4 (TR). Let us assume first that the aircraft is postulated to fly along the prescribed trajectory (3), and that the program motion is ensured by changes in δ_H value (T is assumed to be an arbitrary continuous function, for simplicity $T=$const). Using the mathematical model proposed in Part I, the matrices M, B, A and A^* can be stated as follows:

$$M=\begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & J \end{bmatrix}, \quad B=\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad (22)$$

$$A=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad A^*=\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (23)$$

Then, it can be easily deduced that

$$\begin{bmatrix} P \\ Q \end{bmatrix} = (A, A^*)^{-1}B$$

$$=\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (24)$$

Since $P=0$ and rank $(Q)=1$, a case of tangent realization (TR) has been obtained. Using the mathematical model proposed in Section 3 of Part I, one can easily prove that the equation (19c) will serve only for determination of demanded control, whereas (19a) and (19b) must be combined with the twice differentiated form (4) of the constraint (3). To solve the problem the algorithm introduced in Part I has to be used and the differential condition (4) must be transformed to the following algebraic form

$$w = \frac{1}{2}\rho\, S\, V^2\, c_L - mg\cos\gamma + T\sin\alpha - mV^2\kappa = 0, \quad (25)$$

which follows by substituting $\dot{\gamma}$ from (19b).

The governing equations of the program motion are composed now of (19a), (19b), (20), and (25), and form a set of DAEs with the state vector $[\omega^T, q^T]^T=[V,\gamma,Q,x,z,\theta]^T$. Since the index of the DAE system exceeds one, the following transformation have to be used.

Differentiating with respect to time, and taking into account (20c) and (21), it can be found that

$$\dot{w} = w_v\dot{V} + \left(\tfrac{1}{2}\rho\, S\, V^2\frac{\partial c_L}{\partial\alpha} + T\cos\alpha\right)Q + w_\gamma\dot{\gamma} = 0, \quad (26)$$

where $w_v=\rho\, S\, V\, c_L - 2mV\kappa$, $w_\gamma = -\rho\, S\, V^2(\partial c_L/\partial\alpha) - T\cos\alpha + mg\sin\gamma$. One can easily prove now that the matrix

$$((A^*)^T, w_\omega^T)^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_v & w_\gamma & 0 \end{bmatrix} \tag{27}$$

is singular, the index of the DAE system exceeds one, and that the Algorithm 1 proposed in Part I has to be used to solve the problem. Using the nomenclature proposed there, the matrix R can be defined as

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -w_v & -w_\gamma & 1 \end{bmatrix} \tag{28}$$

to obtain

$$\begin{bmatrix} A_1 \\ 0 \end{bmatrix} = R((A^*)^T, w_\omega^T)^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{29}$$

Premultiplying the set of equations (19*a*), (19*b*), and (26) by R, we can find a new set of DAEs of the form

$$\dot{V} = \frac{1}{m}\left(-\frac{1}{2}\rho\, S\, V^2\, c_D - mg\, \sin\gamma + T\cos\alpha\right), \tag{30a}$$

$$V\dot{\gamma} = \frac{1}{m}\left(\frac{1}{2}\rho\, S\, V^2\, c_L - mg\, \cos\gamma + T\sin\alpha\right), \tag{30b}$$

$$w_1 = -\frac{w_v}{m}\left(\frac{1}{2}\rho\, S\, V^2\, c_D + mg\, \sin\gamma - T\cos\alpha\right) + \frac{w_\gamma}{m_v}\left(\frac{1}{2}\rho\, S\, V^2\, c_L - mg\, \cos\gamma + T\sin\alpha\right) + \left(\frac{1}{2}\rho\, S\, V^2 \frac{\partial c_L}{\partial\alpha} + T\cos\alpha\right) Q = 0. \tag{30c}$$

Now, it is easy to prove that the index of the DAE system (30) is an equal one, since $\partial w_1/\partial Q \neq 0$, as far as $\frac{1}{2}\rho\, S\, V^2(\partial c_L/\partial\alpha) + T\cos\alpha \neq 0$. Obviously, by differentiating (30*c*) once more, the system (30) can be, in principle, standardized to an ODE system. In this case the initial values must satisfy the conditions $w(V_o, \gamma_o, \theta_o) = 0$, and $w_1(V_o, \gamma_o, Q_o, \theta_o) = 0$.

In practical applications of the problem described, the mathematical formulation can be considerably simplified. Since (19*c*) serves only for determination of the actual value of δ_H ensuring the exact trajectory flight, the set of DAEs describing the program motion can be stated as being composed of (19*a*), (4), (25), and (21) (The initial position of the aircraft must satisfy (3) and the vector of V_o must be tangent to the trajectory.) The state vector of the DAE system is $[V, \gamma, \alpha, \theta]^T$ and its index is equal to one. As a consequence, a standard ODE method can be used to solve this DAE system (Gear (1971, 1984), Brenan (1983)). Demanded values of δ_H can then be found from (19*c*), where $\dot{Q} = \ddot{\alpha} - \ddot{\gamma}$, and $\ddot{\alpha}$ and $\ddot{\gamma}$ can be determined by numerical differentiations of α and γ obtained from the solution of the DAE system (for details see Blajer and Parczewski (1987)).

Example 5 (MR). In the last example, the case of MR will be presented. Assume that in the trajectory flight, as in the previous example, an additional constraint (13) is imposed on the value of the aircraft velocity. Now, the control by changes in δ_H and T values must be applied, and the corresponding matrices B, A, and A^* are defined as follows (the matrix M has been already defined in Example 4):

$$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{31}$$

$$A = \begin{bmatrix} \cos\alpha & 0 \\ \sin\alpha & 0 \\ 0 & 1 \end{bmatrix}, \qquad A^* = \begin{bmatrix} -\sin\alpha \\ \cos\alpha \\ 0 \end{bmatrix}, \tag{32}$$

and $C = [T, (1/2)\rho\, S\, V^2\, c_\alpha(\partial c_m/\partial\delta_H)\delta_H]^T$. The matrices P and Q can be defined then as

$$\begin{bmatrix} P \\ Q \end{bmatrix} = (A, A^*)^{-1} B$$

$$= \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ 0 & 0 & 1 \\ -\sin\alpha & \cos\alpha & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sin\alpha & \cos\alpha \\ 0 & 0 \\ \cdots & \cdots \\ \cos\alpha & -\sin\alpha \end{bmatrix}. \tag{33}$$

Since there are two constraints imposed, rank $(P) = 1$ and rank $(Q) = 1$, the case of mixed realization (MR) has been obtained (see Table 1 in Part I).

Considering the theory proposed in Section 3 of Part I, the problem can be solved as follows, defining the full rank matrix (U^*, U) in the form

$$(U^*, U) = \begin{bmatrix} \sin\alpha & \cos\alpha \\ \cos\alpha & -\sin\alpha \end{bmatrix}, \tag{34}$$

one can easily prove that $U^T P^T = 0$. Then, the equations (18) of Part I take the following form

$$-\dot{V}\sin\alpha + V\dot{\gamma}\cos\alpha = \frac{1}{m}\left[\frac{1}{2}\rho\, S\, V^2(c_D\sin\alpha + c_L\cos\alpha) - mg\cos(\gamma+\alpha)\right], \tag{35a}$$

$$\dot{V}\cos\alpha + V\dot{\gamma}\sin\alpha = +V^2\kappa\sin\alpha + b\cos\alpha, \tag{35b}$$

$$\frac{1}{m}\left[\frac{1}{2}\rho\, S\, V^2(c_D\sin\alpha + c_L\cos\alpha) - mg\cos(\gamma+\alpha)\right] - V^2\kappa\cos\alpha - b\sin\alpha = 0. \tag{35c}$$

The aforementioned set of DAEs can be solved using the algorithm described in Section 3 of Part I and used previously in Example 4. The algebraic equation (35*c*) must be differentiated twice to get an ODE system or once to reduce the index of (35) to one. Obviously, appropriate conditions on the initial value problem have to be imposed.

The actual values of control forces (or control parameters) ensuring the exact realization of the program can then be found as

$$C = (A^T M^{-1} A)^{-1}(A^T\dot{\omega} - A^T M^{-1} f). \tag{36}$$

For the case considered, (36) takes the form

$$T = m\dot{V}\cos\alpha + mV\dot{\gamma}\sin\alpha - \frac{1}{2}\rho\, S\, V^2(c_L\sin\alpha - c_D\cos\alpha) + mg\sin(\alpha+\gamma), \tag{37}$$

$$\frac{1}{2}\rho S V^2 c_a \frac{\partial c_m}{\partial \delta_H}\delta_H = J\dot{Q} - \frac{1}{2}\rho S V^2 c_a c_{mo}. \quad (38)$$

In practical applications, the actual values of $\dot{V}$, $\dot{\gamma}$, and $\dot{Q}$ in (37) and (38) can be found numerically, following time histories of V, γ, and $\alpha(Q=\dot{\gamma}+\dot{\alpha})$.

4 Discussion

In this paper the number of control forces C has been assumed to be equal to the number of program constraints Φ imposed on the system. In many applications it may appear that dim $(C)>$ dim $(\Phi)=m$. In such a case, the m "controlling" forces must be chosen in advance, whereas the remaining ones can be taken as arbitrary, suitably differentiable functions of time and state variables.

The problem of "controlling" forces should be discussed more thoroughly. Let us consider again Example 4, which relates to the aircraft program flight along a prescribed trajectory. We have assumed that the aircraft is controlled by changes in δ_H value, and a case of TR has been deduced. However, if the aircraft would be controlled by changes in T value (δ_H is an arbitrary function now), we could find that

$$A=\begin{bmatrix}\cos\alpha\\ \sin\alpha\\ 0\end{bmatrix}, \qquad A^*=\begin{bmatrix}-\sin\alpha & 0\\ \cos\alpha & 0\\ 0 & 1\end{bmatrix}, \quad (39)$$

and then

$$\begin{bmatrix}P\\ Q\end{bmatrix}=(A,A^*)^{-1}B=\begin{bmatrix}\sin\alpha\\ \hline \cos\alpha\\ 0\end{bmatrix}. \quad (40)$$

Now, if $\sin\alpha\neq 0$, rank $(P)=1$ and rank $(Q)=1$, and the case can be classified as NOR. Theoretically then, as far as $\sin\alpha\neq 0$, the aircraft trajectory motion can be controlled by variations in T. However, the problem stated in such a way is ill-conditioned. The angle α is normally small, up to a dozen or so degrees, and T projects mainly in the tangent to the trajectory direction. Even small changes in "normal" projection of T demanded for the trajectory flight will cause big variations in V (see the equation (19a)), and then in other state parameter values. As a consequence, the motion controlled this way may be unstable and, possibly, unrealizable.

Let us assume now that the model of control introduced in Example 4 is slightly modified, i.e., assume that δ_H also affects c_L:

$$c_m=c_{mo}+c'_m\delta_H, \quad c_L=c_{Lo}+c'_L\delta_H, \quad (41)$$

where $c'_m=\partial c_m/\partial\delta_H$, $c'_L=\partial c_L/\partial\delta_H$, and c_{mo}, c_{Lo}, c'_m and c'_L do not depend on δ_H. In fact, c'_L is usually negligible and has been introduced here for theoretical considerations only. Denoting $C=(1/2)\rho S V^2\delta_H$, it follows that

$$A=\begin{bmatrix}0\\ c'_L\\ c_a c'_m\end{bmatrix}, \qquad A^*=\begin{bmatrix}1 & 0\\ 0 & m\,c_a c'_m\\ 0 & -Jc'_L\end{bmatrix}, \quad (42)$$

$$\begin{bmatrix}P\\ Q\end{bmatrix}=(A,A^*)^{-1}B=\frac{1}{d}\begin{bmatrix}0 & J c'_L & x\\ x & 0 & 0\\ 0 & c_\alpha c'_m & x\end{bmatrix}\begin{bmatrix}0\\ 1\\ 0\end{bmatrix}=$$

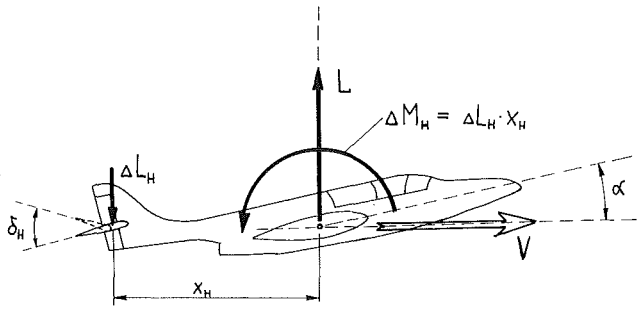


Fig. 3 Model of aircraft δ_H-control

$$=\frac{1}{d}\begin{bmatrix}J c'_L\\ 0\\ c_\alpha c'_m\end{bmatrix}, \quad (43)$$

where $d=\det(A,A^*)=m\,c_\alpha^2 c_m'^2+J\,c_L'^2$ and x denotes nonzero entries being of no use in the following.

Now, rank $(P)=$ rank $(Q)=1$ and, in principle, a case of NOR has been obtained. Then, the demanded values of δ_H ensuring the realization of program can be determined from the relation

$$\frac{1}{m}\left(\frac{1}{2}\rho S V^2 c_{Lo}-mg\cos\gamma+T\sin\alpha\right)+\frac{1}{m}\frac{1}{2}\rho S V^2 c'_L\delta_H-V^2\kappa=0. \quad (44)$$

The problem stated in this way also seems to be ill-conditioned. The value c'_L is usually negligible in comparison to the value of c'_m. Moreover, $c'_L<0$, thus, the variations of δ_H demanded for the trajectory motion through producing a gain in the lift force $\Delta L_H=1/2(\rho S V^2 c'_L\delta_H)$, will cause an increment of the pitching moment $\Delta M_H=1/2(\rho S V^2 c_c c'_m\delta_H)$, which in turn will cause the growth of both α and the lift force in the opposite direction to the "controlling" ΔL_H (see Fig. 3). Then, the simulated flight will be unstable and the proposed way of the program constraint realization is impossible in practice.

This discussion indicates that the choice of possible ways of program control needs a careful consideration, particularly when the orthogonal realization (OR) of program constraints imposed cannot be provided by the system.

References

Blajer, W., 1988a, "Modelling and Control of Systems With Program Constraints," *Advances in Modelling and Simulation*, Vol. 12, No. 2, ASME Press, New York, pp. 53–63.

Blajer, W., 1988b, "Modelling of Aircraft Program Motion With Application to Circular Loop Simulation," *The Aeronautical Journal*, Vol. 92, August-September, pp. 283-286.

Blajer, W., and Parczewski, J., 1987, "The Problem of Aircraft Program Motion Along an Assumed Spatial Trajectory," *Proceedings of the 8th Yugoslav Congress of Aerospace Sciences*, Vol. 2, Mostar, pp. 107–112.

Brenan, K. E., 1983, "Stability and Convergence of Difference Approximations for Higher Index Differential-Algebraic Systems With Applications in Trajectory Control," Ph.D. Thesis, University of California, Los Angeles.

Gear, C. W., "The Simultaneous Numerical Solution of Differential-Algebraic E Equations," *IEEE Transactions on Circuit Theory*, CT-18, pp. 89–95.

Gear, C. W., and Petzold, L. R., "ODE Methods for the Solution of Differential/Algebraic Systems," *SIAM Journal of Numerical Analysis*, Vol. 21, No. 4, pp. 716–728.

Effect of Correlation on the Almost-Sure Asymptotic Stability of Second-Order Linear Stochastic Systems

S. T. Ariaratnam

Wei Chau Xie

Solid Mechanics Division,
Faculty of Engineering,
University of Waterloo,
Waterloo, Ontario, Canada, N2L 3G1

A method of obtaining a sufficient almost-sure (a.s.) asymptotic stability condition for second-order, linear systems with both ergodic damping and stiffness coefficients is presented. The probabilistic property of the correlation between the damping and stiffness coefficients is taken into account. A sufficient condition for a.s. asymptotic stability is derived and numerical results are presented for the case of Gaussian noise coefficients. Results obtained in some of the previous investigations are included in the present study as special cases.

Introduction

The specific system considered is described by the second-order differential equation

$$\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0, \tag{1}$$

where $f(t)$, $g(t)$ are zero mean, ergodic stochastic processes and ζ is the damping coefficient.

When $f(t)$ and $g(t)$ are ergodic, wide-band Gaussian processes which may be approximated by white noise processes, the influence of the correlation between $f(t)$ and $g(t)$ on the almost-sure asymptotic stability of the system has been considered by Mitchell and Kozin (1974), who employed a method of Khas'minskii (1967) to obtain numerically the exact stability boundary. However, for arbitrary random excitation, this method is not applicable. In this paper, a method of obtaining a sufficient condition for a.s. asymptotic stability when the excitations $f(t)$, $g(t)$ are arbitrary ergodic, correlated or independent random processes is presented. Sufficient stability boundaries are obtained numerically in the case of Gaussian excitations.

Basic Equations

Consider a stochastic differential equation of the form

$$\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0, \tag{1}$$

where $f(t)$, $g(t)$ are ergodic processes with mean zero.

Ariaratnam and Ly (1989) considered $f(t)$ and $g(t)$ to be uncorrelated, and were able to get the best available results so far for the stability boundary. However, in general, $f(t)$ and $g(t)$ can be correlated random processes with correlation coefficient ρ.

In order to bring in the correlation effect, we first seek a transformation of the form

$$x=ye^{-\zeta t}, \tag{2}$$

which, when substituted into equation (1), yields

$$\ddot{y}+2f(t)\dot{y}+[c+h(t)]y=0, \tag{3}$$

where

$$c=1-\zeta^2,$$

$$h(t)=g(t)-2\zeta f(t).$$

Equation (3) can be written in the state equation form as

$$\begin{aligned}\dot{y}_1 &= y_2,\\ \dot{y}_2 &= -2f(t)y_2-[c+h(t)]y_1.\end{aligned} \tag{4}$$

The norm of the vector $\mathbf{y}=(y_1, y_2)$, $\|\mathbf{y}\|$, may be defined by

$$\|\mathbf{y}\|^2=V=\mathbf{y}^T\mathbf{A}\mathbf{y}, \tag{5}$$

where $\mathbf{A}$ is a positive-definite matrix given as

$$\mathbf{A}=\begin{bmatrix}\alpha_1^2 & \alpha_2\\ \alpha_2 & 1\end{bmatrix},\quad \alpha_1^2-\alpha_2^2>0.$$

Evaluation of $\dot{V}$ along the trajectories of (4) yields

$$\dot{V}=\mathbf{y}^T\mathbf{B}\mathbf{y}, \tag{6}$$

where

$$\mathbf{B}=\begin{bmatrix}-2\alpha_2(c+h) & \alpha_1^2-2\alpha_2 f-(c+h)\\ \alpha_1^2-2\alpha_2 f-(c+h) & 2\alpha_2-4f\end{bmatrix}.$$

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

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 23, 1988; final revision, December 22, 1988.

Therefore, since **A**, **B** are real symmetric matrices, and **A** is positive-definite,

$$\frac{\dot{V}}{V}=\frac{\mathbf{y}^T\mathbf{B}\mathbf{y}}{\mathbf{y}^T\mathbf{A}\mathbf{y}}\leq\lambda(\mathbf{B}\mathbf{A}^{-1}), \tag{7}$$

where λ is the maximum eigenvalue of $\mathbf{B}\mathbf{A}^{-1}$, i.e., λ is the maximum root of the determinantal equation

$$|\mathbf{B}-\lambda\mathbf{A}|=0. \tag{8}$$

Therefore, a sufficient condition for asymptotic stability with probability 1 (w.p.1) is given by (Infante (1968))

$$-2\zeta+E[\lambda(t)]<-\epsilon,\quad \epsilon>0. \tag{9}$$

Substituting **A** and **B** into equation (8) yields

$$|\mathbf{B}-\lambda\mathbf{A}|=\begin{vmatrix} -2\alpha_2(c+h)-\lambda\alpha_1^2 & \alpha_1^2-2\alpha_2 f-(c+h)-\lambda\alpha_2 \\ \alpha_1^2-2\alpha_2 f-(c+h)-\lambda\alpha_2 & 2\alpha_2-4f-\lambda \end{vmatrix}=0,$$

which gives

$$\lambda^2+4f\lambda-G/(\alpha_1^2-\alpha_2^2)=0, \tag{10}$$

where

$$G=(\alpha_1^2-c)^2+4c\alpha_2^2-4\alpha_2(\alpha_1^2+c)f+4\alpha_2^2f^2-4\alpha_2 fh+(4\alpha_2^2+2c-2\alpha_1^2)h+h^2,$$

so that one obtains the maximum eigenvalue as

$$\lambda=-2f+[4f^2+G/(\alpha_1^2-\alpha_2^2)]^{1/2}. \tag{11}$$

Substituting (11) into (9) leads to a sufficient asymptotic stability condition

$$-2\zeta+E[-2f+(4f^2+G/(\alpha_1^2-\alpha_2^2))^{1/2}]<-\epsilon,\quad \epsilon>0. \tag{12}$$

Since $f(t)$ has zero mean, this simplifies to

$$-2\zeta+E[(4f^2+G/(\alpha_1^2-\alpha_2^2))^{1/2}]<-\epsilon,\quad \epsilon>0. \tag{13}$$

Systems With Arbitrary Ergodic Coefficients

If no further information other than the mean and the variance of the processes $f(t)$ and $g(t)$ and their correlation coefficient ρ is available, one can apply the Schwarz inequality to (13) to obtain a stability boundary

$$E[4f^2+G/(\alpha_1^2-\alpha_2^2)]=4\zeta^2,$$

or

$$F\equiv-4\zeta^2(\alpha_1^2-\alpha_2^2)+4(\alpha_1^2-\alpha_2^2)\sigma_f^2+E[G]=0, \tag{14}$$

where G is given in (10), and $f(t)$ and $g(t)$ have zero mean. Then one can calculate $E[G]$ as

$$E[G]=(\alpha_1^2-c)^2+4c\alpha_2^2+4\alpha_2^2\sigma_f^2+H_1-4\alpha_2H_2, \tag{15}$$

where $H_1=E[h^2(t)]$, $H_2=E[f(t)h(t)]$, and equation (14) then becomes

$$F=-4\zeta^2(\alpha_1^2-\alpha_2^2)+4(\alpha_1^2-\alpha_2^2)\sigma_f^2+(\alpha_1^2-c)^2+4c\alpha_2^2+4\alpha_2^2\sigma_f^2+H_1-4\alpha_2H_2,$$

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

In order to obtain the optimal stability boundary, α_1^2 and α_2 are varied to get maximum σ_f and σ_g, which are given by

$$\frac{\partial\sigma_f}{\partial(\alpha_1^2)}=\frac{\partial\sigma_f}{\partial\alpha_2}=\frac{\partial\sigma_g}{\partial(\alpha_1^2)}=\frac{\partial\sigma_g}{\partial\alpha_2}=0.$$

These conditions turn out to be equivalent to (Ariaratnam and Ly (1989))

$$\frac{\partial F}{\partial(\alpha_1^2)}=\frac{\partial F}{\partial\alpha_2}=0. \tag{17}$$

Substituting (16) into (17) results in

$$\begin{aligned}\frac{\partial F}{\partial(\alpha_1^2)}&:-4\zeta^2+4\sigma_f^2+2(\alpha_1^2-c)=0,\\ \frac{\partial F}{\partial\alpha_2}&:8\zeta^2\alpha_2+8c\alpha_2-4H_2=0,\end{aligned} \tag{18}$$

from which the optimization parameters α_1^2, α_2 are obtained as

$$\alpha_1^2=1+\zeta^2-2\sigma_f^2,\quad \alpha_2=H_2/2. \tag{19}$$

Substituting (19) into (16) results in a sufficient asymptotic stability boundary

$$4(1+\zeta^2)\sigma_f^2-4\sigma_f^4+H_1-H_2^2-4\zeta^2=0. \tag{20}$$

Since $E[f(t)g(t)]=\rho\sigma_f\sigma_g$, one has

$$\begin{aligned}H_1&=E[h^2(t)]=E[g^2(t)-4\zeta f(t)g(t)+4\zeta^2f^2(t)],\\ &=\sigma_g^2-4\zeta\rho\sigma_f\sigma_g+4\zeta^2\sigma_f^2,\\ H_2&=E[f(t)h(t)]=E[f(t)g(t)-2\zeta f^2(t)],\\ &=\rho\sigma_f\sigma_g-2\zeta\sigma_f^2.\end{aligned}$$

Therefore, equation (20) becomes

$$(1-\rho^2\sigma_f^2)\sigma_g^2-4\zeta\rho(1-\sigma_f^2)\sigma_f\sigma_g+4(1-\sigma_f^2)[(1+\zeta^2)\sigma_f^2-\zeta^2]=0. \tag{21}$$

In the following, some particular cases are studied in detail.

(1) The case $f(t)\equiv 0$. Equation (1) becomes

$$\ddot{x}+2\zeta\dot{x}+[1+g(t)]x=0. \tag{22}$$

A sufficient a.s. asymptotic stability condition is

$$\sigma_g^2<4\zeta^2, \tag{23}$$

which is the same as that of Infante (1968).

(2) The case $g(t)\equiv 0$. Equation (1) becomes

$$\ddot{x}+2[\zeta+f(t)]\dot{x}+x=0. \tag{24}$$

A sufficient a.s. asymptotic stability condition is

$$(1-\sigma_f^2)[(1+\zeta^2)\sigma_f^2-\zeta^2]<0,$$

which gives

$$\sigma_f^2<\frac{\zeta^2}{1+\zeta^2}, \tag{25}$$

which is also the same as that of Infante (1968).

(3) The case $f(t)\not\equiv 0$, $g(t)\not\equiv 0$.

(i) $\rho=0$, i.e., f(t) and g(t) are uncorrelated. A sufficient a.s. asymptotic stability condition is, from (21),

$$\sigma_g^2+4(1-\sigma_f^2)[(1+\zeta^2)\sigma_f^2-\zeta^2]<0, \tag{26}$$

which is the same as that found by Ariaratnam and Ly (1989). The stability region defined by (26) is shown in Fig. 1(*a*) for various values of ζ.

(ii) $\rho\neq 0$, i.e., f(t) and g(t) are correlated. A sufficient

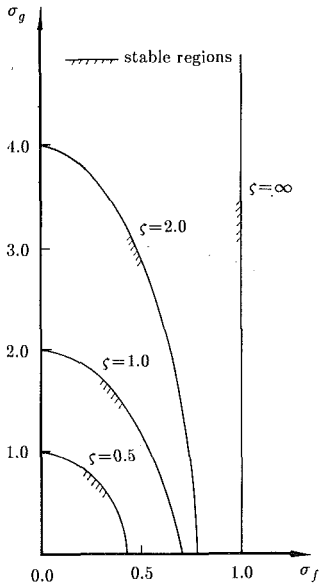


Fig. 1(*a*) Regions of almost-sure asymptotic stability for $\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0$ via Schwarz Inequality ($\rho_{fg}=0$)

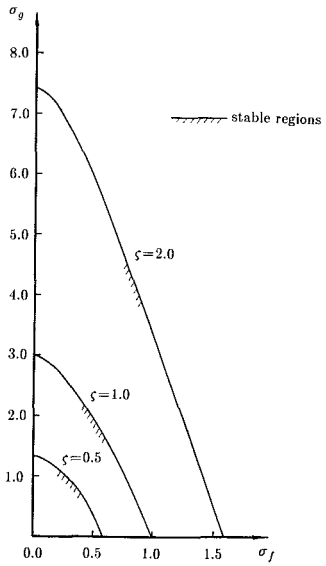


Fig. 1(*b*) Regions of almost-sure asymptotic stability for $\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0$ via Optimization Method ($\rho_{fg}=0$; $f(t)$, $g(t)$ Gaussian)

a.s. asymptotic stability boundary is given by the general form of (21). In Figs. 2 and 3, the stability boundaries are plotted for the cases when $\rho=\pm 0.5$, $\rho=\pm 1$, respectively, for different values of the damping parameter ζ.

It appears that the correlation coefficient of $f(t)$ and $g(t)$, ρ, has a definite influence on the sufficient asymptotic stability boundaries. When $f(t)$ and $g(t)$ are positively correlated, the stability boundaries are enlarged, while when they are negatively correlated, the effect is opposite. The larger the value of $|\rho|$, the larger is this effect.

It should be mentioned that the present results are only sufficient asymptotic stability conditions and can be improved further. In the special case when $\zeta=1$ and $f(t)=0.5g(t)$, it has been shown recently by Kozin (1988) in a private communication that the system is a.s. stable for any value of the variance parameter σ_g^2. Kozin's analysis is presented (with permission) in the Appendix.

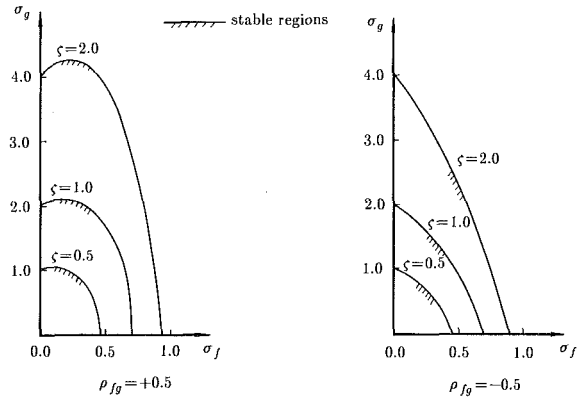


Fig. 2 Regions of almost-sure asymptotic stability for $\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0$ via Schwarz Inequality

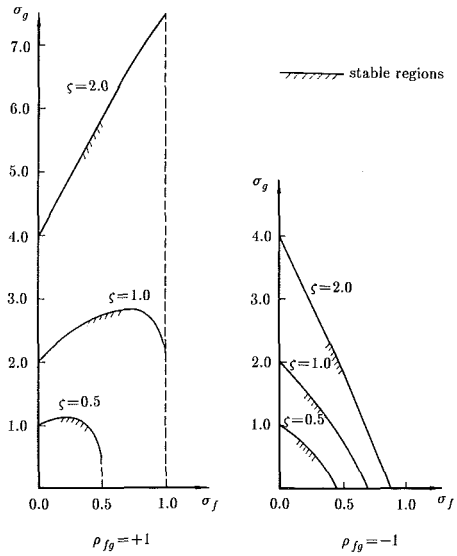


Fig. 3 Regions of almost-sure asymptotic stability for $\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0$ via Schwarz Inequality

These stability regions can be enlarged if the distributional properties of $f(t)$ and $g(t)$ are known. To show this improvement, the particular case of Gaussian excitation is considered in the following section.

Systems with Ergodic Gaussian Coefficients

(1) Optimization Model. Suppose that the joint probability density $p(f, g)$ of the random processes $f(t)$ and $g(t)$ is available, so that for any integrable function $F[f(t), g(t)]$,

$$E[F(f,g)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}F(f,g)p(f,g)dfdg, \tag{27}$$

and (13) can then be calculated numerically.

To obtain the maximum stability boundary, it is necessary that for fixed σ_f (at a given ζ), the parameters α_1 and α_2 be optimally chosen so as to get the maximum value of σ_g. Then one can construct the following optimization model subject to nonlinear constraints, namely

Maximize: $V=\sigma_g^2$,

Subject to Constraints:
$$\begin{cases}\alpha_1>0, \\ -\alpha_1<\alpha_2<\alpha_1, \\ -2\zeta+E[\{4f^2+ \\ \quad +G/(\alpha_1^2-\alpha_2^2)\}^{1/2}]>-\epsilon.\end{cases} \tag{28}$$

One convenient method of carrying out this optimization numerically is described in the following subsection.

(2) The Complex Method for Constrained Optimization. The complex method for constrained optimization, due to Box (1965), enables one to find the maximum of a multivariable, nonlinear function subject to nonlinear constraints, i.e.,

$$\textit{Maximize}: \quad F(X_1, X_2, \ldots, X_N),$$
$$\textit{Subject to Constraints}: \quad G_k \le X_k \le H_k,\ k = 1,2, \ldots, M. \tag{29}$$

The implicit variables $X_{N+1}, \ldots, X_M$ are dependent functions of the explicit independent variables $X_1, X_2, \ldots, X_N$. The upper and lower constraints H_k and G_k are either constants or functions of the independent variables.

The algorithm proceeds as follows (Richardson and Kuester (1973)):

(1) An original "complex" of $K \ge N+1$ points is generated consisting of a feasible starting point and $(K-1)$ additional points generated from random numbers and constraints for each of the independent variables as follows:

$$X_{i,j} = G_i + r_{i,j}(H_i - G_i), \quad i = 1,2, \ldots, N; \quad j = 1,2, \ldots, K-1,$$

where $r_{i,j}$ are random numbers between 0 and 1.

(2) The selected points must satisfy both the explicit and the implicit constraints. If at any step the explicit constraints are violated, the point is moved a small distance δ inside the violated limit. If an implicit constraint is violated, the point is moved one half of the distance to the centroid of the remaining points:

$$X_{i,j}(\text{new}) = [X_{i,j}(\text{old}) + \bar{X}_{i,c}]/2, \quad i = 1,2, \ldots, N,$$

where the coordinates of the centroid of the remaining points, $\bar{X}_{i,c}$, are defined by

$$\bar{X}_{i,c} = \frac{1}{K-1}\left[\sum_{j=1}^{K} X_{i,j} - X_{i,j}(\text{old})\right], \quad i = 1,2, \ldots, N.$$

This process is repeated until all the implicit constraints are satisfied.

(3) The objective function is evaluated at each point. The point having the lowest function value is replaced by its reflected point, chosen as follows:

The centroid of the remaining points

$$\bar{X}_{i,c} = \frac{1}{K-1}\left[\sum_{j=1}^{K} X_{i,j} - X_{i,j}(\text{lowest value})\right], \quad i = 1,2, \ldots, N,$$

is calculated so that the reflected point corresponding to the point having the lowest function value is given by

$$X_{i,j}(\text{new}) = \gamma(\bar{X}_{i,c} - X_{i,j}(\text{old})) + \bar{X}_{i,c}, \quad i = 1,2, \ldots, N.$$

A recommended value for γ is 1.3.

(4) If a point repeats in giving the lowest function value on consecutive trials, it is moved one half the distance to the centroid of the remaining points.

(5) The new point is checked against the constraints and is adjusted as before if the constraints are violated.

(6) Convergence is assumed when the objective function value at each point is within ϵ units for m consecutive iterations.

This method is a sequential search technique, which has been proven to be effective in solving problems with nonlinear objective functions subject to nonlinear inequality constraints. No derivatives are required. The procedure attempts to find the global maximum because the initial set of points is randomly scattered throughout the feasible region.

(3) Numerical Solution for Systems With Ergodic Gaussian Coefficients. Assume that $f(t)$ and $g(t)$ are jointly distributed Gaussian random processes with jointly probability density of the form

$$p(f,g) = \frac{1}{2\pi\sigma_f\sigma_g(1-\rho^2)^{1/2}} \exp\left\{-\frac{f^2}{2\sigma_f^2(1-\rho^2)} - \frac{g^2}{2\sigma_g^2(1-\rho^2)} + \frac{\rho fg}{\sigma_f\sigma_g(1-\rho^2)}\right\}, \tag{30}$$

so that for any integrable function $F(f,g)$

$$E[F(f,g)] = \int_{-\infty}^{+\infty} df \int_{-\infty}^{+\infty} F(f,g)p(f,g)\,dg. \tag{31}$$

Changing to new variables defined by

$$\xi = \frac{f}{[2(1-\rho^2)]^{1/2}\sigma_f}, \qquad \eta = \frac{g}{[2(1-\rho^2)]^{1/2}\sigma_g}, \tag{32}$$

equation (31) becomes

$$\frac{\pi}{(1-\rho^2)^{1/2}} E[F(f,g)] = \int_{-\infty}^{+\infty} \exp(-\xi^2)d\xi \int_{-\infty}^{+\infty} F(\sigma_f,\sigma_g,\rho,\xi,\eta)\exp(2\rho\xi\eta)\exp(-\eta^2)d\eta. \tag{33}$$

The R.H.S. of (33) can be calculated numerically by the double Gauss-Hermite integration formula. Therefore, the sufficient asymptotic stability condition (9) becomes

$$\frac{-2\pi\zeta}{(1-\rho^2)^{1/2}} + \int_{-\infty}^{+\infty} \exp(-\xi^2)d\xi \int_{-\infty}^{+\infty} \lambda(\sigma_f,\sigma_g,\rho,\xi,\eta) \times \exp(2\rho\xi\eta)\exp(-\eta^2)d\eta < -\epsilon,\ \epsilon > 0. \tag{34}$$

One can then construct the optimization model that may be solved by the complex method as follows (for any given σ_f):

$$\begin{aligned} &\textit{Maximize}: \quad V = \sigma_g^2, \\ &\textit{Subject to Constraints}: \\ &\quad 0.01 \le X_1 = \sigma_g \le 10.0, \\ &\quad 0.01 \le X_2 = \alpha_1 \le 5.0, \\ &\quad -|\alpha_1| + 0.01 \le X_3 = \alpha_2 \le |\alpha_1| - 0.01, \\ &\quad 0.0 \le X_4 \le 2\pi\zeta/(1-\rho^2)^{1/2}, \end{aligned} \tag{35}$$

where

$$X_4 = \frac{2\pi\zeta}{(1-\rho^2)^{1/2}} - \int_{-\infty}^{+\infty} \exp(-\xi^2)d\xi \int_{-\infty}^{+\infty} \lambda(\sigma_f,\sigma_g,\rho,\xi,\eta)\exp(2\rho\xi\eta)\exp(-\eta^2)d\eta,$$

and λ is given by (11).

In the optimization model (35), the values of the upper and lower constraints have been chosen for convenience of computation. Then one can solve the problem numerically by the complex method for constrained optimization.

For the case $\rho = 0$, i.e., $f(t)$ and $g(t)$ are independent stochastic processes, the results of numerical computation are plotted in Fig. 1(*b*) which are the same as those of Ariaratnam and Ly (1989), and Kozin and Milstead (1979).

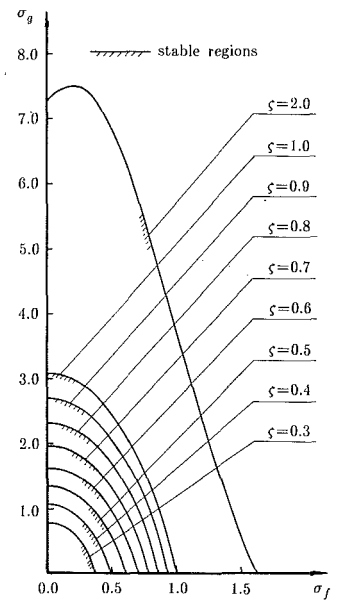


Fig. 4 Regions of almost-sure asymptotic stability for $\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0$ via Opimization Method ($\rho_{fg}=+0.5$; $f(t)$, $g(t)$ Gaussian)

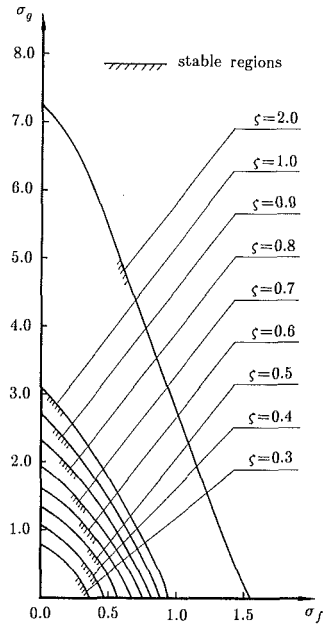


Fig. 5 Regions of almost-sure asymptotic stability for $\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0$ via Opimization Method ($\rho_{fg}=-0.5$; $f(t)$, $g(t)$ Gaussian)

For the case $\rho=\pm 0.5$, the results of the numerical computation are plotted in Fig. 4 and Fig. 5.

Since the results obtained by using the Schwarz inequality are sufficient stability conditions for arbitrary ergodic coefficients, they can be used as the original "complex" points, i.e., initial values in the numerical computation.

Obviously, within the stability domain, the positive-definiteness of matrix **A** is always satisfied, since it is guaranteed in the explicit constraints of the optimization model.

Conclusion

A method of obtaining a sufficient a.s. asymptotic stability condition for second-order systems with ergodic coefficients, which takes into account the correlation between the damping and stiffness coefficients, has been presented. A sufficient condition for stability has been derived and numerical results have been presented for the case of Gaussian noise coefficients, where some of the previous investigations were included as special cases.

It is obvious that the correlation coefficient, ρ, of $f(t)$ and $g(t)$, has a definite influence on the sufficient asymptotic stability boundaries. When $f(t)$ and $g(t)$ are positively correlated, the stability boundaries are enlarged, while when they are negatively correlated, the effect is opposite. The larger the value of $|\rho|$, the larger is this effect.

Mitchell and Kozin (1974) also found a definite effect on the stability boundaries due to correlation between excitations. Their results, which are necessary and sufficient, pertain to excitation by white noise, and cannot therefore be directly compared to the present results. For instance, the behavior near $\zeta=1$ and $\zeta>1$ for the case $\rho=+1$ is not obtainable by the present approach in the case of arbitrary ergodic excitations with finite variance.

Acknowledgments

This research was supported by the National Sciences and Engineering Research Council of Canada through Grant No. A-1815.

References

Ariaratnam, S. T., and Ly, B. L., 1989, "Almost-Sure Stability of Some Linear Stochastic Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 56, pp. 175–178.

Box, M. J., 1965, "A New Method of Constrained Optimization and a Comparison with Other Methods," *Comp. J.*, Vol. 8, pp. 42–52.

Infante, E. F., 1968, "On the Stability of Some Linear Nonautonomous Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 7–12.

Khas'minskii, R. Z., 1967, "Necessary and Sufficient Conditions for the Asymptotic Stability of Linear Stochastic Systems," *Theory of Probability and Its Applications*, Vol. 12, No. 1, pp. 144–147, (English translation).

Kozin, F., 1988, "Stability Region of Oscillator $\ddot{x}+2[\zeta+f(t)]\dot{x}+[1+g(t)]x=0$ for $\rho=1$," private communication.

Kozin, F., and Milstead, R. M., 1979, "The Stability of a Moving Elastic Strip Subjected to Random Parametric Excitation," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 404–410.

Kozin, F., and Wu, C.-M., 1973, "On the Stability of Linear Stochastic Differential Equations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 40, pp. 87–92.

Mitchell, R. R., and Kozin, F., 1974, "Sample Stability of Second-Order Linear Differential Equations with Wide-Band Noise Coefficients," *SIAM Journal of Applied Mathematics*, Vol. 27, No. 4, pp. 571–604.

Richardson, J. A., and Kuester, J. L., 1973, "The Complex Method for Constrained Optimization," *Communications of the ACM*, Vol. 16, No. 8, pp. 487–489.

APPENDIX

Consider system (1) with parameter $\zeta=1$ and $f(t)=0.5g(t)$, namely

$$\ddot{x}+[2+g(t)]\dot{x}+[1+g(t)]x=0, \qquad (A1)$$

where $g(t)$ is an ergodic process with zero mean and variance σ_g^2. System (A1) can be written as

$$\ddot{x}+\dot{x}+[1+g(t)](\dot{x}+x)=0,$$

which is reduced to a single first-order linear equation by setting $y=\dot{x}+x$, with $y_0=\dot{x}_0+x_0$,

$$\dot{y}(t)+[1+g(t)]y(t)=0. \qquad (A2)$$

The solution of (A2) is given by

$$y(t)=y_0e^{-t-G(t)}, \qquad (A3)$$

where $G(t)=\int_0^t g(s)ds$. By definition of y, one has

$$\dot{x}+x=y_0+e^{-t-G(t)},$$

whose solution is

$$\begin{aligned} x(t) &= e^{-t}\left[x_0+y_0\int_0^t e^{-G(s)}ds\right], \\ \dot{x}(t) &= e^{-t}\left[-x_0+y_0e^{-G(s)}-y_0\int_0^t e^{-G(t)}ds\right]. \end{aligned} \tag{A4}$$

It can be shown that $x(t)$ and $\dot{x}(t)$ approach zero as $t\to\infty$, w.p.1. Clearly x_0e^{-t} approaches zero. The integral $\int_0^t e^{-G(s)}ds$ is monotonic nondecreasing since $e^{-G(s)}\geq 0$. Therefore, the integral will approach a finite limit or $+\infty$, samplewise, w.p.1. If the sample limit is finite, then clearly $e^{-t}\int_0^t e^{-G(s)}ds$ approaches zero. On the other hand, if the sample integral approaches infinity, L'Hospital's rule is employed to yield

$$\lim_{t\to\infty}\frac{\int_0^t e^{-G(s)}ds}{e^t}=\lim_{t\to\infty}\frac{e^{-G(t)}}{e^t}=\lim_{t\to\infty}e^{-t-\int_0^t g(s)ds}.$$

Since the stochastic process $g(t)$ is ergodic with zero mean, one has a sample limit

$$\lim_{t\to\infty}e^{-t[1+\frac{1}{t}\int_0^t g(s)ds]}=0. \tag{A5}$$

Hence, from (A5) it is found that $\lim_{t\to\infty}x(t)=0$, w.p.1; and a similar results holds for $\dot{x}(t)$, which establishes the asymptotic stability for system (A1). Furthermore, it may be noted from (A5) that for any mean value $-1<E[g(t)]<\infty$, the system (A1) is almost surely stable for any value of σ_g^2.

Benson H. Tongue
Assistant Professor,
Department of Mechanical Engineering,
University of California,
Berkeley, CA 94720
Assoc. Mem. ASME

David Smith
Graduate Student,
School of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332

Determining Lyapunov Exponents by Means of Interpolated Mapping[1]

The method of Interpolated Mapping is extended to encompass the calculation of transient system behavior, specifically the efficient determination of the Lyapunov exponents for a simple nonlinear system. Both the continuous Lyapunov exponents as well as the corresponding Lyapunov exponents of the Poincaré map for a forced Duffing's oscillator are found. The use of Interpolated Mapping is compared to straightforward numerical integration and is shown to offer distinct computational advantages.

Introduction

Interpolated Mapping has been shown in previous works (Tongue, 1987; Tongue and Gu, 1988) to be a very efficient tool for the global analysis of nonlinear systems. The technique was shown to be suitable for determining the basins of attraction for a system's attractors and for conducting accurate fractal determinations. One of the attractive advantages of the method that has not yet been fully exploited is the ability to analyze the transient motion of any trajectory in phase space, thus permitting an examination of local stability characteristics. This paper will focus on the identification of a system's Lyapunov exponents (Benettin et al., 1980; Wolf et al., 1985) to illustrate this capability. As is well known, Lyapunov exponents measure the exponential rates of divergence or convergence associated with an attractor of a system. For periodic attractors, one obtains only negative and zero exponents, indicating convergence to a highly predictable motion, whereas a chaotic system will exhibit at least one positive exponent. A positive exponent is significant because it gives an indication of the rate at which one loses the ability to predict the system response. This is closely tied to the property of sensitive dependence on initial conditions which is present in chaotic systems. Therefore, one way to determine if a system is behaving in a chaotic manner is to calculate the Lyapunov exponents. A further motivation for calculating these exponents is that a knowledge of the full spectrum of Lyapunov exponents can be used to calculate an approximate value of the fractal dimension of the attractor (Farmer, Ott, and Yorke, 1983).

The chief problem with Lyapunov exponents is that they can be costly to calculate. Because the individual exponent calculations can vary widely over short time intervals, the exponents are defined as a long time average over the entire attractor. This forces long computer simulations that serve only to give the exponents for a given set of parameters with a particular group of initial conditions. This is obviously limiting when there is more than one attractor in the same region of phase space. Thus, there exists a need for calculating Lyapunov exponents cheaply and quickly. Ideally, one would want to be able to determine the exponents analytically from the differential equation. However, as this is not generally possible, an approximate technique would be helpful.

In this paper, it will be shown how some characteristics of the Lyapunov exponents can be calculated using both analytical and approximate methods. The methods will be applied to a general second-order Duffing's equation. In the first section, an analytical method for calculating the Lyapunov exponents of linear systems will be presented. Following this, some of these same techniques will be applied to nonlinear systems in order to determine what information about the exponents is available. Next, some of the numerical difficulties involved in a calculation of the exponents for nonlinear systems will be discussed. In the final sections, methods for calculating a Lyapunov exponent spectrum using the Interpolated Mapping technique will be proposed and the results compared to straightforward numerical integration.

Predicting Lyapunov Exponents

Linear Systems. It is quite simple to calculate the Lyapunov exponents of a linear system. Linear differential equations can frequently be solved exactly, and the exponents determined by an inspection of the solution. For example, the unforced problem

[1]This work was supported by the National Science Foundation, Grant No. MSM-8451186.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for presentation at the Winter Annual Meeting, San Francisco, Calif., December 10–15, 1989.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, September 21, 1988; final revision, December 22, 1988.

Paper No. 89-WA/APM-11.

$$\ddot{x} + 5\dot{x} + 6x = 0 \tag{1}$$

has the solution

$$x = Ae^{\lambda_1 t} + Be^{\lambda_2 t} \tag{2}$$

where the Lyapunov exponents, λ_1 and λ_2, are equal to -2 and -3, respectively.

Even for the case of a forced system such as

$$\ddot{x} + 5\dot{x} + 6x = G\cos(\omega t), \tag{3}$$

the same exponents are obtained, but with an additional exponent that has zero as its real part. This fact is obvious from the general solution of the problem, which is easily recognized as:

$$x = Ae^{\lambda_1 t} + Be^{\lambda_2 t} + F(e^{i\omega t} + e^{-i\omega t}). \tag{4}$$

Notice that the exponents are constants at all points in phase space (i.e., they do not depend upon the value of x). Further, because the forcing function does not alter the complementary solution, λ_1 and λ_2 are independent of G and ω. This makes the computation of the Lyapunov exponents for linear systems very easy.

Nonlinear Systems. Nonlinearities can cause a repeated stretching and folding of even small regions of phase space (Guckenheimer and Holmes, 1983) which causes the locally determined Lyapunov exponents to vary widely over the attractor. Thus, one must examine the long time average of the exponents. Furthermore, unlike the case for linear systems, the Lyapunov exponent behavior changes with forcing amplitude and frequency. Even for a relatively simple nonlinear differential equation, such as the Duffing equation, it is known that periodic as well as chaotic responses can be obtained for a given set of system parameters merely by changing the forcing amplitude and frequency. For example, the Duffing's equation:

$$\ddot{x} + .1\dot{x} - x + x^3 = 3.2\cos(\omega t) \tag{5}$$

yields a periodic response (two negative Lyapunov exponents) for $\omega = 0.482$, but exhibits chaotic behavior for $\omega = .475$ (Tongue, 1987). This indicates that the exponent has shifted from negative to positive over a very small change in ω.

Even though there is no analytical way to determine the Lyapunov exponents for a general system of equations, one can still obtain some information about the local rate of divergence (or convergence) experienced by perturbed trajectories from an attractor by examining the differential equations of motion. Consider the general Duffing's equation:

$$\ddot{x} + \alpha\dot{x} + \beta x + \gamma x^3 = G\cos(\omega t). \tag{6}$$

A standard linearization about a solution involves expressing x as the sum of $n(t)$ and $p(t)$, where n represents the nominal trajectory and p represents a local perturbation from this trajectory.

In this case the equation for p is

$$\ddot{p} + \alpha\dot{p} + (\beta + 3\gamma n^2(t))p = 0. \tag{7}$$

One can easily transform this into an undamped equation by expressing p as

$$p = ye^{-\frac{\alpha}{2}t} \tag{8}$$

in which case the equation for y is

$$\ddot{y} + \left(\beta - \frac{\alpha^2}{4} + 3\gamma n^2\right)y = 0. \tag{9}$$

For small intervals of time, the stiffness term of equation (9) is essentially constant, leading to solutions for y of the form

$$y_{1,2} = Ae^{\pm ct}. \tag{10}$$

It is therefore clear that the sum of the Lyapunov exponents for the given system are

$$\frac{-\alpha}{2} + b - \frac{\alpha}{2} - b = -\alpha. \tag{11}$$

This is true everywhere locally and so will be true over the long time average. This implies that no matter how widely the exponents vary, they must always vary symmetrically about $-.5\alpha$. The fact that $\lambda_1 + \lambda_2$, which governs the rate of phase space contraction, is equal to the negative of the damping coefficient can be used as a convenient means of checking the accuracy of the numerical routines. This property is a specific case of the more general observation that the sum of the eigenvalues of a system having a characteristic polynomial equal to

$$s^n + a_n s^{n-1} + \ldots + a_1$$

is simply $-a_n$.

The foregoing will be illustrated by a numerical calculation of the Lyapunov exponents for the following Duffing's equation:

$$\ddot{x} + .1\dot{x} - x + x^3 = 3.2\cos(.475t). \tag{12}$$

The general method of numerically calculating the exponents of a dynamical system proceeds in the following manner. First, a point that lies on the steady-state attractor of the system of interest is selected. Initially, a vector of magnitude ϵ and arbitrary direction is formed and placed with its base at the point on the trajectory. It is important that ϵ be small because nonlinearities cause, in general, a repeated stretching and folding of phase space, and only the stretching of the space is of interest here. A small test vector, as would be expected, is better able to avoid any effects of folding. A second vector, perpendicular to the first, but with equal magnitude, is also constructed at the test point. Additional vectors are added in a similar fashion until the vector set forms an orthonormal basis for the space in the region of the test point. The test point and the vector set are then integrated a short time into the future. The largest vector is used to calculate the largest local exponent from the equation

$$\lambda_1 = \frac{1}{\Delta t}\ln\left(\frac{l_f}{\epsilon}\right) \tag{13}$$

where Δt is the time interval over which the system was integrated and l_f is the length of the largest vector after integration. This vector will automatically tend toward the direction of maximum divergence (or minimum convergence). The second vector, however, is not free to tend toward the second greatest direction of divergence because of the effect of the largest exponent upon its direction. Thus, the second exponent is calculated through the calculation of the sum of the first two exponents, which measures the rate of contraction of a two-dimensional box in state space. This is governed by a similar equation:

$$\lambda_1 + \lambda_2 = \frac{1}{\Delta t}\ln\left(\frac{A_f}{\epsilon^2}\right) \tag{14}$$

where A_f is the final area of the space covered by the first two vectors. Subsequent exponent sums would be computed in a similar fashion for higher dimensional systems. The largest vector is then renormalized to a magnitude of ϵ, while its direction is preserved, enabling this vector to continue converging to the direction of the largest exponent. The remaining vectors are again constructed perpendicular to the first. This process is repeated over a long time interval and the exponents are calculated as a long time average over the steady-state motion. This long time average is extremely important, as even very close trajectories leading to periodic orbits can diverge from each other over short time intervals. Indeed, this phenomenon is what characterizes transient chaos.

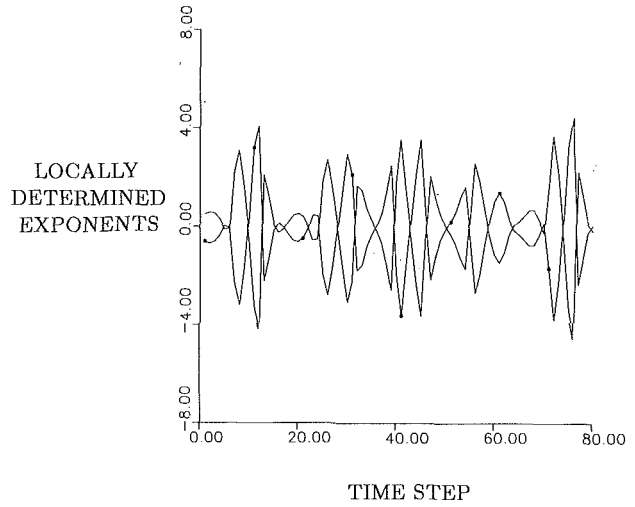


Fig. 1 Local variation of Lyapunov exponents for $\ddot{x} + .1\dot{x} - x + x^3 = 3.2 \cos(.475t)$

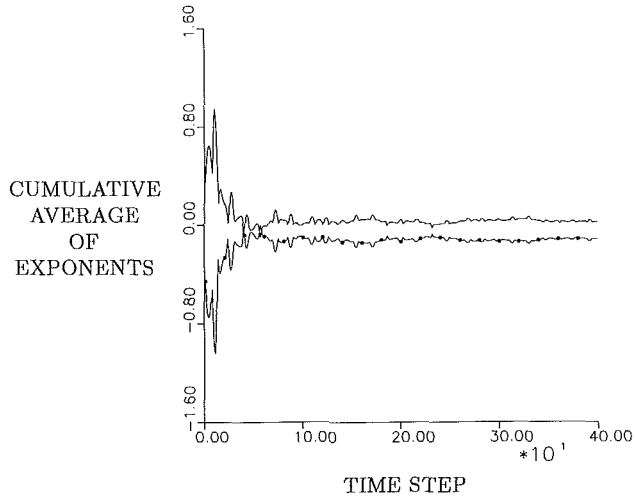


Fig. 2 Cumulative average of Lyapunov exponents for $\ddot{x} + .1\dot{x} - x + x^3 = 3.2 \cos(.475t)$

The above method is conceptually similar to that of Wolf et al. Figure 1 shows the local behavior of λ_1 and λ_2 for the system given by (12) as a function of time. For this graph, $\frac{1}{2}(\lambda_1 + \lambda_2) = -.0500$. Figure 2 shows the cumulative averages of λ_1 and λ_2. Note that, as expected, the sum is not changed by averaging. The analytical method derived here would indicate that $\frac{1}{2}(\lambda_1 + \lambda_2) = -.05$, a precise match.

Behavior of Phase Space in Lyapunov Exponent Calculation

Before attempting to calculate Lyapunov exponents for a point-to-point mapping such as the Interpolated Mapping technique, it is helpful to know what sort of behavior to expect from small regions of phase space. To examine this question, three different cases, each exhibiting period one behavior, were examined. The cases considered were: a forced linear oscillator, a forced Duffing's oscillator with positive linear stiffness, and a forced Duffing's oscillator with negative linear stiffness. Each case selected has a Poincaré map with a stable fixed point in phase space, and, thus, each orbit possesses two negative exponents.

The method used to examine the local region around the fixed point of the Poincaré map was as follows. A set of initial conditions was chosen and this point was numerically integrated for 1000 full forcing periods to ensure that the point was very close to the fixed point. This point was then surrounded with 32 points in a circular pattern at a small radius of ϵ. Each of these points was numerically integrated over one full forcing period and its final position reproduced on the same plot as the original circle. Thus, one is examining the behavior of the Poincaré map of the system. It is natural to expect that, for small enough ϵ, the mapped image will lie entirely within the original circle. Figures 3–6 show the actual results. Note that the axes have been shifted so that the fixed point is located at the origin.

For the linear case (Fig. 3), the first iterate map of the circular region lies well within the original circle. Even for the forced Duffing oscillator with positive linear stiffness (Fig. 4), the image does not leave the boundary of the original set. However, in the case of a forced Duffing's oscillator with negative linear stiffness (Fig. 5), most of the points map outside of the original circle. Smaller values of ϵ yield qualitatively identical results. As Fig. 6 shows, the entire region is asymptotically stable to the fixed point, so it must have two negative Lyapunov exponents. However, these exponents can only be determined as a long time effect.

At first it seems inconsistent to state that a stable system has points arbitrarily close to it that diverge. However, stability requires only that for every region U there exists a region W such that all future iterates of U remain in W (Lefschetz, 1977). Clearly one can draw a larger circle around the ellipse in Fig. 5 within which all subsequent iterates of the original circle will remain. This is the reason that the numerical technique used here requires that the direction of the largest exponent be allowed to evolve over time. Although this effect does not run counter to any established stability theory, it is certainly counterintuitive. An awareness of this sort of behavior is important to an understanding of the use of point-to-point mapping techniques in finding a set of exponents for a system.

Generating Lyapunov Exponents Using an Interpolated Trajectory

Interpolated Cell Mapping (ICM). This technique provides an efficient means of generating continuous system trajectories and allows one to obtain a variety of system characteristics (such as a system's attractor or a plot of the corresponding basins of attraction). The method has been discussed at length (Tongue, 1987), so only an overview will be given here. To utilize Interpolated Mapping, a region of phase space which is to be investigated is defined. One then overlays an array of points over the phase space. Each grid point is used as an initial condition. The system is then numerically integrated for a given length of time and the final location of the trajectory is recorded. Since the object of previous work was to examine Poincaré maps, the length of this integration was set equal to the period of the forcing function. This restriction will be removed in the work to be presented later in this paper. Once a mapping array has been generated, an arbitrary initial condition is chosen and its mapped location is found by interpolating between the terminal points of the trajectories emanating from the four initial condition grid points that surround the chosen point.

Calculation of the Lyapunov Exponents. To calculate the Lyapunov exponents of a full-time series using Interpolated Mapping requires a knowledge of the system behavior at all phases of the forcing function, not just at periodic intervals, as has been the case previously, when considering a system's Poincaré map. To accomplish this it is necessary to record the position of each point in the ICM array, relative to its previous location, at each numerical time step, so that each position

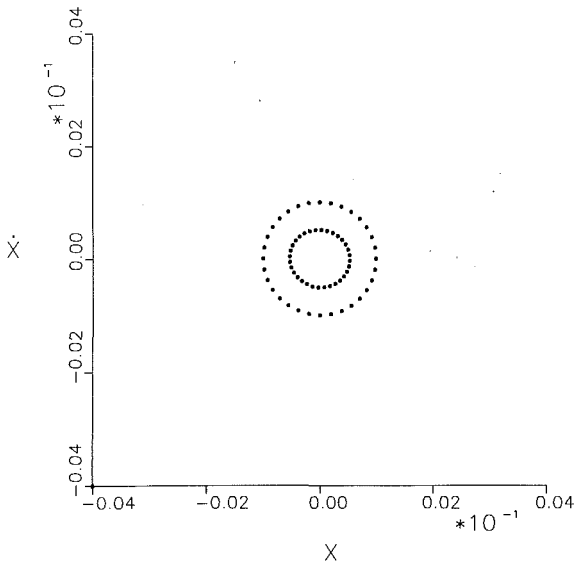


Fig. 3 Mapped image: $\ddot{x} + .1\dot{x} + x = 3.2 \cos(.482t)$, 1st iterate

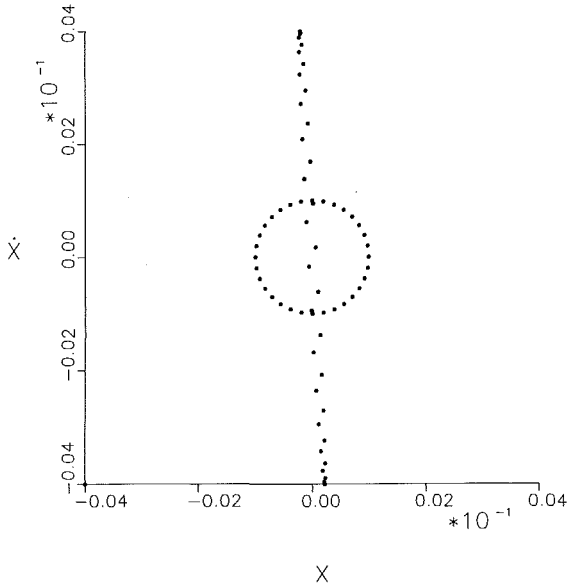


Fig. 5 Mapped image: $\ddot{x} + .1\dot{x} - x + x^3 = 3.2 \cos(.482t)$, 1st iterate

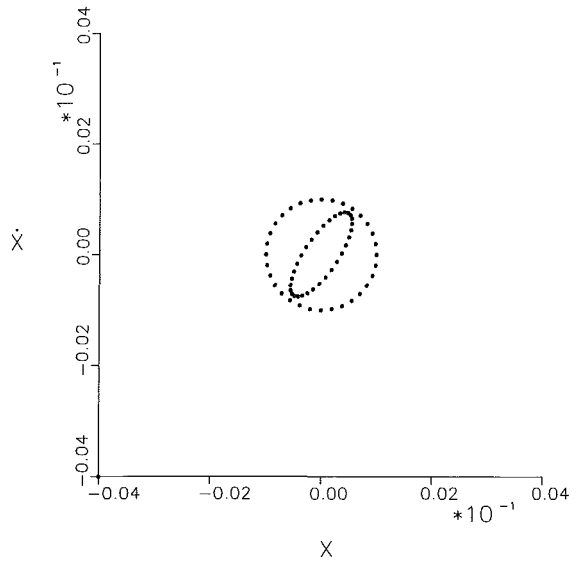


Fig. 4 Mapped image: $\ddot{x} + .1\dot{x} + x + x^3 = 3.2 \cos(.482t)$, 1st iterate

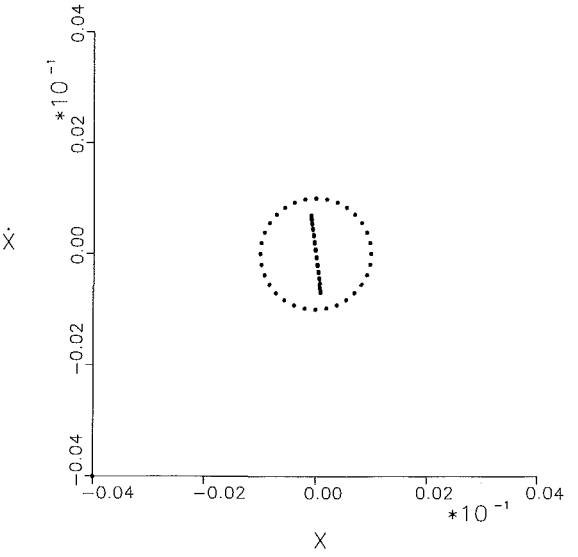


Fig. 6 Mapped image: $\ddot{x} + .1\dot{x} + x - x^3 = 3.2 \cos(.482t)$, 20th iterate

can be used to interpolate to the next. If an Interpolated Mapping grid that is reasonable for a Poincaré map is used, however, this method fails to converge due to the wanderings of the trajectory. Specifically, a typical region in the initial array maps outside of the array bounds after less than ten interpolations. Expanding the array to the degree needed to encompass all of the wanderings of the system for all phases is infeasible, as this would require a grid of such magnitude that all computational advantage would be lost. It would appear, therefore, that Interpolated Mapping is impractical (in its standard form) for finding the exact exponents of a system. Thus, a modification to the technique is in order.

Sequentially Generated Mapping. To obtain the quantity of transient information necessary to determine the Lyapunov exponents, while preserving the computational advantage of Interpolated Mapping, the following technique is used. A test point in phase space is selected and numerically integrated until it has converged onto an attractor. Once the generated trajectory has converged to an attractor, the interpolation process begins. When the point begins its next forcing period, it is surrounded with four other points. Each of these four points are then numerically integrated to their position after just one time step in the numerical integration routine, and their final positions are recorded. The test point is interpolated to its next location at the next time Δt in the future, and the process is repeated. Each time the point advances a step forward, a test is made to see if the point is surrounded by any existing array points. If it is, then the point is interpolated to its next position and the process continues. Otherwise, another group of surrounding points is created in the manner just described before continuing. The procedure used in finding the exponents by this technique is identical to the general method except that at each time step the vectors are interpolated, rather than integrated, to their next positions at some point Δt in the future. The calculations are then performed in the same way as before. In this way, the minimum covering set of array points is calculated that encloses the attractor.

Because of the transient information recorded, the array can be used to recreate an entire trajectory for the system. The

Table 1 Comparison of exact and approximate exponents

	Exact	Approximate
Period 1	−.00860 −.09139	−.00865 −.09135
Period 2	−.04995 −.05004	−.04996 −.05004
Period 4	−.03218 −.06780	−.03287 −.06713
Chaos	.01755 −.11754	.01713 −.11713

Table 2 Comparison of CPU requirements

	Exact	Approximate
Period 1	486.1	75.1
Period 2	487.2	77.0
Period 4	486.2	77.6
Chaos	486.9	283.0

Table 3 Comparison of exact and discrete map exponents

	Exact	Mapping
Period 1	−.00860 −.09139	−.01004 −.08617
Period 2	−.04995 −.05004	−.04024 −.23666
Period 4	−.03218 −.06780	−.01339 −.07438
Chaos	.01755 −.11754	.01612 −.10830

Table 4 Comparison of CPU requirements

	Exact	Approximate
Period 1	486.1	70.2
Period 2	487.2	70.9
Period 4	486.2	71.2
Chaos	486.9	71.6

results of the Lyapunov exponent calculations obtained with this approach are compared in Table 1 with those found through direct numerical integration. Equation (12) was used in the calculations, except that the forcing frequency was chosen to be {0.482, 0.4776, 0.476, 0.475} to allow the different period responses to be studied.

These results indeed show that the exponents found using the Sequentially Generated Mapping are close to the exponents calculated through exact numerical integration. The practicality of the method is further illustrated by a comparison of CPU times required to complete the calculations. Table 2 shows the CPU time required by a CYBER 855 computer to perform the calculations illustrated in Table 1. In each case, 200 forcing periods at 80 time steps per period were used to eliminate the transient behavior. The calculations were then based upon 5000 additional forcing periods.

It is clear from Table 2 that this approximate method is less computationally intensive than "exact" numerical integration for the periodic cases. Even for the chaotic case, the Sequentially Generated Mapping technique took only about half as long as the direct integration method. Because all of the necessary grid points have been generated by the time 5000 forcing periods have been followed, the savings in computation will increase if longer times are examined. Therefore, this sort of approach to Lyapunov exponent calculation appears to be practical and inexpensive, especially in the case of periodic orbits.

Exponents of the Mapping

When dealing with a previously calculated array for Poincaré map generation, transient information is not available. However, the exponents themselves may not be as important as their sign and how they compare with other exponents. It is therefore of interest to see what information can be gained from the use of the entire mapping array and considering the system to be represented by a point-to-point map. That is, the Δt used in the exponent calculation becomes a much larger period of time, such as the full period of the forcing function for the system. This approach was undertaken for the same cases as was done previously for the Sequentially Generated Array. An exponent was considered to have converged when its average exhibited a change of less than 10^{-4} over 10 iterations. In order to allow a fair comparison with the previous results, the number of forcing cycles was chosen to be equal to 400,000. This was because the Sequentially Generated data used 5000 forcing periods with 80 time steps per period, a total of 400,000 individual time increments. In the present case one jumps forward an entire period at a time, thus the 400,000 total periods of interpolation. The computational results are shown in Table 3 and the associated costs in Table 4.

It must be noted that an Interpolated Mapping array must first be found for this method and that the cost of this is 114 CPU seconds. Presumably the array would have been found in both cases if a global determination of the various attractors had been desired. The results show that, generally, there is not an exact correlation between the exponents found from numerical integration and their discrete-mapped counterparts. The smallest exponent for the exact period-two motion, for example, is smaller in magnitude than the corresponding value for the period-one motion. However, the reverse of this trend is evident in the exponents calculated from the mapping. At other times, an increase in the exact system exponents is met with an increase in the mapping exponents. It is also very interesting to note that the exponents found for the chaotic system were almost identical for the two methods. This suggests that more information is available in the strange attractor of a system than in the periodic attracting sets. This would seem to be reasonable in view of the fact that the dimension of a chaotic attractor is higher than that of a periodic attractor. Also, it should be noted that a positive exponent remained positive and a negative exponent remained negative whether the system was viewed as a continuous dynamical system or as a discrete mapping. This sort of calculation, therefore, can serve as a means of classifying the global behavior of a system as periodic or chaotic in nature.

Conclusions

The method of Interpolated Mapping has been extended to the case of transient analyses, specifically that of Lyapunov exponent determinations. Using the newly presented Sequentially Generated Mapping, results that were extremely close to those found from numerical integrations were generated at reduced computational costs. The results of determining Lyapunov exponents from a point-to-point mapping were presented and shown to preserve the sign of the actual exponents, thus permitting an efficient identification of a system as being chaotic or not.

The results shown indicate that more extensive transient analyses, stability calculations, optimal trajectory planning, etc., may well be efficiently obtained through the use of Interpolated Mapping. Furthermore, the examination of higher

dimensional systems would seem to be readily accomplished. These topics shall be addressed in a future paper.

References

Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J.-M., 1980, "Lyapunov Characteristic Exponents for Smooth Dynamical Systems and for Hamiltonian Systems; A Method for Computing All of Them. Part 1: Theory," *Meccanica*, Vol. 15, No. 1, pp. 9–20.

Farmer, J. D., Ott, E., and Yorke, J. A., 1983, "The Dimension of Chaotic Attractors," *Physica D*, Vol. 7, pp. 153–180.

Guckenheimer, J., and Holmes, P. J., 1983, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York.

Lefschetz, S., 1987, *Differential Equations: Geometric Theory*, Dover Publications, New York.

Tongue, B. H., 1987, "Characteristics of Numerical Simulations of Chaotic Motions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 54, No. 3, pp. 695–699.

Tongue, B. H., 1987, "On the Global Analysis of Nonlinear Systems Through Interpolated Cell Mapping," *Physica D*, Vol. 28, No. 3, pp. 401–408.

Tongue, B. H., and Gu, K., 1988, "Interpolated Cell Mapping of Nonlinear Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, No. 2, pp. 461–466.

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A., 1985, "Determining Lyapunov Exponents from a Time Series," *Physica D*, Vol. 16, pp. 285–317.

Improvement of a Nonparametric Identification Procedure Used in Nonlinear Dynamics

P. Argoul
Service de Mecanique,
Jeune Equipe C.N.R.S.,
Laboratoire Central des Ponts
et Chaussees,
Noisy-le-Grand 93167, France

L. Jezequel
Département de Mécanique des Solides,
Ecole Centrale de Lyon,
Ecully 69131, France

The advantage of nonparametric identification methods based on the use of approximations of the restoring forces is that they do not require the a priori *knowledge of a model for the nonlinear behavior of the structure. However, the main difficulty encountered with this type of methods is the fitting of nonlinear forces in the force-state mapping fields where there are not sufficient experimental data. In this paper, an improvement of the regression technique in conjunction with the use of two-dimensional Chebyshev orthogonal polynomials by introducing an interative computation process is presented. It is shown that the proposed method can properly identify the discretized model even in the case of high cross-product displacement-velocity terms and that this method can be used for structures presenting important nonlinear modal coupling.*

1 Introduction

A real structure always presents nonlinear characteristics, and the modal extraction and synthesis procedure based on the theory of linear systems often gives erroneous results. However, the use of modal analysis can inform the experimenter to the fact that the structure is nonlinear and can give a rough idea about its behavior (Busby, 1986). In most cases, the nonlinear characteristics of a structure are low and its dynamic behavior can be considered as linear. Therefore, most mechanical structural identification methods (Caughey, 1963; Iwan, 1972; Rakheja, 1985; Fang, 1986) are based on equivalent linearization models able to provide a good approximation of forced steady-state responses of quasi-linear systems. Nevertheless, such models cannot reproduce such classical phenomena observed in the case of nonlinear structures as jumps in frequency response plots, and secondary, internal or combined resonance.

Therefore, many studies (Tomizuka, 1977; Wysocky, 1979) have compared tested structure response with the response of models of nonlinear "plants," made of few branches, where linear subsystems are present. These models can be identified from harmonic study of responses obtained by varying the frequency and the magnitude of the harmonic excitations (Wysocky, 1979). This type of procedure consists in approximating the Volterra kernels that represent the nonlinear behavior of the studied structure. Using the properties of separable processes, Billings and Fakhouri (1978, 1982) have proposed a method allowing to separately identify the linear part and the memoryless nonlinear characteristics of a system. Such a procedure appears especially suitable for identifying large structures often presenting nonlinear joints (Jezequel, 1984). However, when dealing with systems that incorporate commonly encountered nonlinearities such as polynomial nonlinearities, the evaluation of higher degree terms requires a great amount of tests often extremely difficult to perform as well as extensive computer storage requirments and high computation costs.

The use of the Hilbert transform (Simon, 1984) also allows to extract the linear part of a system and to detect nonlinearities in a frequency response curve. Even if the structure presents nonlinear characteristics, its dynamic behavior in a given frequency range can be approximated by a discrete model of a few degrees-of-freedom. Therefore, many studies have been devoted to the extension of modal synthesis methods to the nonlinear case (Jezequel, 1985). Szemplinska-Stupnicka (1983) have shown that the introduction of nonlinear normal coordinates in relation to the notion of nonlinear modes permits to improve the classical Rayleigh-Ritz method using linear system modes. In the same way, the efficiency of the use of local modes in relation to a tangent stiffness matrix has been shown (Morris, 1977; Almroth, 1978) during the computation of the nonlinear system response. Recently Ibrahim (1984) has proposed to use the I.T.D. (Ibrahim Time Domain) method for identifying the modal characteristics depending to the excitation magnitude from transient responses.

However, within the modal synthesis framework, it might be more appropriate to use a model based on a discretization in relation to fixed trials functions. The methods of control and optimization theory can then be applied to correct a set of discrete parameters of the model. Further, in most cases the knowledge of basic equations is not sufficient to have a "good" discrete model. Very often in the literature, the

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

Discussion on this paper should be addressed to the 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 15, 1988; final revision, August 16, 1988.

knowledge of the connectivity of the discrete model is assumed. This knowledge allows, by means of an appropriate coordinate change, an uncoupling of the nonlinear cross-product terms involving displacement and velocity and to facilitate their identification. In accordance with modal synthesis methods, it would be consistent to use a normal model with a linear part identified from low-level vibration tests. Thus, the nonlinear part appears as a function of the modal coordinates and their derivatives in time. This part obtained experimentally by substracting the inertial forces from the applied forces is often assumed to be of polynomial form. Therefore, it is identified by means of fitting techniques of force-state mapping (Distephano, 1975; Yang, 1985; Crawley, 1986). Such a nonparametric identification procedure is attractive, because it does not require the knowledge of nonlinear characteristics appearing in the dynamic behavior of the tested structures. The fitting method proposed by Masri et al. (1979, 1982a, 1982b, 1987) is particularly interesting because it is based on the use of a regression technique based on Chebyshev polynomials. This method has been used successfully in several cases. In relation with this method, this paper presents a new interpolation procedure which improves the identification of the nonlinear part of lumped parameter systems in areas of the state-space fields where experimental data are insufficient. When the connectivity of the model is unknown, the proposed modifications of the procedure become essential to properly identify the nonlinear modal coupling for any structural discrete model. The technique is applied to two simple systems presenting high velocity and displacement cross-product terms or high modal coupled terms. Analysis with the proposed modifications yields good agreement with actual structural behavior.

2 Identification Procedure

For a discretized dynamic system with N-degree-of-freedom, the restoring forces f are defined from the equation of motion by

$$f(\mathbf{x},\dot{\mathbf{x}}) = p(t) - \mathbf{m}.\ddot{\mathbf{x}}(t) \tag{1}$$

where $\mathbf{m}$ is the mass matrix $(N\times N)$ which can be estimated using a finite element procedure and where $\ddot{\mathbf{x}}(t)$ $(N\times 1)$ the acceleration vector and $p(t)$ $(N\times 1)$ the excitation forces vector are assumed to be available from measurements over a period $T_{\max}$. No assumptions are made on the discrete model connectivity as are often assumed in the literature—for example, a chainlike structure that consists of a lumped mass model with elementary masses being connected to one another by unknown nonlinear elements (Masri, 1982a) or a more complex model called a branched system with $N\times N$ restoring forces elements (Yang, 1985).

According to modal synthesis methods (Jezequel, 1985), a more appropriate modal representation is used. Let

$$\mathbf{x} = \Phi\, u \tag{2}$$

where Φ is the modal matrix $(N\times r)$. Its r columns represent estimations of the r "linearized" normal modes. Φ is identified by means of an usual method of linear identification from low-level magnitude experimental tests.

Equation (1) can be expressed in the form

$$h(u,\dot{u}) = P(t) - \ddot{u}(t) \tag{3}$$

with $\qquad h = \Phi^t f; \quad P = \Phi^t p \quad \text{and} \quad \ddot{u} = \Phi^t \mathbf{m}\,\ddot{\mathbf{x}}.$

Thus, for a given excitation $p(t)$, and r coordinates h_i of h are known for the experimental points $Q_e(u, \dot{u})$ of state plane $\mathbf{E}$ of dimension $2r$. The modal coordinate vectors $u(t)$ and $\dot{u}(t)$ can be obtained from measurements or successive integrations of the modal acceleration vector $\ddot{u}(t)$.

An estimation of the h_i over the whole plane $\mathbf{E}$ must then be made. It is assumed, in relation to the modal representation, that $h_i(u, \dot{u})$ can be expressed as the sum of a main term $h_i^{(1)}(u_i, \dot{u}_i)$ representing the contribution of mode i to h_i and some other terms resulting from the interaction of modes $j(j\neq i)$ with mode i.

From the knowledge of the measurements of $h_i(u, \dot{u})$, $h_i^{(1)}(u_i, \dot{u}_i)$ is approximated in the phase plane $(u_i, \dot{u}_i)$ by a function $\tilde{h}_i^{(1)}$ expressed in terms of two-dimensional Chebyshev polynomials.

$$h_i^{(1)}(u_i, \dot{u}_i) \approx \tilde{h}_i^{(1)}(u_i, \dot{u}_i) = \sum_k \sum_l C1_{kl}^{(i)}\, T_k(u_i)\, T_l(\dot{u}_i) \tag{4}$$

It is assumed that the residual error $[h_i(u, \dot{u}) - \tilde{h}_i^{(1)}(u_i, \dot{u}_i)]$ due to the previous approximation constitutes an estimation of $h_i^{(2)}(u_i, u_j)$, which is also developed in a Chebyshev polynomial expansion.

$$h_i^{(2)}(u_i, u_j) \approx \tilde{h}_i^{(2)}(u_i, u_j) = \sum_k \sum_l C2_{kl}^{(i)}\, T_k(u_i)\, T_l(u_j). \tag{5}$$

As just indicated, $[h_i(u, \dot{u}) - \tilde{h}_i^{(1)}(u_i, \dot{u}_i) - \tilde{h}_i^{(2)}(u_i, u_j)]$ allows an approximation of $h_i^{(3)}(\dot{u}_i, \dot{u}_j)$

$$h_i^{(3)}(\dot{u}_i, \dot{u}_j) \approx \tilde{h}_i^{(3)}(\dot{u}_i, \dot{u}_j) = \sum_k \sum_l C3_{kl}^{(i)}\, T_k(\dot{u}_i)\, T_l(\dot{u}_j). \tag{6}$$

The convergence criterion taken for the calculation is as follows

$$\left\| h_i(u, \dot{u}) - \sum_s \tilde{h}_i^{(s)} \right\| < \epsilon \tag{7}$$

where ϵ is a small positive quantity and where the $\|.\|$ norm is the root-mean-square (rms) value over the K experimental points, defined by

$$\|f\| = \frac{1}{K} \sum_{e=1}^{K} f^2 (Q_e)^{1/2}. \tag{8}$$

Finally, an estimation of $h_i(u, \dot{u})$ is obtained by

$$h_i(u, \dot{u}) \approx \tilde{h}_i(u, \dot{u}) = \tilde{h}_i^{(1)}(u_i, \dot{u}_i) + \sum_{j\neq i} \tilde{h}_i^{(2)}(u_i, u_j) + \tilde{h}_i^{(3)}(\dot{u}_i, \dot{u}_j) + \tilde{h}_i^{(4)}(u_i, \dot{u}_j) + \tilde{h}_i^{(5)}(\dot{u}_i, u_j). \tag{9}$$

The $\tilde{h}_i^{(s)}$ terms due to the interaction between modes j and mode i (j very higher than i) are usually neglected in comparison with $h_i^{(1)}$. This simplification is valid due to the change of representation to modal coordinates. Thus, only $(i-1)$ and $(i+1)$ indexes are usually taken into account in equation (9).

3 Chebyshev Polynomial Approximation

The estimation of $h_i(u, u)$ is based on two-dimensional least-squares orthogonal polynomial approximation. The choice of Chebyshev polynomials of first kind is particularly interesting partly because the least-squares approximation with Chebyshev polynomials of first kind is known to be the most rapidly convergent of all ultraspherical (or Gegenbauer) polynomial approximations (Denman, 1969). Also, the error associated with Chebyshev approximation will tend to oscillate with uniform amplitude over $[-1., 1.]$ interval (equal ripple), whereas the error afforded by the Legendre polynomial approximations (with uniform weighting) will tend to oscillate with an amplitude which increases toward the ends of the interval.

Let $h(a, b)$ be the function known experimentally for a set

S_Q of $Q_{e,e=1,K}$ points of a domain Ω. h is then approximated by Chebyshev polynomials

$$h(a, b) \approx \tilde{h}(a', b') = \sum_{k=1}^{N_a} \sum_{l=1}^{N_b} C_{kl} T_k(a') T_l(b') \quad (10)$$

where (a', b') are the normalized values of (a, b) over $\Omega' = [-1, 1]^2$ and T's are Chebyshev polynomials.

The C_{kl} coefficients are given by

$$C_{kl} = \frac{4}{(\delta_{0l}+1)(\delta_{0k}+1)\Pi^2} \int_{-1}^{1} \int_{-1}^{1} h(a,b) T_k(a') T_l(b') \omega(a') \omega(b') da' db' \quad (11)$$

where δ_{ij} is the Kronecker symbol and

$$\omega(x') = \frac{1}{\sqrt{1-x'^2}} \text{ for } x' \in [-1, 1].$$

The two-variable integral in (11) is computed by a Gauss-Chebyshev quadrature with $(M_a \times M_b)$ P_{ii} Chebyshev points with the following coordinates

$$a_i = \cos\left(\frac{2i-1}{2M_a}\Pi\right) \quad \text{and} \quad b_j = \cos\left(\frac{2j-1}{2M_b}\Pi\right): \quad (12)$$

The C_{kl} are given by

$$C_{kl} = \frac{4}{(\delta_{0l}+1)(\delta_{0k}+1)M_a M_b} \sum_{i=1}^{M_a} \sum_{j=1}^{M_b} h(P_{ij}) \cos\left(k \frac{2i-1}{2M_a}\Pi\right) \cos\left(l \frac{2j-1}{2M_b}\Pi\right). \quad (13)$$

Equation (13) becomes exact when h is a polynomial of degree $(2M_a - 1 - k)$ for the variable u and of degree $(2M_b - 1 - l)$ for the variable $\dot{u}$.

The application of equation (13) for the computation of the C_{kl} coefficients implies the estimation of h at P_{ij} Chebyshev points which are located symmetrically with respect to the central point of the square Ω' in the phase plane. A two-dimensional interpolation scheme is elaborated to generate h at the prescribed P_{ij} points by means of data Q_e points in the vicinity of P_{ij}. It is first necessary to choose a type of excitation and a period $T_{\max}$ adapted to sweep the whole domain as uniformly as possible. However, even with a proper excitation and a large enough period $T_{\max}$, there may still be locations in domain Ω where the density of data points is too low to allow a "good" bilinear interpolation. Moreover, when M_a and M_b increase, the P_{ij} quadrature points concentrate near the edges of Ω' where experimental points are missing. A domain smaller than Ω and where the interpolation procedure would be properly made could be then chosen. But it would not be advisable to stop the computation at that point for two reasons:

- first, the value of h at points of the state plane for u or $\dot{u}$ sufficiently important must be known so that the effects of the nonlinearity be taken into account.
- then, the points far away from the origin of Ω domain are less sensitive to noise pollution present in all experimental tests than the points near the origin.

Therefore, the following interpolation technique has been built up to take into account the remarks mentioned previously.

4 Interpolation Procedure

Two cases are considered, when P is close to S_Q or P is not close to S_Q. Thus, a critical distance r_c is introduced in relation to the number K and the position of the data points Q_e. A point P of Ω is said "close" to S_Q if it is possible to find a triangle including P with end-points R_i belonging to S_Q and verifying

$$\forall i \in [1,3] \quad \text{distance}(P, R_i) \le r_c. \quad (14)$$

To determine whether P is "close" to S_Q or not, the points of S_Q which verify or which do not verify equation (14) are looked for in the four quadrants around P.

(A) P is "Close" to Experimental Points. The optimal triangle for the interpolation must be chosen among previously found triangles including P. To obtain a proper bilinear interpolation, the use of a too "flat" triangle must be avoided. For this purpose, a ratio defined by (perimeter)^2/(surface) is computed for each triangle and the triangle with the smaller ratio is kept for the interpolation. $h(P)$ is then computed using a bilinear interpolation at the three end-points of this triangle where the value of h is experimentally known.

(B) P is not "Close" to Experimental Points. For P points far removed from measured response pairs $(u, \dot{u})$, an extrapolation procedure must be found for properly estimating the nonlinear restoring forces even in the case of cross-product terms.

First, a domain Ω_1 ($\Omega_1 = [a^{(1)}{}_{\min}, a^{(1)}{}_{\max}] \times [b^{(1)}{}_{\min}, b^{(1)}{}_{\max}]$, $\Omega \supset \Omega_1$) must be found where the information on h is sufficient to make a good bilinear interpolation of $h(P_{ij})$—that is to say on Ω_1 domain, sufficient experimental points can be found close to the P_{ij} quadrature-points. It is therefore possible to compute the Chebyshev approximation $h_1{}^*$ of h on Ω_1. Then, for a larger domain Ω_2 ($\Omega_2 \supset \Omega_1$, $\Omega_2 = [a^{(2)}{}_{\min}, a^{(2)}{}_{\max}] \times [b^{(2)}{}_{\min}, b^{(2)}{}_{\max}]$) for $P \in \Omega_2$, P being not "close" to experimental points, the following extrapolation scheme has been tested

$$h(P) \approx h_1{}^*(g(P)) \quad (15)$$

where g is the affinity defined from Ω_2 to Ω_1 by

$$g(a^{(2)}, b^{(2)}) = (a^{(1)}, b^{(1)}) \quad \text{with} \quad (16)$$

$$a^{(1)} = \frac{a^{(2)}(a^{(1)}_{\max} - a^{(1)}_{\min}) + a^{(2)}_{\max} a^{(1)}_{\min} - a^{(2)}_{\min} a^{(1)}_{\max}}{a^{(2)}_{\max} - a^{(2)}_{\min}}$$

and

$$b^{(1)} = \frac{b^{(2)}(b^{(1)}_{\max} - b^{(1)}_{\min}) + b^{(2)}_{\max} b^{(1)}_{\min} - b^{(2)}_{\min} b^{(1)}_{\max}}{b^{(2)}_{\max} - b^{(2)}_{\min}}$$

where the index 1 (or 2) indicates that the indexed terms are relative to Ω_1 (or Ω_2), respectively.

It is now possible to compute Chebyshev polynomial approximation $h_2{}^*$ of h on Ω_2. Then, this process is being iterated until Ω_n covers entirely the domain Ω. In most cases, the experimental points envelope has a form near the elliptic (or circular) one as shown in Fig. 1. For the extension of the successive domains Ω_j $(1 \le j \le n)$ of the proposed extrapolation procedure, the ratio s_n between two successive subdomains areas, is defined by

$$s_n = \frac{S_{\Omega_n}}{S_{\Omega_{n-1}}} \approx \frac{\pi}{2}. \quad (17)$$

s_n ratio is taken equal to $\pi/2$ because it corresponds to the ratio between an elliptic (or circular) surface and the surface of the largest rectangle inscribed within the ellipse as shown in Fig. 1. Finally, an approximation $h_n{}^*$ of h upon Ω is obtained.

5 Applications of the Proposed Method

(1) First, the new extrapolation process has been compared to the one proposed by Masri et al. (1979) with the example of the Van der Pol oscillator which involves cross-coupling nonlinearity in $u^2\dot{u}$

$$h(u, \dot{u}) = -\alpha(1-u^2)\dot{u} + u \quad (18)$$

where $\mathbf{m} = 1$, $\omega = 1$ and $\alpha = 0.4$.

Table I Comparison between Caughey-Masri method (standard-typed) and new proposed method (heavy-typed) for calculation of C_{kl} coefficients for the Van der Pol oscillator identification

k	1	$T_0(u')$	$T_1(u')$	$T_2(u')$	T_3	$T_4(u')$	$T_5(u')$	$T_6(u')$
$T_0(u')$		0.27	12.83	0.18	−2.79	−0.18	1.15	−0.25
		−0.18	−1.28	−0.17	−2.79	−0.16	0.82	−0.50
$T_1(u')$		9.47	0.49	3.41	−0.64	−1.20	0.02	−0.10
		7.20	0.23	−0.52	0.24	−0.66	0.81	0.76
$T_2(u')$		−0.94	7.70	−0.05	−0.93	−0.001	0.89	−0.16
		0.02	−2.66	−0.23	0.28	−0.35	0.48	−0.21
$T_3(u')$		−1.93	−1.27	−0.46	0.60	0.67	−0.25	−0.41
		−2.97	−0.32	−1.25	−1.02	1.69	−0.47	−0.67
$T_4(u')$		0.19	−5.33	0.63	2.81	−0.03	−0.74	0.20
		0.04	−2.18	−0.57	2.84	−0.65	−1.72	0.15
$T_5(u')$		−0.77	1.87	0.45	−0.33	0.52	0.23	−0.46
		0.60	0.30	2.60	0.19	−0.39	−0.002	−1.07
$T_6(u')$		0.64	2.64	0.09	0.64	0.06	−0.24	0.37
		0.16	3.14	0.91	−0.40	0.57	−0.81	0.24

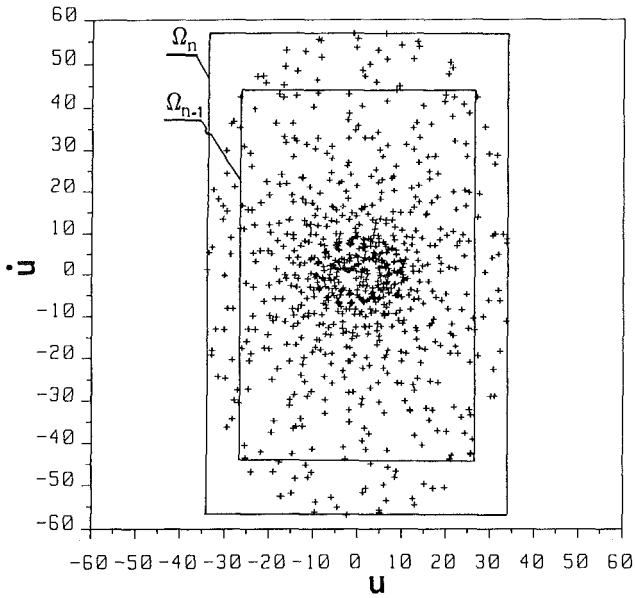


Fig. 1 Illustration of domain-extending ratio s_n for the new extrapolation scheme

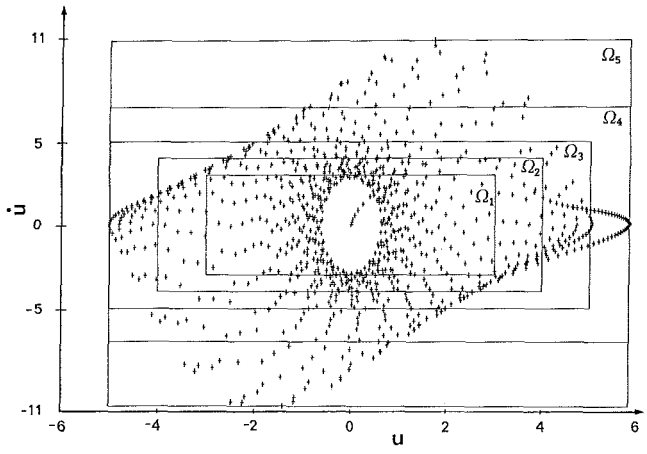


Fig. 2 Data points in the state-plane for the Van der Pol oscillator identification

To identify this oscillator, 1000 data points have been used in the state-plane $\Omega = [-5.02,\ 5.83] \times [-10.71,\ 10.87]$, as shown in Fig. 2. These points have been generated by subjecting during a period $T_{max} = 100$, the mass **m** to an harmonic excitation, with a linearly varying frequency, given by

$$p(t) = 13 \sin\left(\frac{1}{15.\pi}\, t^2 + 0.5\, t\right). \tag{19}$$

For the points far removed from experimental data points, and by analogy with linear systems behavior, Masri et al. have used the following extrapolation scheme

$$h(a,b) \approx h_1(a) + h_2(b) \approx \sum_{i=0}^{m_1} c_i T_i(a') + \sum_{j=0}^{m_2} d_j T_j(b') \tag{20}$$

where the c's are Chebyshev polynomial coefficients for a one-dimensional least-squares fit of the experimental data for which $|a| \approx 0$ and the d's are similar coefficients for data points for which $|b| \approx 0$. It is obvious that if the structure has polynomial nonlinearities without cross-products terms, the approximation of h given by relation (20) is sufficient to give a correct identification of h. Nevertheless, this extrapolation procedure will involve some difficulties to take into account cross-coupling effects because its application for the points far away from experimental points will strongly disturb the two-dimensional Chebyshev approximation. Table 1 give the two-dimensional C_{kl} coefficients obtained, respectively, with the proposed iterative process and with Masri et al. process. Figures 3–4 show a comparison between the constant u or constant $\dot{u}$ sections of $h(u,\dot{u})$ computed with Masri et al. extrapolation scheme and the ones computed with the new extrapolation scheme. For the considered oscillatory, the plots of constant u or constant $\dot{u}$ sections of the restoring force h surface identified with the Masri et al. extrapolation scheme in Figs. 3–4 deviate from exact ones even at locations where sufficient experimental information is available. When the new iterative extrapolation process is used, the identified curves are closer to the exact ones and a good estimation of nonlinear

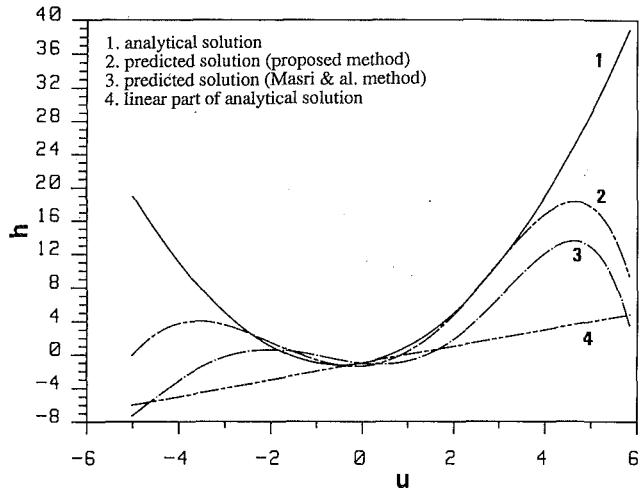


Fig. 3 Section of constant $\dot{u} = 2.5$ $h(u, \dot{u})$ surface for the Van der Pol oscillator

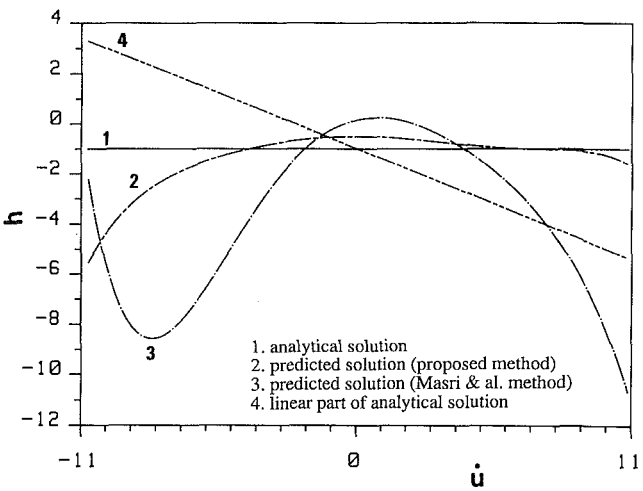


Fig. 4 Section of constant $u = -1$ $h(u, \dot{u})$ surface for the Van der Pol oscillator

cross-product terms effects is achieved in the force-state fields where experimental points are present. The successive subdomains used for the Van der Pol oscillator identification, shown in Fig. 2 are, respectively, $\Omega_1 = [-3., 3.]^2$; $\Omega_2 = [-4., 4.]^2$; $\Omega_3 = [-5.02, 5.] \times [-5., 5.]$; $\Omega_4 = [-5.02, 5.83] \times [-7., 7.]$; $\Omega_5 = \Omega$ and the C_{kl} have been computed with 15×15 quadrature points.

(2) Since noise pollution cannot be prevented in any experimental measurment, the sensitivity of the proposed procedure to measurement noise has been studied. To do this, the $h(Q_e)$ initial data of the previous oscillator have been contaminated by adding Gaussian zero mean white noise with a standard deviation equal to 10 percent of the maximum value of $h(Q_e)$.

In the force-state mapping, the overall $h(u,\dot{u})$ surface identified with noise-perturbed data still yields a good estimate of the physical structure characteristics obtained without noise pollution as shown by Figs. 6(*a*) and 6(*b*). On the contrary, the surface $h_1(u,\dot{u})$ identified with noised data points in the first subdomain Ω_1 is very different from the one identified with noiseless points as shown by Figs. 5(*a*) and 5(*b*).

The influence of points far removed from experimental points is of great interest in the new extrapolation scheme because these points are generally located far away from the state-plane origin point and their value of h is high and low-sensitive to noise pollution.

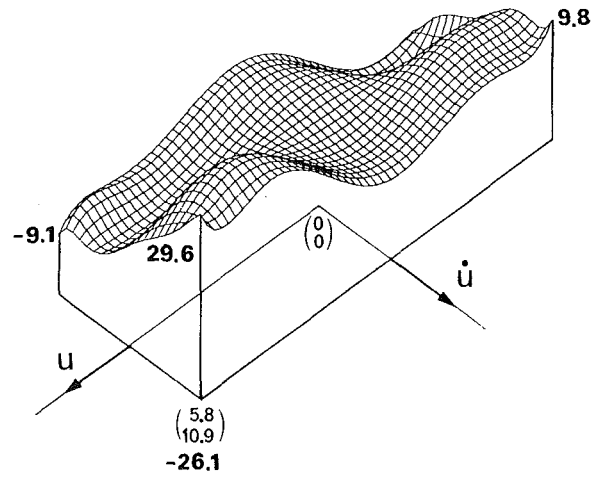


Fig. 5(*a*) Chebyshev approximation $h_1^*(u, \dot{u})$ on Ω_1 for the Van der Pol oscillator obtained with noised data

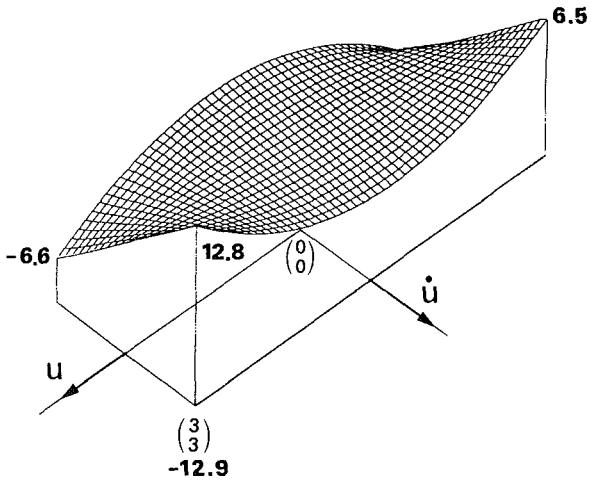


Fig. 5(*b*) Chebyshev approximation $h_1(u, \dot{u})$ on Ω_1 for the Van der Pol oscillator obtained with noiseless data

3 The proposed identification procedure has been tested on a two-degree-of-freedom model presenting a Duffing-type nonlinear element between the two masses M_1, and M_2, illustrated in Fig. 7. After the change into a modal representation where the modal matrix is

$$\Phi = \begin{bmatrix} 0.526 & 0.851 \\ -0.851 & 0.526 \end{bmatrix} \tag{21}$$

and with the natural pulsations $\omega_1 = 1.39$ and $\omega_2 = 1.47$ being very close to each other, relation (3) gives

$$h_1 = 2.162\, u_1 + 0.718\, u_1^3 + 0.508\, u_1^2 u_2 + 0.12\, u_1 u_2^2 + 0.034\, u_2^3 + 0.029\, \dot{u}_1 + 0.002\, \dot{u}_2$$

and

$$h_2 = 1.938\, u_2 + 0.002\, u_2^3 + 0.028\, u_2^2 u_1 + 0.12\, u_2 u_1^2 + 0.169\, u_1^3 + 0.002\, \dot{u}_1 + 0.021\, \dot{u}_2. \tag{22}$$

Figures 8 and 9 show the modal displacement time history $u_1(t)$ and $u_2(t)$ for the two modes. These plots have been generated by subjecting the mass M_2 during a period $T_{\max} = 200$, to a swept-sine excitation given by

$$p(t) = 0.6 \sin\left(\frac{1}{70.\pi}\, t^2 + 0.5\, t\right). \tag{23}$$

1000 Q_e data points have been used for the identification of each of the two modal restoring forces h_1 and h_2 in the force-state mapping. Figure 10 shows a three-dimensional plot of

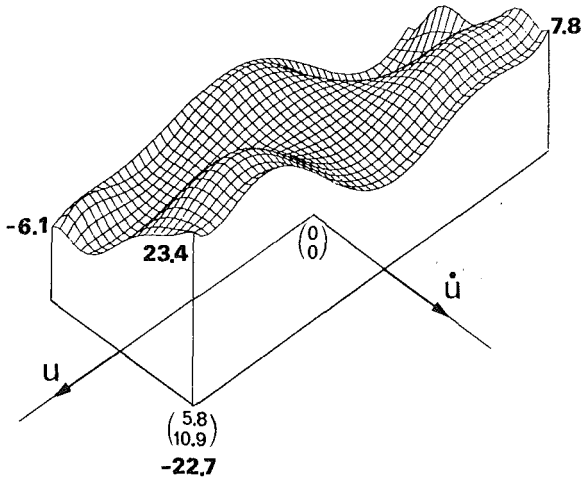


Fig. 6(*a*) Overall Chebyshev approximation $h_s^*(u, \dot{u})$ on Ω for the Van der Pol oscillator obtained with noised data

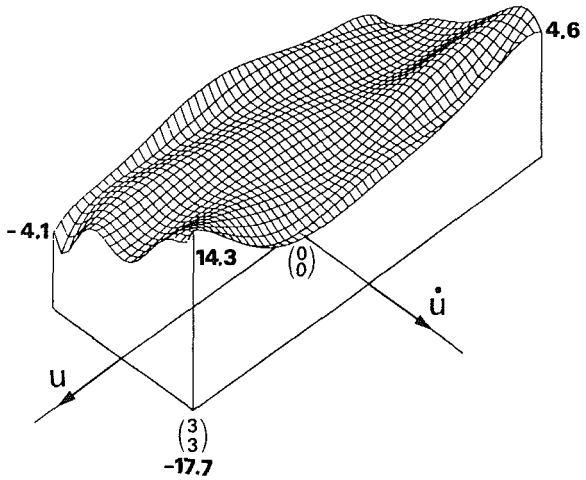
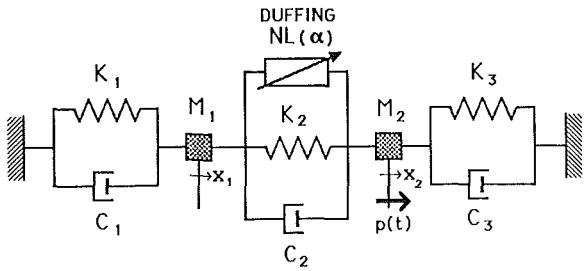


Fig. 6(*b*) Overall Chebyshev approximation $h_s(u, \dot{u})$ on Ω for the Van der Pol oscillator obtained with noiseless data



$M_1 = M_2 = 1$

$K_1 = 1.9\ ,\ K_2 = 0.1\ ,\ K_3 = 2.$

$C_1 = C_3 = 0.02\ ,\ C_2 = 0.005$

and $\alpha = 0.2$

Fig. 7 Example nonlinear 2DOF system (Duffing nonlinearity)

the estimated restoring force surface $h_1(u_1, u_2)$ where $\dot{u}_1$ and $\dot{u}_2$ are taken equal to zero. Figures 11 and 12 show constant u_1 or constant u_2 sections of this surface. The $h_1(u_1, u_2)$ surface in Fig. 10 clearly indicates the cubic variation of the nonlinearity terms along u_1. This result is confirmed by the constant u_1 or constant u_2 sections of h_1 plotted in Figs. 11–12, and the identified curves are very close to the exact ones. The proposed procedure allows to emphasize the nonlinear coupling terms between the modes.

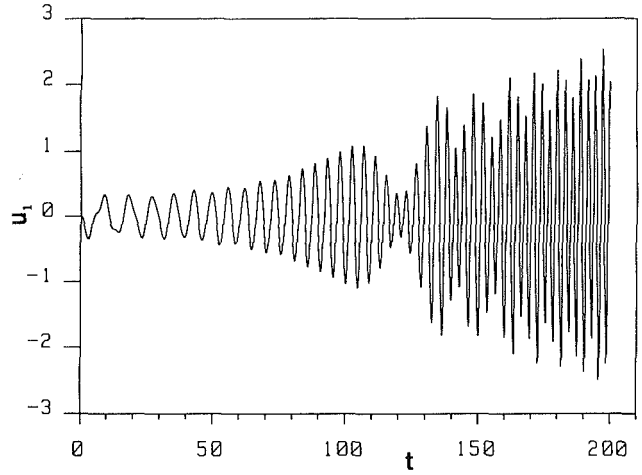


Fig. 8 Modal displacement time history for the 2DOF system first mode

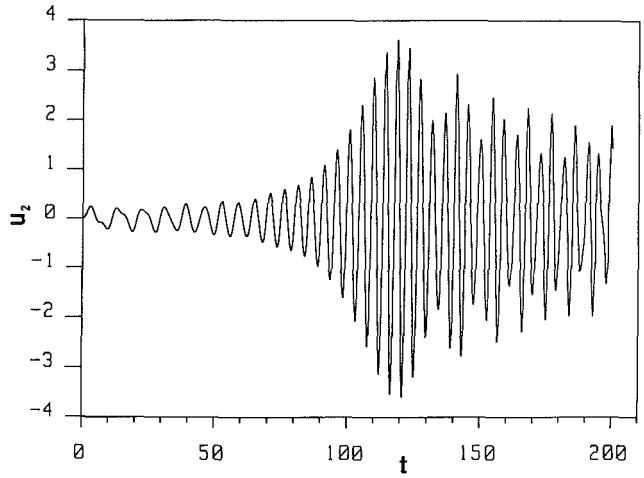


Fig. 9 Modal displacement time history for the 2DOF system second mode

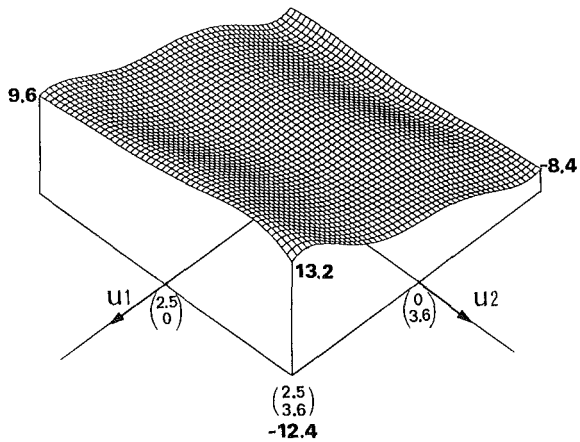


Fig. 10 Constant $u_1 = 0$ and $u_2 = 0$ section of least-squares Chebyshev polynomial approximated surface $h_1(h_1, h_2, \dot{u}_1, \dot{u}_2)$ for the 2DOF system

6 Conclusion

Among nonparametric identification techniques, a relatively simple and efficient technique based on force-state mapping has been presented that is suitable for use with discretized systems without any restriction with regard to the connectivity of the discrete model. It proves to be effective to identify the

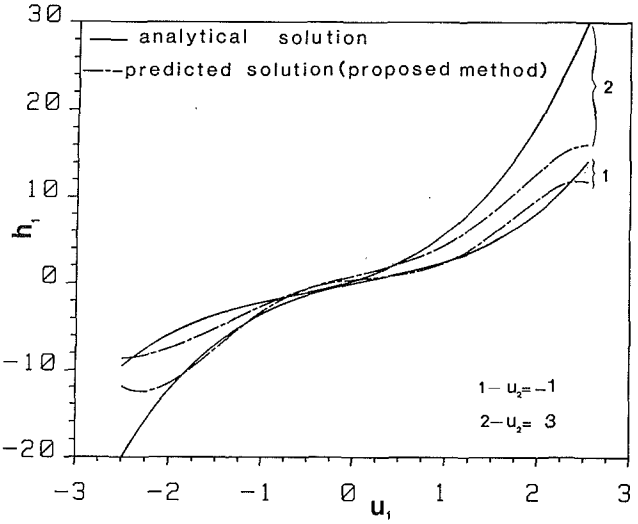


Fig. 11 Constant $u_2 = -1$, $u_2 = 3$ sections of the surface $h_1(u_1, u_2, 0, 0)$ for the 2DOF system

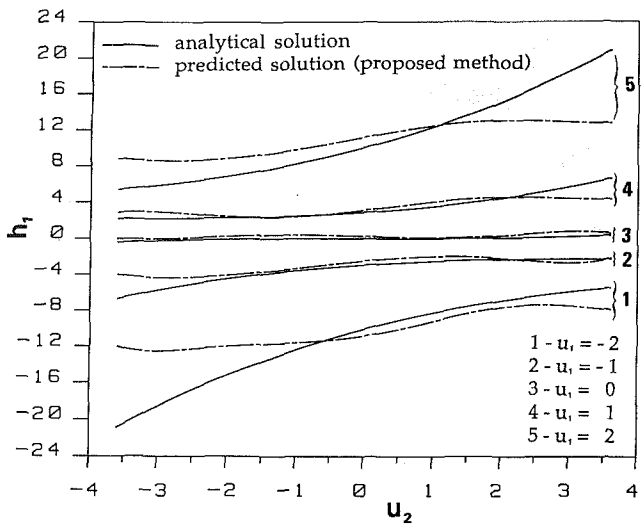


Fig. 12 Constant $u_1 = -2$, $u_1 = -1$, $u_1 = 0$, $u_1 = 1$ or $u_1 = 2$ sections of the surface $h_1(h_1, u_2, 0, 0)$ for the 2DOF system

nonlinear system behavior without making any assumption on the type of nonlinearities of the system. The proposed method requires information on the identification of the linear behavior of the structure—its pertinent mode shapes—which allows the change to a modal representation and on its dynamic nonlinear response. Then, an approximated expression for each of the modal restoring functions is determined by means of regression techniques involving two-dimensional orthogonal functions, in terms of the corresponding modal state variables. The main difficulty with this type of methods is the interpolation procedure in the state-plane fields where there are not sufficient experimental data. The proposed interpolation procedure improved the results by introducing an iterative process which proved to be more efficient when the encountered nonlinearities include velocity and displacement cross-product terms. Analysis indicates that identification results obtained with this new procedure are relatively insensitive to noise pollution of the data and that this new method allows to the proper identification of high nonlinear coupling terms between the modes.

References

Almroth, B. O., Stern, P., and Brogan, F. A., 1978, "Automatic Choice of Global Shape Functions in Structural Analysis," *A.I.A.A. Journal*, Vol. 16, pp. 525–528.

Billings, S. A., and Fakhouri, S. Y., 1978, "Theory of Separable Processes with Applications to the Identification of Nonlinear Systems," *Proceedings I.E.E.E.*, Vol. 125, pp. 1051–1058.

Billings, S. A., and Fakhouri, S. Y., 1982, "Identification of Systems Containing Linear Dynamic and Static Nonlinear Elements," *Automatica*, Vol. 18, pp. 15–26.

Busby, H. R., Nopporn, C., and Singh, R., 1986, "Experimental Modal Analysis of Nonlinear Systems: A Feasibility Study," *Journal of Sound and Vibration*, Vol. 180, No. 3, pp. 415–427.

Caughey, T. K., V 1963, "Equivalent Linearization Techniques," *Journal of Acoustical Society of America*, Vol. 35, pp. 1706–1711.

Crawley, E. F., and Aubert, A. C., 1986, "Identification of Nonlinear Structural Elements by Force-State Mapping," *A.I.A.A. Journal*, Vol. 24, pp. 155–162.

Denman, H. H., 1969, "An Approximate Equivalent Linearization Technique for Nonlinear Oscillations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 36, pp. 358–360.

Distephano, N., and Rath, A., 1975, "System Identification in Nonlinear Structural Seismic Dynamics," *Computer Method in Applied Mechanics and Engineering*, Vol. 5, pp. 353–372.

Fang, T., and Wang, Z. N., 1986, "A Generalization of Caughey's Normal Mode Approach to Nonlinear Random Vibration Problems," *AIAA Journal*, Vol. 24, pp. 531–534.

Ibrahim, S. R., 1984, "Time Domain Quasilinear Identification of Nonlinear Systems," *A.I.A.A. Journal*, Vol. 22, pp. 817–823.

Iwan, W. D., and Yang, I. M., 1972, "Application of Statical Linearization Techniques to Nonlinear MDOF Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 545–549.

Jezequel, L., 1984, "Modal Synthesis of Large Structures with Nonlinear Joints from Vibration Tests," *Proceedings of the Second International Conference on Recent Advances in Structural Dynamics*, Southampton, U.K., Vol. 2, pp. 281–295.

Jezequel, L., 1985, *Synthèse Modale, Théorie et Extensions*, Student Press, Ecole Centrale de Lyon, France, (in French).

Masri, S. F., and Caughey, T. K., 1979, "A Nonparametric Identification technique for Nonlinear Dynamic Problems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 46, pp. 433–447.

Masri, S. F., Bekey, G. A., Sassi, H., and Caughey, T. K., 1982, "Nonparametric Identification of Class of Nonlinear Multidegree Dynamic Systems," *Journal of Earthquake and Structural Dynamics*, Vol. 10, pp. 1–30.

Masri, S. F., Sassi, H., and Caughey, T. K., 1982, "Nonparametric Identification of Nearly Arbitrary Nonlinear Systems," ASME JOURNAL OF APPLIED MECHANICS, Vol. 49, pp. 619–628.

Masri, S. F., and Miller, R. K., Saud, A. F., and Caughey, T. K., 1987, "Identification of Nonlinear Vibrating Systems: Part I and II," ASME JOURANL OF APPLIED MECHANICS, Vol. 54, pp. 918–929.

Morris, N. F., 1977, "The Use of Modal Superposition in Nonlinear Dynamics," *Computers and Structures*, Vol. 7, pp. 65–72.

Rakheja, S., Van Vliet, M., and Sankar, S., 1985, "A Discrete Harmonic Linearization Technique for Simulating Nonlinear Mechanical Systems," *Journal of Sound and Vibration*, Vol. 100, pp. 511–526.

Simon, M., and Tomlinson, G. R., 1984, "Use of the Hilbert Transform in Modal Analysis of Linear and Nonlinear Structures," *Journal of Sound and Vibration*, Vol. 96, pp. 421–436.

Szemplinska-Stupnicka, W., 1983, "Nonlinear Normal Modes and the Generalized Ritz Method in the Problems of Vibrations of Nonlinear Elastic Continuous Systems," *International Journal of Nonlinear Mechanics*, Vol. 18, pp. 149–165.

Tomizuka, M., 1977, "Series-Parallel and Parallel Identification Schemes for a Class of Continuous Nonlinear Systems," *Journal of Dynamic Systems, Measurement and Control*, June, pp. 137–140.

Wysocki, E. M., and Rugh, W. J., 1979, "An Approximation Approach to the Identification of Nonlinear Systems Based on Frequency Response Measurements," *International Journal of Control*, Vol. 29, pp. 113–123.

Yang, Y., and Ibrahim, S. R., 1985, "A Nonparametric Identification Technique for a Variety of Discrete Nonlinear Vibration Systems," *ASME Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 107, pp. 60–66.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N. Y. 10017, or to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

On the Sufficiency of the Principle of Virtual Work for Mechanical Equilibrium: A Critical Reexamination

John G. Papastavridis[1]

Starting from the general kinetic principle of d'Alembert/Lagrange, an energetic proof of the sufficiency conditions for equilibrium (known as Principle of Virtual Work) is presented. It is clearly demonstrated why to maintain equilibrium requires that, in addition to the familiar vanishing of the virtual work of the impressed forces on the originally motionless system, its geometrical (holonomic) constraints be explicitly time independent (stationary) and its nonintegrable kinematical (nonholonomic) ones be linear and homogeneous in the generalized velocities (catastatic).

1 Introduction

In analytical statics, the Principle of Virtual Work (PVW) states that the vanishing of the (first order) virtual work of all the (internal and external) impressed forces on an originally motionless mechanical system, relative to an inertial frame of reference, and subject to *bilateral* geometrical and kinematical (nonholonomic or not) constraints, is a necessary and sufficient condition for that system to remain in equilibrium in that frame. Here analytical means *deductive*, as opposed to synthetic or inductive, and therefore the PVW is taken as the sole and simplest axiom: With it alone one can determine all of the configurations of equilibrium of even the most general mechanical system. Also, and this makes the PVW superior to the elementary force/moment-free body diagram approach, one does not need to include the reactions/constraint forces in the analysis since their virtual work is *independently* zero; the latter is actually an equivalent formulation of the PVW. In concrete mechanical problems what one really employs is the *sufficiency* condition of the principle: If the virtual work is zero, then the originally motionless system remains in equilibrium. The purpose of this paper is to reexamine this static sufficiency condition not as an independent axiom, but starting from the general variational equation known as "Lagrange's Principle" (LP) or "d'Alembert's Principle in Lagrange's Form," specialize it to statics and deduce the PVW as a theorem, i.e., *start with LP as an axiom, set the virtual work equal to zero, and then derive (sufficient) conditions to maintain equilibrium.* It will be clearly demonstrated that to maintain equilibrium it is required that, in addition to the stated conditions, the system also be *scleronomic* in its *holonomic* constraints and catastatic in its *nonholonomic* ones. In practical terms this means that *all support surfaces of the system must be fixed* in the inertial frame. Actually, the entire argument of the paper applies intact, even if some or all of the kinematical constraints are additional holonomic ones in disguise, i.e., integrable constraints in differential form; possible limitations to truly nonintegrable constraints should not be hard to spot. The author is aware of only one rigorous formulation of the PVW stressing the importance of constraint stationarity for equilibrium: that by Hamel (1912). However, no explicit mention of nonholonomic constraints is made there. Most authors proceed from the PVW to LP, i.e., from statics to kinetics; this makes the detection of the importance of the stationarity of holonomic constraints rather difficult. Hamel, on the other hand, in a truly analytical fashion *deduces the PVW from LP*; this allows a clear focusing on the effect of the kinematic constitution of the system on its equilibrium. Here, Hamel's approach is enlarged to include both holonomic and (linear) nonholonomic constraints, in order to show *quantitatively* the precise effect of nonstationarity and acatastaticity on the violation of equilibrium. In the Appendix an alternative formulation of the PVW for nonstationary constraints by Gantmacher (1970) will also be discussed and compared with that of Hamel and with the findings of this paper.

2 Kinematical Background

For a full coverage see, e.g., Hamel (1912, 1949), and Gantmacher (1970). Consider an arbitrary finite mechanical system consisting of discrete or continuous material bodies, and subject to *bilateral*, holonomic and/or nonholonomic constraints. The *position* vectors of its particles, relative to an "origin" fixed in an inertial frame of reference R, are

$$\mathbf{r} = \mathbf{r}(P; q_1, \ldots, q_n, t) \equiv \mathbf{r}(q, t). \tag{1}$$

Here, P is the label identifying the various particles of the system, t is the time, and $q = q(t) = \{q_1, \ldots, q_n\}$ are a set of n-generalized *Lagrangean* (or true), *independent*, positional, *system* coordinates; P is independent of t; from now on the explicit dependence of $\mathbf{r}$ on P will be dropped. Usually the particle label is shown by a subscript on $\mathbf{r}$: $\mathbf{r}_k(q, t)$; $k = 1, \ldots, N$ = number of particles; the practice (1), reminiscent of

[1]School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

Manuscript received by the ASME Applied Mechanics Division, July 29, 1988; final revision, October 26, 1988.

continuum mechanics, simplifies the notation. The representation (1), with $n \equiv 3N-h$, takes care of the h holonomic constraints. Next, (*i*) the *velocity* $\mathbf{v}$, and *acceleration* $\mathbf{a}$, of a typical particle P (relative to R) are defined, respectively, by

$$\mathbf{v} \equiv \frac{d\mathbf{r}}{dt} = \sum_{i=1}^{n} \frac{\partial \mathbf{r}}{\partial q_i}\,\dot{q}_i + \frac{\partial \mathbf{r}}{\partial t}\,, \tag{2a}$$

and

$$\mathbf{a} \equiv \frac{d\mathbf{v}}{dt} = \frac{d}{dt}\left(\frac{d\mathbf{r}}{dt}\right) = \frac{d^2\mathbf{r}}{dt^2}\,, \tag{2b}$$

whereas, (*ii*) a *kinematically possible or admissible* displacement of P is

$$d\bar{r} = \sum_{i=1}^{n} \frac{\partial \mathbf{r}}{\partial q_i}\,dq_i + \frac{\partial \mathbf{r}}{\partial t}\,dt, \tag{3}$$

and (*iii*) a *virtual* displacement of the same P (again at a given instant t and configuration q, as for $d\mathbf{r}$) is

$$\delta\mathbf{r} = \sum_{i=1}^{n} \frac{\partial \mathbf{r}}{\partial q_i}\,\delta q_i \quad (\delta t = 0). \tag{4}$$

Clearly, $\delta\mathbf{r}$ in (4) is the *linear* and, since $\delta t = 0$, *homogeneous* part in the arbitrary increments $\delta q = \{\delta q_1, \ldots, \delta q_n\}$ of the q's of the expansion

$$\mathbf{r}(q+\delta q, t) - \mathbf{r}(q, t). \tag{5}$$

When only the h holonomic constraints are present, the $(n+1)$ differentials dq and dt in (3) are independent of each other, and so are the n, q's, and t. When, however, nonholonomic constraints are imposed on the system, the q's and t are not affected, but now the members of the two sets of dq's, dt, and δq's are no longer independent but have to satisfy the nonholonomic constraints in their $d-$ (= admissible/possible) or $\delta-$ (= virtual) forms. This interdependence of the dq's, dt, and of the δq's, respectively, matters not only in the derivation of equations of motion (equilibrium) from LP (PVW), but also in the equilibrium argument that follows. When the holonomic constraints are *stationary*:

$$\frac{\partial \mathbf{r}}{\partial t} = \mathbf{0}. \tag{6}$$

Then, as (3) and (4) show, *the classes of* $d\mathbf{r}$ *and* $\delta\mathbf{r}$ *are equivalent*; it is this $(\partial\mathbf{r}/\partial t)-$proportional term in (2*a*) and (3)$-$and *corresponding terms for nonholonomic constraints*$-$that necessitate a sharper formulation of the PVW. When $m(<n)$, additional nonholonomic constraints are imposed, then *additional* dt-proportional terms due to those new constraints "widen" the difference between $d\mathbf{r}$ and $\delta\mathbf{r}$. The "particle" forms of these constraints are Gantmacher (1970):

$$\mathcal{S}(\mathbf{B}_j \cdot \mathbf{v}) + B_j = 0 \quad [j = 1, \ldots, m\ (<n)], \tag{7}$$

where $\mathcal{S}$ (. . .) denotes discrete and/or continuous *summation over the entire material system* (à la Stieltjes), and $\mathbf{B}_j = \mathbf{B}_j(P; \mathbf{r}, t)$, $B_j = B_j(P, \mathbf{r}, t)$ are given or known functions. Substituting $\mathbf{r}$ and $\mathbf{v}$ from (1) and (2*a*) into (7) transforms it to

$$\mathcal{S}_{\mathbf{B}_j} \cdot \left(\sum_{i=1}^{n} \frac{\partial \mathbf{r}}{\partial q_i}\,\dot{q}_i + \frac{\partial \mathbf{r}}{\partial t}\right) + B_j = 0,$$

or the *system* form of the constraints:

$$\sum_{i=1}^{n} a_{ji}\dot{q}_i + a_j = 0 \quad [j = 1, \ldots, m(<n)], \tag{8}$$

where

$$a_{ji} = a_{ji}(q,t) \equiv \mathcal{S}\left(\mathbf{B}_j \cdot \frac{\partial \mathbf{r}}{\partial q_i}\right), \tag{9}$$

$$a_j = a_j(q,t) \equiv \mathcal{S}\left(\mathbf{B}_j \cdot \frac{\partial \mathbf{r}}{\partial t}\right) + B_j; \tag{10}$$

the a_{ji} and a_j are *known* functions of the q's and t. If

$$\frac{\partial a_{ji}}{\partial t} = 0, \quad \textit{and } a_j = 0, \tag{11}$$

the NH constraints are called *stationary*, and the system scleronomic in them. Here, however, another classification seems more useful: If

$$a_j = 0, \tag{12}$$

the NH constraints are called *catastatic*, otherwise *acatastatic*; in catastatic constraints, however, one may still have $a_{ji} = a_{ji}(q, t)$! As seen from (10), the $\{a_k\}$ consist of the nonstationary contributions of the holonomic constraints as well as the nonhomogeneous (acatastatic) contributions of the NH ones. Equation (8) holds for the actual *system* generalized velocities $\dot{q}$. For the *kinematically possible/admissible* and *virtual* (generalized) *system* displacements dq, dt, and δq, one has, respectively, the constraints

$$\sum_{i=1}^{n} a_{ji}dq_i + a_j dt = 0, \tag{13}$$

and

$$\sum_{i=1}^{n} a_{ji}\delta q_i = 0 \quad [j = 1, \ldots, m(<n)]. \tag{14}$$

Equation (13) shows that now the dq's depend on each other and on dt; *this is the reason for the additional* dt-*proportional contributions to* $d\mathbf{r}$. To incorporate (13) into (3) the following (special case of the method of *quasi-coordinates*$-$see Hamel (1949)) scheme is chosen: The m equations (13) are used to express the *first* m dq's $\{dq_1, \ldots, dq_m\}$ in terms of the remaining, independent, $(n-m)$ dq's $\{dq_{m+1}, \ldots, dq_n\}$ *and* dt. Thus

$$dq_j = \sum_{k=m+1}^{n} b_{jk}dq_k + b_j dt, \tag{15a}$$

$$b_{jk} = b_{jk}(q,t),\ b_j = b_j(q,t),\ [j = 1, \ldots, m;\ k = m+1, \ldots, n]; \tag{15b}$$

this idea, used in connection with the method of "embedding" of the constraints to LP, dates back to Hadamard (1895, 1899), Chaplygin (1895, 1897), Voronets (1901): Neimark and Fufaev (1972). Inserting (15) into (3), one successfully finds

$$d\mathbf{r} = \sum_{i=1}^{n} \frac{\partial \mathbf{r}}{\partial q_i}\,dq_i + \frac{\partial \mathbf{r}}{\partial t}\,dt \tag{16}$$

$$= \sum_{j=1}^{m} \frac{\partial \mathbf{r}}{\partial q_j}\,dq_j + \sum_{k=m+1}^{n} \frac{\partial \mathbf{r}}{\partial q_k}\,dq_k + \frac{\partial \mathbf{r}}{\partial t}\,dt$$

$$= \sum_{j=1}^{m} \frac{\partial \mathbf{r}}{\partial q_j}\left(\sum_{k=m+1}^{n} b_{jk}dq_k + b_j dt\right)$$

$$+ \sum_{k=m+1}^{n} \frac{\partial \mathbf{r}}{\partial q_k}dq_k + \frac{\partial \mathbf{r}}{\partial t}\,dt$$

$$= \sum_{k=m+1}^{n} \left(\sum_{j=1}^{m} \frac{\partial \mathbf{r}}{\partial q_j}\,b_{jk} + \frac{\partial \mathbf{r}}{\partial q_k}\right)dq_k$$

$$+ \left(\sum_{j=1}^{m} \frac{\partial \mathbf{r}}{\partial q_j}b_j + \frac{\partial \mathbf{r}}{\partial t}\right)dt,$$

or, in terms of the holonomic and nonholonomic (particle *and* system) base vectors, respectively,

$$\mathbf{e}_i \equiv \frac{\partial \mathbf{r}}{\partial q_i}\ ;\ \mathbf{e}_{n+1} \equiv \frac{\partial \mathbf{r}}{\partial t}\ (i=1, \ldots, n;\ q_{n+1} \equiv t,$$
$$\dot{q}_{n+1}=1,\ dq_{n+1}=dt); \tag{17}$$

$$\mathbf{a}_k \equiv \sum_{j=1}^{m} b_{jk}\mathbf{e}_j + \mathbf{e}_k;\ \mathbf{a}_{n+1} \equiv \sum_{j=1}^{m} b_j \mathbf{e}_j + \mathbf{e}_{n+1}$$
$$(k=m+1, \ldots, n), \tag{18}$$

finally

$$d\mathbf{r} = \sum_{k=m+1}^{n} \mathbf{a}_k dq_k + \mathbf{a}_{n+1}dt; \tag{19}$$

the $[(n-m)+1]$ dq's and dt are now independent. Similarly, for the particle and system virtual displacements, $\delta\mathbf{r}$ and δq: from (14)

$$\delta q_j = \sum_{k=m+1}^{n} b_{jk}\delta q_k, \tag{20}$$

instead of (15), and

$$\delta\mathbf{r} = \sum_{k=m+1}^{n} \mathbf{a}_k \delta q_k, \tag{21}$$

instead of (19) – the $(n-m)$ δq's are now independent. Equations (21) and (19) clearly show that *in the NH case, the classes of* $\mathrm{d}\mathbf{r}$ *and* $\delta\mathbf{r}$ *differ by*

$$\mathbf{a}_{n+1}dt = \left(\sum_{j=1}^{m} b_j\mathbf{e}_j + \mathbf{e}_{n+1}\right)dt = \left(\sum_{j=1}^{m} \frac{\partial \mathbf{r}}{\partial q_j} b_j + \frac{\partial \mathbf{r}}{\partial t}\right)dt. \tag{22}$$

Therefore, if (13) is catastatic, i.e., $a_j = 0$ $(j=1, \ldots, m)$, then $b_j = 0$ and (15) is also "catastatic," and $\mathbf{a}_{n+1} \rightarrow \mathbf{e}_{n+1}$ as before. The vectors $\{\mathbf{e}_i, \mathbf{e}_{n+1}\}$ and $\{\mathbf{a}_k, \mathbf{a}_{n+1}\}$ are fundamental to constrained system mechanics (equations of motion, etc.) – see Hamel (1949) and Gantmacher (1970).

3 Kinetics Background

According to the "Principle of Lagrange" (LP), the motion of a mechanical system subject to bilateral and "ideal" constraints is uniquely determined by the following differential variational equation

$$\mathcal{S} dm\mathbf{a}\cdot\delta\mathbf{r} = \mathcal{S} d\mathbf{f}\cdot\delta\mathbf{r}. \tag{23}$$

Here, $d\mathbf{f}$ represents the total (given) *impressed* force on the particle P of mass dm. On the other hand, Newton's "second law of motion" for P is

$$dm\mathbf{a} = d\mathbf{f} + d\mathbf{R}, \tag{24}$$

where $d\mathbf{R}$ = total (contact) *reaction* force on dm. LP states that the system of the "lost" forces $\{-d\mathbf{R}\} = \{d\mathbf{f} - dm\mathbf{a}\}$ is in "equilibrium," not in the elementary sense, but in that *the* $\{-d\mathbf{R}\}$*'s do not affect the acceleration state of the system*, or in work terms:

$$\mathcal{S}(-d\mathbf{R})\cdot\delta\mathbf{r} = 0; \tag{25}$$

combination of (24) with (25) produces the standard form (23).

4 From LP to the PVW/Equilibrium

Conditions for equilibrium: (*i*) *Necessary Conditions*: If the system is in equilibrium (relative to R), then $\mathbf{a} = \mathbf{0}$ and therefore from (7) the (first order) virtual work, $\delta' W$, vanishes:

$$\delta' W \equiv \mathcal{S} d\mathbf{f}\cdot\delta\mathbf{r} = 0. \tag{26}$$

(*ii*) *Sufficiency Conditions*: If $\delta' W = 0$ from an initial moment t_o and for some time thereafter, then as (23) shows

$$\mathcal{S} dm\mathbf{a}\cdot\delta\mathbf{r} = 0, \quad \text{for } t \geqq t_o. \tag{27}$$

The consequences of (26) and (27) for equilibrium will now be examined: Substituting

$$d\mathbf{r} - \mathbf{a}_{n+1}dt = (\mathbf{v} - \mathbf{a}_{n+1})dt, \tag{28}$$

for $\delta\mathbf{r}$ into (27) and cancelling dt $(\neq 0)$, yields

$$\mathcal{S} dm\mathbf{a}\cdot\mathbf{v} = \mathcal{S} dm\mathbf{a}\cdot\mathbf{a}_{n+1}, \tag{29}$$

or, since

$$T \equiv \mathcal{S}\frac{1}{2}\, dm\bar{v}^2$$
$$= \text{kinetic energy of system (relative to } R), \tag{30}$$

finally

$$\frac{dT}{dt} = \mathcal{S} dm\mathbf{a}\cdot\mathbf{v} = \mathcal{S} dm\mathbf{a}\cdot\mathbf{a}_{n+1} \tag{31}$$
$$(= \mathcal{S} d\mathbf{f}\cdot\mathbf{a}_{n+1} + \mathcal{S} d\mathbf{R}\cdot\mathbf{a}_{n+1}).$$

Integrating both sides of the aforementioned equation between t_o and t $(\geqq t_o)$, and with $T_o \equiv T(t_o)$, $T \equiv T(t)$, yields

$$T - T_o = \int_{t_o}^{t} [\mathcal{S} dm\mathbf{a}\cdot\mathbf{a}_{n+1}]dt;\ \mathbf{a}_{n+1} \equiv \sum_{j=1}^{m} \frac{\partial \mathbf{r}}{\partial q_j} b_j + \frac{\partial \mathbf{r}}{\partial t}. \tag{32}$$

One notes that (32) follows from (26) and also from the more general (and perhaps more realistic) condition

$$\int_{t_o}^{t} \delta' W dt = 0. \tag{33}$$

Equation (32) (with (22)) holds the key to the understanding of the constraints/equilibrium connection:

$$(i)\ \text{if } \mathbf{a}_{n+1} = \mathbf{0}, \text{ then } T = T_o, \tag{34}$$

and since $\mathbf{v}_o \equiv \mathbf{v}(t_o) = \mathbf{0} \rightarrow T_o = 0$, it follows that $T = 0$ or, since T is a *positive-definite function* $(-\mathrm{al})$ *in the* $\mathbf{v}^2 = v^2$, *all* the $\mathbf{v}$'s are zero for $t \geqq t_o$, i.e., after t_o the system remains in equilibrium in R. As mentioned in the Introduction, the meaning of (17) is that all the system support surfaces are *fixed* in R. For most structures the earth can serve as R; then a system fixed relative to the earth is *scleronomic*, and the forces from the earth to it are *external reactions* to be calculated – of course, this refers to the *contact* forces, not the gravitational/field ones. However, if the earth is assumed to have a *given* motion (unaffected by the system's action), then the system is *rheonomic* and the corresponding forces are again unknown reactions. If, finally, the earth *interacts* with the system, then *together* they constitute a new scleronomic system and the forces between them are *impressed*.

(*ii*) if $\mathbf{a}_{n+1} \neq \mathbf{0}$, then as (32) shows, in general,

$$\Delta T \equiv T - T_o \neq 0, \tag{35}$$

i.e., the system moves away from the original equilibrium configuration, *even though* $\delta' W = 0$ for $t \geqq t_o$, and $\mathbf{v}_o = \mathbf{0}$! Weaker (special) conditions for equilibrium result:

$$(a)\ \text{if } \mathbf{a}_{n+1} \neq \mathbf{0}, \quad \text{but } \int_{t_o}^{t} [\mathcal{S} dm\mathbf{a}\cdot\mathbf{a}_{n+1}]dt = 0, \tag{36}$$

$$(b)\ \text{if } \mathbf{a}_{n+1} \neq \mathbf{0}, \text{ but } \mathbf{a}\cdot\mathbf{a}_{n+1} = 0, \quad \text{for } t \geqq t_o. \tag{37}$$

The PVW now reads: An originally motionless mechanical system, relative to an inertial frame, remains in equilibrium if and only if: (*i*) $\delta' W = 0$, and (*ii*) its holonomic and nonholonomic constraints are such that *the*

right side of (32) vanishes. Here are some additional possible consequences of the vanishing of $\delta' W$: Substituting (28) for $\delta\mathbf{r}$ into (26) yields

$$\mathcal{S}d\mathbf{f}\cdot\mathbf{v}=\mathcal{S}d\mathbf{f}\cdot\mathbf{a}_{n+1}. \tag{38}$$

This shows that, again, unless $\mathbf{a}_{n+1} = \mathbf{0}$, or some other weaker condition that ends up annihilating the right side of (38), the "power of the impressed forces" (= left side of (38)) is not zero! Finally, it should be noted that the foregoing discussion *avoids the use of unknown Lagrangean multipliers*, and uses virtual work, instead of power, which is more in line with standard tradition of analytical statics (see e.g., Sommerfeld (1964)).

References

Gantmacher, F., 1970, *Lectures in Analytical Mechanics*, Mir, Moscow, (originally in Russian, 1966), pp. 9-12, 25-26, 60-61.

Hamel, G., 1912, *Elementare Mechanik*, B. G. Teubner, Leipzig. (reprinted by Johnson Reprint Co., New York, 1965), pp. 469-471, 474-478.

Hamel, G., 1949, *Theoretische Mechanik*, Springer, Berlin and New York, pp. 361-363, 473-479.

Neimark, Ju. I., and Fufaev, N. A., 1972, *Dynamics of Nonholonomic Systems*, American Mathematical Society, Providence, R.I. (translated from the Russian, 1967), pp. 100-120.

Sommerfeld, A., 1964, *Mechanics*, Academic Press, New York (translated from the fourth German edition, 1948), pp. 66-69.

APPENDIX

Gantmacher's Formulation and its Relation With the Present Work

The distinguished applied mathematician F. Gantmacher (1970) formulates the PVW as follows: "For some position (compatible with constraints) of a system to be an equilibrium position, it is necessary and sufficient that in this position, the sum of the works of effective forces on any virtual displacements of the system be zero." And he adds "If the constraints are nonstationary, then the term "compatible with constraints" signifies that they are satisfied for any t if in them we put (in our notation) $\mathbf{r} = \mathbf{r}_o$ and $\mathbf{v} = \mathbf{0}$," and "It is then assumed that (our) equation (26) holds for any value of t if in the expression for $\mathbf{d}\bar{f}$ we put all $\mathbf{r} = \mathbf{r}_o$ and all $\mathbf{v} = \mathbf{0}$." Gantmacher gives no explanation for his last "rule," and presents the PVW as an independent axiom *before* LP. The connection of the aforementioned with the findings of this paper is worth examining: if $\mathbf{v} = \mathbf{0}$, then from (2*a*).

$$\sum_{i=1}^{n}\frac{\partial\mathbf{r}}{\partial q_i}\dot{q}_i+\frac{\partial\mathbf{r}}{\partial t}=\mathbf{0}, \quad \text{or} \quad \sum_{i=1}^{n}\mathbf{e}_i\dot{q}_i+\mathbf{e}_{n+1}=\mathbf{0}. \tag{A1}$$

If the system is *holonomic* the n $\dot{q}$'s are *independent*, and since $\partial\mathbf{r}/\partial q_i \equiv \mathbf{e}_i \neq \mathbf{0}$,

$\dot{q}_i = 0\ (i = 1, \ldots, n)$ *and*

$$\bar{e}_{n+1}\equiv\frac{\partial\mathbf{r}}{\partial t}=\mathbf{0}\rightarrow\text{stationary constraints.} \tag{A2}$$

This shows that, since $q_i(t) = q_{i,o} =$ constant, the system will remain in equilibrium. If, on the other hand, the system is *nonholonomic*, then from (19),

$$\mathbf{v}=\sum_{k=m+1}^{n}\mathbf{a}_k\dot{q}_k+\mathbf{a}_{n+1}=\mathbf{0}, \tag{A3}$$

and thus since the $(n-m)$ "base" vectors $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_n$ are *independent*,

$$\dot{q}_k=0 \quad (k=m+1,\ldots,n) \quad \text{and} \quad \mathbf{a}_{n+1}=\mathbf{0}, \tag{A4}$$

and noting (15) and (18), finally

$\dot{q}_i = 0\ (i = 1, \ldots, n)$, $b_j = 0\ [\rightarrow a_j = 0$, since $\mathbf{e}_j \neq \mathbf{0}]$

$$(j=1,\ldots,m),\ \mathbf{e}_{n+1}\equiv\frac{\partial\mathbf{r}}{\partial t}=\mathbf{0}, \tag{A5}$$

i.e., the holonomic constraints are stationary, the nonholonomic ones are catastatic, and $q_i = q_i(t) =$ constant. These findings are equivalent with those of this paper. In sum: (*i*) one sets either $\delta' W = 0$ *and* $\mathbf{v} = \mathbf{0}$ (in the constraints and forces), or (*ii*) $\delta' W = 0$ *and* $\partial\mathbf{r}/\partial t = \mathbf{0}$, a_j (or B_j) $= 0$. Therefore, in both the holonomic and nonholonomic cases the formulations of Gantmacher and that of Hamel and this paper are fundamentally equivalent, although the second approach is clearer both logically and physically.

Analysis of Thermal Stresses in a Ceramic-to-Metal Cylindrical Joint With a Homogeneous Elastic Medium

Osamu Kimura[2] and Toshio Kawashima[3]

It is well known that large thermal stresses arise in a ceramic-to-metal joint because of thermal expansion mismatch. Therefore, to avoid fracture of a joint, it has been recommended a joint in which thermal expansion coefficient of a ceramic member is smaller than that of a metal member. The joint is called a "compressive" joint, because it was believed that only compressive stresses occur in the ceramic member. Since ceramic is about ten times as strong in compression as in tension, this joint is considered stronger than a tensile joint.

A finite element calculation (FEC), however, has shown that a large tensile stress is generated even in ceramic part of a compressive joint. The aim of this study is to clarify stress distribution in a compressive joint by solving analytically the stress equation of the joint.

For convenience, we deal with a special joint in which the elastic constants of the ceramic member are assumed to be the same as those of the metal member. The result of the analysis gives the same stress distribution obtained by FEC, and its value of the tensile stress agrees fairly well with that of FEC.

I Introduction

Ceramic-to-metal joints have been investigated using different bonding methods due to their industrial importance. Generally, thermal expansion coefficients of members to be bonded are different. Therefore, when a joint is cooled to room temperature, thermal stresses are caused in the joint by the resulting thermal expansion mismatch.

Several stress calculations were done on joints of other types of Preist and Talcott (1959) or by Cole and Inge (1961). Also stress analyses were carried out both experimentally and theoretically on glass-to-metal seals of coaxial cylinders used in vacuum tubes by Sutton (1958), and of very thin sheets by Oel and Fréchette (1967). In these cases, internal stresses in glass are measurable using photoelastic techniques, so that detailed information is available on stress distribution in seals. No detailed stress calculation, however, has been carried out in this butt joint, except in the discussion by Bogy (1968), who pointed out the occurrence of singularities of the type $r^{-\alpha}$ in the stress field in a wedge with dissimilar media, where r is distance from the bonded edge.

[2]Department of Materials Science and Ceramic Technology, Sagami Institute of Technology, Tsujido, Fujisawa, Kanagawa, 251, Japan.

[3]Department of Common Education, Ashikaga Institute of Technology, Ashikaga Tochigi, 326, Japan.

Manuscript received by the ASME Applied Mechanics Division, April 8, 1988; final revision, October 25, 1988.

right side of (32) vanishes. Here are some additional possible consequences of the vanishing of $\delta' W$: Substituting (28) for $\delta\mathbf{r}$ into (26) yields

$$\mathcal{S}d\mathbf{f}\cdot\mathbf{v}=\mathcal{S}d\mathbf{f}\cdot\mathbf{a}_{n+1}. \tag{38}$$

This shows that, again, unless $\mathbf{a}_{n+1} = \mathbf{0}$, or some other weaker condition that ends up annihilating the right side of (38), the "power of the impressed forces" (= left side of (38)) is not zero! Finally, it should be noted that the foregoing discussion *avoids the use of unknown Lagrangean multipliers*, and uses virtual work, instead of power, which is more in line with standard tradition of analytical statics (see e.g., Sommerfeld (1964)).

References

Gantmacher, F., 1970, *Lectures in Analytical Mechanics*, Mir, Moscow, (originally in Russian, 1966), pp. 9-12, 25-26, 60-61.

Hamel, G., 1912, *Elementare Mechanik*, B. G. Teubner, Leipzig. (reprinted by Johnson Reprint Co., New York, 1965), pp. 469-471, 474-478.

Hamel, G., 1949, *Theoretische Mechanik*, Springer, Berlin and New York, pp. 361-363, 473-479.

Neimark, Ju. I., and Fufaev, N. A., 1972, *Dynamics of Nonholonomic Systems*, American Mathematical Society, Providence, R.I. (translated from the Russian, 1967), pp. 100-120.

Sommerfeld, A., 1964, *Mechanics*, Academic Press, New York (translated from the fourth German edition, 1948), pp. 66-69.

APPENDIX

Gantmacher's Formulation and its Relation With the Present Work

The distinguished applied mathematician F. Gantmacher (1970) formulates the PVW as follows: "For some position (compatible with constraints) of a system to be an equilibrium position, it is necessary and sufficient that in this position, the sum of the works of effective forces on any virtual displacements of the system be zero." And he adds "If the constraints are nonstationary, then the term "compatible with constraints" signifies that they are satisfied for any t if in them we put (in our notation) $\mathbf{r} = \mathbf{r}_o$ and $\mathbf{v} = \mathbf{0}$," and "It is then assumed that (our) equation (26) holds for any value of t if in the expression for $\mathbf{d}\bar{f}$ we put all $\mathbf{r} = \mathbf{r}_o$ and all $\mathbf{v} = \mathbf{0}$." Gantmacher gives no explanation for his last "rule," and presents the PVW as an independent axiom *before* LP. The connection of the aforementioned with the findings of this paper is worth examining: if $\mathbf{v} = \mathbf{0}$, then from (2*a*).

$$\sum_{i=1}^{n}\frac{\partial\mathbf{r}}{\partial q_i}\dot{q}_i+\frac{\partial\mathbf{r}}{\partial t}=\mathbf{0}, \quad \text{or} \quad \sum_{i=1}^{n}\mathbf{e}_i\dot{q}_i+\mathbf{e}_{n+1}=\mathbf{0}. \tag{A1}$$

If the system is *holonomic* the n $\dot{q}$'s are *independent*, and since $\partial\mathbf{r}/\partial q_i \equiv \mathbf{e}_i \neq \mathbf{0}$,

$\dot{q}_i = 0\ (i = 1, \ldots, n)$ *and*

$$\bar{e}_{n+1}\equiv\frac{\partial\mathbf{r}}{\partial t}=\mathbf{0}\rightarrow\text{stationary constraints.} \tag{A2}$$

This shows that, since $q_i(t) = q_{i,o} =$ constant, the system will remain in equilibrium. If, on the other hand, the system is *nonholonomic*, then from (19),

$$\mathbf{v}=\sum_{k=m+1}^{n}\mathbf{a}_k\dot{q}_k+\mathbf{a}_{n+1}=\mathbf{0}, \tag{A3}$$

and thus since the $(n-m)$ "base" vectors $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_n$ are *independent*,

$$\dot{q}_k=0 \quad (k=m+1,\ldots,n) \quad \text{and} \quad \mathbf{a}_{n+1}=\mathbf{0}, \tag{A4}$$

and noting (15) and (18), finally

$\dot{q}_i = 0\ (i = 1, \ldots, n)$, $b_j = 0\ [\rightarrow a_j = 0$, since $\mathbf{e}_j \neq \mathbf{0}]$

$$(j=1,\ldots,m),\ \mathbf{e}_{n+1}\equiv\frac{\partial\mathbf{r}}{\partial t}=\mathbf{0}, \tag{A5}$$

i.e., the holonomic constraints are stationary, the nonholonomic ones are catastatic, and $q_i = q_i(t) =$ constant. These findings are equivalent with those of this paper. In sum: (*i*) one sets either $\delta' W = 0$ *and* $\mathbf{v} = \mathbf{0}$ (in the constraints and forces), or (*ii*) $\delta' W = 0$ *and* $\partial\mathbf{r}/\partial t = \mathbf{0}$, a_j (or B_j) $= 0$. Therefore, in both the holonomic and nonholonomic cases the formulations of Gantmacher and that of Hamel and this paper are fundamentally equivalent, although the second approach is clearer both logically and physically.

Analysis of Thermal Stresses in a Ceramic-to-Metal Cylindrical Joint With a Homogeneous Elastic Medium

Osamu Kimura[2] and Toshio Kawashima[3]

It is well known that large thermal stresses arise in a ceramic-to-metal joint because of thermal expansion mismatch. Therefore, to avoid fracture of a joint, it has been recommended a joint in which thermal expansion coefficient of a ceramic member is smaller than that of a metal member. The joint is called a "compressive" joint, because it was believed that only compressive stresses occur in the ceramic member. Since ceramic is about ten times as strong in compression as in tension, this joint is considered stronger than a tensile joint.

A finite element calculation (FEC), however, has shown that a large tensile stress is generated even in ceramic part of a compressive joint. The aim of this study is to clarify stress distribution in a compressive joint by solving analytically the stress equation of the joint.

For convenience, we deal with a special joint in which the elastic constants of the ceramic member are assumed to be the same as those of the metal member. The result of the analysis gives the same stress distribution obtained by FEC, and its value of the tensile stress agrees fairly well with that of FEC.

I Introduction

Ceramic-to-metal joints have been investigated using different bonding methods due to their industrial importance. Generally, thermal expansion coefficients of members to be bonded are different. Therefore, when a joint is cooled to room temperature, thermal stresses are caused in the joint by the resulting thermal expansion mismatch.

Several stress calculations were done on joints of other types of Preist and Talcott (1959) or by Cole and Inge (1961). Also stress analyses were carried out both experimentally and theoretically on glass-to-metal seals of coaxial cylinders used in vacuum tubes by Sutton (1958), and of very thin sheets by Oel and Fréchette (1967). In these cases, internal stresses in glass are measurable using photoelastic techniques, so that detailed information is available on stress distribution in seals. No detailed stress calculation, however, has been carried out in this butt joint, except in the discussion by Bogy (1968), who pointed out the occurrence of singularities of the type $r^{-\alpha}$ in the stress field in a wedge with dissimilar media, where r is distance from the bonded edge.

[2]Department of Materials Science and Ceramic Technology, Sagami Institute of Technology, Tsujido, Fujisawa, Kanagawa, 251, Japan.

[3]Department of Common Education, Ashikaga Institute of Technology, Ashikaga Tochigi, 326, Japan.

Manuscript received by the ASME Applied Mechanics Division, April 8, 1988; final revision, October 25, 1988.

Recently, a finite element calculation (FEC) was conducted on a Steel-Alumina joint by Suganuma et al. (1984). They showed that a maximum tensile stress occurs in the alumina part of the joint. The aim of this study is, therefore, to calculate analytically the thermal stress distribution in the ceramic-metal joint to check the calculation by FEC. Because of some difficulties in treating general cases, we treat a special ceramic-metal joint in which the elastic constants of the ceramic member are the same as those of the metal member, namely, a joint with uniform elastic medium (hereafter, referred as a "homogeneous" joint). The present model is, of course, of limited applicability to actual cases, however, the analysis of the joint may serve as a starting point for the analysis of dissimilar joints. Because of the elastic homogeneity of the joint, the stresses in it can be treated as those generated in a single rod using the formal elastic theory, and no singularity in its stress field is supposed to occur, according to Bogy's treatment (Bogy, 1968).

II Analysis

The procedure in analysis proceeds as follows, which is analogous to that proposed by Barton (1941) for the thermal stress problem of a circular cylinder.

We treat the joint as a rod with thermal expansion coefficient $\alpha-\beta$ and with no thermal conductivity at the interface, where α and β correspond to the thermal expansion coefficient of the lower half and the upper half of the joint, respectively. Thus, the thermal stresses in the joint are exactly the same as those stresses generated in the rod when its upper half is heated to t, while its lower half is still maintained at temperature t_0. This process corresponds to establish thermal expansion mismatch of $\xi\ [=(\alpha-\beta)\ (t-t_0)]$ between the upper and lower half of the rod. This mismatch may be diminished by plastic deformation or creep during cooling process of the joint.

It is easily seen that angular displacement V vanishes and other displacements U and W are independent of θ according to the cylindrical symmetry of strains.

(1) The First Deformation. In this step, the radial expansion of the upper half of the rod is reduced to zero by applying an external pressure of $P = E\xi/(1-\nu)$, where E = Young's modulus and ν = Poisson's ratio. The stresses in the upper half due to this pressure are $\sigma_r = \sigma_\theta = -P$, $\sigma_z = \tau_{rz} = 0$ and axial deformation $W = (1+\nu)\xi/(1-\nu)$, while all stresses and deformations are zero in the lower half.

(2) The Second Deformation. In this step, the pressure is removed from the rod surface by superimposing on the upper half of the rod an external tensile load of $P = E\xi/(1-\nu)$. Then, radial displacement U occurs as well as W. The final stresses and displacements are given by the algebraic sum of the first and the second ones. Then, we obtain the relations for the second surface tractions

$$(\sigma_r)_{r=a} = \begin{cases} P \text{ for } z>0 & (\tau_{rz})_{r=a}=0 \text{ everywhere,} \\ 0 \text{ for } z<0. \end{cases} \tag{1}$$

This condition can be divided into following conditions (*a*) and (*b*).

Condition (*a*).

$$(\sigma_r)_{r=a}=P/2, \quad (\tau_{rz})_{r=a}=0 \text{ everywhere,}$$

Condition (*b*).

$$(\sigma_r)_{r=a} = \begin{cases} P/2 \text{ for } z>0 & (\tau_{rz})_{r=a}=0 \text{ everywhere.} \\ -P/2 \text{ for } z<0. \end{cases} \tag{2}$$

For solving the problem, we utilize Love's biharmonic stress function (see text by Timoshenko and Goodier (1970a)),

$$\nabla^4\Phi=0 \tag{3}$$

where stresses are given by

$$\sigma_r=\partial/\partial z(\nu\nabla^2\Phi-\partial^2\Phi/\partial r^2),$$

$$\sigma_\theta=\partial/\partial z(\nu\nabla^2\Phi-1/r\partial\Phi/\partial r)$$

$$\sigma_z=\partial/\partial z((2-\nu)\nabla^2\Phi-\partial^2\Phi/\partial z^2),$$

$$\tau_{rz}=\partial/\partial r((1-\nu)\nabla^2\Phi-\partial^2\Phi/\partial z^2)$$

with

$$\nabla^2=\partial^2/\partial r^2+1/r\partial/\partial r+\partial^2/\partial z^2. \tag{4}$$

Since condition (*a*) corresponds to the problem for uniform tensile load of $P/2$ applied on the outer surface of the rod, Φ_a is given by

$$\Phi_a=-(1-\nu)Pr^2z/4(1+\nu)+(2-\nu)Pz^3/6(1+\nu). \tag{5}$$

This gives

$$\sigma_r=\sigma_\theta=P/2, \quad \sigma_z=\tau_{rz}=0 \text{ everywhere.} \tag{6}$$

On the other hand, the solution for condition (*b*) was already given by Barton (1941) and Rankin (1944). Following Timoshenko and Goodier (1970b), the potential Φ_b is given by

$$\Phi_b=P/\pi\cdot\int_0^\infty [\rho I_0(kr)-krI_1(kr)]\cos(kz)/k^4F_k(a)\cdot dk$$

with

$$\rho=2(1-\nu)+ka\,I_0(ka)/I_1(ka)$$

$$F_k(a)=ka[I_0^2(ka)-I_1^2(ka)]/I_1(ka)-2(1-\nu)I_1(ka)/ka, \tag{7}$$

where $I_0(x)$ and $I_1(x)$ are modified Bessel function of zero order and first order, respectively.

Since stresses are the algebraic sum of the first and the secondary stresses, we have

$$\sigma_r=P/\pi\cdot\int_0^\infty [F_k(r)/F_k(a)-1]\sin(kz)/k\cdot dk$$

$$\sigma_\theta=P/\pi\cdot\int_0^\infty [G_k(r)/F_k(a)-1]\sin(kz)/k\cdot dk$$

$$\sigma_z=P/\pi\cdot\int_0^\infty [H_k(r)/F_k(a)]\sin(kz)/k\cdot dk$$

$$\sigma_{rz}=P/\pi\cdot\int_0^\infty [J_k(r)/F_k(a)]\cos(kz)\cdot dk \tag{8}$$

where, $F_k(r)$, $G_k(r)$, $H_k(r)$, and $J_k(r)$ are given by

$$\begin{aligned} F_k(r) &= [1+kaI_0(ka)/I_1(ka)]I_0(kr) \\ &\quad -[aI_0(ka)I_1(kr)/rI_1(ka)]-krI_1(kr)-2(1-\nu)I_1(kr)/kr, \\ G_k(r) &= -(1-2\nu)I_0(kr)+2(1-\nu)I_1(kr)/kr \\ &\quad +aI_0(ka)I_1(kr)/rI_1(ka), \\ H_k(r) &= [2-kaI_0(ka)/I_1(ka)]I_0(kr)+krI_1(kr), \\ J_k(r) &= aI_0(ka)I_1(kr)/I_1(ka)-rI_0(kr). \end{aligned}$$

III Results and Discussion

(1) Maximum Values of Stresses and Some Limiting Cases. All the stresses should increase as z approaches to zero, namely, near the joined interface. In this case, by changing the integral variable from k to $K\ (=kz)$, we find that argument of $F(k)$ goes to infinity as z approaches to zero. Thus, we can utilize asymptotic expansion for $I_0(x)$ and $I_1(x)$ (see Whittacker and Watson (1935), for example)

$$I_0(x)\sim\exp(x)/\sqrt{2\pi x}\cdot[1+1/8x+\ldots]$$

$$I_1(x)\sim\exp(x)/\sqrt{2\pi x}\cdot[1-3/8x+\ldots]. \tag{9}$$

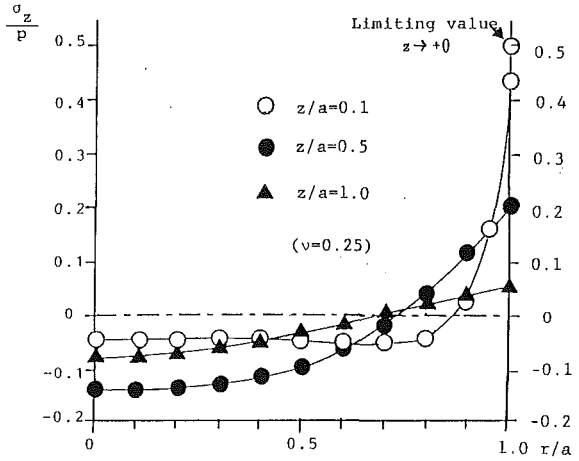


Fig. 1 Radial distribution of σ_z for various z

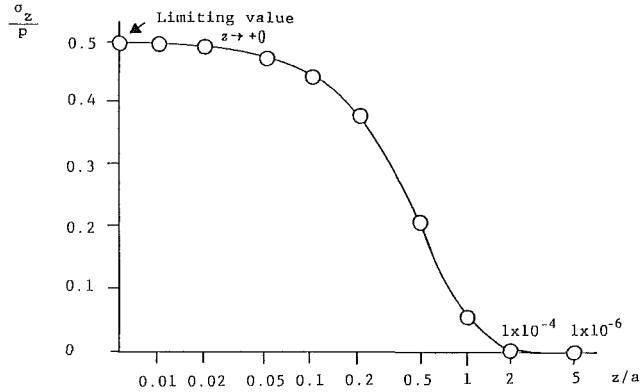


Fig. 2 Change of σ_z along axial direction at the surface of the rod ($r = a$)

So, we get maximum absolute values for σ_r and σ_θ at $r=0$ and $z\rightarrow +0$ by virtue of equation (8)

$$\lim_{\substack{z\rightarrow +0\\ r\rightarrow 0}} \sigma_r = \lim_{\substack{z\rightarrow +0\\ r\rightarrow 0}} \sigma_\theta = -\lim P/\pi \cdot \int_0^\infty \sin(kz)/k \cdot dk = -P/2. \tag{2}$$

These stress are compressive. In the same way, we find a maximum value for σ_z at $r=a$ and $z\rightarrow +0$ to be $P/2$. This stress is tensile and most critical for ceramics as already mentioned. Its existence was first shown by Suganuma et al. (1984) using FEC. Also, we can find that shear stress τ_{rz} reaches a maximum value of P/π at $r\rightarrow a$, $z\rightarrow 0$.

(2) Comparison of the Theoretical Value With the Value by Finite Element Calculation. Suganuma et al. calculated the maximum tensile stress to be 0.94 GPa in Steel-Alumina joint at $r=a$ and $z\rightarrow +0$ (in alumina side). They used the following values for elastic constants shown in Table 1, on the condition that t = 1000°C and t_0 = 25°C. Substituting these values for alumina or steel into the value of $P/2$, the maximum tensile stress gives the values $\sigma_{z\max} = P/2$ = [1.47 GPa for alumina and 0.75 GPa for steel].

The value of alumina is about 60 percent larger than that by FEC and that of steel is 20 percent smaller. The numerical agreement of our theory with FEC is found fairly good.

Table 1 Values of Young's modulus (E), Poisson's ratio (ν) and thermal expansion coefficient (α) for alumina and steel

Material	E(MPa)	ν	$\alpha(\times 10^{-6})$
Alumina	372000	0.27	8.1
Steel	196000	0.25	14.0

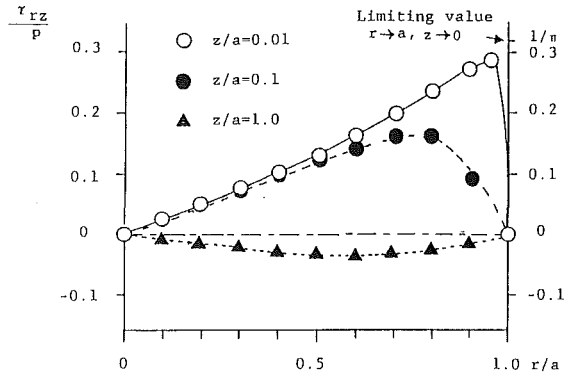


Fig. 3 Radial distribution of τ_{rz} for various z

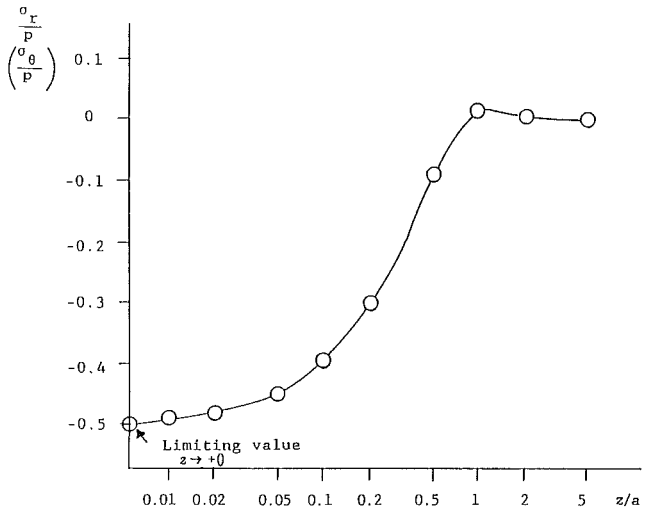


Fig. 4 Change of σ_r and of σ_θ along axial direction at the center axis of the rod ($r = 0$)

(3) Stress Distribution. Figure 1 shows radial distribution of σ_z for positive axial direction, that is, for positive z. Hereafter, we put Poisson's ratio ν = 0.25, which is usual for ceramics. As can be seen for small r, namely, in the inner part of the rod, this stress is compressive, while it becomes tensile for larger r, namely, at the outer side of the rod. Also, it becomes very small for z over $2a$. It means that the effect of bonding becomes negligibly small for z over $2a$, corresponding to Saint Venin's principle.

Figure 2 shows changes in axial direction of σ_z at the surface of the rod, namely, at $r=a$. We can see that it approaches to $P/2$ at the boundary surface and decreases rapidly for z over $2a$.

Figure 3 shows radial distribution of τ_{rz} for various z. As can be seen, it increases very rapidly as z goes to zero and reaches a maximum as r approaches to a. Also, it changes its sign at z nearly equal to a. Then, it becomes very small for z over $2a$.

Figure 4 shows change in axial direction of σ_r at the center axis of the rod, which coincides in value with σ_θ in the case of ν = 0.25. They reach a minimum (maximum absolute value) of $-P/2$ at the boundary surface and beyond $2a$ become negligibly small, as other stresses. Both σ_r and σ_θ are compressive almost everywhere, and their absolute values reach maximum at the center axis of the rod and decrease with increasing r, contrary to σ_z.

As can be seen in the figures, no singularity is found to occur in the stress field, as already mentioned.

IV Conclusion

The exact analytical solution is found for the thermal

stresses in the joint with elastically homogeneous medium by treating it as a single rod with abrupt temperature gradient. The obtained result is found reasonable concerning stress distributions. Also, the maximum value of tensile stress agrees fairly well with that by finite element calculation conducted by Suganuma et al. (1984). In another paper (to be published), the theory will be extended to thermal stresses in a usual joint with dissimilar elastic medium.

Acknowledgment

The authors would like to express their sincere thanks to Prof. T. Dezaki at Ashikaga Institute of Technology for his encouragement and to Dr. S. Ishioka at Tohoku University for his guidance and discussion.

References

Barton, M. V., 1941, "The Circular Cylinder With a Band of Uniform Pressure on a Finite Length of the Surface," ASME JOURNAL OF APPLIED MECHANICS, Vol. 8, pp. 97–104.

Bogy, D. B., 1968, "Edge-Bonded Dissimilar Orthogonal Wedge Under Normal and Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 460–466.

Cole, Jr., S. S., and Inge, J. E., 1961, "Calculation and Measurement of Stress in a Ceramic-to-Metal Seal," *American Ceramic Society Bulletin*, Vol. 40, pp. 738–743.

Oel, H. J., and Fréchette, V. D., 1967, "Stress Distribution in Multiphase System: I, Composites With Planar Interfaces," *Journal of the American Ceramic Society*, Vol. 50, pp. 542–549.

Preist, D. H., and Talcott, R., 1959, "Thermal Stresses in Ceramic Cylinders Used in Vacuum Tubes," *American Ceramic Society Bulletin*, Vol. 38, pp. 99–105.

Rankin, A. W., 1944, "Shrink-Fit Stresses and Deformations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 11, pp. 77–85.

Suganuma, K., Okamoto, T., Koizumi, M., and Shimada, M., 1984, "Effect of Interlayers in Ceramic-Metal Joints with Thermal Expansion Mismatches," *Journal of the American Ceramic Society*, Vol. 67, p. C256.

Sutton, P. M., 1958, "Stress Measurement in Circular Cylinders," *Journal of American Ceramic Society*, Vol. 41, pp. 103–109.

Timoshenko, S. P., and Goodier, J. N., 1970a, *Theory of Elasticity*, 3rd ed., McGraw-Hill, Singapore, pp. 380–383.

Timoshenko, S. P., and Goodier, J. M., 1970b, Ibid., pp. 425–428.

Whittaker, E. T., and Watson, G. N., 1935, *A Course of Modern Analysis*, 4th ed., Cambridge University Press, London, pp. 372–373.

A Study of the Effects of Baffles on Rotating Compressible Flows

Max D. Gunzburger,[4] Houston G. Wood,[5] and Rosser L. Wayland[6]

Onsager's pancake equation for the fluid dynamics of a gas centrifuge is modified for the case of centrifuges with baffles which render the flow domain doubly connected. A finite element algorithm is used for solving the mathematical model and to compute numerical examples for flow fields induced by thermal boundary conditions and by mass injection and extraction.

I Introduction

In the past decade numerous papers have appeared which address the dynamics of rotating compressible flows in a centrifuge. The primary purpose of these efforts has been directed towards modeling the diffusive processes in binary mixtures which enable the separation of species, with the chief application being separating the isotopes of uranium hexafluoride. The isotopically-altered uranium is then used as fuel for nuclear power reactors.

The purpose of the work presented in this paper is to study the effects on the flow field, due to the inclusion of baffles inside the centrifuge. The flow through the baffle holes and the effect of the baffle on the flow field have also been analyzed by Kai (1977), Lopez (1977), and Soubbaramayer (1979).

The inclusion of the baffle in the analysis of the flow field renders the domain to be doubly connected. The authors cited in the previous paragraph have used finite difference methods to solve the equations of motion in primitive variable form so that they encountered no new problems due to the multiple connectivity of the computational domain. However, if the equations of motion are formulated in terms of a stream function or a potential function as is the case for Onsager's model, care must be taken to assure that physical quantities such as the pressure are single-valued functions. This type of problem has been addressed for the incompressible case by Sood and Elrod (1974), Israeli and Ungarish (1977), Girault and Raviart (1979), and Gunzburger and Peterson (1985), and for the compressible case by Greenspan (1981) and Viecelli (1984).

II Mathematical Models

In an unpublished report, Lars Onsager showed that within a rapidly rotating centrifuge the equations of motion can be reduced to a single differential equation to be solved for a master potential function χ. This equation is

$$(e^x(e^x\chi_{xx})_{xx})_{xx}+B^2\chi_{yy}=F(x,y) \tag{1}$$

where $B^2 = \mathrm{Re}S^{1/2}/4A^6$, $Re = \rho_w\Omega a^2/\mu$, $S = 1 + \mathrm{Pr}A^2(\gamma - 1)/2\gamma$, $A^2 = \Omega^2a^2/2RT_o$, ρ_w is the density, at the outer cylinder wall, of the undisturbed fluid in solid body rotation, Ω is the rotation frequency, a is the cylinder radius, γ is the ratio of specific heats, Pr is the Prandtl number, μ is the bulk viscosity, R is the specific gas constant, and T_o is the temperature of the undisturbed fluid. The radial variable x measures scale heights, i.e., e-folding heights, of the ambient density and is given in terms of the physical radial variable r by $x = A^2\,[1 - (r/a)^2]$. The variable y is the axial variable nondimensionalized with respect to a. The function $F(x, y)$ is identically zero unless internal sources and sinks of mass, momentum, and energy are present. The derivation of (1) is described in detail in the papers by Wood and Morton (1980), Wood and Sanders (1983), Gunzburger and Wood (1982), and Gunzburger, Wood, and Jordan (1984). These references also give the details concerning the specific relation between the function $F(x, y)$ and the internal sources. All of the physical variables of the flow field can be obtained from the master potential and these formulas are also described in the earlier references.

The baffle is assumed to be the rectangle $x_l \le x \le x_r$, $y_{Bb} \le y \le y_{Bt}$. We assume that the baffle is very thin; in fact, in our computations, y_{Bt} differs from y_{Bb} by one grid length, i.e., the baffle is one finite element thick. The domain $\mathfrak{D}$ in which (1) holds is the interior of the centrifuge excluding the baffle, i.e.,

$$\mathfrak{D} = \{0<x<x_T,\ 0<y<y_t\}$$
$$\text{with } \{x_r \le x \le x_l,\ y_{Bb} \le y \le y_{Bt}\} \text{ removed.}$$

Therefore, we not only have to prescribe boundary conditions on the boundary of the cylinder, but also on the baffle itself.

The boundary conditions at $x = 0$, the outer wall of the centrifuge, are prescribed temperature and the no slip condition; and the boundary conditions at $x = x_t$, the rarefied inner radius position, are adiabatic and no shear. The details of how these boundary conditions are related to the master potential are given in Gunzburger, Wood, and Wayland (1987) as well as most of the earlier references by these authors.

Ekman boundary layers form on the horizontal boundaries,

[4]Department of Mathematics, Carnegie-Mellon University, Pittsburgh, Penn. 15213.

[5]Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, Va. 22901.

[6]Academic Computing Center, University of Virginia, Charlottesville, Va. 22901.

Manuscript received by the ASME Applied Mechanics Division, August 25, 1986; final revision, December 13, 1988.

stresses in the joint with elastically homogeneous medium by treating it as a single rod with abrupt temperature gradient. The obtained result is found reasonable concerning stress distributions. Also, the maximum value of tensile stress agrees fairly well with that by finite element calculation conducted by Suganuma et al. (1984). In another paper (to be published), the theory will be extended to thermal stresses in a usual joint with dissimilar elastic medium.

Acknowledgment

The authors would like to express their sincere thanks to Prof. T. Dezaki at Ashikaga Institute of Technology for his encouragement and to Dr. S. Ishioka at Tohoku University for his guidance and discussion.

References

Barton, M. V., 1941, "The Circular Cylinder With a Band of Uniform Pressure on a Finite Length of the Surface," ASME JOURNAL OF APPLIED MECHANICS, Vol. 8, pp. 97–104.

Bogy, D. B., 1968, "Edge-Bonded Dissimilar Orthogonal Wedge Under Normal and Shear Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 460–466.

Cole, Jr., S. S., and Inge, J. E., 1961, "Calculation and Measurement of Stress in a Ceramic-to-Metal Seal," *American Ceramic Society Bulletin*, Vol. 40, pp. 738–743.

Oel, H. J., and Fréchette, V. D., 1967, "Stress Distribution in Multiphase System: I, Composites With Planar Interfaces," *Journal of the American Ceramic Society*, Vol. 50, pp. 542–549.

Preist, D. H., and Talcott, R., 1959, "Thermal Stresses in Ceramic Cylinders Used in Vacuum Tubes," *American Ceramic Society Bulletin*, Vol. 38, pp. 99–105.

Rankin, A. W., 1944, "Shrink-Fit Stresses and Deformations," ASME JOURNAL OF APPLIED MECHANICS, Vol. 11, pp. 77–85.

Suganuma, K., Okamoto, T., Koizumi, M., and Shimada, M., 1984, "Effect of Interlayers in Ceramic-Metal Joints with Thermal Expansion Mismatches," *Journal of the American Ceramic Society*, Vol. 67, p. C256.

Sutton, P. M., 1958, "Stress Measurement in Circular Cylinders," *Journal of American Ceramic Society*, Vol. 41, pp. 103–109.

Timoshenko, S. P., and Goodier, J. N., 1970a, *Theory of Elasticity*, 3rd ed., McGraw-Hill, Singapore, pp. 380–383.

Timoshenko, S. P., and Goodier, J. M., 1970b, Ibid., pp. 425–428.

Whittaker, E. T., and Watson, G. N., 1935, *A Course of Modern Analysis*, 4th ed., Cambridge University Press, London, pp. 372–373.

A Study of the Effects of Baffles on Rotating Compressible Flows

Max D. Gunzburger,[4] Houston G. Wood,[5] and Rosser L. Wayland[6]

Onsager's pancake equation for the fluid dynamics of a gas centrifuge is modified for the case of centrifuges with baffles which render the flow domain doubly connected. A finite element algorithm is used for solving the mathematical model and to compute numerical examples for flow fields induced by thermal boundary conditions and by mass injection and extraction.

I Introduction

In the past decade numerous papers have appeared which address the dynamics of rotating compressible flows in a centrifuge. The primary purpose of these efforts has been directed towards modeling the diffusive processes in binary mixtures which enable the separation of species, with the chief application being separating the isotopes of uranium hexafluoride. The isotopically-altered uranium is then used as fuel for nuclear power reactors.

The purpose of the work presented in this paper is to study the effects on the flow field, due to the inclusion of baffles inside the centrifuge. The flow through the baffle holes and the effect of the baffle on the flow field have also been analyzed by Kai (1977), Lopez (1977), and Soubbaramayer (1979).

The inclusion of the baffle in the analysis of the flow field renders the domain to be doubly connected. The authors cited in the previous paragraph have used finite difference methods to solve the equations of motion in primitive variable form so that they encountered no new problems due to the multiple connectivity of the computational domain. However, if the equations of motion are formulated in terms of a stream function or a potential function as is the case for Onsager's model, care must be taken to assure that physical quantities such as the pressure are single-valued functions. This type of problem has been addressed for the incompressible case by Sood and Elrod (1974), Israeli and Ungarish (1977), Girault and Raviart (1979), and Gunzburger and Peterson (1985), and for the compressible case by Greenspan (1981) and Viecelli (1984).

II Mathematical Models

In an unpublished report, Lars Onsager showed that within a rapidly rotating centrifuge the equations of motion can be reduced to a single differential equation to be solved for a master potential function χ. This equation is

$$(e^x(e^x\chi_{xx})_{xx})_{xx}+B^2\chi_{yy}=F(x,y) \tag{1}$$

where $B^2 = \mathrm{Re}S^{1/2}/4A^6$, $Re = \rho_w\Omega a^2/\mu$, $S = 1 + \mathrm{Pr}A^2(\gamma - 1)/2\gamma$, $A^2 = \Omega^2a^2/2RT_o$, ρ_w is the density, at the outer cylinder wall, of the undisturbed fluid in solid body rotation, Ω is the rotation frequency, a is the cylinder radius, γ is the ratio of specific heats, Pr is the Prandtl number, μ is the bulk viscosity, R is the specific gas constant, and T_o is the temperature of the undisturbed fluid. The radial variable x measures scale heights, i.e., e-folding heights, of the ambient density and is given in terms of the physical radial variable r by $x = A^2\,[1 - (r/a)^2]$. The variable y is the axial variable nondimensionalized with respect to a. The function $F(x, y)$ is identically zero unless internal sources and sinks of mass, momentum, and energy are present. The derivation of (1) is described in detail in the papers by Wood and Morton (1980), Wood and Sanders (1983), Gunzburger and Wood (1982), and Gunzburger, Wood, and Jordan (1984). These references also give the details concerning the specific relation between the function $F(x, y)$ and the internal sources. All of the physical variables of the flow field can be obtained from the master potential and these formulas are also described in the earlier references.

The baffle is assumed to be the rectangle $x_l \le x \le x_r$, $y_{Bb} \le y \le y_{Bt}$. We assume that the baffle is very thin; in fact, in our computations, y_{Bt} differs from y_{Bb} by one grid length, i.e., the baffle is one finite element thick. The domain $\mathfrak{D}$ in which (1) holds is the interior of the centrifuge excluding the baffle, i.e.,

$$\mathfrak{D} = \{0<x<x_T,\ 0<y<y_t\}$$
$$\text{with } \{x_r\le x\le x_l,\ y_{Bb}\le y\le y_{Bt}\} \text{ removed.}$$

Therefore, we not only have to prescribe boundary conditions on the boundary of the cylinder, but also on the baffle itself.

The boundary conditions at $x = 0$, the outer wall of the centrifuge, are prescribed temperature and the no slip condition; and the boundary conditions at $x = x_t$, the rarefied inner radius position, are adiabatic and no shear. The details of how these boundary conditions are related to the master potential are given in Gunzburger, Wood, and Wayland (1987) as well as most of the earlier references by these authors.

Ekman boundary layers form on the horizontal boundaries,

[4]Department of Mathematics, Carnegie-Mellon University, Pittsburgh, Penn. 15213.

[5]Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, Va. 22901.

[6]Academic Computing Center, University of Virginia, Charlottesville, Va. 22901.

Manuscript received by the ASME Applied Mechanics Division, August 25, 1986; final revision, December 13, 1988.

and the boundary conditions appropriate for Onsager's pancake equation have been derived using the Carrier-Maslen matching conditions. The boundary conditions for the top and bottom end caps of the centrifuge can be found, for example, in Gunzburger, Wood, and Jordan (1984).

Similar to the top and bottom end caps of the cylinder, Ekman boundary layers form on both the top and bottom of the baffle. Thus, at these locations we impose Carrier-Maslen type conditions similar to those found in the references. Along the baffle bottom $y = y_{Bb}$ we have that

$$B^2\chi_y - 2AB^{3/2}(e^{x/2}\chi_x)_x = \frac{B^{3/2}}{2A}e^{x/2}\bar{\psi} - \frac{Re}{32A^{10}}\frac{d}{dx}\theta_{Bb}(x) \tag{2a}$$

and along the baffle top $y = y_{Bt}$ we have that

$$B^2\chi_y + 2AB^{3/2}(e^{x/2}\chi_x)_x = -\frac{B^{3/2}}{2A}e^{x/2}\bar{\psi} - \frac{Re}{32A^{10}}\frac{d}{dx}\theta_{Bt}(x) \tag{2b}$$

for $x_r \le x \le x_l$. Here, $\theta_{Bb}(x)$ and $\theta_{Bt}(x)$ denote the prescribed temperatures along the bottom and top of the baffle, respectively, and $\bar{\psi}$ is the stream function on the baffle. Since we assume that no fluid enters or leaves the centrifuge through the boundary of the baffle, $\bar{\psi}$ is a constant. However, due to the multiple connectivity of $\mathfrak{D}$ it is not possible to prescribe $\bar{\psi}$, since we have already set $\psi(x_T, 0) = 0$. Thus, $\bar{\psi}$ is unknown and must be determined as part of the solution. An auxiliary condition is needed which determines $\bar{\psi}$; this condition is given next.

Before doing so we note that no mention has been made of boundary conditions at the ends of the baffle $x = x_l$ and $x = x_r$. Some authors, e.g., Viecelli (1984), assume that the baffle has a finite axial thickness and then impose solid wall conditions at those locations. They thus ignore any interaction between the flow on the sides and the top and bottom of the baffle. We instead assume that the baffle is thin and, following Greenspan (1981), simply require the continuity of the flow variables at the ends of the baffle.

As has been noted in (2), the value of $\bar{\psi}$ may not be chosen arbitrarily; an extra condition must be imposed to ensure the single-valuedness of the pressure. Because the problem for χ and $\bar{\psi}$ is linear, the correct physical solution may be obtained in a straightforward manner. The relation between $\bar{\psi}$ and other flow variables is given by (see Gunzburger and Peterson (1985) and, especially, Gunzburger, Wood, and Wayland (1987) for detailed derivations)

$$\bar{\psi} = -\frac{1}{e^{x_r/2} - e^{x_l/2}}\left(\frac{A^2}{2}\int_{x_r}^{x_l} e^{x/2}\Big(\chi_x(x,y_{Bt}) + \chi_x(x,y_{Bb})\Big)\,dx + \frac{1}{16Re^{1/2}S^{3/4}}\int_{x_r}^{x_l}\Big(\bar{\theta}_{Bt}(x) - \bar{\theta}_{Bb}(x)\Big)\,dx\right) \tag{3}$$

This relation between the unknown stream function $\bar{\psi}$ on the baffle and the master potential χ is the necessary auxiliary condition which selects the proper physical value of $\bar{\psi}$.

In deriving (2) it has been assumed that all distributed internal sources vanish in the neighborhood of the baffle. Also note that we make a distinction between $-2A^2\chi_x(x, y_{Bb}) = \psi(x, y_{Bb})$ and $\bar{\psi}$ and between $\phi(x, y_{Bb})$ and $\bar{\theta}_{Bb}(x) = \bar{\phi}_{Bb}(x)$, and analogously along the top of the baffle. This, of course, is because the barred quantities refer to values at the baffle while the unbarred quantities refer to values at the outer edge of the Ekman layers along the baffle. The latter we do not explicitly resolve; indeed, the Carrier-Maslen conditions enable the avoidance of such resolution.

It will be convenient when discussing the solution algorithm for a doubly-connected domain to consider a more symbolic representation of our governing system. To this end we write $L(\chi) = \mathfrak{F}$ for the differential equation (1), $B_1(\chi) = \mathfrak{G}$ for boundary conditions not involving $\bar{\psi}$; $B_2(\chi) = C(\bar{\psi}) + \mathfrak{K}$ for equation (2) which represents boundary conditions involving $\bar{\psi}$; $D(\chi) = \bar{\psi}$ for equation (3) which is the auxiliary continuity condition. We note that all of the operators L, B_1. B_2, C, and D are linear and that $\mathfrak{F}$, $\mathfrak{G}$, and $\mathfrak{K}$ represent inhomogeneous data, e.g., $\mathfrak{F} = F(x, y)$.

III Solution Strategy for a Doubly-Connected Domain

The fact that $\bar{\psi}$ is unknown makes it cumbersome to discretize directly the previously described problem, although this certainly can be accomplished (e.g., see Girault and Raviart, 1979; Gunzburger and Peterson, 1985). Here we describe an alternate solution strategy described in the papers by Gunzburger and Peterson (1985), Viecelli (1984), and Gunzburger, Wood, and Wayland (1987). We divide the problem into a series of two problems as follows. First choose $k \neq 0$ and solve for χ_0 where

$$L(\chi_0) = 0, \text{ with } B_1(\chi_0) = 0 \text{ and } B_2(\chi_0) = C(k). \tag{4}$$

Second, solve for χ_1 where

$$L(\chi_1) = \mathfrak{F}, \text{ with } B_1(\chi_1) = \mathfrak{G} \text{ and } B_2(\chi_1) = \mathfrak{K}. \tag{5}$$

Next, set $k_0 = D(\chi_0)$ and $k_1 = D(\chi_1)$. It can now be easily shown that the actual solution is $\chi = \chi_1 + \alpha\chi_0$ where $\alpha = k_1/(k - k_0)$.

We note that (4) and (5) define, for any $k \neq 0$, well-posed boundary value problems. Furthermore, (4) and (5) have the same left-hand sides so that we may solve for χ_0 and χ_1 with a single linear system solve. Also, the aforementioned strategy depends, in a crucial way, on the linearity of our problem; for nonlinear problems an iterative procedure in which one successively guesses new values of k in (4) can be used. (See, e.g., Gunzburger and Peterson (1985).) Finally, we have only treated doubly-connected domains; the extension of the aforementioned strategy to general multiply-connected domains is straightforward. (Again, see Gunzburger and Peterson (1985).)

IV Examples

We have chosen to present numerical examples of flows for centrifuges with baffles using the same parameters as in our previous work. This will allow the interested reader to make direct comparisons and thereby ascertain the effects of the geometric modifications. The numerical results were obtained using a finite element method. For details, see Gunzburger, Wood, and Wayland (1987). The centrifuge parameters used are: length 335.3 cm; radius 9.145 cm; wall pressure 13.3 kPa; average temperature 300 K; peripheral speed 700 m/s. These parameters correspond to values of the dimensionless parameters $\mathrm{Re} = 1.92 \times 10^6$ and $A = 5.88$.

The baffle is usually located near an end of the centrifuge in order to shield the gas removal scoop. However, in our example, we have located the baffle at the axial midplane.This permits the features of the flow field to be examined without the necessity of refining the grid to a level required by a configuration wherein the baffle is located near an end cap. Therefore, the computational effort is kept at an affordable level. In the radial direction, the baffle extends from $x_r = 1$ to $x_l = 4$.

Figure 1 shows the streamlines for the case of a linear temperature profile along the outer wall with constant temperature on each end equal to the corresponding outer wall temperature. The streamlines clearly show the effects of the baffle; and, for this arrangement, the flow goes over the inner boundary of the baffle. The approximate width of the stream near the wall

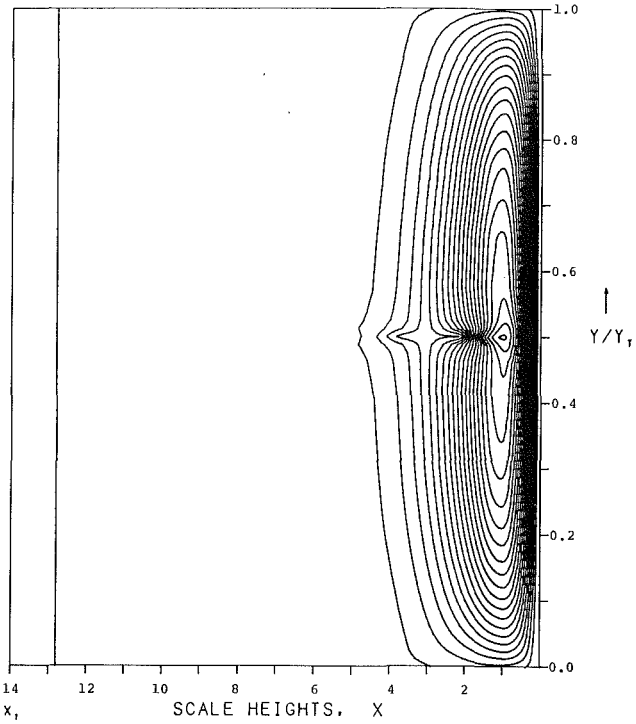


Fig. 1 Streamlines for linear wall temperature on a centrifuge with a baffle at the midplane. Uniform mass, 5 percent, is contained between each streamline.

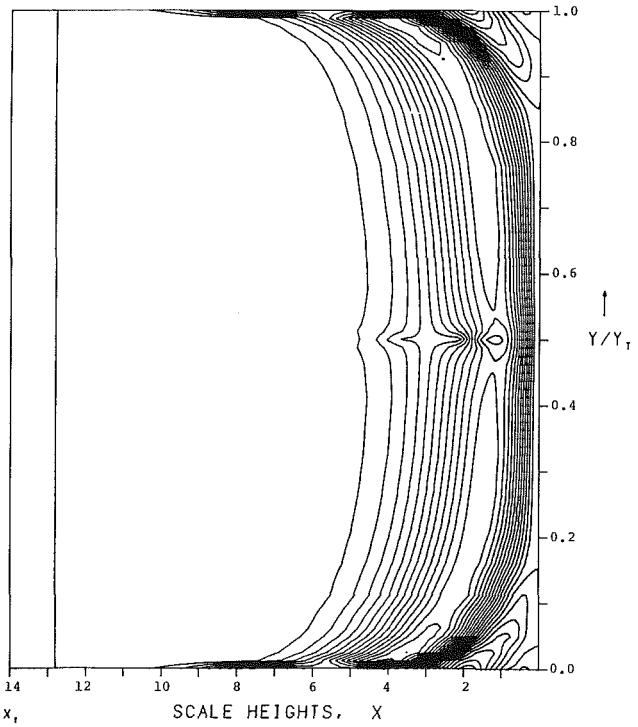


Fig. 2 Streamlines for mass throughput in a centrifuge with a midplane baffle. Uniform mass, 5 percent, is contained between each streamline.

is one scale height so the baffle does little to constrict the flow. Other calculations with the baffle located closer to the wall show that the stream can be essentially stopped. Also, for a baffle extending inward beyond $x_r = 4$, the flow cannot go around the inner edge of the baffle. Figure 2 shows the streamlines for the flow induced by a four port mass drive. Mass is introduced at $y = 0$ between $x = 7$ and $x = 8$ and at $y = y_t$ between $x = 1$ and $x = 2$; mass is removed at $y = 0$ between $x = 1$ and $x = 2$ and at $y = y_t$ between $x = 7$ and $x = 8$. This calculation, without the baffle, was presented in Gunzburger and Wood (1982). Again, because the stream near the wall is about one scale height wide, the baffle does little to impede the flow. Also, the radial extent of the baffle is small enough so that the flow can go around the inner end of the baffle. Other calculations were performed in which the temperature of the baffle was assumed to be different than that of the surrounding gas. It was found that these conditions drive very weak flows that decay almost totally before reaching the ends of the cylinder.

References

Girault, V., and Raviart, P.-A., 1979, *Finite Element Approximation of the Navier Stokes Equations*, Springer-Verlag, Berlin.

Greenspan, H., 1981, "Adaptation of Pancake Theory to Confurations with Slotted Barriers," *Report No. UVA-ER-670-81U*, University of Virginia, Charlottesville.

Gunzburger, M., and Wood, H., 1982, "A Finite Element Method for the Onsager Pancake Equation," *Comp. Meth. Appl. Mech. Engng.*, Vol. 31, pp. 43–59.

Gunzburger, M., Wood, H., and Jordan, J., 1984, "A Finite Element Method for Gas Centrifuge Problems," *SIAM J. Sci. Stat. Comp.*, Vol. 5, pp. 78–94.

Gunzburger, M., and Peterson, J., 1985, "Finite Element Methods for the Streamfunction-vorticity Equations: Boundary Condition Treatments and Multiply Connected Domains," Technical Report ICMA-85-89, University of Pittsburgh, Pittsburgh.

Gunzburger, M., Wood, H., and Wayland, R., 1987, "Analysis of Geometric Effects on Rotating Compressible Flows: Part II," *Proceedings of the Workshop on Separation Phenomena in Liquids and Gases*, K. G. Roesner, ed., Darmstadt Technische Hochschule, West Germany.

Israeli, M., and Ungarish, M., 1977, "Approximate Computation of Nonlinear, Rotating, Axisymmetric Flows," *Stromungsmechanik and Stromungsmaschinene*, Verlag G. Braun, Karlsruhe.

Kai, T., 1977, "Basic Characteristics of Centrifuges, IV," *J. Nuc. Sci. Tech*, Vol. 14, pp. 506–518.

Lopez, S., 1977, "Comparison Between the Flow in a Mechanically Driven Centrifuge with a Baffle and that in an Externally Driven Centrifuge," *Proc. of the 2nd Workshop on Gases in Strong Rotation*, France.

Sood, D., and Elrod, H., 1974, "Numerical Solution of the Incompressible Navier-Stokes Equations in Doubly-Connected Regions," *AIAA J.*, Vol. 12, pp. 636–641.

Soubbaramayer, 1979, "Centrifugation," *Uranium Enrichment*, Springer-Verlag, Berlin.

Viecelli, J., 1984, "Solution of the Onsager Equation in Doubly-Connected Regions," *J. Comput. Phys.*, Vol. 56, pp. 530–536.

Wood, H., and Morton, J., 1980, "Onsager's Pancake Approximation for the Fluid Dynamics of a Gas Centrifuge," *J. Fluid Mech.*, Vol. 101, pp. 1–31.

Wood, H., and Sanders, G., 1983, "Rotating Compressible Flows with Internal Sources and Sinks," *J. Fluid Mech.*, Vol. 127, pp. 200–313.

The Effect of Wall Compliance on the Behavior of a Confined Elastic Ring

W. J. Bottega[7]

1 Introduction

The problem of a ring contained within a cavity or cylindrical structure is germane to a variety of applications, particularly to those concerned with protective linings for insulation or environmental purposes. Most studies to date have considered the wall of the cavity as "rigid" and neglected the compliance of the substrate (see, for example, Hsu, Elkon, and Pian, 1964; McMinn and Chan, 1966; Bucciarelli and Pian, 1967; Herrmann and Zagustin, 1967; Chicurel, 1968; McGhie and Brush, 1971; El-Bayoumy, 1972; Lardner, 1980; Kyriakides and Youn, 1984; Bottega, 1988).

In the present work we explore the effect of nonvanishing compliance of the cavity wall in the context of an elastic ring contained within a circular cavity and subject to a radially-directed point load, with the wall of the cavity modeled as an elastic foundation. As for the corresponding problem involv-

[7]Department of Mechanics and Materials Science, College of Engineering, Rutgers University, Piscataway, N.J. 08855-0909. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, October 10, 1988; final revision, January 28, 1989.

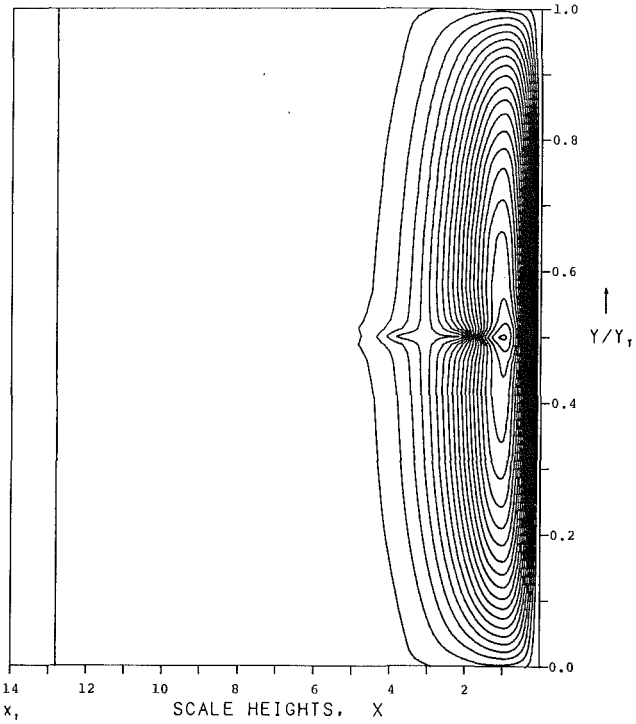


Fig. 1 Streamlines for linear wall temperature on a centrifuge with a baffle at the midplane. Uniform mass, 5 percent, is contained between each streamline.

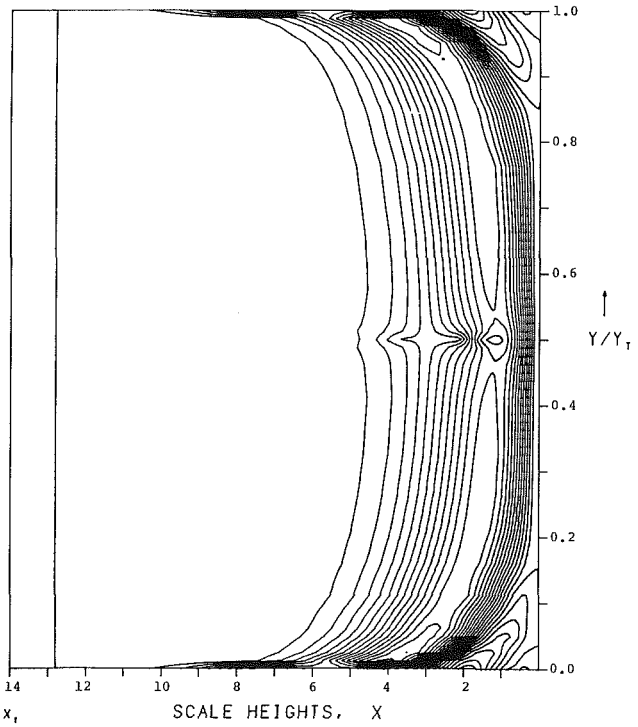


Fig. 2 Streamlines for mass throughput in a centrifuge with a midplane baffle. Uniform mass, 5 percent, is contained between each streamline.

is one scale height so the baffle does little to constrict the flow. Other calculations with the baffle located closer to the wall show that the stream can be essentially stopped. Also, for a baffle extending inward beyond $x_r = 4$, the flow cannot go around the inner edge of the baffle. Figure 2 shows the streamlines for the flow induced by a four port mass drive. Mass is introduced at $y = 0$ between $x = 7$ and $x = 8$ and at $y = y_t$ between $x = 1$ and $x = 2$; mass is removed at $y = 0$ between $x = 1$ and $x = 2$ and at $y = y_t$ between $x = 7$ and $x = 8$. This calculation, without the baffle, was presented in Gunzburger and Wood (1982). Again, because the stream near the wall is about one scale height wide, the baffle does little to impede the flow. Also, the radial extent of the baffle is small enough so that the flow can go around the inner end of the baffle. Other calculations were performed in which the temperature of the baffle was assumed to be different than that of the surrounding gas. It was found that these conditions drive very weak flows that decay almost totally before reaching the ends of the cylinder.

References

Girault, V., and Raviart, P.-A., 1979, *Finite Element Approximation of the Navier Stokes Equations*, Springer-Verlag, Berlin.

Greenspan, H., 1981, "Adaptation of Pancake Theory to Confurations with Slotted Barriers," *Report No. UVA-ER-670-81U*, University of Virginia, Charlottesville.

Gunzburger, M., and Wood, H., 1982, "A Finite Element Method for the Onsager Pancake Equation," *Comp. Meth. Appl. Mech. Engng.*, Vol. 31, pp. 43–59.

Gunzburger, M., Wood, H., and Jordan, J., 1984, "A Finite Element Method for Gas Centrifuge Problems," *SIAM J. Sci. Stat. Comp.*, Vol. 5, pp. 78–94.

Gunzburger, M., and Peterson, J., 1985, "Finite Element Methods for the Streamfunction-vorticity Equations: Boundary Condition Treatments and Multiply Connected Domains," Technical Report ICMA-85-89, University of Pittsburgh, Pittsburgh.

Gunzburger, M., Wood, H., and Wayland, R., 1987, "Analysis of Geometric Effects on Rotating Compressible Flows: Part II," *Proceedings of the Workshop on Separation Phenomena in Liquids and Gases*, K. G. Roesner, ed., Darmstadt Technische Hochschule, West Germany.

Israeli, M., and Ungarish, M., 1977, "Approximate Computation of Nonlinear, Rotating, Axisymmetric Flows," *Stromungsmechanik and Stromungsmaschinene*, Verlag G. Braun, Karlsruhe.

Kai, T., 1977, "Basic Characteristics of Centrifuges, IV," *J. Nuc. Sci. Tech*, Vol. 14, pp. 506–518.

Lopez, S., 1977, "Comparison Between the Flow in a Mechanically Driven Centrifuge with a Baffle and that in an Externally Driven Centrifuge," *Proc. of the 2nd Workshop on Gases in Strong Rotation*, France.

Sood, D., and Elrod, H., 1974, "Numerical Solution of the Incompressible Navier-Stokes Equations in Doubly-Connected Regions," *AIAA J.*, Vol. 12, pp. 636–641.

Soubbaramayer, 1979, "Centrifugation," *Uranium Enrichment*, Springer-Verlag, Berlin.

Viecelli, J., 1984, "Solution of the Onsager Equation in Doubly-Connected Regions," *J. Comput. Phys.*, Vol. 56, pp. 530–536.

Wood, H., and Morton, J., 1980, "Onsager's Pancake Approximation for the Fluid Dynamics of a Gas Centrifuge," *J. Fluid Mech.*, Vol. 101, pp. 1–31.

Wood, H., and Sanders, G., 1983, "Rotating Compressible Flows with Internal Sources and Sinks," *J. Fluid Mech.*, Vol. 127, pp. 200–313.

The Effect of Wall Compliance on the Behavior of a Confined Elastic Ring

W. J. Bottega[7]

1 Introduction

The problem of a ring contained within a cavity or cylindrical structure is germane to a variety of applications, particularly to those concerned with protective linings for insulation or environmental purposes. Most studies to date have considered the wall of the cavity as "rigid" and neglected the compliance of the substrate (see, for example, Hsu, Elkon, and Pian, 1964; McMinn and Chan, 1966; Bucciarelli and Pian, 1967; Herrmann and Zagustin, 1967; Chicurel, 1968; McGhie and Brush, 1971; El-Bayoumy, 1972; Lardner, 1980; Kyriakides and Youn, 1984; Bottega, 1988).

In the present work we explore the effect of nonvanishing compliance of the cavity wall in the context of an elastic ring contained within a circular cavity and subject to a radially-directed point load, with the wall of the cavity modeled as an elastic foundation. As for the corresponding problem involv-

[7]Department of Mechanics and Materials Science, College of Engineering, Rutgers University, Piscataway, N.J. 08855-0909. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, October 10, 1988; final revision, January 28, 1989.

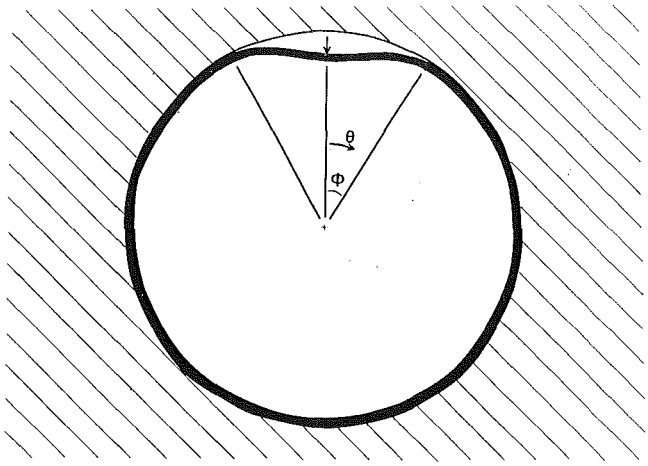


Fig. 1 Confined elastic ring subjected to a point load

ing a rigid substrate (Bottega, 1988), the present problem is approached as a moving intermediate boundary value problem in the calculus of variations with the boundary separating the portion of the ring lifted away from the cavity wall from its counterpart in contact with the "substrate" being sought as part of the solution. As in the aforementioned reference, the arch equations of El-Bayoumy (1972) are employed as the mathematical model for the ring, and the theorem of stationary potential energy is applied with the resulting transversality condition yielding the condition which defines the intermediate boundary. A closed-form analytical solution is obtained and numerical solutions are performed for several cases, in order to assess the influence of compliance of the cavity wall on the behavior of the confined ring.

2 Problem Formulation and Analytical Solution

Consider a thin elastic ring confined within the smooth wall of a cavity contained within a compliant substrate, such that the radius of the undeformed cavity coincides with the radius of the undeformed ring, and the ring is subjected to a radially-directed point load (Fig. 1). Let $Q_0 = \bar{Q}_0 R^2/D$ correspond to the normalized counterpart of the magnitude of the applied load $\bar{Q}_0$, where R and D represent the radius of the undeformed cavity and the bending stiffness of the ring, respectively. As in (Bottega, 1988), the ring shall be modeled using an arch-type theory (El-Bayoumy, 1972), and thus each point on the ring centerline may be defined by its corresponding angular coordinate in the undeformed configuration. Paralleling the aforementioned reference, the ring may be partitioned into two regions corresponding to a "lift zone" $\mathfrak{D}_1$ defined on $[0, \phi]$ and a "contact zone" $\mathfrak{D}_2$ defined on $[\phi, \pi]$, where the angular coordinate θ is measured clockwise from the line of action of the applied load. Only half of the ring is considered due to the symmetry of the problem. The compliance of the substrate shall be incorporated by introducing a "smooth elastic foundation" around the perifery of the cavity. In what follows, all lengths are normalized with respect to the initial radius of the cavity.

We begin by defining the strain energy of the substrate, U_s, as

$$U_s = \frac{1}{2}\int_{\mathfrak{D}_2} K w_2^2 d\theta, \tag{1}$$

where $K = E^* C/h \gg 1$ corresponds to the normalized radial stiffness of the substrate, E^* represents the ratio of the "equivalent" modulus of the substrate to the elastic modulus of the ring, $C = 12/h^2$ represents the normalized membrane stiffness of the ring, and $h \ll 1$ corresponds to the normalized thickness of the ring. The quantity $w_i(\theta)$, $i = 1, 2$ corresponds to the radial component (positive inward) of the deflection of a point on the ring's centerline.

Incorporating the functional U_s into the energy functional II defined in (Bottega, 1988), deleting the corresponding constraint functional Λ, and paralleling the development therein leads to the moving intermediate boundary value problem defined by

$$w_i^{iv} + (N_0 + 2)w_i'' + (1 + K_{(i)})w_i = -N_0, \quad \theta \in \mathfrak{D}_i \ (i = 1,2) \tag{2a,b}$$

where

$$K_{(1)} = 0, \qquad K_{(2)} = K, \tag{3a,b}$$

N_0 corresponds to the normalized (compressive) membrane force for the ring (found to be uniform on $[0, \pi]$), and superimposed primes denote differentiation with respect to θ.

The corresponding boundary and matching conditions for w_i are found to be

$$w_1'''(0) + (N_0 + 1)w_1'(0) = Q_0, \ w_1'(0) = 0, \tag{4a,b}$$

$$w_1(\phi) = w_2(\phi), \ w_1'(\phi) = w_2'(\phi), \tag{5a,b}$$

$$w_1''(\phi) + w_1(\phi) = w_2''(\phi) + w_2(\phi), \tag{5c}$$

$$w_1'''(\phi) + (N_0 + 1)w_1'(\phi) = w_2'''(\phi) + (N_0 + 1)w_2'(\phi), \tag{5d}$$

$$w_2'''(\pi) + (N_0 + 1)w_2'(\pi) = 0, \ w_2'(\pi) = 0, \tag{6a,b}$$

with the associated "integrability" and transversality conditions given, respectively, by

$$\int_0^{\phi}\left(w_1 - \frac{1}{2}w_1'^2\right)d\theta + \int_{\phi}^{\pi}\left(w_2 - \frac{1}{2}w_2'^2\right)d\theta = \pi N_0/C, \tag{7}$$

and

$$w(\phi) = 0. \tag{8}$$

The system (2) together with the conditions (4)–(8) constitute a moving intermediate boundary value problem for the deflections $w_1(\theta)$ and $w_2(\theta)$, the resultant membrane force N_0, and the interface angle ϕ. Solving the aforementioned system, while neglecting the nonlinear terms associated with the ring's deformation on $[\phi, \pi]$, we obtain the solution for $K \gg 1$ given by

$$w_1(\theta) = A_0(\alpha,\phi,Q_0,K)\cos\alpha\theta - B_0(\alpha,\phi,Q_0,K)\cos\theta/\alpha + Q_0F(\alpha,\theta) - N_0, \quad \theta \in \mathfrak{D}_1 \tag{9}$$

and

$$w_2(\theta) = \frac{N_0}{K}\left[\frac{G_2(\beta,\theta)}{G_2^{(0)}} - 1\right] + [Q_0 b_1(\alpha,\phi,K) + N_0 b_2(\alpha,\phi,K)]\frac{Z(\beta,\theta)}{Z^{(1)}}, \quad \theta \in \mathfrak{D}_2 \tag{10}$$

where

$$\alpha^2 = \frac{1}{2}[N_0 + 2 + \sqrt{N_0(N_0+4)}] > 1 \text{ or } N_0 = (\alpha^2 - 1)^2/\alpha^2, \tag{11a,b}$$

$$\beta^2 = K^{1/2}/2, \tag{12}$$

$$F(\alpha,\theta) = (\alpha^2\sin\theta/\alpha - \sin\alpha\theta)\alpha/(\alpha^4 - 1), \tag{13}$$

$$Z(\beta,\theta) = G_1(\beta,\theta) - G_2(\beta,\theta)G_1^{(0)}/G_2^{(0)}, \tag{14}$$

$$G_1(\beta,\theta) = e^{-\beta\theta}\cos\beta\theta - e^{-\beta(2\pi-\theta)}[f_1(\beta\pi)\cos\beta\theta + \sin 2\beta\pi\sin\beta\theta], \tag{15}$$

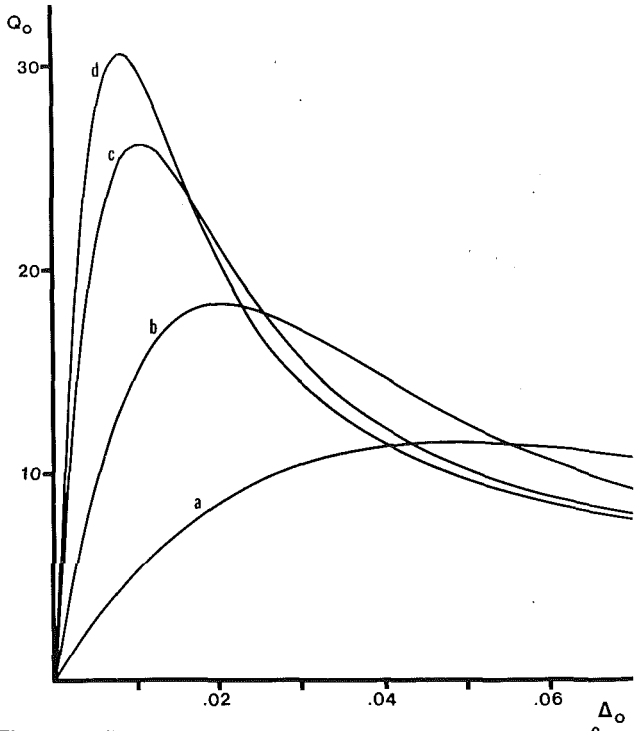


Fig. 2 Applied load versus crown point deflection; (a) $K/C = 10^{-2}$, (b) $K/C = 10^{-1}$, (c) $K/C = 1$, (d) $K/C = 10^3$, ($C = 4 \times 10^6$)

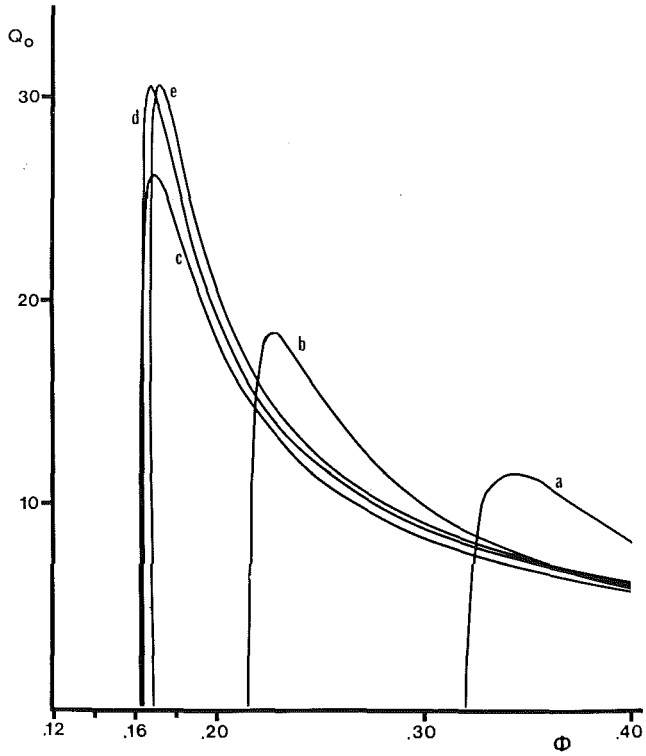


Fig. 3 Applied load versus lift zone/contact zone interface angle; (a) $K/C = 10^{-2}$, (b) $K/C = 10^{-1}$, (c) $K/C = 1$, (d) $K/C = 10^3$, (e) $K/C = 10^6$, ($C = 4 \times 10^6$)

$$G_2(\beta,\theta) = e^{-\beta\theta}\sin\beta\theta + e^{-\beta(2\pi-\theta)}[f_2(\beta\pi)\cos\beta\theta + \cos2\beta\pi\sin\beta\theta], \tag{16}$$

and

$$\mathfrak{F}^{(j)} \equiv \frac{d^{(j)}}{d\theta^{(j)}}\mathfrak{F}(\theta)\Big|_{\theta=\phi}$$

for all functions of the form $\mathfrak{F}(\theta)$.

In addition,

$$A_0(\alpha,\phi,Q_0,K) = Q_0\mu_1 + N_0\mu_2, \tag{17a}$$

$$B_0(\alpha,\phi,Q_0,K) = [Q_0F^{(0)} + (Q_0\mu_1 + N_0\mu_2)\cos\alpha\phi - N_0]/\cos\phi/\alpha, \tag{17b}$$

$$b_1(\alpha,\phi,K) = F^{(1)} + F^{(0)}\frac{\tan\phi/\alpha}{\alpha} + \mu_1 V^{(1)}, \tag{17c}$$

$$b_2(\alpha,\phi,K) = \mu_2 V^{(1)} - \frac{\tan\phi/\alpha}{\alpha} - \frac{1}{K}\frac{G_2^{(1)}}{G_2^{(0)}}, \tag{17d}$$

$$\mu_1 = \mu_1(\alpha,\phi,K) = \frac{1}{\mu^*}\left\{\left[F^{(1)} + F^{(0)}\frac{\tan\phi/\alpha}{\alpha}\right]Z^{(2)} - \frac{\sin\alpha\phi}{\alpha}Z^{(1)}\right\}, \tag{17e}$$

$$\mu_2 = \mu_2(\alpha,\phi,K) = \frac{1}{\mu^*}\left\{\left[\frac{1}{K}\frac{G_2^{(2)}}{G_2^{(0)}} + \frac{1}{\alpha^2}\right]Z^{(1)} - \left[\frac{\tan\phi/\alpha}{\alpha} + \frac{1}{K}\frac{G_2^{(1)}}{G_2^{(0)}}\right]Z^{(2)}\right\}, \tag{17f}$$

$$\mu^* = \mu^*(\alpha,\phi,K) = V^{(2)}Z^{(1)} - V^{(1)}Z^{(2)}, \tag{17g}$$

$$f_1(\beta\pi) = 2\cos^2\beta\pi + 1. \tag{17h}$$

$$f_2(\beta\pi) = \cos^3\beta\pi/\sin\beta\pi, \tag{17i}$$

and

$$V(\alpha,\theta) = \cos\alpha\theta - \frac{\cos\alpha\phi}{\cos\phi/\alpha}\cos\theta/\alpha. \tag{18}$$

The aforesaid solution had been obtained using the condition (8) in lieu of (5*d*), along with conditions (4*a*,*b*), (5*a*,*b*,*c*) and (6*a*,*b*). Substitution of (18) into the remaining conditions (5*d*) and (7) results in a pair of coupled nonlinear algebraic equations in the parameters α (or equivalently N_0), ϕ, and Q_0 which may be solved to yield values corresponding to equilibrium states of the system for given K (or equivalently β) and C. Numerical simulations corresponding to specific ring-substrate systems will be presented in the next section.

3 Numerical Results and Discussion

As discussed in Section 2, the coupled algebraic equations resulting from (5*d*) and (7) may be solved simultanenously to yield values of α, ϕ, and Q_0 which correspond to states of equilibrium of the ring-substrate system for given values of K and C. The "crown point" deflection $\Delta_0 \equiv w(0)$, associated with each equilibrium state may then be obtained by evaluating (9) at $\theta = 0$, thus generating the corresponding image of each "equilibrium path" in the space of the applied load and the crown point deflection. This is done numerically in a manner similar to that done by Bottega (1988). Results are displayed in Figs. 2 and 3, for a ring with normalized membrane stiffness $C = 4 \times 10^6$, for various values of the normalized substrate stiffness K ranging from $C \times 10^{-2}$ to $C \times 10^6$.

Upon inspection of each of the load-deflection and load-"interface angle" curves displayed in Figs. 2 and 3, we note the existence of a "critical load," Q_c, corresponding to the relative maximum of each curve, indicating the onset of "snap-thru" type buckling of the ring with the crown point deflection, Δ_0, and associated interface angle, ϕ, increasing to relatively large value in an unstable manner. We also note the existence of a minimum value of the interface angle ϕ_{min} below which no equilibrium states of the system exist. Both of these

characteristics were observed for the limiting case of a rigid substrate considered by Bottega (1988) and, as was pointed out therein, for related problems involving rigid cavity walls considered by Herrmann and Zagustin (1967), and McGhie and Brush (1971).

It may be observed that the load-deflection curves flatten and Q_c decreases with decreasing values of the substrate stiffness K, and also that the unstable portion of each curve lies above those with higher K. This would seem to indicate that although buckling of the ring occurs at a lower load for a more compliant substrate, it appears to be accompanied by a less extensive "jump" in the deflection. This may be attributed to the smaller build up of strain energy preceding buckling for the more compliant system. We note here that, within the resolution of the figure, the curve corresponding to $K/C=10^3$ in Fig. 2 is effectively colinear with the analogous curve for the case of a rigid substrate (Bottega, 1988). Further insight into the "mechanisms" involved may be obtained upon consideration of the load-interface angle curves displayed in Fig. 3.

Let us first separate the systems considered into two categories, the first such that $K \geq C$ the substrates of which shall be referred to as "stiff" substrates, and the second category such that $K<C$ the substrates of which shall be said to be "compliant." One may first observe from Fig. 3 that the curves corresponding to the stiff substrates shift to the right as K increases with the curve corresponding to the stiffest substrate considered tending toward the corresponding curve associated with a rigid substrate (see Bottega, 1988). It may be next observed that as Q_0 increases, and $K \geq C$, ϕ initially decreases slightly (from a finite value as $Q_0 \to 0^+$) to $\phi_{\min}$, then increases until $Q_0 = Q_c$ at which point buckling occurs and ϕ increases in an unstable manner. This is analogous to what was seen to occur for the case of a rigid substrate and offers the interpretation that for a stiff substrate, the ring predominantly "shrinks" at the initial application of the load such that a finite value of ϕ (as $Q_0 \to 0^+$) is achieved. This is followed by ϕ decreasing slightly with increasing Q_0 due to the initial bending as a result of the stiff constraint. Further increases in load are accompanied by increased flexure with the ring bending away from the substrate slightly (ϕ increasing) and ultimately buckling and its consequences as discussed earlier. The leftward tendency of the $K \geq C$ curves with decreasing substrate stiffness would indicate that slightly more flexure is allowed by the more compliant of the stiff substrates during the "shrink dominated phase" of the ring's deformation.

For the compliant substrates ($K<C$) we observe a shifting to the right, of the corresponding load-deflection and load-interface angle curves, as K decreases. This phenomenon offers the explanation that as the ring stiffness to substrate stiffness ratio increases, the system (initially) tends toward the behavior of a rigid ring confined by a compliant substrate, where for a rigid ring $\phi = \pi/2$ for all finite K. We also observe, for $K<C$, that $\phi \to \phi_{\min}$ as $Q_0 \to 0^+$ and that the positive slopes of the corresponding $Q_0 - \phi$ curves decrease with increasing K, indicating that the ring bends aways from the substrate with increasing Q_0, immediately following the initial shrink dominated phase and that this behavior is more pronounced as K decreases.

References

Bottega, W. J., 1988, "On the Constrained Elastic Ring," *J. Eng. Math.*, Vol. 22, No. 1, pp. 43–51.

Bucciarelli, Jr., L. L., and Pian, T. H. H., 1967, "Effect of Initial Imperfections on the Instability of a Ring Confined to an Imperfect Rigid Boundary," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 979–984.

Chicurel, R., 1968, "Shrink Buckling of Circular Rings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 608–610.

El-Bayoumy, L., 1972, "Buckling of a Circular Elastic Ring Confined to a Uniformly Contracting Boundary," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 758–766.

Hsu, P. T., Elkon, J., and Pian, T. H. H., 1964, "Note on the Instability of Circular Rings Confined to a Rigid Boundary," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 559–562.

Herrmann, G., and Zagustin, E. A., 1967, "Stability of an Elastic Ring in a Rigid Cavity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 263–270.

Kyriakides, S., and Youn, S. K., 1984, "On the Collapse of Circular Confined Rings Under External Pressure," *Int. J. Solids and Structures*, Vol. 51, pp. 699–713.

Lardner, T. J., 1980, "On the Nonbuckling of a Circular Ring Under a 'Wrapping' Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 973–974.

McGhie, R. D., and Brush, D. O., 1971, "Deformation of an Inertia-Loaded Thin Ring in a Rigid Cavity With Initial Clearance," *Int. J. Solids and Structures*, Vol. 7, pp. 1539–1553.

McMinn, S. J., and Chan, H. C., 1966, "The Stability of a Uniformly Compressed Ring Surrounded by a Rigid Circular Surface," *Int. J. Mech. Sci.*, Vol. 8, pp. 433–442.

Approximations for Steady Waves in Viscoelastic Materials

G. T. Warhola[8] and A. C. Pipkin[9]

Propagation of steady waves and shocks in viscoelastic materials has been reviewed by Nunziato et al. (1974). With appropriate interpretation of the particle velocity v, strain ϵ, and stress σ, a variety of different kinds of waves can be discussed in a common formalism. When each of these quantities is a function of the one variable $t-(x/U)$, where U is the wave speed, integration of the strain-velocity compatibility condition and the momentum equation gives (Warhola and Pipkin (1988))

$$-U\epsilon(t) = v(t), \quad -Uv(t) = \sigma(t), \tag{1}$$

where we have now set $x=0$. The variables ϵ, v, and σ are required to vanish as $t \to -\infty$. The quantity σ in (1) is now the stress divided by the constant mass density.

Because of the small range of strains and times involved in laboratory measurements of steady waves in real viscoelastic materials (Nunziato et al., 1974; Schuler, 1970; Kolsky, 1969) such observations are explained very adequately by only a slight generalization of the linear viscoelastic constitutive equation (Pipkin, 1986), in which the strain is replaced by a nonlinear function of strain,

$$f(\epsilon) = J' * \sigma = J * \sigma'. \tag{2}$$

The function $f(\epsilon)$ is asymptotic to ϵ when the strain is small. For the present discussion we suppose that $f(\epsilon)/\epsilon$ is an increasing function of ϵ, and for illustrative purposes we use the form

$$f(\epsilon) = \epsilon[1 + (\epsilon/\epsilon_c)^P] (P>0). \tag{3}$$

The asterisk in (2) denotes convolution over the interval $(-\infty, +\infty)$, and the derivatives J' and σ' are treated as generalized functions. $J(t)$ is the compliance multiplied by the mass density. It is identically zero for $t<0$, with initial value $J_0 = J(0+) \geqq 0$. For $t>0$, J is an increasing function of t, with J' decreasing. Experimentally determined compliances for some real materials are given in Ferry's book (Ferry, 1970). Such data are given graphically in plots of $\log J(t)$ versus $\log(t)$. At the short times involved in laboratory experiments (microseconds (Schuler, 1970) or milliseconds (Kolsky, 1969)), the slope p of such a log-log plot is still very small. Approximating J by its tangent on such a plot leads to the power-law

[8]Air Force Institute of Technology, Wright-Patterson A. F. B., Ohio 45433.

[9]Division of Applied Mathematics, Brown University, Providence, R. I. 02912.

Manuscript received by the ASME Applied Mechanics Division, February 9, 1988; final revision, January 12, 1989.

characteristics were observed for the limiting case of a rigid substrate considered by Bottega (1988) and, as was pointed out therein, for related problems involving rigid cavity walls considered by Herrmann and Zagustin (1967), and McGhie and Brush (1971).

It may be observed that the load-deflection curves flatten and Q_c decreases with decreasing values of the substrate stiffness K, and also that the unstable portion of each curve lies above those with higher K. This would seem to indicate that although buckling of the ring occurs at a lower load for a more compliant substrate, it appears to be accompanied by a less extensive "jump" in the deflection. This may be attributed to the smaller build up of strain energy preceding buckling for the more compliant system. We note here that, within the resolution of the figure, the curve corresponding to $K/C=10^3$ in Fig. 2 is effectively colinear with the analogous curve for the case of a rigid substrate (Bottega, 1988). Further insight into the "mechanisms" involved may be obtained upon consideration of the load-interface angle curves displayed in Fig. 3.

Let us first separate the systems considered into two categories, the first such that $K \geq C$ the substrates of which shall be referred to as "stiff" substrates, and the second category such that $K<C$ the substrates of which shall be said to be "compliant." One may first observe from Fig. 3 that the curves corresponding to the stiff substrates shift to the right as K increases with the curve corresponding to the stiffest substrate considered tending toward the corresponding curve associated with a rigid substrate (see Bottega, 1988). It may be next observed that as Q_0 increases, and $K \geq C$, ϕ initially decreases slightly (from a finite value as $Q_0 \to 0^+$) to $\phi_{\min}$, then increases until $Q_0 = Q_c$ at which point buckling occurs and ϕ increases in an unstable manner. This is analogous to what was seen to occur for the case of a rigid substrate and offers the interpretation that for a stiff substrate, the ring predominantly "shrinks" at the initial application of the load such that a finite value of ϕ (as $Q_0 \to 0^+$) is achieved. This is followed by ϕ decreasing slightly with increasing Q_0 due to the initial bending as a result of the stiff constraint. Further increases in load are accompanied by increased flexure with the ring bending away from the substrate slightly (ϕ increasing) and ultimately buckling and its consequences as discussed earlier. The leftward tendency of the $K \geq C$ curves with decreasing substrate stiffness would indicate that slightly more flexure is allowed by the more compliant of the stiff substrates during the "shrink dominated phase" of the ring's deformation.

For the compliant substrates ($K<C$) we observe a shifting to the right, of the corresponding load-deflection and load-interface angle curves, as K decreases. This phenomenon offers the explanation that as the ring stiffness to substrate stiffness ratio increases, the system (initially) tends toward the behavior of a rigid ring confined by a compliant substrate, where for a rigid ring $\phi = \pi/2$ for all finite K. We also observe, for $K<C$, that $\phi \to \phi_{\min}$ as $Q_0 \to 0^+$ and that the positive slopes of the corresponding $Q_0 - \phi$ curves decrease with increasing K, indicating that the ring bends aways from the substrate with increasing Q_0, immediately following the initial shrink dominated phase and that this behavior is more pronounced as K decreases.

References

Bottega, W. J., 1988, "On the Constrained Elastic Ring," *J. Eng. Math.*, Vol. 22, No. 1, pp. 43–51.

Bucciarelli, Jr., L. L., and Pian, T. H. H., 1967, "Effect of Initial Imperfections on the Instability of a Ring Confined to an Imperfect Rigid Boundary," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 979–984.

Chicurel, R., 1968, "Shrink Buckling of Circular Rings," ASME JOURNAL OF APPLIED MECHANICS, Vol. 35, pp. 608–610.

El-Bayoumy, L., 1972, "Buckling of a Circular Elastic Ring Confined to a Uniformly Contracting Boundary," ASME JOURNAL OF APPLIED MECHANICS, Vol. 39, pp. 758–766.

Hsu, P. T., Elkon, J., and Pian, T. H. H., 1964, "Note on the Instability of Circular Rings Confined to a Rigid Boundary," ASME JOURNAL OF APPLIED MECHANICS, Vol. 31, pp. 559–562.

Herrmann, G., and Zagustin, E. A., 1967, "Stability of an Elastic Ring in a Rigid Cavity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 263–270.

Kyriakides, S., and Youn, S. K., 1984, "On the Collapse of Circular Confined Rings Under External Pressure," *Int. J. Solids and Structures*, Vol. 51, pp. 699–713.

Lardner, T. J., 1980, "On the Nonbuckling of a Circular Ring Under a 'Wrapping' Loading," ASME JOURNAL OF APPLIED MECHANICS, Vol. 47, pp. 973–974.

McGhie, R. D., and Brush, D. O., 1971, "Deformation of an Inertia-Loaded Thin Ring in a Rigid Cavity With Initial Clearance," *Int. J. Solids and Structures*, Vol. 7, pp. 1539–1553.

McMinn, S. J., and Chan, H. C., 1966, "The Stability of a Uniformly Compressed Ring Surrounded by a Rigid Circular Surface," *Int. J. Mech. Sci.*, Vol. 8, pp. 433–442.

Approximations for Steady Waves in Viscoelastic Materials

G. T. Warhola[8] and A. C. Pipkin[9]

Propagation of steady waves and shocks in viscoelastic materials has been reviewed by Nunziato et al. (1974). With appropriate interpretation of the particle velocity v, strain ϵ, and stress σ, a variety of different kinds of waves can be discussed in a common formalism. When each of these quantities is a function of the one variable $t-(x/U)$, where U is the wave speed, integration of the strain-velocity compatibility condition and the momentum equation gives (Warhola and Pipkin (1988))

$$-U\epsilon(t) = v(t), \quad -Uv(t) = \sigma(t), \tag{1}$$

where we have now set $x=0$. The variables ϵ, v, and σ are required to vanish as $t \to -\infty$. The quantity σ in (1) is now the stress divided by the constant mass density.

Because of the small range of strains and times involved in laboratory measurements of steady waves in real viscoelastic materials (Nunziato et al., 1974; Schuler, 1970; Kolsky, 1969) such observations are explained very adequately by only a slight generalization of the linear viscoelastic constitutive equation (Pipkin, 1986), in which the strain is replaced by a nonlinear function of strain,

$$f(\epsilon) = J' * \sigma = J * \sigma'. \tag{2}$$

The function $f(\epsilon)$ is asymptotic to ϵ when the strain is small. For the present discussion we suppose that $f(\epsilon)/\epsilon$ is an increasing function of ϵ, and for illustrative purposes we use the form

$$f(\epsilon) = \epsilon[1 + (\epsilon/\epsilon_c)^P] (P>0). \tag{3}$$

The asterisk in (2) denotes convolution over the interval $(-\infty, +\infty)$, and the derivatives J' and σ' are treated as generalized functions. $J(t)$ is the compliance multiplied by the mass density. It is identically zero for $t<0$, with initial value $J_0 = J(0+) \geqq 0$. For $t>0$, J is an increasing function of t, with J' decreasing. Experimentally determined compliances for some real materials are given in Ferry's book (Ferry, 1970). Such data are given graphically in plots of $\log J(t)$ versus $\log(t)$. At the short times involved in laboratory experiments (microseconds (Schuler, 1970) or milliseconds (Kolsky, 1969)), the slope p of such a log-log plot is still very small. Approximating J by its tangent on such a plot leads to the power-law

[8]Air Force Institute of Technology, Wright-Patterson A. F. B., Ohio 45433.

[9]Division of Applied Mathematics, Brown University, Providence, R. I. 02912.

Manuscript received by the ASME Applied Mechanics Division, February 9, 1988; final revision, January 12, 1989.

form Ct^p (Pipkin, 1986). For illustrative purposes we use the more general form

$$J(t) = J_0 + J_1(t/t_r)^p \quad (0 \le p \le 1). \tag{4}$$

Combining (1) and (2) gives an integral equation for $\epsilon(t)$,

$$f(\epsilon) = U^2 J' * \epsilon. \tag{5}$$

Note from (1) that $v(t)$, which is more likely to be the observed quantity, is proportional to $\epsilon(t)$. Warhola (1988) has devised a numerical algorithm for the solution of (5), which is capable of being used with functions $f(\epsilon)$ and $J(t)$ given in the form of data. Existence, uniqueness, and monotonicity of solutions are discussed elsewhere (Warhola and Pipkin, 1988, Warhola, 1988). Here we assess the accuracy of two kinds of approximate solutions of (5) by using the forms (3) and (4).

When $U > U_0$, where U_0 is the instantaneous wave speed defined by $U_0^2 J_0 = 1$, the strain is zero until some time $t = 0$, say, at which it jumps discontinuously to a value ϵ_0. For $t > 0$, $\epsilon(t)$ increases continuously (Warhola and Pipkin, 1988). Let $\epsilon_q(t)$ be the intersection of the Rayleigh line $\sigma = U^2\epsilon$ with the *stress-strain isochrone* $f(\epsilon) = J(t)\sigma$. This *quasi-elastic* approximation is given by

$$f(\epsilon_q)/\epsilon_q = U^2 J(t). \tag{6}$$

The quasi-elastic approximation furnishes a rigorous upper bound on the exact solution, and in cases of the type that are of main physical interest, the error in $\epsilon_q(t)$ is very small (Warhola and Pipkin, 1988). When $f(\epsilon)$ has the form (3), (6) gives

$$\epsilon_q(t) = \epsilon_c[U^2 J(t) - 1]^{1/P}. \tag{7}$$

When $P = 1$ (a quadratic nonlinearity in f,) the wave form has the same shape as $J(t)$. Schuler's (1970) observed wave forms are like those given by using (7) with a power-law compliance (4) with small p (Warhola and Pipkin, 1988).

The basis of the quasi-elastic approximation is that with $J(T)$ defined by

$$J' * \epsilon = J[T(t)]\epsilon(t), \tag{8}$$

$T(t)$ must be approximately equal to t when the strain history is nearly a simple step at time zero. A closer approximation $\epsilon_a(t)$ can be obtained by using a more refined estimate of T. Since $J(T)$ is the average value of $J(s)$ with respect to the weight function $\epsilon'(t-s)/\epsilon(t)$, we estimate T as the average value of s itself with respect to the same weight function. With an integration by parts, this leads to the prescription

$$\epsilon_a(t)T(t) = \int_{-\infty}^{t} \epsilon_a(s)ds \tag{9}$$

or

$$dt = dT + Td\epsilon_a/\epsilon_a, \tag{10}$$

with

$$f(\epsilon_a)/\epsilon_a = U^2 J(T). \tag{11}$$

Integration of (10) with appropriate initial conditions then gives t as a function of ϵ_a. This procedure gives the exact solution if J has the form (4) with $p = 0$ or $p = 1$. For real materials with J concave, ϵ_a lies between ϵ_q and the exact solution ϵ if $U > U_0$.

To test the accuracy of ϵ_q and ϵ_a, in some cases in which the exact solution is known, we use the forms (3) and (4) in (5) to obtain

$$\epsilon(1 - U^2 J_0) + \epsilon(\epsilon/\epsilon_c)^P = U^2 J_1 (t/t_r)^p * \epsilon'. \tag{12}$$

For *critical* waves (Warhola and Pipkin, 1988) with $U^2 J_0 = 1$, the strain is zero for $t \lesseqgtr 0$ (say). For $t \gtrsim 0$, the quasi-elastic approximation is

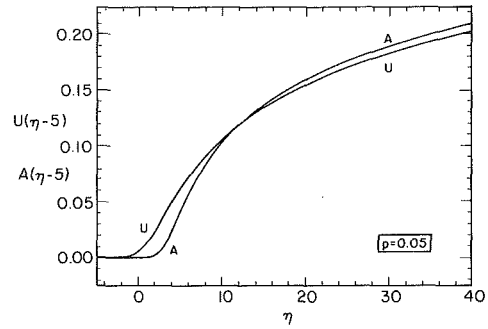


Fig. 1 Subcritical wave in a power-law material with $p = 0.05$; *u*: exact, *A*: approximate

$$\epsilon_q(t) = \epsilon_c(U^2 J_1)^{1/P}(t/t_r)^q, \; q = p/P. \tag{13}$$

Both $\epsilon(t)$ and $\epsilon_a(t)$ are proportional to $\epsilon_q(t)$:

$$\epsilon(t) = K^{1/P}\epsilon_q(t) = K^{1/P}(1+q)^q \epsilon_a(t), \tag{14}$$

where

$$K = p!q!/(p+q)!. \tag{15}$$

For $P = 1$, the factor K is unity at $p = 0$ but $1/2$ at $p = 1$, so $\epsilon_q(t)$ is accurate only at small p. In contrast, the ratio ϵ/ϵ_a is unity at both $p = 0$ and $p = 1$, and it is never less than about 0.96 for p between these values.

For subcritical ($U^2 J_0 < 1$) or supercritical ($U^2 J_0 > 1$) waves, we use the scaling

$$\epsilon = \epsilon_c |1 - U^2 J_0|^{1/P} u(\eta), \; \eta = rt, \tag{16}$$

where r is defined by

$$(rt_r)^p = U^2 J_1 p!/|1 - U^2 J_0|. \tag{17}$$

This brings (12) into the form

$$\pm u + u^{P+1} = Q * u', \tag{18}$$

with the upper sign for subcritical waves and the lower sign for supercritical waves. $Q(\eta)$ is defined as $\eta^p/p!$. To test the approximation $\epsilon_a(t)$ in a subcritical case, let $A(\eta)$ be the correspondingly scaled version of $\epsilon_a(t)$, and let τ be the scaled version of T. Then, in place of (10) and (11) we have

$$d\eta = d\tau + \tau dA/A, \; A^P = Q(\tau) - 1. \tag{19}$$

The integration can be carried out in finite form if $p = 1/N$ with N an integer. To lighten the notation, let us take $P = 1$. Then integration gives

$$\eta/(p!)^N = (A+1)^N + \sum_{n=1}^{N} \binom{N}{n} \frac{1}{n} A^n + lnA. \tag{20}$$

This is the exact solution when $N = 1 (p = 1)$. For $N = 20 (p = 0.05)$, the exact and approximate solutions are compared in Fig. 1. The exact solution was obtained by numerical computation (Warhola, 1988).

Acknowledgment

We are grateful for the support from NSF grant DMS-8702866 and from the U.S. Air Force.

References

Ferry, J. D., 1970, *Viscoelastic Properties of Polymers*, John Wiley and Sons, New York.

Kolsky, H., 1969, "Production of Tensile Shock Waves in Stretched Natural Rubber," *Nature*, Vol. 224, p. 1301.

Nunziato, J. W., Walsh, E. K., Schuler, K. W., and Barker, L. M., 1974, "Wave Propagation in Nonlinear Viscoelastic Solids," *Enc. of Physics VIa/4 Mechanics of Solids*, Springer-Verlag, New York.

Pipkin, A. C., 1986, *Lectures on Viscoelasticity Theory*, Springer-Verlag, New York.
Schuler, K. W., 1970, "Propagation of Steady Shock Waves in Polymethyl Methacrylate," *J. Mech. Phys. Solids*, Vol. 18, pp. 277-293.
Warhola, G. T., 1988, "Steady Waves in a Nonlinear Theory of Viscoelasticity," Ph.D. dissertation, Brown University.
Warhola, G. T., and Pipkin, A. C., 1988, "Shock Structure in Viscoelastic Materials," *IMA Journal of Applied Mathematics*, Vol. 41, pp. 47-66.

Interference of a Uniform Open Ring With a Rigid Cylinder

W. W. King[10]

Introduction

A common connector for fiber-optic communciation systems has the ends of fibers embedded in otherwise solid cylindrical plugs. The plugs, and hence the fibers, are aligned by inserting then into a split sleeve which expands to accommodate the plugs. The mechanical designer would like to know, for various combinations of diameters of plugs and inside diameter of sleeve, the extent of contact and the distribution of contact pressure between plugs and sleeve, as well as the stresses in the sleeve. This problem is formidable for other than numerical analysis. However, some guidance may be expected from solution of the plane-strain problem of longitudinally independent interference of a plug and sleeve or from the related strength-of-materials problem of interference between a rigid plug and a thin-wall open ring.

The desire to maximize plug-sleeve contact has motivated J. M. Anderson of AT&T Bell Laboratories to design an alignment sleeve of circumferentially varying thickness (Hogan, 1985). This is suggested by the piston-ring problem discussed by Timoshenko (1955) in which an appropriate nonuniformity of thickness yields circle-to-circle deformation of a thin ring under the action of uniform pressure; that is, complete contact is achieved between the ring and interfering rigid cylinder, and the interface pressure is uniform. More commonly, alignment sleeves of uniform thickness are employed, but, surprisingly, the writer has been unable to find in the literature an analytical solution to the corresponding ring-plug problem. Thus, the purpose of this note is communication of a solution to the problem of interference between a rigid cylinder and an elastic, thin, open ring of uniform thickness.

Analysis

Assuming frictionless contact between rigid cylinder and thin elastic ring, only radial loading on the ring need be considered. For this case the governing equations are taken to be effectively those of Fettahlioglu and Mayers (1977).

$$N - \frac{dQ}{d\theta} = Rp \tag{1}$$

$$\frac{dN}{d\theta} + Q = 0 \tag{2}$$

$$\frac{dM}{d\theta} + RQ = 0 \tag{3}$$

$$N = \frac{EA}{R}\left(u + \frac{dv}{d\theta}\right) \tag{4}$$

$$M = \frac{EI}{R^2}\left(\frac{d^2u}{d\theta^2} - \frac{dv}{d\theta}\right). \tag{5}$$

The first three of these express balances of forces and of moments. (See Fig. 2 for definitions of the stress resultants.) Equations (4) and (5) are constitutive relations in which u and v are the radial and transverse components of displacement. E, A, I, and R are Young's modulus, cross-sectional area, area moment of inertia, and radius, respectively.

If the ring is taken to be of rectangular cross-section and of unit length and if E is replaced by $E/(1-\nu^2)$, ν being Poisson's ratio, then equations (1)-(5) govern the behavior of a thin cylindrical shell in plane strain (see Timoshenko and Woinowsky-Krieger (1959) or Kraus (1967)).

Elimination of N, Q, and M from equations (1)-(5) yields the pair of differential equations in u and v deduced by Fettahlioglu and Mayers (1977) by use of Hamilton's Principle. It would be entirely appropriate to follow their lead and generate a general solution to these equations in each of the regions of contact and noncontact and ultimately proceed to a solution of the present problem. However, some special features of the problem allow a less formal approach to be more efficient.

Figure 1 depicts the ring with the opening angle 2β and the interfering cylinder. Because of symmetry, only the range $0<\theta<\pi-\beta$ need be considered. Continuous contact is impossible in the neighborhood of an end, $\theta=0$, of the ring; this is because continuous contact must be associated with a bending moment which is constant, as will be seen shortly. The unknown angle ϕ identifies the region over which the ring and cylinder are not in contact. We assume continuous contact for $\theta>\phi$.

The free-body diagram (Fig. 2), which shows the concentrated contact force, F, at $\theta=0$, allows us to conclude from a balance of moments that $M=RN$ everywhere. The same result could equivalently be deduced by eliminating Q from equations (2) and (3), integrating, and then using the fact that $N=M=0$ at $\theta=0$. Combining the constitutive relations with this result,

$$\frac{1}{R^2}\frac{d^2u}{d\theta^2} - \frac{1}{R^2}\frac{dv}{d\theta} = \frac{M}{EI}$$

$$\frac{1}{R^2}\frac{d^2u}{d\theta^2} - \frac{1}{R^2}\left(\frac{RN}{EA} - u\right) = \frac{M}{EI}$$

or

$$\frac{d^2u}{d\theta^2} + u = \frac{MR^2}{EI}\left(1 + \frac{I}{AR^2}\right). \tag{6}$$

Now let δ be the radial interference, i.e., the difference between the radius of the rigid cylinder and the (inner) radius of the undeformed ring. Then in the region of continuous contact, $\theta>\phi$, we specify $u=\delta$ and the bending moment is then, by equation (6), the constant

$$M_o = \left(\frac{EI\delta}{R^2}\right)\Big/\left(1 + \frac{I}{AR^2}\right). \tag{7}$$

From Fig. 2 we can see that

$$M_o = FR\sin\phi. \tag{8}$$

In the region of noncontact, $\theta<\phi$,

$$M = FR\sin\theta \tag{9}$$

or

$$M = \frac{M_o\sin\theta}{\sin\phi}.$$

Therefore, in this region equation (6) becomes

$$\frac{d^2u}{d\theta^2} + u = \frac{\delta\sin\theta}{\sin\phi}. \tag{10}$$

[10]Professor, Georgia Institute of Technology, Atlanta, Ga. 30332-0355. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, July 19, 1988; final revision, December 8, 1988.

Pipkin, A. C., 1986, *Lectures on Viscoelasticity Theory*, Springer-Verlag, New York.
Schuler, K. W., 1970, "Propagation of Steady Shock Waves in Polymethyl Methacrylate," *J. Mech. Phys. Solids*, Vol. 18, pp. 277-293.
Warhola, G. T., 1988, "Steady Waves in a Nonlinear Theory of Viscoelasticity," Ph.D. dissertation, Brown University.
Warhola, G. T., and Pipkin, A. C., 1988, "Shock Structure in Viscoelastic Materials," *IMA Journal of Applied Mathematics*, Vol. 41, pp. 47-66.

Interference of a Uniform Open Ring With a Rigid Cylinder

W. W. King[10]

Introduction

A common connector for fiber-optic communciation systems has the ends of fibers embedded in otherwise solid cylindrical plugs. The plugs, and hence the fibers, are aligned by inserting then into a split sleeve which expands to accommodate the plugs. The mechanical designer would like to know, for various combinations of diameters of plugs and inside diameter of sleeve, the extent of contact and the distribution of contact pressure between plugs and sleeve, as well as the stresses in the sleeve. This problem is formidable for other than numerical analysis. However, some guidance may be expected from solution of the plane-strain problem of longitudinally independent interference of a plug and sleeve or from the related strength-of-materials problem of interference between a rigid plug and a thin-wall open ring.

The desire to maximize plug-sleeve contact has motivated J. M. Anderson of AT&T Bell Laboratories to design an alignment sleeve of circumferentially varying thickness (Hogan, 1985). This is suggested by the piston-ring problem discussed by Timoshenko (1955) in which an appropriate nonuniformity of thickness yields circle-to-circle deformation of a thin ring under the action of uniform pressure; that is, complete contact is achieved between the ring and interfering rigid cylinder, and the interface pressure is uniform. More commonly, alignment sleeves of uniform thickness are employed, but, surprisingly, the writer has been unable to find in the literature an analytical solution to the corresponding ring-plug problem. Thus, the purpose of this note is communication of a solution to the problem of interference between a rigid cylinder and an elastic, thin, open ring of uniform thickness.

Analysis

Assuming frictionless contact between rigid cylinder and thin elastic ring, only radial loading on the ring need be considered. For this case the governing equations are taken to be effectively those of Fettahlioglu and Mayers (1977).

$$N - \frac{dQ}{d\theta} = Rp \tag{1}$$

$$\frac{dN}{d\theta} + Q = 0 \tag{2}$$

$$\frac{dM}{d\theta} + RQ = 0 \tag{3}$$

$$N = \frac{EA}{R}\left(u + \frac{dv}{d\theta}\right) \tag{4}$$

$$M = \frac{EI}{R^2}\left(\frac{d^2u}{d\theta^2} - \frac{dv}{d\theta}\right). \tag{5}$$

The first three of these express balances of forces and of moments. (See Fig. 2 for definitions of the stress resultants.) Equations (4) and (5) are constitutive relations in which u and v are the radial and transverse components of displacement. E, A, I, and R are Young's modulus, cross-sectional area, area moment of inertia, and radius, respectively.

If the ring is taken to be of rectangular cross-section and of unit length and if E is replaced by $E/(1-\nu^2)$, ν being Poisson's ratio, then equations (1)-(5) govern the behavior of a thin cylindrical shell in plane strain (see Timoshenko and Woinowsky-Krieger (1959) or Kraus (1967)).

Elimination of N, Q, and M from equations (1)-(5) yields the pair of differential equations in u and v deduced by Fettahlioglu and Mayers (1977) by use of Hamilton's Principle. It would be entirely appropriate to follow their lead and generate a general solution to these equations in each of the regions of contact and noncontact and ultimately proceed to a solution of the present problem. However, some special features of the problem allow a less formal approach to be more efficient.

Figure 1 depicts the ring with the opening angle 2β and the interfering cylinder. Because of symmetry, only the range $0<\theta<\pi-\beta$ need be considered. Continuous contact is impossible in the neighborhood of an end, $\theta=0$, of the ring; this is because continuous contact must be associated with a bending moment which is constant, as will be seen shortly. The unknown angle ϕ identifies the region over which the ring and cylinder are not in contact. We assume continuous contact for $\theta>\phi$.

The free-body diagram (Fig. 2), which shows the concentrated contact force, F, at $\theta=0$, allows us to conclude from a balance of moments that $M=RN$ everywhere. The same result could equivalently be deduced by eliminating Q from equations (2) and (3), integrating, and then using the fact that $N=M=0$ at $\theta=0$. Combining the constitutive relations with this result,

$$\frac{1}{R^2}\frac{d^2u}{d\theta^2} - \frac{1}{R^2}\frac{dv}{d\theta} = \frac{M}{EI}$$

$$\frac{1}{R^2}\frac{d^2u}{d\theta^2} - \frac{1}{R^2}\left(\frac{RN}{EA} - u\right) = \frac{M}{EI}$$

or

$$\frac{d^2u}{d\theta^2} + u = \frac{MR^2}{EI}\left(1 + \frac{I}{AR^2}\right). \tag{6}$$

Now let δ be the radial interference, i.e., the difference between the radius of the rigid cylinder and the (inner) radius of the undeformed ring. Then in the region of continuous contact, $\theta>\phi$, we specify $u=\delta$ and the bending moment is then, by equation (6), the constant

$$M_o = \left(\frac{EI\delta}{R^2}\right)\Big/\left(1 + \frac{I}{AR^2}\right). \tag{7}$$

From Fig. 2 we can see that

$$M_o = FR\sin\phi. \tag{8}$$

In the region of noncontact, $\theta<\phi$,

$$M = FR\sin\theta \tag{9}$$

or

$$M = \frac{M_o\sin\theta}{\sin\phi}.$$

Therefore, in this region equation (6) becomes

$$\frac{d^2u}{d\theta^2} + u = \frac{\delta\sin\theta}{\sin\phi}. \tag{10}$$

[10]Professor, Georgia Institute of Technology, Atlanta, Ga. 30332-0355. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, July 19, 1988; final revision, December 8, 1988.

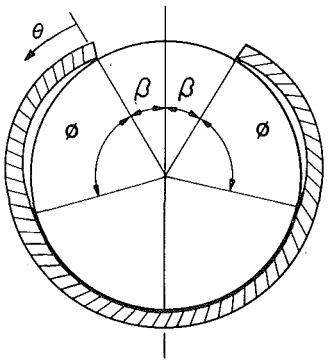


Fig. 1 Interfering ring and cylinder

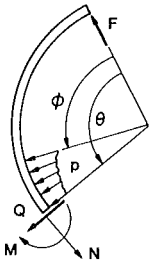


Fig. 2 Free-body diagram

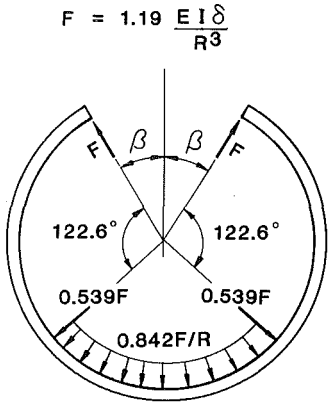
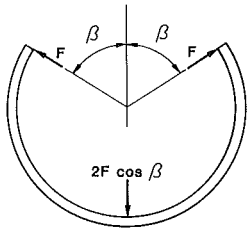


Fig. 3 External forces on ring ($\beta < 57.4$ deg)



$$F = 2\left(\frac{1+\cos\beta}{\pi-\beta+\sin\beta\cos\beta}\right)\frac{EI\delta}{R^3}$$

Fig. 4 External forces on ring ($\beta > 57.4$ deg)

The general solution is

$$u = -\frac{\delta\,\theta\cos\theta}{2\cos\phi} + B\cos\theta + C\sin\theta. \tag{11}$$

The constants B, C, and ϕ are determined from the conditions

$$u(0)=\delta$$
$$u(\phi)=\delta$$
$$\frac{du}{d\theta}(\phi) = 0.$$

Applying these in (11), there results the following transcendental equation establishing the extent of contact,

$$\phi + (\cos\phi - 2)\sin\phi = 0. \tag{12}$$

The only positive root of (12) is, to four significant figures, 2.139 radians (122.6 deg). This is the angle of noncontact on one side of the ring provided, of course, that the opening angle β is less than 180 deg − 122.6 deg = 57.4 deg.

It should be noted that it has not been necessary to resort to the assumption of inextensibility, so common in the analysis of ring problems. On the other hand, the response of the ring in this problem is for all practical purposes inextensible; in the foregoing, wherever the term $(1+I/AR^2)$ has been encountered, unity would have appeared had we made the *a priori* assumption of inextensibility. For a ring of rectangular cross-section and thickness h,

$$\frac{I}{AR^2} = \frac{1}{12}\frac{h^2}{R^2},$$

and the neglect of such terms is already embedded in the governing equations. So for the remainder of this paper I/AR^2 is neglected in comparison to unity.

Returning to stress resultants and external forces, from (7) and (8) the concentrated force at $\theta = 0$ is

$$F = \frac{EI\delta}{R^3\sin\phi} = 1.19\frac{EI\delta}{R^3}. \tag{13}$$

In the region of continuous contact, the stress resultants have the constant values

$$M = FR\sin\phi = 0.842\,FR$$
$$N = 0.842F$$
$$Q = 0.$$

And thus the distributed contact force is

$$p = 0.842\,F/R.$$

From the free-body diagram, Fig. 2, $Q = -F\cos\phi = 0.539\,F$ at $\theta = \phi^-$, and $Q = 0$ at $\theta = \phi^+$. Thus, the plug must exert there a radially outward concentrated force of 0.539F. Figure 3 summarizes the forces exerted by the rigid plug on the ring.

When the opening half-angle β exceeds 57.4 deg there is no region of continous contact; rather, there is contact only at $\theta = 0$ and $\theta = \pi - \beta$ (within the half ring considered because of symmetry). Combining, for this case,

$$\frac{d^2u}{d\theta^2} + u = \frac{MR^2}{EI} \tag{6}$$

and $$M = FR\sin\theta, \tag{9}$$

then $$u = -\frac{FR^3}{2EI}\theta\cos\theta + B\cos\theta + C\sin\theta.$$

Letting $\gamma = \pi - \beta$ for compactness, the boundary conditions are

$u = \delta$ at $\theta = 0$,

$u = \delta$ and $\dfrac{du}{d\theta} = 0$ at $\theta = \gamma$.

Applying these, there results

$$\frac{FR^3}{2EI} = \frac{(1-\cos\gamma)\delta}{\gamma - \sin\gamma\cos\gamma}. \tag{14}$$

Forces on the ring for this case are shown in Fig. 4.

Conclusions

Currently, prevalent alignment sleeves have narrow slits. For the corresponding ring problem, forces of interaction bet-

ween a thin elastic ring and a rigid cylinder are summarized in Fig. 3. The interaction is quite nonuniform—two pairs of concentrated forces and a region of uniform pressure. Neither these forces nor the magnitude of the pressure depend upon the ring-opening angle 2β. The extent of contact depends upon this angle by virtue of the fact that the zone of no contact is independent of the angle.

While actual concentrated forces are in conflict with results from elasticity theory (Dundurs and Comninou, 1979), such forces arise here because of the rigidity of the plug and the structural mechanics theory used to characterize behavior of the ring (see Timoshenko (1955), for example, for similar situations arising out of beam theory).

The case of $\beta > 57.4$ deg (Fig. 4), for which there is only three-point contact, is probably of lesser importance as a mechanical device, but it is interesting to observe that the three forces are equal when $\beta = 60$ deg.

Real finite-length alignment sleeves will necessarily conform to plugs in ways that vary longitudinally in accordance with both sleeve and plug geometries. However, the present analytical solution can, at least, provide a benchmark for numerical analyses designed for the more practical problem.

Acknowledgment

The author's interest in this problem arose in the course of his work at AT&T Bell Laboratories, Atlanta, during the summer of 1987. Thanks are particularly due to J. M. Anderson for many helpful discussions and for sharing the results of his finite element analysis of similar alignment-sleeve problems.

References

Dundurs, J., and Comninou, M., 1979, "Some Consequences of the Inequality Conditions in Contact and Crack Problems," *Journal of Elasticity*, Vol. 9, pp. 71–82.

Fettahlioglu, O. A., and Mayers, J., 1977, "Consistent Treatment of Extensional Deformations for the Bending of Arches, Curved Beams and Rings," ASME *Journal of Pressure Vessel Technology*, Vol. 99, No. 1, pp. 2–11.

Hogan, B. J., 1985, "Flexible Sleeve is Heart of Fiber-Optic Connector," *Design News*, Vol. 41, No. 14, p. 120.

Kraus, H., 1967, *Thin Elastic Shells*, John Wiley and Sons, New York, p. 200.

Timoshenko, S., 1955, *Strength of Materials*, Vol. 1, 3rd ed., D. Van Nostrand, New York, p. 391.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, 2nd ed., McGraw-Hill, New York, p. 512.

Shear Stress Concentration Between Holes

Paul S. Steif[11]

Introduction

The longitudinal shear strength of unidirectionally reinforced fiber composites is occasionally found to be rather low. High concentrations of stress between nearly contacting fibers has been one suggested explanation. The problem of a large body containing two fibers which is subjected to longitudinal shear was studied by Goree and Wilson (1967), who obtained numerical results for the stresses at the fiber-matrix interface. They found that the stress became large as the fibers approached one another. Recently, this problem has been revisited by Budiansky and Carrier (1984); they used complex function theory to show explicitly that the stress becomes unbounded as *rigid* fibers come together.

For many composite systems, however, the fiber and matrix are not well bonded. As can be expected, and as shown explicitly by Steif and Dollar (1988), such a composite responds as if it had holes instead of fibers when the remote shearing becomes large compared with the interfacial shear resistance. Accordingly, this note is devoted to computing the average stress concentration between two holes in an infinite matrix subjected to antiplane shearing. This is done by rigorously proving an analogy between the solutions to problems involving rigid inclusions and those involving holes. Then, the average stress concentration between the holes can be found purely by analogy with Budiansky and Carrier's (1984) explicit result for the average stress between rigid inclusions. The stress concentration right at the hole is also considered, and it is compared with the analogous plane-strain problem that was solved by Savin (1961).

Analysis and Results

The problem considered here is shown schematically in Fig. 1. Two holes of the same radius (taken to be unity) are situated a distance 2ϵ apart in an infinite, homogeneous, linear elastic, isotropic body subjected to remote antiplane shearing $\sigma_{yz} = \tau_o$. We are interested in the distribution of stress in the body. It is useful to employ the complex variable representation which features the stress potential $\omega(Z)$ (analytic in the domain in which the equations of elasticity are satisfied) which is defined by

$$\omega = \phi + iGw \tag{1}$$

where $Z = x + iy$, G is the elastic shear modulus, w is the out-of-plane displacement, and ϕ is the Prandtl stress function often used in torsion problems. The nonzero stresses are derived from ϕ according to

$$\sigma_{xz} = -\frac{\partial \phi}{\partial y} \qquad \sigma_{yz} = \frac{\partial \phi}{\partial x}$$

and, hence, can be computed from ω by

$$\omega' = \sigma_{yz} + i\sigma_{xz} \tag{2}$$

where $(\,)'$ denotes complex differentiation with respect to Z.

The result of interest here is derived by exploiting an analogy with the problem featuring an infinite body, subjected to a remote stress $\sigma_{xz} = \tau_o$, which has rigid inclusions instead of holes. One might surmise such an analogy from the results of Goree and Wilson (1967), who found the stress component $\sigma_{\theta z}$ at the hole to be identical to the component $\sigma_{\rho z}$ at the rigid inclusion, where ρ and θ refer to a polar coordinate system with origin coincident with the center of the hole (inclusion). In fact, as we now demonstrate, there is a very general analogy between the solutions to problems involving rigid inclusions and the solutions to problems involving holes. The proof for this analogy now follows.

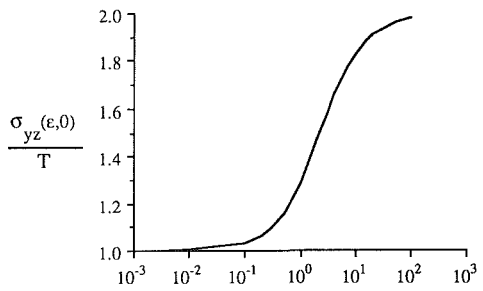


Fig. 1 Schematic of a solid with two holes

[11]Department of Mechanical Engineering, Carnegie-Mellon University, Pittsburgh, Penn. 15213. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, August 16, 1988; final revision, January 23, 1989.

ween a thin elastic ring and a rigid cylinder are summarized in Fig. 3. The interaction is quite nonuniform—two pairs of concentrated forces and a region of uniform pressure. Neither these forces nor the magnitude of the pressure depend upon the ring-opening angle 2β. The extent of contact depends upon this angle by virtue of the fact that the zone of no contact is independent of the angle.

While actual concentrated forces are in conflict with results from elasticity theory (Dundurs and Comninou, 1979), such forces arise here because of the rigidity of the plug and the structural mechanics theory used to characterize behavior of the ring (see Timoshenko (1955), for example, for similar situations arising out of beam theory).

The case of $\beta > 57.4$ deg (Fig. 4), for which there is only three-point contact, is probably of lesser importance as a mechanical device, but it is interesting to observe that the three forces are equal when $\beta = 60$ deg.

Real finite-length alignment sleeves will necessarily conform to plugs in ways that vary longitudinally in accordance with both sleeve and plug geometries. However, the present analytical solution can, at least, provide a benchmark for numerical analyses designed for the more practical problem.

Acknowledgment

The author's interest in this problem arose in the course of his work at AT&T Bell Laboratories, Atlanta, during the summer of 1987. Thanks are particularly due to J. M. Anderson for many helpful discussions and for sharing the results of his finite element analysis of similar alignment-sleeve problems.

References

Dundurs, J., and Comninou, M., 1979, "Some Consequences of the Inequality Conditions in Contact and Crack Problems," *Journal of Elasticity*, Vol. 9, pp. 71–82.

Fettahlioglu, O. A., and Mayers, J., 1977, "Consistent Treatment of Extensional Deformations for the Bending of Arches, Curved Beams and Rings," ASME *Journal of Pressure Vessel Technology*, Vol. 99, No. 1, pp. 2–11.

Hogan, B. J., 1985, "Flexible Sleeve is Heart of Fiber-Optic Connector," *Design News*, Vol. 41, No. 14, p. 120.

Kraus, H., 1967, *Thin Elastic Shells*, John Wiley and Sons, New York, p. 200.

Timoshenko, S., 1955, *Strength of Materials*, Vol. 1, 3rd ed., D. Van Nostrand, New York, p. 391.

Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, 2nd ed., McGraw-Hill, New York, p. 512.

Shear Stress Concentration Between Holes

Paul S. Steif[11]

Introduction

The longitudinal shear strength of unidirectionally reinforced fiber composites is occasionally found to be rather low. High concentrations of stress between nearly contacting fibers has been one suggested explanation. The problem of a large body containing two fibers which is subjected to longitudinal shear was studied by Goree and Wilson (1967), who obtained numerical results for the stresses at the fiber-matrix interface. They found that the stress became large as the fibers approached one another. Recently, this problem has been revisited by Budiansky and Carrier (1984); they used complex function theory to show explicitly that the stress becomes unbounded as *rigid* fibers come together.

For many composite systems, however, the fiber and matrix are not well bonded. As can be expected, and as shown explicitly by Steif and Dollar (1988), such a composite responds as if it had holes instead of fibers when the remote shearing becomes large compared with the interfacial shear resistance. Accordingly, this note is devoted to computing the average stress concentration between two holes in an infinite matrix subjected to antiplane shearing. This is done by rigorously proving an analogy between the solutions to problems involving rigid inclusions and those involving holes. Then, the average stress concentration between the holes can be found purely by analogy with Budiansky and Carrier's (1984) explicit result for the average stress between rigid inclusions. The stress concentration right at the hole is also considered, and it is compared with the analogous plane-strain problem that was solved by Savin (1961).

Analysis and Results

The problem considered here is shown schematically in Fig. 1. Two holes of the same radius (taken to be unity) are situated a distance 2ϵ apart in an infinite, homogeneous, linear elastic, isotropic body subjected to remote antiplane shearing $\sigma_{yz} = \tau_o$. We are interested in the distribution of stress in the body. It is useful to employ the complex variable representation which features the stress potential $\omega(Z)$ (analytic in the domain in which the equations of elasticity are satisfied) which is defined by

$$\omega = \phi + iGw \tag{1}$$

where $Z = x + iy$, G is the elastic shear modulus, w is the out-of-plane displacement, and ϕ is the Prandtl stress function often used in torsion problems. The nonzero stresses are derived from ϕ according to

$$\sigma_{xz} = -\frac{\partial \phi}{\partial y} \qquad \sigma_{yz} = \frac{\partial \phi}{\partial x}$$

and, hence, can be computed from ω by

$$\omega' = \sigma_{yz} + i\sigma_{xz} \tag{2}$$

where $(\,)'$ denotes complex differentiation with respect to Z.

The result of interest here is derived by exploiting an analogy with the problem featuring an infinite body, subjected to a remote stress $\sigma_{xz} = \tau_o$, which has rigid inclusions instead of holes. One might surmise such an analogy from the results of Goree and Wilson (1967), who found the stress component $\sigma_{\theta z}$ at the hole to be identical to the component $\sigma_{\rho z}$ at the rigid inclusion, where ρ and θ refer to a polar coordinate system with origin coincident with the center of the hole (inclusion). In fact, as we now demonstrate, there is a very general analogy between the solutions to problems involving rigid inclusions and the solutions to problems involving holes. The proof for this analogy now follows.

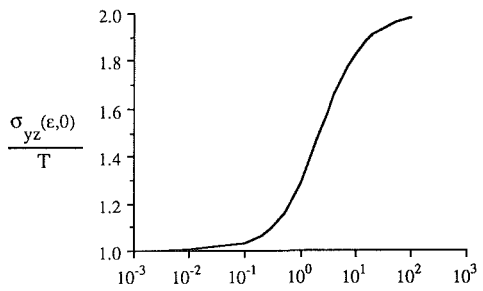


Fig. 1 Schematic of a solid with two holes

[11]Department of Mechanical Engineering, Carnegie-Mellon University, Pittsburgh, Penn. 15213. Assoc. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division, August 16, 1988; final revision, January 23, 1989.

Let $\omega_H(Z)$ denote the stress function which is the solution to the general antiplane problem of an infinite solid that is subjected to remotely applied stresses $\sigma_{yz}=\tau_o$, $\sigma_{xz}=0$ containing any finite number of traction-free holes. Note that $\omega_H(Z)$ is analytic in the multiply connected region (outside the holes). The remote stress implies that $\omega_H(Z)\to\tau_o Z$ as $|Z|\to\infty$. One only needs to specify the values of the real or imaginary parts of $\omega_H(Z)$ on the holes to determine $\omega_H(Z)$ throughout the multiply-connected region (except for an arbitrary constant). The traction on the hole can be expressed in the form

$$T_z=\sigma_{xz}n_x+\sigma_{yz}n_y=-\mathrm{Re}\left[\frac{d\omega_H}{ds}\right] \tag{3}$$

where s denotes arc length along each of the holes, and Re [ ] denotes the real part of the enclosed quantity. Hence, the holes will be free of traction provided

$$\mathrm{Re}\left[\frac{d\omega_H}{ds}\right]=0 \tag{4}$$

at each point of each hole. Clearly, we must also insist that $\omega_H(Z)$ is single valued.

Now contemplate an infinite solid subjected to the remote stress state $\sigma_{xz}=\tau_o$, $\sigma_{yz}=0$ that has rigid inclusions wherever the solid considered above had holes. Let $\omega_R(Z)$, which is analytic in the region exterior to the rigid inclusions, be the solution to this problem. The correct remote stress is obtained if $\omega_R(Z)\to i\tau_o Z$ as $|Z|\to\infty$. As the rigid inclusion is approached from the matrix, the displacement w—and, of course, its tangential derivative dw/ds—must vanish. From (1), the boundary condition for $\omega_R(Z)$ is

$$\mathrm{Im}\left[\frac{d\omega_R}{ds}\right]=0 \tag{5}$$

as each inclusion is approached from the matrix. (Im[ ] denotes the imaginary part.)

From these equations it can be readily seen that $\omega_R(Z)=i\omega_H(Z)$ for all Z in the matrix (actually, they differ by at most a constant) no matter what the number, position, and shape of the holes; therefore, from (2), the stresses in the two problems are related according to

$$(\sigma_{yz}+i\sigma_{xz})_{\text{Rigid Inclusions}}=(-\sigma_{xz}+i\sigma_{yz})_{\text{Holes}}. \tag{6}$$

Given the result of Budiansky and Carrier (1984) for the average stress concentration between two rigid inclusions, one can immediately write down the stress concentration we seek here. Consider the quantity T, which is the average stress σ_{yz} between the holes; i.e.,

$$T=\frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon}\sigma_{yz}dx.$$

This must be equal to the average value of σ_{xz} between rigid inclusions; therefore,

$$T=\tau_o\sqrt{\frac{2+\epsilon}{\epsilon}}. \tag{7}$$

As an alternative means of derivation, one can use an analysis quite similar to Budiansky and Carrier (1984) to obtain a solution for the entire plane. The solution involves the transformation

$$\zeta=\frac{Z-a}{Z+a}\qquad a=\sqrt{\epsilon(2+\epsilon)}$$

which maps the region under consideration into the interior of the annulus $\rho<|\zeta|<1/\rho$ where

$$\rho=\frac{a-\epsilon}{a+\epsilon}.$$

The end result is that $\omega_H(\zeta)$ is given by

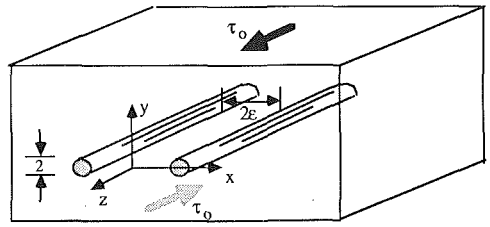


Fig. 2 Maximum shear stress at hole relative to the average ligament shear stress

$$\omega_H(\zeta)=\tau_o\left[\frac{a(1+\zeta)}{1-\zeta}+2a\sum_{n=1}^{\infty}\frac{\rho^{2n}(\zeta^n-1/\zeta^n)}{1-\rho^{2n}}\right]. \tag{8}$$

It is interesting to consider the stress concentration right at the holes, since it is there that one surmises the stresses to be highest. (Note that Goree and Wilson have calculated the stresses at the holes.) In fact, one can use the properties of harmonic functions to show that this is the case. Note that the stresses σ_{xz} and σ_{yz} are the real and imaginary parts of the analytic function $\omega_H'(Z)$; therefore, they are both harmonic in the entire plane outside the two holes. Harmonic functions always take on their maximum and minimum values on the boundary of the region in which they are defined. Hence, the maximum and minimum values of σ_{xz} and σ_{yz} must be found either at infinity or right at the holes.

In particular, consider the stress at $x=\epsilon$; it may be calculated from

$$\sigma_{yz}(\epsilon,0)=\tau_o\left[1-\frac{(1+\rho)^2}{\rho}\sum_{n=1}^{\infty}n(-\rho)^n\frac{\rho^n+\dfrac{1}{\rho^n}}{\dfrac{1}{\rho^n}-\rho^n}\right] \tag{9}$$

which we have evaluated numerically. The stress concentration right at the hole relative to the average stress between the holes is shown in Fig. 2, where we have plotted $\sigma_{yz}(\epsilon,0)/T$ versus ϵ. As the holes become close together, the stress between them becomes uniform. From the distributions of stress found by Goree and Wilson, it may be seen that the maximum stress around the hole is at the point $x=\epsilon$ that we have just considered. Thus, from the fact that the maximum must be found on the hole (the stress at $x=\epsilon$ is higher than the stress at infinity) and from the stress distributions given by Goree and Wilson, it is extremely likely that the absolute maximum is at $x=\epsilon$, although we could not prove this from (8) directly.

Finally, a comparison is made with the analogous plane-strain problem solved by Savin (1961). For $\epsilon=0.5$, Savin finds the stress concentration at $x=\epsilon$ to be 3.264 in plane strain, while we find the stress concentration in antiplane strain to be 2.594. (As far as can be seen from Fig. 2 in Goree and Wilson, this agrees with their result.) Recall, however, that the stress concentration associated with an isolated hole is 3.0 in plane strain and 2.0 in antiplane shear. Hence, the *relative* effect of a nearby hole is greater for antiplane shear than for plane strain.

Acknowledgments

The author appreciates the comments made by anonymous reviewers and by Associate Editor C. O. Horgan.

References

Budiansky, B., and Carrier, G. F., 1984, "High Shear Stress in Stiff Fiber Composites," ASME JOURNAL OF APPLIED MECHANICS, Vol. 51, pp. 733–735.

Goree, J. G., and Wilson, H. B., 1967, "Transverse Shear Loading in an Elastic Matrix Containing Two Circular Cylindrical Inclusions," ASME JOURNAL OF APPLIED MECHANICS, Vol. 34, pp. 511–513.

Savin, G. N., 1961, *Stress Concentration Around Holes*, Pergamon Press, London.

Steif, P. S., and Dollar, A., 1988, "Longitudinal Shearing of a Weakly Bonded Fiber Composite," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 618–623.

A Note on the Efficiency of the Boundary Element Method for Inelastic Axisymmetric Problems With Large Strains

H. Rajiyah[12] and S. Mukherjee[13]

1 Introduction

A comparative study of inelastic axisymmetric problems in the presence of large strains by the boundary element method (BEM) and the finite element method (FEM) is presented here. The FEM (for example, Hibbitt, Marcal, and Rice (1970) and Yamada and Hirakawa (1978)) has been the most popular method of choice for the analysis of these complex class of problems that include both geometric and material nonlinearities. The work done in this area using the FEM is reasonably well understood at present. The FEM formulation presented in this paper is based on a Galerkin-type weighted residual approach.

As for the BEM, Chandra and Mukherjee (1983) and Okada, Rajiyah, and Atluri (1988) have proposed BEM formulations for material as well as geometrically nonlinear problems. As noted by Sarihan and Mukherjee (1982) among others, the modeling of axisymmetric problems by the BEM is far more challenging than its planar counterpart. Recently, Rajiyah and Mukherjee (1987) have competed a BEM study of inelastic axisymmetric problems with large strains and rotations. This has been the first attempt of its kind to study such a class of problems by the BEM. An updated Lagrangian approach is adopted for both the BEM and FEM approaches.

2 Constitutive Assumptions

The key assumptions used in both the FEM and BEM formulations are presented as follows: The rate of deformation tensor can be linearly decomposed into an elastic or a nonelastic part

$$d_{ij} = d_{ij}^{(e)} + d_{ij}^{(n)}. \tag{1}$$

A hypoelastic material law relates $d_{ij}^{(e)}$ to an objective stress rate

$$\sigma_{ij}^{*} = \lambda d_{kk}^{(e)} \delta_{ij} + 2G d_{ij}^{(e)} \tag{2}$$

where σ_{ij}^{*} is the Jaumann rate of the Cauchy stress and λ and G are Lame constants. The Jaumann rate is related to the material rate by the formula

$$\overset{*}{t}_{ij} = \overset{\circ}{t}_{ij} + t_{ik}\omega_{kj} - \omega_{ik}t_{kj} \tag{3}$$

where the material derivative is defined as

$$\overset{\circ}{t}_{ij} = \frac{\partial t_{ij}}{\partial t} + \frac{\partial t_{ij}}{\partial x_k} v_k. \tag{4}$$

Here $\partial t_{ij}/\partial t$ is the time derivative, $\partial t_{ij}/\partial x_k$ is the spacial derivative, and v_k is the velocity of the material point under consideration. The nonelastic part of the deformation gradient $d_{ij}^{(n)}$ must be determined from an appropriate constitutive model. The constitutive model proposed by Anand (1982) is used for illustrative purposes here.

3 Axisymmetric Boundary Element Formulation

In the cylinder polar coordinate system R, θ, and Z, the nonzero components of displacements, stresses, and strains are u_R, u_Z and ϵ_{RR}, $\epsilon_{\theta\theta}$, ϵ_{ZZ}, ϵ_{RZ} $(=\epsilon_{ZR})$, σ_{RR}, $\sigma_{\theta\theta}$, σ_{ZZ} and σ_{RZ} $(=\sigma_{ZR})$. All dependent variables are functions of R, Z, and t.

The notation used here is the same as in Rajiyah and Mukherjee (1987). The source point is denoted by (R, θ, Z) and the field point by (ρ, θ, ζ). Since the problem is axisymmetric, it is sufficient to choose the source point in the $x_1 - x_3$ plane.

The axisymmetric BEM equations are derived by integrating the three-dimensional kernels U_{ij}, T_{ij}, etc., for the field point moving around a ring keeping the source point fixed. Integrating θ from 0 to 2π results in ($j = 1$ and 3, no sum over ρ or ζ)

$$\begin{aligned} v_j &= \int_{\partial B^o} [U_{\rho j}\dot{\tau}_\rho + U_{\zeta j}\dot{\tau}_\zeta - T_{\rho j}\dot{u}_\rho - T_{\zeta j}\dot{u}_\zeta]\rho_o dc_o \\ &+ 2G\int_{B^o} \Big[U_{\rho j,\rho} d_{\rho\rho}^{(n)} + U_{\rho j,\zeta} d_{\rho\zeta}^{(n)} + U_{\zeta j,\rho} d_{\zeta\rho}^{(n)} \\ &+ U_{\zeta j,\zeta} d_{\zeta\zeta}^{(n)} + \frac{U_{\rho j} d_{\theta\theta}^{(n)}}{\rho_o}\Big]\rho_o d\rho_o d\zeta_o \\ &+ \int_{B^o} \Big[U_{\rho j,\rho}[\sigma_{\rho\rho} d_{\rho\rho} + \sigma_{\rho\zeta}(d_{\rho\zeta} - \omega_{\rho\zeta})] + U_{\rho j,\zeta}[\sigma_{\rho\rho} d_{\rho\zeta} \\ &+ \sigma_{\rho\zeta} d_{\zeta\zeta} - \sigma_{\zeta\zeta}\omega_{\rho\zeta}] + U_{\zeta j,\rho}[\sigma_{pp}\omega_{p\zeta} + \sigma_{\rho\zeta} d_{\rho\rho} + \sigma_{\zeta\zeta} d_{\rho\zeta}] \\ &+ U_{\zeta j,\zeta}[\sigma_{\rho\zeta}(d_{\rho\zeta} + \omega_{\rho\zeta}) + \sigma_{\zeta\zeta} d_{\zeta\zeta}] + \frac{U_{\rho j}\sigma_{\theta\theta} d_{\theta\theta}}{\rho_o}\Big]\rho_\vartheta d\rho_\vartheta d\zeta_o \end{aligned} \tag{5}$$

where, because of axisymmetry, $\dot{u}_R = \dot{u}_1$, $\dot{u}_Z = \dot{u}_3$ and $dc_o = \sqrt{d\rho^2 + d\zeta^2}$ is an element on the boundary of the $\rho - \zeta$ plane. The domain B^o and boundary ∂B^o in the equation (5) now refer to a generator plane of the axisymmetric solid and its boundary (excluding the portion on the x_3 axis), respectively, (Fig. 1 of Rajiyah and Mukherjee (1987)) so that the three-dimensional problem is effectively reduced to a two-dimensional one. Equation (5) is valid for the velocity at an internal source point. The corresponding boundary integral equation is obtained, as usual, by taking the limit $p \to P$ which introduces the corner tensor C_{ij}.

The kernels T_{ij} and U_{ij}, when $p \notin$ axis of symmetry, become singular when the source point coincides with the axis of symmetry. Hence, special T_{ij} and U_{ij} kernels need to be evaluated when $p \in$ axis of symmetry. The explicit form of these kernels which appear in equation (5) are given in Rajiyah and Mukherjee (1987).

The Lagrangian traction rates from equation (5) become

$$\begin{aligned} \dot{\tau}_\rho &= \dot{t}_\rho - n_\rho[\sigma_{\rho\rho} d_{\rho\rho} + \sigma_{\rho\zeta}(d_{\rho\zeta} - \omega_{\rho\zeta})] - n_\zeta[\sigma_{\rho\rho} d_{\rho\zeta} + \sigma_{\rho\zeta} d_{\zeta\zeta} - \sigma_{\zeta\zeta}\omega_{\rho\zeta}] \\ \dot{\tau}_\zeta &= \dot{t}_\zeta - n_\rho[\sigma_{\rho\rho}\omega_{\rho\zeta} + \sigma_{\rho\zeta} d_{\rho\rho} + \sigma_{\zeta\zeta} d_{\rho\zeta}] - n_\zeta[\sigma_{\rho\zeta} + \omega_{\rho\zeta}) + \sigma_{\zeta\zeta} d_{\zeta\zeta}] \end{aligned} \tag{6}$$

where $\dot{t}_\rho$ and $\dot{t}_\xi$ are the prescribed traction rates which take the following form

$$\begin{aligned} \dot{t}_\rho &= \overset{*}{\sigma}_{\rho\rho} n_\rho + \overset{*}{\sigma}_{\rho\xi} n_\xi \\ \dot{t}_\xi &= \overset{*}{\sigma}_{\xi\rho} n_\rho + \overset{*}{\sigma}_{\xi\xi} n_\xi. \end{aligned} \tag{7}$$

Velocity gradients at an internal point are obtained, as usual, by differentiating equation (5) at an internal source point p. The domain integrals in equation (5), in general, are $1/r$ singular and special care must be taken when differentiating them analytically. Free terms result from these integrals are these must be taken into consideration when evaluating these expressions. All such free-term contributions are listed in Rajiyah and Mukherjee (1987).

[12]Assistant Professor, Georgia Institute of Technology, School of Civil Engineering, Atlanta, GA 30332.

[13]Professor, Cornell University, Ithaca, N.Y. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division March 16, 1988; final revision, November 16, 1988.

Savin, G. N., 1961, *Stress Concentration Around Holes*, Pergamon Press, London.

Steif, P. S., and Dollar, A., 1988, "Longitudinal Shearing of a Weakly Bonded Fiber Composite," ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, pp. 618–623.

A Note on the Efficiency of the Boundary Element Method for Inelastic Axisymmetric Problems With Large Strains

H. Rajiyah[12] and S. Mukherjee[13]

1 Introduction

A comparative study of inelastic axisymmetric problems in the presence of large strains by the boundary element method (BEM) and the finite element method (FEM) is presented here. The FEM (for example, Hibbitt, Marcal, and Rice (1970) and Yamada and Hirakawa (1978)) has been the most popular method of choice for the analysis of these complex class of problems that include both geometric and material nonlinearities. The work done in this area using the FEM is reasonably well understood at present. The FEM formulation presented in this paper is based on a Galerkin-type weighted residual approach.

As for the BEM, Chandra and Mukherjee (1983) and Okada, Rajiyah, and Atluri (1988) have proposed BEM formulations for material as well as geometrically nonlinear problems. As noted by Sarihan and Mukherjee (1982) among others, the modeling of axisymmetric problems by the BEM is far more challenging than its planar counterpart. Recently, Rajiyah and Mukherjee (1987) have competed a BEM study of inelastic axisymmetric problems with large strains and rotations. This has been the first attempt of its kind to study such a class of problems by the BEM. An updated Lagrangian approach is adopted for both the BEM and FEM approaches.

2 Constitutive Assumptions

The key assumptions used in both the FEM and BEM formulations are presented as follows: The rate of deformation tensor can be linearly decomposed into an elastic or a nonelastic part

$$d_{ij} = d_{ij}^{(e)} + d_{ij}^{(n)}. \tag{1}$$

A hypoelastic material law relates $d_{ij}^{(e)}$ to an objective stress rate

$$\sigma_{ij}^{*} = \lambda d_{kk}^{(e)} \delta_{ij} + 2G d_{ij}^{(e)} \tag{2}$$

where σ_{ij}^{*} is the Jaumann rate of the Cauchy stress and λ and G are Lame constants. The Jaumann rate is related to the material rate by the formula

$$\overset{*}{t}_{ij} = \overset{\circ}{t}_{ij} + t_{ik}\omega_{kj} - \omega_{ik}t_{kj} \tag{3}$$

where the material derivative is defined as

$$\overset{\circ}{t}_{ij} = \frac{\partial t_{ij}}{\partial t} + \frac{\partial t_{ij}}{\partial x_k} v_k. \tag{4}$$

Here $\partial t_{ij}/\partial t$ is the time derivative, $\partial t_{ij}/\partial x_k$ is the spacial derivative, and v_k is the velocity of the material point under consideration. The nonelastic part of the deformation gradient $d_{ij}^{(n)}$ must be determined from an appropriate constitutive model. The constitutive model proposed by Anand (1982) is used for illustrative purposes here.

3 Axisymmetric Boundary Element Formulation

In the cylinder polar coordinate system R, θ, and Z, the nonzero components of displacements, stresses, and strains are u_R, u_Z and ϵ_{RR}, $\epsilon_{\theta\theta}$, ϵ_{ZZ}, ϵ_{RZ} $(=\epsilon_{ZR})$, σ_{RR}, $\sigma_{\theta\theta}$, σ_{ZZ} and σ_{RZ} $(=\sigma_{ZR})$. All dependent variables are functions of R, Z, and t.

The notation used here is the same as in Rajiyah and Mukherjee (1987). The source point is denoted by (R, θ, Z) and the field point by (ρ, θ, ζ). Since the problem is axisymmetric, it is sufficient to choose the source point in the $x_1 - x_3$ plane.

The axisymmetric BEM equations are derived by integrating the three-dimensional kernels U_{ij}, T_{ij}, etc., for the field point moving around a ring keeping the source point fixed. Integrating θ from 0 to 2π results in ($j = 1$ and 3, no sum over ρ or ζ)

$$\begin{aligned} v_j &= \int_{\partial B^o} [U_{\rho j}\dot{\tau}_\rho + U_{\zeta j}\dot{\tau}_\zeta - T_{\rho j}\dot{u}_\rho - T_{\zeta j}\dot{u}_\zeta]\rho_o dc_o \\ &+ 2G\int_{B^o} \Big[U_{\rho j,\rho} d_{\rho\rho}^{(n)} + U_{\rho j,\zeta} d_{\rho\zeta}^{(n)} + U_{\zeta j,\rho} d_{\zeta\rho}^{(n)} \\ &+ U_{\zeta j,\zeta} d_{\zeta\zeta}^{(n)} + \frac{U_{\rho j} d_{\theta\theta}^{(n)}}{\rho_o}\Big]\rho_o d\rho_o d\zeta_o \\ &+ \int_{B^o} \Big[U_{\rho j,\rho}[\sigma_{\rho\rho} d_{\rho\rho} + \sigma_{\rho\zeta}(d_{\rho\zeta} - \omega_{\rho\zeta})] + U_{\rho j,\zeta}[\sigma_{\rho\rho} d_{\rho\zeta} \\ &+ \sigma_{\rho\zeta} d_{\zeta\zeta} - \sigma_{\zeta\zeta}\omega_{\rho\zeta}] + U_{\zeta j,\rho}[\sigma_{pp}\omega_{p\zeta} + \sigma_{\rho\zeta} d_{\rho\rho} + \sigma_{\zeta\zeta} d_{\rho\zeta}] \\ &+ U_{\zeta j,\zeta}[\sigma_{\rho\zeta}(d_{\rho\zeta} + \omega_{\rho\zeta}) + \sigma_{\zeta\zeta} d_{\zeta\zeta}] + \frac{U_{\rho j}\sigma_{\theta\theta} d_{\theta\theta}}{\rho_o}\Big]\rho_\vartheta d\rho_\vartheta d\zeta_o \end{aligned} \tag{5}$$

where, because of axisymmetry, $\dot{u}_R = \dot{u}_1$, $\dot{u}_Z = \dot{u}_3$ and $dc_o = \sqrt{d\rho^2 + d\zeta^2}$ is an element on the boundary of the $\rho - \zeta$ plane. The domain B^o and boundary ∂B^o in the equation (5) now refer to a generator plane of the axisymmetric solid and its boundary (excluding the portion on the x_3 axis), respectively, (Fig. 1 of Rajiyah and Mukherjee (1987)) so that the three-dimensional problem is effectively reduced to a two-dimensional one. Equation (5) is valid for the velocity at an internal source point. The corresponding boundary integral equation is obtained, as usual, by taking the limit $p \to P$ which introduces the corner tensor C_{ij}.

The kernels T_{ij} and U_{ij}, when $p \notin$ axis of symmetry, become singular when the source point coincides with the axis of symmetry. Hence, special T_{ij} and U_{ij} kernels need to be evaluated when $p \in$ axis of symmetry. The explicit form of these kernels which appear in equation (5) are given in Rajiyah and Mukherjee (1987).

The Lagrangian traction rates from equation (5) become

$$\begin{aligned} \dot{\tau}_\rho &= \dot{t}_\rho - n_\rho[\sigma_{\rho\rho} d_{\rho\rho} + \sigma_{\rho\zeta}(d_{\rho\zeta} - \omega_{\rho\zeta})] - n_\zeta[\sigma_{\rho\rho} d_{\rho\zeta} + \sigma_{\rho\zeta} d_{\zeta\zeta} - \sigma_{\zeta\zeta}\omega_{\rho\zeta}] \\ \dot{\tau}_\zeta &= \dot{t}_\zeta - n_\rho[\sigma_{\rho\rho}\omega_{\rho\zeta} + \sigma_{\rho\zeta} d_{\rho\rho} + \sigma_{\zeta\zeta} d_{\rho\zeta}] - n_\zeta[\sigma_{\rho\zeta} + \omega_{\rho\zeta}) + \sigma_{\zeta\zeta} d_{\zeta\zeta}] \end{aligned} \tag{6}$$

where $\dot{t}_\rho$ and $\dot{t}_\xi$ are the prescribed traction rates which take the following form

$$\begin{aligned} \dot{t}_\rho &= \overset{*}{\sigma}_{\rho\rho} n_\rho + \overset{*}{\sigma}_{\rho\xi} n_\xi \\ \dot{t}_\xi &= \overset{*}{\sigma}_{\xi\rho} n_\rho + \overset{*}{\sigma}_{\xi\xi} n_\xi. \end{aligned} \tag{7}$$

Velocity gradients at an internal point are obtained, as usual, by differentiating equation (5) at an internal source point p. The domain integrals in equation (5), in general, are $1/r$ singular and special care must be taken when differentiating them analytically. Free terms result from these integrals are these must be taken into consideration when evaluating these expressions. All such free-term contributions are listed in Rajiyah and Mukherjee (1987).

[12]Assistant Professor, Georgia Institute of Technology, School of Civil Engineering, Atlanta, GA 30332.

[13]Professor, Cornell University, Ithaca, N.Y. Mem. ASME.

Manuscript received by the ASME Applied Mechanics Division March 16, 1988; final revision, November 16, 1988.

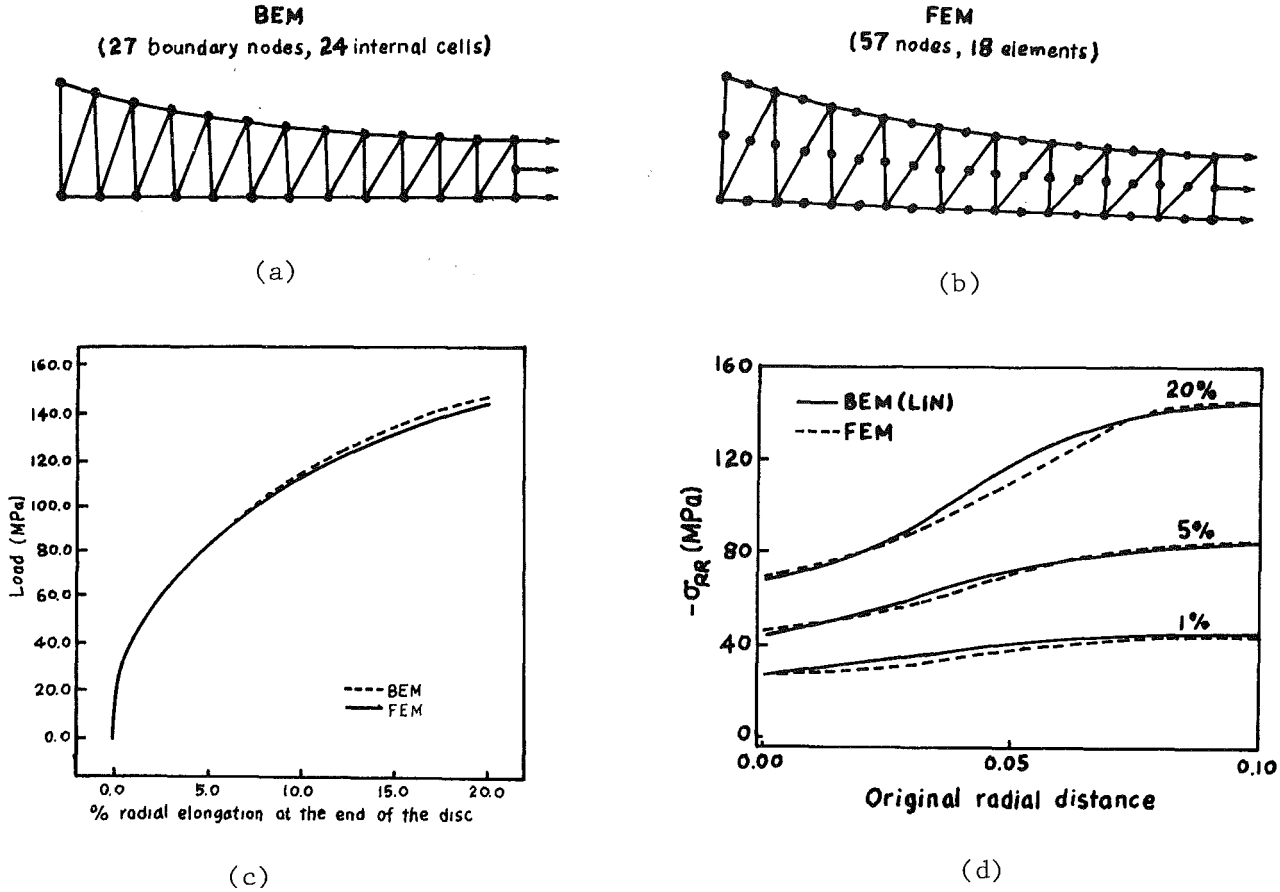


Fig. 1 (*a*) BEM mesh descretization, (*b*) FEM mesh descretization, (*c*) load versus percent of radial elongation at the end of the disc (up to 20 percent), and (*d*) radial stresses in the beam at various stages of radial elongation.

Stress Rates and Spin on the Boundary. The boundary stress rates, at any time, are best obtained from a boundary algorithm. The normal and tangential components of the traction rate vector are first calculated at some point P on ∂B (P is assumed to lie at a point on ∂B where it is locally smooth). Now, in the rotated local coordinate system $c-n$ (where c is tangential to ∂B at P)

$$\overset{*}{\sigma}_{nn} = -\dot{t}_n, \quad \overset{*}{\sigma}_{nc} = -\dot{t}_c \tag{8}$$

where $\overset{*}{\sigma}_{nn}$ and $\overset{*}{\sigma}_{nc}$ are the normal and shearing components of the Jaumann rates of the Cauchy stress at P. Next, the normal and tangential components of the displacement rate vector are calculated at P and the tangential derivatives of $\dot{u}_c$ and $\dot{u}_n$, $[\partial \dot{u}_c/\partial c$ and $\partial \dot{u}_n/\partial c]$ are obtained at P by numerical differentiation along the boundary element. The constitutive equations are written as

$$\frac{\partial \dot{u}_c}{\partial c} = d_{cc} = \frac{1}{E}[\overset{*}{\sigma}_{cc} - \nu(\overset{*}{\sigma}_{nn} + \overset{*}{\sigma}_{\theta\theta})] + d_{cc}^{(n)}$$

$$\frac{\dot{u}_R}{R} = d_{\theta\theta} = \frac{1}{E}[\overset{*}{\sigma}_{\theta\theta} - \nu(\overset{*}{\sigma}_{nn} + \overset{*}{\sigma}_{cc})] + d_{\theta\theta}^{(n)}$$

$$\omega_{cn} = \omega_{RZ} = \overset{*}{\sigma}_{nc}/2G + d_{nc}^{(n)} - \frac{\partial \dot{u}_n}{\partial c} \tag{9}$$

4 Finite Element Formulation

Using a Galerkin-type weighted residual of the equilibrium equation with a virtual velocity field as the test function, together with the divergence theorem, and relationships between various stress tensors, one can easily show that (Yamada and Hirakawa (1978))

$$\int_{B^o} [\overset{*}{\tau}_{ij}\delta(d_{ij}) - \frac{1}{2}\sigma_{ij}\delta(2d_{ik}d_{kj} - v_{k,i}v_{k,j})]dV^o$$

$$- \int_{\partial B_F^o} \dot{\tau}_j \delta v_j dS^o - \int_{B^o} \rho_o \dot{F}_j^o \delta v_j dV^o = 0 \tag{10}$$

where B^o is the domain of the body and ∂B_F^o is the part of the boundary of the body on which tractions are applied. Introducing interpolating functions for velocities at the element level, one can rewrite equation (10) in matrix form as

$$([\mathbf{K}] + [K_G] + [K_L])\{V\} = \{P^{(n)}\} + \{P^{(B)}\} \tag{11}$$

where $[K]$ is the stiffness matrix defined as in the case of small deformations and $[K_G]$ and $[K_L]$ are the geometric and load correction matrices, respectively. $\{P^{(n)}\}$ is the inelastic load vector and $\{P^{(B)}\}$ is the load vector defined in the usual manner. A six-noded triangular (T6) element was considered for this formulation.

5 Comparison of BEM and FEM Results

5.1 Radial Expansion of a Tapered Disc. A constant distributed rate of loading (5 MPa/s) is applied at the outer edge. The BEM and FEM meshes for this problem are shown in Fig. 1. Both quadratic and linear variation of tractions and displacements on the boundary are tried for the BEM mesh and variation between the results are found to be negligible. The BEM and FEM agree very well up to a deformation of 20 percent of radial elongation at the inner surface. The BEM took 246 seconds of CPU time as compared to the FEM which took 98 seconds of CPU time.

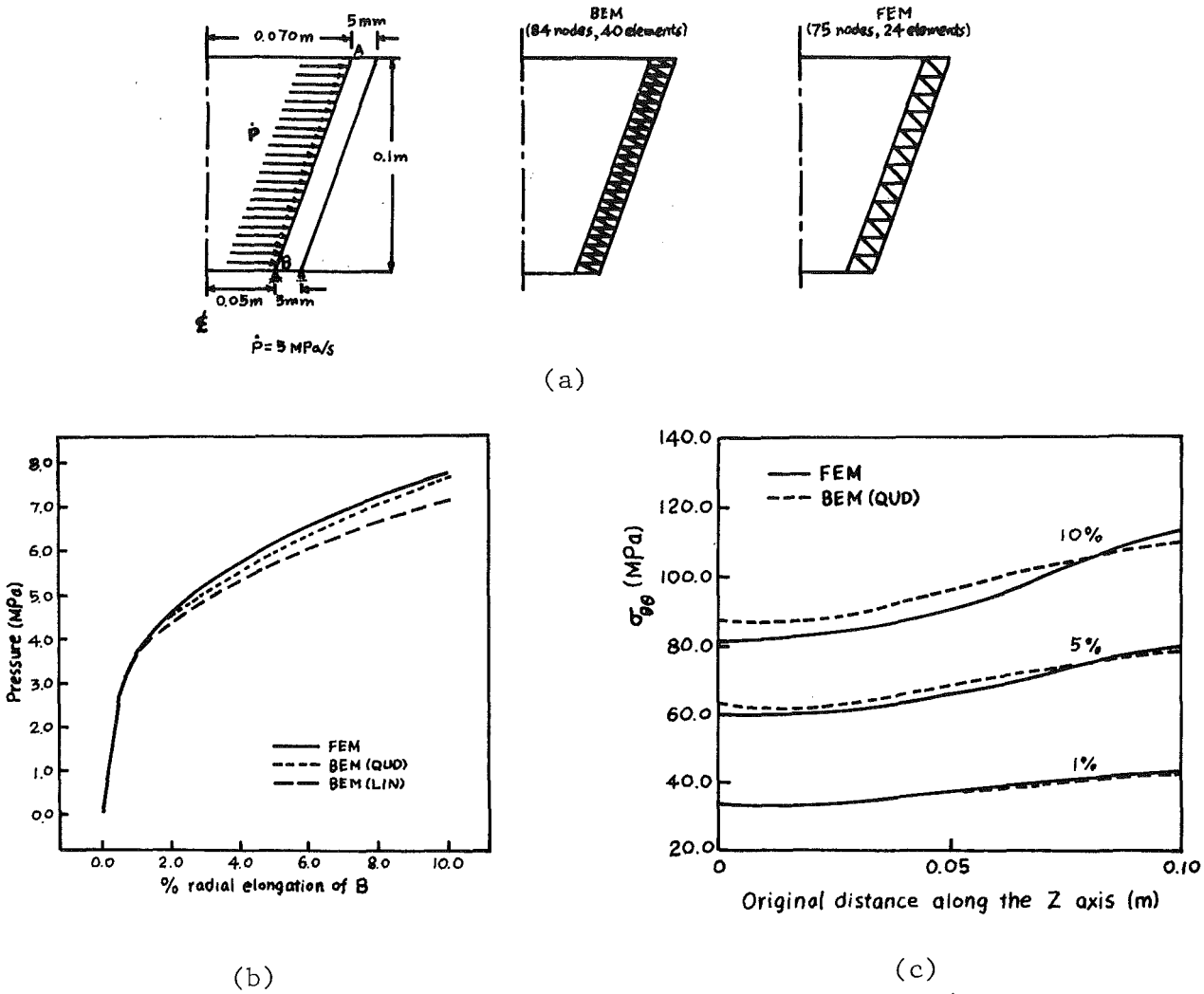


Fig. 2 (a) BEM and FEM mesh descretizations, (b) pressure versus the percent of radial elongation at B (see (a)), (c) hoop stresses versus the original distance along the z-axis at different stages of radial elongation of B

5.2 Radial Expansion of a Cone. A constant distributed rate of loading (5 MPs/s) is applied at the inner surface of the cone. The BEM and FEM mesh descretizations are given in Fig. 2. Both quadratic and linear variations of displacement and tractions on the boundary and linear variation of nonelastic deformation rates inside the body were considered for this problem. Comparisons were made up to 10 percent of radial deflection at point A. The BEM took 217 seconds of CPU time as compared to the FEM which took 69 seconds of CPU time.

Both of these problems were run on an IBM 3090/400 computer.

6 Conclusions

Although the BEM is comparable in terms of accuracy to the FEM in the aforementioned class of problems, the efficiency of the BEM (unlike in the linear case) seems poor when compared to the FEM. One obvious reason is that domain descretization now becomes necessary with the unknown velocities being present inside as well as on the boundary in equation (5). Also, the singular sensitive and complex axisymmetric kernels of the BEM must be repeatedly evaluated and integrated with changing geometry when using an updated Lagrangian approach. In earlier work on planar problems (Chandra and Mukherjee (1983)) the kernels were simpler and could be evaluated in closed form. This led to the BEM computer times being about half of those from the FEM calculations for planar problems when closed-form integrations were performed on kernel functions. The situation is obviously different for axisymmetric problems.

Another difficulty with the BEM is that the presence of internal points further increases the computational burden. Thus, the BEM computer times are expected to increase substantially if the number of internal points increases as compared to boundary points. Alternate mechanics formulations and specialized computing algorithms should be explored if the BEM is to become a viable numerical method for such a class of problems. Ideas along these lines are the use of a total Lagrangian formulation and/or the use of vector and parallel computing.

Acknowledgments

The research reported here has been supported by NSF Grant No. MSM-8609391 to Cornell University. The computing has been supported by the Cornell National Supercomputer Facility.

References

Anand, L., 1982, "Constitutive Equations for the Rate Dependent Deformation of Metals," *ASME Journal of Engineering Materials and Technology*, Vol. 104, No. 1, pp. 12–17.

Chandra, A., and Mukherjee, S., 1983, "Applications of Boundary Element Method to Large Strain Large Deformation Problems of Viscoplasticity," *Journal of Strain Analysis*, Vol. 8, No. 4, pp. 261-270.

Hibbit, H. D., Marcal, P. V., and Rice, J. R., 1970, "A Finite Element Formulation for Problems of Large Strain and Large Displacements," *International Journal of Solids and Structures*, Vol. 6, No. 8, pp. 1069-1086.

Okada, H., Rajiyah, H., ad Atluri, S. N., 1988, "Some Recent Developments in Finite Strain Elastoplasticity Using the Field-Boundary Element Method," *Computers and Structures*, Vol. 30, No. 1-2, pp. 275-288.

Rajiyah, H., and Mukherjee, S., 1987, "Boundary Element Analysis of Inelastic Axisymmetric Problems with Large Strains and Rotations," *International Journal of Solids and Structures*, Vol. 23, No. 12, pp. 1679-1698.

Sarihan, V., and Mukherjee, S., 1982, "Axisymmetric Viscoplastic Deformation by the Boundary Element Method," *International Journal of Solids and Structures*, Vol. 18, No. 12, pp. 1113-1128.

Yamada, Y., and Hirakawa, H., 1978, "Large Deformation and Instability Analysis in Metal Forming Processes," *Applications of Numerical Methods to Forming Process*, H. Armen, and R. G. Jones, eds., AMD-Vol. 28, ASME, New York, pp. 27-38.

On the Range of Applicability of von Karman Plate Equations

You-He Zhou[14] and Xiao-Jing Zheng[14]

Introduction

In the derivation of the von Karman plate equations (von Karman, 1910), as is known, there are two essential assumptions: one of which is the Kirchhoff's hypothesis and the other is the approximate calculations of the bending curvatures. The former requires that the thickness of a plate be small in comparison with its smallest lateral dimension. In most applications, the span-to-thickness ratio of a plate is greater than 15, so that the Kirchhoff's hypothesis may be valid. The latter requires that the plate slope, or rotations, should be less than 0.2. According to this requirement, many scholars (Timoshenko and Woinwsky-Krieoger, 1959; Chien et al., 1954; Chia, 1980) qualitatively consider that the ratio of deflection-to-thickness is of the same order with the unit, that is, $w/h \sim 0(1)$, which cannot explain why the solutions of the von Karman plate equations agree closely with the experiment data up to the deflection-to-thickness of six (Chia, 1980). Strictly speaking, the range should be decided by comparing the solutions of the von Karman theory with those of the more accurate Reissner plate equations (Reissner, 1958). However, it is difficult if not impossible to solve the Reissner equations for each problem of a thin plate. The best way is to check whether or not the prerequisites of the von Karman plate equations are satisfied by their solutions (see in Fig. 3). If the prerequisites are satisfied, we should be convinced of the feasibility of the von Karman theory.

It follows, in this paper, that a new quantitative criterion of the fitting range of the von Karman equations is to be derived from the analysis of errors caused by the approximate calculations based on the prerequisites, which shows that the von Karman theory in large deflection is reasonable. Finally, the range of the rigidly clamped, circular thin plate under uniform pressure will be given.

For the sake of simplicity, we discuss here the range of applicability of the axisymmetrical equations of the von Karman plate theory, which the dimensionless and independent differential equations can be written in the form (Zheng and Zhou, 1987; Zheng and Zhou, 1988):

$$y^2 \frac{d^2\varphi}{dy^2} = \varphi(y)S(y) + f(y), \tag{1}$$

$$y^2 \frac{d^2S}{dy^2} = -\frac{1}{2}(\varphi(y))^2, \quad y\in(0,1). \tag{2}$$

And the boundary conditions are

$$y=0:\ \varphi(y)=0;\ S(y)=0; \tag{3}$$

$$y=1:\ \varphi(y)=\frac{\lambda}{\lambda-1}\frac{d\varphi}{dy};\ S(y)=\frac{\mu}{\mu-1}\frac{dS}{dy}. \tag{4}$$

Here,

$$y=\left(\frac{r}{a}\right)^2;\ W=(3(1-\nu^2))^{1/2}\frac{w}{h};\ \varphi(y)=y\cdot\frac{dW}{dy};$$

$$S(y)=\frac{3(1-\nu^2)a^2N_r y}{Eh^3}$$

in which w is the deflection of a plate, r is the radial coordinate, h is the thickness of a plate, a is the radius of outer edge of a plate, N_r is the radial stress per unit length, E is the Young's modulus; ν is the Poisson's ratio, $f(y)$ is a function represented the distribution of axisymmetric loads, λ and μ are constants dependent on the boundary condition, e.g., for the rigidly-clamped edge, $\lambda=0$ and $\mu=2/1-\nu$.

Analysis of Errors

Since the Kirchhoff's hypothesis is held for most applications in engineering, thus, we merely discuss the errors caused by the approximate calculations imposed on the bending curvatures, etc. From the process of establishing equations (1)-(2), the approximate calculations, except for those from the Kirchhoff's hypothesis, are in the form

$$\frac{1}{\rho_r}=\frac{-\frac{d^2w}{dr^2}}{\left(1+\left(\frac{dw}{dr}\right)^2\right)^{3/2}}\approx-\frac{d^2w}{dr^2}, \tag{5}$$

$$\frac{1}{\rho_t}=\frac{\sin\theta}{r}\approx\frac{1}{r}tg\theta=-\frac{1}{r}\frac{dw}{dr}, \tag{6}$$

$$\sin\theta\approx\theta, \tag{7}$$

$$\cos\theta\approx 1, \tag{8}$$

where θ is the angle of rotation of a radial line element of the middle plane. Let ϵ_r, ϵ_t, ϵ_s, and ϵ_c be the maximum relative errors of the equations (5)-(8), respectively. Then we have

$$\epsilon_r=\max_r\left|\frac{1}{\left(1+\left(\frac{dw}{dr}\right)^2\right)^{3/2}}-1\right|=\left|\frac{1}{(1+(tg\theta_{\max})^2)^{3/2}}-1\right|, \tag{9}$$

$$\epsilon_t=\max_r\left|\frac{\sin\theta}{tg\theta}-1\right|=\left|\cos\theta_{\max}-1\right|=\epsilon_c \tag{10}$$

$$\epsilon_s=\max_r\left|\frac{\sin\theta}{\theta}-1\right|=\left|\frac{\sin\theta_{\max}}{\theta_{\max}}-1\right|. \tag{11}$$

From equations (9)-(11) we find that ϵ_r, $\epsilon_t(=\epsilon_c)$, and ϵ_s vary with $tg\theta_{\max}(=|dw/dr|_{\max})$, which their numeric values are listed in Table 1. In this table, we can see that ϵ_r is the greatest of these errors, next is $\epsilon_t(\epsilon_c)$, and the smallest is ϵ_s. Thus, it is feasible that ϵ_r may be taken as a standard of precision to criticize how the prerequisites are satisfied.

The Range of Applicability of the von Karman Theory

According to the equation (9), we can obtain

[14]Department of Mechanics, Lanzhou University, Lanzhou, Gansu 730000, People's Republic of China.

Manuscript received by the ASME Applied Mechanics Division, March 25, 1988; January 1, 1989.

Chandra, A., and Mukherjee, S., 1983, "Applications of Boundary Element Method to Large Strain Large Deformation Problems of Viscoplasticity," *Journal of Strain Analysis*, Vol. 8, No. 4, pp. 261-270.

Hibbit, H. D., Marcal, P. V., and Rice, J. R., 1970, "A Finite Element Formulation for Problems of Large Strain and Large Displacements," *International Journal of Solids and Structures*, Vol. 6, No. 8, pp. 1069-1086.

Okada, H., Rajiyah, H., ad Atluri, S. N., 1988, "Some Recent Developments in Finite Strain Elastoplasticity Using the Field-Boundary Element Method," *Computers and Structures*, Vol. 30, No. 1-2, pp. 275-288.

Rajiyah, H., and Mukherjee, S., 1987, "Boundary Element Analysis of Inelastic Axisymmetric Problems with Large Strains and Rotations," *International Journal of Solids and Structures*, Vol. 23, No. 12, pp. 1679-1698.

Sarihan, V., and Mukherjee, S., 1982, "Axisymmetric Viscoplastic Deformation by the Boundary Element Method," *International Journal of Solids and Structures*, Vol. 18, No. 12, pp. 1113-1128.

Yamada, Y., and Hirakawa, H., 1978, "Large Deformation and Instability Analysis in Metal Forming Processes," *Applications of Numerical Methods to Forming Process*, H. Armen, and R. G. Jones, eds., AMD-Vol. 28, ASME, New York, pp. 27-38.

On the Range of Applicability of von Karman Plate Equations

You-He Zhou[14] and Xiao-Jing Zheng[14]

Introduction

In the derivation of the von Karman plate equations (von Karman, 1910), as is known, there are two essential assumptions: one of which is the Kirchhoff's hypothesis and the other is the approximate calculations of the bending curvatures. The former requires that the thickness of a plate be small in comparison with its smallest lateral dimension. In most applications, the span-to-thickness ratio of a plate is greater than 15, so that the Kirchhoff's hypothesis may be valid. The latter requires that the plate slope, or rotations, should be less than 0.2. According to this requirement, many scholars (Timoshenko and Woinwsky-Krieoger, 1959; Chien et al., 1954; Chia, 1980) qualitatively consider that the ratio of deflection-to-thickness is of the same order with the unit, that is, $w/h \sim 0(1)$, which cannot explain why the solutions of the von Karman plate equations agree closely with the experiment data up to the deflection-to-thickness of six (Chia, 1980). Strictly speaking, the range should be decided by comparing the solutions of the von Karman theory with those of the more accurate Reissner plate equations (Reissner, 1958). However, it is difficult if not impossible to solve the Reissner equations for each problem of a thin plate. The best way is to check whether or not the prerequisites of the von Karman plate equations are satisfied by their solutions (see in Fig. 3). If the prerequisites are satisfied, we should be convinced of the feasibility of the von Karman theory.

It follows, in this paper, that a new quantitative criterion of the fitting range of the von Karman equations is to be derived from the analysis of errors caused by the approximate calculations based on the prerequisites, which shows that the von Karman theory in large deflection is reasonable. Finally, the range of the rigidly clamped, circular thin plate under uniform pressure will be given.

For the sake of simplicity, we discuss here the range of applicability of the axisymmetrical equations of the von Karman plate theory, which the dimensionless and independent differential equations can be written in the form (Zheng and Zhou, 1987; Zheng and Zhou, 1988):

$$y^2 \frac{d^2\varphi}{dy^2} = \varphi(y)S(y) + f(y), \tag{1}$$

$$y^2 \frac{d^2S}{dy^2} = -\frac{1}{2}(\varphi(y))^2, \quad y\in(0,1). \tag{2}$$

And the boundary conditions are

$$y=0:\ \varphi(y)=0;\ S(y)=0; \tag{3}$$

$$y=1:\ \varphi(y)=\frac{\lambda}{\lambda-1}\frac{d\varphi}{dy};\ S(y)=\frac{\mu}{\mu-1}\frac{dS}{dy}. \tag{4}$$

Here,

$$y=\left(\frac{r}{a}\right)^2;\ W=(3(1-\nu^2))^{1/2}\frac{w}{h};\ \varphi(y)=y\cdot\frac{dW}{dy};$$

$$S(y)=\frac{3(1-\nu^2)a^2N_r y}{Eh^3}$$

in which w is the deflection of a plate, r is the radial coordinate, h is the thickness of a plate, a is the radius of outer edge of a plate, N_r is the radial stress per unit length, E is the Young's modulus; ν is the Poisson's ratio, $f(y)$ is a function represented the distribution of axisymmetric loads, λ and μ are constants dependent on the boundary condition, e.g., for the rigidly-clamped edge, $\lambda=0$ and $\mu=2/1-\nu$.

Analysis of Errors

Since the Kirchhoff's hypothesis is held for most applications in engineering, thus, we merely discuss the errors caused by the approximate calculations imposed on the bending curvatures, etc. From the process of establishing equations (1)-(2), the approximate calculations, except for those from the Kirchhoff's hypothesis, are in the form

$$\frac{1}{\rho_r}=\frac{-\frac{d^2w}{dr^2}}{\left(1+\left(\frac{dw}{dr}\right)^2\right)^{3/2}}\approx-\frac{d^2w}{dr^2}, \tag{5}$$

$$\frac{1}{\rho_t}=\frac{\sin\theta}{r}\approx\frac{1}{r}tg\theta=-\frac{1}{r}\frac{dw}{dr}, \tag{6}$$

$$\sin\theta\approx\theta, \tag{7}$$

$$\cos\theta\approx 1, \tag{8}$$

where θ is the angle of rotation of a radial line element of the middle plane. Let ϵ_r, ϵ_t, ϵ_s, and ϵ_c be the maximum relative errors of the equations (5)-(8), respectively. Then we have

$$\epsilon_r=\max_r\left|\frac{1}{\left(1+\left(\frac{dw}{dr}\right)^2\right)^{3/2}}-1\right|=\left|\frac{1}{(1+(tg\theta_{\max})^2)^{3/2}}-1\right|, \tag{9}$$

$$\epsilon_t=\max_r\left|\frac{\sin\theta}{tg\theta}-1\right|=\left|\cos\theta_{\max}-1\right|=\epsilon_c \tag{10}$$

$$\epsilon_s=\max_r\left|\frac{\sin\theta}{\theta}-1\right|=\left|\frac{\sin\theta_{\max}}{\theta_{\max}}-1\right|. \tag{11}$$

From equations (9)-(11) we find that ϵ_r, $\epsilon_t(=\epsilon_c)$, and ϵ_s vary with $tg\theta_{\max}(=|dw/dr|_{\max})$, which their numeric values are listed in Table 1. In this table, we can see that ϵ_r is the greatest of these errors, next is $\epsilon_t(\epsilon_c)$, and the smallest is ϵ_s. Thus, it is feasible that ϵ_r may be taken as a standard of precision to criticize how the prerequisites are satisfied.

The Range of Applicability of the von Karman Theory

According to the equation (9), we can obtain

[14]Department of Mechanics, Lanzhou University, Lanzhou, Gansu 730000, People's Republic of China.

Manuscript received by the ASME Applied Mechanics Division, March 25, 1988; January 1, 1989.

Table 1 Relative errors versus $|dw/dr|_{max}$

$\left\|\frac{dw}{dr}\right\|_{max}$	$\epsilon_r(\%)$	$\epsilon_t(\%)$	$\epsilon_s(\%)$	$\left\|\frac{dw}{dr}\right\|_{max}$	$\epsilon_r(\%)$	$\epsilon_t(\%)$	$\epsilon_s(\%)$
0.0819	1.0	0.334	0.112	0.2052	6.0	2.041	0.681
0.1164	2.0	0.671	0.224	0.2226	7.0	2.389	0.798
0.1432	3.0	1.010	0.337	0.2391	8.0	2.741	0.915
0.1661	4.0	1.253	0.454	0.2547	9.0	3.095	1.034
0.1865	5.0	1.695	0.566	0.2698	10.0	3.397	1.153

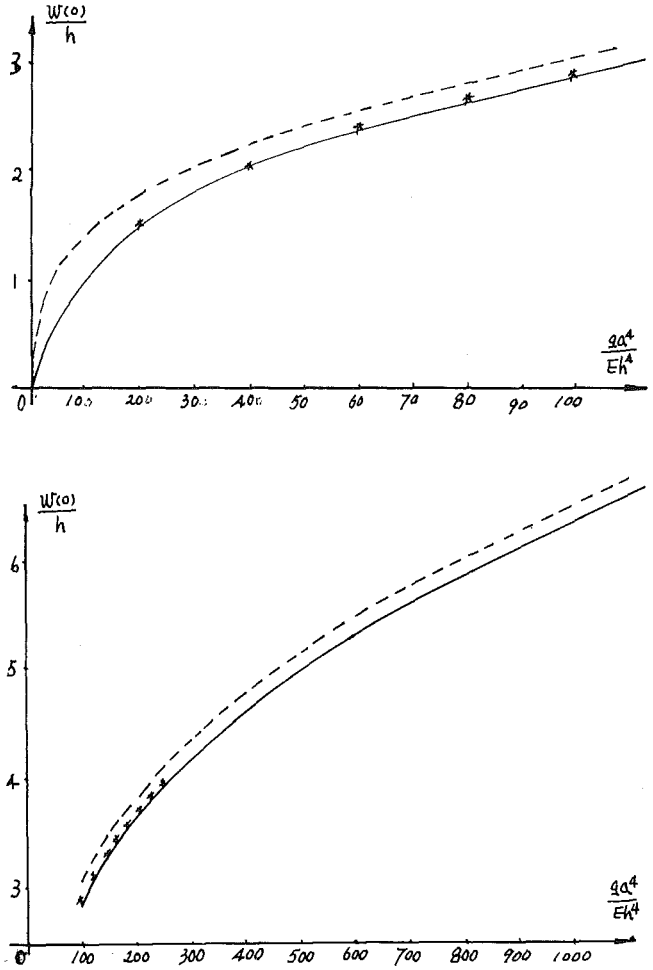


Fig. 1 The central deflection-load curves ($\nu = 0.3$); "_____": The von Karman theory of the rigidly-clamped circular plate under uniformly distributed pressure; "----": The Hencky's membrane theory (Chien et al., 1954); "*": The experiment data (McPherson et al., 1942)

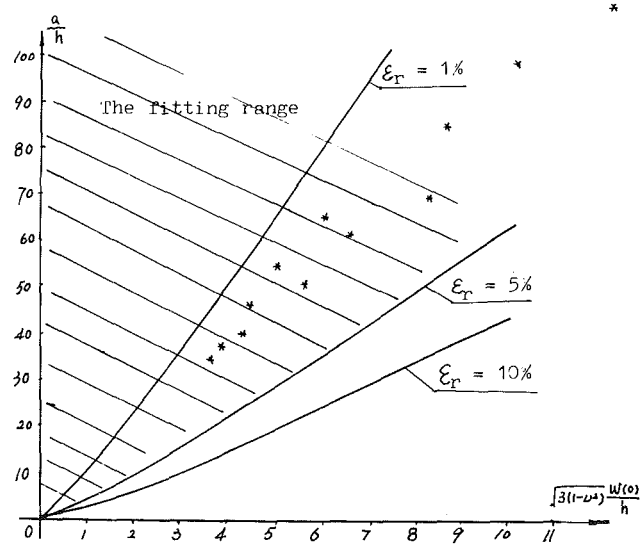


Fig. 2 The $\bar{W}_m \cdots a/h|_{min}$ curves and the fitting range of the von Karman equations of the rigidily-clamped circular plate under uniformly distributed pressure: $\nu = 0.3$; "*": The maximum central deflection of the circular plate for a pregiven *a/h* in the experiments (McPherson et al., 1942)

$$\left|\frac{dw}{dr}\right| \le \left|\frac{dw}{dr}\right|_{max} = \left|\frac{1}{(1-\epsilon_r)^{2/3}} - 1\right|^{1/2}. \tag{12}$$

Let $x = \frac{r}{a}$. Then we get

$$\frac{dW}{dx} = (3(1-\nu^2))^{1/2}\frac{a}{h}\frac{dw}{dr}, \tag{13}$$

that is,

$$\frac{dw}{dr} = (3(1-\nu^2))^{-1/2}\frac{h}{a}\frac{dW}{dx}. \tag{14}$$

Substituting the formula (14) into the inequality (12), we have

$$(3(1-\nu^2))^{-1/2}\frac{h}{a}\left|\frac{dW}{dx}\right| \le \left|\frac{1}{(1-\epsilon_r)^{2/3}} - 1\right|^{1/2}. \tag{15}$$

Further, we gain the following inequality

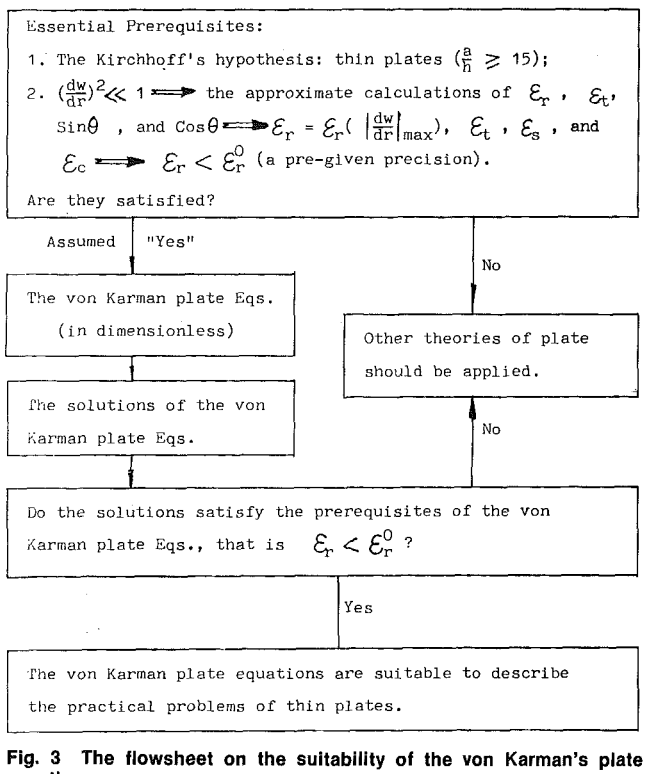


Fig. 3 The flowsheet on the suitability of the von Karman's plate equations

$$\frac{a}{h} \geq \left(\frac{(1-\epsilon_r)^{2/3}}{3(1-\nu^2)(1-(1-\epsilon_r)^{2/3})}\right)^{1/2} \left|\frac{dW}{dx}\right| \tag{16}$$

for all values of dW/dx in $x \in [0,1]$. Consequently, the inequality (16) also hold for $|dW/dx|_{max} (= \max_{x\in[0,1]} |dW/dx|)$, that is the inequality

$$\frac{a}{h} \geq \left(\frac{(1-\epsilon_r)^{2/3}}{3(1-\nu^2)(1-(1-\epsilon_r)^{2/3})}\right)^{1/2} \left|\frac{dW}{dx}\right|_{max} = \frac{a}{h}\Big|_{min} \tag{17}$$

hold, in which dW/dx is decided by the formula

$$\frac{dW}{dx} = 2y^{-1/2}\varphi(y)|_{y=x^2}. \tag{18}$$

Here, $\varphi(y)$ is obtained by solving the boundary value problem (1)–(4). On account of the prerequisites of the Kirchhoff's hypothesis, we have

$$\frac{a}{h} \geq \max\left(\frac{a}{h}\Big|_{min}, 15\right). \tag{19}$$

The inequalities (17) and (19) make us know that the ratio of radius-to-thickness of a plate must satisfy them in order to use the von Karman plate theory to describe practical problems under a pregiven precision, ϵ_r, which give a quantitative criterion on the fitting range of application of the von Karman plate equations.

Example of Applications

Here, we give the range of the von Karman theory of the rigidly-clamped circular plate under uniform pressure. The solutions of this problem have been, in detail, obtained by means of the analytically interpolated-iterative method (Zheng and Zhou, 1988), where the maximum relative errors between the nth and the $(n+1)$th iterative solutions are accurate to 1 percent. The central deflection-load curve and the experimental data (McPherson et al., 1942) are shown in Fig. 1. According to the obtained values of $\varphi(y)$ and $\bar{W}_m (= w(0)/h)$, we have made a drawing of $\bar{W}_m$ - - - $a/h|_{min}$ curve in Fig. 2 in which the asterisks "*" represent the maximum central deflection of the circular plate for a pregiven a/h in the experiments. It is obvious that the extents of the experiments are all in the fitting range for $\epsilon_r = 5$ percent. Hence, the von Karman plate equations may be applied to describe the elastic deformation of the circular plates in these experiments.

Conclusions

In order to apply the von Karman plate theory to designing the structures of thin plates, both the values of a/h and the ratio of deflection-to-thickness of a plate must be in the fitting range of themselves; that is, the inequalities (17) and (19) have to be satisfied. In the meantime, there is no difficulty in generalizing the idea presented in this paper to other cases of thin plates.

References

Chia, C. Y., 1980, *Nonlinear Analysis of Plates*, McGraw-Hill Book Co., New York.

Chien, W. Z., Lin, H. S., Hu, H. C., and Yeh, K. Y., 1954, *Large Deflection Problems of Elastic Thin Circular Plates*, Science Press, Beijing, (in Chinese).

McPherson, A. E., Ramberg, W., and Levy, S., 1942, "Normer Pressure Tests of Circular Plates with Clamped Edges," NACA Report No. 744.

Reissner, E., 1958, "Rotationally Symmetric Problems in the Theory of Thin Elastic Shells," *Proc. 3rd. U.S. Nat. Congr. Appl. Mech.*, pp. 51–68.

Timoshenko, S., and Woinwsky-Krieoger, S., 1959, *Theory of Plates and Shells*, McGraw-Hill Book Co., New York.

von Karman, T., 1910, "Festigketis Problem in Maschinbau," *Enzyklopadie der Mathematischen Wissenschaften*, Vol. IV, p. 348.

Zheng, X. J., and Zhou, Y. H., 1987, "Exact Solution to Large Deflection of Circular Plate under Compound Loads," *Scinetia Sinica* (Ser. A), Vol. 30, pp. 391–404.

Zheng, X. J., and Zhou, Y. H., 1988, "Analytical Computerization Method of Solving a Kind of Nonlinear Equation of Plates and Shells," *Computers and Structures*, (in press).

On the Corner Singularities in Reissner's Theory for the Bending of Elastic Plates

D. H. Y. Yen[15,17] and Mingru Zhou[16,17,18]

Introduction

In a recent work Burton and Sinclair (1986) studied the singularities at corner points along an edge in Reissner's theory for the bending of wedge-shaped plates. These authors began their singularity analysis by observing that the governing equations for the plate deflection, w, and the stress potential, χ, have asymptotic solutions of the form $w = 0(r^{\lambda+1})$, $\chi = 0(r^{\lambda+1})$ as $r \to 0$, where r is the distance from a corner point and λ is an eigenvalue parameter to be determined from the plate boundary conditions. The corner singularity is described by λ.

As we shall next illustrate, the asymptotic solutions assumed above are incomplete. By beginning with the same $w = 0(r^{\lambda+1})$-type solutions one derives $\chi = 0(r^{\lambda-1})$-type solutions instead as $r \to 0$. When homogeneous boundary conditions are imposed along the straight edges of the plate, however, terms resulting from the $0(r^{\lambda-1})$ terms in χ drop out in all but the simply-supported cases, leaving those resulting from the $0(r^{\lambda+1})$ terms in χ as the dominant ones. Thus most of the results by Burton and Sinclair (1986), i.e., with the exception when two simply-supported edges are involved, remain correct. The purpose of this note is to present details that will clarify the analysis and justify and supplement the results by Burton and Sinclair.

The issue here is one of completeness of the eigensolutions under various boundary conditions. Whereas we have demonstrated that the analysis of Burton and Sinclair (1986) to be incomplete, there remains the question as to whether the analysis presented here is complete. This latter question will be addressed later in this note.

Governing Equations and Asymptotic Solutions

We use the same notations as those in Burton and Sinclair (1986) and consider the wedge defined by $0 < r < \infty$ and $-\alpha/2 < \theta < \alpha/2$. The governing equations in Reissner's plate theory reduce to the following Cauchy-Riemann equations for the conjugate harmonic functions $\chi - \gamma\nabla^2\chi$ and $D\nabla^2 w$

$$\frac{\partial}{\partial r}(\chi - \gamma\nabla^2\chi) = \frac{1}{r}\frac{\partial}{\partial\theta}(D\nabla^2 w) \tag{1}$$

$$\frac{1}{r}\frac{\partial}{\partial\theta}(\chi - \gamma\nabla^2\chi) = -\frac{\partial}{\partial r}(D\nabla^2 w) \tag{2}$$

where $\gamma = h^2/10$. As in Burton and Sinclair (1988), the biharmonic function w is taken as

$$w = r^{\lambda+}[a_1\cos(\lambda+1)\theta + a_2\sin(\lambda+1)\theta + a_3\cos(\lambda-1)\theta + a_4\sin(\lambda-1)\theta] + 0(r^{\lambda+3}) \tag{3}$$

where the arbitrary constants are now named differently.

With w given in (3) we can easily verify that $D\nabla^2 w$ has the conjugate harmonic function

$$\psi = 4D\lambda r^{\lambda-1}[a_3\sin(\lambda-1)\theta - a_4\cos(\lambda-1)\theta], \tag{4}$$

[15]Professor, Mem. ASME.

[16]Visiting Associate Professor.

[17]Department of Mathematics, Michigan State University, East Lansing, Mich. 48824.

[18]Permanent address: Department of Mathematics, Xuzhou Teachers' College, Xuzhou, Jiangsu, People's Republic of China.

Manuscript received by the ASME Applied Mechanics Division, June 23, 1988; final revision, January 3, 1989.

$$\frac{a}{h} \geq \left(\frac{(1-\epsilon_r)^{2/3}}{3(1-\nu^2)(1-(1-\epsilon_r)^{2/3})}\right)^{1/2} \left|\frac{dW}{dx}\right| \tag{16}$$

for all values of dW/dx in $x \in [0,1]$. Consequently, the inequality (16) also hold for $|dW/dx|_{max} (= \max_{x\in[0,1]} |dW/dx|)$, that is the inequality

$$\frac{a}{h} \geq \left(\frac{(1-\epsilon_r)^{2/3}}{3(1-\nu^2)(1-(1-\epsilon_r)^{2/3})}\right)^{1/2} \left|\frac{dW}{dx}\right|_{max} = \frac{a}{h}\Big|_{min} \tag{17}$$

hold, in which dW/dx is decided by the formula

$$\frac{dW}{dx} = 2y^{-1/2}\varphi(y)|_{y=x^2}. \tag{18}$$

Here, $\varphi(y)$ is obtained by solving the boundary value problem (1)–(4). On account of the prerequisites of the Kirchhoff's hypothesis, we have

$$\frac{a}{h} \geq \max\left(\frac{a}{h}\Big|_{min}, 15\right). \tag{19}$$

The inequalities (17) and (19) make us know that the ratio of radius-to-thickness of a plate must satisfy them in order to use the von Karman plate theory to describe practical problems under a pregiven precision, ϵ_r, which give a quantitative criterion on the fitting range of application of the von Karman plate equations.

Example of Applications

Here, we give the range of the von Karman theory of the rigidly-clamped circular plate under uniform pressure. The solutions of this problem have been, in detail, obtained by means of the analytically interpolated-iterative method (Zheng and Zhou, 1988), where the maximum relative errors between the nth and the $(n+1)$th iterative solutions are accurate to 1 percent. The central deflection-load curve and the experimental data (McPherson et al., 1942) are shown in Fig. 1. According to the obtained values of $\varphi(y)$ and $\bar{W}_m (= w(0)/h)$, we have made a drawing of $\bar{W}_m$ - - - $a/h|_{min}$ curve in Fig. 2 in which the asterisks "*" represent the maximum central deflection of the circular plate for a pregiven a/h in the experiments. It is obvious that the extents of the experiments are all in the fitting range for $\epsilon_r = 5$ percent. Hence, the von Karman plate equations may be applied to describe the elastic deformation of the circular plates in these experiments.

Conclusions

In order to apply the von Karman plate theory to designing the structures of thin plates, both the values of a/h and the ratio of deflection-to-thickness of a plate must be in the fitting range of themselves; that is, the inequalities (17) and (19) have to be satisfied. In the meantime, there is no difficulty in generalizing the idea presented in this paper to other cases of thin plates.

References

Chia, C. Y., 1980, *Nonlinear Analysis of Plates*, McGraw-Hill Book Co., New York.

Chien, W. Z., Lin, H. S., Hu, H. C., and Yeh, K. Y., 1954, *Large Deflection Problems of Elastic Thin Circular Plates*, Science Press, Beijing, (in Chinese).

McPherson, A. E., Ramberg, W., and Levy, S., 1942, "Normer Pressure Tests of Circular Plates with Clamped Edges," NACA Report No. 744.

Reissner, E., 1958, "Rotationally Symmetric Problems in the Theory of Thin Elastic Shells," *Proc. 3rd. U.S. Nat. Congr. Appl. Mech.*, pp. 51–68.

Timoshenko, S., and Woinwsky-Krieoger, S., 1959, *Theory of Plates and Shells*, McGraw-Hill Book Co., New York.

von Karman, T., 1910, "Festigketis Problem in Maschinbau," *Enzyklopadie der Mathematischen Wissenschaften*, Vol. IV, p. 348.

Zheng, X. J., and Zhou, Y. H., 1987, "Exact Solution to Large Deflection of Circular Plate under Compound Loads," *Scinetia Sinica* (Ser. A), Vol. 30, pp. 391–404.

Zheng, X. J., and Zhou, Y. H., 1988, "Analytical Computerization Method of Solving a Kind of Nonlinear Equation of Plates and Shells," *Computers and Structures*, (in press).

On the Corner Singularities in Reissner's Theory for the Bending of Elastic Plates

D. H. Y. Yen[15,17] and Mingru Zhou[16,17,18]

Introduction

In a recent work Burton and Sinclair (1986) studied the singularities at corner points along an edge in Reissner's theory for the bending of wedge-shaped plates. These authors began their singularity analysis by observing that the governing equations for the plate deflection, w, and the stress potential, χ, have asymptotic solutions of the form $w = 0(r^{\lambda+1})$, $\chi = 0(r^{\lambda+1})$ as $r \to 0$, where r is the distance from a corner point and λ is an eigenvalue parameter to be determined from the plate boundary conditions. The corner singularity is described by λ.

As we shall next illustrate, the asymptotic solutions assumed above are incomplete. By beginning with the same $w = 0(r^{\lambda+1})$-type solutions one derives $\chi = 0(r^{\lambda-1})$-type solutions instead as $r \to 0$. When homogeneous boundary conditions are imposed along the straight edges of the plate, however, terms resulting from the $0(r^{\lambda-1})$ terms in χ drop out in all but the simply-supported cases, leaving those resulting from the $0(r^{\lambda+1})$ terms in χ as the dominant ones. Thus most of the results by Burton and Sinclair (1986), i.e., with the exception when two simply-supported edges are involved, remain correct. The purpose of this note is to present details that will clarify the analysis and justify and supplement the results by Burton and Sinclair.

The issue here is one of completeness of the eigensolutions under various boundary conditions. Whereas we have demonstrated that the analysis of Burton and Sinclair (1986) to be incomplete, there remains the question as to whether the analysis presented here is complete. This latter question will be addressed later in this note.

Governing Equations and Asymptotic Solutions

We use the same notations as those in Burton and Sinclair (1986) and consider the wedge defined by $0 < r < \infty$ and $-\alpha/2 < \theta < \alpha/2$. The governing equations in Reissner's plate theory reduce to the following Cauchy-Riemann equations for the conjugate harmonic functions $\chi - \gamma\nabla^2\chi$ and $D\nabla^2 w$

$$\frac{\partial}{\partial r}(\chi - \gamma\nabla^2\chi) = \frac{1}{r}\frac{\partial}{\partial\theta}(D\nabla^2 w) \tag{1}$$

$$\frac{1}{r}\frac{\partial}{\partial\theta}(\chi - \gamma\nabla^2\chi) = -\frac{\partial}{\partial r}(D\nabla^2 w) \tag{2}$$

where $\gamma = h^2/10$. As in Burton and Sinclair (1988), the biharmonic function w is taken as

$$w = r^{\lambda+}[a_1\cos(\lambda+1)\theta + a_2\sin(\lambda+1)\theta + a_3\cos(\lambda-1)\theta + a_4\sin(\lambda-1)\theta] + 0(r^{\lambda+3}) \tag{3}$$

where the arbitrary constants are now named differently.

With w given in (3) we can easily verify that $D\nabla^2 w$ has the conjugate harmonic function

$$\psi = 4D\lambda r^{\lambda-1}[a_3\sin(\lambda-1)\theta - a_4\cos(\lambda-1)\theta], \tag{4}$$

[15]Professor, Mem. ASME.

[16]Visiting Associate Professor.

[17]Department of Mathematics, Michigan State University, East Lansing, Mich. 48824.

[18]Permanent address: Department of Mathematics, Xuzhou Teachers' College, Xuzhou, Jiangsu, People's Republic of China.

Manuscript received by the ASME Applied Mechanics Division, June 23, 1988; final revision, January 3, 1989.

and thus

$$\chi - \gamma\nabla^2\chi = -\psi \,. \tag{5}$$

The solution for χ consists of a particular solution χ_p and a general solution χ_c of the homogeneous equation. An obvious choice for χ_p, since χ is harmonic, is

$$\chi_p = -\psi \,. \tag{6}$$

The homogeneous equation for χ_c

$$\chi_c - \gamma\nabla^2\chi_c = 0 \tag{7}$$

has the solutions

$$\chi_c = \{I_\nu(ar), K_\nu(ar)\}\{\sin\nu\theta, \cos\nu\theta\} \,, \tag{8}$$

where $a = \gamma^{-1/2} = 10^{1/2}/h$, ν is a complex number with non-negative real part, and I_ν and K_ν denote modified Bessel functions of order ν. General solutions of (7) can be constructed in the form of integrals, with respect to ν, involving the solutions in (8) multiplied by arbitrary functions of ν. However, as only specific values of ν must be chosen, such general integrals are not needed.

For χ_c to remain bounded as $r \to \infty$ we must discard the $I_\nu(ar)$ terms in (8). For small r we have

$$K_\nu(ar) \sim \left(\frac{1}{2}\right)^{-\nu+1}\Gamma(\nu)\,(ar)^{-\nu} - \left(\frac{1}{2}\right)^{-\nu+3}\Gamma(\nu-1)(ar)^{-\nu+2} \tag{9}$$

where Γ denotes the Gamma function. For the leading term in (9) to match that in $\chi_p = -\psi$ for small r, we must take

$$-\nu = \lambda - 1 \,. \tag{10}$$

Combining the results of (6) through (10) we now have

$$\chi = \left(\frac{1}{2}\right)^{\lambda}\Gamma(1-\lambda)a^{\lambda-1}r^{\lambda-1}\,[-a_5\sin(\lambda-1)\theta + a_6\cos(\lambda-1)\theta] + 4D\lambda r^{\lambda-1}[-a_3\sin(\lambda-1)\theta + a_4\cos(\lambda-1)\theta] + 0(r^{\lambda+1}). \tag{11}$$

Thus, with w being given in (3), the asymptotic solution for χ for small r begins with terms of the order $0(r^{\lambda-1})$.

Eigensolutions for Various Boundary Conditions

The following boundary conditions are considered:

Clamped:	$\beta_r = \beta_\theta = 0,\ w = 0$	
Free:	$M_\theta = M_{r\theta} = 0,\ V_\theta = 0$	(12)
Simply-supported:	$M_\theta = 0,\ \beta_r = 0,\ w = 0$	

where β_r and β_θ are the plate rotations in the r- and θ-directions respectively, M_θ is the bending moment in the θ-direction, $M_{r\theta}$ is the twisting moment, and V_θ is the shear force on a θ-face. When these quantities are expressed in terms of w and χ, contributions from the leading terms in χ may cancel out when the homogeneous boundary conditions on $\theta = \pm\alpha/2$ are imposed as we mentioned before. To prepare for this we shall include the next higher-order terms in the solutions for w and χ given below

$$w = r^{\lambda+1}\,[a_1\cos(\lambda+1)\theta + a_2\sin(\lambda+1)\theta + a_3\cos(\lambda-1)\theta + a_4\sin(\lambda-1)\theta] + r^{\lambda+3}[b_1\cos(\lambda+3)\theta + b_2\sin(\lambda+3)\theta + b_3\cos(\lambda+1)\theta + b_4\sin(\lambda+1)\theta] \tag{13}$$

$$\chi = r^{\lambda-1}[-(a_3'+a_5')\sin(\lambda-1)\theta + (a_4'+a_6')\cos(\lambda-1)\theta] + r^{\lambda+1}[-(b_3'+b_5')\sin(\lambda+1)\theta + (b_4'+b_6')\cos(\lambda+1)\theta - a_5''\sin(\lambda-1)\theta + a_6''\cos(\lambda-1)\theta] \tag{14}$$

where

$$a_i' = 4D\lambda a_i,\ a_j' = \Gamma(1-\lambda)a^{\lambda-1}a_j \,,$$
$$a_j'' = -\left(\frac{1}{2}\right)^{\lambda+2}\Gamma(-\lambda)a^{\lambda+1}a_j \,,$$
$$b_i' = 4D(\lambda+2)b_i \,,\ b_j' = \left(\frac{1}{2}\right)^{\lambda+2}\Gamma(-\lambda-1)a^{\lambda+1}b_j \,,$$
$$i = 3,4, \qquad j = 5,6 \,. \tag{15}$$

In terms of w and χ, we now have

$$\begin{aligned}\beta_r &= \frac{2\gamma}{D(1-\mu)}\frac{1}{r}\frac{\partial\chi}{\partial\theta} - \frac{\partial w}{\partial r}\\ &= -r^{\lambda-2}\,c(\lambda-1)[(a_3'+a_5')\cos(\lambda-1)\theta + (a_4'+a_6')\sin(\lambda-1)\theta]\\ &\quad - r^{\lambda}\{(\lambda+1)[c(b_3'+b_5') + a_1]\cos(\lambda+1)\theta + (\lambda+1)[c(b_4'+b_6') + a_2]\sin(\lambda+1)\theta\\ &\quad + [c(\lambda-1)a_5'' + (\lambda+1)a_3]\cos(\lambda-1)\theta + [c(\lambda-1)a_6'' + (\lambda+1)a_4]\sin(\lambda-1)\theta\}\end{aligned} \tag{16}$$

$$\begin{aligned}\beta_\theta &= -\frac{2\gamma}{D(1-\mu)}\frac{\partial\chi}{\partial r} - \frac{1}{r}\frac{\partial w}{\partial\theta}\\ &= r^{\lambda-2}c(\lambda-1)[a_3'+a_5')\sin(\lambda-1)\theta - (a_4'+a_6')\cos(\lambda-1)\theta]\\ &\quad + r^{\lambda}\{(\lambda+1)[c(b_3'+b_5') + a_1]\sin(\lambda+1)\theta - (\lambda+1)[c(b_4'+b_6') + a_2]\cos(\lambda+1)\theta\\ &\quad + [c(\lambda+1)a_5'' + (\lambda-1)a_3]\sin(\lambda-1)\theta - [c(\lambda+1)a_6'' + (\lambda-1)a_4]\cos(\lambda-1)\theta\}\end{aligned} \tag{17}$$

$$\begin{aligned}M_\theta &= 2\gamma\left(\frac{1}{r^2}\frac{\partial\chi}{\partial\theta} - \frac{1}{r}\frac{\partial^2\chi}{\partial r\partial\theta}\right) - D\left(\mu\frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{1}{r^2}\frac{\partial^2 w}{\partial\theta^2}\right)\\ &= r^{\lambda-3}Dc(\lambda-1)(\lambda-2)(1-\mu)[(a_3' + a_5')\cos(\lambda-1)\theta + (a_4' + a_6')\sin(\lambda-1)\theta]\\ &\quad + r^{\lambda-1}D\lambda\{(\lambda+1)(1-\mu)[c(b_3' + b_5') + a_1]\cos(\lambda+1)\theta + (\lambda+1)(1-\mu)[c(b_4' + b_6') + a_2]\sin(\lambda+1)\theta\\ &\quad + [c(\lambda-1)(1-\mu)a_5'' + (\lambda-3-\lambda\mu-\mu)a_3]\cos(\lambda-1) + [c(\lambda-1)(1-\mu)a_6'' + (\lambda-3-\lambda\mu-\mu)a_4]\sin(\lambda-1)\theta\}\end{aligned} \tag{18}$$

$$\begin{aligned}M_{r\theta} &= \gamma\left(\frac{1}{r^2}\frac{\partial^2\chi}{\partial\theta^2} - r\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\chi}{\partial r}\right)\right) - (1-\mu)D\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial w}{\partial\theta}\right)\\ &= r^{\lambda-3}Dc(\lambda-1)(\lambda-2)(1-\mu)[(a_3'+a_5')\sin(\lambda-1)\theta - (a_4'+a_6')\cos(\lambda-1)\theta]\\ &\quad + r^{\lambda-1}D\lambda(1-\mu)\{(\lambda+1)[c(b_3'+b_5')+a_1]\sin(\lambda+1)\theta - (\lambda+1)[c(b_4'+b_6')+a_2]\cos(\lambda+1)\theta\\ &\quad + (\lambda-1)(ca_5''+a_3)\sin(\lambda-1)\theta + (\lambda-1)(ca_6''+a_4)\cos(\lambda-1)\theta\}\end{aligned} \tag{19}$$

$$\begin{aligned}V_\theta &= -\frac{\partial\chi}{\partial\theta}\\ &= r^{\lambda-2}(\lambda-1)[(a_3'+a_5')\sin(\lambda-1)\theta - (a_4'+a_6')\cos(\lambda-1)\theta]\end{aligned}$$

$$+ r^{\lambda}(\lambda+1)[(b_3'+b_5')\sin(\lambda+1)\theta - (b_4'+b_6')\cos(\lambda+1)\theta + a_5'' \sin(\lambda-1)\theta - a_6'' \cos(\lambda-1)\theta] , \quad (20)$$

where

$$c = \frac{12(1+\mu)}{5Eh} = \frac{2}{Da^2(1-\mu)} . \quad (21)$$

We observe that the dominant terms in (16) through (20) as well as in (14) for χ depend on the constants a_3', a_4', a_5', and a_6', only through the combinations $a_3' + a_5'$ and $a_4' + a_6'$. At each of the edges $\theta = \pm \alpha/2$, we have at least two boundary conditions not involving w. If such non-w boundary conditions from both edges lead to two or more linearly-independent homogeneous equations for $a_3' + a_5'$ and $a_4' + a_6'$ for any λ, then we must have $a_3' + a_5' = 0$ and $a_4' + a_6' = 0$. The dominant terms in χ become $0(r^{\lambda+1})$ and those in (16) through (20) all increase their order by two. One can then proceed as in Burton and Sinclair (1986). It turns out that this is the situation in all cases of boundary conditions except the one in which both edges are simply-supported. Thus, the first five entries in Table 1 of Burton and Sinclair (1986) are correct.

The case of a plate with two simply-supported edges at $\theta = \pm \alpha/2$ will now be examined.

We apply the conditions $w = 0$, $\beta_r = 0$ and $M_\theta = 0$ at $\theta = \pm \alpha/2$. From the leading terms in (13), (16), and (18) we obtain

$$a_1 \cos(\lambda+1)\frac{\alpha}{2} + a_2 \sin(\lambda+1)\frac{\alpha}{2} + a_3 \cos(\lambda-1)\frac{\alpha}{2} + a_4 \sin(\lambda-1)\frac{\alpha}{2} = 0$$

$$a_1 \cos(\lambda+1)\frac{\alpha}{2} - a_2 \sin(\lambda+1)\frac{\alpha}{2} + a_3 \cos(\lambda-1)\frac{\alpha}{2} - a_4 \sin(\lambda-1)\frac{\alpha}{2} = 0$$

$$(a_3'+a_5')\cos(\lambda-1)\frac{\alpha}{2} + (a_4'+a_6')\sin(\lambda-1)\frac{\alpha}{2} = 0$$

$$(a_3'+a_5')\cos(\lambda-1)\frac{\alpha}{2} - (a_4'+a_6')\sin(\lambda-1)\frac{\alpha}{2} = 0 \quad (22)$$

$$(a_3'+a_5')\cos(\lambda-1)\frac{\alpha}{2} + (a_4'+a_6')\sin(\lambda-1)\frac{\alpha}{2} = 0$$

$$(a_3'+a_5')\cos(\lambda-1)\frac{\alpha}{2} - (a_4'+a_6')\sin(\lambda-1)\frac{\alpha}{2} = 0$$

The third member in (22) is identical with the fifth, and the fourth member is identical with the sixth. If $a_3' + a_5' = 0$, $a_4' + a_6' = 0$, we may proceed as before and obtain the same sixth entry in Table 1 of Burton and Sinclair (1986), i.e.,

$$\cos \lambda\alpha = \pm \cos \alpha . \quad (23)$$

On the other hand, if $a_3' + a_5' \neq 0$ we obtain from the third and fourth (or the fifth and sixth) members of (22) the eigenequation

$$\cos(\lambda-1)\frac{\alpha}{2} = 0 \quad (24)$$

which has the solutions for λ

$$\lambda = \frac{(2k+1)\pi}{\alpha} + 1, \qquad k = \text{integer} \quad (25)$$

along with the eigensolutions

$$w = r^{\lambda+1}\left[\frac{-\sin(\lambda-1)\frac{\alpha}{2}}{\sin(\lambda+1)\frac{\alpha}{2}}\sin(\lambda+1)\theta + \sin(\lambda-1)\theta\right] a_4 + a_3 \cos(\lambda-1)\theta + 0(r^{\lambda+3})$$

$$\beta_r = r^{\lambda-2}c(1-\lambda)(a_3' + a_5')\cos(\lambda-1)\theta + 0(r^{\lambda}) \quad (26)$$

$$M_\theta = r^{\lambda-3} Dc(\lambda-1)(\lambda-2)(1-\mu)(a_3' + a_5')\cos(\lambda-1)\theta + 0(r^{\lambda-1}) .$$

Since $\lambda \geq -1$ for $0 \leq \alpha \leq \pi$ for w to be bounded, we take $k = 0$ in (25) and obtain

$$\lambda_{\text{critical}} = \frac{\pi}{\alpha} + 1. \quad (27)$$

Substituting this $\lambda_{\text{critical}}$ into (26) yields singularities of the solutions that are not covered by Burton and Sinclair (1986).

Concluding Remarks

By assuming the same biharmonic function for w as in Burton and Sinclair (1986) given in (3), we obtained the asymptotic expansion (11) for χ via its general solution. The leading terms are of the order $0(r^{\lambda-1})$. When homogeneous boundary conditions of the type given in (12) are imposed along $\theta = \pm \alpha/2$, contributions from the $0(r^{\lambda-1})$ terms in χ drop out except in the case with two simply-supported edges, leaving those from the $0(r^{\lambda+1})$ terms in χ as the dominant ones as assumed by Burton and Sinclair. The eigensolutions given in (26) and (27) thus supplement the results of Burton and Sinclair. Numerical solutions of the transcendental equations for λ have been discussed in Burton and Sinclair (1986) and the literature referred to therein.

The analysis given here suggests that the eigensolutions here are complete within the class of biharmonic functions for w, as assumed in (3), as we have examined the general solutions for χ governed by (1) and (2) with this assumed w. On the other hand, it is known that biharmonic functions other than those given in (3) for w are possible (for example, see Dempsey and Sinclair (1979)) that contain logarithmic functions. Particular solutions χ_p generated by such w must then be combined with the complementary solutions χ_c given in (8) to satisfy the homogeneous edge conditions. It is not clear whether new eigensolutions may result and thus the general question of completeness of the eigenfunctions obtained here remains open.

The singularities considered here describe the behavior of solutions at corner points along the boundary of general plate bending problems. One may imagine that a small portion of a plate near a corner point has been isolated. Transverse load is neglected but the portion is under the action of shear forces and bending and twisting moments along the cut boundary. For asymptotic solutions for small r near the corner point the portion is treated as a wedge. A knowledge of such singularity solutions is useful in obtaining accurate solutions near a corner point by numerical methods.

A more comprehensive treatment of the corner singularity problem in Mindlin's formulation has also been given by us and will be reported elsewhere (Yen and Zhou, 1989). The problem has been formulated so that a small parameter ϵ is involved that is related to the shear deformation. By letting $\epsilon \to 0$ the more complete solutions obtained there enable us to analyze and explain the difference in the corner singularities between the classical plate theory and the Reissner-Mindlin plate theory.

References

Burton, W. S., and Sinclair, G. B., 1986, "On the Singularities in Reissner's Theory for the Bending of Elastic Plates," ASME Journal of Applied Mechanics, Vol. 53, pp. 220-222.

Dempsey, J. P., and Sinclair, A. B., 1979, "On the Stress Singularities in the Plane Elasticity of the Composite Wedge," *Journal of Elasticity*, Vol. 9, pp. 373-391.

Yen, D. H. Y., and Zhou, M., 1988, "On the Singularity at Corner Points of Solutions of Plate Bending Problems," Preliminary Report, Department of Mathematics, Michigan State University, (submitted to the *Journal of Elasticity*).

New Integral Equation Approach for a Curved Rigid Line Problem in an Infinite Plate

Y. Z. Chen[19]

1 Introduction

There has been a resurgence of interest in the rigid line problem in plane elasticity. Many problems in this field have been investigated by several authors: Erdogan and Gupta (1972), Wang, Zhang, and Chou (1985), Chen (1986), and Chen and Cheung (1987). Like the crack problem in plane elasticity, the stress components in the vicinity of the rigid line tip also possess a singular character and can be expressed as (Chen, 1986):

$$\sigma_x=(2\pi r)^{-1/2}\left[S_1\left(\cos(\theta/2)-\frac{1}{3+\kappa}\sin(\theta)\sin(3\theta/2)\right)\right.$$
$$\left.-S_2\left(\frac{3-\kappa}{1-\kappa}\sin(\theta/2)+\frac{1}{1-\kappa}\sin(\theta)\cos(3\theta/2)\right)\right]+0(1),$$
$$\sigma_y=(2\pi r)^{-1/2}\left[S_1\left(\frac{1-\kappa}{3+\kappa}\cos(\theta/2)+\frac{1}{3+\kappa}\sin(\theta)\sin(3\theta/2)\right)\right.$$
$$\left.-S_2\left(\frac{1+\kappa}{1-\kappa}\sin(\theta/2)-\frac{1}{1-\kappa}\sin(\theta)\cos(3\theta/2)\right)\right]+0(1),$$
$$\sigma_{xy}=(2\pi r)^{-1/2}\left[S_1\left(\frac{1+\kappa}{3+\kappa}\sin(\theta/2)+\frac{1}{3+\kappa}\sin(\theta)\cos(3\theta/2)\right)\right.$$
$$\left.+S_2\left(\cos(\theta/2)-\frac{1}{1-\kappa}\sin(\theta)\sin(3\theta/2)\right)\right]+0(1), \quad (1)$$

where $\kappa=(3-\nu)/(1+\nu)$ (under plane-stress conditions), $\kappa=3-4\nu$ (under plane-strain conditions) and ν is the Poisson's ratio in elasticity. In equation (1), S_1 and S_2 are the so-called stress singularity coefficients and can be defined by

$$S_1=\lim_{r\to 0}(2\pi r)^{1/2}\sigma_{x,0},\quad S_2=\lim_{r\to 0}(2\pi r)^{1/2}\sigma_{xy,0} \quad (2)$$

where $\sigma_{y,0}$ and $\sigma_{xy,0}$ are the stresses in the front of rigid line tip. The aim of this note is to formulate a new integral equation with a logarithmic kernel for the curved rigid line problem in an infinite plate. In the formulation we take the traction difference between the upper and lower borders of the rigid line to be an unknown function, and the displacement to be the right-hand term of the integral equation. Therefore, the resulting integral equation has a logarithmic kernel.

2 New Integral Equation in the Curved Rigid Line Problem

It is well known that the resultant forces and displacement in plane elasticity can be expressed by (Muskhelishvili, 1953):

$$f=i(X+iY)=\varphi(z)+z\overline{\varphi'(z)}+\overline{\psi(z)}+C_1 \quad (3)$$

$$2G(u+iv)=\kappa\varphi(z)-z\overline{\varphi'(z)}-\overline{\psi(z)}+C_2 \quad (4)$$

where $\varphi(z)$ and $\psi(z)$ are the complex potential, $z=x+iy$ and G is the shear modululs of elasticity.

Now we want to solve the curved rigid line problem as shown in Fig. 1(*a*). Assume in the problem that the remote rotation will be $\epsilon^\infty=0$ and the remote stresses are expressed by σ_x^∞, σ_y^∞, and σ_{xy}^∞. If the curved rigid line is in a floating state, then the conditions along the curve rigid line L in Fig. 1(*a*) will be:

(*a*) There are no resultant forces applied on the whole rigid line.

(*b*) The rigid line has a rotation

$$2G(u+iv)=i\gamma t_o,\ (t_o\epsilon L) \quad (5)$$

where the rotation γ is determined from the zero-moment condition imposed on the whole rigid line.

Obviously, the problem shown in Fig. 1(*a*) can be separated into two problems as shown in Fig. 1(*b*) and Fig. 1(*c*).

The problem shown in Fig. 1(*b*) has a solution (Muskhelishvili, 1953):

$$2G(u+iv)=(\kappa-1)\Gamma z-\overline{\Gamma'}\bar{z}, \quad (6)$$

where

$$\Gamma=(\sigma_x^\infty+\sigma_y^\infty)/4,\quad \Gamma'=(\sigma_y^\infty-\sigma_x^\infty)/2+i\sigma_{xy}^\infty. \quad (7)$$

It is easy to see that the conditions imposed on the problem shown in Fig. 1(*c*) can be stated as follows:

(*a*) There are no resultant forces applied on the whole rigid line.

(*b*) The displacements along the curved rigid line should take the form

$$2G(u+iv)=-(\kappa-1)\Gamma t_o+\overline{\Gamma'}\bar{t}_o+i\gamma t_o,\ (t_o\epsilon L) \quad (8)$$

where the rotation γ is still determined from the zero-moment condition.

It is easy to verify that the appropriate potentials in the curved rigid line problem have the following form:

$$\varphi(z)=\int_0^t\mu(s)\log(z-t)ds,$$
$$\psi(z)=-\kappa\int_0^t\overline{\mu(s)}\log(z-t)ds-\int_0^t(\mu(s)\bar{t}/(z-t))ds. \quad (9)$$

Since there are no resultant forces applied on the whole rigid line L, from equations (3) and (9) we obtain

$$\int_0^t\mu(s)ds=0. \quad (10)$$

Let z approach the upper or lower border of the curved rigid line L in Fig. 1, i.e., $z\to t_o^+$ or $z\to t_o^-$. In either case, we have

$$2G(u(t_o)+iv(t_o))=\kappa\int_0^t 2\log|t-t_o|\mu(s)ds$$
$$-\int_0^t((t-t_o)/(\bar{t}-\bar{t}_o))\overline{\mu(s)}ds+C_2. \quad (11)$$

After using equations (8) and (11), and by letting

$$t-t_o=r(t,t_o)\exp(i\theta(t,t_o)),$$
$$\mu(s)=\mu_1(s)+i\mu_2(s), \quad (12)$$

the following system of the integral equations is obtainable

$$\kappa\int_0^t 2\log(r(t,t_o))\mu_1(s)ds+\int_0^t[-\mu_1(s)\cos(2\theta(t,t_o))$$
$$-\mu_2(s)\sin(2\theta(t,t_o))]ds+\mathrm{Re}(C_2)=H_1(t_o),\ (t_o\epsilon L) \quad (13a)$$

[19]Professor, Division of Engineering Mechanics, Jiangsu Institute of Technology, Zhengjian, Jiangsu, 212013, People's Republic of China.

Manuscript received by the ASME Applied Mechanics Division, December 4, 1987; final revision, March 1, 1989.

Dempsey, J. P., and Sinclair, A. B., 1979, "On the Stress Singularities in the Plane Elasticity of the Composite Wedge," *Journal of Elasticity*, Vol. 9, pp. 373-391.

Yen, D. H. Y., and Zhou, M., 1988, "On the Singularity at Corner Points of Solutions of Plate Bending Problems," Preliminary Report, Department of Mathematics, Michigan State University, (submitted to the *Journal of Elasticity*).

New Integral Equation Approach for a Curved Rigid Line Problem in an Infinite Plate

Y. Z. Chen[19]

1 Introduction

There has been a resurgence of interest in the rigid line problem in plane elasticity. Many problems in this field have been investigated by several authors: Erdogan and Gupta (1972), Wang, Zhang, and Chou (1985), Chen (1986), and Chen and Cheung (1987). Like the crack problem in plane elasticity, the stress components in the vicinity of the rigid line tip also possess a singular character and can be expressed as (Chen, 1986):

$$\sigma_x=(2\pi r)^{-1/2}\left[S_1\left(\cos(\theta/2)-\frac{1}{3+\kappa}\sin(\theta)\sin(3\theta/2)\right)\right.$$
$$\left.-S_2\left(\frac{3-\kappa}{1-\kappa}\sin(\theta/2)+\frac{1}{1-\kappa}\sin(\theta)\cos(3\theta/2)\right)\right]+0(1),$$
$$\sigma_y=(2\pi r)^{-1/2}\left[S_1\left(\frac{1-\kappa}{3+\kappa}\cos(\theta/2)+\frac{1}{3+\kappa}\sin(\theta)\sin(3\theta/2)\right)\right.$$
$$\left.-S_2\left(\frac{1+\kappa}{1-\kappa}\sin(\theta/2)-\frac{1}{1-\kappa}\sin(\theta)\cos(3\theta/2)\right)\right]+0(1),$$
$$\sigma_{xy}=(2\pi r)^{-1/2}\left[S_1\left(\frac{1+\kappa}{3+\kappa}\sin(\theta/2)+\frac{1}{3+\kappa}\sin(\theta)\cos(3\theta/2)\right)\right.$$
$$\left.+S_2\left(\cos(\theta/2)-\frac{1}{1-\kappa}\sin(\theta)\sin(3\theta/2)\right)\right]+0(1), \quad (1)$$

where $\kappa=(3-\nu)/(1+\nu)$ (under plane-stress conditions), $\kappa=3-4\nu$ (under plane-strain conditions) and ν is the Poisson's ratio in elasticity. In equation (1), S_1 and S_2 are the so-called stress singularity coefficients and can be defined by

$$S_1=\lim_{r\to 0}(2\pi r)^{1/2}\sigma_{x,0},\quad S_2=\lim_{r\to 0}(2\pi r)^{1/2}\sigma_{xy,0} \quad (2)$$

where $\sigma_{y,0}$ and $\sigma_{xy,0}$ are the stresses in the front of rigid line tip. The aim of this note is to formulate a new integral equation with a logarithmic kernel for the curved rigid line problem in an infinite plate. In the formulation we take the traction difference between the upper and lower borders of the rigid line to be an unknown function, and the displacement to be the right-hand term of the integral equation. Therefore, the resulting integral equation has a logarithmic kernel.

2 New Integral Equation in the Curved Rigid Line Problem

It is well known that the resultant forces and displacement in plane elasticity can be expressed by (Muskhelishvili, 1953):

$$f=i(X+iY)=\varphi(z)+z\overline{\varphi'(z)}+\overline{\psi(z)}+C_1 \quad (3)$$

$$2G(u+iv)=\kappa\varphi(z)-z\overline{\varphi'(z)}-\overline{\psi(z)}+C_2 \quad (4)$$

where $\varphi(z)$ and $\psi(z)$ are the complex potential, $z=x+iy$ and G is the shear modululs of elasticity.

Now we want to solve the curved rigid line problem as shown in Fig. 1(*a*). Assume in the problem that the remote rotation will be $\epsilon^\infty=0$ and the remote stresses are expressed by σ_x^∞, σ_y^∞, and σ_{xy}^∞. If the curved rigid line is in a floating state, then the conditions along the curve rigid line L in Fig. 1(*a*) will be:

(*a*) There are no resultant forces applied on the whole rigid line.

(*b*) The rigid line has a rotation

$$2G(u+iv)=i\gamma t_o,\ (t_o\epsilon L) \quad (5)$$

where the rotation γ is determined from the zero-moment condition imposed on the whole rigid line.

Obviously, the problem shown in Fig. 1(*a*) can be separated into two problems as shown in Fig. 1(*b*) and Fig. 1(*c*).

The problem shown in Fig. 1(*b*) has a solution (Muskhelishvili, 1953):

$$2G(u+iv)=(\kappa-1)\Gamma z-\overline{\Gamma'}\bar{z}, \quad (6)$$

where

$$\Gamma=(\sigma_x^\infty+\sigma_y^\infty)/4,\quad \Gamma'=(\sigma_y^\infty-\sigma_x^\infty)/2+i\sigma_{xy}^\infty. \quad (7)$$

It is easy to see that the conditions imposed on the problem shown in Fig. 1(*c*) can be stated as follows:

(*a*) There are no resultant forces applied on the whole rigid line.

(*b*) The displacements along the curved rigid line should take the form

$$2G(u+iv)=-(\kappa-1)\Gamma t_o+\overline{\Gamma'}\bar{t}_o+i\gamma t_o,\ (t_o\epsilon L) \quad (8)$$

where the rotation γ is still determined from the zero-moment condition.

It is easy to verify that the appropriate potentials in the curved rigid line problem have the following form:

$$\varphi(z)=\int_0^t\mu(s)\log(z-t)ds,$$
$$\psi(z)=-\kappa\int_0^t\overline{\mu(s)}\log(z-t)ds-\int_0^t(\mu(s)\bar{t}/(z-t))ds. \quad (9)$$

Since there are no resultant forces applied on the whole rigid line L, from equations (3) and (9) we obtain

$$\int_0^t\mu(s)ds=0. \quad (10)$$

Let z approach the upper or lower border of the curved rigid line L in Fig. 1, i.e., $z\to t_o^+$ or $z\to t_o^-$. In either case, we have

$$2G(u(t_o)+iv(t_o))=\kappa\int_0^t 2\log|t-t_o|\mu(s)ds$$
$$-\int_0^t((t-t_o)/(\bar{t}-\bar{t}_o))\overline{\mu(s)}ds+C_2. \quad (11)$$

After using equations (8) and (11), and by letting

$$t-t_o=r(t,t_o)\exp(i\theta(t,t_o)),$$
$$\mu(s)=\mu_1(s)+i\mu_2(s), \quad (12)$$

the following system of the integral equations is obtainable

$$\kappa\int_0^t 2\log(r(t,t_o))\mu_1(s)ds+\int_0^t[-\mu_1(s)\cos(2\theta(t,t_o))$$
$$-\mu_2(s)\sin(2\theta(t,t_o))]ds+\mathrm{Re}(C_2)=H_1(t_o),\ (t_o\epsilon L) \quad (13a)$$

[19]Professor, Division of Engineering Mechanics, Jiangsu Institute of Technology, Zhengjian, Jiangsu, 212013, People's Republic of China.

Manuscript received by the ASME Applied Mechanics Division, December 4, 1987; final revision, March 1, 1989.

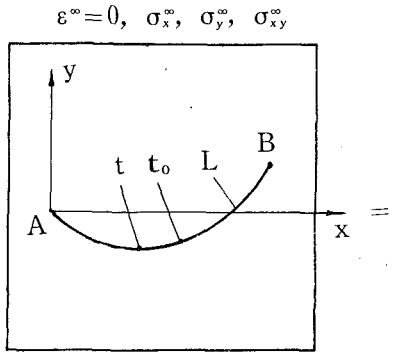


Fig. 1(a)

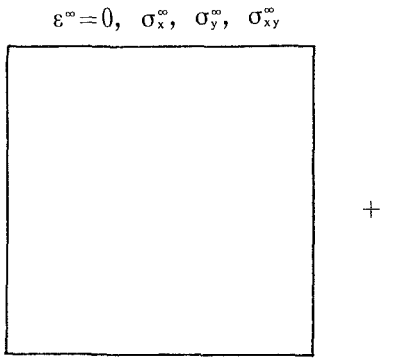


Fig. 1(b)

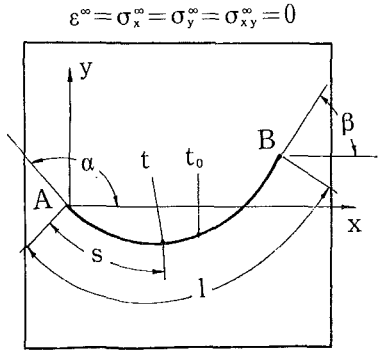


Fig. 1(c)

Fig. 1 The original problem of Fig. (a) consists of two particular problems, (b) and (c). The boundary condition along L has been shown by equation (5) (for (a)) and equation (8) (for (c)).

$$\kappa\int_0^\iota 2\log(r(t,t_o))\mu_2(s)ds+\int_0^\iota[-\mu_1(s)\sin(2\theta(t,t_o))+\mu_2(s)\cos(2\theta(t,t_o))]ds+\mathrm{Im}(C_2)=H_2(t_o)\ (t_o\epsilon L) \quad (13b)$$

where $H_1(t_o)$ and $H_2(t_o)$ are derived from equation (8) and can be expressed as

$$H_1(t_o)=\mathrm{Re}(-(\kappa-1)\Gamma t_o+\bar{\Gamma}'\bar{t}_o+i\gamma t_o),\ (t_o\epsilon L)$$
$$H_2(t_o)=\mathrm{Im}(-(\kappa-1)\Gamma t_o+\bar{\Gamma}'\bar{t}_o+i\gamma t_o),\ (t_o\epsilon L). \quad (14)$$

Under the condition (10) and from equation (9), the complex potential $\psi(z)$ at the remote place can be expressed as

$$\psi(z)=(a_1+ib_1)/z+(a_2+ib_2)/z^2+\ldots, \quad (15)$$

where a_1, a_2, and b_2 are some integrals and

$$b_1=(\kappa+1)\mathrm{Im}\left(\int_0^\iota t\overline{\mu(s)}ds\right). \quad (16)$$

Meanwhile, the moment applied on a large circle or on the curved rigid line is equal to $M=2\pi b_1$ (Muskhelishvili, 1953). From the fact that the curved rigid line is in a floating state, then the following equation should be satisfied

$$\mathrm{Im}\left(\int_0^\iota t\overline{\mu(s)}ds\right)=0. \quad (17)$$

From the eigenexpansion form of the stress-displacement field at the vicinity of the rigid line tip (Chen, 1986), one should take

$$\mu(s)=h_1 s^{-1/2},\ \text{(at the vicinity of the left tip } A)$$
$$\mu(s)=h_2(\iota-s)^{-1/2},\ \text{(at the vicinity of the right tip } B). \quad (18)$$

After some manipulation, the stress singularity coefficient at the left and right crack tips can be expressed as

$$(S_1+i((\kappa+3)/(\kappa-1))S_2)_A=-(3+\kappa)\pi(2\pi)^{1/2}\exp(-i\alpha)h_1,$$
$$(S_1+i((\kappa+3)/(\kappa-1))S_2)_B=(3+\kappa)\pi(2\pi)^{1/2}\exp(-i\beta)h_2. \quad (19)$$

3 Numerical Example

In the example, the infinite plate contains a straight rigid line with length $2a=2$ and the remote conditions are $\epsilon^\infty=0$, $\sigma_x^\infty=1$, $\sigma_y^\infty=0$ and $\sigma_{xy}^\infty=0$. Under these conditions, the integral equations are reduced to

$$\int_{-1}^{1}\log|t-t_o|\mu(t)dt=-(\kappa+1)t_o/8\kappa,\ |t_o|<1 \quad (20)$$

Considering equation (10), equation (20) has a solution as follows (Cheung and Chen, 1987):

$$\mu(t)=((\kappa+1)/8\pi\kappa)t(1-t^2)^{-1/2},$$
$$S_1=((\kappa+1)(\kappa+3)/8\kappa)(\pi)^{1/2}. \quad (21)$$

The structure of the integral equations (13a) and (13b) is essentially the same as mentioned in the curve crack problem (Cheung and Chen, 1987). Therefore, the same interpolation formulae which was cited in the paper (Cheung and Chen, 1987) can be used for the function $\mu(s)$ in this paper. If 20 intervals with equal space are divided along the rigid line, the maximum deviation of the calculated $\mu(t)$ at the discrete points from the exact solution is less than 2 percent. In addition, the calculated stress singularity coefficient is

$$S_1=1.0176((\kappa+1)(\kappa+3)/8\kappa(\pi)^{1/2}. \quad (22)$$

From (22) we see that, even for a coarse division in computation, the numerical solution of the integral equation already provides a very accurate result.

Acknowledgment

The research project is supported by National Natural Science Foundation of China.

References

Chen, Y. Z., 1986, "Singular Behaviour at Fixed Rigid Line Tip in Plane Plane Elasticity," *Engineering Fracture Mechanics*, Vol. 25, No. 1, pp. 11-16.

Chen, Y. Z., and Cheung, Y. K., 1987, "Stress Singularity Coefficients in an Infinite Plate Containing Rigid Line and Applied by Concentrated Forces," *Engineering Fracture Mechanics*, Vol. 26, No. 5, pp. 729-739.

Cheung, Y. K., and Chen, Y. Z., 1987, "New Integral Equation for Plane Elasticity Crack Problems," *Theoretical and Applied Fracture Mechanics*, Vol. 7, No. 3, pp. 177-184.

Erdogan, F., and Gupta, G. D., 1972, "Stresses Near Flat Inclusion in Bonded Dissimilar Materials," *International Journal of Solids and Structures*, Vol. 8, No. 4, pp. 533-548.

Muskhelishvili, N. L., 1953, *Some Basic Problems of Mathematical Theory of Elasticity*, P. Noordhoof, The Netherlands.

Wang, Z. Y., Zhang, H. T., and Chou, Y. T., 1985, "Characteristics of Elastic Field of a Rigid Line Inhomogeneity," ASME JOURNAL OF APPLIED MECHANICS, Vol. 52, No. 4, pp. 818-812.

Journal of Applied Mechanics

Discussion

Approximate Decoupling of the Equations of Motion of Linear Underdamped Systems[1]

Stephen F. Felszeghy.[2] The authors examined the motions of linear underdamped discrete systems with n degrees-of-freedom when these motions are expanded in the modal vectors of the undamped system. They considered the error that is introduced when, in solving for the temporal parts of these expansions, that is, the normal coordinates from the governing equations of motion, the generalized damping matrix with nonzero off-diagonal elements in the equations of motion is approximated with a diagonal matrix. The authors derived "tight" normed error bounds for the normal coordinates q_i, $i=1, \ldots, n$. It can be shown that under certain circumstances the error bounds can be tightened even further.

For let us suppose that a modal vector $u^{(j)}$, which satisfies the equation representing the eigenvalue problem of the undamped system

$$\mathbf{K}u^{(j)} = \omega_j^2 \mathbf{M}u^{(j)}, \tag{1}$$

also satisfies the equation representing the eigenvalue problem of the unsprung system (Felszeghy, 1989)

$$\mathbf{C}u^{(j)} = \gamma \mathbf{M}u^{(j)}. \tag{2}$$

Then, the jth row of the symmetric matrix $\tilde{\mathbf{C}}_r$ is zero, and it follows from the authors' equation (2.7) that the error $e_j(t)$, associated with normal coordinate q_j, is identically zero. Hence, the error bounds can be interpreted as applying exclusively to those normal coordinates q_i for which (2) is *not* satisfied.

Lastly, Fig. 1 appears to show a zero-state response instead of a steady-state response.

References

Felszeghy, S. F., 1989, "The Development of Natural Modes of Free Vibration for Linear Discrete Systems from the Synchronous Motion Assumption," *ASME Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 111, No. 1, pp. 77–81.

Authors' Closure

We have read Professor Felszeghy's discussion and would like to thank him for his comments. The point brought up in his discussion appears to be an easy consequence of the results presented in our paper.

Professor Felszeghy remarked that, if for some j, the modal vectors $u^{(j)}$ satisfy *assumption* (2) in his discussion, then the upper bound on the approximation error can be tightened. The general results presented in our paper incorporate the special case mentioned by Professor Felszeghy. If for some j, the modal vectors $u^{(j)}$ satisfy *assumption* (2), then in the normalized damping matrix the off-diagonal elements on the rows corresponding to those j are zero, and, hence, the quantities σ_j defined in equation (2.11) in our paper are zero for those j. Therefore, equation (2.15) in our paper obviously yields the same error bound as suggested by Professor Felszeghy. The error bounds derived in our paper remain the tightest possible in a functional form specified by equation (2.16) in our paper.

Lastly, Fig. 1 in our paper depicts a zero-state response which has reached its steady state after a short transient behavior.

[1] By S. M. Shahruz and F. Ma, published in the September 1988 issue of the ASME JOURNAL OF APPLIED MECHANICS, Vol. 55, No. 3, pp. 716–720.

[2] Professor of Mechanical Engineering, School of Engineering and Technology, California State University, Los Angeles, 5151 State University Drive, Los Angeles, Calif. 90032. Mem. ASME.

Journal of Applied Mechanics

Book Reviews

Buckling and Postbuckling, by J. Arbocz, M. Potier-Ferry, J. Singer, and V. Tvergaard. Springer-Verlag, Berlin, 1987. 246 pages.

REVIEWED BY I. ELISHAKOFF[1]

The volume contains four lectures delivered at the International Centre for Mechanical Sciences (CISM) in Udine, Italy, in 1985. It opens with a preface by Profs. W. T. Koiter and M. Potier-Ferry, who note that "these lectures are to give a comprehensive account for the theoretical, numerical, and experimental methods which are useful to analyze the buckling and post-buckling behavior of structures." Indeed, four internationally recognized specialists – leaders in the field, contributed lectures on different topics of buckling. These are "Foundations of Elastic Postbuckling Theory" (by M. Potier-Ferry), "Post-Buckling Behavior of Structures: Numerical Techniques for More Complicated Structures" (by J. Arbocz), "Effect of Plasticity on Post-Buckling Behaviour" (by V. Tvergaard), and "Experimental Techniques and Comparison with Theoretical Results" (by J. Singer).

The first chapter (Potier-Ferry) represents a sound and definitive introduction to the elastic postbuckling theory, with particular emphasis on classification of singularities according to their "robustness" and on Koiter's imperfection sensitivity theory. The second chapter (Arbocz) deals with the stability of axially-compressed imperfect orthotropic cylindrical shells via numerical techniques. Complementing the simple models used in the first chapter to elucidate basic ideas, initial imperfections are represented here by double Fourier series and various approximations are illustrated. Multimode analysis, introduced by the late Professor Babcock and by the author, is compared with the large finite difference code STAGS. The chapter concludes with a detailed description of stochastic stability analysis as developed by the present reviewer and Arbocz for shell structures. The author also mentions the development of DISDECO, the Delft Interactive Shell Design Code, an optimal combination of accumulated theoretical knowledge in the field of shell stability with advanced interactive and computational facilities. One would want the code to be "very user-friendly."

Whereas the second chapter is devoted exclusively to elastic instability, plastic effects are studied in depth in Tvergaard's contribution, which includes inter alia Hill's general theory of uniqueness and bifurcation of elastic-plastic solids and Hutchinson's asymptotic procedures for estimating the initial post-bifurcation behavior. Hypoelastic estimates, first introduced by Hutchinson and Budiansky, are discussed for elastic-plastic cylindrical shells. Other important topics are numerical analysis of plastic buckling, localization of the buckling pattern, and the effect of rate sensitivity.

The three theoretical and/or numerical lectures are appropriately complemented by the fourth (Singer), which examines the role of experiments in buckling and post-buckling. The article first deals with the motivation for experiments, whose aims may be summed up as follows:

(*a*) better understanding of buckling and post-buckling behavior and of the primary factors affecting it,
(*b*) discovery of new phenomena,
(*c*) better inputs for computation,
(*d*) correlation factors between analysis and tests,
(*e*) creation of confidence in multipurpose computer programs,
(*f*) novel ideas of construction of highly complex structural elements, and
(*g*) certification tests of full-scale structures.

These first-class lectures are a timely contribution to better understanding of this ancient topic – buckling of structures. (There are those who believe that the first engineering structure – the Tower of Babel – buckled under its own weight.) This volume is, therefore, a must at every engineering department library for the staff and students to have access to this elegant review of both the state-of-the-art and the latest developments in buckling and post-buckling.

Elastic-Plastic Problems, by B. D. Annin and G. P. Cherepanov. ASME Press, New York, 1988. 250 pages. Price: $31.20 (ASME members); $39.00 (nonmembers).

REVIEWED BY W. J. DRUGAN[2]

This book is a survey of analytical (and some numerical) solutions to selected two and three-dimensional boundary value problems for elastic-plastic materials (primarily metals). It is written by two members of the U.S.S.R. Academy of Sciences who are accomplished researchers in solid mechanics. The present version is an "ASME Press Translation" of the Russian edition published by Nauka Publishers in 1983. The translation is generally quite good. I believe I speak for many when I encourage ASME Press to continue this valuable service of translating books by distinguished foreign researchers; for example, several additional Soviet books on solid mechanics merit serious consideration.

The present book is aimed at researchers and practitioners who are experienced in solid mechanics and applied mathematics; it focuses on the mathematical formulation and solution of problems. Little contact is made with the behavior of actual materials and the physical motivation and justifica-

[1] Professor, Faculty of Aerospace Engineering, Technion – I.I.T., Haifa, 32000, Israel.

[2] Associate Professor, Department of Engineering Mechanics, University of Wisconsin, Madison, Wis. 53706, Assoc. Mem. ASME.

tion for the constitutive equations employed; indeed, the reader may get the impression in some problems that a specific constitutive equation was selected primarily because it facilitated a closed-form analytical solution. It would be difficult to *begin* learning elastoplasticity from the book, as the authors tacitly acknowledge, by referencing a number of introductory texts on the subject. However, for a knowledgeable reader, the book has a number of strengths: The nonstandard viewpoint occasionally adopted by the authors is illuminating. Analytical derivations and solutions to some quite complex problems are presented unflinchingly, if sometimes without much tutelage. Helpful comparisons are made between the structure of resulting equations and equations from other branches of applied science. The presentation is at times very general and ambitious, which this reader found interesting and instructive. Many different mathematical methods are illustrated, such as the theory of functions of complex variables, conformal mapping, perturbation methods, integral equation theory, variational approaches, etc. The book contains about 250 references to books and articles, both Soviet and Western and, hence, can provide a roadmap to some of the Soviet literature in this area, although much of it is untranslated. For many of these untranslated works, the book serves the valuable function of explaining and summarizing their contributions. I was also pleased to see several important Western works summarized. Additionally, the authors attempt in several places to give an historical perspective.

The book does have drawbacks. It is not up-to-date, and the selection of topics is strongly influenced by the interests of the authors. The entire book is restricted to a quasi-static, "small strain," rate-independent formulation, and most of the research papers referenced are over 15 or 20 years old. The authors devote extensive space to their own work, thereby excluding some important work of other researchers; this governs the types of problems covered, of which crack, hole, and notch problems are heavily favored. Even within the realm of crack problems, much important work is unmentioned, such as the Hutchinson-Rice-Rosengren plane near-tip singular fields for stationary cracks in power-law hardening elastic-plastic materials, and the quasi-static and dynamic growing crack near-tip elastic-ideally plastic solutions of L. I. Slepyan (and related work of Western authors). Another criticism of the book is that in many places it is obvious the authors have not strived for lucid explanations; sometimes I was hard pressed even to understand the rationale behind a problem's formulation. When the solution of problems gets difficult, especially as in certain plane stress and plane strain cases, the explanations provided are not always up to conveying the material. This is exacerbated by absent or unclear figures that often show a bare minimum of information. The book also lacks an index. Finally, there are a number of misleading or erroneous statements. For example, in the chapter on torsion, it is explained that the out-of-plane displacement w is an unknown function of all three Cartesian coordinates x, y, z, while it is also stated that the out-of-plane normal strain component vanishes. In Chapter 6, it is stated that a general three-dimensional problem of ideal elastoplasticity is statically determinate whenever the loads on the boundary of the body are given.

As regards specific content, the book has an introductory chapter that tersely summarizes the governing field equations for elastic-plastic problems and some general methods of their solution. Following are rather extensive chapters on complex (anti-plane) shear, torsion, and plane strain, and far briefer ones on plane stress and three-dimensional problems. There are two short, mathematical appendices, one on a method of solving problems with unknown boundaries, and one on nonlinear Riemann problems.

In summary, the book should interest, and occasionally fascinate, fairly experienced solid mechanicians and applied mathematicians who know some plasticity and would like to see some important, difficult problems elegantly solved, but who are not expecting a current, comprehensive treatment of the title subject area.

Wave Propagation in Solids and Fluids, by J. L. Davis. Springer-Verlag, New York, 1988. 386 pages. Price: $75.00.

REVIEWED BY D. MINTZER[3]

This volume, the first of two on wave propagation, emphasizes the mathematical structure of the subject rather than laying a foundation for application to various physical phenomena. (This volume is concerned with continua which do not involve electromagnetic phenomena; a second volume will consider waves involving a variety of electric and magnetic interactions, as well as quantum mechanical waves.) The book touches on a wide variety of areas of wave propagation (strings, water waves, sound waves, waves in elastic media), connecting the underlying mathematical foundations of the various areas. Although he does not delve deeply into any particular problem, the author gives good insight into the models used in the derivations, so that the reader does not have the feeling that it is all just a mathematical exercise. Nevertheless, the generality of approach, the lack of specific examples, and the neglect of even simplified approaches to practical problems makes this unsuitable as a text for any but an advanced student in applied mathematics. It will be useful, however, as a reference work in its chosen field.

The book starts with a chapter on "Oscillatory Phenomena," going rapidly through the usual expositions of free and forced damped vibrations of a linear mass-spring system, and then touching briefly on Fourier series; a brief section on two-dimensional motion follows. After a short exposition of two-body coupled oscillations, the general N-body (coupled) problem is treated by means of Lagrange's equations. The eigenvalue equations for small oscillations are thereby derived, the matrix formulation is then introduced, and the normal coordinate equations are found by means of a similarity transformation. The paragraph which closes the chapter then shows that this general technique gives the same result for a two-body system as did the earlier exposition of the two-body coupled oscillation case. Although each step is carefully explained, this is clearly a chapter for the reader who has already been exposed to most of the material.

The second and fourth chapters, both on vibrations of a string, again point up both the strengths and weaknesses of this book. The second chapter gives a brief description of waves on a string, emphasizing the physics of the situation. The fourth chapter contains the standard (Newton's law) derivation of the wave equation for a string undergoing small oscillations; the general solution to the equation is derived, and a few special cases are discussed; finally, the derivation of the wave equation as the limit of an N-body mass-spring system is developed. Yet nowhere is there any discussion of the cases of discontinuities in the physical properties of the string, masses attached to the string, or reactive end-supports. This approach, with its emphasis in the mathematical generalities to the virtual exclusion of subjects with a view towards eventual applications, is carried on throughout most of the book.

However, the two most mathematical chapters do deserve

[3]Professor, Department of Mechanical Engineering, Northwestern University, Evanston, Ill. 60208.

special mention. Chapter 3, on the partial differential equations of wave propagation, develops the theory of characteristics, gives a good discussion of the Cauchy initial value problem, and derives the corresponding integral equation. Chapter 9, entitled "Variational Method in Wave Phenomena" seems out of place in this book, being more suitable to a book on advanced mechanics. The application of the variational methods given here is deriving the wave equation, but the major content of the chapter is on deriving Lagrange's equations, Hamilton's equations, and Hamilton-Jacobi theory. A short section on the relationship between the wave equation and the ray equations, and a few words about quantum mechanics end the chapter.

In all, the book does reasonably well with what the author sets out as his goal – an exposition of the mathematical methods of wave phenomena, but the presentation appears to restrict it to only an audience of physicists and engineers who already have a good background in the subject.