Transactions of the ASME®

Journal of Applied Mechanics

Published Bimonthly by The American Society of Mechanical Engineers

VOLUME 70 • NUMBER 5 • SEPTEMBER 2003

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published bimonthly (Jan., Mar., May, July, Sept., Nov.) The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing office. 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.
STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B7.1, Para. 3).

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc. Canadian Goods & Services Tax Registration #126148048. Reduced-Order Models for Nonlinear Dynamical

TECHNICAL PAPERS

(Contents continued on inside back cover)

This journal is printed on acid-free paper, which exceeds the ANSI Z39.48-1992 specification for permanence of paper and library materials. ™
85% recycled content, including 10% post-consumer fibers.

(**Contents continued**)

Journal of Applied Mechanics **Volume 70, Number 5** **SEPTEMBER 2003**

BRIEF NOTES

DISCUSSION

The ASME Journal of Applied Mechanics is abstracted and indexed in the following:

Alloys Index, Aluminum Industry Abstracts, Applied Science & Technology Index, AMR Abstracts Database, Ceramic Abstracts, Chemical Abstracts, Civil Engineering Abstracts, Compendex (The electronic equivalent of Engineering Index), Computer & Information Systems Abstracts, Corrosion Abstracts, Current Contents, EEA (Earthquake Engineering Abstracts Database), Electronics & Communications Abstracts Journal, Engineered Materials Abstracts, Engineering Index, Environmental Engineering Abstracts, Environmental Science and Pollution Management, Fluidex, Fuel & Energy Abstracts, GeoRef, Geotechnical Abstracts, INSPEC, International Aerospace Abstracts, Journal of Ferrocement, Materials Science Citation Index, Mechanical Engineering Abstracts, METADEX (The electronic equivalent of Metals Abstracts and Alloys Index), Metals Abstracts, Nonferrous Metals Alert, Polymers Ceramics Composites Alert, Referativnyi Zhurnal, Science Citation Index, SciSearch (Electronic equivalent of Science Citation Index), Shock and Vibration Digest, Solid State and Superconductivity Abstracts, Steels Alert, Zentralblatt MATH

D. L. Marshall[1]
M. E. Taylor
Y. Ivshin
G. Diderrich
Johnson Controls, Inc.,
P.O. Box 591,
Milwaukee, WI 53201-0591

T. J. Pence[2]
Department of Mechanical Engineering,
Michigan State University,
East Lansing, MI 48824-1226
e-mail: pence@egr.msu.edu

On Corrosion-Induced Creep and the Shedding of Oxide Pest From Metal Plates

We consider a boundary value problem for a flat plate of an originally ductile metal that is subject to surface corrosion. Corrosion is modeled with the aid of two mobile interfaces. The leading interface is an oxidation front where volumetric expansion generates stress. This interface is initially present, with its location taken as prescribed. The trailing interface is a failure front associated with the shedding of oxide pest. This interface is not initially present, rather it first appears at some finite time into the process, and is determined via a stress-based failure criterion. The problem is reduced to a set of ordinary differential equations whose form changes depending on whether or not shedding has initiated. Numerical treatment shows a substantial increase in creep rate due to oxide shedding. [DOI: 10.1115/1.1604837]

1 Introduction

We present a simple treatment for certain stress-related issues pertaining to the formation, growth, and degradation of a corrosion surface layer on an infinite plate. The formulation is in terms of a boundary value problem wherein the metal is ductile and the oxide is brittle. Corrosion near each surface is modeled with the aid of two mobile interfaces. The leading interface is an oxidation front where volumetric expansion (described by the Pilling-Bedworth ratio) generates stress. This interface is taken to coincide with the external boundary at $t=0$ with subsequent motion into the plate interior. The trailing interface is a failure front associated with the shedding of oxide pest. This front also first appears at the external boundary, but at some time $t=t_{ff}>0$ associated with the satisfaction of an appropriate failure condition. Passage of this second front sheds material from the plate thereby generating corrosion pest found either as loose surface scale or as external debris.

Since oxide formation involves volumetric expansion, the formation of oxide on originally unstressed metal will generate stress if the oxide is constrained so that this volume is not attained, [1]. Such a constraint occurs when transformation from metal to oxide takes place in the interior of an oxide layer as for example will be the case if oxygen ions have sufficiently higher mobility than metal ions within the oxide. The advance of oxidation fronts can then be modeled on the basis of the continuum mechanics of diffusion, [2]. The free volume change associated with oxidation is described with the aid of the Pilling-Bedworth ratio (PBr) which is regarded as a material parameter. The PBr is tabulated for standard metal-to-oxide transformations of engineering significance, and associated eigenstrain-type problems have been analyzed for the constraint stresses associated with oxide formation, [3–5]. Oxide degradation due to such stress is addressed in [6,7]. The combined effects of diffusion and stress in the setting of carefully posed boundary value problems is discussed in [8].

For ductile metals, creep provides a mechanism for stress relaxation. Analysis of creep relaxation during oxide growth is provided by Bernstein [9] and by Touati et al. [10] via the consideration of boundary value problems in the setting of infinitesimal strain. This provides a framework for determining the evolution of the stress fields in the oxide. Here we extend this type of creep analysis so as to consider the possibility of oxide failure in the event that the oxide stresses become tensile.

We consider an appropriate boundary value problem for a flat plate of original thickness $2H$ subject to advancing oxidation fronts $X=\pm D_{\mathrm{oxy}}(t)$ beginning from each of the two plate surfaces $X=\pm H$. In view of the system symmetry, the problem is formulated on $0\leqslant X\leqslant H$ paying special attention to an infinitesimal strain description that allows each location X to creep prior to oxidation and to undergo free volume expansion at its instant of oxidation. After oxidation this inelastic strain is locked in, whereupon all additional strain for locations in the oxide ($D_{\mathrm{oxy}}(t)<X<H$) is due solely to elastic effects. Due to the in-plane constraint afforded by the parallel metal and oxide zones (i.e., displacement continuity), the in-plane stress is originally tensile in the metal and compressive in the oxide. The in-plane compressive stress in the oxide is minimized at the external surface $X=H$, since this location does not undergo creep prior to oxidation.

The simple flat plate geometry permits a boundary value problem treatment that reduces to a system of ordinary differential equations with time as the independent variable. On this basis we show that the stress fields in the oxide can evolve with time from a state of global compression to a state with localized tensions. On the assumption that tension leads to oxide failure in the form of corrosion pest, this naturally leads to a description involving both the oxidation interface $X=D_{\mathrm{oxy}}(t)$ and a shedding interface $X=D_{\mathrm{pest}}(t)$. The latter is described by a failure condition in rate form, and, as we show, the mathematical formulation gives rise to differential equations with retarded time arguments. In the example considered here these equations are treated by procedures that are standard in the study of differential equations with retarded time arguments.

Section 2 gives the general formulation of the governing equations describing these processes. Section 3 treats these equations so long as the oxide remains in a state of compressive stress. During this period the retarded time is the instant of first oxide formation, and the associated retarded time terms evaluate to zero. A numerical example is presented for this case demonstrating that the stress in the oxide will change from compressive to tensile, with the change first occurring at the plate's external surface. Section 4 extends the treatment to account for shedding of the surface oxide when the in-plane stress becomes sufficiently tensile. For this shedding analysis, the retarded time is the time at which the

[1]Deceased. This article is dedicated to him by his co-authors.
[2]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 3, 2002; final revision, Mar. 27, 2003. Associate Editor: E. Arruda. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

currently failing oxide had experienced its prior transformation from metal to oxide. This time value is not known in advance and becomes an additional system variable. The associated increase in the number of system unknowns is compensated by a supplemental differential equation that follows from the rate form of the failure condition. The resulting formulation remains well-posed, and the numerical example is continued into the shedding regime. After a brief transient, the oxide layer is found to assume a thickness that is relatively constant. Namely, the rate of thickness variation is approximately an order of magnitude less than the rate of corrosion advance. Further, the metal core carries additional tensile load, and this significantly increases the overall creep rate. Section 5 then provides some brief summary comments and conclusions.

2 The General Formulation

The geometrical description distinguishes between locations in a flat plate described with respect to the reference configuration (at $t=0$) and with respect to the evolving current configuration (for $t>0$). For the plate geometry under consideration here, we let X give the through thickness coordinate, and (Y,Z) provide the in-plane coordinates with respect to the reference configuration. Symmetry is assumed with respect to the midplane $X=0$ and we consider locations $0\leqslant X\leqslant H$. This region is partitioned between base metal $0\leqslant X\leqslant D_{\rm oxy}(t)$, intact oxide $D_{\rm oxy}(t)\leqslant X\leqslant D_{\rm pest}(t)$, and corrosion debris $D_{\rm pest}(t)\leqslant X\leqslant H$. Corrosion debris is oxide that has been degraded due to stress and effectively shed from the system so that $X=D_{\rm pest}(t)$ is regarded as a shedding front. The transition from intact oxide to corrosion debris is described by a failure condition in terms of stress. There is no oxide at the initial time $t=0$ when the system is all metal so that $D_{\rm oxy}(0)=H$, $D_{\rm pest}(0)=H$. Oxidation begins immediately so that $D_{\rm oxy}(t)$ is decreasing with time. For some initial time period $0<t<t_{ff}$ the stress in the oxide is such that no failure occurs, so that $D_{\rm pest}(t)=H$ for $0<t<t_{ff}$. First failure occurs at $X=H$, $t=t_{ff}$ whereupon $D_{\rm pest}(t)$ is decreasing with time for $t>t_{ff}$. Locations will be described with respect to the current configuration using a lower case letter convention for distance quantities $(x,y,z,d_{\rm oxy}(t),d_{\rm pest}(t),h)$.

In order to separately denote field quantities in the oxide from those in the metal, a superposed ˆ is used to distinguish field quantities in the oxide. Let u_i, ε_{ij}, σ_{ij}, $\hat{u}_i$, $\hat{\varepsilon}_{ij}$, $\hat{\sigma}_{ij}$ denote components of displacement, strain, and stress in the metal and oxide, respectively. The $\{X,Y,Z\}$ system provides a principle frame by geometrical symmetry and uniformity in Y and Z of the corrosion process. There is no inherent in-plane length scale, consequently the strain and stress fields are independent of Y and Z. The equivalence of the Y and Z-directions allows us to eliminate Z in terms of Y. Stress and strain are therefore described in terms of ε_{xx}, ε_{yy}, σ_{xx}, σ_{yy}, $\hat{\varepsilon}_{xx}$, $\hat{\varepsilon}_{yy}$, $\hat{\sigma}_{xx}$, $\hat{\sigma}_{yy}$, where lower case indices are employed for notational ease. All strains are measured with respect to the reference configuration. The metal undergoes homogeneous deformation, so that $\varepsilon_{xx}=\varepsilon_{xx}(t)$, $\varepsilon_{yy}=\varepsilon_{yy}(t)$. Displacement continuity across the interface $X=D_{\rm oxy}(t)$ requires that $\hat{\varepsilon}_{yy}=\varepsilon_{yy}(t)$. The remaining strain component $\hat{\varepsilon}_{xx}$ is dependent on both X and t as explained in what follows, so that $\hat{\varepsilon}_{xx}=\hat{\varepsilon}_{xx}(X,t)$. These conditions yield a single-valued displacement field upon integration. Taking the plate origin as fixed gives: $u_y=\hat{u}_y=Y\varepsilon_{yy}(t)$, $u_z=\hat{u}_z=Z\varepsilon_{yy}(t)$, $u_x=X\varepsilon_{xx}(t)$, $\hat{u}_x=D_{\rm oxy}(t)\varepsilon_{xx}(t)+\int_{D_{\rm oxy}(t)}^{X}\hat{\varepsilon}_{xx}(\xi,t)d\xi$.

The strains in both the metal and the oxide are decomposed into the sum of an elastic strain (superscript E) and an inelastic strain (superscript I) as follows:

$$\varepsilon_{xx}^{E}(t)+\varepsilon_{xx}^{I}(t)=\varepsilon_{xx}(t),\quad \hat{\varepsilon}_{xx}^{E}(X,t)+\hat{\varepsilon}_{xx}^{I}(X)=\hat{\varepsilon}_{xx}(X,t),$$
$$\varepsilon_{yy}^{E}(t)+\varepsilon_{yy}^{I}(t)=\varepsilon_{yy}(t)=\hat{\varepsilon}_{yy}(t)=\hat{\varepsilon}_{yy}^{E}(X,t)+\hat{\varepsilon}_{yy}^{I}(X). \tag{1}$$

Observe that $\hat{\varepsilon}_{xx}^{I}$ and $\hat{\varepsilon}_{yy}^{I}$ do not depend on t. This is explained as follows. The inelastic strain in the metal is due to creep. The oxide, however, is not able to creep. The inelastic strain in the oxide is instead due to the combination of two effects, both of which cease after oxidation occurs. The first effect is the creep prior to oxidation during which time significant metal flow can take place. The second effect is the volumetric expansion from the metal-to-oxide transformation. Consequently, for the oxide, the inelastic strain is locked in by the oxidation process and so depends on the time at which oxidation occurred. This time depends on X and is independent of t, and so accounts for $\hat{\varepsilon}_{xx}^{I}=\hat{\varepsilon}_{xx}^{I}(X)$ and $\hat{\varepsilon}_{yy}^{I}=\hat{\varepsilon}_{yy}^{I}(X)$. Also, since $\hat{\varepsilon}_{yy}^{E}=\hat{\varepsilon}_{yy}-\hat{\varepsilon}_{yy}^{I}$, with $\hat{\varepsilon}_{yy}=\hat{\varepsilon}_{yy}(t)$ and $\hat{\varepsilon}_{yy}^{I}=\hat{\varepsilon}_{yy}^{I}(X)$ it follows that $\hat{\varepsilon}_{yy}^{E}=\hat{\varepsilon}_{yy}^{E}(X,t)$. Neither the oxide elastic strains $\hat{\varepsilon}^{E}$ by themselves, nor the oxide inelastic strains $\hat{\varepsilon}^{I}$ by themselves, need satisfy the standard compatibility conditions that ensure an associated displacement field. Thus it is generally not possible to decompose the displacement in the oxide into an elastic displacement and an inelastic displacement.

We now turn to consider the stresses. The traction-free boundary condition at the plate surface, the stress equations of equilibrium and the continuity of traction across the metal/oxide interface give $\hat{\sigma}_{xx}=\sigma_{xx}=0$. The remaining stresses $\hat{\sigma}_{yy}$, σ_{yy} are connected to the elastic strains by the standard relations of isotropic linear elasticity, giving

$$\varepsilon_{xx}^{E}=-\frac{2\nu}{(1-\nu)}\varepsilon_{yy}^{E},\quad \sigma_{yy}=\frac{E}{(1-\nu)}\varepsilon_{yy}^{E},$$

$$\hat{\varepsilon}_{xx}^{E}=-\frac{2\hat{\nu}}{(1-\hat{\nu})}\hat{\varepsilon}_{yy}^{E},\quad \hat{\sigma}_{yy}=\frac{\hat{E}}{(1-\hat{\nu})}\hat{\varepsilon}_{yy}^{E}.$$

Here E, ν, $\hat{E}$, $\hat{\nu}$ denote the Young's modulus and Poisson's ratio in the metal and oxide, respectively. Thus equibiaxial stress prevails, with $\hat{\sigma}_{yy}=\hat{\sigma}_{yy}(X,t)$, $\sigma_{yy}=\sigma_{yy}(t)$.

It is assumed that the edges of the plate are traction free. Consequently, the tractions on any internal plane of constant Y or constant Z are self-equilibrated. In view of the system symmetry this gives

$$\int_{0}^{D_{\rm oxy}(t)}\sigma_{yy}(t)(1+\varepsilon_{xx}(t))d\xi+\int_{D_{\rm oxy}(t)}^{D_{\rm pest}(t)}\hat{\sigma}_{yy}(\xi,t)(1+\hat{\varepsilon}_{xx}(\xi,t))d\xi$$
$$=0, \tag{2}$$

where the factors $(1+\varepsilon_{xx})$ and $(1+\hat{\varepsilon}_{xx})$ account for plate thinning prior to oxidation and volumetric expansion upon oxidation. Notice also that only the intact portion of the oxide $D_{\rm oxy}(t)\leqslant X\leqslant D_{\rm pest}(t)$ is regarded as capable of supporting load.

Returning to the inelastic strains, creep is assumed to be volume preserving so that $\varepsilon_{xx}^{I}(t)=-2\varepsilon_{yy}^{I}(t)$. Making appropriate replacements in Eq. (2) now gives

$$\frac{E}{1-\nu}\varepsilon_{yy}^{E}(t)(1-2\varepsilon_{yy}^{I}(t))D_{\rm oxy}(t)+\frac{\hat{E}}{1-\hat{\nu}}\int_{D_{\rm oxy}(t)}^{D_{\rm pest}(t)}\hat{\varepsilon}_{yy}^{E}(\xi,t)$$
$$\times\left(1+\hat{\varepsilon}_{xx}^{I}(\xi)+\left(\frac{-2\hat{\nu}}{1-\hat{\nu}}\right)\hat{\varepsilon}_{yy}^{E}(\xi,t)\right)d\xi=0, \tag{3}$$

where we have ignored the contribution of the elastic strain in the metal ε_{xx}^{E} as a negligible effect in comparison to the creep strain ε_{xx}^{I}.

Turning to consider the inelastic strains in the oxide, let α_x and α_y give the out-of-plane and in-plane stress-free strain associated with the transformation from metal to oxide. These transformation strains are regarded as constants obeying $\alpha_x+2\alpha_y+1=\Delta V_{\mathrm{oxide}}/\Delta V_{\mathrm{metal}}$, which is known as the Pilling-Bedworth ratio (PBr) in the oxidation literature, [1]. The inelastic strains in the oxide register the combined effect of creep prior to oxidation and transformation strain upon oxidation. For the in-plane directions Y and Z, this inelastic strain results in an in-plane line segment of original length l_o first creeping to the length $(1+\varepsilon_{yy}^I)l_o$ and second transforming to the length $(1+\alpha_y)(1+\varepsilon_{yy}^I)l_o$ at the time of oxidation. The difficulty in describing this process stems from the fact that the creep strain value ε_{yy}^I that must be used in this description is not the current value of creep strain, but rather is the creep strain in effect at the time that the location transformed from metal to oxide. Accordingly, define the function $T_D(X)$ that, for each location in the oxide, gives the prior time at which that location had transformed from metal to oxide. Since oxide locations are here described in terms of reference location X, it follows that the prior time function $T_D(X)$ is the functional inverse to $D_{\mathrm{oxy}}(t)$. Thus $T_D(D_{\mathrm{oxy}}(t))=t$ and $D_{\mathrm{oxy}}(T_D(X))=X$. Once the location X has oxidized, the inelastic strain in the oxide remains fixed giving $\hat{\varepsilon}_{yy}^I(X)=\alpha_y+(\alpha_y+1)\varepsilon_{yy}^I(T_D(X))$. A similar expression holds for $\hat{\varepsilon}_{xx}^I$, which, upon use of the volume preservation condition for creep strains, gives $\hat{\varepsilon}_{xx}^I(X)=\alpha_x-2(\alpha_x+1)\varepsilon_{yy}^I(T_D(X))$. The terms involving $\varepsilon_{yy}^I(T_D(X))$ are a new feature of the present treatment so as to validate the analysis to times for which $\varepsilon_{yy}^I(t)$ is no longer small compared to α_y and α_x.

In what follows, it is assumed that α_y and α_x are each small compared with one, so that Eqs. (1) and (3) give

$$\varepsilon_{yy}^E(t)+\varepsilon_{yy}^I(t)=\hat{\varepsilon}_{yy}^E(X,t)+\alpha_y+\varepsilon_{yy}^I(T_D(X)),\quad \text{for } D_{\mathrm{oxy}}(t)<X<D_{\mathrm{pest}}(t), \tag{4}$$

$$\beta D_{\mathrm{oxy}}(t)\varepsilon_{yy}^E(t)(1-2\varepsilon_{yy}^I(t))+\int_{D_{\mathrm{oxy}}(t)}^{D_{\mathrm{pest}}(t)}\hat{\varepsilon}_{yy}^E(\xi,t)(1-2\varepsilon_{yy}^I(T_D(\xi))-\chi\hat{\varepsilon}_{yy}^E(\xi,t))d\xi=0, \tag{5}$$

where we have introduced the nondimensional parameters

$$\frac{(1-\hat{\nu})E}{(1-\nu)\hat{E}}\equiv\beta,\qquad \frac{2\hat{\nu}}{(1-\hat{\nu})}\equiv\chi.$$

Eliminating $\hat{\varepsilon}_{yy}^E(X,t)$ between Eqs. (4) and (5) gives

$$\begin{aligned}&\beta D_{\mathrm{oxy}}(t)\varepsilon_{yy}^E(t)(1-2\varepsilon_{yy}^I(t))\\&\quad+(\varepsilon_{yy}^E(t)+\varepsilon_{yy}^I(t)-\alpha_y)(D_{\mathrm{pest}}(t)-D_{\mathrm{oxy}}(t)-2\phi(t))\\&\quad-(\phi(t)-2\psi(t))-\chi(\varepsilon_{yy}^E(t)+\varepsilon_{yy}^I(t)-\alpha_y)^2(D_{\mathrm{pest}}(t)\\&\quad-D_{\mathrm{oxy}}(t))-2\chi(\varepsilon_{yy}^E(t)+\varepsilon_{yy}^I(t)-\alpha_y)\phi(t)+\chi\psi(t)=0\end{aligned} \tag{6}$$

where the following auxiliary functions

$$\phi(t)=\int_{D_{\mathrm{oxy}}(t)}^{D_{\mathrm{pest}}(t)}\varepsilon_{yy}^I(T_D(\xi))d\xi,\qquad \psi(t)=\int_{D_{\mathrm{oxy}}(t)}^{D_{\mathrm{pest}}(t)}(\varepsilon_{yy}^I(T_D(\xi))^2d\xi,$$

register the integrated effect of the different prior creep cutoff times. Note that $\dot{\phi}(t)=\varepsilon_{yy}^I(T_D(D_{\mathrm{pest}}(t)))\dot{D}_{\mathrm{pest}}(t)-\varepsilon_{yy}^I(t)\dot{D}_{\mathrm{oxy}}(t)$ where the superposed dot denotes time differentiation. A corresponding expression holds for $\dot{\psi}(t)$. It is convenient to express these rate forms as

$$\dot{\phi}=(\varepsilon_{yy}^I\circ T_D\circ D_{\mathrm{pest}})\dot{D}_{\mathrm{pest}}-\varepsilon_{yy}^I\dot{D}_{\mathrm{oxy}}, \tag{7}$$

$$\dot{\psi}(t)=(\varepsilon_{yy}^I\circ T_D\circ D_{\mathrm{pest}})^2\dot{D}_{\mathrm{pest}}-(\varepsilon_{yy}^I)^2\dot{D}_{\mathrm{oxy}}. \tag{8}$$

Here $\circ$ denotes functional composition. The argument $T_D(D_{\mathrm{pest}}(t))$ is a retarded time. Namely, $D_{\mathrm{pest}}(t)>D_{\mathrm{oxy}}(t)$ implies $T_D(D_{\mathrm{pest}}(t))<T_D(D_{\mathrm{oxy}}(t))=t$ so that $\varepsilon_{yy}^I\circ T_D\circ D_{\mathrm{pest}}$ is determined from previously obtained past values of ε_{yy}^I.

The rate form for Eq. (6) is

$$\begin{aligned}&(-2\beta D_{\mathrm{oxy}}\varepsilon_{yy}^E+D_{\mathrm{pest}}-D_{\mathrm{oxy}}-2\phi-2\chi((\varepsilon_{yy}^E+\varepsilon_{yy}^I-\alpha_y)(D_{\mathrm{pest}}-D_{\mathrm{oxy}})-\phi))\dot{\varepsilon}_{yy}^I\\&\quad+(\beta D_{\mathrm{oxy}}(1-2\varepsilon_{yy}^I)+D_{\mathrm{pest}}-D_{\mathrm{oxy}}-2\phi-2\chi((\varepsilon_{yy}^E+\varepsilon_{yy}^I-\alpha_y)(D_{\mathrm{pest}}-D_{\mathrm{oxy}})-\phi))\dot{\varepsilon}_{yy}^E\\&\quad+(-2\varepsilon_{yy}^E-2\varepsilon_{yy}^I-1+2\chi(\varepsilon_{yy}^E+\varepsilon_{yy}^I-\alpha_y))\dot{\phi}+(2-\chi)\dot{\psi}\\&=(-\varepsilon_{yy}^E-\varepsilon_{yy}^I+\alpha_y+\chi(\varepsilon_{yy}^E+\varepsilon_{yy}^I-\alpha_y)^2)\dot{D}_{\mathrm{pest}}+(-\beta\varepsilon_{yy}^E+2\beta\varepsilon_{yy}^E\varepsilon_{yy}^I+\varepsilon_{yy}^E+\varepsilon_{yy}^I-\alpha_y-\chi(\varepsilon_{yy}^E+\varepsilon_{yy}^I-\alpha_y)^2)\dot{D}_{\mathrm{oxy}}.\end{aligned} \tag{9}$$

If D_{pest}, D_{oxy}, and T_D are prescribed, then the set (7)–(9) comprises three quasi-linear ordinary differential equations relating the four functions of time ε_{yy}^E, ε_{yy}^I, ϕ, and ψ. A fourth equation to complete the formulation is now obtained from the creep constitutive law.

A standard model for isochoric creep at high homologous temperature involves a creep strain rate $d\varepsilon_{ij}^I/dt$ that is related to stress σ_{ij} via $d\varepsilon_{ij}^I/dt=3C/2(\sigma_{\mathrm{eff}})^{n-1}s_{ij}$ where C and n are material creep parameters, [11]. Here $\mathbf{s}$ is the stress deviator $s_{ij}=\sigma_{ij}-(\sigma_{kk}/3)\delta_{ij}$, and σ_{eff} is the effective stress

$$\sigma_{\mathrm{eff}}=((\sigma_{xx}-\sigma_{yy})^2+(\sigma_{yy}-\sigma_{zz})^2+(\sigma_{zz}-\sigma_{xx})^2+6(\sigma_{xy}^2+\sigma_{yz}^2+\sigma_{zx}^2))^{1/2}/\sqrt{2}.$$

For uniaxial load this reduces to $d\varepsilon^I/dt=C\sigma^n$, enabling the determination of C and n via standard experimental procedures. For the states of equibiaxial stress encountered in the present planar setting, the effective stress is $\sigma_{\mathrm{eff}}=\sigma_{yy}=E/(1-\nu)\varepsilon_{yy}^E$, and the stress deviator components are $s_{xx}=-2/3\sigma_{yy}$, $s_{yy}=1/3\sigma_{yy}$. This then gives

$$\dot{\varepsilon}_{yy}^I=\frac{C}{2}\left(\frac{E}{(1-\nu)}\varepsilon_{yy}^E\right)^n, \tag{10}$$

as an additional differential equation to complement the previous set of Eqs. (7)–(9) relating ε_{yy}^E, ε_{yy}^I, ϕ, and ψ.

It remains to specify rules for the determination of $D_{\mathrm{oxy}}(t)$ and $D_{\mathrm{pest}}(t)$. Recall that $D_{\mathrm{pest}}(t)=H$ prior to shedding. Shedding begins once a failure criterion is met and the subsequent determination of the now free boundary $D_{\mathrm{pest}}(t)$ involves reference to this failure criterion. Section 3 gives the treatment prior to shedding and Section 4 gives the treatment during shedding.

The specification of $D_{\mathrm{oxy}}(t)$ is most naturally accomplished with respect to the current configuration location $d_{\mathrm{oxy}}(t)=(1-2\varepsilon_{yy}^I(t))D_{\mathrm{oxy}}(t)$. We shall eventually consider an example

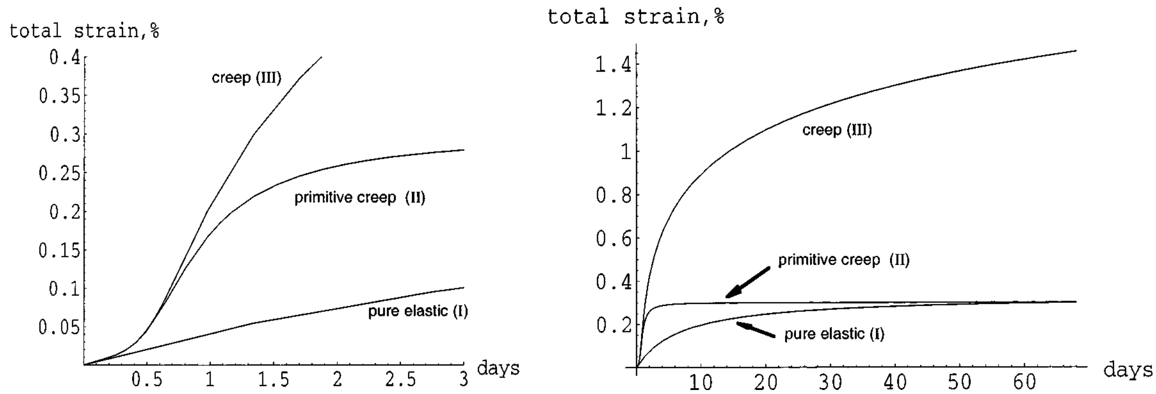


Fig. 1 Overall in-plane strain $\varepsilon_{yy}=\hat{\varepsilon}_{yy}(\times 100)$ is the same in the metal and the oxide due to displacement continuity (Eq. (1)). After the first few days, accounting for the additional oxide strain due to prior creep (III) gives rise to significantly larger total strain than either an elastic analysis (I) or the primitive analysis that neglects prior creep (II). In particular, the total strain under (III) is not bounded by the value $\alpha_y=0.003$, the asymptote for (I) and (II).

where the corrosion front advance with respect to the metal in the current configuration is described by a rate parameter μ so that $\dot{d}_{\text{oxy}}(t)=-\mu$.

In general, descriptions of oxidation front movement in terms of chemical species diffusion can lead to mathematical issues that require detailed treatment even in the absence of the creep mechanisms under study here, [12]. Rate factors such as μ are dependent on the local electrochemical environment (voltage, temperature, etc.) and the influence of the oxide as a corrosion shield. If shielding is significant, then the corrosion rate depends on the oxide layer thickness. If, for example, the corrosion rate involves a contribution that is proportional to the concentration of the diffusing oxygen species at the oxidation front, and a linear concentration gradient is in place with respect to distance from the shedding front, then the rate factor involves a contribution that is inversely proportional to the oxide thickness $d_{\text{pest}}(t)-d_{\text{oxy}}(t)$. Prior to shedding, any such contribution provides a continual slowdown in the corrosion rate (which essentially goes like $1/\sqrt{t}$ with some possible departure due to oxidation expansion). Once shedding begins, this slowdown is abated, and the determination of the corrosion rate would then become coupled to the shedding analysis. Evans [1] can be consulted for additional insight into the associated metallurgical issues. For the purposes of the examples presented in what follows, these effects will not be considered, and the rate factor μ will be regarded as constant. This is in keeping with the previously mentioned numerical studies of oxidation induced creep, [9,10].

3 Creep Prior to Oxide Shedding

Initially, the oxide expansion gives rise to a tensile stress in the metal and a compressive stress in the oxide. The stress in the oxide is dependent on X. For any location X this stress remains compressive until the in-plane expansion matches the natural (stress-free) length of the oxide at the location under consideration. Since the oxide at the surface $X=H$ was the first to transform (it experiences no prior creep) it has the smallest natural length. Consequently, the oxide at the surface is the first to bear a tensile load. The brittle nature of many oxides often renders them incapable of sustaining a significant tensile load and we regard this as leading to shedding. In this section we analyze the creep process prior to any such tensile failure induced shedding. Accordingly, $D_{\text{pest}}(t)=H$, $T_D(D_{\text{pest}}(t))=T_D(H)=0$, $\varepsilon^I_{yy}(T_D(D_{\text{pest}}(t)))=0$, so that the integral definition of $\phi(t)$ and $\psi(t)$ prior to shedding involves the full range of prior times $[0,t]$. Namely, Eqs. (7) and (8) give

$$\dot{\phi}(t)=-\varepsilon^I_{yy}(t)\dot{D}_{\text{oxy}}(t),\quad \phi(t)=-\int_0^t \varepsilon^I_{yy}(t)\dot{D}_{\text{oxy}}(t)dt,$$

$$\dot{\psi}(t)=-(\varepsilon^I_{yy}(t))^2\dot{D}_{\text{oxy}}(t),\quad \psi(t)=-\int_0^t (\varepsilon^I_{yy}(t))^2\dot{D}_{\text{oxy}}(t)dt.$$

In particular, initial conditions on ε^E_{yy}, ε^I_{yy}, ϕ, ψ are given by

$$\varepsilon^E_{yy}(0)=0,\quad \varepsilon^I_{yy}(0)=0,\quad \phi(0)=0,\quad \psi(0)=0.$$

We have numerically solved the set of four differential Eqs. (7)–(10) subject to these specializations and initial conditions for the case in which $\dot{d}_{\text{oxy}}(t)=-\mu$ so that $D_{\text{oxy}}(t)=(H-\mu t)/(1-2\varepsilon^I_{yy}(t))$. Parameter values employed in this example are representative of a high-temperature metal/oxide system using units of time in days, length in mm, force in Newtons: $H=0.0635$, $\mu=9.26\times 10^{-4}$, $E=1.67\times 10^4$, $\nu=0.434$, $\hat{E}=2.45\times 10^5$, $\hat{\nu}=0.25$, $\alpha_y=0.003$, $C=3.93\times 10^{-6}$, $n=3.78$. The transformation strain value $\alpha_y=0.003$ gives results that calibrate well with creep rates in proprietary studies. It is conjectured, [9,13], that only a small fraction of the transformation strain is expressed in the in-plane directions, i.e., $\alpha_y \ll \alpha_x$.

The curve labeled (III) in Fig. 1 gives the in-plane strain $\varepsilon_{yy}(t)$ for the numerical solution to Eqs. (7)–(10). It can be read from this figure that $\varepsilon_{yy}(t)<0.003=\alpha_y$ for $t<1.35$ and $\varepsilon_{yy}(t)>0.003=\alpha_y$ for $t>1.35$. It is the effect of creep prior to oxidation that permits ε_{yy} to exceed α_y. For comparison sake, two additional curves are shown. Curve (I) is the result of a pure elastic analysis in which $\varepsilon^I_{xx}=\varepsilon^I_{yy}=0$. This analysis gives closed-form algebraic expressions. Curve (II) is for an analysis that incorporates creep in the metal, but which does not include the effect of prior creep in computing the inelastic strain in the oxide, i.e., $\hat{\varepsilon}^I_{yy}=\alpha_y$, $\hat{\varepsilon}^I_{xx}=\alpha_x$. Thus the analysis giving (II) involves no explicit prior time reference and is similar in spirit to that of [9,10]. Such a "primitive creep analysis" is reasonable for some short period of time during which the creep strain is small compared to the corresponding transformation strain.

In both (I) and (II) the strain $\varepsilon_{yy}(t)$ does not exceed α_y, which provides the large time asymptote for the corresponding ε_{yy} curves. Figure 2 shows the stress in the metal for all three treatments. A pure elastic treatment involves a continually increasing biaxial tensile stress σ_{yy}. This stress is relieved by creep relaxation in both creep treatments with treatment (II) providing the quickest relief. This is because the inelastic strain in the oxide, which is driving the creep, is artificially low in the treatment (II)

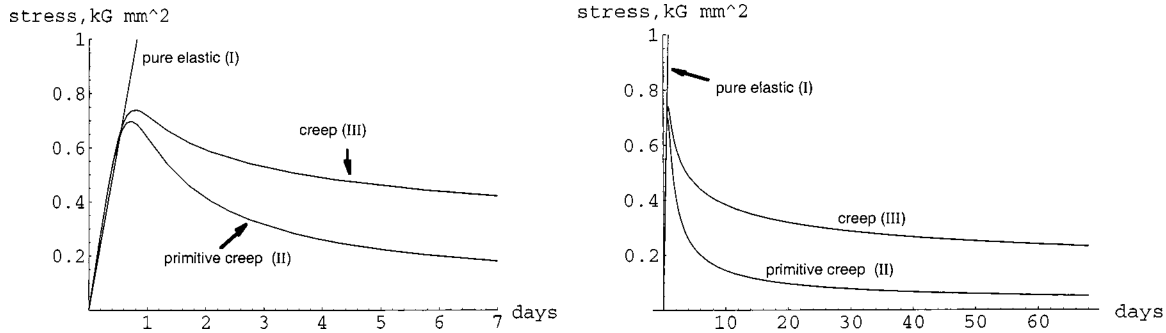


Fig. 2 Creep allows the stress in the metal σ_{yy} to relax from that predicted by a pure elastic analysis (I). Accounting for oxide strain magnification due to prior creep (III) gives higher stress than an analysis that ignores this effect (II).

due to the neglect of prior creep. The treatment of interest here (III) provides a biaxial tensile stress σ_{yy} that is intermediate between the simplified treatments (I) and (II).

The variation of the stress $\hat{\sigma}_{yy}$ with position in the oxide is an essential feature of the present treatment (III). All of the stresses $\hat{\sigma}_{yy}$ for these different treatments (I,II,III) match the pure elastic value $-\hat{E}/(1-\hat{\nu})\alpha_y$ at $t=0$. The stress $\hat{\sigma}_{yy}$ in the simpler treatments (I) and (II) is uniform and tends to zero as the metal is consumed. However, for the treatment (III) of interest here, the stress $\hat{\sigma}_{yy}$ is nonuniform with maximum compression at $X=D_{\text{met}}(t)$. At the outer surface, the stress $\hat{\sigma}_{yy}$ transitions from compressive to tensile at $t=1.35$ corresponding to the time at which $\varepsilon_{yy}=\alpha_y$. Figure 3 shows the $\hat{\sigma}_{yy}$ stress variation with X for increasing five-day intervals. After oxidation, the stress $\hat{\sigma}_{yy}$ at each material location X is increasing with time because the total strain ε_{yy} is approaching the locked-in portion $\hat{\varepsilon}^I_{yy}(X)=\alpha_y+\varepsilon^I_{yy}(T_D(X))$. Accordingly, each such material location transitions from compressive $\hat{\sigma}_{yy}$ to tensile $\hat{\sigma}_{yy}$ at the time t that solves $\alpha_y+\varepsilon^I_{yy}(T_D(X))=\varepsilon_{yy}(t)$.

4 Oxide Shedding

The numerical example of the previous section (e.g., Fig. 3) shows that $\hat{\sigma}_{yy}(D_{\text{oxy}}(t),t)<0$ with $\partial\hat{\sigma}_{yy}/\partial t>0$ and $\partial\hat{\sigma}_{yy}/\partial X>0$ on $D_{\text{oxy}}(t)<X<H$. In particular, each location X eventually obeys $\hat{\sigma}_{yy}>0$. If the oxide is unable to sustain such tensile loading, then the analysis must be modified as described next.

Since the stress in the oxide is completely determined by the equibiaxial stress $\hat{\sigma}_{yy}(X)$, such failure can be characterized by a critical equibiaxial stress $\hat{\sigma}_{\text{shed}}\geqslant 0$ that is regarded here as a material property. This section presents the analytical development for describing this shedding process and gauging its effect on system growth. Analytically, the reference configuration location of the shedding front $D_{\text{pest}}(t)$ becomes an unknown that must be determined from the condition

$$\hat{\varepsilon}^E_{yy}(D_{\text{pest}}(t),t)=\frac{(1-\hat{\nu})}{\hat{E}}\hat{\sigma}_{\text{shed}}. \tag{11}$$

Shedding begins at the time of first failure $t=t_{ff}$ as determined on the basis of the pre-shedding analysis via $(1-\hat{\nu})/\hat{E}\hat{\sigma}_{\text{shed}}=\hat{\varepsilon}^E_{yy}(H,t_{ff})=\varepsilon^E_{yy}(t_{ff})+\varepsilon^I_{yy}(t_{ff})-\alpha_y$. The initial values of $\varepsilon^E_{yy}(t_{ff})$, $\varepsilon^I_{yy}(t_{ff})$, $\phi(t_{ff})$, $\psi(t_{ff})$ for the shedding analysis are then determined from the pre-shedding analysis via continuity at $t=t_{ff}$.

The location $D_{\text{pest}}(t)$ can be tracked in terms of an additional, and as yet undetermined, auxiliary function $\tau(t)$ defined via

$$\tau(t)\equiv T_D(D_{\text{pest}}(t)),\quad \Rightarrow D_{\text{pest}}(t)=D_{\text{oxy}}(\tau(t)).$$

Note that $\tau(t)<t$ and that $\tau(t)$ is the time at which the oxide location shed at t had previously oxidized from the original metal. In particular, $\tau(t_{ff})=0$. Evaluating (4) at $X=D_{\text{pest}}(t)$ and eliminating $\hat{\varepsilon}^E_{yy}(D_{\text{pest}}(t),t)$ with the aid of Eq. (11) now gives

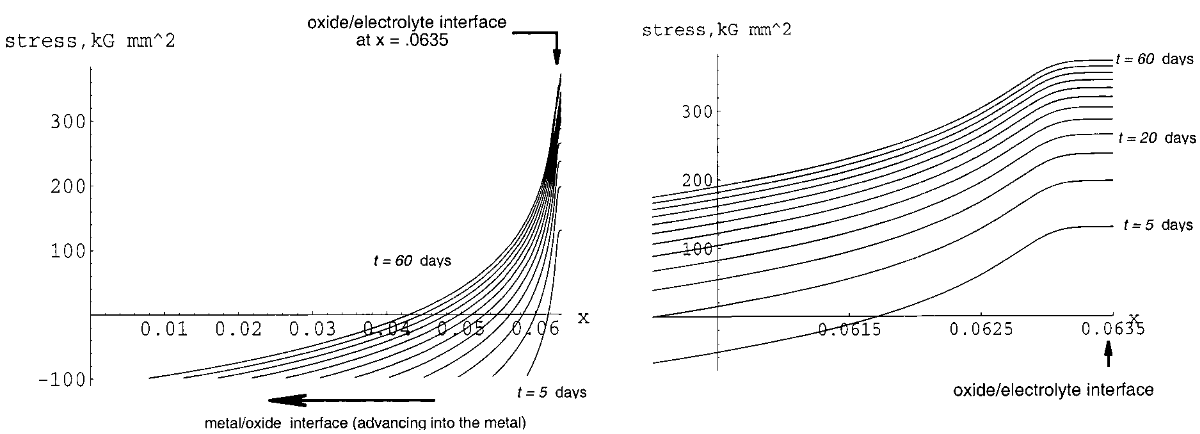


Fig. 3 Variation of the oxide stress $\hat{\sigma}_{yy}$ through the oxide thickness as the oxide layer progresses into the base metal. Different curves represent different five day intervals, ranging from $t=5$ to $t=60$. Graph on right provides magnification near the external surface, $x=0.0635$.

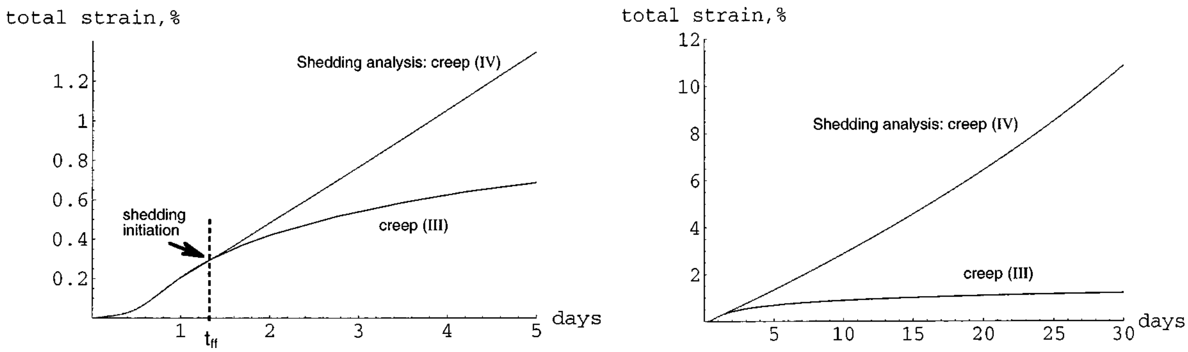


Fig. 4 Once shedding begins at t_{ff} the strain increases dramatically (IV) compared to a situation in which oxide under tension does not fail (III). This strain increase for (IV) is due to the loss of tensile load carrying capacity in the oxide, which increases the tensile stress in the metal as shown in Fig. 5.

$$\frac{(1-\hat{\nu})}{\hat{E}}\hat{\sigma}_{\rm shed}=\varepsilon^E_{yy}(t)+\varepsilon^I_{yy}(t)-\alpha_y-\varepsilon^I_{yy}(\tau(t)).$$

The rate form of this equation, in conjunction with the four previous rate Eqs. (7), (8), (9), and (10) under the replacement $D_{\rm pest}(t)=D_{\rm oxy}(\tau(t))$, now provides five quasi-linear ordinary differential equations relating the five functions of time ε^E_{yy}, ε^I_{yy}, ϕ, ψ, and τ:

$$\dot{\varepsilon}^I_{yy}=\frac{C}{2}\left(\frac{E}{(1-\nu)}\varepsilon^E_{yy}\right)^n,$$

$$\begin{aligned}
&(-2\beta D_{\rm oxy}\varepsilon^E_{yy}+(D_{\rm oxy}\circ\tau)-D_{\rm oxy}-2\phi-2\chi((\varepsilon^E_{yy}+\varepsilon^I_{yy}-\alpha_y)((D_{\rm oxy}\circ\tau)-D_{\rm oxy})-\phi))\dot{\varepsilon}^I_{yy}\\
&\quad+(\beta D_{\rm oxy}(1-2\varepsilon^I_{yy})+(D_{\rm oxy}\circ\tau)-D_{\rm oxy}-2\phi-2\chi((\varepsilon^E_{yy}+\varepsilon^I_{yy}-\alpha_y)((D_{\rm oxy}\circ\tau)-D_{\rm oxy})-\phi))\dot{\varepsilon}^E_{yy}\\
&\quad+(-2\varepsilon^E_{yy}-2\varepsilon^I_{yy}-1+2\chi(\varepsilon^E_{yy}+\varepsilon^I_{yy}-\alpha_y))\dot{\phi}+(2-\chi)\dot{\psi}+(\varepsilon^E_{yy}+\varepsilon^I_{yy}-\alpha_y-\chi(\varepsilon^E_{yy}+\varepsilon^I_{yy}-\alpha_y)^2)(\dot{D}_{\rm oxy}\circ\tau)\dot{\tau}\\
&=(-\beta\varepsilon^E_{yy}+2\beta\varepsilon^E_{yy}\varepsilon^I_{yy}+\varepsilon^E_{yy}+\varepsilon^I_{yy}-\alpha_y-\chi(\varepsilon^E_{yy}+\varepsilon^I_{yy}-\alpha_y)^2)\dot{D}_{\rm oxy},
\end{aligned}$$

$$\dot{\phi}-(\dot{\varepsilon}^I_{yy}\circ\tau)(\dot{D}_{\rm oxy}\circ\tau)\dot{\tau}=-\varepsilon^I_{yy}\dot{D}_{\rm oxy},$$

$$\dot{\psi}-(\dot{\varepsilon}^I_{yy}\circ\tau)^2(\dot{D}_{\rm oxy}\circ\tau)\dot{\tau}=-(\dot{\varepsilon}^I_{yy})^2\dot{D}_{\rm oxy},$$

$$\dot{\varepsilon}^I_{yy}+\dot{\varepsilon}^E_{yy}-(\dot{\varepsilon}^I_{yy}\circ\tau)\dot{\tau}=0.$$

These five equations are to be solved for $t>t_{ff}$ subject to $\tau(t_{ff})=0$ and ε^E_{yy}, ε^I_{yy}, ϕ, ψ continuity at $t=t_{ff}$, where the prescription for $D_{\rm oxy}(t)$ continues from the pre-shedding analysis. Note also that

$$\phi(t)=-\int_{\tau(t)}^{t}\varepsilon^I_{yy}(t)\dot{D}_{\rm oxy}(t)dt,$$

$$\psi(t)=-\int_{\tau(t)}^{t}(\varepsilon^I_{yy}(t))^2\dot{D}_{\rm oxy}(t)dt.$$

We have extended the numerical solution presented in Section 3 into the shedding regime for the particular case $\hat{\sigma}_{\rm shed}=0$, i.e., an oxide that is not capable of supporting any tensile load whatsoever. From the analysis of the previous section it follows that $t_{ff}=1.35$. Figure 4 provides a comparison between ε_{yy} for the present shedding analysis (IV) with ε_{yy} for the analysis in which shedding is ignored (III). Shedding increases both the strain ε_{yy} and the stress σ_{yy} (Fig. 5). This is because the tensile stress that would have been carried by the shed oxide must be carried instead by the metallic core, increasing σ_{yy} and hence $\dot{\varepsilon}^I_{yy}$. In the numerical example, stress relaxation had already ensued prior to shedding. The onset of shedding arrests this stress decline, after which the stress σ_{yy} is found to increase at a mild rate. Figure 6 charts the change in oxide thickness with time. Prior to shedding, the oxide thickness increases monotonically. Immediately after shedding initiates, this thickness exhibits a rapid drop, followed by a brief transient prior to a period of relatively slow decline in thickness. The rate of the oxide thickness decline after this transient is found to be more than an order of magnitude lower than the rate of corrosion advance.

5 Concluding Remarks

A treatment has been presented for creep that is driven solely by oxidation. The main physical assumptions of the treatment concern the mechanical behavior of the base metal, the oxidation process that transforms metal to oxide, and the mechanical properties of the resulting oxide. The base metal is ductile and described by a conventional viscoplastic model. It does not fail in the sense that it is always able to support load. Oxidation corrosion proceeds via the movement of an interface into the base metal. Here this rate of advance is specified a priori. Oxidation is localized at this front and the material experiences volumetric expansion upon conversion from metal to oxide. Oxide is regarded as brittle elastic, failing in tension but not in compression. Failed oxide ceases to support any mechanical load and can be regarded as physically scaling off.

Analyzing this process via an associated boundary value problem requires a description for inelastic strain in the oxide that

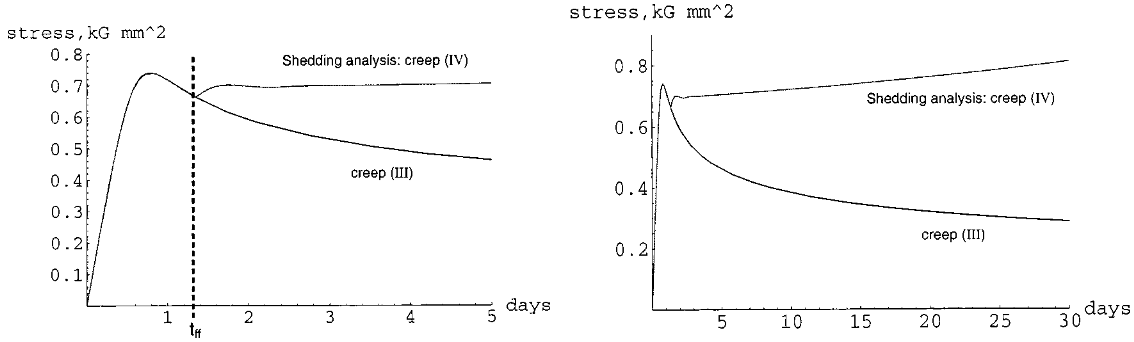


Fig. 5 In this example, oxide shedding begins after stress relaxation has already begun in the metal. Shedding leads to additional loading of the metal, and the post-shedding stress experiences a mild increase with time.

captures both creep prior to oxidation and volumetric strain of the metal-to-oxide transformation. This gives rise to the first new feature of the present treatment: tensile stress at the outer surface of the oxide layer. This tensile stress first appears at a finite time after the beginning of oxidation and leads to the second new feature of the present treatment: a shedding front that is not specified a priori.

The boundary value problem for the case of a flat-plate geometry has a different formulation before and after the initiation of the shedding front. Prior to the development of the shedding front, the associated boundary value problem can be reduced to a system of four ordinary differential equations for four dependent variables with time as the independent variable. The four dependent variables are: an elastic strain in the metal, a creep strain in the metal, and two auxiliary functions that describe the integrated effect of creep prior to oxidation. Numerical solution determines the time at which the oxide on the external surface first becomes tensile, which in turn initiates the shedding front.

After the development of the shedding front, the associated boundary value problem can be reduced to a system of five ordinary differential equations for five dependent variables. Four of these five dependent variables are the four that describe the process prior to failure. The new dependent variable is a retarded time argument that describes the time interval between oxide formation and oxide shedding. Four of the five differential equations are continuations of the previous set and include new terms involving the four original dependent variables evaluated at the retarded time. The fifth differential equation is a rate form of the oxide failure condition. The system can be solved numerically. We find that the main features of the numerical solution to this boundary value problem are as follows.

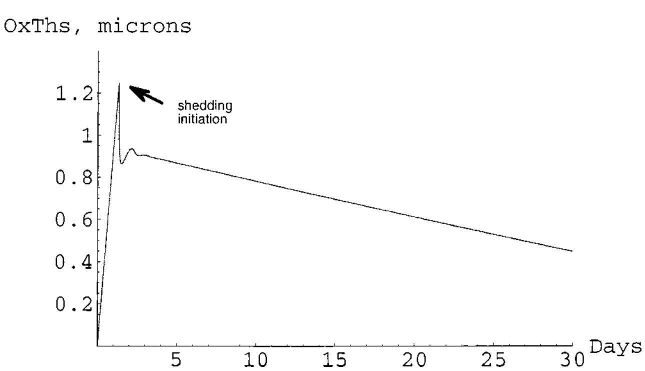


Fig. 6 The oxide thickness increases at the corrosion rate prior to shedding. Once shedding begins, the oxide thickness experiences an abrupt transient prior to a period of only mild thickness decrease.

1 The volumetric expansion gives rise to stress. Initially this stress is tensile in the metal and compressive in the oxide. This is well known.

2 As the oxidation front proceeds into the base metal, the volumetric expansion strain is leveraged against the strain due to prior creep. For the simple flat-plate geometry, the resulting inelastic strain in the oxide is position dependent and time independent. Higher inelastic strains correlate with later locations of oxidation. The stress at a fixed location in the oxide is most compressive upon first formation.

3 The compressive stress at each location in the oxide diminishes with time due to creep elongation of the metal. When the creep elongation matches the inelastic strain for a particular oxide location, then this oxide location is fully stress relieved. Additional creep elongation would put such an oxide site into tension. The external surface of the oxide is the first to become tensile.

4 If the oxide is unable to support tensile load then a failure front forms in the oxide. This front proceeds into the interior of the oxide beginning from the external surface and can be regarded as a location for the shedding of oxide pest. This shedding front lags the oxidation front, but could eventually advance at approximately the same rate as the oxidation front, leading to an essentially constant thickness oxide layer. This shedding increases the tensile stress in the metal and so increases the creep rate.

While the first of these four features is widely appreciated, the remaining three features have received much less attention especially with regard to their interaction. In particular, the formation and advance of a shedding front in the oxide, while creep is ongoing in the metal, gives rise to interactions that can treated by formulating and solving a boundary value problem of the type presented here. Presumably, similar treatments could apply to geometries other than the flat plate presented here.

References

[1] Evans, H. E., 1995, "Stress Effects in High Temperature Oxidation of Metals," Int. Met. Rev., **40**, pp. 1–40.
[2] Lagoudas, D. C., Ma, X., Miller, D. A., and Allen, D. H., 1995, "Modeling of Oxidation in Metal Matrix Composites," Int. J. Eng. Sci., **33**, pp. 2327–2343.
[3] Srolovitz, D. J., and Ramanarayanan, T. A., 1984, "An Elastic Analysis of Growth Stresses During Oxidation," Oxid. Met., **22**, pp. 133–146.
[4] Grinfeld, M. A., 1994, "Stress Corrosion and Stressed Induced Surface Morphology of Epitaxial Films," Scanning Microsc., **8**, pp. 869–882.
[5] Bull, S. J., 1998, "Modeling of Residual Stress in Oxide Scales," Oxid. Met., **49**, pp. 1–17.
[6] Evans, H. E., 1988, "Cavity Formation and Metallurgical Changes Induced by Growth of Oxide Scale," Mater. Sci. Technol., **4**, pp. 1089–1098.
[7] Evans, A. G., and Cannon, R. M., 1989, "Stresses in Oxide Films and Relationships With Cracking and Spalling," Mater. Sci. Forum, **43**, pp. 243–268.

[8] Entchev, P. B., Lagoudas, D. C., and Slattery, J. C., 2001, ''Effects of Non-Planar Geometries and Volumetric Expansion in the Modeling of Oxidation in Titanium,'' Int. J. Eng. Sci., **39**, pp. 695–714.
[9] Bernstein, H. L., 1987, ''A Model for the Oxide Growth Stress and Its Effect on the Creep of Metals,'' Metall. Trans. A, **18**, pp. 975–986.
[10] Touati, A., Roelandt, J. M., Armanet, F., Lambertin, M., and Beranger, G., 1993, ''Un Modele pour le Calcul des Deformations et des Contraintes Residuelles dans les Couches d'Oxyde lors de Chargements Thermiques a Regime Variable,'' J. Phys. IV, **C9**, pp. 1023–1029.
[11] Kraus, H., 1980, *Creep Analysis*, John Wiley and Sons, New York.
[12] Lagoudas, D. C., and Ding, Z., 1998, ''Numerical Computation of Metal Oxidation Problems on Bounded Domains,'' Int. J. Eng. Sci., **36**, pp. 367–381.
[13] Holmes, D. R., and Pascoe, R. T., 1972, ''Strain/Oxidation Interactions in Steels and Model Alloys,'' Werkst. Korros., **23**, pp. 859–870.

Z. Dursunkaya
Professor,
Mechanical Engineering Deparatment,
Middle East Technical University,
Inonu Bulvari,
Ankara 06531, Turkey

S. Nair
Professor,
Mechancial, Materials and Aerospace
Engineering Department,
Illinois Institute of Technology,
Engineering 1 Building,
10 West 32nd Street,
Chicago, IL 60616-3793
Fellow ASME

Solidification of a Finite Medium Subject to a Periodic Variation of Boundary Temperature

The motion of a solid-liquid interface in a finite one-dimensional medium, subject to a fluctuating boundary temperature, is analyzed. The fluctuations are assumed to be periodic. The solution method involves a semi-analytic approach in which, at any given time, the spatial temperature distributions are represented in infinite series. The effect of the solid, liquid Stefan numbers and the unsteady boundary temperature variation is investigated. The results showed a retrograde motion of the solidification front for large liquid Stefan numbers. [DOI: 10.1115/1.1604836]

Introduction

A phase change in an initially single-phase medium can be started by changing the temperature at the boundaries to a value which is higher or lower than the melting temperature depending on the state of the initial phase. The newly created phase advances into the original phase forming an interface between the two regions. The interface remains at the melting temperature but its location and velocity of propagation are not known a priori, and are a part of the solution of the problem. The transient interface location is a function of the physical properties of the two phases, as well as the temperature gradients at the interface and the latent heat of fusion. The diffusion equations in the two domains are coupled through the interface temperature, interface location, and the interface heat fluxes. The moving nature of the interface and the coupling of the energy equations through the interface relations render the mathematical formulation of the problem complex. The location of the interface forms a part of the solution to the problem, and the interface energy balance in turn is a differential equation for the interface location. Therefore, the formulation is comprised of two second-order partial differential equations for the energy equations, and a first-order differential equation for the interface equation, for constant physical properties.

The complex nature of the governing equations restricts the general analytical solution of the problem to a limited number of specific cases. Neumann's solution (Carslaw and Jaeger [1]) to the melting-freezing problem in a semi-infinite medium is the oldest solution to a phase-change moving boundary problem. This problem considers the diffusion into a medium, which is initially at a constant temperature, different than the melting temperature, and the boundary temperature is suddenly altered by a constant amount, which is higher or lower than the melting temperature. Neumann's solution has been extended to problems with constant temperature boundary condition in cylindrical and spherical coordinate systems. In these solutions the initial phase is at a uniform temperature at the time the boundary condition is changed. Other analytical solutions of moving boundary problems in semi-infinite domains with arbitrary initial and boundary conditions have been found by Tao [2,3]. So far general analytical solutions of the problem for finite geometries have not been published in the literature.

Various approximate analytical and numerical techniques have been used in the solution of moving boundary problems. Although most of the conventional methods can be applied to these problems, the presence of the moving boundary results in complex formulations, whose solutions are difficult to obtain. Methods used for solving moving boundary problems include the approximate integral method, enthalpy formulation, variational formulation, formulation with moving heat sources, and the discretization methods, such as finite difference and finite element methods. In addition, perturbation techniques have also been used. Surveys of these methods have been published in a number of books (Crank [4], Rubinstein [5], Ockendon and Hogkins [6], and Wilson, Solomon, and Boggs [7]) and survey articles (Fukusako and Seki [8]).

Weinbaum and Jiji [9] applied the singular perturbation theory for the problem in a finite slab, where the medium initially is not at the melting temperature. The location of the interface was found and the results, however, are valid for small Stefan numbers. Charach and Zoglin [10] used the heat balance integral method and time-dependent perturbation theory to solve the same problem. They found the interface location and temperature distributions for small liquid and solid Stefan numbers. Small Stefan numbers correspond to small temperature differences, and large temperature differences require higher-order perturbation expansions, which are complicated. This limits the utilization of perturbation techniques to small temperature differences. Nair [11] attempted an integral equation formulation of the problem with a limited set of boundary conditions. Rizwan-uddin [12] solved a one-dimensional phase-change problem subject to periodic boundary conditions. A semi-analytical numerical scheme was used in a one-domain problem. Results were reported for a number of different frequencies of boundary temperatures and Stefan numbers. Dursunkaya and Nair [13] solved the solidification of a two-phase problem, where a Fourier series expansion was used to represent the spatial temperature distribution. The resulting infinite system of ordinary differential equations is solved iteratively for the interface location and the temperature distributions. One advantage of this method is that the interface location can be computed without the detailed information about the temperature distribution. This approach will be used in the current study.

Formulation

The problem involves the formulation and solution of one-dimensional solidification in a finite slab as shown in Fig. 1. The medium is initially a liquid (at a temperature above freezing), and is suddenly exposed to a temperature below freezing on one surface, which also varies in time. The energy equations in the newly formed solid and the originally existing liquid phases are given by

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 4, 2002; final revision, Feb. 5, 2003. Associate Editor: D. A. Siginer. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

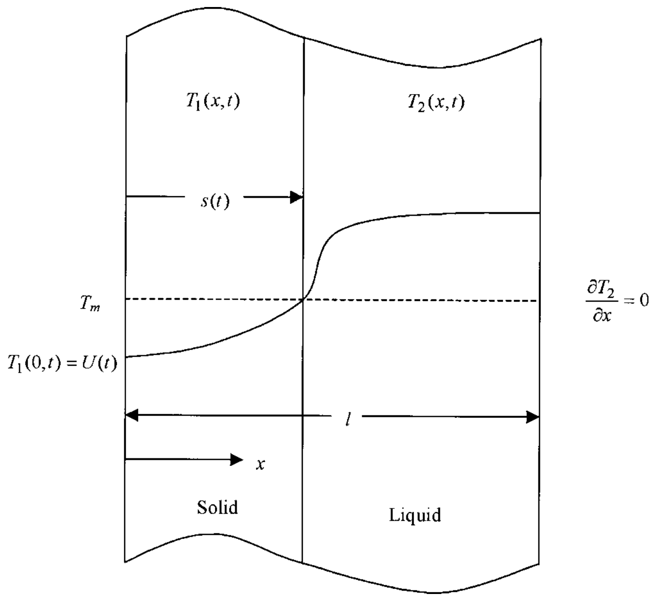


Fig. 1 Solidification domain and temperatures

$$\alpha_1 \frac{\partial^2 T_1}{\partial x^2} = \frac{\partial T_1}{\partial t}, \quad 0 < x < s(t), \tag{1a}$$

$$\alpha_2 \frac{\partial^2 T_2}{\partial x^2} = \frac{\partial T_2}{\partial t}, \quad s(t) < x < l, \tag{1b}$$

where T is the temperature, α is the thermal diffusivity, s is the interface location, and l is the length of the slab. The subscripts 1 and 2 denote the newly formed solid, and the original liquid phases, respectively. The boundaries of the medium are subject to the following initial and boundary conditions.

$$\begin{gathered} t=0: \quad s(0)=0, \quad T_2(x,0)=V \\ x=0: \quad T(0,t)=U(t) \geqslant U_0; \quad x=l: \quad \frac{\partial T}{\partial x}(l,t)=0, \end{gathered} \tag{2}$$

where U_0 is the minimum temperature encountered at the boundary. At the phase-change interface the temperature is equal to the fusion temperature, T_m,

$$T_1(s,t) = T_2(s,t) = T_m, \tag{3}$$

and a heat balance at the interface between the conduction in the two phases and the liberation of heat due to phase change gives

$$k_1 \frac{\partial T_1}{\partial x} - k_2 \frac{\partial T_2}{\partial x} = \rho L \frac{ds}{dt}, \tag{4}$$

where k is the thermal conductivity, ρ is the density, and L is the latent heat of fusion. The following dimensionless quantities are introduced:

$$\begin{gathered} \xi = x/l, \quad \sigma = s/l, \quad \tau = \alpha_1 t/l^2, \quad V_0 = (V - T_m)/(T_m - U_0) \\ \theta_1(x,t) = \frac{T_1 - U(t)}{T_m - U_0} - \frac{T_m - U(t)}{T_m - U_0} \frac{x}{s}, \\ \theta_2(x,t) = \frac{T_2 - T_m}{V - T_m}, \quad \psi(t) = \frac{T_m - U(t)}{T_m - U_0}. \end{gathered} \tag{5}$$

V_0 and $\psi(t)$ are the dimensionless initial and boundary temperatures, respectively. Substitution in the energy equations gives

$$\frac{\partial^2 \theta_1}{\partial \xi^2} = \frac{\partial \theta_1}{\partial \tau} - \frac{\xi}{\sigma^2} \frac{d\sigma}{d\tau} \psi - \left(1 - \frac{\xi}{\sigma}\right) \frac{d\psi}{d\tau}, \tag{6a}$$

$$A \frac{\partial^2 \theta_2}{\partial \xi^2} = \frac{\partial \theta_2}{\partial \tau}, \tag{6b}$$

where $A = \alpha_2/\alpha_1$, is the ratio of the thermal diffusivities. The energy Eqs. (6a) and (6b) are subject to the following boundary and interface conditions:

$$\begin{gathered} \tau=0: \quad \sigma(0)=0, \quad \theta_2(\xi,0)=V_0 \\ \xi=0: \quad \theta_1(0,\tau)=0; \quad \xi=1: \quad \frac{\partial \theta_2}{\partial \xi}(1,\tau)=0 \\ \xi=\sigma: \quad \theta_1(\sigma,t)=\theta_2(\sigma,\tau)=0. \end{gathered} \tag{7}$$

The heat balance at the interface, using dimensionless quantities, becomes

$$\frac{d\sigma}{d\tau} = \mathrm{St}_1\left(\psi + \frac{\partial \theta_1}{\partial \xi}\right) - A\mathrm{St}_2 \frac{\partial \theta_2}{\partial \xi}, \tag{8}$$

where, $\mathrm{St}_1 = C_1(T_m - U_0)/L$ is the solid and $\mathrm{St}_2 = C_2(V - T_m)/L$ is the liquid Stefan number, with C representing the heat capacity. The solution methodology is based on fitting an infinite series of functions for the temperature distributions. Specifically,

$$\begin{gathered} \theta_1(\xi,\tau) = \sum_{m=1}^{\infty} \beta_{1m}(\tau)\phi_{1m}(\xi,\sigma(\tau)), \\ \theta_2(\xi,\tau) = \sum_{m=1}^{\infty} \beta_{2m}(\tau)\phi_{2m}(\xi,\sigma(\tau)). \end{gathered} \tag{9}$$

Here the β-functions reflect the temporal variation, whereas the ϕ-functions are related to the spatial variation of the temperatures. Note that the ϕ's also account for a temporal effect due to the fact that they are functions of the interface location. In this study the β's are unknown functions and ϕ's are harmonic functions that satisfy the boundary and interface conditions, specifically,

$$\phi_{1m} = \sin\left(m\pi \frac{\xi}{\sigma}\right), \quad \phi_{2m} = \sin\left(\bar{m}\pi \frac{\xi - \sigma}{1 - \sigma}\right), \quad \bar{m} = m - 1/2. \tag{10}$$

This choice of functions form orthogonal sequences in $0 \leqslant \xi \leqslant \sigma$, and in $\sigma \leqslant \xi \leqslant 1$, respectively, that satisfy the boundary conditions. For different type of boundary conditions, orthogonal functions that satisfy the boundary conditions must be used. Although any orthogonal series could be used in the analysis, using harmonic functions in the Cartesian geometry has the advantage of being easier to manipulate. With the above choice of the functions for the temperature variation, the substitution of Eqs. (9) and (10) in energy Eqs. (6a) and (6b) gives two first-order ordinary differential equations for the unknown functions β's. Because of the fact that the differential equation is first order, the nature of the solution for the β functions is exponential. Specifically for the first iteration

$$\begin{aligned} \beta_{1m}^{(1)} = &- \frac{2(-1)^m}{m\pi\sqrt{\sigma}} \int_0^{\tau} \left(\psi \frac{\dot{\sigma}}{\sigma} - (-1)^m \frac{d\psi}{d\tau}\right)\sqrt{\sigma} \\ &\times \exp\left[-\int_{\tau'}^{\tau} \frac{m^2\pi^2}{\sigma^2} d\tau''\right] d\tau', \end{aligned} \tag{11a}$$

$$\beta_{2m}^{(1)} = -\frac{2}{\bar{m}\pi}\sqrt{\frac{1}{1-\sigma}} \exp\left[-\int_0^{\tau} \frac{A\bar{m}^2\pi^2}{(1-\sigma)^2} d\tau'\right] \times D_m, \tag{11b}$$

where the constants D_m are such that the initial temperature distribution in the liquid phase is satisfied, and the superscript denotes iteration. After the second iteration for β's,

$$\beta_{1m}^{(2)} = -\frac{2(-1)^m}{m\pi\sqrt{\sigma}} \int_0^\tau \left(\left(\psi + m^2\pi \sum_{\substack{n=1 \\ n\neq m}}^{\infty} \frac{n(-1)^n}{m^2-n^2} \beta_{1n}^{(1)} \right) \frac{\dot{\sigma}}{\sigma} - (-1)^m \frac{d\psi}{d\tau} \right) \sqrt{\sigma} \exp\left[-\int_{\tau'}^{\tau} \frac{m^2\pi^2}{\sigma^2} d\tau'' \right] d\tau', \quad (12a)$$

$$\beta_{2m}^{(2)} = -\frac{2}{\overline{m}\pi} \sqrt{\frac{1}{1-\sigma}} \exp\left[-\int_0^\tau \frac{A_2 \overline{m}^2 \pi^2}{(1-\sigma)^2} d\tau' \right] \times \left\{ D_m + 2\overline{m}^2 \int_0^\tau \frac{\dot{\sigma}}{1-\sigma} \times \sum_{\substack{n=1 \\ n\neq m}}^{\infty} \left(\frac{D_n}{\overline{m}^2 - \overline{n}^2} \times \exp\left[-\int_0^{\tau'} \frac{A(\overline{m}^2 - \overline{n}^2)\pi^2}{(1-\sigma)^2} d\tau'' \right] d\tau' \right) \right\}. \quad (12b)$$

The disadvantage lies in the fact that the convergence of the series is slow. Nevertheless, it was observed that the slow convergence of the harmonic functions does not create a problem in this case. The interface energy balance given in Eq. (8) written in terms of ϕ's and β's becomes

$$\frac{d\sigma}{d\tau} = \mathrm{St}_1 \left(\frac{\psi(\tau)}{\sigma} + \frac{\pi}{\sigma} \sum_{m=1}^{\infty} m(-1)^m \beta_{1m}^{(2)} \right) - A\mathrm{St}_2 \frac{\pi}{1-\sigma} \sum_{m=1}^{\infty} \overline{m} \beta_{2m}^{(2)}. \quad (13)$$

Having computed the β functions, this equation can be integrated to find the new interface location $\sigma(\tau)$. Further details of the analytical and numerical approach are similar to Dursunkaya and Nair [13].

In the current problem the boundary is subjected to a sinusoidal variation of temperature in the form,

$$U(t) = T_0 + E\sin(\omega t) = U_0 + E(1 + \sin(\omega t)), \quad (14)$$

where ω denotes the frequency. When nondimensionalized using, $\overline{\omega} = \omega l^2/\alpha_1$, Eq. (14) gives

$$\psi(\tau) = 1 - \gamma(1 + \sin(\overline{\omega}\tau)), \quad \gamma = E/(T_m - U_0), \quad (15)$$

where γ is the amplitude of boundary temperature variation.

Although the dimensionless temperatures defined in Eq. (9) are better suited for formulation and the solution of the governing differential equations, they are not convenient for plotting. For this purpose the dimensionless temperature $\varphi(\xi,\tau) = (T_i - U_0)/(T_m - U_0)$ is introduced, which, when used in the two phases renders the dimensionless interface temperature unity. In terms of the dimensionless quantities used in the formulation,

$$\varphi(\xi,\tau) = \theta_1(\xi,\tau) + \psi(\tau)\left(\frac{\xi}{\sigma} - 1 \right) + 1, \quad 0 \leq \xi \leq \sigma, \quad (16a)$$

$$\varphi(\xi,\tau) = V_0 \theta_2(\xi,\tau) + 1, \quad \sigma \leq \xi \leq 1. \quad (16b)$$

Results

The sensitivity of the numerical results to the number of terms used in the summation of the harmonic series was tested. For small time, the interface heat balance is dominated by the first term on the right-hand side of Eq. (13) that reflects the effect of the boundary temperature. Therefore, for small time using a small number of terms for the summation of the β functions (Eqs. (12*a*) and (12*b*)) does not adversely affect the computations for the interface location. For large time, due to the exponentially decaying temporal nature of the β functions, only the first few terms of the series for the temperatures (Eqs. (9) and (12)) contribute to the interface equation, hence the interface location. On the other hand, the effect of the number of terms used in the summation has a pronounced effect on the temperatures (Eq. (9)), especially for small time. In a typical run, approximately 10,000 terms are needed to obtain five-digit accuracy in temperatures for small time. This, nevertheless, is only needed if temperature distribution is to be plotted, since the interface location can be solved independent of the actual computation of temperature profiles, as explained above. In this study 250 terms were used in the summations. The comparison of the predictions of the current approach with the data available in the literature is given in [13] and the agreement was found to be good.

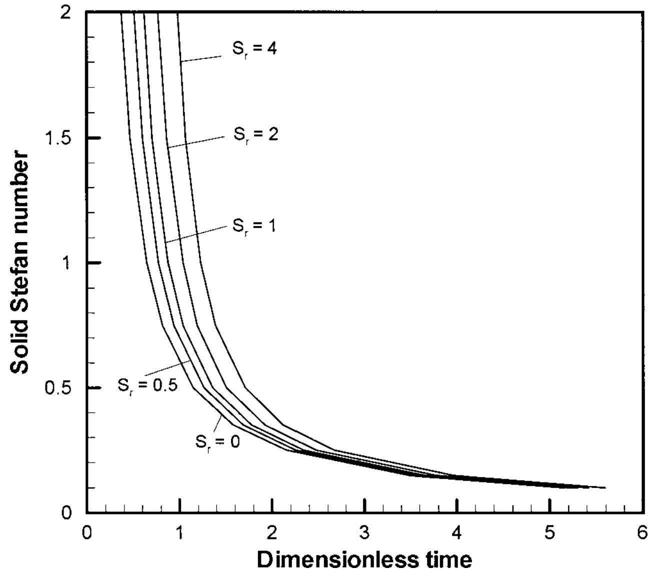


Fig. 2 Time for complete solidification in a finite slab

The effect of the boundary temperature variation, solid, and liquid Stefan numbers on solidification were investigated. In the first case to be reported the boundary temperature has a constant value. Figure 2 shows the solidification time for a range of solid Stefan numbers for this case. In this figure S_r denotes the ratio of Stefan numbers, $\mathrm{St}_2/\mathrm{St}_1$. Increasing the solid Stefan number results in a fast movement of the interface and the solidification time decreases. The liquid Stefan number has an opposite effect and an increase in the liquid Stefan number increases the total time for solidification. The interface motion is only slightly affected by the liquid Stefan number, if the latter is small. In such a case the temperature in the liquid phase drops to the constant melting temperature, after which the advance of the solidification front is similar to a single-phase problem. For small solid Stefan numbers, a change in the Stefan number ratio does not have a pronounced effect on time for complete solidification. For large solid Stefan numbers, however, the solidification time increases many folds by increasing Stefan number ratio, i.e., for increasing liquid Stefan numbers.

The effect of a sinusoidally varying boundary temperature was investigated for a fluctuation frequency, $\overline{\omega} = 20$. Results for two different fluctuating temperature amplitudes, γ, are reported. In one case the amplitude of temperature fluctuations, $\gamma = 0.25$, which results in a boundary temperature (ψ) variation between 0 and 0.5. The other case is for a boundary temperature variation amplitude of $\gamma = 0.499$, and the resulting boundary temperature variation is between 0 and 0.998. Note that throughout the text this case will be referred to as $\gamma = 0.5$ rather than $\gamma = 0.499$.

The motion of the interface for a solid Stefan number, $\mathrm{St}_1 = 0.1$ and boundary temperature amplitude variation $\gamma = 0.25$ is given in Fig. 3. In order to avoid crowding in the figure, the boundary temperature variation is plotted only for the first few cycles. The dimensionless interface location is plotted for two different Stefan number ratios. It can be seen that in this case the liquid Stefan number has a small effect on the time for total solidification. Due to fluctuating boundary temperature, the advance

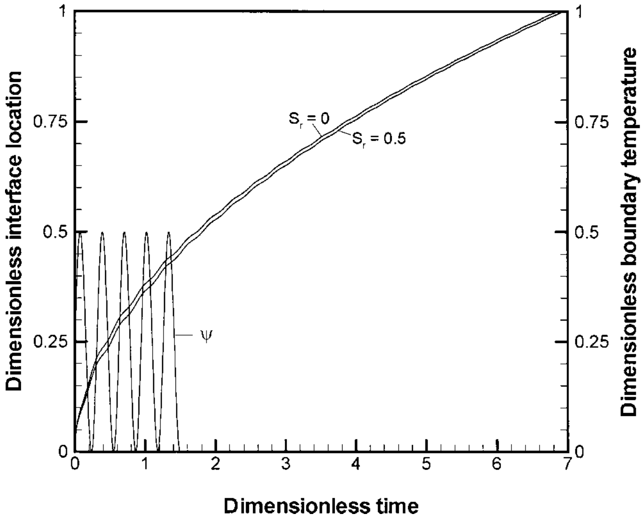


Fig. 3 Interface motion and boundary temperature variation for $St_1=0.1$ and $\gamma=0.25$

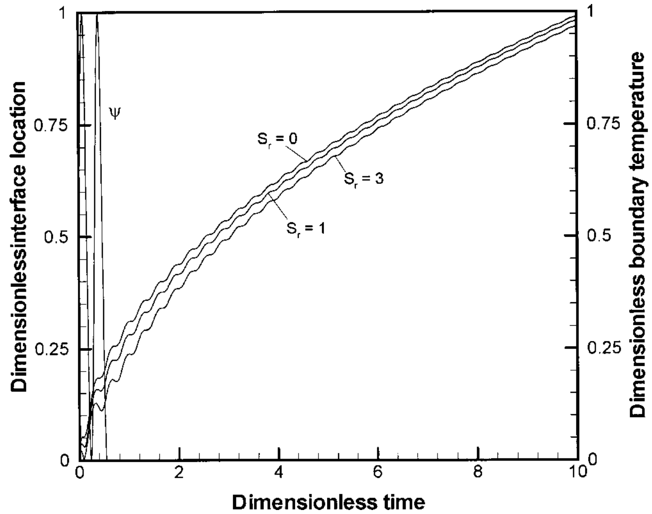


Fig. 5 Interface motion and boundary temperature variation for $St_1=0.1$ and $\gamma=0.5$

of the solidification front accelerates and decelerates in time. This effect is more pronounced during the initial phases of the problem when the solid-liquid interface is closer to the boundary. As the solidification interface advances deeper into the liquid medium, the oscillations of the interface advance are damped due to the increased delaying effect of a thicker solid layer. Figure 4 shows the same phenomenon for a higher solid Stefan number, namely, $St_1=1$. In this case, the interface advances faster and as the ratio of the Stefan numbers, S_r increases—i.e., for large liquid Stefan numbers—the interface moves slower. When the case for $S_r=4$ is analyzed, for small time there is little lag between the interface motion and the boundary temperature variation. As the boundary temperature increases, the speed of the interface advance decreases and *vice versa*. For large time, this effect is damped. In addition, the acceleration and deceleration of the interface motion is large, and for large liquid Stefan numbers the interface starts to retreat towards the solid domain. It can be observed that for a Stefan number ratio $S_r=4$ after a full cycle of boundary temperature variation, the interface moves backwards towards the boundary. In this case the liquid Stefan number is high and the interface heat balance is dominated by the large temperature gradient in the liquid phase. This results in a retrograde motion of the interface. When the liquid Stefan number is smaller (for example $S_r=3$) the backward motion is less important, and for even smaller Stefan number ratios this phenomenon does not exist. The only effect is a slowing down of the forward advance of the interface. The retrograde motion is observed only once, after which the effect is reduced to a slower advance of the solidification front. This effect is even more prominent for larger values of the boundary temperature amplitude. The case when $\gamma=0.5$ is given for a liquid Stefan number, $St_1=0.1$ in Fig. 5. The boundary temperature variation is again plotted only for small time to avoid crowding. When the liquid Stefan number is large $(S_r=3)$, the retrograde motion persists for numerous cycles and it is observed even for smaller liquid Stefan numbers $(S_r=1$ and $S_r=0.5)$. An expanded view of the same case for small time is given in Fig. 6. This figure shows that in the case when $S_r=3$ the interface motion slows down and is arrested around $\tau=0.03$, when the boundary temperature is still substantially lower than the interface temperature. Due to the thin solid layer, the temperature in the solid assumes a linear shape. At the same time due to the large liquid Stefan number, heat capacity of the liquid phase is also large, and the interface heat balance is dominated by the liquid phase, resulting in the backward motion

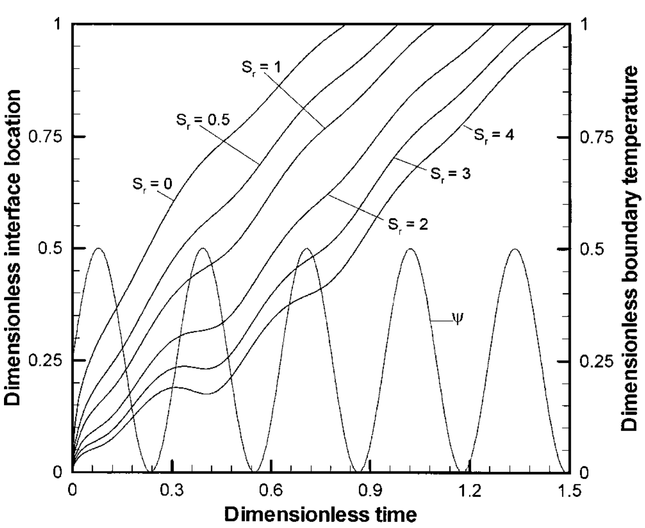


Fig. 4 Interface motion and boundary temperature variation for $St_1=1$ and $\gamma=0.25$

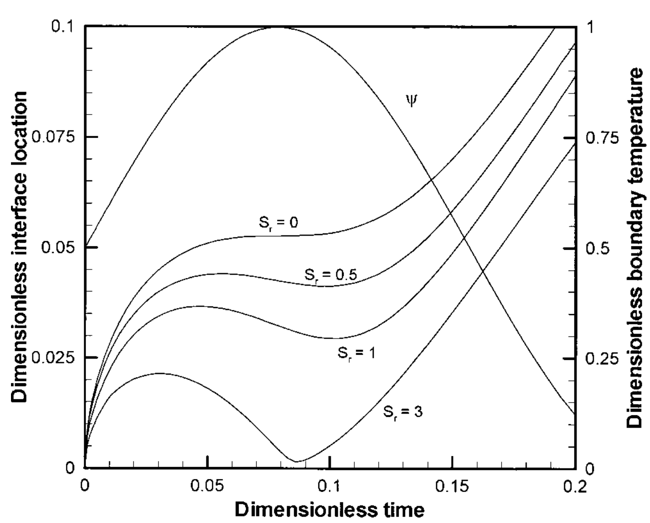


Fig. 6 Interface motion and boundary temperature variation for $St_1=0.1$ and $\gamma=0.5$ for small time

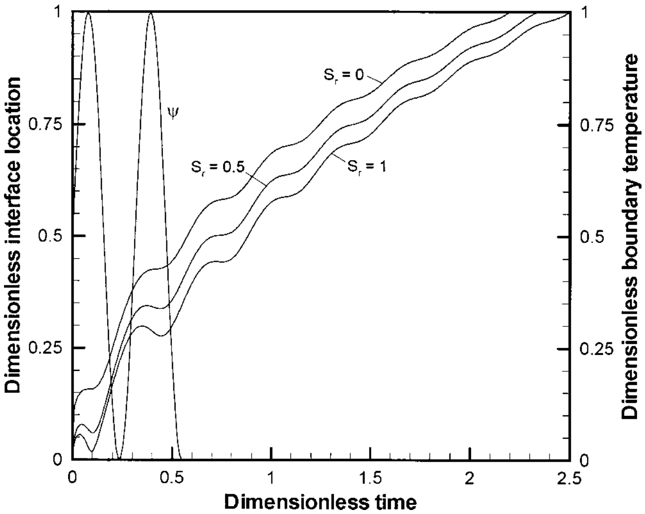


Fig. 7 Interface motion and boundary temperature variation for $St_1 = 0.5$ and $\gamma = 0.5$

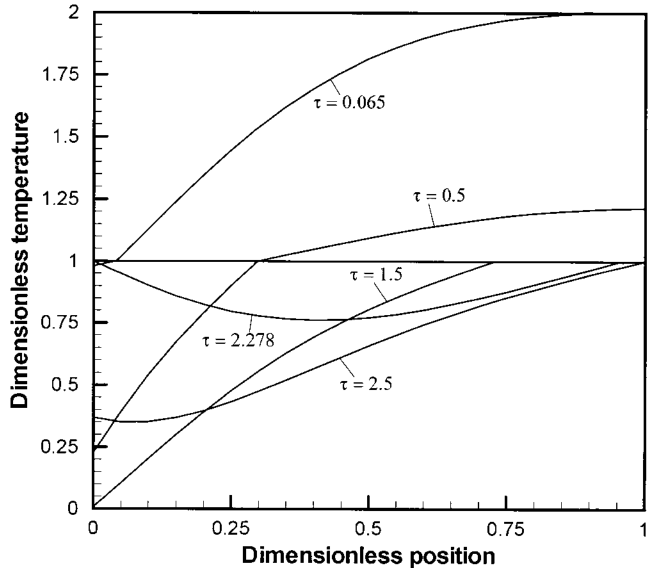


Fig. 8 Temperature distribution during solidification for $St_1 = 0.5$, $St_2 = 0.5$, and $\gamma = 0.5$

of the solidification front, i.e., melting rather than solidification. The forward motion resumes after the boundary temperature reaches its minimum value at the interface. After this point ($\tau \approx 0.075$) due to an increasing temperature difference between the boundary and the interface, and a thin solid layer, the temperature gradient in the solid phase gets large, resulting in a faster advance of the solidification front. It should be noted that even for the case when the liquid is initially at the melting temperature ($S_r = 0$) the forward motion is almost halted.

In most practical applications solid and liquid Stefan numbers are small. It is, nevertheless, instructive to see the effect of large Stefan numbers on solidification for the problem in question. Figure 7 shows the interface motion for a solid Stefan number, $St_1 = 0.5$. The general features of the interface motion are similar to the previous case, the retrograde motion, however, even occurs for smaller values of Stefan number ratios. The forward motion is arrested for a longer duration even when the liquid Stefan number is zero.

Figure 8 shows the transient temperatures (Eq. (16)) for the case when $St_1 = 0.5$, $St_2 = 0.5$, and $\gamma = 0.5$. When $\tau = 0.065$, the interface is in the first retrograde motion as given in Fig. 8. The solid layer is thin; the temperature gradient in the liquid layer is small. On the other hand, although the effect of boundary temperature has penetrated well into the liquid phase, there is a large temperature gradient on the liquid side of the interface. As time advances, heat is removed from the liquid phase and when $\tau = 0.5$ the temperature difference between the insulated boundary and the interface drops to a fraction of its original value. This corresponds to the end of the second retrograde motion and the beginning of the new advance of the interface as shown in Fig. 8. One can see that the magnitude of the backward motion is smaller and its duration is dispersed over longer time interval. For large time ($\tau = 1.5$) the temperature of liquid phase has already reached the interface value. At this point a retrograde motion is no more possible, since the heat flux term that results in a backward motion of the interface vanishes. Comparing this with the interface advance of the same problem given in Fig. 8, it can be seen that the interface motion, indeed, is not halted; only the speed of the advance slows down. When $\tau = 2.278$, prior to complete solidification, the boundary temperature reaches the interface value, and the temperature in the solid phase has a local minimum in the middle of the domain. The temperature distribution given at $\tau = 2.5$ is at the instant of complete solidification.

Conclusion

In this study the solidification of a finite medium was analyzed using a semi analytical approach. The results showed that if the initial temperature of the medium is different than melting, a possibility of a reversal of the interface motion exists. With increasingly large differences between the initial medium temperature and the phase-change temperature, the retrograde motion is more prominent, with a longer duration and persists for more cycles. It was also observed that with increasing solid Stefan numbers the interface advance is faster and with increasing liquid Stefan numbers the interface advance is slower.

References

[1] Carslaw, H. S., and Jaeger, J. C., 1954, *Conduction of Heat in Solids*, Oxford University Press, Oxford, UK.
[2] Tao, L. N., 1978, "The Stefan Problem With Arbitrary Initial and Boundary Conditions," Q. Appl. Math., **36**, pp. 223–233.
[3] Tao, L. N., 1981, "The Exact Solutions of Some Stefan Problems With Prescribed Heat Flux," ASME J. Appl. Mech., **48**, pp. 732–736.
[4] Crank, J., 1988, *Free and Moving Boundary Problems*, Oxford University Press, Oxford, UK.
[5] Rubinstein, L. I., 1971, *The Stefan Problem*, American Mathematical Society, Providence, RI (English translation).
[6] Ockendon, J. R., and Hogkins, W. R., 1975, *Moving Boundary Problems in Heat Flow and Diffusion*, Clarendon Press, Oxford, UK.
[7] Wilson, D. G., Solomon, A. D., and Boggs, P. T., 1978, *Moving Boundary Problems*, Academic Press, San Diego, CA.
[8] Fukusako, S., and Seki, N., 1987, "Fundamental Aspects of Analytical and Numerical Methods on Freezing and Melting Heat-Transfer Problems," *Annual Review of Numerical Fluid Mechanics and Heat Transfer*, T. C. Chawla, ed., Hemisphere, Washington, DC, **1**, pp. 351–402.
[9] Weinbaum, S., and Jiji, L. M., 1977, "Singular Perturbation Theory for Melting and Freezing in Finite Domains Initially not at the Fusion Temperature," ASME J. Appl. Mech., **44**, pp. 25–30.
[10] Charach, Ch., and Zoglin, P., 1985, "Solidification in a Finite, Initially Overheated Slab," Int. J. Heat Mass Transfer, **28**, pp. 2261–2268.
[11] Nair, S., 1994, "Numerical Solution of Moving Boundary Problems Using Integral Equations," ASME AMD 182, *Transport Phenomena in Solidification*, ASME, New York, AMD-182, pp. 109–118.
[12] Rizwan-uddin, 1999, "One-Dimensional Phase Change With Periodic Boundary Conditions," Numer. Heat Transfer, Part A, **35**, pp. 361–372.
[13] Dursunkaya, Z., and Nair, S., 1990, "A Moving Boundary Problem in a Finite Domain," ASME J. Appl. Mech., **57**, pp. 50–56.

A. E. Giannakopoulos[1]
Department of Materials Science and Engineering,
Massachusetts Institute of Technology,
Cambridge, MA 02139
Mem. ASME

The Influence of Initial Elastic Surface Stresses on Instrumented Sharp Indentation

The present work examines the influence of pre-existing, elastic surface stresses on instrumented sharp indentation. The surface is modeled as a homogeneous and isotropic elastoplastic solid in the context of linear elasticity and Mises plasticity with isotropic strain hardening and associative flow rule. Prior to indentation, a homogeneous, biaxial, elastic stress state may exist in the substrate. The influence of the initial elastic surface stresses on the force-depth response of instrumented sharp indentation tests, such as Vickers and Berkovich, was analyzed. The unique connection of the indentation loading response and the average initial stress is based on a relation of the stresses right under the indenter, which holds universally for nearly incompressible materials. [DOI: 10.1115/1.1485756]

1 Introduction

The effect of surface residual stresses is important in many applications of scientific and technological interest. Regarding instrumented indentation, surface residual stresses may affect considerably the interpretation of indentation response with regard to inferring the material properties of the indented substrate. Of particular concern is the micro and nano-indentation analysis of films which have initial stresses due to thermal, mechanical, or other processing treatment. Recently, Suresh and Giannakopoulos [1] proposed a general methodology for the determination of equal-biaxial surface residual stresses using instrumented sharp indentation. However, surface initial stresses are not always equal-biaxial. Complex thermal and/or mechanical loading often produce residual stresses that are unequal. It is the purpose of this short paper to examine the general case of the influence of surface initial stresses on instrumented sharp indentation. In the following, the general analytic approach is presented first, followed by experimental validation from available experiments in the open literature and from experiments conducted in this work. The analysis is based on a fundamental relation of contact mechanics which relates the components of the stress tensor directly under the contact area of an indentation by a sharp indenter. The final outcome of the analysis shows how the surface initial stresses influence the force-depth response of instrumented sharp indentation tests where load and penetration depth are recorded simultaneously.

2 Assumptions and Approach

Consider frictionless, quasi-static sharp indentation of an elastoplastic substrate by an elastic indenter which is much harder than the substrate. The substrate is assumed homogeneous and isotropic.[2] Around the contact area, the indented substrate plastifies and undergoes isotropic strain hardening. The indentation induced plastic deformation is assumed to be isochoric (incompressible) and the plastic strains to be much higher than the elastic strains (this is usually the case for Vickers or Berkovich indentation of soft metals with moderate strain hardening). A normal load P produces an indent of area A and penetration depth h.

The indented surface may have been subjected to prior mechanical and/or thermal loading. Such prior loading results in a biaxial state of residual stresses. Therefore, without loss of generality, a Cartesian coordinate system (O, x, y, z) is chosen so that the x and y-axes are the principal directions of the surface initial stresses, with the origin attached at the center of the contact area, and z is the axis which is positive inside the substrate. The principal stresses at the surface are $\sigma^R_{x,0}$, $\sigma^R_{y,0}$ along the x and y components of coordinate system, respectively, as shown in Fig. 1. The surface stresses $(\sigma^R_{x,0}, \sigma^R_{y,0})$ are assumed to be homogeneous (independent of the x, y, z coordinates), or approximately uniform at a distance of $3\sqrt{\pi A}$ in all directions from the center of the contact area. The biaxial elastic residual strains are derived from a state of plane stress

$$\epsilon^R_{x,0} = (\sigma^R_{x,0} - \nu\sigma^R_{y,0})/E, \quad \epsilon^R_{y,0} = (\sigma^R_{y,0} - \nu\sigma^R_{x,0})/E, \tag{1}$$

where E and ν are, respectively, the Young's modulus and Poisson ratio of the substrate. Since the residual stress field is elastic, the limits of $\sigma^R_{x,0}$ and $\sigma^R_{y,0}$ are determined by the critical Mises yield condition

$$\sigma_y^2 < (\sigma^R_{x,0})^2 + (\sigma^R_{y,0})^2 - (\sigma^R_{x,0})(\sigma^R_{y,0}), \tag{2}$$

where σ_y is the yield strength of the substrate.

In the absence of initial stresses, contact mechanics analysis of nearly incompressible linear elastic behavior ([2]) and axisymmetric slip-line theory of rigid-perfectly plastic response ([3]) gives a relation among the stresses just under the contact area ($z=0$)

$$\sigma_{xx} + \sigma_{yy} = 2\sigma_{zz}. \tag{3}$$

Note that σ_{zz} in Eq. (3) is the normal contact stress distribution. The shear stresses at the surface are zero, due to the frictionless type of contact. Equation (3) is true for normal contact of any punch shape with a nearly incompressible linear elastic material and is approximately true for Vickers, Berkovich, and cone indentation of elastoplastic substrates in the absence of surface initial stresses. Finite element analysis of axisymmetric cone indentation of elastoplastic materials, performed by the present author, confirm Eq. (3) within 5% and the results deviate significantly for low Poisson ratio, $\nu < 0.33$, high yield strength, $\sigma_y/E > 0.005$, and high strain hardening. Equation (3) is based on the nearly isoch-

[1]Currently at the Department of Civil Engineering, University of Thessaly, 38336 Volos, Greece.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, February 26, 2001; final revision, February 19, 2002. Associate Editor: H. Gao. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

[2]Under certain symmetry conditions of the material and the contact area, the isotropy assumption may be relaxed and the substrate may be considered to be orthotropic.

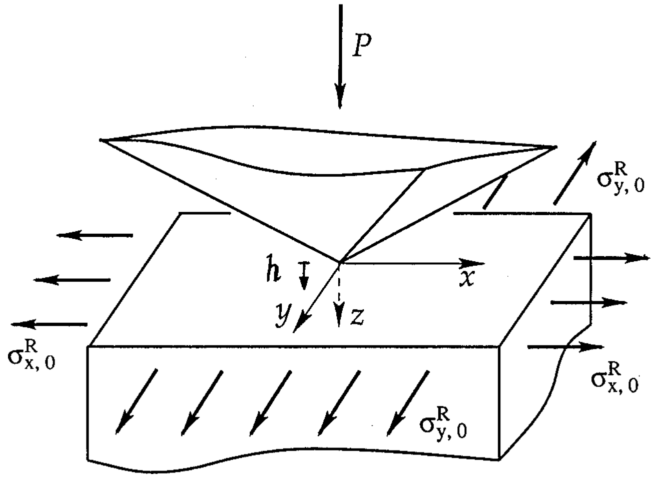


Fig. 1 Schematic of sharp indentation with the associated nomenclature

oric deformation under the contact region and is important for solving the problem of the influence of initial stresses in the sharp indentation response (the $P-h$ loading curve) of many metals and metal alloys.

To continue the analysis, the surface initial stresses are decomposed into an equal-biaxial part and a pure plane shear part, Fig. 2. The magnitude of the equal-biaxial component is

$$\sigma^R=(\sigma^R_{x,0}+\sigma^R_{y,0})/2 \tag{4}$$

and the magnitude of the shear component is

$$\tau^R=|\sigma^R_{x,0}-\sigma^R_{y,0}|/2. \tag{5}$$

The shear component, τ^R, does not affect the macroscopic indentation response, because it does not affect the contact stress distribution. This can be shown by recasting the shear initial stress into a biaxial initial stressing of equal but opposite in sign components. It can then be immediately seen that after superposition of the initial stresses to the indentation induced stresses (e.g., in the absence of residual stresses), Eq. (3) is preserved without changing the σ_{zz} stress component (the contact stress distribution)

$$[\sigma_{xx}+(\sigma^R_{x,0}-\sigma^R_{y,0})/2]+[\sigma_{yy}+(\sigma^R_{y,0}-\sigma^R_{x,0})/2]=2\sigma_{zz}. \tag{6}$$

Equation (6) proves that the plane shear initial stress leaves the contact stress distribution invariant. Therefore, the only influence of the surface initial stresses to the force-depth indentation response is due only to the equal-biaxial part of the initial stresses, σ^R, given by Eq. (4).

A key aspect in the present analysis is the invariance of the hardness when initial elastic stresses are present. This invariance can be proven in the context of linear elasticity since the residual stresses do not contribute to the indentation deformation work. However, this is not the case for large deformation elasticity. Dhaliwal and Singh [4] obtained an analytic solution for the axisymmetric conical indentation of a prestressed neo-Hookean (incompressible) half-space and their results indicate that the average contact pressure (hardness) is a function of the amount of prestrain. Nevertheless, if the residual stress, σ^R, is sufficiently small compared to the elastic modulus, E,

$$3\sigma^R/E\ll 1, \tag{7}$$

then the large deformation analysis also predicts invariant average contact pressure. Regarding elastic-plastic material behavior, Bolshakov et al. [5] used finite elements to simulate cone indentation of an aluminum alloy and found that equal-biaxial initial stress leaves the average contact pressure invariant. The negligible effect of pre-existing stress on hardness was shown numerically by Mesarovic and Fleck [6] for spherical indentation of elastic-plastic solids, provided that the loading is high enough for plastic strains to dominate over the whole contact area. For the cone and pyramid indentation, a low bound condition for the elastic strains to dominate over the plastic strains is

$$E/\tan\alpha<2.31\sigma_y(1-\nu^2)/(1-2\nu), \tag{8}$$

where $\alpha=\pi/2-\gamma$, with 2γ being the included angle of the indenter tip ($\alpha=22$ deg for Vickers tetragonal pyramid indenter, 24.7 deg for the Berkovich trigonal pyramid indenter, and 19.7 deg for the equivalent circular conical indenter).[3] Combination of (7) and (8) gives an interesting constrain for the maximum initial tension that preserves hardness

$$\sigma^R/\sigma_y<0.77\tan\alpha(1-\nu^2)/(1-2\nu). \tag{9}$$

For example, for $\alpha=22$ deg and $\nu=0.3$, $\sigma^R<0.64\sigma_y$ representing a low bound limit for the magnitude of the residual stresses.

A general explanation of the hardness invariance can be derived by observing that in the principal stress space the origin for the stressing of each material point is shifted either to the positive quarter or to the negative quarter depending to the sign of σ^R and that the stress increments due to the compressive nature of inden-

[3]Proof of Eq. (8) can be given using the stress state at the contact perimeter of a cone indentation ([2]), $\sigma_{xx}=-\sigma_{yy}=(1-2\nu)E/[4(1-\nu^2)\tan\alpha]$ and all other stresses are zero. Therefore, the Mises stress at the contact perimeter is $\sigma_{xx}\sqrt{3}$, so elasticity prevails if $\sigma_{xx}\sqrt{3}<\sigma_y$ and therefore Eq. (8) holds true.

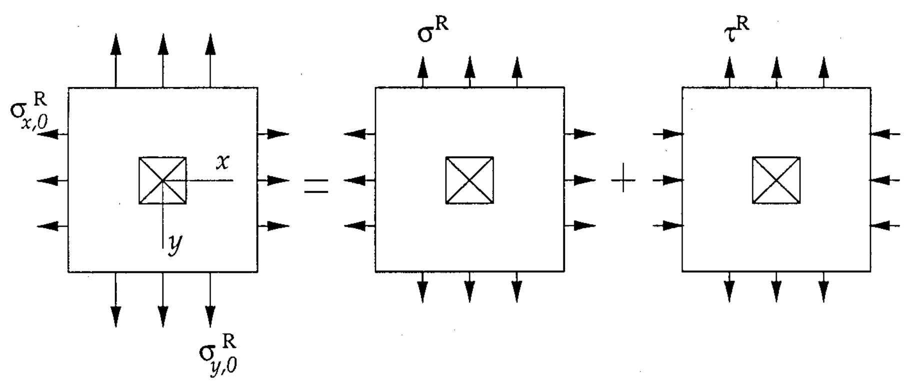


Fig. 2 Decomposition of the surface initial stresses to an equal-biaxial part of magnitude $\sigma^R=(\sigma^R_{x,0}+\sigma^R_{y,0})/2$ and a shear part of magnitude $\tau^R=|\sigma^R_{x,0}-\sigma^R_{y,0}|/2$

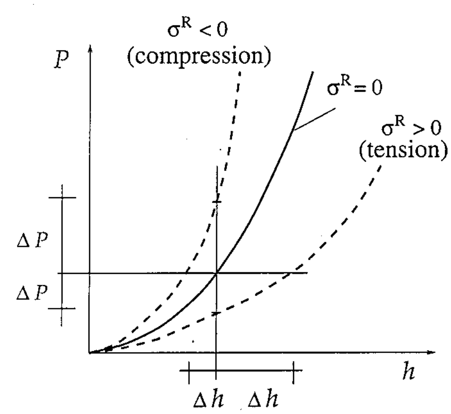


Fig. 3 The change of the loading part of the indentation $P-h$ curve due to the presence of surface initial stresses

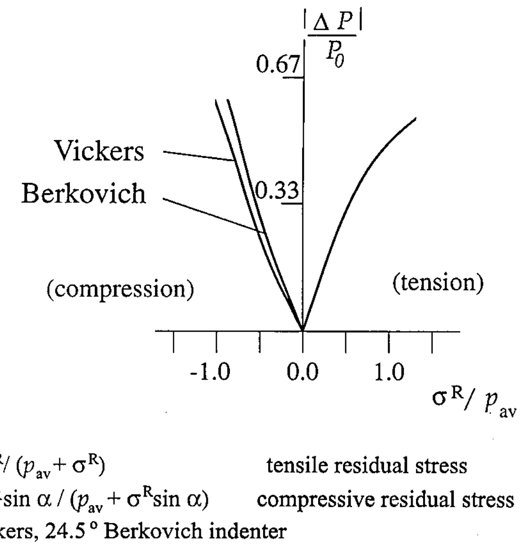


Fig. 4 The relative change of load, $|\Delta P|/P_0$, as a function of the normalized initial stress, σ^R/p_{av}. Comparison at constant indentation depth.

tation are always pointing to the negative direction. It is then clear that the material points at the surface will yield and deform plastically much easier under initial compression, whereas the opposite is true under initial tension. This implies that the contact area will tend to be larger under initial compression and smaller under initial tension. For classic elastoplastic behavior, the yield surface in the principal stress space is an inclined cylinder with the symmetry line the line of equal principal stresses. Since the out-of-plane initial stress is zero and the contact stresses are negative (compressive), negative stress increments due to contact will produce higher contact stresses for compressive initial stresses and lower contact stresses for tensile initial stresses. Therefore, since both load and contact area vary in the same way, similarity requires that the hardness is invariant with respect to the initial stresses. This result will not be valid, if yielding is pressure sensitive.

3 Effect on the Loading Part of the Force-Depth Response

Denote by P_0 and A_0 the load and real (vertically projected) contact area corresponding to the sharp indentation in the absence of equal-biaxial component, $\sigma^R=0$, and by P and A the corresponding quantities when $\sigma^R\neq 0$. From the invariance of the average contact pressure (and, equivalently, the invariance of the hardness)

$$p_{av}=P/A=P_0/A_0. \tag{10}$$

If the only length in the problem comes from the depth of penetration, Kick's law holds regarding the load-depth relation

$$P=Ch^2, \quad P_0=C_0h_0^2. \tag{11}$$

In Eq. (11), C signifies the indentation compliance and the subscript 0 refers to the $\sigma^R=0$ case. Sinking-in and pile-up at the indenter contact perimeter are included in the relations of the projected contact areas

$$A=Dh^2, \quad A_0=D_0h_0^2, \tag{12}$$

for $\sigma^R\neq 0$ and $\sigma^R=0$, respectively.

The influence of the initial stresses to the loading part of the load-depth response depends on the sign of σ^R (positive when tensile and negative when compressive), see Fig. 3. The comparison of the load-depth response of indentation of a substrate with and without initial stresses can be made either at constant applied load or at constant penetration depth (see [1] for details). A thermodynamic explanation of Fig. 3 is given in the Appendix.

If the comparison between the $\sigma^R=0$ and $\sigma^R\neq 0$ cases is made at constant penetration depth, then

$$\frac{A}{A_0}=\left(1+\frac{\sigma^R}{p_{av}}\right)^{-1}; \quad \sigma^R>0, \tag{13}$$

$$\frac{A}{A_0}\approx\left(1+\frac{\sigma^R\sin\alpha}{p_{av}}\right)^{-1}; \quad \sigma^R<0. \tag{14}$$

If the comparison between the $\sigma^R=0$ and the $\sigma^R\neq 0$ cases is made at constant applied load, then

$$\frac{h^2}{h_0^2}=\left(1-\frac{\sigma^R}{p_{av}}\right)^{-1}; \quad \sigma^R>0, \tag{15}$$

$$\frac{h^2}{h_0^2}\approx\left(1-\frac{\sigma^R\sin\alpha}{p_{av}}\right)^{-1}; \quad \sigma^R<0. \tag{16}$$

When comparing at constant indentation depth, Eqs. (13), (14) can be recasted as relative change of the applied force, $\Delta P=P-P_0$,

$$\Delta P/P_0=-\sigma^R/(p_{av}+\sigma^R); \quad \sigma^R>0 \tag{17}$$

$$\Delta P/P_0\approx-\sigma^R\sin\alpha/(p_{av}+\sigma^R\sin\alpha); \quad \sigma^R<0. \tag{18}$$

On the other hand, in case of comparison at constant load, Eqs. (15), (16) can be recasted as a relative change of the penetration depth of the indenter, $\Delta h=h-h_0$,

$$\Delta h/h_0=\left(\frac{p_{av}}{p_{av}-\sigma^R}\right)^{1/2}-1; \quad \sigma^R>0 \tag{19}$$

$$\Delta h/h_0\approx\left(\frac{p_{av}}{p_{av}-\sigma^R\sin\alpha}\right)^{1/2}-1; \quad \sigma^R<0. \tag{20}$$

Equations (17) and (18) are plotted in Fig. 4 and Eqs. (19) and (20) are plotted in Fig. 5. It can be seen that the highest change in the load-depth response appears when the comparison is made at constant indentation depth, Fig. 4. It is therefore recommended that for experimental deduction of the initial stresses from the $P-h$ loading response, the constant depth comparison be used, i.e., Eqs. (17) and (18) and Fig. 4, because it maximizes resolution. An obvious additional limit for the present analysis can be shown from Eqs. (14) and (15)

$$0<\sigma^R<p_{av}, \tag{21}$$

when comparison is made at constant applied load and

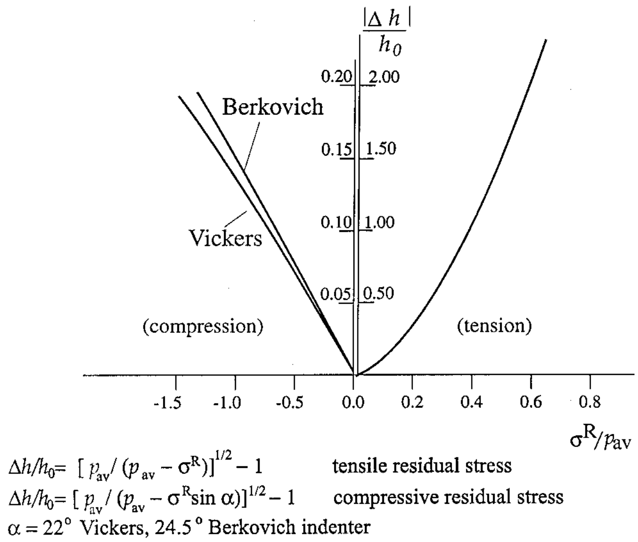


Fig. 5 The relative change of penetration depth, $|\Delta h|/h_0$, as a function of the normalized initial stress, σ^R/p_{av}. Comparison at constant applied load.

$$0 < -\sigma^R \sin\alpha < p_{av}, \tag{22}$$

when comparison is made at constant penetration depth.

Figures 4 and 5 can be used to assess the experimental error of the residual stresses provided that the accuracy of either the load or depth measurements is known.

4 Effect on the Unloading Part of the Force-Depth Response

Denote by E^* the indentation modulus which includes the elastic deformation of the indenter,

$$E^* = \left(\frac{1-\nu^2}{E} + \frac{1-\nu_{in}^2}{E_{in}}\right)^{-1}, \tag{23}$$

where E_{in}, ν_{in} are the Young's modulus and Poisson ratio of the sharp indenter and E, ν are the corresponding elastic constants for the substrate. The real contact areas at maximum load are related to the initial unloading portion of the $P-h$ curve

$$A = \left(\frac{dP}{dh}\frac{1}{C_u E^*}\right)^2, \quad A_0 = \left(\frac{dP_0}{dh_0}\frac{1}{C_u E^*}\right)^2, \tag{24}$$

where $C_u = 1.142$ for the Vickers and 1.167 for the Berkovich indenter [7,8].

Comparing the cases of $\sigma^R = 0$ and $\sigma^R \neq 0$ at fixed depth of penetration, Eq. (24) gives

$$\frac{A}{A_0} = \left(\frac{dP}{dh}\right)^2 \left(\frac{dP_0}{dh_0}\right)^{-2}, \tag{25}$$

see Fig. 6. Comparing the cases of $\sigma^R = 0$ and $\sigma^R \neq 0$ at fixed applied load, Eqs. (24) and (10) give

$$\frac{dP}{dh} = \frac{dP_0}{dh_0}. \tag{26}$$

It can be concluded that when comparison is made at fixed applied load, the unloading part of the $P-h$ curve cannot give any information on the initial stress. The unloading part of the curve can be useful in the determination of the initial stresses, if comparison is made at fixed indentation depth by combining Eqs. (25), (13), and (14).

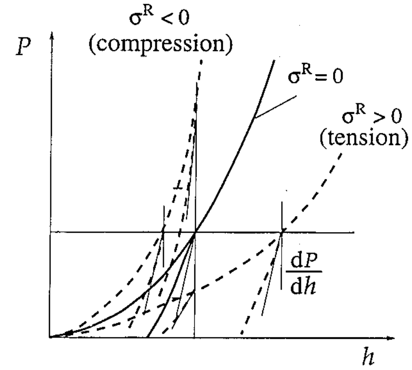


Fig. 6 The change of the unloading part of the indentation $P-h$ curve due to the presence of surface initial stresses

5 Comparison of Predictions With Experiments

Early experiments of Fink and Van Horn [9] investigated experimentally the influence of uniaxial tensile bending stress on the Rockwell E hardness of the 17S aluminum alloy. They found that for $\sigma^R/\sigma_y < 0.5$, the hardness decreased less than 2% from the hardness of the unstressed material. Later, Sines and Carlson [10] investigated experimentally the influence of uniaxial bending stress on the Rockwell B hardness of an annealed high-carbon steel (the surface stresses were below the elastic limit). They found that for compressive initial stress, the hardness increased less than 1% and for tensile initial stress the hardness decreased less than 5% from the hardness of the unstressed material.

More recently, Simes et al. [11] examined experimentally the influence of uniform, tensile biaxial stress on the Vickers hardness of an annealed bright drawn mild steel. They found that for $\sigma^R/\sigma_y < 0.4$, the hardness decreased less than 3% from the hardness of the unstressed material and for $1 > \sigma^R/\sigma_y > 0.4$ the hardness decreased dramatically. This result is similar to that found by Fink and vanHorn for the uniaxial tensile residual stress and can be explained by the constraint given by Eq. (9) (for the Vickers geometry $\tan\alpha \approx 0.4$).

In order to further validate the theory, available experimental results were found in the work of Tsui et al. [12]. Tsui et al. give adequate information regarding nanoindentation tests where trigonal Berkovich diamond pyramids indented a rapidly solidified 8009 aluminum alloy (this is a very fine-grained alloy with a yield strength of $\sigma_y = 353$ MPa), at room temperature. Prior to indentation, the substrate was subjected to pure bending that produced an initial uniaxial (elastic) stress of different magnitude (tensile as well as compressive). A fixed grip attachment locked in a surface initial stress, $\sigma^R_{x,0}$, due to bending. Therefore, taking the y direction as the bending moment direction with $\sigma^R_{y,0} = 0$, Eq. (4) gives

$$\sigma^R = \sigma^R_{x,0}/2. \tag{27}$$

Figure 7 shows the experimental values of the apparent contact areas, at different values of initial (residual) stress $\sigma^R_{x,0}$. From Tsui et al. [12], the reported values of the mechanical constants are: $E = 82.1$ GPa, $\nu = 0.31$, $\sigma_{0.29} = 426$ MPa[4] for 8008 aluminum and $E_{\rm in} = 1006$ GPa, $\nu_{\rm in} = 0.07$ for the diamond indenter. According to Giannakopoulos and Suresh [13], at $P_0 = 110$ mN, the average pressure is $p_{av} = 1.3$ GPa, the contact area is $A_0 = 85\ \mu{\rm m}^2$, the unloading slope is $dP_0/dh_0 = 0.89$ mN/mm and the residual penetration depth is $h_{r,0}/h_{\max,0} = 0.913$. Tsui et al. reported average pressure $p_{av} = 1.2$ GPa, contact area $A_0 = 92\ \mu{\rm m}^2$, unloading slope $dP_0/dh_0 = 0.89$ mN/mm, residual penetration depth $h_{r,0}/h_{\max,0} = 0.913$ and contact stiffness $C_0 = 37$ GPa. Figure 7

[4] $\sigma_{0.29}$ is the uniaxial stress at 0.29 compressive strain.

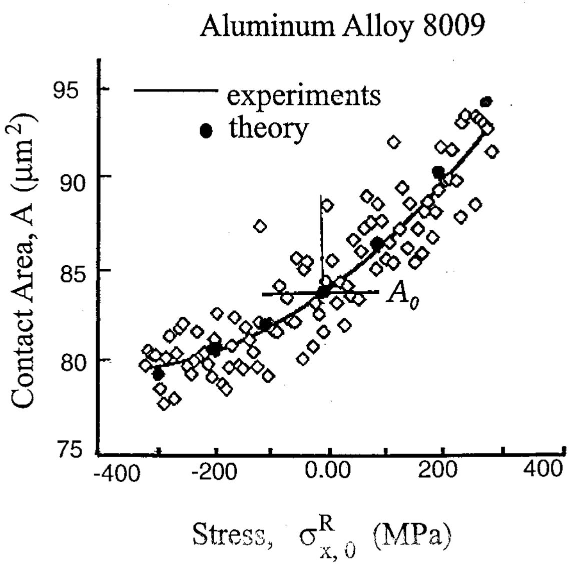


Fig. 7 Nanoindentation measurements of the apparent contact area for uniaxially stressed specimen of aluminum alloy 8009, after Tsui et al. [12]. The prediction of the present analysis is shown by dark dots. The comparison of apparent areas is done at constant load. The apparent contact area is a simple scale of the maximum indentation load due to the elastic residual stresses, ignoring the hardness invariance.

shows the experimental values of the apparent contact area (comparison of the $P-h$ curves at constant load), together with the predictions from the present analysis, Eqs. (15) and (16) with $\alpha=24.7\text{deg}$ and σ^R from Eq. (27). The agreement is very good.

A key result of the present analysis is that an initial shear stress will neither affect the hardness nor the loading force-depth relation. It is then important to experimentally validate this result, since such initial stress state has never been tested in connection to instrumented indentation. To that end, instrumented microindentation tests were performed on a cylindrical specimen of Al 7075-T6 which was prestressed under torsion. The material properties of this aluminum alloy are $E=70$ GPa, $\nu=0.3$, shear yield strength $\tau_y=240$ MPa and hardness $p_{av}=1.9$ GPa. A locking device was designed to apply to the specimen an elastic shear stress by locking in an appropriate twisting angle, as shown in Fig. 8. The dimensions of the cylindrical specimen were length $L=10.1$ cm and diameter $D=1.27$ cm. An applied twist of $\Theta^R=6$ deg gave an initial elastic shear stress $\tau^R=177$ MPa. The specimen was polished with 0.3 μm alumina paste and cleaned with acetone. A Vickers diamond indenter was used with a Wilson MicroRockwell indentation device provided by INSTRON. The test was load controlled with constant loading rate of 0.25 N/s. The device was able to record simultaneously the normal load and the vertical displacement provided by a capacitance SMU-9000 from Kaman Instrumentation Corporation. The resolution was 0.1 N for the load and 0.1 μm for the displacement. Four test were conducted in the unstressed and four tests in the twisted specimen in ambient (25°C and 60% relative humidity). In all tests the maximum normal load was 10 N. Under such loading, the assumption of constant shear stress was satisfied within -5%. and the maximum deviation from normal alignment of the indenter was 2 deg. The $P-h$ curves were recorded and the average maximum indentation depth was 15 μm (standard deviation 2.5 μm) for the unstressed specimen and 15.1 μm (standard deviation 3.1 μm) for the stressed specimen. This result proves the invariance of the $P-h$ loading curve under initial shear stress. The imprints were subsequently observed with the Olympus BH2-UMA optical microscope with magnification 3×10^4. The contact area was used

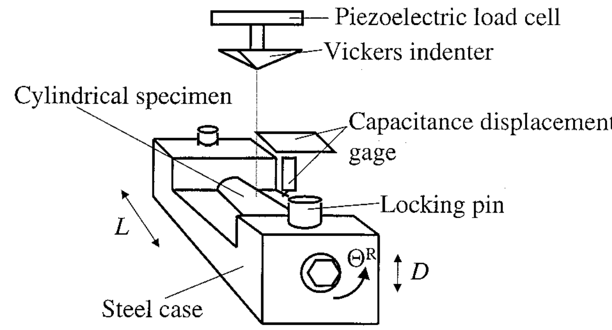


Fig. 8 Schematic of the experimental procedure used to examine the influence of the initial elastic shear stress in the instrumented indentation loading response. The applied twist Θ^R is locked in and produces an elastic shear stress $\tau^R=0.25E\Theta^R D/(L(1+\nu))$ at the surface of the specimen.

to calculate the hardness (average pressure) which was found to be 1.86 GPa, both for the unstressed and the stressed specimen. This result proves the invariance of p_{av} under initial shear stress.

6 Conclusions

A theoretical investigation of the influence of surface initial stresses in the force-depth response of instrumented sharp indentation was presented. The initial stresses are elastic, but not equal-biaxial. The analysis investigated the conditions under which the hardness remains independent of the initial stresses. Both analytic and experimental results indicate that the loading part of the $P-h$ response contains only information about the average surface initial stress, $(\sigma^R_{x,0}+\sigma^R_{x,0})/2$. It is then important to assess the influence of the initial stresses when the macroscopic indentation response is used to extract mechanical properties of the substrates. This may be particularly troublesome in the evaluation of thin film properties. The situation is no better even when the finite element method is used without including the initial stresses in the modelling. If the material properties are known, the force-depth indentation response can be used to assess the average biaxial surface initial stress, but not each initial stress component separately.

In the absence of any information about the initial stresses (e.g., their ratio), only the equal-biaxial and the uniaxial initial stress can be resolved conclusively by the force-depth response. Equal-biaxial stresses are often present in thin films and uniaxial surface residual stress often appears in metal line deposition on ceramic wafers and in uniaxial bending or stretching of plates. In such cases, indentation can be used with the present theory to assess the residual stresses. The limits of the analysis must be also kept in mind, especially regarding homogeneity and isotropy of the material properties, uniformity of the initial stresses, low friction coefficient, and geometric sharpness of the indenter geometry. Moderate deviations from the previous limits may lead to strong deviations from the present results. The method can be applied when the stress-free indentation response is measured or can be constructed from known material properties.

Acknowledgments

The author would like to thank Krystyn Van Vliet for her assistance with the experiments and George Labonde for his help with the loading stage. The experimental work was completed at MIT in the Laboratory for Experimental and Computational Mechanics of Prof. Subra Suresh who the author gratefully acknowledges.

Appendix

Thermodynamic Proof of the Contact Pressure Invariance The force-depth response shown in Fig. 3 can be explained starting from the convexity of the internal energy relation with respect to entropy, S, and volume, V [14]. In thermodynamic terms, the stability of equilibrium during indentation at constant temperature ($T=\text{const}$) requires the Le Chatelier-Braun inequality

$$(\partial p/\partial V)_T < 0, \tag{A1}$$

which connects the pressure change, dp, with the volume change, dV. Similarity implies that the pressure change is proportional to the applied load over the contact area, P/A, assumed for the time invariant. Then, at constant indentation depth ($h=\text{const}$), (A1) gives

$$(\partial P/\partial V)_{T,h} < 0. \tag{A2}$$

Therefore, compared to the stress-free state, the applied load is lower, if the surface is in initial tension ($dV>0$) and higher, if the surface is in initial compression ($dV<0$).

For classic elastoplastic materials, when an initial surface shear stress is applied the change of volume is zero, $dV=0$, therefore the force-depth curve remains invariant. For pressure sensitive materials, such as certain metals, ceramics, glasses, and rocks, shear stresses induce volume changes and for these materials the present analysis does not apply.

The invariance of the contact pressure can be investigated by examining separately the cases of pileup and sinking in of the contact perimeter. Assume that the contact pressure is not invariant but has to remain positive semi-definite in all cases. Consider the constant load comparison. Then, overall equilibrium requires the contact pressure distribution to expand (or recede) from its initial region corresponding to the tensile (or compressive) character of the initial stresses. However, compatibility at the contact perimeter would require the volume close to the surface perimeter to increase or decrease in violation of the LeChatelier-Braun inequality. Therefore, at constant load, both the contact area and the contact pressure distribution must remain invariant in the presence of initial stresses. As a result, the average contact pressure remains also invariant.

References

[1] Suresh, S., and Giannakopoulos, A. E., 1998, "A New Method for Estimating Residual Stresses by Instrumented Sharp Indentation," Acta Mater., **46**, pp. 5755–5767.

[2] Johnson, K.-L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

[3] Shield, R. T., 1955, "On the Plastic Flow of Metals Under Conditions of Axial Symmetry," Proc. R. Soc. London, Ser. A, **A233**, pp. 267–287.

[4] Dhaliwal, R. S., and Singh, B. M., 1978, "The Axisymmetric Boussinesq Problem of an Initially Stressed Neo-Hookean Half-Space for a Punch of Arbitrary Profile," Int. J. Eng. Sci., **16**, pp. 379–385.

[5] Bolshakov, A., Oliver, W. C., and Pharr, G. M., 1996, "Influences of Stress on the Measurement of Mechanical Properties Using Nanoindentation: Part II. Finite Element Simulations," J. Mater. Res., **11**, pp. 760–768.

[6] Mesarovich, S. D., and Fleck, N. A., 1999, "Spherical Indentation of Elastic-Plastic Solids," Proc. R. Soc. London, Ser. A, **A455**, pp. 2707–2728.

[7] Giannakopoulos, A. E., Larsson, P.-L., and Vestergaard, R., 1994, "Analysis of Vickers Indentation," Int. J. Solids Struct., **31**, pp. 2679–2708.

[8] Larsson, P.-L., Giannakopoulos, A. E., Soderlund, E., Rowcliffe, D., and Vestergaard, R., 1996, "Analysis of Berkovich Indentation," Int. J. Solids Struct., **33**, pp. 221–248.

[9] Fink, W. L., and VanHorn, K. R., 1930, "Lattice Distortion as a Factor in the Hardening of Metals," J. Inst. Met., **44**, pp. 241–254.

[10] Sines, G., and Carlson, R., 1952, "Hardness Measurements for Determination of Residual Stresses," ASTM Bull., Feb., pp. 35–37.

[11] Simes, T. R., Mellor, S. G., and Hills, D. A., 1984, "A Note on the Influence of Residual Stress on Measured Hardness," J. Strain Anal., **19**, pp. 135–137.

[12] Tsui, T. Y., Oliver, W. C., and Pharr, G. M., 1996, "Influence of Stress on the Measurement of Mechanical Properties Using Nanoindentation: Part I. Experimental Studies in an Aluminum Alloy," J. Mater. Res., **11**, pp. 752–759.

[13] Giannakopoulos, A. E., and Suresh, S., 1999, "Determination of Elastoplastic Properties by Instrumented Sharp Indentation," Scr. Mater., **40**, pp. 1191–1198.

[14] Kestin, J., 1992, "Local-Equilibrium Formalism Applied to Mechanics of Solids," Int. J. Solids Struct., **29**, pp. 1827–1836.

S. Mukherjee
Fellow ASME
Department of Theoretical and Applied Mechanics,
Cornell University,
Kimball Hall,
Ithaca, NY 14853
e-mail: sm85@cornell.edu

C.-S. Liu
Department of Mechanical and Marine Engineering,
Taiwan Ocean University,
Keelung, Taiwan
e-mail: csliu@mail.ntou.edu.tw

Computational Isotropic-Workhardening Rate-Independent Elastoplasticity

A novel formulation for elastoplasticity has been recently proposed by Liu and Hong. These authors have explored the internal symmetry of the constitutive model for perfect plasticity to ensure that the consistency condition is satisfied at each time step. Moreover, for perfect plasticity, they have converted the usual nonlinear elastoplastic constitutive model into a linear system of ordinary differential equations in redefined variables. The present paper is concerned with general isotropic workhardening. With the present formulation, it is still possible to satisfy the elastoplastic consistency condition at every time step, without the need for iterations even for nonlinear workhardening. The resulting system of ordinary differential equations, however, is, in general, nonlinear. Different strategies for obtaining numerical solutions of these equations are proposed in this paper, one of them based on group theory. Numerical solutions from the different schemes, for a simple illustrative example, are presented in the paper. [DOI: 10.1115/1.1607356]

1 Introduction

Computational rate-independent elastoplasticity is a mature subject. Return mapping algorithms provide an effective integration scheme for the constitutive equations. This procedure carries out a discrete enforcement of the consistency condition that was, to the best of the authors' knowledge, first suggested by Wilkins [1]. The consistency condition ensures that the changing stress state always remains on the (in general) changing yield surface. Radial return mapping algorithms have been employed in conjunction with the continuum (or elastoplastic) tangent by, among others, Hinton and Owen [2] and Pinsky et al. [3]. These elastoplastic moduli are obtained from the continuum (rate) constitutive model by enforcing the above mentioned consistency condition. Nagtegaal [4] observed, in the context of a linear isotropic hardening rule, that use of the continuum tangent results in the loss of quadratic convergence of the associated iterative method. In a seminal paper, Simo and Taylor [5] proposed the consistent tangent elastoplastic operator for rate-independent elastoplasticity. This approach assures consistency between the integration algorithm and the tangent modulus, thereby preserving the quadratic rate of convergence of iterative solution schemes based upon Newton's method. The consistent tangent operator (CTO) has been routinely employed in the context of the finite element method (FEM) for solving problems in elastoplasticity. It has also been employed within the boundary element method (BEM) by Mukherjee and his co-authors, [6,7].

Liu and Hong [8–10] have recently revisited this problem and have proposed a novel formulation for it. *They have explored the internal symmetry of the constitutive model for perfect plasticity to ensure that the consistency condition is satisfied at every time step*. Moreover, for perfect plasticity, they are able to convert a physically nonlinear constitutive model into a *linear* system of first order ordinary differential equations (ODEs) of the form $\dot{x}=Ax$ (with suitably defined variables x_k).

The present paper is concerned with small-strain small-rotation rate-independent elastoplastic problems with general isotropic workhardening. The constitutive model is again written in the form of a system of ODEs in suitably redefined variables. It is again possible to satisfy the consistency condition at every time step, without the need for iterations even for nonlinear workhardening plasticity, but the system of ODEs is no longer linear in this case. Three different strategies for obtaining numerical solutions to the system of equations are proposed herein. These are a direct strategy, one based on converting the system of ODEs into an equivalent nonlinear Volterra integral equation, and, finally, a solution strategy based on group theory. Numerical solutions from the different methods, for a simple illustrative example, are presented next. A concluding remarks section completes the paper.

2 Constitutive Equations

2.1 Separation of Strain Rates.

$$\dot{\boldsymbol{\epsilon}}=\dot{\boldsymbol{\epsilon}}^{(e)}+\dot{\boldsymbol{\epsilon}}^{(p)} \tag{1}$$

$$\mathbf{d}=\mathbf{d}^{(e)}+\mathbf{d}^{(p)} \tag{2}$$

where

$$\mathbf{d}\equiv\dot{\mathbf{e}}=\dot{\boldsymbol{\epsilon}}-(1/3)\mathrm{tr}(\dot{\boldsymbol{\epsilon}})\mathbf{I} \tag{3}$$

and similarly for $\mathbf{d}^{(e)}$ and $\mathbf{d}^{(p)}$.

2.2 Elasticity.

$$\mathbf{d}^{(e)}=\dot{\mathbf{s}}/(2G) \tag{4}$$

$$tr(\dot{\boldsymbol{\epsilon}}^{(e)})=\mathrm{tr}(\dot{\boldsymbol{\sigma}})/(3K)\equiv\dot{p}/K \tag{5}$$

where

$$\mathbf{s}=\boldsymbol{\sigma}-(1/3)\mathrm{tr}(\boldsymbol{\sigma})\mathbf{I}. \tag{6}$$

2.3 Plasticity.

$$\mathbf{d}^{(p)}=0 \quad \text{for } \sigma<\kappa(\epsilon^{(p)}) \tag{7}$$

$$\mathbf{d}^{(p)}=\frac{3d_0^{(p)}}{2\sigma}\mathbf{s} \quad \text{for } \sigma=\kappa(\epsilon^{(p)}) \tag{8}$$

$$\mathrm{tr}(\dot{\boldsymbol{\epsilon}}^{(p)})=0 \tag{9}$$

where

$$\sigma=\sqrt{(3/2)\mathbf{s}:\mathbf{s}} \tag{10}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 17, 2002; final revision, Apr. 24, 2003. Associate Editor: B. M. Moran. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$d_0^{(p)} = \dot{\epsilon}^{(p)} = \sqrt{(2/3)\mathbf{d}^{(p)}:\mathbf{d}^{(p)}}. \tag{11}$$

For the case of plastic flow, use of (4) and (8) in (2) gives

$$\dot{\mathbf{s}} + \frac{3Gd_0^{(p)}}{\kappa(\epsilon^{(p)})}\mathbf{s} = 2G\mathbf{d}. \tag{12}$$

In Eqs. (1) through (12), the tensor $\boldsymbol{\epsilon}$ is the strain, with a superscript (e) or (p) denoting its elastic and plastic parts, respectively. A superscribed dot over a variable denotes its derivative with respect to (pseudo) time. The stress tensor is $\boldsymbol{\sigma}$ and the deviatoric parts of the stress and strain tensors are $\mathbf{s}$ and $\mathbf{e}$, respectively. The elastic material properties are the bulk modulus K and the shear modulus G, while the plastic hardening function is $\kappa(\epsilon^{(p)})$. Finally, for a second rank tensor $\mathbf{a}$, $\mathbf{a}:\mathbf{a} = a_{ij}a_{ij}$.

3 Integration of Eq. (12)

3.1 Integrating Factor. Define the integrating factor:

$$J = \exp\left[3G\int_0^t \frac{\dot{\epsilon}^{(p)}dt}{\kappa(\epsilon^{(p)})}\right] = \exp\left[3G\int_0^{\epsilon^{(p)}} \frac{d\epsilon^{(p)}}{\kappa(\epsilon^{(p)})}\right]. \tag{13}$$

Note that

$$\dot{J} = 3Gd_0^{(p)}J/\kappa(\epsilon^{(p)}). \tag{14}$$

Using the integrating factor J from (13), Eq. (12) can be written as

$$\frac{d}{dt}[J(t)\mathbf{s}] = 2GJ(t)\mathbf{d}. \tag{15}$$

3.2 Examples of Hardening Functions.

Perfect Plasticity. In this case, $\kappa(\epsilon^{(p)}) = Y$, where Y is the constant yield stress in tension. Now

$$J(\epsilon^{(p)}) = \exp\left[\frac{3G\epsilon^{(p)}}{Y}\right]. \tag{16}$$

Linear Workhardening. In this case, $\kappa(\epsilon^{(p)}) = Y + C\epsilon^{(p)}$ with Y and C constant. Now

$$J(\epsilon^{(p)}) = \left[1 + \frac{C\epsilon^{(p)}}{Y}\right]^{3G/C}. \tag{17}$$

3.3 Determination of s:d. Rearranging (12):

$$\mathbf{d} - \frac{\dot{\mathbf{s}}}{(2G)} = \frac{3d_0^{(p)}\mathbf{s}}{2\kappa(\epsilon^{(p)})} \tag{18}$$

Taking the inner product of (18) with $\mathbf{s}$, one has

$$\mathbf{s}:\mathbf{d} = \frac{\mathbf{s}:\dot{\mathbf{s}}}{2G} + \frac{3d_0^{(p)}(\mathbf{s}:\mathbf{s})}{2\kappa(\epsilon^{(p)})}. \tag{19}$$

First Term of (19). From (10) and $(8)_2$

$$\mathbf{s}:\mathbf{s} = (2/3)\kappa^2. \tag{20}$$

Differentiating (20) with respect to time,

$$\mathbf{s}:\dot{\mathbf{s}} = (2/3)\kappa\kappa' d_0^{(p)}. \tag{21}$$

Second Term of (19). Using (20)

$$\frac{3d_0^{(p)}(\mathbf{s}:\mathbf{s})}{2\kappa(\epsilon^{(p)})} = \kappa d_0^{(p)}. \tag{22}$$

Finally, using (19), (21), and (22), one gets

$$\mathbf{s}:\mathbf{d} = \left(\frac{\kappa\kappa'}{3G} + \kappa\right)d_0^{(p)}. \tag{23}$$

4 Minkowski Space-Time

4.1 System of Differential Equations. Define the $m+1$ dimensional vector:

$$x = [x_1, x_2, \ldots, x_m, J]^T = J[s_1, s_2, \ldots, s_m, 1]^T \tag{24}$$

where, for two-dimensional problems, $m=4$ and the tensor $\mathbf{s}$ has components

$$[s] = \begin{bmatrix} s_1 & s_3 \\ s_4 & s_2 \end{bmatrix} \tag{25}$$

and, for three-dimensional problems, $m=9$ and the tensor $\mathbf{s}$ has components

$$[s] = \begin{bmatrix} s_1 & s_6 & s_5 \\ s_9 & s_2 & s_4 \\ s_8 & s_7 & s_3 \end{bmatrix}. \tag{26}$$

Of course, the tensor $\mathbf{s}$ is symmetric, so that $s_3 = s_4$ in (25) and similarly in (26).

The equations for plastic flow can now be written in the form

$$\dot{x} = Ax \tag{27}$$

where, for the two-dimensional case (for example), the explicit form of the matrix A is

$$A = 2G\begin{bmatrix} 0 & 0 & 0 & 0 & d_1 \\ 0 & 0 & 0 & 0 & d_2 \\ 0 & 0 & 0 & 0 & d_3 \\ 0 & 0 & 0 & 0 & d_4 \\ \lambda d_1 & \lambda d_2 & \lambda d_3 & \lambda d_4 & 0 \end{bmatrix} \tag{28}$$

with λ determined in the next subsection. Also, the components of the tensor $\mathbf{d}$ in (28) are written in the same manner as those of $\mathbf{s}$ in (25).

The first four equations of (27) represent equations of (15), while the last one represents (14).

4.2 Determination of λ in Eq. (28). Using (14), the last equation in (27) can be written in explicit form as

$$2GJ\lambda(\mathbf{s}:\mathbf{d}) = \dot{J} = 3GJd_0^{(p)}/\kappa(\epsilon^{(p)}). \tag{29}$$

Finally, using (23) and (29),

$$\lambda(\epsilon^{(p)}) = \frac{3\hat{\gamma}}{2\kappa^2(\epsilon^{(p)})}, \tag{30}$$

where, as in Simo and Taylor [5]

$$\hat{\gamma} = \frac{3G}{3G+\kappa'}. \tag{31}$$

Note that for perfect plasticity $\lambda = 3/(2Y^2)$, and, for this case, (27) is linear!

4.3 Final Form of Differential Equations. With the Minkowski space-time formalism, the system of differential equation for elastoplasticity can be written as

$$\dot{x} = Ax \tag{32}$$

where

$$x = J[\mathbf{s}, 1]^T \tag{33}$$

$$A = 2G\begin{bmatrix} 0_{m\times m} & \mathbf{d} \\ 0_{1\times m} & 0 \end{bmatrix} \quad \text{if } x^T g x < 0 \quad \text{(elastic state)} \tag{34}$$

$$A = 2G\begin{bmatrix} 0_{m\times m} & \mathbf{d} \\ \lambda\mathbf{d}^T & 0 \end{bmatrix} \quad \text{if } x^T g x = 0 \quad \text{(plastic state)} \tag{35}$$

$$g=\begin{bmatrix} I_{m\times m} & 0_{m\times 1} \\ 0_{1\times m} & -(2/3)\kappa^2 \end{bmatrix} \tag{36}$$

with $m=4$ for two-dimensional and $m=9$ for three-dimensional problems. The space consisting of all x endowed with the indefinite metric tensor g is called the Minkowski space-time.

5 Solution Strategies

5.1 Scheme One: A Direct Solution Strategy. Given the state at time t_n, i.e., $\boldsymbol{\sigma}_n$, $\mathbf{d}_n$, $\epsilon_n^{(p)}$ and J_n, and also $\mathbf{d}_{n+1}$ at time t_{n+1}, one can write the matrix A in (32) as

$$A_n=2G\begin{bmatrix} 0_{m\times m} & \mathbf{d}_{n+1} \\ 0_{1\times m} & 0 \end{bmatrix} \quad \text{if } x_n^T g_n x_n<0 \quad (\text{elastic state}) \tag{37}$$

$$A_n=2G\begin{bmatrix} 0_{m\times m} & \mathbf{d}_{n+1} \\ \lambda_n \mathbf{d}_{n+1}^T & 0 \end{bmatrix} \quad \text{if } x_n^T g_n x_n=0 \quad (\text{plastic state}). \tag{38}$$

Solving (32) with the Cayley transform, one has

$$x_{n+1}=J_{n+1}[\mathbf{s}_{n+1},1]^T=G_n x_n=G_n J_n[\mathbf{s}_n,1]^T \tag{39}$$

where

$$G_n=[I-\tau A_n]^{-1}[I+\tau A_n]. \tag{40}$$

With the help of the following formula

$$\begin{bmatrix} I_m & \mathbf{a} \\ \mathbf{b}^T & 1 \end{bmatrix}^{-1}=\begin{bmatrix} I_m+\dfrac{\mathbf{a}\mathbf{b}^T}{1-\mathbf{a}\cdot\mathbf{b}} & \dfrac{-\mathbf{a}}{1-\mathbf{a}\cdot\mathbf{b}} \\ \dfrac{-\mathbf{b}^T}{1-\mathbf{a}\cdot\mathbf{b}} & \dfrac{1}{1-\mathbf{a}\cdot\mathbf{b}} \end{bmatrix} \tag{41}$$

and the use of A_n from (38), one gets

$$G_n=\begin{bmatrix} I+\dfrac{8\tau^2G^2\lambda_n\mathbf{d}_{n+1}\mathbf{d}_{n+1}^T}{1-4\tau^2G^2\lambda_n\mathbf{d}_{n+1}:\mathbf{d}_{n+1}} & \dfrac{4\tau G\mathbf{d}_{n+1}}{1-4\tau^2G^2\lambda_n\mathbf{d}_{n+1}:\mathbf{d}_{n+1}} \\ \dfrac{4\tau G\lambda_n\mathbf{d}_{n+1}^T}{1-4\tau^2G^2\lambda_n\mathbf{d}_{n+1}:\mathbf{d}_{n+1}} & \dfrac{1+4\tau^2G^2\lambda_n\mathbf{d}_{n+1}:\mathbf{d}_{n+1}}{1-4\tau^2G^2\lambda_n\mathbf{d}_{n+1}:\mathbf{d}_{n+1}} \end{bmatrix}. \tag{42}$$

Here $n\in Z^+$, $\mathbf{s}_n$ denotes the value of $\mathbf{s}$ at a discrete time t_n and so on, and τ is one half of the time increment, that is, $\tau=\Delta t/2=(t_{n+1}-t_n)/2$.

Equation (39) gives J_{n+1} and $\mathbf{s}_{n+1}$. Next, $\epsilon_{n+1}^{(p)}$ is obtained from J_{n+1}.

Finally, from (6), (5), (9), and (1), one has

$$\boldsymbol{\sigma}_{n+1}=\mathbf{s}_{n+1}+K\,\mathrm{tr}(\boldsymbol{\epsilon}_{n+1})\mathbf{I}. \tag{43}$$

5.2 Scheme Two: A Numerical Scheme Based on a Nonlinear Volterra Integral Equation. The solution of Eq. (15) is

$$\mathbf{s}(t)=\frac{J(t_i)}{J(t)}\mathbf{s}(t_i)+2G\int_{t_i}^{t}\frac{J(\zeta)}{J(t)}\mathbf{d}(\zeta)d\zeta \tag{44}$$

where one needs to specify the initial value $\mathbf{s}(t_i)$ at the initial time t_i.

From (14), (23), and (44) it follows that

$$\dot{J}(t)=\frac{3G\hat{\gamma}}{\kappa^2}\left[J(t_i)\mathbf{s}(t_i)+2G\int_{t_i}^{t}J(\zeta)\mathbf{d}(\zeta)d\zeta\right]:\mathbf{d}(t). \tag{45}$$

Furthermore, let

$$\dot{Z}(t)=\frac{\kappa^2}{3G\hat{\gamma}}\dot{J}(t)=\frac{\kappa}{\hat{\gamma}}Jd_0^{(p)}. \tag{46}$$

Therefore

$$\dot{Z}(t)=\left[J(t_i)\mathbf{s}(t_i)+2G\int_{t_i}^{t}J(\zeta)\mathbf{d}(\zeta)d\zeta\right]:\mathbf{d}(t). \tag{47}$$

Integrating (47), one gets

$$Z(t)=Z(t_i)+J(t_i)\mathbf{s}(t_i):[\mathbf{e}(t)-\mathbf{e}(t_i)]+2G\int_{t_i}^{t}[\mathbf{e}(t)-\mathbf{e}(\zeta)]:\mathbf{d}(\zeta)J(\zeta)d\zeta, \tag{48}$$

a nonlinear Volterra integral equation for Z.

The three parameters $\epsilon^{(p)}$, J, and Z are homeomorphic. For example, for the linear workhardening material in Section 3.2, it follows that

$$J(\epsilon^{(p)})=\left(1+\frac{C\epsilon^{(p)}}{Y}\right)^{3G/C},$$

$$Z(\epsilon^{(p)})=\frac{(3G+C)Y^2}{3(3G+2C)G}\left[\left(1+\frac{C\epsilon^{(p)}}{Y}\right)^{(3G+2C)/C}-1\right]$$

$$Z(J)=\frac{(3G+C)Y^2}{3(3G+2C)G}[J^{(3G+2C)/(3G)}-1],$$

$$J(Z)=\left[\frac{3(3G+2C)GZ}{(3G+C)Y^2}+1\right]^{3G/(3G+2C)} \tag{49}$$

$$\epsilon^{(p)}(J)=\frac{Y}{C}[J^{C/(3G)}-1],$$

$$\epsilon^{(p)}(Z)=\frac{Y}{C}\left(\left[\frac{3(3G+2C)GZ}{(3G+C)Y^2}+1\right]^{C/(3G+2C)}-1\right).$$

A numerical scheme based on the above formulation is derived next. The discretizations of Eqs. (48) and (44) are obtained by applying the trapezoidal rule for the integrals as follows:

$$Z_{n+1}=Z_n+J_n\mathbf{s}_n:[\mathbf{e}_{n+1}-\mathbf{e}_n]+G\Delta tJ_n[\mathbf{e}_{n+1}-\mathbf{e}_n]:\mathbf{d}_n \tag{50}$$

$$\mathbf{s}_{n+1}=\frac{J_n}{J_{n+1}}\mathbf{s}_n+G\Delta t\left[\mathbf{d}_{n+1}+\frac{J_n}{J_{n+1}}\mathbf{d}_n\right]. \tag{51}$$

The above J_n and J_{n+1} can be calculated through the function $J=J(Z)$ as demonstrated, for example, by Eq. (49) for the linear workhardening material.

5.3 Scheme Three: A Numerical Method Based on Group Theory. The third method is derived as follows. Let

$$\hat{\mathbf{s}}=\frac{\mathbf{s}}{\sqrt{2/3}\kappa} \tag{52}$$

be the normalized stress deviator. From (12), the governing equation for $\hat{\mathbf{s}}$ is found to be

$$\dot{\hat{s}}+\frac{3G+\kappa'}{\kappa}d_0^{(p)}\hat{\mathbf{s}}=\frac{2G}{\sqrt{2/3}\kappa}\mathbf{d} \tag{53}$$

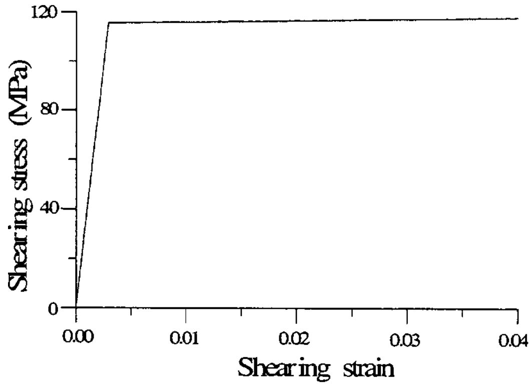
which, upon defining the integrating factor

$$L=\exp\left[\int_0^{\epsilon^{(p)}}\frac{3G+\kappa'(\epsilon^{(p)})}{\kappa(\epsilon^{(p)})}d\epsilon^{(p)}\right]=\frac{\kappa J}{\kappa(0)}, \tag{54}$$

becomes

$$\frac{d}{dt}(L(t)\hat{\mathbf{s}})=\frac{2G}{\sqrt{2/3}\kappa}L(t)\mathbf{d}. \tag{55}$$

For example, for the linear workhardening material in Section 3.2, one has



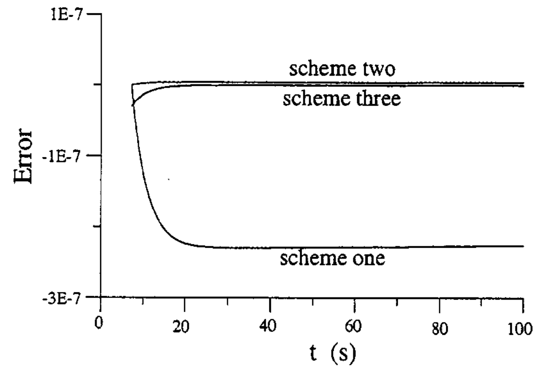


Fig. 1 (*a*) **Shearing stress as a function of shearing strain and** (*b*) **error in satisfying the consistency condition for the three proposed numerical schemes**

$$L(\epsilon^{(p)}) = \left(1 + \frac{C\epsilon^{(p)}}{Y}\right)^{(3G+C)/C}. \tag{56}$$

Also, from (54), (23) and (52), one has

$$\dot{L} = \frac{2G}{\sqrt{2/3}\kappa} L \hat{\mathbf{s}} : \mathbf{d}. \tag{57}$$

Equations (55) and (57) can be jointly written as

$$\dot{X} = \hat{A} X \tag{58}$$

in which

$$X = \begin{bmatrix} L\hat{\mathbf{s}} \\ L \end{bmatrix} \tag{59}$$

is called the augmented deviatoric stress, and

$$\hat{A} = \frac{2G}{\sqrt{2/3}\kappa} \begin{bmatrix} 0_{m\times m} & \mathbf{d} \\ \mathbf{d}^T & 0 \end{bmatrix} \tag{60}$$

satisfying

$$\hat{A}^T \hat{g} + \hat{g} \hat{A} = 0 \tag{61}$$

which is the Lie algebra of $SO_o(m,1)$. Here

$$\hat{g} = \begin{bmatrix} I_{m\times m} & 0_{m\times 1} \\ 0_{1\times m} & -1 \end{bmatrix} \tag{62}$$

is a constant metric of Minkowski spacetime M^{m+1}. Thus, the one-parameter group generated by $\hat{A}$ gives the following transformation formula for X:

$$X(t) = \hat{G}(t) X(0) \tag{63}$$

where $\hat{G}$ is an element of the proper orthochronous Lorentz group $SO_o(m,1)$ satisfying

$$\hat{G}^T \hat{g} \hat{G} = \hat{g} \tag{64}$$

$$\det \hat{G} = 1 \tag{65}$$

$$\hat{G}_0^0 > 0. \tag{66}$$

From (59) and (62) it follows that

$$X^T \hat{g} X = (L)^2 [\hat{\mathbf{s}} : \hat{\mathbf{s}} - 1] \tag{67}$$

which, in view of (52), (7), and (8), leads to

$$X^T \hat{g} X \leqslant 0. \tag{68}$$

The equation $X^T \hat{g} X = 0$ is called the cone condition, which corresponds to the yield condition $\mathbf{s}:\mathbf{s} = (2/3)\kappa^2$. On the other hand, $X^T \hat{g} X < 0$ corresponds to the elastic state.

A numerical scheme based on the group properties can be utilized to enhance computational accuracy and efficiency. The time-centered Euler scheme for Eq. (58) is

$$\begin{bmatrix} L_{n+1}\hat{\mathbf{s}}_{n+1} \\ L_{n+1} \end{bmatrix} = \hat{G}_n \begin{bmatrix} L_n \hat{\mathbf{s}}_n \\ L_n \end{bmatrix} \tag{69}$$

where

$$\hat{G}_n = [I - \tau \hat{A}_n]^{-1} [I + \tau \hat{A}_n]. \tag{70}$$

In the above, from (60), one writes

$$\hat{A}_n = \frac{2G}{\sqrt{2/3}\kappa_n} \begin{bmatrix} 0_{m\times m} & \mathbf{d}_{n+1} \\ \mathbf{d}_{n+1}^T & 0 \end{bmatrix}. \tag{71}$$

It is very important to point out that the $\hat{G}_n$ in (70) fulfills the group properties (64)–(66) as discussed in Liu [11]. It has the following form:

$$\hat{G}_n = \begin{bmatrix} I + \dfrac{12\tau^2 G^2 \mathbf{d}_{n+1} \mathbf{d}_{n+1}^T}{\kappa_n^2 - 6\tau^2 G^2 \mathbf{d}_{n+1} : \mathbf{d}_{n+1}} & \dfrac{2\sqrt{6}\tau G \kappa_n \mathbf{d}_{n+1}}{\kappa_n^2 - 6\tau^2 G^2 \mathbf{d}_{n+1} : \mathbf{d}_{n+1}} \\ \dfrac{2\sqrt{6}\tau G \kappa_n \mathbf{d}_{n+1}^T}{\kappa_n^2 - 6\tau^2 G^2 \mathbf{d}_{n+1} : \mathbf{d}_{n+1}} & \dfrac{\kappa_n^2 + 6\tau^2 G^2 \mathbf{d}_{n+1} : \mathbf{d}_{n+1}}{\kappa_n^2 - 6\tau^2 G^2 \mathbf{d}_{n+1} : \mathbf{d}_{n+1}} \end{bmatrix}. \tag{72}$$

6 A Numerical Example

In order to compare the merits of the different schemes, consider a simple numerical example of a linear workhardening material subjected to simple shearing deformation. In this problem, the strain rate tensor $\mathbf{d}$ is

$$\mathbf{d} = \frac{\dot{\gamma}}{2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{73}$$

where γ is the engineering shearing strain.

Figure 1 compares the results calculated by the three different numerical schemes described in Sections 5.1, 5.2, and 5.3, respectively. The material constants used in these calculations are $G = 20{,}000$ MPa, $Y = 200$ MPa, and $C = 100$ MPa. The time step used in all three calculations is $\Delta t = 0.0001$ s. The three schemes give the same shearing stress-strain curve within plotting accuracy. This (bilinear) curve is shown in Fig. 1(a). Their errors in satisfying the consistency condition, defined as $e=3\mathbf{s}:\mathbf{s}/(2\kappa^2)-1$, are shown in Fig. 1($b$). It can be seen from this figure that all the schemes are very accurate, with the group-preserving scheme giving the smallest error in satisfying the consistency condition, closely followed by the scheme based on the Volterra integral equation.

7 Concluding Remarks

This paper presents a new formulation for isotropic-workhardening rate-independent plasticity, in which the governing equations are reduced to a system of (in general) nonlinear differential equations. Numerical results for a simple illustrative problem demonstrate the accuracy of three different numerical schemes that are proposed here for solving the problem. The authors feel that the scheme based on group theory is particularly promising for the solution of this class of problems.

References

[1] Wilkins, M. L., 1964, ''Calculation of Elastoplastic Flow,'' *Methods of Computational Physics*, **3**, Academic Press, New York.

[2] Hinton, E., and Owen, D. R. J., 1980, *Finite Elements in Plasticity: Theory and Practice*, Pineridge Press, Swansea, Wales.

[3] Pinsky, P. M., Pister, K. S., and Taylor, R. L., 1981, ''Formulation and Numerical Integration of Elasoplastic and Elastoviscoplastic Rate Constitutive Equations,'' Report No. UCB/SESM-81/05, Department of Civil Engineering, University of California, Berkeley, CA.

[4] Nagtegaal, J. C., 1982, ''On the Implementation of Inelastic Constitutive Equations With Special Reference to Large Deformation Problems,'' Comput. Methods Appl. Mech. Eng., **33**, pp. 469–484.

[5] Simo, J. C., and Taylor, R. L., 1985, ''Consistent Tangent Operators for Rate-Independent Elastoplasticity,'' Comput. Methods Appl. Mech. Eng., **48**, pp. 101–118.

[6] Bonnet, M., and Mukherjee, S., 1996, ''Implicit BEM Formulations for Usual and Sensitivity Problems in Elasto-Plasticity Using the Consistent Tangent Operator Concept,'' Int. J. Solids Struct., **33**, pp. 4461–4480.

[7] Poon, H., Mukherjee, S., and Bonnet, M., 1998, ''Numerical Implementation of a CTO-Based Implicit Approach for the BEM Solution of Usual and Sensitivity Problems in Elasto-Plasticity,'' Eng. Anal. Boundary Elem., **22**, pp. 257–269.

[8] Hong, H.-K., and Liu, C.-S., 1999, ''Lorentz Group $SO_o(5,1)$ for Perfect Elastoplasticity With Large Deformation and a Consistency Numerical Scheme,'' Int. J. Non-Linear Mech., **34**, pp. 1113–1130.

[9] Hong, H.-K., and Liu, C.-S., 2000, ''Internal Symmetry in the Constitutive Model of Perfect Elastoplasticity,'' Int. J. Non-Linear Mech., **35**, pp. 447–466.

[10] Liu, C.-S., and Hong, H.-K., 2001, ''Using Comparison Theorem to Compare Corotational Stress Rates in the Model of Perfect Elastoplasticity,'' Int. J. Solids Struct., **38**, pp. 2969–2987.

[11] Liu, C.-S., 2001, ''Cone of Non-Linear Dynamical System and Group Preserving Schemes,'' Int. J. Non-Linear Mech., **36**, pp. 1047–1068.

Y.-Q. Zhang[1]
e-mail: cyqzhang@ntu.edu.sg
School of Civil and Environmental Engineering,
Nanyang Technological University,
Singapore 639798, Singapore

H. Hao
Department of Civil and Resource Engineering,
University of Western Australia,
Crawley, Western Australia 6009

M.-H. Yu
School of Civil Engineering and Mechanics,
Xi'an Jiaotong University,
Xi'an 710049, Xi'an, China

A Unified Characteristic Theory for Plastic Plane Stress and Strain Problems

Based on the unified strength criterion, a characteristic theory for solving the plastic plane stress and plane strain problems of an ideal rigid-plastic body is established in this paper, which can be adapted for a wide variety of materials. Through this new theory, a suitable characteristic method for material of interest can be obtained and the relations among different sorts of characteristic methods can be revealed. Those characteristic methods on the basis of different strength criteria, such as Tresca, von Mises, Mohr-Coulomb, twin shear (TS) and generalized twin shear (GTS), are the special cases (Tresca, Mohr-Coulomb, TS, and GTS) or linear approximation (von Mises) of the proposed theory. Moreover, a series of new characteristic methods can be easily derived from it. Using the proposed theory, the influence of yield criterion on the limit analysis is analyzed. Two examples are given to illustrate the application of this theory.
[DOI: 10.1115/1.1602484]

1 Introduction

The theory of plasticity deals with the methods of calculating stresses and strains in a deformed body after part or all of the body has yielded, and it has been applied to lots of practical problems. As a large number of plastic plane problems exist in engineering practice, they have been drawing much attention from researchers throughout the world ([1–6]).

For plane problems, characteristic methods can be used to solve the quasilinear differential equation systems of stress and velocity fields. Judgements on the types of these differential equation systems can be made using the theory of characteristics. They may be elliptic or hyperbolic, depending on the considered stress state. The methods of characteristics based on the von Mises and Tresca criteria can be found in the literature of Kachanov [7] and Yan [8] for plane strain and plane stress problems. Mandel [9] discussed the method of characteristics based on the Mohr-Coulomb criteria. Yan and Bu [10,11] established the method based on the maximum stress deviator yield criterion (also known as twin-shear (TS) yield criterion ([12]).

The methods of characteristics based on the von Mises, Tresca, and twin-shear criteria can be applied to the limit analysis of the plane problems. However, they are only adapted for the non-SD (strength-differential) materials with $\tau_0 \approx 0.58\sigma_t$, $\tau_0 = 0.5\sigma_t$, and $\tau_0 = 2\sigma_t/3$, respectively, but fail for the SD materials. Although the method of characteristics based on the Mohr-Coulomb criteria takes account of the SD effect, it is only adapted for the material with $\tau_0 = \sigma_t\sigma_c/(\sigma_t + \sigma_c)$.

Normally, the plane strain and plane stress characteristic theories are studied independently. In this paper, a generalized characteristic theory system for solving the plastic plane problem of an ideal rigid-plastic body is established based on the unified strength criterion ([13,14]), which includes the strength-differential effect (SD effect) and can be used for a wide variety of materials. Those characteristic methods on the basis of different strength criteria, such as Tresca, von Mises, Mohr-Coulomb, TS, and generalized TS (GTS) ([15,16]), are the special cases (Tresca, Mohr-Coulomb, TS, and GTS) or linear approximation (von Mises) of the proposed theory. Besides, a series of new characteristic methods can be obtained from it. The theory put forward in this paper can be used conveniently in all sorts of plane strain and plane stress problems.

2 The Unified Strength Criterion

Based on orthogonal octahedron of the twin shear element model ([12]), a unified strength criterion (USC) was developed which specifies that the failure occurs when a certain function of the two larger principal shear stresses and their corresponding normal stresses reach a limit value ([13,14]). The mathematical expression of the USC is

$$F = \tau_{13} + b\tau_{12} + \beta(\sigma_{13} + b\sigma_{12}) = C \quad \text{when} \quad \tau_{12} + \beta\sigma_{12} \geqslant \tau_{23} + \beta\sigma_{23}, \tag{1a}$$

$$F = \tau_{13} + b\tau_{23} + \beta(\sigma_{13} + b\sigma_{23}) = C \quad \text{when} \quad \tau_{12} + \beta\sigma_{12} \leqslant \tau_{23} + \beta\sigma_{23}, \tag{1b}$$

where τ_{13}, τ_{12}, and τ_{23} are principal shear stresses and $\tau_{13} = \frac{1}{2}(\sigma_1 - \sigma_3)$, $\tau_{12} = \frac{1}{2}(\sigma_1 - \sigma_2)$, and $\tau_{23} = \frac{1}{2}(\sigma_2 - \sigma_3)$. σ_{13}, σ_{12}, and σ_{23} are normal stresses corresponding to the three principal shear stresses and $\sigma_{13} = \frac{1}{2}(\sigma_1 + \sigma_3)$, $\sigma_{12} = \frac{1}{2}(\sigma_1 + \sigma_2)$, and $\sigma_{23} = \frac{1}{2}(\sigma_2 + \sigma_3)$, in which σ_1, σ_2, and σ_3 are the principal stresses and $\sigma_1 \geqslant \sigma_2 \geqslant \sigma_3$. β and C are material parameters, and they can be expressed as

$$\beta = \frac{\sigma_c - \sigma_t}{\sigma_c + \sigma_t} = \frac{1-\alpha}{1+\alpha}, \quad C = \frac{(1+b)\sigma_t\sigma_c}{\sigma_c + \sigma_t} = \frac{1+b}{1+\alpha}\sigma_t, \tag{2}$$

where σ_t is uniaxial tensile strength, σ_c is uniaxial compressive strength, and $\alpha = \sigma_t/\sigma_c$ is the ratio of the tensile to the compressive strengths and $0 < \alpha \leqslant 1$. The ratio is an index of the material strength differential effect (SD effect). If σ_t, σ_c and the shear strength τ_0 are chosen as the basic material parameters, through Eq. (2) for pure shear loading, the parameter b can be expressed as

[1]Corresponding author. Tel: +65-67906199; fax: +65-67910676

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 12, 2002; final revision, Jan. 10, 2003. Associate Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

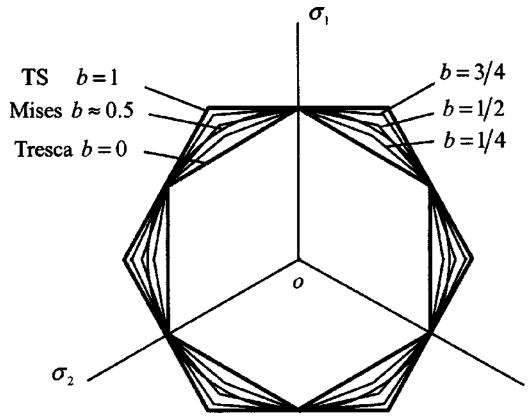


Fig. 1 Different yield loci on the π plane for non-SD materials (α=1)

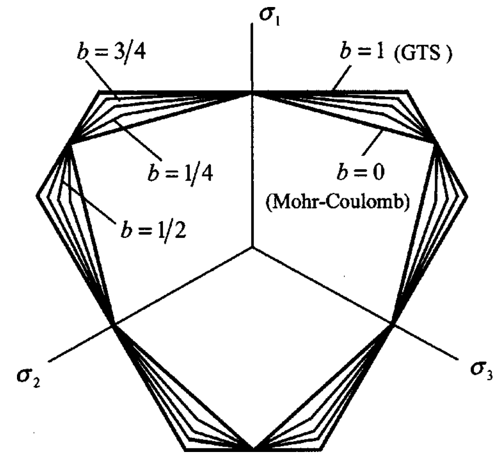


Fig. 2 Different yield loci on the π plane for SD materials (0 <α<1)

$$b=\frac{(\sigma_c+\sigma_t)\tau_0-\sigma_t\sigma_c}{(\sigma_t-\tau_0)\sigma_c}. \tag{3}$$

The USC also can be expressed in terms of principal stresses as follows:

$$F=\sigma_1-\frac{\alpha}{1+b}(b\sigma_2+\sigma_3)=\sigma_t, \quad \text{when } \sigma_2\leqslant\frac{\sigma_1+\alpha\sigma_3}{1+\alpha}, \tag{4a}$$

$$F'=\frac{1}{1+b}(\sigma_1+b\sigma_2)-\alpha\sigma_3=\sigma_t, \quad \text{when } \sigma_2\geqslant\frac{\sigma_1+\alpha\sigma_3}{1+\alpha}. \tag{4b}$$

It should be noted that the parameter b plays an important role in the USC. It builds a bridge among different strength criteria. It is this parameter that distinguishes one criterion from another. On the other hand, the scope of application of each criterion is also represented by this parameter. Hence the USC is not a single strength criterion but a theoretical system including a series of regular strength criteria, and it can be applied to more than one kind of material. In practice, when basic material parameters are obtained by experiments, the value of b can be determined through Eq. (3). Whenever parameter b is obtained, the yield criterion for this sort of material is determined and the application is possible. Consequently, b can be regarded as a parameter by which the suitable yield criterion for material of interest can be determined.

The USC is a series of piecewise linear yield criteria on the π plane as shown in Figs. 1 and 2. The exact form of expression depends on the choice of parameter b. With different choices of parameter b, the USC can be simplified to the Tresca ($\alpha=1$ and $b=0$), the linear approximations of Mises ($\alpha=1$ and $b=1/2$ or $\alpha=1$ and $b=1/(1+\sqrt{3})$), the Mohr-Coulomb ($0<\alpha<1$ and $b=0$), the TS ($\alpha=1$ and $b=1$), the GTS ($0<\alpha<1$ and $b=1$), and a series of new strength criteria. In the stress space, the lower and upper bounds of the yield surfaces on the π plane are special cases of the USC, i.e., $b=0$ ($\alpha=1$ for the Tresca or $0<\alpha<1$ for the Mohr-Coulomb) and $b=1$ ($\alpha=1$ for the TS or $0<\alpha<1$ for the GTS), respectively. When the parameter b varies between 0 and 1, a series of yield surfaces between the two limiting surfaces can be obtained.

3 Characteristics for Plastic Plane Problems

For the cases of plane stress and plane strain, σ_{I} and σ_{III} are assumed to be two principal stresses in the xy plane and $\sigma_{\mathrm{I}}\geqslant\sigma_{\mathrm{III}}$, and σ_{II} is assumed to be the out-of-plane principal stress. Assuming $A=(\sigma_x+\sigma_y)/2=(\sigma_{\mathrm{I}}+\sigma_{\mathrm{III}})/2$ and $B=\sqrt{[(\sigma_x-\sigma_y)/2]^2+\tau_{xy}^2}=(\sigma_{\mathrm{I}}-\sigma_{\mathrm{III}})/2$, the USC in plane state can be expressed as

$$F=mA+nB=\sigma_t, \tag{5}$$

where m and n are material parameters.

In the case of plane stress, the out-of-plane principal stress σ_{II} vanishes. Then, there are three cases to be distinguished in the state of plane stress.

Case A. When $\sigma_{\mathrm{I}}\geqslant\sigma_{\mathrm{III}}\geqslant 0$, it has $\sigma_1=\sigma_{\mathrm{I}}$, $\sigma_2=\sigma_{\mathrm{III}}$, and $\sigma_3=0$. From Eq. (4), we can know that

$$m=\frac{1+b-\alpha b}{1+b}, \quad n=\frac{1+b+\alpha b}{1+b}, \quad \text{when } B\leqslant A\leqslant\frac{2+\alpha}{\alpha}B, \tag{6a}$$

$$m=1, \quad n=\frac{1-b}{1+b}, \quad \text{when } A\geqslant\frac{2+\alpha}{\alpha}B.$$

Case B. When $\sigma_{\mathrm{I}}\geqslant 0\geqslant\sigma_{\mathrm{III}}$, it has $\sigma_1=\sigma_{\mathrm{I}}$, $\sigma_2=0$, and $\sigma_3=\sigma_{\mathrm{III}}$. From Eq. (4), we can know that

$$m=\frac{1+b-\alpha}{1+b}, \quad n=\frac{1+b+\alpha}{1+b}, \quad \text{when } \frac{\alpha-1}{1+\alpha}B\leqslant A\leqslant B, \tag{6b}$$

$$m=\frac{1-\alpha-\alpha b}{1+b}, \quad n=\frac{1+\alpha+\alpha b}{1+b}, \quad \text{when } -B\leqslant A\leqslant\frac{\alpha-1}{1+\alpha}B.$$

Case C. When $0\geqslant\sigma_{\mathrm{I}}\geqslant\sigma_{\mathrm{III}}$, it has $\sigma_1=0$, $\sigma_2=\sigma_{\mathrm{I}}$, and $\sigma_3=\sigma_{\mathrm{III}}$. From Eq. (4), we can know that

$$m=-\alpha, \quad n=\alpha\frac{1-b}{1+b}, \quad \text{when } A\leqslant-(1+2\alpha)B, \tag{6c}$$

$$m=\frac{b-\alpha-\alpha b}{1+b}, \quad n=\frac{b+\alpha+\alpha b}{1+b},$$

$$\text{when } -(1+2\alpha)B\leqslant A\leqslant -B$$

In the case of plane strain of an ideal rigid-plastic body, since the strain rate in the z direction (perpendicular to the xy plane) vanishes, the relation $\sigma_z=\sigma_{\mathrm{II}}=(\sigma_x+\sigma_y)/2=(\sigma_{\mathrm{I}}+\sigma_{\mathrm{III}})/2$ exists. Since the principal stresses $\sigma_1\geqslant\sigma_2\geqslant\sigma_3$, it has $\sigma_1=\sigma_{\mathrm{I}}$, $\sigma_2=\sigma_{\mathrm{II}}$, and $\sigma_3=\sigma_{\mathrm{III}}$. Thus in the plane strain condition, it has

$$m=1-\alpha, \quad n=(1+b+\alpha)/(1+b). \tag{7}$$

As $0<\alpha\leqslant 1$ and $0\leqslant b\leqslant 1$, from Eqs. (6) and (7) it has $n\geqslant 0$.

3.1 Characteristics of Stress Field. Neglecting body force, the equations of equilibrium for plane problems are

$$\frac{\partial\sigma_x}{\partial x}+\frac{\partial\tau_{xy}}{\partial y}=0,\quad \frac{\partial\tau_{xy}}{\partial x}+\frac{\partial\sigma_y}{\partial y}=0 \tag{8}$$

and the stress components can be expressed as

$$\sigma_x=A+B\cos 2\varphi,\quad \sigma_y=A-B\cos 2\varphi,\quad \tau_{xy}=B\sin 2\varphi, \tag{9}$$

where φ is the angle from x axis to the direction of principal stress σ_1 in an anticlockwise way. Combining Eqs. (9), (8), and (5) gives

$$\frac{\partial A}{\partial x}\left(1-\frac{m}{n}\cos 2\varphi\right)+\frac{\partial A}{\partial y}\left(-\frac{m}{n}\sin 2\varphi\right)+2B\left(\frac{\partial\varphi}{\partial y}\cos 2\varphi-\frac{\partial\varphi}{\partial x}\sin 2\varphi\right)=0, \tag{10a}$$

$$\frac{\partial A}{\partial x}\left(-\frac{m}{n}\sin 2\varphi\right)+\frac{\partial A}{\partial y}\left(1+\frac{m}{n}\cos 2\varphi\right)+2B\left(\frac{\partial\varphi}{\partial x}\cos 2\varphi+\frac{\partial\varphi}{\partial y}\sin 2\varphi\right)=0, \tag{10b}$$

Equation (10) is a quasilinear partial differential equation system of the first order. When it has two different real roots, it is of the hyperbolic type and two families of characteristics can be obtained. When it has two equal real roots, it is of the parabolic type and only one family of characteristics exists. When it has no real root, it is of the elliptic type and no characteristics exist.

To utilize the method of characteristics, two supplementary incremental expressions are needed,

$$\frac{\partial A}{\partial x}dx+\frac{\partial A}{\partial y}dy=dA, \tag{11a}$$

$$\frac{\partial\varphi}{\partial x}dx+\frac{\partial\varphi}{\partial y}dy=d\varphi. \tag{11b}$$

Equation (10) together with Eq. (11) make an algebraic equation system with $\partial A/\partial x$, $\partial A/\partial y$, $\partial\varphi/\partial x$, and $\partial\varphi/\partial y$ as unknowns. Let the determinant of coefficients vanish, and obtain

$$\frac{dy}{dx}=\frac{-n\sin 2\varphi\pm\sqrt{n^2-m^2}}{-n\cos 2\varphi+m}. \tag{12}$$

Assuming

$$\cos 2\psi=-\frac{m}{n} \tag{13}$$

then Eq. (12) can be rewritten as

$$\frac{dy}{dx}=\tan(\varphi\mp\psi). \tag{14}$$

As can be seen, the two families of characteristics make angles $\mp\psi$ with the direction of principal stress σ_1. Here those corresponding to the minus sign are assigned as family α and those corresponding to the plus sign as family β. On replacing any column of the coefficient determinant by the right-hand-side terms of Eqs. (10) and (11), it has

$$\pm\sqrt{n^2-m^2}dA-2(\sigma_t-mA)d\varphi=0. \tag{15}$$

When $m\neq 0$, with Eq. (15) it has

$$2m\varphi+\sqrt{n^2-m^2}\ln(\sigma_t-mA)=\text{const},\quad \text{along } \alpha \text{ line}, \tag{16a}$$

$$2m\varphi-\sqrt{n^2-m^2}\ln(\sigma_t-mA)=\text{const},\quad \text{along } \beta \text{ line}. \tag{16b}$$

When $m=0$, with Eq. (15) it has

$$nA-2\sigma_t\varphi=\text{const},\quad \text{along } \alpha \text{ line}, \tag{17a}$$

$$nA+2\sigma_t\varphi=\text{const},\quad \text{along } \beta \text{ line}. \tag{17b}$$

Equations (16) and (17) express the properties of characteristics of the stress field for the plane problems.

3.2 Characteristics of Velocity Field. From the associated flow rule, it has

$$\xi=\eta\frac{\partial F}{\partial\sigma}, \tag{18}$$

where ξ and σ denote strain rate and stress tensors, respectively, and η is a non-negative constant. Under the conditions of small deformation and ideal rigid plasticity, it has

$$\xi_x=\frac{\partial v_x}{\partial x},\quad \xi_y=\frac{\partial v_y}{\partial y},\quad \xi_{xy}=\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}, \tag{19}$$

where v_x and v_y are velocity components in the x and y directions, respectively.

To find the characteristics of velocity field, from Eqs. (18) and (19), we know that

$$\frac{\partial v_x/\partial x}{\partial F/\partial\sigma_x}=\frac{\partial v_y/\partial y}{\partial F/\partial\sigma_y}=\frac{\partial v_x/\partial y+\partial v_y/\partial x}{\partial F/\partial\tau_{xy}}=\eta.$$

Combining the above equation with the yield condition (5) and Eq. (9) gives

$$(m-n\cos 2\varphi)\frac{\partial v_x}{\partial x}-(m+n\cos 2\varphi)\frac{\partial v_y}{\partial y}=0, \tag{20a}$$

$$\cos 2\varphi\left(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}\right)-\sin 2\varphi\left(\frac{\partial v_x}{\partial x}-\frac{\partial v_y}{\partial y}\right)=0. \tag{20b}$$

Similarly, two incremental expressions are

$$\frac{\partial v_x}{\partial x}dx+\frac{\partial v_x}{\partial y}dy=dv_x, \tag{21a}$$

$$\frac{\partial v_y}{\partial x}dx+\frac{\partial v_y}{\partial y}dy=dv_y. \tag{21b}$$

Thus Eqs. (20) and (21) make an algebraic equation system with $\partial v_x/\partial x$, $\partial v_x/\partial y$, $\partial v_y/\partial x$, and $\partial v_y/\partial y$ as unknowns. On assuming the determinant of coefficients equal to zero, it has

$$\frac{dy}{dx}=\frac{-n\sin 2\varphi\pm\sqrt{n^2-m^2}}{m-n\cos 2\varphi}. \tag{22}$$

It is shown that the characteristics of the velocity field just coincide with those of the stress field. The following properties of characteristics of the velocity field can also be obtained by a simple derivation:

$$dv_x+dv_y\tan(\varphi-\psi)=0,\quad \text{along } \alpha \text{ line}, \tag{23a}$$

$$dv_x+dv_y\tan(\varphi+\psi)=0,\quad \text{along } \beta \text{ line}. \tag{23b}$$

The present unified characteristic theory is adapted for the plane stress problems when the parameters m and n are determined from Eq. (6), and it is adapted for the plane strain problems when the parameters m and n are determined from Eq. (7). With different values of parameters b and α, the theory can be simplified to those characteristic methods based on Tresca, Mohr-Coulomb, TS, and GTS criteria and a series of new characteristic methods. When $b=1/(1+\sqrt{3})$ or 1/2, the characteristic field simplified from the unified characteristic field is the linear approximations of that based on von Mises criterion.

4 Application

On the basis of the obtained characteristics of stress and velocity fields, many sorts of plastic plane problems can be studied.

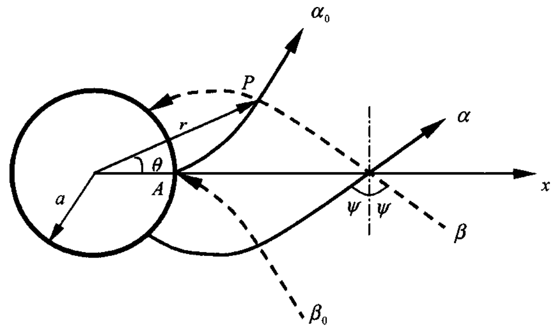


Fig. 3 An infinite thin plate with a circular hole under a two-directional uniform tension at infinity

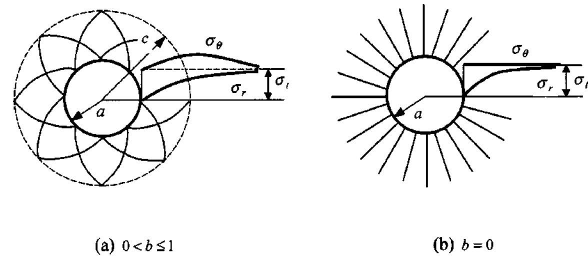


Fig. 4 Characteristics and stress distribution

Example 1: An infinite thin plate, having a circular hole with radius a (Fig. 3), is subjected to a two-directional uniform tension q at infinity. Find the limit load q_s and the corresponding stress distribution.

It is obvious that this is a plane stress problem. As the hole is free and the plate experiences two-directional uniform tension at infinity, on the edge of the hole there will be $\sigma_\theta>0$ and $\sigma_r=0$, and at infinity there will be $\sigma_\theta=\sigma_r>0$. Thus $\sigma_\theta\geqslant\sigma_r\geqslant 0$ holds in the whole plate.

According to the stress state of this problem and from Eqs. (5) and (6), the following yield conditions are available:

$$F=mA+nB=\sigma_t,\quad m=\frac{1+b-\alpha b}{1+b},$$

$$n=\frac{1+b+\alpha b}{1+b}\quad \text{when}\ \ B\leqslant A\leqslant\frac{2+\alpha}{\alpha}B, \tag{24}$$

$$F=mA+nB=\sigma_t,\quad m=1,\quad n=\frac{1-b}{1+b}\quad \text{when}\ \ A\geqslant\frac{2+\alpha}{\alpha}B. \tag{25}$$

Near the hole edge, with the available yield condition (24) and Eq. (13), it has

$$\psi=\frac{1}{2}\cos^{-1}\left(\frac{\alpha b-b-1}{\alpha b+b+1}\right). \tag{26}$$

Then the differential equations of characteristics are

$$\frac{dr}{r\,d\theta}=\pm tg\,\psi=\pm\sqrt{\frac{1+b}{\alpha b}} \tag{27}$$

and equations of characteristics passing through the point A $(r=a,\theta=0)$ will be

$$\theta=\pm\sqrt{\frac{\alpha b}{1+b}}\ln\frac{r}{a}\quad (\text{plus for } \alpha \text{ line, minus for } \beta \text{ line}). \tag{28}$$

Using Eq. (16a) along α_0 line AP (see Fig. 3), we have

$$(1+b-\alpha b)\theta+\sqrt{\alpha b(1+b)}\ln\frac{B_P}{B_A}=0.$$

As $B_A=\sigma_t/2$, thus

$$B_P=\frac{\sigma_t}{2}\left(\frac{a}{r}\right)^{(1+b-\alpha b)/(1+b)},$$

$$A_P=\frac{(1+b)\sigma_t}{2(1+b-\alpha b)}\left[2-\frac{1+b+\alpha b}{1+b}\left(\frac{a}{r}\right)^{(1+b-\alpha b)/(1+b)}\right],$$

and the stress distribution can be obtained as

$$\sigma_\theta=A_P+B_P=\frac{(1+b)\sigma_t}{1+b-\alpha b}\left[1-\frac{\alpha b}{1+b}\left(\frac{a}{r}\right)^{(1+b-\alpha b)/(1+b)}\right],$$

$$\sigma_r=A_P-B_P=\frac{(1+b)\sigma_t}{1+b-\alpha b}\left[1-\left(\frac{a}{r}\right)^{(1+b-\alpha b)/(1+b)}\right] \tag{29}$$

when $a\leqslant r\leqslant c$. With r increasing, the difference between σ_θ and σ_r will decrease and, up to $A=[(2+\alpha)/\alpha]B$, yield condition (24) is to be in the limit of availability. Designate c to denote this radius, it has

$$c=\left[1+\frac{1+b-\alpha b}{\alpha(1+b)}\right]^{(1+b)/(1+b-\alpha b)}a \tag{30}$$

Substituting Eq. (30) into Eq. (29) yields

$$(\sigma_\theta)_c=\frac{(1+\alpha)(1+b)}{1+b+\alpha}\sigma_t,\quad (\sigma_r)_c=\frac{1+b}{1+b+\alpha}\sigma_t. \tag{31}$$

When $r>c$, the yield condition (25) is then available. However, it is obvious that the equation $\psi=\frac{1}{2}\cos^{-1}[(1+b)/(b-1)]$ has no real solution when $0<b\leqslant 1$, and thus the characteristics field is not available. The stress distribution cannot be derived by the method of characteristics. It must be solved directly from the equation of equilibrium

$$\frac{d\sigma_r}{dr}+\frac{\sigma_r-\sigma_\theta}{r}=0 \tag{32}$$

and the yield condition (25). The obtained stress distribution is

$$\sigma_r=(p-\sigma_t)\left(\frac{c}{r}\right)^{(1+b)}+\sigma_t,\quad \sigma_\theta=-b(p-\sigma_t)\left(\frac{c}{r}\right)^{(1+b)}+\sigma_t, \tag{33}$$

where $r>c$, and $p=(\sigma_r)_c$. Thus the limit load

$$q_s=\sigma_r|_{r=\infty}=\sigma_t. \tag{34}$$

For the case of $b=0$, from equation $\psi=\frac{1}{2}\cos^{-1}(-1)=\pi/2$, it can be concluded that two families of characteristics are reduced to a family of characteristics whose direction is the same as that of principal stress σ_r.

The characteristics and stress distribution are shown in Fig. 4, in which variable c denotes the maximum radial of the characteristics, and its value can be calculated from Eq. (30). When $b=1$ and $\alpha=1$, it has $c=2.25a$, and the unified characteristic field is reduced to that based on the TS ([10,11]); When $b=1/(1+\sqrt{3})$ and $\alpha=1$, it has $c=2.117a$, and the characteristic field simplified from the unified characteristic field is the linear approximation of that based on von Mises criterion $(c=2.07a)$ ([7,11]); When $b=0$, the unified characteristic field is reduced to that based on the Mohr-Coulomb criterion, which is the same as that based on Tresca criterion for this special example ([7,11]).

Example 2: An acute wedge is in a plane strain state, with constituent material being ideal rigid plasticity, angle $\gamma<\pi/2$, and surface AB subjected to uniform pressure P_u, as shown in Fig. 5. Determine the limit pressure P_u on surface AB.

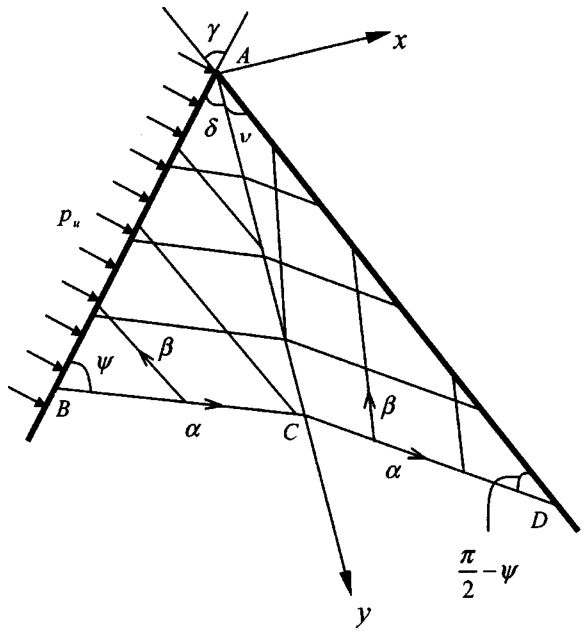


Fig. 5 Acute wedge under unilateral pressure

When $0<\gamma<\pi/2$, a stress discontinuous line in the wedge will appear ([7,8,17,18]). The characteristic field is shown in Fig. 5, where $\angle BAC=\delta$, $\angle CAD=\nu$, $\delta+\nu=\gamma$. The regions ABC and ACD are regions of constant biaxial compression and uniaxial compression, respectively ([17,18]). The constant stress regions ABC and ACD are separated by the line of stress discontinuity AC which is inclined to AB at an angle δ to be determined. The different values a quantity may assume in the regions ABC and ACD will be distinguished by subscripts 1 and 2, respectively. The angle between characteristics α and β is 2ψ.

Using the stress condition on the stress discontinuous line ([7,8,17,18]), it has

$$A_1+B_1\cos 2\varphi_1=A_2+B_2\cos 2\varphi_2,\quad B_1\sin 2\varphi_1=B_2\sin 2\varphi_2. \tag{35}$$

From Eq. (5) it has

$$mA_1+nB_1=mA_2+nB_2. \tag{36}$$

Hence, with Eqs. (35) and (36), it has

$$m\cos(\varphi_1-\varphi_2)=n\cos(\varphi_1+\varphi_2). \tag{37}$$

From the stress boundary condition of the wedge, we have $\sigma_3=-p_u$ in region ABC, and $\sigma_1=0$ in region ACD. Thus it has

$$\varphi_1=\frac{\pi}{2}-\delta,\quad \varphi_2=\nu, \tag{38}$$

and

$$B_1=(\sigma_t+mp_u)/(n+m),\quad B_2=\sigma_t/(n-m). \tag{39}$$

Substituting Eq. (38) into Eq. (37) yields

$$m\sin(\delta+\nu)=n\sin(\delta-\nu). \tag{40}$$

It also has

$$\delta+\nu=\gamma. \tag{41}$$

Using Eqs. (40) and (41), the values of δ and ν can be derived.

Substituting Eqs. (38) and (39) into Eq. (35) gives the unified limit load

$$p_u=\frac{(n+m)\sigma_t}{m(n-m)}\frac{\sin 2\nu}{\sin 2\delta}-\frac{\sigma_t}{m}, \tag{42}$$

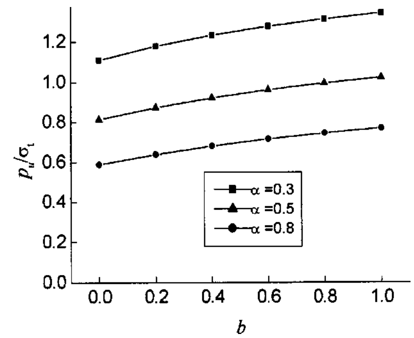


Fig. 6 The relation between the limit load p_u and the parameter b

where parameters m and n are determined by Eq. (7). Since it is given by Ref. [15] that

$$\alpha=\frac{1-\sin\varphi}{1+\sin\varphi},\quad \sigma_t=\frac{2c\cos\varphi}{1+\sin\varphi}, \tag{43}$$

where φ and c are angle of internal friction and cohesion, respectively, substituting Eqs. (7) and (43) into Eq. (42) and introducing two parameters φ_t and c_t, the unified limit load can be rewritten as

$$p_u=c_t\cot\varphi_t\left(\frac{1+\sin\varphi_t}{1-\sin\varphi_t}\frac{\sin 2\nu}{\sin 2\delta}-1\right), \tag{44}$$

where φ_t and c_t are defined as

$$\sin\varphi_t=\frac{2(b+1)\sin\varphi}{2+b(1+\sin\varphi)},\quad c_t=\frac{2(b+1)c\cos\varphi}{2+b(1+\sin\varphi)}\cdot\frac{1}{\cos\varphi_t}. \tag{45}$$

For the case of $b=0$ and $\alpha\neq 1$, the unified limit load is reduced to

$$p_u'=c\cot\varphi\left(\frac{1+\sin\varphi}{1-\sin\varphi}\frac{\sin 2\nu}{\sin 2\delta}-1\right). \tag{46}$$

This is the solution on the basis of Mohr-Coulomb criterion ([17,18]). For the case of $b=0$ and $\alpha=1$, the unified limit load is reduced to

$$p_u''=\lim_{\substack{\alpha\to 1\\(\varphi\to 0)}} p_u'=2c(1-\cos\gamma). \tag{47}$$

This is the solution on the basis of Tresca criterion ([18]).

When $\gamma=\pi/3$, the relation between the limit load p_u and α are shown in Fig. 6. It can be found that the SD effect of material and the influence of intermediate principal stress on the limit load are significant. Through the parameters α and b, the dependence of the result of the limit load on yield criterion is also reflected. As shown, at the same α, the Mohr-Coulomb criterion $(b=0)$ leads to the minimum value of p_u/σ_t while the GTS $(b=1)$ leads to the maximum value of p_u/σ_t.

5 Conclusions

On the basis of the unified strength criterion, a new characteristic theory for solving the plane stress and strain problems of an ideal rigid-plastic body is established in this paper, which includes the SD effect and can be applied to a wide variety of materials.

This new theory can be conveniently applied to the plane strain and the plane stress problems. It builds a bridge among different sorts of characteristic methods. Those characteristic methods on the basis of different strength criteria, such as Tresca, von Mises, Mohr-Coulomb, TS, and GTS, are the special cases (Tresca, Mohr-Coulomb, TS, and GTS) or linear approximation (von

Mises) of the proposed theory. The influence of yield criterion on the limit analysis is also reflected. Besides, a series of new characteristic methods can be obtained from it.

In practice, when basic material parameters are obtained by experiments, the value of b can be determined. With the parameter b, the characteristic method suitable for the material of interest can be determined.

References

[1] Johnson, W., Sowerby, R., and Venter, R. D., 1982, *Plane Strain Slip Line Fields for Metal Deformation Processes—A Source Book and Bibliography*, Pergamon Press, Oxford.

[2] Lee, Y. K., 1989, ''Conditions for Shear Banding and Material Instability in Finite Elastoplastic Deformation,'' Int. J. Plast., **5**, pp. 197–226.

[3] Runesson, K., Ottosen, N., and Peric, D., 1991, ''Discontinuous Bifurcations of Elastic-Plastic Solutions at Plane Stress and Plane Strain,'' Int. J. Plast., **7**, pp. 99–121.

[4] Collins, I. F., and Dewhurst, P., 1993, ''A Matrix Technique for Constructing Slip-Line Field Solutions to a Class of Plane-Strain Plasticity Problems,'' Int. J. Numer. Methods Eng., **7**, pp. 357–378.

[5] Panayotounakos, D., 1996, ''Ad Hoc Exact Solutions for the Stress and Velocity Fields in Rigid Perfectly Plastic Materials Under Plane-Strain Conditions,'' Int. J. Plast., **12**, pp. 1069–1190.

[6] Panayotounakos, D., 1999, ''Ad Hoc Exact Solutions for the Stress and Velocity Fields in Rigid Perfectly Plastic Materials Under Plane-Stress Conditions,'' Int. J. Non-Linear Mech., **34**, pp. 71–84.

[7] Kachanov, L. M., 1971, *Foundations of the Theory of Plasticity*, North-Holland Publication Co., Amsterdam.

[8] Yan, Z., 1988, *Theory of Plasticity*, Tianjin University Publishing House, Tianjin (in Chinese).

[9] Lubliner, J., 1990, *Plasticity Theory*, Macmillan Publishing Company, New York.

[10] Yan, Z., and Bu, X., 1992, ''The Method of Characteristics for Solving the Plane Stress Problem of Ideal Rigid-Plastic Body on the Basis of Twin Shear Stress Yield Criterion,'' *Proc. of the Asia-Pacific Symposium on Advances in Engineering Plasticity and Its Application*, Elsevier, Hong Kong.

[11] Yan, Z., and Bu, X., 1996, ''An Effective Characteristic Method for Plastic Plane Stress Problems,'' J. Eng. Mech., **122**, pp. 502–506.

[12] Yu, M., 1983, ''Twin Shear Stress Yield Criterion,'' Int. J. Mech. Sci., **25**, pp. 71–74.

[13] Yu, M., 2002, ''Advance in Strength Theory of Material and Complex Stress State in the 20th Century,'' Appl. Mech. Rev., **55**, pp. 169–218.

[14] Fan, S. C., Yu, M. H., and Yang, S. Y., 2001, ''On the Unification of Yield Criteria,'' J. Appl. Mech., **68**, pp. 341–343.

[15] Yu, M., He, L., and Song, L., 1985, ''Twin Shear Strength Theory and its Generalization,'' Sci. Sin., Ser. A, **28**, pp. 1174–1183.

[16] Yu, M., and He, L., 1991, ''A New Model and Theory on Yield and Failure of Materials Under Complex Stress State,'' *Mechanical Behavior of Materials* ~6, Pergamon Press, New York, **3**, pp. 841–846.

[17] Shield, R. T., 1954, ''Stress and Velocity Fields in Soil Mechanics,'' J. Math. Phys., **33**, pp. 144–156.

[18] Chen, W. F., 1975, *Limit Analysis and Soil Plasticity*, Elsevier, Amsterdam.

J. H. Davies
e-mail: jdavies@elec.gla.ac.uk;
URL: www.elec.gla.ac.uk/~jdavies
Department of Electronics
and Electrical Engineering,
Glasgow University,
Glasgow, G12 8QQ, U.K.

Elastic Field in a Semi-Infinite Solid due to Thermal Expansion or a Coherently Misfitting Inclusion

It is shown that the elastic field due to nonuniform temperature or a coherently misfitting inclusion in a semi-infinite region can be derived simply from the corresponding field in an infinite region. This follows from the work of Mindlin and Cheng [J. Appl. Phys. **21**, *931 (1950)] but it is not necessary to calculate the thermoelastic potential itself. In particular, the displacement of the free surface is the same as that of the equivalent plane in an infinite solid, increased by a factor of 4(1−ν). The change in volume associated with the distortion of the surface is reduced by a factor of 2(1+ν)/3 from the free expansion of the inclusion. A rectangular inclusion is used to illustrate the theory.*
[DOI: 10.1115/1.1602481]

1 Introduction

The elastic field due to nonuniform thermal expansion in a semi-infinite solid is of considerable practical importance as well as being a basic problem of thermoelasticity ([1,2]). Another widely occurring issue is stress due to a coherently misfitting inclusion, which can be treated in the same way provided that the elastic constants are identical in the inclusion and its surroundings. Semiconductor technology provides plentiful examples of inclusions. Selective oxidation is an important aspect of complementary metal-oxide semiconductor (CMOS) technology where stress arises from the mismatch between silicon and its oxide, an isolation trench of rectangular cross section ([3,4]) being one geometry. Most advanced III–V devices are pseudomorphic, containing mismatched active regions ([5,6]). The composition of nominally uniform layers may fluctuate both laterally ([7,8]) and normal to the layers ([9,10]). In particular, there is great interest in the elastic field around both wires ([11–13]) and dots ([14–18]). A direct approach for investigating this strain is scanning probe microscopy of a surface cleaved through the structure ([10,19]). The results should be compared with a solution of the elastic field in a semi-infinite solid, rather than the infinite solid that is used in most calculations and which is appropriate for the structure before cleavage.

Several approaches have been used for an analytical solution of the thermoelastic problem in a semi-infinite solid. Chiu [20,21] solved the problem of a rectangular parallelepiped by using the method of images to satisfy most of the boundary conditions on the free surface and adding a further elastic field to satisfy the remainder. He considered a more general initial strain but the solution is cumbersome. Glas [7,8] used the same general approach for the pure thermoelastic problem with much more tractable results, which were applied to the rectangular parallelepiped and a step. The basic problem is the elastic field of a center of dilatation, which was solved by Mindlin and Cheng [22,23] and Sen [24]. The former solution is particularly attractive because it can be written in terms of the thermoelastic potential in an *infinite* solid ([25]). Unfortunately the potential itself tends to be difficult to calculate.

In this paper Mindlin and Cheng's solution is developed to show that the elastic quantities for the semi-infinite solid can be written in terms of those for the infinite solid and their derivatives normal to the surface. For example, the displacement in the semi-infinite solid follows directly from the displacement and strains in the infinite solid, which are usually easier to calculate than the potential. There is a remarkably simple result for the surface itself: The displacement has the same form as the equivalent plane in an infinite solid, increased by a factor of $4(1-\nu)$, where ν is Poisson's ratio. The strains at the surface follow trivially from this result and the absence of traction. A simple result for the change in volume also follows and can be used to deduce the volume of a buried inclusion. The example of a buried rectangular region will be considered in some detail, including the limit of a semi-infinite slab, because of its numerous applications.

2 Theory

Consider the half space $z \geqslant 0$ where the plane $z=0$ is free of traction. Stress may arise from the linear thermal expansion $\alpha T(\mathbf{r})$ of a region within this half space relative to its surroundings. Alternatively, $\alpha T(\mathbf{r})$ may be replaced by the linear fractional mismatch $\varepsilon_0(\mathbf{r})$ for an inclusion and this terminology will be used henceforth. Some examples are shown in Fig. 1, either completely buried or exposed on the free surface. The displacement and strains are measured with respect to the unheated state or the surroundings of the inclusion; the misfit $\varepsilon_0(\mathbf{r})$ should be subtracted inside an inclusion from the tensile strains calculated here if they are to be measured with respect to the natural state of the local material. It is assumed that the elastic response is linear and isotropic with the same constants everywhere, and piezoelectric effects are neglected.

2.1 Infinite Solid. For the same inclusion in an *infinite* solid, Goodier [25,26] showed that the displacement $\mathbf{u}^{(\infty)}(\mathbf{r})$ can be written as $4\pi\mathbf{u}^{(\infty)} = -\nabla\varphi$ where the potential φ obeys Poisson's equation,

$$\nabla^2\varphi = -4\pi\frac{1+\nu}{1-\nu}\varepsilon_0 = -4\pi\varepsilon^{(\infty)}. \tag{1}$$

The usual integral solution is

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 12, 2002; final revision, Jan. 21, 2003. Associate Editor: J. R. Barber. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

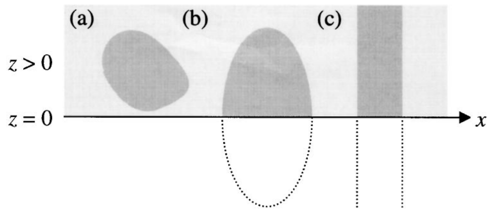


Fig. 1 Inclusions (dark gray) within the half space $z>0$ (light gray). Inclusions may be (*a*) fully buried or (*b*) exposed on the surface, and (*c*) shows a semi-infinite slab.

$$\varphi(\mathbf{r})=\frac{1+\nu}{1-\nu}\int\frac{\varepsilon_0(\mathbf{r}')d^3\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}. \tag{2}$$

This integral is performed over the region where $\varepsilon_0(\mathbf{r}')\neq 0$, which must lie in the half space $z\geqslant 0$. In the case of an inclusion that has been fractured by the surface (Fig. 1(b)), only the region that remains inside the body must be included.

This elastic field has some simple properties because it is derived from a scalar potential.

1. The dilatation $\varepsilon^{(\infty)}$ is proportional to the *local* value of ε_0 as shown in Eq. (1). It therefore vanishes outside the inclusion. This includes the region $z<0$, which means that $\nabla^2\varphi=0$ for $z<0$.
2. The derivatives of the displacement are symmetric, so that $\partial u_x^{(\infty)}/\partial y=(-1/4\pi)\partial^2\varphi/\partial x\partial y=\partial u_y^{(\infty)}/\partial x$ for example.
3. Derivatives of the strain can similarly be rewritten, such as $\partial\varepsilon_{xx}^{(\infty)}/\partial y=(-1/4\pi)\partial^3\varphi/\partial x^2\partial y=\partial\varepsilon_{xy}^{(\infty)}/\partial x$.

These properties will be used to simplify the results for the semi-infinite solid.

2.2 Semi-Infinite Solid. Mindlin and Cheng [22,23] showed that the displacement $\mathbf{u}(\mathbf{r})$ in the semi-infinite region $z\geqslant 0$ with a free surface at $z=0$ can be written in the form

$$4\pi\mathbf{u}=-\boldsymbol{\nabla}\varphi-\boldsymbol{\nabla}_2\varphi_2. \tag{3}$$

Here φ is the potential for the infinite system, defined in the previous section. The second potential φ_2 is equal to φ evaluated at the mirror image of the point in the plane $z=0$. Thus $\varphi_2(x,y,z)=\varphi(x,y,-z)\equiv\overline{\varphi}(x,y,z)$, where the bar denotes that the sign of z is changed. Finally, the second vector operator is defined by

$$\boldsymbol{\nabla}_2=(3-4\nu)\boldsymbol{\nabla}+2\boldsymbol{\nabla}z\frac{\partial}{\partial z}-4(1-\nu)\hat{\mathbf{k}}\nabla^2 z, \tag{4}$$

where $\hat{\mathbf{k}}$ is a unit vector in the z direction.

The first step is to simplify $\boldsymbol{\nabla}_2$ using the property that $\nabla^2\overline{\varphi}=0$ in the region of interest (property 1 above). This gives $\nabla^2 z\overline{\varphi}=2\partial\overline{\varphi}/\partial z$. The middle term can also be reordered using $\boldsymbol{\nabla}z=z\boldsymbol{\nabla}+\hat{\mathbf{k}}$. Regrouping terms gives

$$\boldsymbol{\nabla}_2\overline{\varphi}=(3-4\nu)\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},-\frac{\partial}{\partial z}\right)\overline{\varphi}+2z\frac{\partial}{\partial z}\boldsymbol{\nabla}\overline{\varphi}. \tag{5}$$

The first term resembles a "twinned gradient" ([27]).

The derivatives can now be expressed in terms of $\overline{\mathbf{u}}^{(\infty)}$. This is defined in the same way as the potential by the displacement at the image point, $\overline{\mathbf{u}}^{(\infty)}(x,y,z)=\mathbf{u}^{(\infty)}(x,y,-z)$. Note that this is *not* the same as the mirror image of the displacement itself, which would entail a change in sign of the component along z. Care is needed with the signs when derivatives of z are taken; $\overline{u}_x^{(\infty)}=(-1/4\pi)\partial\overline{\varphi}/\partial x$ but the changed sign of z in $\overline{\varphi}$ means that $\overline{u}_z^{(\infty)}=(+1/4\pi)\partial\overline{\varphi}/\partial z$. The displacement within the semi-infinite region is thus found to be

$$\mathbf{u}=\mathbf{u}^{(\infty)}+(3-4\nu)\overline{\mathbf{u}}^{(\infty)}+2z\frac{\partial}{\partial z}(\overline{u}_x^{(\infty)},\overline{u}_y^{(\infty)},-\overline{u}_z^{(\infty)}) \tag{6}$$

$$=\mathbf{u}^{(\infty)}+(3-4\nu)\overline{\mathbf{u}}^{(\infty)}-2z(\overline{\varepsilon}_{xz}^{(\infty)},\overline{\varepsilon}_{yz}^{(\infty)},-\overline{\varepsilon}_{zz}^{(\infty)}). \tag{7}$$

The second expression is written in terms of the strains at the image point, which are defined in the same way as the other functions. This is the central result of this paper, and shows that the displacement in the semi-infinite solid follows directly from the displacement in the infinite solid and its derivative normal to the surface.

2.3 Distortion of Free Surface. The distortion of the free surface follows by setting $z=0$ in Eq. (7). The term with the strains vanishes, the plain and barred displacements in the infinite region coincide on this plane, and the displacement of the surface is therefore

$$\mathbf{u}(x,y,0)=4(1-\nu)\mathbf{u}^{(\infty)}(x,y,0). \tag{8}$$

Thus the displacement of the free surface is given by that of the same plane within an infinite medium, increased by a factor of $4(1-\nu)$. This factor is greater than 2 (except for $\nu=1/2$), a curious feature that will be discussed in Sec. 4.1. The strains in the plane of the surface, ε_{xx}, ε_{xy}, and ε_{yy}, are related to those in an infinite sample by the same factor. The slope of the surface is

$$\frac{\partial u_z}{\partial x}=4(1-\nu)\frac{\partial u_z^{(\infty)}}{\partial x}=4(1-\nu)\varepsilon_{xz}^{(\infty)}(x,y,0), \tag{9}$$

with a similar result for $\partial u_z/\partial y$. These results contain the shear strain in the *infinite* region; the corresponding strains at the surface of the semi-infinite region vanish because there is no traction. The remaining strain, ε_{zz}, follows from the absence of traction and the stress-strain relations for thermoelasticity. These take the usual form

$$E(\varepsilon_{xx}-\varepsilon_0)=\sigma_{xx}-\nu(\sigma_{yy}+\sigma_{zz}) \tag{10}$$

and permutations, where E is Young's modulus. The result is

$$\varepsilon_{zz}(x,y,0)=4\nu\varepsilon_{zz}^{(\infty)}+(1-4\nu)\varepsilon^{(\infty)}. \tag{11}$$

Finally, addition of these strains shows that the dilatation on the surface is

$$\varepsilon(x,y,0)=(5-8\nu)\varepsilon^{(\infty)}-4(1-2\nu)\varepsilon_{zz}^{(\infty)}. \tag{12}$$

There would be no dilatation in the infinite region if the plane of the surface lay outside the inclusion, but this is not true at the surface of the semi-infinite region because of the second term.

2.4 Strain. Expressions for the components of the strain throughout the region $z\geqslant 0$ can be derived by differentiation of the displacement in Eq. (7). The results are

$$\varepsilon_{xx}=\varepsilon_{xx}^{(\infty)}+(3-4\nu)\overline{\varepsilon}_{xx}^{(\infty)}+2z\frac{\partial\overline{\varepsilon}_{xx}^{(\infty)}}{\partial z}, \tag{13}$$

$$\varepsilon_{zz}=\varepsilon_{zz}^{(\infty)}-(1-4\nu)\overline{\varepsilon}_{zz}^{(\infty)}+2z\frac{\partial\overline{\varepsilon}_{zz}^{(\infty)}}{\partial z}, \tag{14}$$

$$\varepsilon_{xz}=\varepsilon_{xz}^{(\infty)}-\overline{\varepsilon}_{xz}^{(\infty)}-2z\frac{\partial\overline{\varepsilon}_{xz}^{(\infty)}}{\partial z}, \tag{15}$$

$$\varepsilon=\varepsilon^{(\infty)}-4(1-2\nu)\overline{\varepsilon}_{zz}^{(\infty)}. \tag{16}$$

Of the other strains, ε_{xy} and ε_{yy} follow the same pattern as ε_{xx}, and ε_{yz} has the same form as ε_{xz}. The freedom to interchange derivatives and subscripts on the strain in an infinite region (property 3 above) has been used so that all derivatives are taken with respect to z. The result for the dilatation is simpler because $\overline{\varepsilon}^{(\infty)}=0$. This expression shows clearly that a dilatation of $-4(1-2\nu)\overline{\varepsilon}_{zz}^{(\infty)}$ arises from the influence of the surface. Thus the dilatation no longer vanishes outside the inclusion, nor is it constant within an inclusion of constant mismatch.

These values of the strain should agree at the free surface with those given previously. This is straightforward to confirm for ε_{xx} and other components in the xy plane. It is also clear that ε_{xz}

$=\varepsilon_{yz}=0$ when $z=0$, which ensures the absence of shear traction. Care is needed for ε_{zz}, however, if the inclusion is exposed on the surface (Fig. 1(b)). In this case $\varepsilon_{zz}^{(\infty)}$ should be the limiting value as $z \to 0$ *inside* the inclusion, while $\overline{\varepsilon}_{zz}^{(\infty)}$ always lies *outside* the inclusion. These two values differ by

$$\overline{\varepsilon}_{zz}^{(\infty)}(x,y,0)=\varepsilon_{zz}^{(\infty)}(x,y,0)-\varepsilon^{(\infty)}, \tag{17}$$

where $\varepsilon^{(\infty)}=[(1+\nu)/(1-\nu)]\varepsilon_0$ from Eq. (1). Thus Eqs. (11) and (14) are consistent for the normal strain on the surface, which can also be written as

$$\varepsilon_{zz}(x,y,0)=4\nu\overline{\varepsilon}_{zz}^{(\infty)}+\varepsilon^{(\infty)}. \tag{18}$$

This expression is useful because it is often easier to easier to calculate the elastic field of the infinite system outside the inclusion. Likewise, Eqs. (12) and (16) are consistent for the dilatation on the surface.

2.5 Stress. Similar expressions can be derived for components of the stress in terms of those for the same inclusion in an infinite region. The results are simplified by using relations of the form $\sigma_{xx}^{(\infty)}=2G(\varepsilon_{xx}^{(\infty)}-\varepsilon^{(\infty)})$ for the thermoelastic field in an infinite region, where $G=E/2(1+\nu)$ is the shear modulus. This leads to

$$\sigma_{xx}=\sigma_{xx}^{(\infty)}+(3-4\nu)\overline{\sigma}_{xx}^{(\infty)}-4\nu\overline{\sigma}_{zz}^{(\infty)}+2z\frac{\partial\overline{\sigma}_{xx}^{(\infty)}}{\partial z}, \tag{19}$$

$$\sigma_{zz}=\sigma_{zz}^{(\infty)}-\overline{\sigma}_{zz}^{(\infty)}+2z\frac{\partial\overline{\sigma}_{zz}^{(\infty)}}{\partial z}, \tag{20}$$

$$\sigma_{xy}=\sigma_{xy}^{(\infty)}+(3-4\nu)\overline{\sigma}_{xy}^{(\infty)}+2z\frac{\partial\overline{\sigma}_{xy}^{(\infty)}}{\partial z}, \tag{21}$$

$$\sigma_{xz}=\sigma_{xz}^{(\infty)}-\overline{\sigma}_{xz}^{(\infty)}-2z\frac{\partial\overline{\sigma}_{xz}^{(\infty)}}{\partial z}. \tag{22}$$

The expression for σ_{yy} is similar to that for σ_{xx} and σ_{yz} is similar to σ_{xz}. Addition shows that the sum of the normal stresses, $\Theta=\sigma_{xx}+\sigma_{yy}+\sigma_{zz}$, is

$$\Theta=\Theta^{(\infty)}-4(1+\nu)\overline{\sigma}_{zz}^{(\infty)}. \tag{23}$$

As in the case of the dilatation, Θ would vanish outside an inclusion in an infinite region but the surface causes the second term to appear.

Care is again needed at the free surface of an exposed inclusion because of a possible discontinuity in $\sigma_{xx}^{(\infty)}$, given by

$$\overline{\sigma}_{xx}^{(\infty)}=\sigma_{xx}^{(\infty)}-\frac{1}{2}\Theta^{(\infty)}=\sigma_{xx}^{(\infty)}+\frac{E\varepsilon_0}{1-\nu}. \tag{24}$$

There is a similar expression for $\overline{\sigma}_{yy}^{(\infty)}$. The stresses on the surface are thus found to be

$$\begin{aligned}\sigma_{xx}(x,y,0)&=4(1-\nu)\sigma_{xx}^{(\infty)}-4\nu\sigma_{zz}^{(\infty)}-\frac{1}{2}(3-4\nu)\Theta^{(\infty)}\\&=4(1-\nu)\overline{\sigma}_{xx}^{(\infty)}-4\nu\overline{\sigma}_{zz}^{(\infty)}-\frac{E\varepsilon_0}{1-\nu},\end{aligned} \tag{25}$$

$$\sigma_{xy}(x,y,0)=4(1-\nu)\sigma_{xy}^{(\infty)}, \tag{26}$$

$$\Theta(x,y,0)=\Theta^{(\infty)}-4(1+\nu)\sigma_{zz}^{(\infty)}=-\frac{2E\varepsilon_0}{1-\nu}-4(1+\nu)\overline{\sigma}_{zz}^{(\infty)}. \tag{27}$$

The expression for σ_{yy} is similar to that for σ_{xx} while $\sigma_{xz}=\sigma_{yz}=\sigma_{zz}=0$ because of the absence of traction.

3 Center of Dilatation

To verify the above results, consider a center of dilatation located at $\mathbf{r}_0=(0,0,c)$ with $\varepsilon_0(\mathbf{r})=S\delta(\mathbf{r}-\mathbf{r}_0)$. The potential is given by the solution of Eq. (1),

$$\varphi(\mathbf{r})=\frac{1+\nu}{1-\nu}\frac{S}{R_1}, \tag{28}$$

where $R_1^2=|\mathbf{r}-\mathbf{r}_0|^2=x^2+y^2+(z-c)^2$. The displacement in an infinite region and that at the image point are

$$\mathbf{u}^{(\infty)}=\frac{1+\nu}{1-\nu}\frac{S}{4\pi}\frac{(x,y,z-c)}{R_1^3}, \tag{29}$$

$$\overline{\mathbf{u}}^{(\infty)}=\frac{1+\nu}{1-\nu}\frac{S}{4\pi}\frac{(x,y,-z-c)}{R_2^3}, \tag{30}$$

where $R_2^2=x^2+y^2+(z+c)^2$. The corresponding strains are given by expressions of the form

$$\varepsilon_{zz}^{(\infty)}=\frac{1+\nu}{1-\nu}\frac{S}{4\pi}\left[\frac{1}{R_2^3}-\frac{3(z-c)^2}{R_1^5}\right], \tag{31}$$

$$\varepsilon_{xz}^{(\infty)}=\frac{1+\nu}{1-\nu}\frac{S}{4\pi}\frac{-3x(z-c)}{R_1^5}. \tag{32}$$

Again the barred strains are obtained by changing the sign of z and writing R_2 instead of R_1. The displacement in the semi-infinite region can then be obtained from Eq. (7), using the displacement and strains above for the infinite region, which gives

$$u_x=\frac{1+\nu}{1-\nu}\frac{S}{4\pi}\left[\frac{x}{R_1^3}+(3-4\nu)\frac{x}{R_2^3}-\frac{6xz(z+c)}{R_2^5}\right], \tag{33}$$

$$u_z=\frac{1+\nu}{1-\nu}\frac{S}{4\pi}\left[\frac{z-c}{R_1^3}-(3-4\nu)\frac{z+c}{R_2^3}+\frac{2z}{R_2^3}-\frac{6z(z+c)^2}{R_2^5}\right]. \tag{34}$$

This agrees with Mindlin and Cheng [23], Sen [24], and Hu [3].

The displacement at the surface reduces to

$$\mathbf{u}(x,y,0)=\frac{(1+\nu)S}{\pi}\frac{(x,y,-c)}{(x^2+y^2+c^2)^{3/2}}. \tag{35}$$

A more direct derivation of this result was given by Barber [28], who used potential functions ([29]) rather than the Galerkin vector. Comparison with the result for an infinite region, Eq. (29), at $z=0$ confirms that the displacement of the surface of the semi-infinite region is larger by a factor of $4(1-\nu)$, in agreement with Eq. (8).

3.1 Total Change in Volume. The total change in volume Ω due to distortion of the free surface is given by integration of the displacement along z in Eq. (35) over the plane $z=0$. This gives $\delta\Omega=2(1+\nu)S$, which is independent of the depth of the center of dilatation. The result for an extended inclusion is therefore ([28])

$$\Delta\Omega=\frac{2}{3}(1+\nu)\int 3\varepsilon_0(\mathbf{r})d^3\mathbf{r}. \tag{36}$$

The integral without the prefactor is the change in volume if the inclusion could expand freely, and it is known that a bounded body containing an inclusion expands by the same amount ([30,31]). The additional factor of $2(1+\nu)/3$ applies to the surface of a semi-infinite body and shows that this change in volume is smaller except in the case of an incompressible medium with $\nu=1/2$.

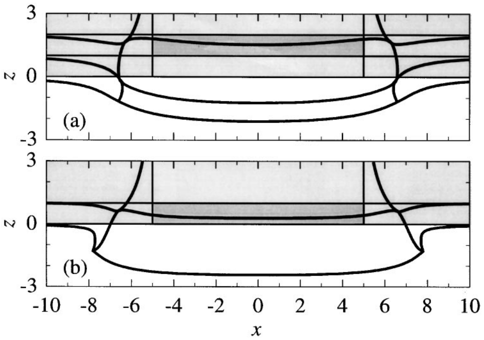


Fig. 2 Displacement of a semi-infinite region due to rectangular wires with $\varepsilon_0=1$, $\nu=1/3$, width 10 units, and thickness 1 unit. Wire (*a*) is buried to a depth of 1 unit while wire (*b*) meets the surface. Light gray shows the region $z>0$ and dark gray shows the wire before the surroundings are strained. Thick lines show the displacement of the surface and of planes that define the edge of the wire with thin lines for their original positions.

4 Examples

4.1 Rectangular Inclusion. Consider first a rectangular "wire" of constant misfit ε_0, parallel to the surface, of infinite extent along y, with cross section $L<x<R$, $B<z<T$ where $B\geqslant 0$. Particular geometries are shown in Fig. 2. Its elastic field in an infinite region was one of the first examples solved using the thermoelastic potential ([25,26]). The elastic quantities can be written as sums of the form

$$\varphi(x,z)=f_\varphi(x-L,z-B)-f_\varphi(x-L,z-T)-f_\varphi(x-R,z-B)+f_\varphi(x-R,z-T), \tag{37}$$

where for the potential

$$f_\varphi(x,z)=-\frac{1+\nu}{1-\nu}\varepsilon_0\left[xz\log(x^2+z^2)-3xz+x^2\arctan\frac{z}{x}+z^2\arctan\frac{x}{z}\right]. \tag{38}$$

The arctangents are principal values, which leads to singularities at the edges of the wire. Sums of the same form as Eq. (37) over the following functions give $u_x^{(\infty)}$, $\varepsilon_{xx}^{(\infty)}$, and $\varepsilon_{xz}^{(\infty)}$:

$$f_x(x,z)=\frac{1+\nu}{1-\nu}\frac{\varepsilon_0}{4\pi}\left[z\ln(x^2+z^2)+2x\arctan\frac{z}{x}\right], \tag{39}$$

$$f_{xx}(x,z)=\frac{1+\nu}{1-\nu}\frac{\varepsilon_0}{2\pi}\arctan\frac{z}{x}, \tag{40}$$

$$f_{xz}(x,z)=\frac{1+\nu}{1-\nu}\frac{\varepsilon_0}{4\pi}\ln(x^2+z^2). \tag{41}$$

There are similar expressions for $u_z^{(\infty)}$ and $\varepsilon_{zz}^{(\infty)}$, and the tensile strains can be visualized in terms of angles subtended by the edges of the wire ([25,26]).

The elastic field for the wire in a semi-infinite medium can now be deduced using the results in Sec. 2. For example, the normal displacement of the free surface is given by Eq. (8) as

$$-u_z(x,0)=\frac{(1+\nu)\varepsilon_0}{\pi}\left[(x-L)\ln\frac{(x-L)^2+T^2}{(x-L)^2+B^2}-(x-R)\ln\frac{(x-R)^2+T^2}{(x-R)^2+B^2}+2T\left(\arctan\frac{x-L}{T}-\arctan\frac{x-R}{T}\right)-2B\left(\arctan\frac{x-L}{B}-\arctan\frac{x-R}{B}\right)\right]. \tag{42}$$

There is no complication if the inclusion reaches the surface and $B=0$. It is possible to measure this displacement with a scanning probe such as an atomic force microscope for submicron structures ([10,19]) or a stylus for larger structures, and this approach has been used to characterize waveguides induced by irradiation of silica ([32]). The displacement of the surface near the middle of a wide, thin, shallow wire with $(R-L)\gg T$ reduces to

$$-u_z=2(1+\nu)\varepsilon_0(T-B). \tag{43}$$

A more physical derivation of this will be given shortly.

Figure 2 shows the displacement around two grossly misfitting wires. There is severe distortion around the corners of the wire, where the shear strain diverges logarithmically (Eq. (41)). A surprising result is that the deeper edge of the wire is displaced towards the surface, which is particularly clear in the wire at the surface shown in Fig. 2(b). This increases the displacement of the surface and can be understood as follows for a wide, thin, shallow wire ([28]).

Consider first a wire in an infinite region. The material on either side constrains the wire so that $\varepsilon_{xx}=\varepsilon_{yy}=0$, while its thin shape allows relaxation so that $\sigma_{zz}=0$ away from the narrow edges. This gives $\varepsilon_{zz}=[(1+\nu)/(1-\nu)]\varepsilon_0$, which pushes out the wide faces of the wire by

$$u_z^{(\infty)}=\pm\frac{1}{2}\frac{1+\nu}{1-\nu}\varepsilon_0(T-B). \tag{44}$$

Now suppose that there is a free surface close to the wire as in Figs. 2(a) or (b). A thin layer between the wire and the free surface will be unstrained and have no effect on the argument. The general result in Eq. (8) for the displacement of the surface shows that the relaxation given by Eq. (44) is enhanced by a factor of $4(1-\nu)$. The outcome is consistent both with the result already derived, Eq. (43), and with the change in volume given by Eq. (36). However, the obvious approach is to argue that the displacement in Eq. (44) will be directed entirely toward the free surface, instead of equally on either side. The surface is therefore pushed outward by

$$-u_z^{(1)}=\frac{1+\nu}{1-\nu}\varepsilon_0(T-B). \tag{45}$$

This is less than the correct result in Eq. (43) because the "amplification factor" $4(1-\nu)>2$.

The extra displacement arises from the effect of the inclusion on its surroundings ([28]). The constraint $\varepsilon_{xx}=\varepsilon_{yy}=0$ induces a compressive stress within the wire, $\sigma_{xx}=-E\varepsilon_0/(1-\nu)$. This in turn gives rise to an outward force per unit length of $\bar{F}=-\sigma_{xx}(T-B)$ at each end. For a shallow wire these can be treated as tangential (Flamant) line forces on the surface, which causes the surface between them to swell outward by ([33])

$$-u_z^{(2)}=\frac{(1-2\nu)(1+\nu)\bar{F}}{E}=\frac{(1-2\nu)(1+\nu)}{1-\nu}\varepsilon_0(T-B). \tag{46}$$

This pulls the deeper edge of the wire toward the surface, as in Fig. 2, and brings the total expansion $u_z^{(1)}+u_z^{(2)}$ into agreement with Eq. (43). The line forces also provide a good approximation to the elastic field outside a thin, shallow wire, particularly at large distances.

4.2 Surface Cleaved Through Infinite Slab. A useful limit is obtained by setting $B=0$ and $T\rightarrow\infty$. This describes a sample with a uniform misfitting slab such as a quantum well as shown in Fig. 1(c), which has been cleaved perpendicular to the layer. This is a standard method of preparing specimens for scanning probe microscopy. For convenience set $L=-a$ and $R=+a$ for a slab of width $2a$ centered on the origin. A difficulty arises in taking the limit of Eq. (42) for the normal displacement of the surface because a term diverges as ln T; it does not depend on x and will be written as a constant $-C$. The displacement is then

$$-u_z(x,0)=C-\frac{2(1+\nu)\varepsilon_0}{\pi}\left[(x+a)\ln\left|\frac{x+a}{a}\right|-(x-a)\ln\left|\frac{x-a}{a}\right|\right]. \tag{47}$$

The surface is unstrained (but distorted!) outside the misfitting layer. Within the layer, $|x|<a$, the strains are constant and given by

$$\varepsilon_{xx}(x,0)=2(1+\nu)\varepsilon_0, \tag{48}$$

$$\varepsilon_{zz}(x,0)=\frac{(1+\nu)(1-2\nu)}{1-\nu}\varepsilon_0, \tag{49}$$

$$\varepsilon(x,0)=\frac{(1+\nu)(3-4\nu)}{1-\nu}\varepsilon_0. \tag{50}$$

These results can also be obtained directly by the strain suppression method because of the simple geometry ([34]).

4.3 Other Geometries. It is often the case that the inclusion had a highly symmetric shape when it was embedded in an infinite region, but this symmetry is destroyed when the sample is reduced to a semi-infinite region. An example is shown in Fig. 1(b), where an ellipsoidal inclusion has been cut on a plane through its polar axis. The elastic field around an ellipsoid in an infinite region is well known ([35]), but that of a half ellipsoid is much more complicated. However, the displacement and strain in the plane of the free surface can be deduced from the results from the symmetric inclusion in an infinite region provided that the surface passes through a mirror plane. The displacement within this plane is simply half that of the symmetric body and Eq. (8) is replaced by

$$u_x(x,y,0)=2(1-\nu)u_x^{(\mathrm{sym})}(x,y,0), \tag{51}$$

where $\mathbf{u}^{(\mathrm{sym})}$ is the displacement due to the full, symmetric inclusion in an infinite region. There are similar relations for u_y, ε_{xx}, ε_{yy}, and ε_{xy} at the surface, while ε_{zz} can be found from the thermoelastic stress-strain relations. Unfortunately the normal distortion u_z, which is the easiest to measure, cannot be obtained this way.

These results can be used to check the calculations for a semi-infinite slab in Sec. 4.2. Take x as the normal to the slab in an infinite region, in which case $\varepsilon_{yy}^{(\infty)}=\varepsilon_{zz}^{(\infty)}=0$ and $\sigma_{xx}^{(\infty)}=0$. This leads to $\varepsilon_{xx}^{(\infty)}=[(1+\nu)/(1-\nu)]\varepsilon_0$ within the slab. There is no stress or strain outside. It follows from Eq. (51) that the strain on the surface of the semi-infinite body should be $\varepsilon_{xx}=2(1+\nu)\varepsilon_0$ within the slab and zero elsewhere, in agreement with Eq. (48).

This approach can also be applied to cross-sectional scanning tunneling microscopy of quantum dots ([19]) by modeling them as oblate ellipsoids of revolution about the x axis. The dots were inclusions of InAs in GaAs, about 4 nm high and 26 nm across, with a mismatch $\varepsilon_0=7\%$. The strains within an ellipsoid are constant ([35]) and these values lead to $\varepsilon_{xx}^{(\infty)}=11\%$ within the dot. According to Eq. (51) this rises to 16% on an exposed surface, taking $\nu=0.3$. This strain is measured with respect to the surroundings (GaAs); subtracting ε_0 gives a strain of 9% with respect to the dot itself (InAs). The measurements ([19]) gave a lower value of around 6%. Many factors could contribute to the discrepancy: the shape of the dot may not be ellipsoidal and the cleavage plane may not be through its center; the composition may not be pure InAs; the elastic properties have cubic symmetry rather than being isotropic; and the elastic response may be nonlinear at such high strain.

The elastic field due to a misfitting rectangular parallelepiped in a semi-infinite solid can be solved in a similar way to the rectangle, starting from the effect in an infinite region ([2,36]); results will not be given because this problem has been solved directly ([2,3,8,21]). An ellipsoid in a semi-infinite solid ([37]) could be treated in a similar way using the field of a free ellipsoid ([35]). Quantum dots are often assumed to be pyramidal and the elastic field around such a dot near a surface ([38]) could be deduced from the results for a pyramid in an infinite region ([17,39]). The strain on the surface is of particular importance because it encourages successive dots to grow in correlated stacks ([40]).

5 Conclusions

It has been shown that the elastic field due to a coherently misfitting inclusion or nonuniform temperature in a semi-infinite region can be derived in a straightforward way from the corresponding field in an infinite region. This follows from Mindlin and Cheng's approach ([23]) but it is *not* necessary to calculate the thermoelastic potential itself, which is usually more difficult than the strain or displacement. The displacement of the free surface is enhanced by a factor of $4(1-\nu)$ over that of the equivalent plane in an infinite region. Part of this enhancement arises because the inclusion is less contained in its outward expansion but the remaining effect is more subtle and is due to distortion of the surroundings, which pushes the inclusion toward the surface ([28]). Results for the center of dilatation agree with previous work and yield a useful relation for the total change in volume due to distortion of the surface by the inclusion. A rectangular inclusion was described in detail because it has many applications, including the limit of a semi-infinite slab.

References

[1] Boley, B. A., and Weiner, J. H., 1997, *Theory of Thermal Stresses*, Dover, Mineola, NY.
[2] Nowacki, W., 1986, *Thermoelasticity*, Pergamon, Oxford, UK, 2nd edition.
[3] Hu, S. M. 1989, "Stress From a Parallelepipedic Thermal Inclusion in a Half-space," J. Appl. Phys., **66**, pp. 2741–2743.
[4] Hu, S. M. 1990, "Stress From Isolation Trenches in Silicon Substrates," J. Appl. Phys., **67**, pp. 1092–1101.
[5] Freund, L. B. 2000, "The Mechanics of Electronic Materials," Int. J. Solids Struct., **37**, pp. 185–196.
[6] Jain, S. C., Maes, H. E., Pinardi, K., and De Wolf, I. 1996, "Stresses and Strains in Lattice-Mismatched Stripes, Quantum Wires, Quantum Dots, and Substrates in Si Technology," J. Appl. Phys., **79**, pp. 8145–8165.
[7] Glas, F. 1987, "Elastic State of the Thermodynamic Properties of Inhomogeneous Epitaxial Layers: Application to Immiscible III-V Alloys," J. Appl. Phys., **62**, pp. 3201–3208.
[8] Glas, F. 1991, "Coherent Stress Relaxation in a Half Space: Modulated Layers, Inclusions, Steps, and a General Solution," J. Appl. Phys., **70**, pp. 3556–3571.
[9] Pinnington, T., Sanderson, A., Tiedje, T., Pearsall, T. P., Kasper, E., and Presting, H. 1992, "Ambient Pressure Scanning Tunneling Microscope Imaging of Hydrogen-Passivated Si/Ge Multilayers," Thin Solid Films, **222**, pp. 259–264.
[10] Chen, H., Feenstra, R. M., Piva, P. G., Goldberg, R. D., Mitchell, I. V., Aers, G. C., Poole, P. J., and Charbonneau, S. 1999, "Enhanced Group-V Intermixing in InGaAs/InP Quantum Wells Studied by Cross-Sectional Scanning Tunneling Microscopy," Appl. Phys. Lett., **75**, pp. 79–81.
[11] Gosling, T. J., and Willis, J. R. 1995, "Mechanical Stability and Electronic Properties of Buried Strained Quantum Well Arrays," J. Appl. Phys., **77**, pp. 5601–5610.
[12] Faux, D. A., Downes, J. R., and O'Reilly, E. P. 1996, "A Simple Method for Calculating Strain Distributions in Quantum-Wire Structures," J. Appl. Phys., **80**, pp. 2515–2517.
[13] Faux, D. A., Downes, J. R., and O'Reilly, E. P. 1997, "Analytic Solutions for Strain Distributions in Quantum-Wire Structures," J. Appl. Phys., **82**, pp. 3754–3762.

[14] Grundmann, M., Stier, O., and Bimberg, D. 1995, ''InAs/GaAs Pyramidal Quantum Dots: Strain Distribution, Optical Phonons, and Electronic Structure,'' Phys. Rev. B, **52**, pp. 11969–11981.

[15] Downes, J. R., Faux, D. A., and O'Reilly, E. P. 1997, ''A Simple Method for Calculating Strain Distributions in Quantum Dot Structures,'' J. Appl. Phys., **81**, pp. 6700–6702.

[16] Pryor, C., Kim, J., Wang, L. W., Williamson, A., and Zunger, A. 1998, ''Comparison of Two Methods for Describing the Strain Profiles in Quantum Dots,'' J. Appl. Phys., **83**, pp. 2548–2554.

[17] Davies, J. H. 1998, ''Elastic and Piezoelectric Fields Around a Buried Quantum Dot: A Simple Picture,'' J. Appl. Phys., **84**, pp. 1358–1365.

[18] Davies, J. H. 1999, ''Quantum Dots Induced by Strain,'' Appl. Phys. Lett., **75**, pp. 4142–4144.

[19] Legrand, B., Grandidier, B., Nuys, J. P., Stiévenard, D., Gérard, J. M., and Thierry-Mieg, V. 1998, ''Scanning Tunneling Microscopy and Scanning Tunneling Spectroscopy of Self-Assembled InAs Quantum Dots,'' Appl. Phys. Lett., **73**, pp. 96–98.

[20] Chiu, Y. P. 1977, ''On the Stress Field Due to Initial Strains in a Cuboid Surrounded by an Infinite Elastic Space,'' ASME J. Appl. Mech., **44**, pp. 587–590.

[21] Chiu, Y. P. 1978, ''On the Stress Field in Surface Deformation in a Half Space With a Cuboidal Zone in Which Initial Strains are Uniform,'' ASME J. Appl. Mech., **45**, pp. 302–306.

[22] Mindlin, R. D., and Cheng, D. H. 1950, ''Nuclei of Strain in the Semi-Infinite Solid,'' J. Appl. Phys., **21**, pp. 926–930.

[23] Mindlin, R. D., and Cheng, D. H. 1950, ''Thermoelastic Stress in the Semi-Infinite Solid,'' J. Appl. Phys., **21**, pp. 931–933.

[24] Sen, B. 1951, ''Note on the Stresses Produced by a Nuclei of Thermoelastic Strain in a Semi-Infinite Elastic Solid,'' Q. Appl. Math., **8**, pp. 365–369.

[25] Goodier, J. N. 1937, ''On the Integration of the Thermo-Elastic Equations,'' Philos. Mag., **23**, pp. 1017–1032.

[26] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York, 3rd edition.

[27] Westergaard, H. M., 1952, *Theory of Elasticity and Plasticity*, Harvard University Press, Cambridge, MA.

[28] Barber, J. R. 1987, ''Thermoelastic Distortion of the Half-Space,'' J. Therm. Stresses, **10**, pp. 221–228.

[29] Barber, J. R., 2002, *Elasticity*, 2nd Ed., Kluwer, Dordrecht, The Netherlands.

[30] Nowacki, W., 1954, ''Thermal Stresses in Anisotropic Bodies (I),'' Arch. Mech. Stos. (Arch. Mech.) **6**, pp. 481–492.

[31] Hieke, M. 1955, Z. Angew. Math. Mech., **35**, pp. 285–294.

[32] Garcia Blanco, S., Glidle, A., Davies, J. H., Aitchison, J. S., and Cooper, J. M. 2001, ''Electron Beam Induced Densification of Ge-Doped Flame Hydrolysis Silica for Waveguide Fabrication,'' Appl. Phys. Lett., **79**, pp. 2889–2891.

[33] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

[34] Davies, J. H., Bruls, D. M., Vugs, J. W. A. M., and Koenraad, P. M. 2002, ''Relaxation of a Strained Quantum Well at a Cleaved Surface,'' J. Appl. Phys., **91**, pp. 4171–4176.

[35] Myklestad, N. O. 1942, ''Two Problems of Thermal Stress in the Infinite Solid,'' ASME J. Appl. Mech., **9**, pp. A136–A143.

[36] Faivre, G. 1964, ''Déformations de Cohérence d'un Précipité Quadratique,'' Phys. Status Solidi, **35**, pp. 249–259.

[37] Seo, K., and Mura, T. 1979, ''The Elastic Field in a Half Space due to Ellipsoidal Inclusions With Uniform Dilational Eigenstrains,'' ASME J. Appl. Mech., **46**, pp. 568–572.

[38] Glas, F. 2001, ''Elastic Relaxation of Truncated Pyramidal Quantum Dots and Quantum Wires in a Half Space: An Analytical Calculation,'' J. Appl. Phys., **90**, pp. 3232–3241.

[39] Pearson, G. S., and Faux, D. A. 2000, ''Analytical Solutions for Strain in Pyramidal Quantum Dots,'' J. Appl. Phys., **88**, pp. 730–736.

[40] Lita, B., Goldman, R. S., Phillips, J. D., and Battacharya, P. K. 1999, ''Interdiffusion and Surface Segregation and Stacked, Self-Assembled InAs/GaAs Quantum Dots,'' Appl. Phys. Lett., **75**, pp. 2797–2799.

Boundary Integral Equation Formulation in Generalized Linear Thermo-Viscoelasticity With Rheological Volume

A. S. El-Karamany
Faculty of Education,
Department of Mathematics,
P.O. Box 272,
Rustaq 329, Sultanate of Oman,
e-mail: aelkaramani@yahoo.com

A general model of generalized linear thermo-viscoelasticity for isotropic material is established taking into consideration the rheological properties of the volume. The given model is applicable to three generalized theories of thermoelasticity: the generalized theory with one (Lord-Shulman theory) or with two relaxation times (Green-Lindsay theory) and with dual phase-lag (Chandrasekharaiah-Tzou theory) as well as to the dynamic coupled theory. The cases of thermo-viscoelasticity of Kelvin-Voigt model or thermoviscoelasticity ignoring the rheological properties of the volume can be obtained from the given model. The equations of the corresponding thermoelasticity theories result from the given model as special cases. A formulation of the boundary integral equation (BIE) method, fundamental solutions of the corresponding differential equations are obtained and an example illustrating the BIE formulation is given.
[DOI: 10.1115/1.1607354]

Introduction

The linear viscoelasticity remains an important area of research not only due the advent and use of polymers, but also because most solids when subjected to dynamic loading exhibit viscous effects (see, e.g., Fredudenthal [1]). The stress-strain law for many materials such as polycrystalline metals and high polymers can be approximated by the linear viscoelasticity theory, [2]. Many works were devoted to the viscoelasticity and thermoviscoelasticity theories, e.g., Bland [3], Gurtin and Sternberg [4], Christensen [5], and Ilioushin and Pobedria [6]. Results of important experiments determining the mechanical properties of viscoelastic materials are included in Koltunov's work, [7]. The association and exploitation of integral equations with viscoelasticity have been given by Rabotnov [8] and Gurtin and Sternberg [4]. Most investigations in thermal viscoelasticity are ignoring the relaxation effects of the volume, although the rheological behavior of the volume in polymers is confirmed experimentally in the work by Kovacs [9].

Biot [10] formulated the theory of coupled thermoelasticity to eliminate the paradox inherent in the classical uncoupled theory that elastic changes have no effect on the temperature. The heat equations for both theories of the diffusion type predicting infinite speeds of propagation for heat waves contrary to physical observations. Five generalizations to the coupled theory were introduced. The first is due to Lord and Shulman [11] who introduced the theory of generalized thermoelasticity with one relaxation time by postulating a new law of heat conduction to replace the classical Fourier's law. This law contains the heat flux vector as well as its time derivative. It contains also a new constant that acts as relaxation time. The heat equation of this theory is of the wave type, ensuring finite speeds of propagation for heat and elastic waves. The remaining governing equations for this theory, namely, the equations of motion and the constitutive relations remain the same as those for the coupled and the uncoupled theories. The second generalization to the coupled theory of thermoelasticity is known as the generalized theory with two relaxation times. Muller [12] proposed an entropy production inequality, with the help of which he considered restrictions on a class of constitutive equations. A generalized of this inequality was given by Green and Laws [13], Green and Lindsay [14], and Suhubi [15]. The third generalization to the coupled theory is known as the dual-phase-lag thermoelasticity, proposed by Chandrasekharaiah and Tzou [16,17] (C-T theory), in which the Fourier law is replaced by an approximation to a modification of the Fourier law with two different translations for the heat flux and the temperature gradient. One can refer to Ignaczak [18] for a review, presentation of the two generalizations L-S and G-L theories and some important results obtained under these two theories of thermoelasticity, and to Hetnarski and Ignaczak [19] for a review and presentation of five generalized theories of thermoelasticity.

The boundary integral equation method (BIEM) has been applied successfully to many branches of physics, applied mathematics and engineering sciences, due to its efficiency and ease of implementation compared with the other numerical methods. Applications in thermal stress problems were given by Cruse and Rizzo [20], Rizzo and Shippy [21], Banerjee and Butterfield [22], Brebbia et al. [23], and Ziegler and Irschik [24]. Sladek and Sladek [25] set up the BIE formulation for the coupled thermoelasticity. A treatment of scalar and vector potential theory directed towards the BIE formulation and description of numerical dealing with these equations, are given by Jaswon and Symm [26].

The present work consists of the introduction, general mathematical model, the formulation of the problem in Laplace transform domain; the fundamental solutions in Laplace transform domain and the BIE formulation for the given model. An initial-mixed boundary value problem is considered as an example illustrating the BIE formulation.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 19, 2003; final revision, Feb. 26, 2003. Associate Editor: D. A. Siginer. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

The Mathematical Model

Assuming a linear thermoviscoelastic material occupies a regular region D with a smooth boundary surface B in the three-dimensional Euclidian space. The material is assumed to be isotropic. Through this paper a rectangular coordinate system (x_1, x_2, x_3) is employed. $\underline{x}$ is the position vector and t the time. All the functions are, considered to be functions of $(\overline{x},t)$, defined on $\overline{D}(=D\cup B)\times[0,\infty)$. A superposed dot denotes differentiation with respect to time, while a comma denotes partial differentiation with respect to the space variables x_i. The summation notation is used. The nonrheological properties of the material are assumed temperature-independent and time-independent, therefore the constitutive equations are of convolution type, [8].

The system of governing equations for the linear thermoviscoelastic solid consists of the following:

The equations of motion on $D\times(0,\infty)$:

$$\sigma_{ji,j}+\rho F_i=\rho\ddot{u}_i. \tag{1}$$

The constitutive equations on $D\times[0,\infty)$, [4,6]:

$$S_{ij}=\check{R}_G(e_{ij}),\quad \sigma=\check{R}_K(e-3\alpha_T\hat{T}), \tag{2}$$

where

$$S_{ij}=\sigma_{ij}-\sigma\delta_{ij},\quad e_{ij}=\varepsilon_{ij}-\frac{e}{3}\delta_{ij},\quad \sigma=\frac{\sigma_{kk}}{3},$$

$$\varepsilon_{ij}=\frac{1}{2}(u_{i,j}+u_{j,i}). \tag{3}$$

Substituting from Eqs. (3) into Eqs. (2) we obtain

$$\sigma_{ij}=\sigma\delta_{ij}+\check{R}_G\left(\varepsilon_{ij}-\frac{e}{3}\delta_{ij}\right). \tag{4}$$

Equation (1) together with Eq. (4) reduces to

$$\rho(\ddot{u}_i-F_i)=\check{R}_G\left(\frac{\nabla^2u_i}{2}+\frac{e_{,i}}{6}\right)+\check{R}_K(e_{,i}-3\alpha_T\hat{T}_{,i}). \tag{5}$$

The generalized heat conduction equation on $D\times(0,\infty)$:

$$k\left(1+t_1\frac{\partial}{\partial t}\right)T_{,ii}=\rho C_E(\dot{T}+\tau_0\ddot{T}+t_2^2\dddot{T})+3T_0\alpha_T\check{R}_K \times(\dot{e}+n_0\tau_0\ddot{e}+t_2^2\dddot{e})-(Q+n_0\tau_0\dot{Q}+t_2^2\ddot{Q}). \tag{6}$$

The operator $\check{R}_a(f)$, $(a=G,K)$ is defined for any function $f(\overline{x},t)$, of class C^1, as

$$\check{R}_a(f)=\check{R}_a(f(\overline{x},t))=\int_0^t R_a(t-\tau)\frac{\partial f(\overline{x},\tau)}{\partial\tau}d\tau,\quad (a=G,K) \tag{7}$$

where $R_G(t)$ and $R_K(t)$ are two relaxation functions (shear and bulk viscoelasticity moduli) The present formulation are valid for any, positive monotonic decreasing, relaxation functions, satisfying the conditions, [5],

$$R_G(t)>0,\quad R_K(t)>0,\quad 3R_K(t)-R_G(t)>0 \tag{8}$$

and the nonretroactivity conditions, [2,27],

$$R_a(t)=0\,\forall t\in(-\infty,0),\quad (a=G,K). \tag{9}$$

Equations (5) and (6) are the *field equations*, (on $D\times(0,\infty)$), of the generalized linear thermo-viscoelasticity, valid for infinitesimal temperature deviations from the reference temperature T_0, [5,28], applicable to the coupled theory, three generalizations, and to several special cases as follows:

1. the equations of the coupled linear thermoviscoelasticity, when
$$t_1=t_2=\tau_0=\nu=0. \tag{10}$$
2. the equations of the generalized linear thermoviscoelasticity with one relaxation time (L-S theory), when
$$n_0=1,\quad t_1=t_2=\nu=0,\quad \tau_0>0 \tag{11}$$
where τ_0 is relaxation time.
3. the equations of the generalized linear thermoviscoelasticity with two relaxation times (G-L theory), when
$$n_0=0,\quad t_1=t_2=0,\quad \nu\geqslant\tau_0>0, \tag{12}$$
where ν and τ_0 are two relaxation times.
4. the equations of the generalized linear thermoviscoelasticity with dual-phase-lag (C-T theory), when
$$n_0=1,\quad t_1=\tau_\theta>0,\quad \tau_0=\tau_q>0,\quad t_2^2=\frac{1}{2}\tau_q^2,\quad \nu=0,$$
$$\tau_q\geqslant\tau_\theta>0. \tag{13}$$
5. the equations of the generalized linear thermoviscoelasticity of Kelvin-Voigt model, [29], can be obtained from the above Eqs. (2) and (4)–(6) replacing the operators $\check{R}_G(f)$ and $\check{R}_K(f)$ by the operators
$$R_G^{(v)}(f(\overline{x},t))=2\mu\left(1+\lambda_v\frac{\partial}{\partial t}\right)f(\overline{x},t)\quad\text{and}$$
$$R_K^{(v)}(f(\overline{x},t))=K\left(1+\lambda_v\frac{\partial}{\partial t}\right)f(\overline{x},t)$$
respectively, where $\lambda_v>0$ is *the retardation period of Kelvin-Voigt model*, [29].
6. the equations of the generalized linear thermoelasticity can be obtained from Eqs. (2) and (4)–(6) replacing the operators $\check{R}_G(f)$ and $\check{R}_K(f)$ by $2\mu f(\overline{x},t)$ and $Kf(\overline{x},t)$, respectively.
7. the corresponding equations of the generalized linear thermoviscoelasticity ignoring the rheological volume properties (the bulk viscoelasticity) can be obtained from Eqs. (2) and (4)–(6) replacing the operator $\check{R}_K(f)$ by $Kf(\overline{x},t)$.

Let us introduce the following nondimensional variables:

$$x_i^*=C_0\eta_0x_i;\quad u_i^*=C_0\eta_0u_i;\quad t^*=C_0^2\eta_0t;$$

$$\tau_0^*=C_0^2\eta_0\tau_0;\quad \nu^*=C_0^2\eta_0\nu$$

$$t_1^*=C_0^2\eta_0t_1;\quad t_2^*=C_0^2\eta_0t_2;\quad \tau_\theta^*=C_0^2\eta_0\tau_\theta;$$

$$\tau_q^*=C_0^2\eta_0\tau_q;\quad F_i^*=\frac{F_i}{C_0^3\eta_0};\quad Q^*=\frac{\gamma Q}{k\rho C_0^4\eta_0^2};$$

$$\theta=\frac{\gamma(T-T_0)}{\rho C_0^2};\quad T_0=\frac{\rho C_0^2\delta_0}{\gamma}=\frac{\delta_0}{3\alpha_T};\quad \hat{\Theta}=\Theta+\nu\dot{\Theta};$$

$$\sigma_{ij}^*=\frac{\sigma_{ij}}{K};\quad R_G^*=\frac{2}{3K}R_G;\quad R_K^*=\frac{R_K}{K}.$$

The nondimensional positive constant δ_0 is introduced to specify the value of T_0 considered. In terms of these nondimensional variables, Eqs. (2), take the form (dropping the asterisks)

$$\sigma=\breve{R}_K(e-\hat{\Theta}),\quad \sigma_{ij}=\sigma\delta_{ij}+\frac{3}{2}\breve{R}_G\left(\varepsilon_{ij}-\frac{e}{3}\delta_{ij}\right),\quad \sigma_{ji,j}=\ddot{u}_i-F. \tag{14}$$

The field Eqs. (5) and (6) take the form

$$\ddot{u}_i-F_i=\frac{1}{4}\breve{R}_G(3\nabla^2u_i+e_{,i})+\breve{R}_K(e_{,i}-\hat{\Theta}_{,i}) \tag{15}$$

$$\left(1+t_1\frac{\partial}{\partial t}\right)\Theta_{,ii}=(\dot{\Theta}+\tau_0\ddot{\Theta}+t_2^2\dddot{\Theta})+\varepsilon\breve{R}_K(\dot{e}+n_0\tau_0\ddot{e}+t_2^2\dddot{e})$$
$$-(Q+n_0\tau_0\dot{Q}+t_2^2\ddot{Q}). \tag{16}$$

The system of Eqs. (5) and (6) is completed by the initial and boundary conditions.

The initial conditions will be assumed homogeneous:

$$u_i(\overline{x},t)=0,\ \ \Theta(\overline{x},t)=0,\quad \frac{\partial^n u_i(\overline{x},t)}{\partial t^n}=0,\quad \frac{\partial^n\Theta(\overline{x},t)}{\partial t^n}=0,$$
$$\overline{x}\in\overline{D},\ \ t\leqslant 0,\ \ (n\geqslant 1). \tag{17}$$

The boundary conditions:

$$\sigma_{ji}n_j=f_i(\overline{x_\sigma},t)\ \text{ on }\ B_\sigma\times(0,\infty);\quad u_i=g_i(\overline{x_{B_u}},t)\ \text{ on }\ B_u\times(0,\infty) \tag{18}$$

$$\Theta=\Phi(\overline{x_{B_1}},t)\ \text{ on }\ B_1\times(0,\infty);\quad \Theta_{,n}=\Theta_{,i}n_i=G(\overline{x_{B_2}},t)\ \text{ on }$$
$$B_2\times(0,\infty) \tag{19}$$

where the functions f_i, g_i, Φ and G are given functions, equal to zero when $t\leqslant 0$ (B_u,B_σ) and (B_1,B_2) are two partitions of the boundary surface B such that $B=B_u\cup B_\sigma=B_1\cup B_2$, $B_u\cap B_\sigma=B_1\cap B_2=\phi$, and $n_i=n_i(\overline{x_B})$ are the components of the outer normal vector to the surface at $\overline{x_B}$.

The formulation of the Problem in the Laplace Transform Domain

Performing the Laplace transform, [30], over Eqs. (14), (15), and (16), taking into consideration the homogeneous initial conditions (17), and omitting the bars, we get

$$\sigma=sR_K(e-\omega_2\Theta),\quad \sigma_{ij}=\sigma\delta_{ij}+\frac{3}{2}sR_G\left(\varepsilon_{ij}-\frac{e}{3}\delta_{ij}\right),$$
$$\sigma_{ji,j}=s^2u_i-F \tag{20}$$

$$s^2u_i-F_i=\frac{3sR_G}{4}\nabla^2u_i+\frac{s(R_G+4R_K)}{4}e_{,i}-sR_K\omega_2\Theta_{,i}, \tag{21}$$

$$\omega_3\Theta_{,ii}=s\omega_1\Theta+\varepsilon R_Ks^2\omega e-\omega Q. \tag{22}$$

$$\begin{aligned}&\sigma_{ij}n_j=f_i(\overline{x_B},s),\ \ \overline{x_B}\in B_\sigma;\quad u_i=g_i(\overline{x_B},s),\ \ \overline{x_B}\in B_u;\\ &\Theta=\Phi(\overline{x_B},s),\ \ \overline{x_B}\in B_1;\quad \Theta_{,n}=G(\overline{x_B},s),\ \ \overline{x_B}\in B_2\end{aligned} \tag{23}$$

where

$$\omega=1+n_0\tau_0s+t_2^2s^2,\quad \omega_1=1+\tau_0s+t_2^2s^2,\quad \omega_2=1+\nu s,$$
$$\omega_3=1+t_1s. \tag{24}$$

According to the Helmholtz theorem, [31] the displacement and the body forces can be expressed in the form

$$u_i=\Omega_{,i}+\epsilon_{ijk}\Psi_{k,j},\quad \Psi_{i,i}=0;\quad F_i=X_{,i}+\epsilon_{ijk}Y_{k,j},\quad Y_{i,i}=0 \tag{25}$$

where Ω, X are the scalar potentials and Ψ_k, Y_k the vector potentials of the vector fields u_i, F_i, respectively. Substituting Eqs. (25) into Eqs. (21) and (22), we get

$$(\nabla^2-P_1^2)\Omega-m\Theta=-\frac{X}{C_1^2}, \tag{26}$$

$$(\nabla^2-P_2^2)\Psi_i=-\frac{Y_i}{C_2^2}, \tag{27}$$

$$(\nabla^2-m_1^2)\Theta-\alpha\nabla^2\Omega=-b_0Q, \tag{28}$$

where

$$C_1^2=s(R_G+R_K);\quad C_2^2=\frac{3sR_G}{4},\quad P_n=\frac{s}{C_n},\ \ (n=1,2) \tag{29}$$

$$m=\frac{\omega_0}{C_1^2},\quad m_1^2=\frac{s\omega_1}{\omega_3},\quad \alpha=\varepsilon s^2b_0R_K,\quad \omega_0=sR_K\omega_2,$$
$$b_0=\frac{\omega}{\omega_3}. \tag{30}$$

Fundamental Solutions in the Laplace Transform Domain

We shall consider two cases, [28],

Case I: An instantaneous source of heat located at $x_i=y_i$ where $\overline{y}\in(D\cup B)$, acting upon a viscoelastic body in the absence of body forces, i.e., we assume $Q=\delta(r)\delta(t)$, $F_i=0$ then

$$L\{Q\}=\delta(r),\quad L\{F\}=0, \tag{31}$$

where

$$r=\sqrt{(x_i-y_i)(x_i-y_i)}.$$

Denoting the corresponding fundamental solutions by $u^{(1)}$, $\Theta^{(1)}$, substituting the Laplace transforms from Eqs. (31) in the governing Eqs. (26)–(28), using the Helmholtz equation, [32],

$$\frac{1}{\nabla^2-k^2}[\delta(r)]=-\frac{1}{4\pi r}e^{-kr}, \tag{32}$$

and introducing the notations

$$E_n=(-1)^{n-1}e^{-k_nr},\quad \xi_n=(-1)^{n-1}\left(k_n+\frac{1}{r}\right)e^{-k_nr},$$

$$V_n=3\xi_n+k_n^2rE_n,\quad C=\frac{mb_0}{4\pi(k_1^2-k_1^2)}, \tag{33}$$

we obtain for an infinite region in view of Eqs. (17)

$$\Psi_i^{(1)}(r,s)=0,\quad \Omega^{(1)}(r,s)=\frac{C}{r}\sum_1^2 E_n,$$

$$\Theta^{(1)}=\frac{C}{mr}\sum_1^2(k_n^2-P_1^2)E_n \tag{34}$$

where k_1^2, k_2^2 are the roots of the characteristic equation

$$k^4-(m_1^2+\alpha m+P_1^2)k^2+P_1^2m_1^2=0. \tag{35}$$

From (34) one obtains

$$u_i^{(1)}(\overline{x},\overline{y},s)=-\frac{Cr_{,i}}{r}\sum_1^2 \xi_n, \quad e^{(1)}=\frac{C}{r}\sum_1^2 k_n^2E_n,$$

$$\varepsilon_{ij}=\frac{C}{r^2}\sum_1^2 (r_{,i}r_{,j}V_n-\xi_n\delta_{ij}). \tag{36}$$

The Laplace transform of the traction vector in this case is obtained from Eqs. (20), (34), and (36). The expression for $\Theta_{,n}^{(1)}=\Theta_{,i}^{(1)}n_i$ can be obtained from (34).

Case II: We assume now that $Q=0$, and an instantaneous concentrated body force $F_i=\delta(\overline{x}-\overline{y})\delta(t)\delta_{ij}$ is acting at the point $x_i=y_i$ where $\overline{y}\in(D\cup B)$, in the direction of x_j-axis. Taking the Laplace transform of F_i, omitting the bars, we have

$$Q=0, \quad F_i=F_i^{(j)}=\delta_{ij}\delta(r). \tag{37}$$

Since $\epsilon_{ilk}Y_{k,li}^{(j)}=0$ and $\epsilon_{iqp}X_{,iq}^{(J)}=0$, Eq. (37) with the second equation of Eqs. (25) leads to: $\nabla^2X^{(j)}=(\delta_{ij}\delta(r))_{,i}$, $\nabla^2Y_p^{(j)}=(\epsilon_{iqp}\delta_{ij}\delta(r))_{,q}$ from which using Eq. (32), one obtains

$$X^{(j)}=-\frac{1}{4\pi}\left(\frac{\delta_{ij}}{r}\right)_{,i}; \quad Y_k^{(j)}=\frac{1}{4\pi}\epsilon_{iqk}\left(\frac{\delta_{qj}}{r}\right)_{,i}. \tag{38}$$

Substituting from Eqs. (37) and (38) into the governing Eqs. (26)–(28), using Eq. (32) and the notations $\beta=1/4\pi s^2$, $A_n=(k_n^2-m_1^2)/4\pi C_1^2k_n^2(k_2^2-k_1^2)$, we obtain

$$\Omega^{(j)}=\frac{\beta\delta_{ij}r_{,i}}{r^2}+\frac{\delta_{ij}r_{,i}}{r}\sum_1^2 A_n\xi_n, \tag{39}$$

$$\Psi_k^{(j)}=\epsilon_{ijk}\left(\frac{\beta r_{,i}}{r^2}\right)[(1+P_2r)e^{-P_2r}-1]. \tag{40}$$

Substituting Eqs. (39) and (40) into Eqs. (25) one obtains

$$u_i^{(j)}(\overline{x},\overline{y},s)=\frac{\beta}{r^2}(\delta_{ij}\xi_r-r_{,i}r_{,j}V_r)+\frac{1}{r^2}\sum_1^2 A_n(\delta_{ij}\xi_n-r_{,i}r_{,j}V_n)$$
$$=u_i^{(j)}(\overline{y},\overline{x},s), \tag{41}$$

where

$$\xi_r=\left(rP_2^2+P_2+\frac{1}{r}\right)e^{-P_2r}, \quad V_r=\left(rP_2^2+3P_2+\frac{3}{r}\right)e^{-P_2r}. \tag{42}$$

Solving Eqs. (26) and (28) taking into consideration Eqs. (37), (38) and using Eq. (32) we get

$$\Theta^{(j)}(\overline{x},\overline{y},s)=\left(-\frac{\varepsilon sC}{\omega_2}\right)\frac{r_i\delta_{ij}}{r}\sum_1^2 \xi_n=\frac{\varepsilon s}{\omega_2}u_j^{(1)}(\overline{x},\overline{y},s). \tag{43}$$

From Eq. (41) we get

$$u_{i,k}^{(j)}=\frac{R_{ijk}^*}{r^3}\left(\beta V_r+\sum_1^2 A_nV_n\right)+\frac{r_{,i}r_{,j}r_{,k}}{r}\left(\sum_1^2 A_nk_n^2\xi_n\right)$$
$$+\frac{(r_{,i}r_{,j}r_{,k}-r_{,k}\delta_{ij})}{r}P_r \tag{44}$$

where

$$R_{ijk}^*=5r_{,i}r_{,j}r_{,k}-(r_{,k}\delta_{ij}+r_{,j}\delta_{ki}+r_{,i}\delta_{jk})$$

and

$$P_r=\beta P_2^2\left(P_2+\frac{1}{r}\right)e^{-P_2r}. \tag{45}$$

The dilatation $e^{(j)}=(r_{,i}\delta_{ij}/r)(\Sigma_1^2A_nk_n^2\xi_n)$ results from Eq. (44). The Laplace transform of the traction vector in this case is obtained from Eqs. (20), (43), and (44). The expression for $\Theta_{,n}^{(j)}=\Theta_{,i}^{(j)}n_i$ can be obtained from (43).

Boundary Integral Equations

The dynamic reciprocity theorem for the model given by the system of Eqs. (20)–(22), supplemented with the boundary conditions (23) and homogenous initial conditions, in Laplace transform domain is, [33],

$$\varepsilon s\omega\int_D F_i^{(1)}u_i^{(2)}dV+\omega\omega_2\int_D Q^{(2)}\Theta^{(1)}dV$$
$$+\varepsilon s\omega\left[\int_{B_u}\sigma_{ij}^{(1)}n_jg_i^{(2)}dA+\int_{B_\sigma}f_i^{(1)}u_i^{(2)}dA\right]$$
$$+\omega_2\omega_3\left[\int_{B_1}\Phi^{(1)}\Theta_{,n}^{(2)}dA+\int_{B_2}G^{(2)}\Theta^{(1)}dA\right]=S_{21}^{12} \tag{46}$$

where S_{21}^{12} indicates the same expression as on the left-hand side except that superscripts (1) and (2) are interchanged, [2].

To obtain the integral representation of the transformed temperature and displacement inside a bounded region D in terms of the prescribed Θ, $\Theta_{,n}$, $f_i=\sigma_{ji}n_j$ and u_i on the surface B, the Green functions $u_i^{(1)}$, $\Theta^{(1)}$, $u_i^{(j)}$, $\Theta^{(j)}$, in infinite region and their values $g_i^{(1)}$, $f_i^{(1)}$, $\Phi^{(1)}$, $G^{(1)}$, $g_i^{(j)}$, $f_i^{(j)}$, $\Phi^{(j)}$, and $G^{(j)}$ on the same surface B, we substitute in Eq. (46) in view of Eqs. (23), the instantaneous heat source $Q^{(1)}=\delta(\overline{x}-\overline{y})$, $F_i^{(1)}=0$, and the corresponding solution $u^{(1)}$ and $\Theta^{(1)}$ we get

$$\omega_2\omega\Delta(\overline{x})\Theta(\overline{x},s)=\omega_2\omega\int_D Q\Theta^{(1)}dV-\varepsilon s\omega\int_D F_iu_i^{(1)}dV$$
$$+\omega_2\omega_3\left[\int_{B_1}\Theta_{,n}\Phi^{(1)}dA+\int_{B_2}G\Theta^{(1)}dA\right]$$
$$+\varepsilon s\omega\left[\int_{B_\sigma}f_i^{(1)}u_idA+\int_{B_u}\sigma_{ij}^{(1)}n_jg_idA\right]$$
$$-\left(\sum_{(1)0}^{0(1)}P4T\right) \tag{47}$$

where $(\Sigma_{(1)0}^{0(1)}P4T)$ indicates the same expression as the preceding four terms except that the superscript (1) is written above the other function in every integrand of the surface integrals and

$$\int_D\delta(\overline{x}-\overline{y})dV(\overline{y})=\Delta(\overline{x})=\begin{cases}1 & \overline{x}\in D\\ 0 & \overline{x}\notin(D\cup B)\\ \frac{1}{2} & \overline{x}\in B\end{cases}. \tag{48}$$

Secondly, we take $F_i^{(j)}=\delta_{ij}\delta(\overline{x}-\overline{y})$, $Q^{(j)}=0$, and the corresponding Green functions $u_i^{(j)}$, $\Theta^{(j)}$ we get

$$\varepsilon s\omega\Delta(\overline{x})u_j(\overline{x},s)=-\omega_2\omega\int_D Q\Theta^{(j)}dV+\varepsilon s\omega\int_D F_iu_i^{(j)}dV$$
$$+\omega_2\omega_3\left[\int_{B_1}\Theta_{,n}^{(j)}\Phi dA+\int_{B_2}G^{(j)}\Theta dA\right]$$
$$+\varepsilon s\omega\left[\int_{B_\sigma}f_iu_i^{(j)}dA+\varepsilon s\omega\int_{B_u}\sigma_{ik}n_kg_i^{(j)}dA\right]$$
$$-\sum_{(j)0}^{0(j)}P4T. \tag{49}$$

Taking into consideration Eqs. (10)–(13) we get for all the considered generalized theories $\nu t_1=\nu t_2=\nu n_0t_0=0$, and therefore

$$\omega\omega_2=(1+\nu_2 s+t_2^2 s^2),\quad \omega_2\omega_3=(1+\nu_1 s),\quad \nu_1=(\nu+t_1),$$
$$\nu_2=(\nu+n_0\tau_0). \tag{50}$$

The inversion, [30], of Eq. (47) in view leads to

$$\Delta(\overline{x})L_1(\Theta(\overline{x},t))=W_1(\overline{x},t) \tag{51}$$

where

$$W_1(\overline{x},t)=\int_0^t\int_D Q(\overline{y},t-\tau)L_1(\Theta^{(1)}(\overline{y},\overline{x},\tau))dV(\overline{y})d\tau-\varepsilon\int_0^t\int_D F_i(\overline{y},t-\tau)\frac{\partial L_2(u_i^{(1)}(\overline{y},\overline{x},\tau))}{\partial\tau}dV(\overline{y})d\tau+\varepsilon\int_0^t\int_{B_\sigma}u_i(\overline{y},t-\tau)$$
$$\times\frac{\partial L_2(f_i^{(1)}(\overline{y},\overline{x},\tau))}{\partial\tau}dA(\overline{y})d\tau+\varepsilon\int_0^t\int_{B_u}g_i(\overline{y},t-\tau)\frac{\partial L_2(\sigma_{ji}^{(1)}(\overline{y},\overline{x},\tau))}{\partial\tau}n_j dA(\overline{y})d\tau+\int_0^t\int_{B_1}\Theta_{,n}(\overline{y},t-\tau)L_3(\Phi^{(1)}$$
$$\times(\overline{y},\overline{x},\tau))dA(\overline{y})d\tau+\int_0^t\int_{B_2}G(\overline{y},t-\tau)L_3(\Theta^{(1)}(\overline{y},\overline{x},\tau))dA(\overline{y})d\tau-\sum_{(1)0}^{0(1)}P4T \tag{52}$$

$$L_1(f(\overline{x},t))=\left(1+\nu_2\frac{\partial}{\partial t}+t_2^2\frac{\partial^2}{\partial t^2}\right)f(\overline{x},t),\quad L_2(f(\overline{x},t))=\left(1+n_0\tau_0\frac{\partial}{\partial t}+t_2^2\frac{\partial^2}{\partial t^2}\right)f(\overline{x},t)$$
$$L_3(f(\overline{x},t))=\left(1+\nu_1\frac{\partial}{\partial t}\right)f(\overline{x},t). \tag{53}$$

The inversion of Eq. (49) leads to

$$\Delta(\overline{x})L_2(u_j(\overline{x},t))=W_2(\overline{x},t) \tag{54}$$

where

$$W_2(\overline{x},t)=\varepsilon\int_0^t\int_D F_i(\overline{y},t-\tau)L_2(u_i^{(j)}(\overline{y},\overline{x},\tau))dV(\overline{y})d\tau-\int_0^t\int_D Q(\overline{y},t-\tau)L_1^*(\Theta^{(j)}(\overline{y},\overline{x},\tau))dV(\overline{y})d\tau$$
$$+\int_0^t\int_{B_1}\Phi(\overline{y},t-\tau)L_3^*(\Theta_{,n}^{(j)}(\overline{y},\overline{x},\tau))dA(\overline{y})d\tau+\int_0^t\int_{B_2}\Theta(\overline{y},t-\tau)L_3^*(G^{(j)}(\overline{y},\overline{x},\tau))dA(\overline{y})d\tau$$
$$+\varepsilon\int_0^t\int_{B_\sigma}f_i(\overline{y},t-\tau)L_2(u_i^{(j)}(\overline{y},\overline{x},\tau))dA(\overline{y})d\tau+\varepsilon\int_0^t\int_{B_u}\sigma_{ki}(\overline{y},t-\tau)n_k L_2(g_i^{(j)}(\overline{y},\overline{x},\tau))dA(\overline{y})d\tau-\left(\sum_{(j)0}^{0(j)}P4T\right) \tag{55}$$

$$L_3^*(f)=\int_0^\tau L_3(f(\overline{y},\overline{x},\zeta))d\zeta,\quad L_1^*(f)=\int_0^\tau L_1(f(\overline{y},\overline{x},\zeta))d\zeta. \tag{56}$$

From Eqs. (51) and (54) we obtain $\Theta(\overline{x},t)$ and $u_j(\overline{x},t)$:

i. For the dynamic coupled theory, in view of Eqs. (10), (50), and (53), we get
$$\Delta(\overline{x})\Theta(\overline{x},t)=W_1(\overline{x},t),\quad \Delta(\overline{x})u_j(\overline{x},t)=W_2(\overline{x},t). \tag{57}$$

ii. For L-S theory, in view of Eqs. (11), (50), and (53), we get
$$\Delta(\overline{x})\Theta(\overline{x},t)=\frac{1}{\tau_0}e^{-t/\tau_0}\int_0^t e^{\tau/\tau_0}W_1(\overline{x},\tau)d\tau, \tag{58}$$
$$\Delta(\overline{x})u_j(\overline{x},t)=\frac{1}{\tau_0}e^{-t/\tau_0}\int_0^t e^{\tau/\tau_0}W_2(\overline{x},\tau)d\tau. \tag{59}$$

For G-L theory, in view of Eqs. (12), (50), and (53), we get

$$\Delta(\overline{x})\Theta(\overline{x},t)=\frac{1}{\nu}e^{-t/\nu}\int_0^t e^{\tau/\nu}W_1(\overline{x},\tau)d\tau,$$
$$\Delta(\overline{x})u_j(\overline{x},t)=W_2(\overline{x},t). \tag{60}$$

For C-T theory, in view of Eqs. (13), (50), and (53), we get

$$\Delta(\overline{x})\Theta(\overline{x},t)=\frac{2}{\tau_q}e^{-t/\tau_q}[\Theta_1\sin(t/\tau_q)-\Theta_2\cos(t/\tau_q)] \tag{61}$$

$$\Delta(\overline{x})u_j(\overline{x},t)=\frac{2}{\tau_q}e^{-t/\tau_q}[u_1\sin(t/\tau_q)-u_2\cos(t/\tau_q)] \tag{62}$$

where

$$\Theta_1=\int_0^t e^{\tau/\tau_q}\cos(\tau/\tau_q)W_1(\overline{x},\tau)d\tau,$$
$$\Theta_2=\int_0^t e^{\tau/\tau_q}\sin(\tau/\tau_q)W_1(\overline{x},\tau)d\tau \tag{63}$$
$$u_1=\int_0^t e^{\tau/\tau_q}\cos(\tau/\tau_q)W_2(\overline{x},\tau)d\tau,$$
$$u_2=\int_0^t e^{\tau/\tau_q}\sin(\tau/\tau_q)W_2(\overline{x},\tau)d\tau. \tag{64}$$

Generalizations of the Green-Somiliana formula result from Eqs. (51)–(64) when $\Delta(\overline{x})=1$.

Letting $\overline{x}\to\overline{\xi}$, $\xi\in B$ and $\Delta(\overline{x})=1/2$ in Eqs. (51) and (54), we get

$$L_1(\Theta(\overline{\xi},t))=2W_1(\overline{\xi},t);\quad L_2(u_j(\overline{\xi},t))=2W_2(\overline{\xi},t). \tag{65}$$

Equation (65), with the boundary conditions (18) and (19) and the limiting behavior of the solutions, can be used to set up the system of linear equations of BIE method.

Example

In this section the mixed boundary conditions are considered to illustrate the BIE formulation. Let the problem to be solved is to determine $u_i(\overline{x},t)$ and $\Theta(\overline{x},t)$, $\overline{x}\in D$, $t>0$ the solution of the field Eqs. (5) and (6), subjected to the homogeneous initial conditions (17) and the following boundary conditions [28]:

$$u_i(\overline{x_B},t)=g_i=0,\quad \Theta_{,n}(\overline{x_B},t)=G=0,\quad x_B\in B_2=B_u \tag{66}$$

$$\sigma_{ji}(\overline{x_B},t)n_j(\overline{x_B})=f_i(\overline{x_B},t),\quad \Theta(\overline{x_B},t)=\Phi(\overline{x_B},t),$$
$$x_B\in B_1=B_\sigma \tag{67}$$

here $f_i(\overline{x_B},t)$ and $\Phi(\overline{x_B},t)$ are given, on $B_1(=B_\sigma)$, functions.

It is important to notice that, the traction vector $f_i(\overline{x_B},t)=\sigma_{ki}(\overline{x_B},t)n_k(\overline{x_B})$, and the surface temperature $\Phi(\overline{x_B},t)=\Theta(\overline{x_B},t)$ are unknown functions on the part $B_2(=B_u)$ of the surface.

Equations (47), (48) taking $\Delta(\overline{x})=1$, will give the solution in a more simple form if the Green functions $u_i^{(1)}$, $\Theta^{(1)}$, $u_i^{(j)}$, $\Theta^{(j)}$ are satisfying the conditions

$$f_i^{(1)}(\overline{x_B},t)=f_i^{(j)}(\overline{x_B},t)=0,\quad \Phi^{(1)}(\overline{x_B},t)=\Phi^{(j)}(\overline{x_B},t)=0,$$
$$x_B\in B_1=B_\sigma. \tag{68}$$

Equations (47) and (49) with Eqs. (66)–(68) lead to the relations

$$\omega\omega_2\Theta(\overline{x},s)=\Theta_0(\overline{x},s)-\omega_3\omega_2\int_{B_2}\Theta(\overline{y},s)G^{(1)}(\overline{y},\overline{x},s)dA(\overline{y})$$
$$-\varepsilon s\omega\int_{B_2}f_i(\overline{y},s)g_i^{(1)}(\overline{y},\overline{x},s)dA(\overline{y}), \tag{69}$$

$$\varepsilon s\omega u_j(\overline{x},s)=u_j^0(\overline{x},s)+\omega_3\omega_2\int_{B_2}\Theta(\overline{y},s)G^{(j)}(\overline{y},\overline{x},s)dA(\overline{y})$$
$$+\varepsilon s\omega\int_{B_2}f_i(\overline{y},s)g_i^{(j)}(\overline{y},\overline{x},s)dA(\overline{y}) \tag{70}$$

where $\Theta_0(\overline{x},s)$ and $u_j^0(\overline{x},s)$ are known functions, given in terms of the Green functions (34), (36), (41), and (43), given on the part $B_1=B_\sigma$ of the surface, functions $\Phi(\overline{x_B},s)$, $f_i(\overline{x_B},s)$, the mass force, and the heat source

$$\Theta_0(\overline{x},s)=\omega\omega_2\int_D Q(\overline{y},s)\Theta^{(1)}(\overline{y},\overline{x},s)dV(\overline{y})$$
$$-\varepsilon s\omega\int_D F_i(\overline{y},s)u_i^{(1)}(\overline{y},\overline{x},s)dV(\overline{y})$$
$$-\omega_3\omega_2\int_{B_1}\Phi(\overline{y},s)\Theta_{,n}^{(1)}(\overline{y},\overline{x},s)dA(\overline{y})$$
$$-\varepsilon s\omega\int_{B_1}f_i(\overline{y},s)u_i^{(1)}(\overline{y},\overline{x},s)dA(\overline{y}), \tag{71}$$

$$u_j^0(\overline{x},s)=\varepsilon s\omega\int_D F_i(\overline{y},s)u_i^{(j)}(\overline{y},\overline{x},s)dV(\overline{y})-\omega\omega_2\int_D Q(\overline{y},s)\Theta^{(j)}$$
$$\times(\overline{y},\overline{x},s)dV(\overline{y})+\varepsilon s\omega\int_{B_1}f_i(\overline{y},s)u_i^{(j)}(\overline{y},\overline{x},s)dA(\overline{y})$$
$$+\omega_3\omega_2\int_{B_1}\Phi(\overline{y},s)\Theta_{,n}^{(j)}(\overline{y},\overline{x},s)dA(\overline{y}). \tag{72}$$

To find the solution given by Eqs. (69) and (70) it is necessary to determine the two unknown functions $f_i(\overline{x_B},t)=\sigma_{ki}(\overline{x_B},t)n_k(\overline{x_B})$, and $\Theta(\overline{x_B},t)=\Phi(\overline{x_B},t)$ on the part B_2 $(=B_u)$ of the surface B. In Eqs. (69) and (70) letting $\overline{x}\to\overline{\xi}\in B_2$, and substituting from Eqs. (66), we get the following system of two singular Fredholm integral equations in the two unknown functions:

$$\begin{aligned}0=&\frac{\partial\Theta_0(\overline{\xi},s)}{\partial n'(\xi)}-\omega_2\omega_3\int_{B_2}\Theta(\overline{y},s)\frac{\partial G^{(1)}(\overline{y},\overline{\xi},s)}{\partial n'(\xi)}dA(\overline{y})\\&-\varepsilon s\omega\int_{B_2}f_i(\overline{y},s)\frac{\partial g_i^{(1)}(\overline{y},\overline{\xi},s)}{\partial n'(\xi)}dA(\overline{y})\\0=&u_j^{(0)}(\overline{\xi},s)+\omega_3\omega_2\int_{B_2}\Theta(\overline{y},s)G^{(j)}(\overline{y},\overline{\xi},s)dA(\overline{y})\\&+\varepsilon s\int_{B_2}f_i(\overline{y},s)g_i^{(j)}(\overline{y},\overline{\xi},s)dA(\overline{y}),\end{aligned} \tag{73}$$

where $n'(\overline{\xi})$ is outer normal vector to B_2.

For general boundary shapes the system of Eqs. (75) do not seem to have analytical solutions, whence the necessity of recurring to numerical techniques. The integrals have to be discretized and the problem reduces to finding the solution of a system of linear algebraic equations.

Conclusion

The direct formulation (applying the Betti-reciprocical theorem) of the boundary integral equation method, in Laplace transform domain, is given for generalized linear thermoviscoelasticity. The Green functions for the corresponding differential equations are obtained. For the mixed boundary value problem a system of two singular Fredholm integral equations, in two unknown functions on a part of the boundary, is obtained and the necessity of recurring to the numerical methods (namely BEM) is shown. For any smooth enough boundary shape (to guarantee existence of unique normal to the boundary at each of its points) the integrals involved in the system of the integral equations have to be discretized and the problem reduces to finding the solution (in Laplace transform domain) of a system of linear algebraic equations. Using a numerical inversion method (see, e.g., Honig and Hirdes [34]) the solutions in the physical domain can be obtained.

Nomenclature

t = time; λ, μ Lame's constants; $K=\lambda+2/3\mu$ bulk modulus
T_0 = reference temperature; T absolute temperature such that $|T-T_0/T_0|\ll 1$
ρ = density; C_E specific heat at constant strain; k thermal conductivity
σ_{ij} = Components of stress tensor; S_{ij} components of stress deviator tensor
ε_{ij} = components of strain tensor; e_{ij} components of strain deviator tensor
u_i = components of displacement vector; $e=u_{i,i}$ dilatation
$R_G(t)$, $R_K(t)$ = relaxation functions
τ_0, ν, t_1, t_2, n_0, δ_0 = constants
τ_θ = phase-lag of temperature gradient; τ_q phase-lag of heat flux
α_T = coefficient of linear thermal expansion; $\gamma=3K\alpha_T$
η_0 = $\rho C_E/k$; $\varepsilon=\delta_0\gamma/\rho C_E$; $C_0^2=K/\rho$; $\hat{T}=T-T_0+\nu\dot{T}$
F_i = components of mass force vector

Q = the intensity of applied heat source per unit mass

ϵ_{ijk} = the permutation tensor; δ_{ij} delta Kronecker; $\delta(\ldots)$ Dirac delta ''function''

References

[1] Fredudenthal, A. M., 1954, ''Effect of Rheological Behavior on Thermal Stresses,'' J. Appl. Phys., **25**, pp. 1110–1117.
[2] Fung, Y. C., 1968, *Foundations of Solid Mechanics*, Prentice-Hall, Englewood Cliffs, NJ.
[3] Bland, D. R., 1960, *The Theory of Linear Viscoelasticity*, Pergamon Press, Oxford, UK.
[4] Gurtin, M. E., and Sternberg, E., 1962, ''On the Linear Theory of Viscoelasticity,'' Arch. Ration. Mech. Anal., **11**, pp. 291–356.
[5] Christensen, R. M., 1971, *Theory of Viscoelasticity—An Introduction*, Academic Press, San Diego, CA.
[6] Ilioushin, A. A., and Pobedria, B. E., 1970, *Fundamentals of the Mathematical Theory of Thermal Visco-Elasticity*, Nauka, Moscow.
[7] Koltunov, M. A., 1976, *Creeping and Relaxation*, Izd. Vyschaya Shkola, Moscow.
[8] Rabotnov, Yu. N., 1980, *Elements of Hereditary Solid Mechanics*, Mir, Moscow.
[9] Kovacs, A. J., 1958, ''La Contraction Isotherme du Volume des Polymeres Amorphes,'' J. Polym. Sci., Part A: Gen. Pap., **30**, p. 131.
[10] Biot, M., 1956, ''Thermoelasticity and Irreversible Thermodynamics,'' J. Appl. Phys., **27**, pp. 240–253.
[11] Lord, H., and Shulman, Y., 1967, ''A Generalized Dynamical Theory of Thermo-Elasticity,'' J. Mech. Phys. Solids, **15**, pp. 299–309.
[12] Muller, I. M., 1971, ''The Coldness, A Universal Function in Thermoelastic Bodies,'' Arch. Ration. Mech. Anal., **41**, pp. 319–332.
[13] Green, A. E., and Laws, N., 1972, ''On the Entropy Production Inequality,'' Arch. Ration. Mech. Anal., **45**, pp. 47–53.
[14] Green, A. E., and Lindsay, K. A., 1972, ''Thermoelasticity,'' J. Elast., **2**, pp. 1–7.
[15] Suhubi, E. S., 1975, ''Thermoelastic Solids,'' *Continuum Physics*, A. C. Eringen, ed., Academic Press, San Diego, CA, **2**, Part 2, Chap. 2.
[16] Tzou, D. Y., 1995, ''A Unified Approach for Heat Conduction From Macro to Micro-Scales,'' ASME J. Heat Transfer, **117**, pp. 8–16.
[17] Chandraskharaiah, D. S., 1998, ''Hyperbolic Thermoelasticity, A Review of Recent Literature,'' Appl. Mech. Rev., **51**, pp. 705–729.
[18] Ignaczak, J., 1989, ''Generalized Thermoelasticity and Its Applications,'' *Mechanical and Mathematical Methods* (Thermal Stresses III), R. B. Hetnarski, ed., North-Holland Amsterdam.
[19] Hetnarski, R. B., and Ignaczak, J., 1999, ''Generalized Thermoelasticity,'' J. Therm. Stresses, **22**, pp. 451–476.
[20] Cruse, T. A., and Rizzo, F. J., eds., 1985, ''Boundary Integral Equation Methods-Computational Applications in Applied Mechanics,'' *Proc. ASME Conf. on Boundary Integral Equation Methods* (ASME, New York).
[21] Rizzo, F. J., and Shippy, D. J., 1979, ''Recent Advances of the Boundary Element Method in Thermoelasticity,'' *Developments in Boundary Element Methods*, P. K. Banerjee and R. Butterfield, eds., Applied Science Publishers, London, **1**, pp. 155–172.
[22] Banerjee, P. K., and Butterfield, R., 1981, *Boundary Element Methods in Engineering Science*, McGraw-Hill, New York.
[23] Brebbia, C. A., Telles, J. C. F., and Wrobel, L. C., 1984, *Boundary Element Techniques*, Springer-Verlag, Berlin.
[24] Ziegler, Franz, and Irschik, Hans, 1987, ''Thermal Stress Analysis Based on Maysel's Formula,'' *Mechanical and Mathematical Methods* (Thermal Stresses II), R. B. Hetnarski, ed., North Holland, Amsterdam.
[25] Sladek, V., and Sladek, J., 1983, ''Boundary Integral Equation Method in Thermoelasticity,'' Appl. Math. Model., **7**, pp. 241–253.
[26] Jaswon, M. A., and Symm, G. T., 1977, *Integral Equation Methods in Potential Theory and Elastostatics*, Academic Press, London.
[27] Coleman, B. D., 1964, ''Thermodynamics of Materials With Memory,'' Arch. Ration. Mech. Anal., **17**, pp. 1–46.
[28] Nowacki, W., 1975, *Dynamic Problems of Thermoelasticity*, Noordhoff, Leyden, The Netherland.
[29] Alfrey, T., and Gurnee, E. F., 1956, *Rheology Theory and Applications*, F. R. Eirich, ed., Academic Press, San Diego, CA.
[30] Cherchill, R. V., 1972, *Operational Mathematics*, 3rd Ed., Mc Graw-Hill, New York.
[31] Nowacki, W., 1962, *Thermoelasticity*, Pergamon Press, London.
[32] Morse, P., and Feshbach, H., 1953, *Methods of Theoretical Physics*, McGraw-Hill, New York.
[33] Ezzat, Magdy A., and El-Karamany, Ahmed S., 2002, ''The Uniqueness and Reciprocity Theorems for Generalized Thermoviscoelasticity for Anisotropic Media,'' J. Therm. Stresses, **25(6)**, pp. 507–522.
[34] Honig, G., and Hirdes, U., 1984, ''A Method for the Numerical Inversion of the Laplace Transform,'' J. Comput. Appl. Math., **10**, pp. 113–132.

Y.-L. Chung
Professor
e-mail: chung@hp.ct.ntust.edu.tw

M.-R. Chen
Graduate student

Department of Construction Engineering,
National Taiwan University
of Science and Technology,
Taipei, Taiwan

Dynamic Analysis of a Mode I Propagating Crack Subjected to a Concentrated Load

This work investigates the phenomenon of mode I central crack propagating with a constant speed subjected to a concentrated load on the crack surfaces. This problem is not a self-similar problem. However, the method of self-similar potential (SSP) in conjunction with superposition can be successfully applied if the time delay and the origin shift are considered. After the complete solution is obtained, attention is stressed on the dynamic stress intensity factors (DSIFs). Analytical results indicate that the DSIF equals the static stress intensity factor if the crack-tip speed is very slow and equal to zero if the crack-tip velocity approaches the Rayleigh-wave speed. However, the dynamic effect becomes obvious only if the crack-tip speed is 0.4 times faster than the S-wave speed. Moreover, the combination of SSP method and the superposition scheme can be applied to the expanding uniformly distributed load acting on a portion of the crack surfaces.
[DOI: 10.1115/1.1600473]

1 Introduction

Dynamic fracture-related studies can be categorized as steady-state dynamics and transient dynamics. In a steady-state solution, the coordinate system moving with a crack tip is used. The spatial field expressed in this coordinate system is assumed not to depend on time. The conventional procedure for the transient analysis, which considers the full dynamic equations and the actual initial and boundary conditions, is a transform method with the Wiener-Hopf and Cagniard techniques [1]. The transient approach frequently encounters mathematical difficulty in deriving an effective solution.

The transient problem is markedly simplified when the crack expands from zero length with a constant crack-tip speed. The resulting problem may be termed a self-similar problem and the method of self-similar potentials (SSP) can be used to effectively solve it. Pioneered by Smirov [2] in the 1930s and later applied to wave propagation problems, the SSP method not only significantly reduces computational effort, but also directly leads to identification of fronts and types of reflected and refracted waves. In 1969, Thompson and Robinson [3] reviewed pertinent literature of this method and applied it to various dynamic indentation problems in a linearly elastic half space. Since the dynamic crack propagation problem is also closely related to the dynamic indentation problem [3], the choice of the SSP approach to solve dynamic crack propagation problems is one that gives considerable promise of success [4,5].

Despite its name, the SSP method is not limited to solutions to problems having a physical similarity. This method has been applied to resolve problems of a subsurface source [6] and suddenly stopping cracks [7], in which they do not exhibit any physical similarity.

Dynamic fracture analysis often considers problems of a semi-infinite running crack subjected to external loads. In a series of investigations, Freund [8–10] concluded that the dynamic stress intensity factor (DSIF) for a semi-infinite running crack with uniform or nonuniform extend rate is given by the product of a "universal function" of crack-tip velocity and the stress intensity factor for an equivalent stationary crack. Lamb's solution to the problem of a suddenly applied concentrated load at a point on a surface of a half space marked a great advance while analyzing the wave motion in a solid, [11]. Freund [12] further used the Lamb's solution to study a semi-infinite stationary crack subjected to a concentrated load. For the crack with a finite length, most researchers considered a uniform load acting on the crack surfaces so that the extension occurs in mode I [13], mode II [14], or mode III [15]. However, the case of a finite crack subjected to a concentrated load has seldom been addressed. Therefore, in this work, we investigate a concentrated load acting on the crack surfaces of a running crack. Although this is not a self-similar problem, this work also applied the SSP method in which the origin shift and time delay are considered.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 9, 1999; final revision, Dec. 17, 2002. Associate Editor: V. K. Kinra. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 The Method of Self-Similar Potentials

The basis of the method is to seek a class of solutions of plane elastodynamic equations in terms of self-similar potentials in which the value of the function is zero for time $t<0$ and is constant on straight lines passing through the origin of the (x,y,t) space for time $t\geqslant 0$. Self-similar problems arise when the initial and boundary conditions are homogeneous functions of the spatial variables and time. The homogeneity of the boundary values makes it possible in the method of self-similar potentials to reduce the three independent variables (x,y,z) to two variables $(x/t,y/t)$. It can be inferred that a self-similar solution to a physical problem can be found if the loading is suddenly applied at the origin, or, with initial conditions zero, if the boundary conditions involve a characteristic length proportional to time.

This section summarizes the general application of self-similar potentials to the solution of some elasticity problems for the half plane [3,6]. For homogeneous, isotropic, and elastic plane problems, if the boundary tractions of the elastic half plane are homogeneous functions of degree zero in space and time, the displacement and stress fields under strain deformation expressed in terms of self-similar potentials $\Phi'(\theta)$ and $\Psi'(\theta)$ are

$$u_x = -\mathrm{Re}\left[\int_0^t \int_0^{\theta_1} \theta\Phi'(\theta)\,d\theta\,d\tau + \int_0^t \int_0^{\theta_2} \sqrt{b^{-2}-\theta^2}\,\Psi'(\theta)\,d\theta\,d\tau\right], \tag{1a}$$

$$u_y = \mathrm{Re}\left[\int_0^t \int_0^{\theta_1} -\sqrt{a^{-2}-\theta^2}\Phi'(\theta)d\theta d\tau + \int_0^t \int_0^{\theta_2} \theta\Psi'(\theta)d\theta\, d\tau\right], \tag{1b}$$

$$\sigma_x = \mu\ \mathrm{Re}\left[\int_0^{\theta_1}(b^{-2}-2a^{-2}+2\theta^2)\Phi'(\theta)d\theta + \int_0^{\theta_2} 2\theta\sqrt{b^{-2}-\theta^2}\Psi'(\theta)d\theta\right], \tag{1c}$$

$$\sigma_y = \mu\ \mathrm{Re}\left[\int_0^{\theta_1}(b^{-2}-2\theta^2)\Phi'(\theta)d\theta - \int_0^{\theta_2} 2\theta\sqrt{b^{-2}-\theta^2}\Psi'(\theta)d\theta\right], \tag{1d}$$

$$\sigma_z = \mu\ \mathrm{Re}\int_0^{\theta_1}(b^{-2}-2a^{-2})\Phi'(\theta)d\theta, \tag{1e}$$

$$\tau_{xy} = \mu\ \mathrm{Re}\left[\int_0^{\theta_1} 2\theta\sqrt{a^{-2}-\theta^2}\Phi'(\theta)d\theta + \int_0^{\theta_2}(b^{-2}-2\theta^2)\Psi'(\theta)d\theta\right], \tag{1f}$$

where μ, a, and b are the shear modulus, longitudinal, and transverse wave velocities of a homogeneous isotropic elastic medium, respectively. The functions $\Phi'(\theta)$ and $\Psi'(\theta)$ indicate the derivatives of the self-similar potentials $\Phi(\theta)$ and $\Psi(\theta)$ which are unknown in general for the boundary-valued problems. The values θ_1 and θ_2 represent the parameters of the characteristic surfaces of the wave equations $\delta_j = t-\theta_j x - y\sqrt{c_j^{-2}-\theta_j^2}=0$, $j=1$, 2 and they are defined as

$$\theta_j = \frac{1}{r^2}(tx+iy\sqrt{t^2-r^2c_j^{-2}}) \quad \text{for } x^2+y^2 \leqslant c_j^2t^2, \tag{2a}$$

$$\theta_j = \frac{1}{r^2}[tx+\mathrm{sgn}(x)y\sqrt{r^2c_j^{-2}-t^2}] \quad \text{for } x^2+y^2 \geqslant c_j^2t^2, \tag{2b}$$

with $j=1$, 2 and $c_1=a$, $c_2=b$. It is easy to verify, from Eq. (2), that on the surface $y=0$, the parameters θ_1 and θ_2 are equal to each other, i.e., $\theta_1=\theta_2=t/x$. Therefore, we denote t/x as θ_0 and then we have $\theta_1=\theta_2=\theta_0$ on the surface $y=0$.

It is worth noting that the combination of the SSP method and Schwarz integral theorem [16] provides an extremely powerful technique for solving the boundary-valued problems, e.g., wave propagation problems, with such less mathematical manipulation [3,6]. However, for crack problems which are mixed-boundary problems, the Schwartz integral theorem cannot be used directly and therefore an approach of function-theoretic integral equation could be applied.

3 Problem Description and the Complete Solution

Consider an unbounded homogeneous isotropic elastic medium. The material is at rest at time $t<0$. A crack begins to extend symmetrically from zero length along the x axis with a constant speed s for time $t \geqslant 0$. The x coordinates of the crack tips at time t are then $\pm st$. A concentrated loading acts on the origin of the crack surfaces such that the state of deformation is a plane strain. A central crack subjected to concentrated loading is no longer a self-similar problem. To solve this problem the method of self-similar potentials is still applicable by a superposition scheme illustrated in Fig. 1. Part (i) assumes that a uniform stress $\sigma_y = -q$ acts on the faces of the crack, the ends of which move out with a constant speed s; denote this case as case A. In part (ii) a uniform stress $\sigma_y = q$ acts on $x \geqslant |x_0|$ of the crack surfaces. The problem of part (ii) is further divided into two cases: case B with stress $\sigma_y = q$ in the cracked portion to the right and case C where it is to the left. Finally, the solution of the original problem can be obtained by summing up the solutions of cases A, B, and C and taking $x_0 \to 0$.

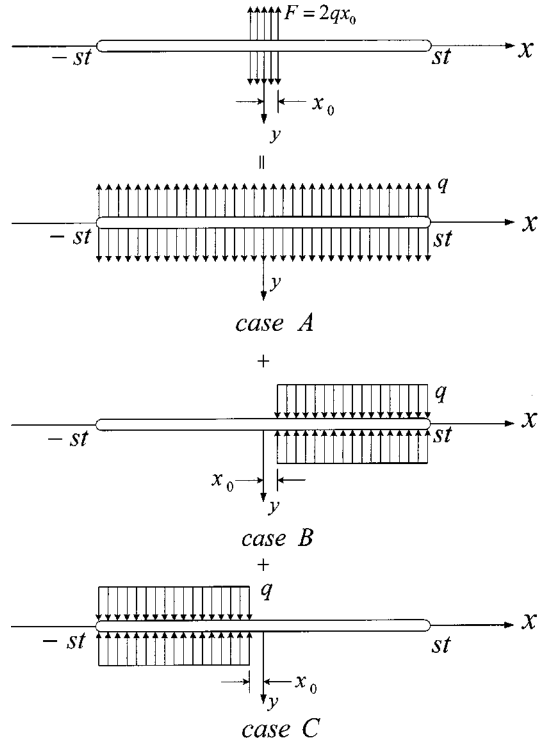


Fig. 1 The superposition scheme for a concentrated load acting on the crack surfaces

3.1 Solution to the Case A Problem. From symmetry about the x axis, consideration can be limited to the region $y \geqslant 0$ while subjected to the following boundary conditions:

$${}_{\mathrm{A}}\sigma_y(x,0,t) = -q, \quad |x|<st \tag{3a}$$

$${}_{\mathrm{A}}\tau_{xy}(x,0,t) = 0, \quad \forall x \tag{3b}$$

$${}_{\mathrm{A}}u_y(x,0,t) = 0, \quad |x|>st \tag{3c}$$

where the subscript A denotes the solution of the case A problem. Notably, the boundary tractions in Eqs. (3) are homogeneous and of degree zero in x and t. The velocities and displacements at every point must be homogeneous functions of degree zero and one, respectively. Equation (3b) reveals that the absence of the shear stress on the entire surface $y=0$. Substituting Eq. (3b) into Eq. (1f) leads to the following relation:

$${}_{\mathrm{A}}\Psi'(\theta) = \frac{-2\theta\sqrt{a^{-2}-\theta^2}}{b^{-2}-2\theta^2}\,{}_{\mathrm{A}}\Phi'(\theta). \tag{4}$$

Substituting Eq. (4) into Eqs. (1d) and (1b) and defining a new unknown function ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$

$${}_{\mathrm{A}}V_y^{*\prime}(\theta) = \frac{-b^{-2}\sqrt{a^{-2}-\theta^2}}{b^{-2}-2\theta^2}\,{}_{\mathrm{A}}\Phi'(\theta). \tag{5}$$

Then the boundary conditions can be rewritten as

$$ {}_{\mathrm{A}}\sigma_y = \mathrm{Re}\int_0^{\theta_0} \frac{-\mu b^2 R(\theta^2)}{\sqrt{a^{-2}-\theta^2}}\, {}_{\mathrm{A}}V_y^{*\prime}(\theta)\, d\theta = -q, \quad |\theta_0| > s^{-1} \tag{6a} $$

$$ {}_{\mathrm{A}}v_y = \mathrm{Re}\int_0^{\theta_0} {}_{\mathrm{A}}V_y^{*\prime}(\theta)\, d\theta = 0, \quad |\theta_0| < s^{-1} \tag{6b} $$

where ${}_{\mathrm{A}}v_y$ is the velocity in the y direction, i.e., ${}_{\mathrm{A}}v_y = \partial_{\mathrm{A}} u_y / \partial t$, and $R(\theta^2) = (b^{-2} - 2\theta^2)^2 + 4\theta^2 \sqrt{a^{-2}-\theta^2}\sqrt{b^{-2}-\theta^2}$ is the Rayleigh function. Obviously the quantity ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ represents the derivative of complex-valued velocity in the y direction. The problem is now to solve the mixed-integral equations of Eqs. (6) which involve an unknown function ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$. To do so, the boundary condition, Eqs. (6), is satisfied if the following conditions are fulfilled:

(i) $-\mu b^2 R(\theta^2)/\sqrt{a^{-2}-\theta^2}\, {}_{\mathrm{A}}V_y^{*\prime}(\theta)$ is analytic for $|\theta_0| > s^{-1}$,
(ii) ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ is analytic for $|\theta_0| < s^{-1}$.

Condition (i) clearly reveals that $-\mu b^2 R(\theta^2)/\sqrt{a^{-2}-\theta^2}$ has branch cuts either from $\pm a^{-1}$ to $\pm\infty$ or from $\pm b^{-1}$ to $\pm\infty$ in the θ plane. To locate all of the branch cuts of ${}_{\mathrm{A}}\sigma_y^{*\prime}(\theta) = -\mu b^2 R(\theta^2)\, {}_{\mathrm{A}}V_y^{*\prime}(\theta)/\sqrt{a^{-2}-\theta^2}$ in the range $|\theta_0| < s^{-1}$ such that condition (i) holds, the quantity ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ should have branch points at $\theta = \pm s^{-1}$. Moreover, the stress is nowhere more singular than that under a concentrated load of finite magnitude, the highest allowable singularity in ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ at a point in the θ plane is a pole of order 2 [3]. It follows directly that ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ must be of the following form:

$$ {}_{\mathrm{A}}V_y^{*\prime}(\theta) = \sum_{n=-\infty}^{\infty} {}_{\mathrm{A}}F_n(\theta)(s^{-2}-\theta^2)^{n+1/2}, \tag{7} $$

where $F_n(\theta)$ is an entire function and n is an integer. However, to have stresses at the origin of the physical plane that are less singular than those corresponding to a concentrated load, the highest order of ${}_{\mathrm{A}}\sigma_y^{*\prime}(\theta)$ should be a power of order -1 when θ approaches infinity. However, $-\mu b^2 R(\theta^2)/\sqrt{a^{-2}-\theta^2}$ approaches θ as $\theta \to \infty$. Therefore, the highest order of ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ is a power of order -2 as $\theta \to \infty$. Hence, we can infer that ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ must be the form of

$$ {}_{\mathrm{A}}V_y^{*\prime}(\theta) = \frac{iA_1}{(s^{-2}-\theta^2)^{3/2}} + \frac{iA_2\theta}{(s^{-2}-\theta^2)^{3/2}}. \tag{8} $$

Substituting Eq. (8) into Eq. (6) reveals that A_1 and A_2 should be real-valued constants. Since the branch cuts of ${}_{\mathrm{A}}V_y^{*\prime}(\theta)$ are taken from $\pm s^{-1}$ to $\pm\infty$, the first and second terms in Eq. (8) are antisymmetric and symmetric functions of θ, respectively, for $|\theta_0| > s^{-1}$. In addition, these terms provide symmetric and antisymmetric normal surface velocity on the crack faces because the integration of an antisymmetric (antisymmetric) function yields a symmetric (antisymmetric) result. However, for the case A problem, the normal surface velocity is symmetric about the y axis. Therefore, $A_2 = 0$, and

$$ {}_{\mathrm{A}}V_y^{*\prime}(\theta) = \frac{iA_1}{(s^{-2}-\theta^2)^{3/2}}. \tag{9} $$

The constant A_1 can be obtained by substituting Eq. (9) into Eq. (6) and is equal to q/J, where

$$ J = \mathrm{Re}\int_0^{\theta_0} \frac{i\mu b^2 R(\theta^2)}{\sqrt{a^{-2}-\theta^2}(s^{-1}-\theta)^{3/2}}\, d\theta \quad \text{for } |\theta_0| > s^{-1}. \tag{10} $$

The branch cuts of the integrand of J are all located in the range $a^{-1} < |\theta| < s^{-1}$ on the real axis of the θ plane. Therefore, the upper and lower limits of the integration in Eq. (10) can be changed to $\pm a^{-1}$ and $\pm s^{-1}$. This observation implies that the value of J is constant for any $|\theta_0| > s^{-1}$.

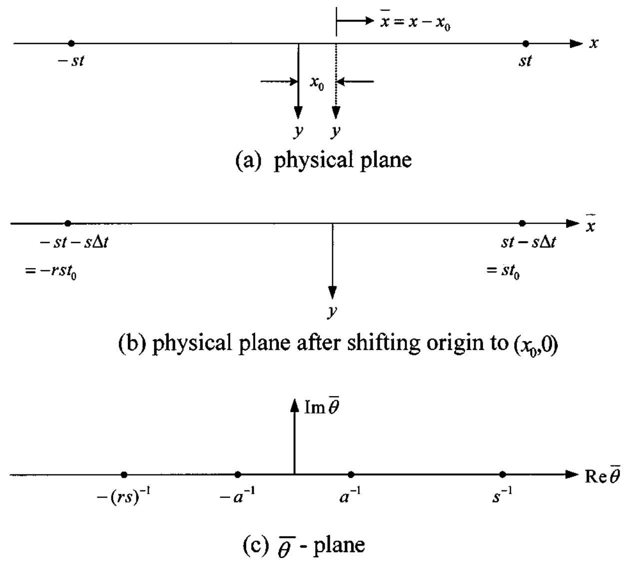


Fig. 2 The case B problem: (*a*) on the physical plane, (*b*) on the physical plane after shifting the origin to (x_0,0), (*c*) on the complex $\bar{\theta}$ plane

The solution here is the same as that obtained by Cherepanov [17]; he expressed the integral of Eq. (10) in terms of incomplete elliptical integrals of first and second kinds. However, doing so would appear to be undesirable. Indeed, the quadrature in Eq. (10) can be easily calculated as accurately as desired by selecting any path off the real axis of the θ plane, while the incomplete elliptic integrals are not well tabulated. A convenient path proceeds along the imaginary axis of the θ plane from the origin to positive infinity.

The displacement and stress fields of the case A problem can be obtained by substituting Eq. (9) into Eqs. (5) and (4) to obtain the self-similar potentials $\Phi'(\theta)$ and $\Psi'(\theta)$ first, and then substituting the self-similar potentials $\Phi'(\theta)$ and $\Psi'(\theta)$ into Eq. (1). Herein, we do not include the results of case A problem, whereas only the y-component stress ${}_{\mathrm{A}}\sigma_y$ is listed here:

$$ {}_{\mathrm{A}}\sigma_y = (-\mu b^2 A_1)\mathrm{Re}\left\{ \int_0^{\theta_1} \frac{i(b^{-2}-2\theta^2)^2}{\sqrt{a^{-2}-\theta^2}(s^{-2}-\theta^2)^{3/2}}\, d\theta + \int_0^{\theta_2} \frac{4i\theta^2\sqrt{b^{-2}-\theta^2}}{(s^{-2}-\theta^2)^{3/2}}\, d\theta \right\}. \tag{11} $$

The first and second terms of Eq. (11) represent the disturbances of the P and S waves, respectively.

3.2 Solution to the Case B Problem. The problem of case B is no longer a self-similar problem. However, the SSP method can be applied if the origin shift and time delay are considered. Define the new coordinate system $(\bar{x},\bar{y})$ which are $\bar{x} = x - x_0$, $\bar{y} = y$, and the new time system $t_0 = t - \Delta t$. Then, the crack-tip coordinate which was $x = st$ in the old coordinate system (x,y) becomes $\bar{x} = st - s\Delta t = st_0$ in the new coordinate system $(\bar{x},\bar{y})$. The crack tip of $x = -st$ in the old coordinate system (x,y) becomes $\bar{x} = -st - s\Delta t = -rst$ in the new coordinate system $(\bar{x},\bar{y})$, where $r = (t+\Delta t)/(t-\Delta t)$ is dimensionless, as illustrated in Fig. 2. Therefore, after the origin shift and time delay are used, the case B problem can be viewed as the situation in which the rightmost crack tip propagates in the positive x direction with speed s and the leftmost crack tip extends in the negative x direction with speed rs. Consequently, the boundary conditions of case B problem are

$$_{\mathrm{B}}\sigma_y(\bar{x},0,t_0)=q, \quad 0<\bar{x}<st_0 \tag{12a}$$

$$_{\mathrm{B}}\sigma_y(\bar{x},0,t_0)=0, \quad -rst_0<\bar{x}<0 \tag{12b}$$

$$_{\mathrm{B}}\tau_{xy}(\bar{x},0,t_0)=0, \quad \forall \bar{x} \tag{12c}$$

$$_{\mathrm{B}}u_y(\bar{x},0,t_0)=0, \quad \bar{x}>st_0 \text{ or } \bar{x}<-rst_0 \tag{12d}$$

where the subscript B denotes the case B problem. Notably, the use of the origin shift and time delay converts the non-self-similar problem into a self-similar problem. Therefore, the case B problem can be treated in the complex $\bar{\theta}$ plane, where $\bar{\theta}$ satisfies the characteristic surface:

$$(t-\Delta t)-\bar{\theta}_j(x-x_0)-y\sqrt{c_j^{-2}-\bar{\theta}_j^2}=0, \quad j=1,2. \tag{13}$$

Hence, the boundary conditions of Eq. (12) can be rewritten as

$$_{\mathrm{B}}\sigma_y=q, \quad \bar{\theta}_0>s^{-1} \tag{14a}$$

$$_{\mathrm{B}}\sigma_y=0, \quad \bar{\theta}_0<-(rs)^{-1} \tag{14b}$$

$$_{\mathrm{B}}\tau_{xy}=0, \quad \forall \bar{\theta}_0 \tag{14c}$$

$$_{\mathrm{B}}u_y=0, \quad -(rs)^{-1}<\bar{\theta}_0<s^{-1} \tag{14d}$$

where $\bar{\theta}_0=\bar{\theta}_1=\bar{\theta}_2=t_0/\bar{x}$ are the parameters of $\bar{\theta}_1$ and $\bar{\theta}_2$ on the surface $\bar{y}=0$. With the aid of the absence of the shear stress on the entire surface $\bar{y}=0$, the boundary conditions of Eqs. (14) can be expressed as

$$_{\mathrm{B}}\sigma_y=\mathrm{Re}\int_0^{\bar{\theta}_0}\frac{-\mu b^2R(\theta^2)}{\sqrt{a^{-2}-\theta^2}}\,{}_{\mathrm{B}}V_y^{*\prime}(\theta)d\theta=q, \quad \bar{\theta}_0>s^{-1} \tag{15a}$$

$$_{\mathrm{B}}\sigma_y=\mathrm{Re}\int_0^{\bar{\theta}_0}\frac{-\mu b^2R(\theta^2)}{\sqrt{a^{-2}-\theta^2}}\,{}_{\mathrm{B}}V_y^{*\prime}(\theta)d\theta=0, \quad \bar{\theta}_0<-(rs)^{-1} \tag{15b}$$

$$_{\mathrm{B}}v_y=\mathrm{Re}\int_0^{\bar{\theta}_0}{}_{\mathrm{B}}V_y^{*\prime}(\theta)d\theta=0, \quad -(rs)^{-1}<\bar{\theta}_0<s^{-1}. \tag{15c}$$

To solve the mixed boundary value problem in Eq. (15), the unknown function ${}_{\mathrm{B}}V_y^{*\prime}(\theta)$ must be obtained first. In a manner similar to the case A problem, the unknown function ${}_{\mathrm{B}}V_y^{*\prime}(\theta)$ is found as follows:

$$_{\mathrm{B}}V_y^{*\prime}(\theta)=\frac{iB_1}{(s^{-1}-\theta)^{3/2}[(rs)^{-1}+\theta]^{3/2}}+\frac{iB_2\theta}{(s^{-1}-\theta)^{3/2}[(rs)^{-1}+\theta]^{3/2}}, \tag{16}$$

where B_1 and B_2 are real-valued constants which can be evaluated from boundary conditions of Eqs. (15a) and (15b). They are

$$B_1=\frac{q\bar{J}_{22}}{J_{11}J_{22}-J_{12}J_{21}}, \quad B_2=\frac{-q\bar{J}_{12}}{J_{11}J_{22}-J_{12}J_{21}}, \tag{17}$$

where

$$\bar{J}_{11}=\mathrm{Re}\int_0^{\bar{\theta}_0}\bar{I}(\theta)d\theta \quad \text{for } \bar{\theta}_0>s^{-1}, \tag{18a}$$

$$\bar{J}_{21}=\mathrm{Re}\int_0^{\bar{\theta}_0}\bar{I}(\theta)\theta d\theta \quad \text{for } \bar{\theta}_0>s^{-1}, \tag{18b}$$

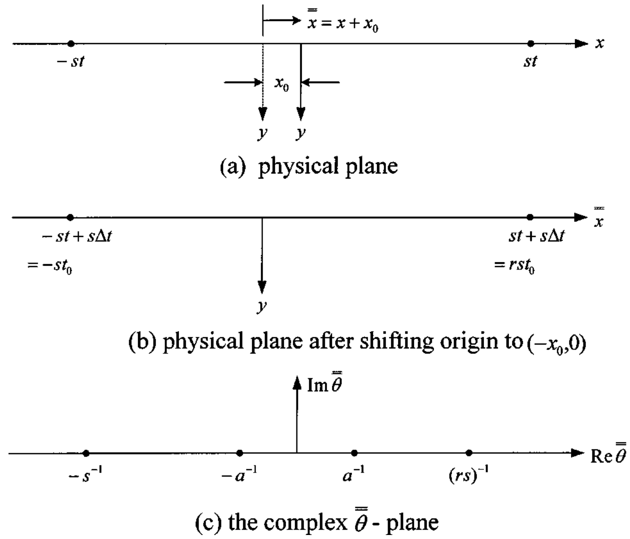


Fig. 3 The case C problem: (*a*) on the physical plane, (*b*) on the physical plane after shifting the origin to $(x_0,0)$, (*c*) on the complex $\bar{\theta}$ plane.

$$\bar{J}_{12}=\mathrm{Re}\int_0^{\bar{\theta}_0}\bar{I}(\theta)d\theta \quad \text{for } \bar{\theta}_0<-(rs)^{-1}, \tag{18c}$$

$$\bar{J}_{22}=\mathrm{Re}\int_0^{\bar{\theta}_0}\bar{I}(\theta)\theta d\theta \quad \text{for } \bar{\theta}_0<-(rs)^{-1}, \tag{18d}$$

with

$$\bar{I}(\theta)=\frac{-i\mu b^2R(\theta^2)}{(a^{-2}-\theta^2)^{1/2}(s^{-1}-\theta)^{3/2}[(rs)^{-1}+\theta]^{3/2}}.$$

The displacements and stresses of the case B problem then can be obtained without difficulty. However, only the stress ${}_{\mathrm{B}}\sigma_y$ is listed here and it is

$$_{\mathrm{B}}\sigma_y=-\mu b^2\,\mathrm{Re}\left\{\int_0^{\bar{\theta}_1}\frac{i(B_1+B_2\theta)(b^{-2}-2\theta^2)^2}{\sqrt{a^{-2}-\theta^2}(s^{-1}-\theta)^{3/2}[(rs)^{-1}+\theta]^{3/2}}d\theta+\int_0^{\bar{\theta}_2}\frac{4i\theta^2(B_1+B_2\theta)\sqrt{b^{-2}-\theta^2}}{(s^{-1}-\theta)^{3/2}[(rs)^{-1}+\theta]^{3/2}}d\theta\right\}. \tag{19}$$

3.3 Solution to the Case C Problem. The case C problem is not a self-similar problem either. However, if we shift the origin to $x=-x_0$ and consider the time delay Δt, the new coordinate system $(\bar{\bar{x}},\bar{\bar{y}})$ and new time system t_0 are defined as $\bar{\bar{x}}=x+x_0$, $\bar{\bar{y}}=y$, and $t_0=t-\Delta t$. Consequently, the crack tip $x=st$ in the new coordinate and time system is $\bar{\bar{x}}=st+s\Delta t=rst_0$, and the left crack tip $x=-st$ is $\bar{\bar{x}}=-st+s\Delta t=-st_0$ (Fig. 3), where $r=(t+\Delta t)/(t-\Delta t)$. The boundary conditions of the case C problem in the complex $\bar{\bar{\theta}}$ plane are

$$_{\mathrm{C}}\sigma_y=q, \quad \bar{\bar{\theta}}_0<-s^{-1} \tag{20a}$$

$$_{\mathrm{C}}\sigma_y=0, \quad \bar{\bar{\theta}}_0>(rs)^{-1} \tag{20b}$$

$$_{\mathrm{C}}\tau_{xy}=0, \quad \forall \bar{\bar{\theta}}_0 \tag{20c}$$

$$_{\mathrm{C}}u_y=0, \quad -s^{-1}<\bar{\bar{\theta}}_0<(rs)^{-1} \tag{20d}$$

where the subscript C indicates the case C problem and $\bar{\bar{\theta}}$ satisfies the characteristic surface:

$$(t-\Delta t)-\bar{\bar{\theta}}_j(x+x_0)-y\sqrt{c_j^{-2}-\bar{\bar{\theta}}_j^2}=0, \quad j=1,2. \tag{21}$$

With the similar procedure in the case B problem, the solution to the case C problem can be obtained. Only the stress ${}_{\mathrm{C}}\sigma_y$ is listed here:

$$
{}_{\mathrm{C}}\sigma_y=-\mu b^2\,\mathrm{Re}\left\{\int_0^{\bar{\bar{\theta}}_1}\frac{i(C_1+C_2\theta)(b^{-2}-2\theta^2)^2}{\sqrt{a^{-2}-\theta^2}[(rs)^{-1}-\theta]^{3/2}(s^{-1}+\theta)^{3/2}}d\theta\right.
$$

$$
\left.+\int_0^{\bar{\bar{\theta}}_2}\frac{4i\theta^2(C_1+C_2\theta)\sqrt{b^{-2}-\theta^2}}{[(rs)^{-1}-\theta]^{3/2}(s^{-1}+\theta)^{3/2}}d\theta\right\}, \tag{22}
$$

where

$$
C_1=\frac{q\bar{\bar{J}}_{22}}{\bar{\bar{J}}_{11}\bar{\bar{J}}_{22}-\bar{\bar{J}}_{12}\bar{\bar{J}}_{21}},\quad C_2=\frac{-q\bar{\bar{J}}_{12}}{\bar{\bar{J}}_{11}\bar{\bar{J}}_{22}-\bar{\bar{J}}_{12}\bar{\bar{J}}_{21}}. \tag{23}
$$

and

$$
\bar{\bar{J}}_{11}=\mathrm{Re}\int_0^{\bar{\bar{\theta}}_0}\bar{\bar{I}}(\theta)d\theta\quad\text{for }\bar{\bar{\theta}}_0<-s^{-1}, \tag{24a}
$$

$$
\bar{\bar{J}}_{21}=\mathrm{Re}\int_0^{\bar{\bar{\theta}}_0}\bar{\bar{I}}(\theta)\theta d\theta\quad\text{for }\bar{\bar{\theta}}_0<-s^{-1}, \tag{24b}
$$

$$
\bar{\bar{J}}_{12}=\mathrm{Re}\int_0^{\bar{\bar{\theta}}_0}\bar{\bar{I}}(\theta)d\theta\quad\text{for }\bar{\bar{\theta}}_0>(rs)^{-1}, \tag{24c}
$$

$$
\bar{\bar{J}}_{22}=\mathrm{Re}\int_0^{\bar{\bar{\theta}}_0}\bar{\bar{I}}(\theta)\theta d\theta\quad\text{for }\bar{\bar{\theta}}_0>(rs)^{-1}, \tag{24d}
$$

with

$$
\bar{\bar{I}}(\theta)=\frac{-i\mu b^2R(\theta^2)}{(a^{-2}-\theta^2)^{1/2}(s^{-1}+\theta)^{3/2}[(rs)^{-1}-\theta]^{3/2}}.
$$

3.4 Complete Solution to the Original Problem. If the solution of case A problem is expressed by $q\phi(x,y,t)$, then the solutions of the cases B and C are denoted by $-q\phi(x-x_0,y,t-\Delta t)$ and $-q\phi(x+x_0,y,t-\Delta t)$, respectively. The relation between the magnitude of the uniformly distributed load q and the concentrated load F is $q=F/2x_0$, for $x_0(=s\Delta t)\to 0$. Therefore, the complete solution of the original problem is

$$
\lim_{\Delta t\to 0}\frac{F\{\phi(x,y,t)-\phi(x-s\Delta t,y,t-\Delta t)-\phi(x+s\Delta t,y,t-\Delta t)\}}{2s\Delta t}. \tag{25}
$$

Hence, summing up the solutions of cases A, B, and C and taking $x_0\to 0$ (or $\Delta t\to 0$) yield the complete solution. To verify the results, Figs. 4 and 5 display the displacement u_y and the stress σ_y on the surface $y=0$ of the original problem.

Notably, the displacements of cases A, B, and C are homogeneous functions of degree one in space and time, and stresses are homogeneous functions of degree zero. After the derivative procedure in Eq. (25), the displacement of the original problem is the homogenous function of degree zero and the stresses are degree -1. Therefore, for the problem of the propagating central crack subjected to the concentrated load, the displacement and stress fields are homogeneous functions of degree zero and minus one, respectively.

4 Dynamic Stress Intensity Factor

Equations (11), (19), and (22) suggest that the integrands are of the type $G(\theta)/(s^{-2}-\theta^2)^{3/2}$ and have high singularities with a power of order 3/2 near the crack tip. The asymptotic solutions near the crack tip can be determined by expanding asymptotically about the point $\theta=s^{-1}$ and then performing the integration. To illustrate the procedure, consider the y component of the stress of the case A problem ${}_{\mathrm{A}}\sigma_y$, which can be rewritten as

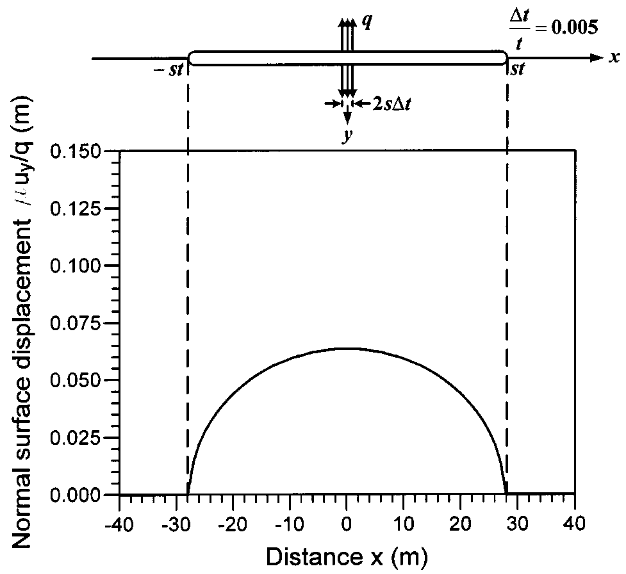


Fig. 4 The normal surface displacement on the x axis for $a=6326$ m/s, $b=3463$ m/s, $s=2770.4$ m/s, $t=0.01$ s

$$
{}_{\mathrm{A}}\sigma_y=\mu b^2A_1\,\mathrm{Im}\left\{\int_0^{\theta_1}\frac{{}_{A}G_1(\theta)}{(s^{-2}-\theta^2)^{3/2}}d\theta+\int_0^{\theta_2}\frac{{}_{\mathrm{A}}G_2(\theta)}{(s^{-2}-\theta^2)^{3/2}}d\theta\right\},
$$

where

$$
{}_{\mathrm{A}}G_1(\theta)=\frac{(b^{-2}-2\theta^2)^2}{\sqrt{a^{-2}-\theta^2}},\quad {}_{\mathrm{A}}G_2(\theta)=4\theta^2\sqrt{b^{-2}-\theta^2}.
$$

By expanding ${}_{\mathrm{A}}G_1(\theta)$ and ${}_{\mathrm{A}}G_2(\theta)$ about $\theta=s^{-1}$ and then taking the dominant term in the series expansion, ${}_{\mathrm{A}}\sigma_y$ near the crack tip is approximated by

$$
{}_{\mathrm{A}}\sigma_y\approx\mu b^2A_1\,\mathrm{Im}\left\{\frac{s^2\theta_1(b^{-2}-2s^{-2})^2}{\sqrt{a^{-2}-s^{-2}}\sqrt{s^{-2}-\theta_1^2}}+\frac{4\theta_2\sqrt{b^{-2}-s^{-2}}}{\sqrt{s^{-2}-\theta_2^2}}\right\}.
$$

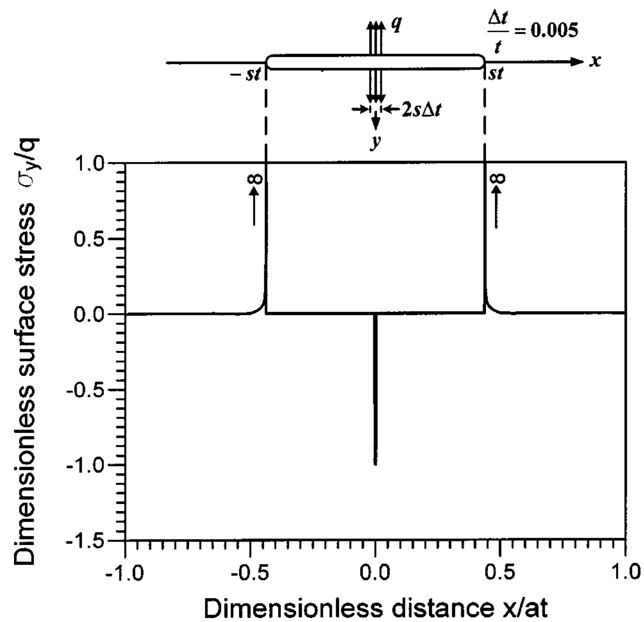


Fig. 5 The normal surface stress σ_y on the x axis for $a=6326$ m/s, $b=3463$ m/s, $s=2770.4$ m/s, $t=0.01$ s

If θ_1 and θ_2 are expressed in terms of t, y, and $\varepsilon = x - st$, then

$$
{}_{A}\sigma_y \approx \frac{\mu b^2 s^2 A_1}{\sqrt{2}} \mathrm{Im}\left\{ \frac{(b^{-2}-2s^{-2})^2}{\sqrt{a^{-2}-s^{-2}}\sqrt{\frac{\varepsilon}{st}-\frac{iy}{st}\sqrt{1-s^2a^{-2}}}} + \frac{4s^{-2}\sqrt{b^{-2}-s^{-2}}}{\sqrt{\frac{\varepsilon}{st}-\frac{iy}{st}\sqrt{1-s^2b^{-2}}}} \right\}. \tag{26}
$$

Similarly, the asymptotic solutions of cases B and C can be found.

The dynamic stress intensity factor (DSIF), which reveals the stress behavior near the crack tip, is an important parameter in fracture mechanics. Therefore, evaluating the DSIF is of relevant concern. According to Eq. (26), the stresses near the crack tips have square root singularity. Therefore, the DSIF of mode I running crack is defined by

$$
K_I(t) = \lim_{\varepsilon \to 0} \{ \sqrt{2\pi\varepsilon}\,\sigma_y(\varepsilon, y=0, t) \}, \tag{27}
$$

where $\sigma_y(\varepsilon, y=0, t)$ is the y-component stress of the original problem near the crack tip. Therefore, summing up the asymptotic solutions of the y-component stresses of case A, B, and C and setting $y=0$ yield

$$
\begin{aligned}
\sigma_y \approx{}& \frac{\mu b^2 s^2 R(s^{-2}) A_1 \sqrt{st}}{\sqrt{2}\sqrt{s^{-2}-a^{-2}}\sqrt{\varepsilon}} + \frac{\mu b^2 (2s) R(s^{-2})(B_1 s + B_2)}{\sqrt{s^{-2}-a^{-2}}(1+r^{-1})^{3/2}} \\
&\times \left(\frac{\sqrt{st_0}}{\sqrt{\bar{\varepsilon}}} - 1 \right) + \frac{\mu b^2 (2s^{3/2}) R(r^{-2}s^{-2})(C_1 + r^{-1}s^{-1}C_2)}{\sqrt{r^{-2}s^{-2}-a^{-2}}(1+r^{-1})^{3/2}} \\
&\times \left(\frac{\sqrt{(rs)^2 t_0}}{\sqrt{\bar{\bar{\varepsilon}}}} - \sqrt{rs} \right),
\end{aligned} \tag{28}
$$

where $\bar{\varepsilon} = \bar{x} - st_0$ and $\bar{\bar{\varepsilon}} = \bar{\bar{x}} - rst_0$. By substituting Eq. (28) into Eq. (27), the DSIF for the original problem can be expressed as

$$
\begin{aligned}
K_I(t) ={}& \frac{\pi\mu b^2 s^2 R(s^{-2})}{2J\sqrt{s^{-2}-a^{-2}}} \frac{t}{\Delta t} \frac{F}{\sqrt{\pi st}} \\
&+ \frac{\sqrt{2}\pi\mu b^2 s R(s^{-2})}{\sqrt{s^{-2}-a^{-2}}(1+r^{-1})^{3/2}} \frac{t}{\Delta t} \sqrt{1-\frac{\Delta t}{t}} \left(\frac{s\bar{J}_{22}}{J_{11}J_{22}-J_{12}J_{21}} \right. \\
&\left. - \frac{\bar{J}_{12}}{J_{11}J_{22}-J_{12}J_{21}} \right) \frac{F}{\sqrt{\pi st}} \\
&+ \frac{\sqrt{2}\pi\mu b^2 s R(r^{-2}s^{-2})}{\sqrt{r^{-2}s^{-2}-a^{-2}}(1+r^{-1})^{3/2}} \frac{t}{\Delta t} \sqrt{1-\frac{\Delta t}{t}} \\
&\times \left(\frac{rs\bar{\bar{J}}_{22}}{\bar{\bar{J}}_{11}\bar{\bar{J}}_{22}-\bar{\bar{J}}_{12}\bar{\bar{J}}_{21}} - \frac{\bar{\bar{J}}_{12}}{\bar{\bar{J}}_{11}\bar{\bar{J}}_{22}-\bar{\bar{J}}_{12}\bar{\bar{J}}_{21}} \right) \frac{F}{\sqrt{\pi st}}.
\end{aligned} \tag{29}
$$

Notably, for the static central crack with length st subjected to a concentrated load F, the static stress intensity factor (SIF), which is denoted as K_0, is equal to $F/\sqrt{\pi st}$. Therefore, the DSIF of a concentrated load is the product of so-called universal function [8] and the static SIF. It is noted that the term $t/\Delta t = st/(s\Delta t) = st/x_0$ denotes the ratio of the crack length and the applied load range. Moreover, if we let x_0 equal $v_L t$, where v_L is the velocity of expanding uniform load, then $t/\Delta t = s/v_L$ can be regarded as the ratio of the crack-tip speed and the expanding load speed. Consequently, the DSIF in Eq. (29) can be directly applied to the problem of a central crack subjected to uniform distributed load expanding on a portion of the running crack surfaces.

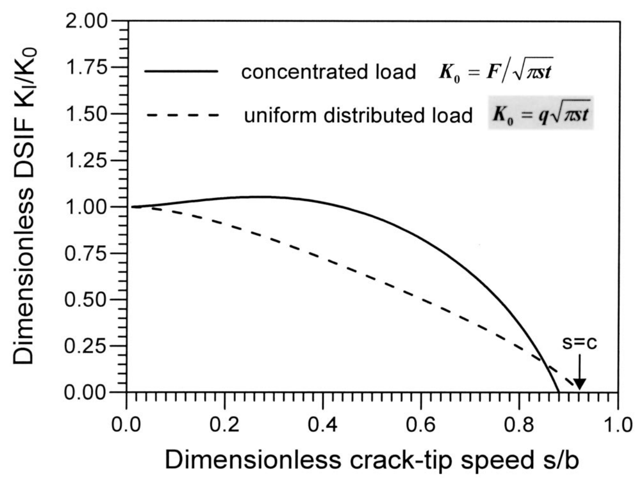


Fig. 6 Comparison of normalized DSIF for a central running crack subjected a concentrated load and uniform distributed load

4.1 Concentrated Load. The DSIF of a central crack subjected to a concentrated load can be evaluated from Eq. (29) by taking a very small Δt. Figure 6 plots the variation of dimensionless DSIF $K_I(t)/K_0$ virus dimensionless crack-tip speed s/b for $\Delta t = 0.0001t$. This figure indicates that when the crack-tip speed s is very slow, the value of $K_I(t)/K_0 = 1$. This phenomenon is owing to the fact that the inertia effect can be neglected when the crack-tip speed is very slow. When the crack-tip velocity s approaches the Rayleigh-wave speed c, the Rayleigh function in the first and second terms in Eq. (29) is zero, i.e., $R(s=c)=0$; but the Rayleigh function in the third term of Eq. (29), $R(rs>c)$, is negative and very small. Therefore, the DSIF for the crack-tip speed approaching the Rayleigh-wave speed is close to zero. Moreover, the solid line in Fig. 6 denotes the DSIF of the central crack subjected to a concentrated load, while the dashed line represents the DSIF of the central crack subjected to uniformly distributed load on the entire crack surfaces. According to Fig. 6, although the DSIF due to a uniformly distributed load decays rapidly with an increase of the crack-tip speed, the DSIF under a concentrated load obviously decreases only when $s > 0.4b$. This occurrence is owing to the fact that when a concentrated load acts on the center of the crack surfaces, there is a distance between the

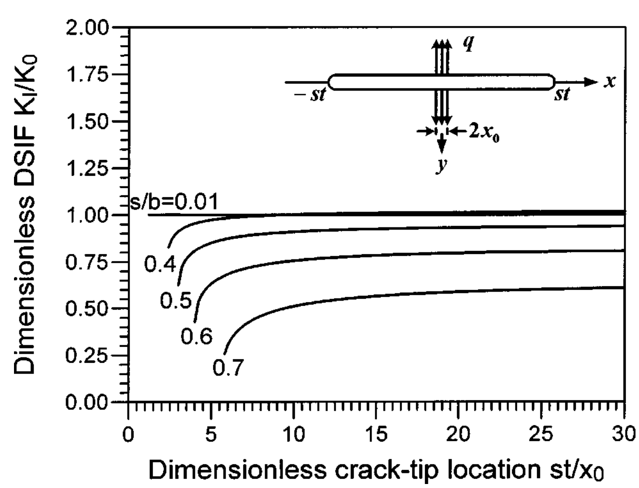


Fig. 7 Normalized DSIF versus crack-tip location for a mode I running crack subjected to a concentrated load with a variation in crack-tip speed

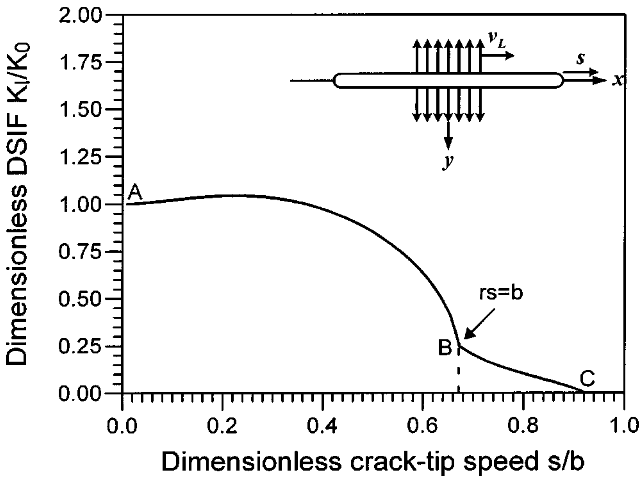


Fig. 8 Normalized DSIF versus crack-tip speed for a mode I running crack subjected to a uniformly portion load with expanding speed $v_L/s=0.2$

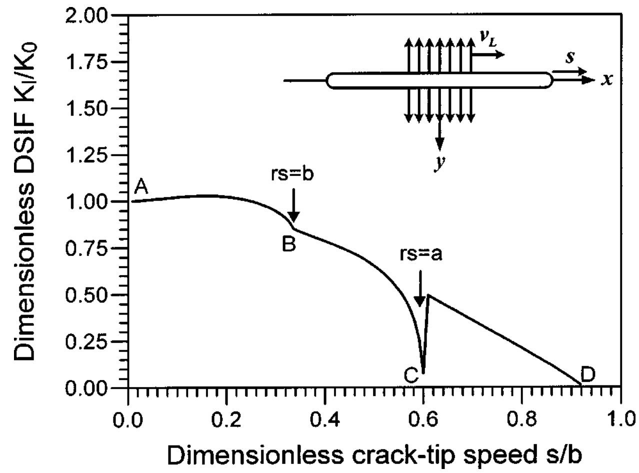


Fig. 10 Normalized DSIF versus crack-tip speed for a mode I running crack subjected to a uniformly portion load with expanding speed $v_L/s=0.5$

applied load and the extending crack tip. Hence, the crack-tip speed must be fast enough to exhibit the inertia effect. However, when the uniformly distributed load is applied on the entire crack surfaces and expands with crack tip (even if the crack-tip speed is slow), there always exists the expanding load on the crack tip, and the dynamic behavior becomes obvious as well. Hence, DSIF due to a uniformly distributed load decays rapidly.

Next, the DSIF at crack tip varies with different time intervals, when the central crack is subjected to a concentrated load. This can be simply achieved by fixing Δt and changing the time t. Figure 7 summarizes those results. According to this figure, when the crack-tip speed is very slow, the inertia effect cannot be ignored if the distance between the applied load and the crack tip is small; hence, the value of $K_I(t)/K_0$ is smaller than one. However, the static state occurs immediately when the crack propagates some more distance. When the crack-tip speed increases, the crack must propagate some more for the static state condition to occur. However, when the crack-tip speed $s>0.4b$, although the crack tip is far away from the applied load, the value of $K_I(t)/K_0$ is less than one. Restated, the stresses near the crack tip is in a dynamic state, which is caused by the inertia effect of the crack propagation.

4.2 Expanding the Uniformly Distributed Load. Assume that the applied load is not the concentrated load. Instead, it is the uniformly distributed load expanding symmetrically on a portion of the crack surfaces with the speed v_L. Consequently, the stress waves are emitted from the ends of the expanding load, propagate, and interact with the stress field near the cracks. The DSIF due to expanding uniform load can be evaluated by setting $\Delta t/t=v_L/s$ in Eq. (29), as explained before. Figure 8 illustrates DSIF virus the crack-tip speed for $v_L/s=0.2$. According to this figure, there is a jump at $s/b=2/3$. This phenomenon is due to the fact that the waves emitted from the left end of the applied load interact with the rightmost crack tip, as shown in Fig. 9. When the rightmost crack-tip speed is smaller than the speed of the S wave, called the subsonic case, the S wave from the left end of the applied load does not interact with the rightmost crack tip. The AB line of Fig. 8 displays the subsonic results. However, when the rightmost crack-tip speed equals the S wave speed, this finding corresponds to the transition point (or the point B) in Fig. 8. Consider a situation in which the rightmost crack-tip speed is larger than the S wave speed but less than the P-wave speed, commonly referred to as a transonic case. The BC line in Fig. 8 represents the DSIF of the transonic case. Indeed, the phenomenon can be explained from Eq. (29). The third term of Eq. (29) is attributed to the case C problem. After the origin is shifted to the left end of the applied load and the time delay is considered, the rightmost crack tip of the case C problem can be regarded as propagating with speed rs. Therefore, the Rayleigh function in the third term of Eq. (29) is a function of rs. When $rs<b$, which is the subsonic case, this finding corresponds to the line AB. When $b<rs<a$ (the transonic case), the Rayleigh function $R(r^{-2}s^{-2})$ has different values for $rs<b$ and $b<rs<a$ due to the branch cuts and, hence, there exists a jump between lines AB and BC in Fig. 8. For this ex-

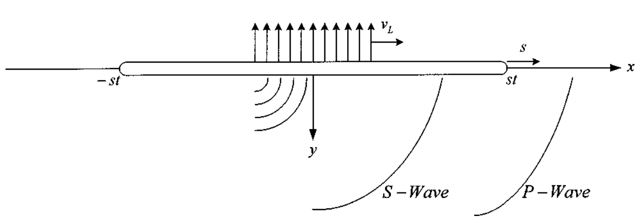


Fig. 9 Geometrical configuration of the disturbance of P and S waves

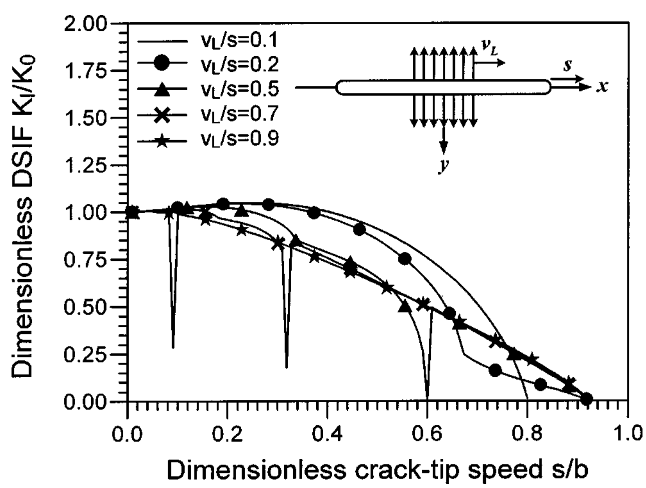


Fig. 11 Normalized DSIF versus crack-tip speed for a mode I running crack subjected to a uniformly portion load with expanding speed $v_L/s=0.1$, 0.2, 0.5, 0.7, 0.9

ample in which $v_L/s=0.2$, the value of r equals 1.5. Consequently, the point $rs=1.5s=b$ (or $s/b=2/3$) is the transition point of subsonic and transonic cases.

Figure 10 plots the DSIF for $v_L/s=0.5$. The lines of AB and BC denote the results of the subsonic and transonic cases, respectively. The line CD represents the solution of supersonic case in which the rightmost crack-tip speed rs is larger than the P-wave speed a, i.e., $a<rs$. For the supersonic case, although the P or S waves do not affect the stresses near the crack tip, the wave fronts of P or S wave interfere with it.

To understand how the speed of expanding load affects the DSIF, the value of v_L/s are taken as 0.1, 0.2, 0.5, 0.7, 0.9. Figure 11 depicts the DSIFs for different v_L/s.

5 Conclusion

This study applies the method of self-similar potentials to a propagating crack subjected to concentrated load leads to the following conclusions:

1. The SSP method in which the origin shift and time delay are considered in conjunction with the superposition can be applied to the non-self-similar problems. This method is not only limited to the problems due to a concentrated load but can also be applied to the problem of a portion of the crack surfaces subjected to a uniformly distributed load.

2. The DSIFs for the problems subjected to a concentrated load and a uniformly distributed load differ from each other. If the concentrated load acts on the crack surface, the dynamic effects become obvious when the crack-tip speed $s>0.4b$; however, if the crack surfaces are subjected to a uniform load, the inertia effect cannot be neglected even though the crack-tip speed is very slow.

3. If a uniform load expands along a portion of the crack surfaces as the crack tips propagate, the ends of the extending load emit waves which will interfere with the stress field of the crack tips. According the different crack-tip speeds, the stress field near the crack tip has the so called subsonic, transonic, and supersonic cases.

Acknowledgment

The authors would like to thank the National Science Council of Taiwan for financially supporting this research under Contract No. NSC 88-2211-E011-008.

References

[1] Achenbach, J. D., 1973, *Wave Propagation in Elastic Solids*, Elsevier, New York.

[2] Smirov, V. I., 1964, *A Course of Higher Mathematics*, Addison-Wesley, London.

[3] Thompson, J. C., and Robinson, A. R., 1969, ''Exact Solution of Some Dynamic Problems of Indentation and Transient Loadings of an Elastic Half Space,'' SRS 350, University of Illinois at Champaign-Urbana.

[4] Chung, Y. L., 1995, ''The Stress Singularity of Mode-I Crack Propagating With Transonic Speed,'' Eng. Fract. Mech., **52**, pp. 977–985.

[5] Chung, Y. L., 1992, ''The Transient Problem of a Mode-III Interface Crack,'' Eng. Fract. Mech., **41**, pp. 321–330.

[6] Johnson, J. J., and Robinson, A. R., 1972, ''Wave Propagation in Half Space Due to an Interior Point Load Parallel to the Surface,'' SRS 388, University of Illinois at Champaign-Urbana.

[7] Chung, Y. L., 1997, ''Analysis of a Running Crack Stopping Suddenly—Solution for Small Time After Stop,'' Chin. J. Mech., **13**, pp. 175–183.

[8] Freund, L. B., 1972, ''Crack Propagation in an Elastic Solid Subjected to General Loading—II Non-Uniform Rate Extension,'' J. Mech. Phys. Solids, **20**, pp. 141–152.

[9] Freund, L. B., 1973, ''Crack Propagation in an Elastic Solid Subjected to General Loading—III Stress Wave Loading,'' J. Mech. Phys. Solids, **21**, pp. 46–61.

[10] Freund, L. B., 1974, ''Crack Propagation in an Elastic Solid Subjected to General Loading—IV Obliquely Incident Stress Pulse,'' J. Mech. Phys. Solids, **22**, pp. 137–146.

[11] Eringen, A. C., and Suhubi, E. S., 1975, *Elastodynamics*, Vol. II, Academic Press, New York.

[12] Freund, L. B., 1974, ''The Stress Intensity Factor Due to Normal Impact Loading of the Faces of a Crack,'' Int. J. Eng. Sci., **12**, pp. 179–189.

[13] Kostrov, B. V., 1964, ''The Axisymmetric Problem of Propagation of a Tension Crack,'' J. Appl. Math. Mech., **28**, pp. 793–803.

[14] Kostrov, B. V., 1964, ''Self-Similar Problems of Propagation of Shear Cracks,'' J. Appl. Math. Mech., **28**, pp. 1077–1087.

[15] Chung, Y. L., 1991, ''The Transient Solutions of Mode-I, Mode-II, Mode-III Cracks Problems,'' J. Chin. Inst. Civil Hydraulic Eng., **3**, pp. 511–61.

[16] Churchill, R. V., 1984, *Complex Variables and Applications*, McGraw-Hill, New York.

[17] Cherepanov, G. P., 1974, ''Some Dynamic Problems of the Theory of Elasticity—A Review,'' Int. J. Eng. Sci., **12**, pp. 665–690.

Y.-S. Wang[1]
Professor
Institute of Engineering Mechanics,
Northern Jiaotong University,
Beijing, 100044,
People's Republic of China
e-mail: yswang@center.njtu.edu.cn

G.-Y. Huang
Ph.D Candidate
Institute of Engineering Mechanics,
Northern Jiaotong University,
Beijing, 100044,
People's Republic of China

D. Dross
Professor
Institute of Mechanics,
Technical University of Darmstadt,
Hochschulstr 1, D-64289,
Darmstadt, Germany

On the Mechanical Modeling of Functionally Graded Interfacial Zone With a Griffith Crack: Anti-Plane Deformation

An analytical model is developed for a functionally graded interfacial zone between two dissimilar elastic solids. Based on the fact that an arbitrary curve can be approached by a continuous broken line, the interfacial zone with material properties varying continuously in an arbitrary manner is modeled as a multilayered medium with the elastic modulus varying linearly in each sublayer and continuous on the interfaces between sublayers. With this new multilayered model, we analyze the problem of a Griffith crack in the interfacial zone. The transfer matrix method and Fourier integral transform technique are used to reduce the mixed boundary-value problem to a Cauchy singular integral equation. The stress intensity factors are calculated. The paper compares the new model to other models and discusses its advantages. [DOI: 10.1115/1.1598476]

1 Introduction

In engineering composites, interfaces usually act as sources of failures because of the discrepancies in mechanical and thermal properties between the component materials. One way to reduce the apparent property mismatch between different materials is to introduce intentionally interfacial zones with continuously varying mechanical properties known as functionally graded materials (FGM's). The absence of sharp interfaces in bonded materials with graded interfacial properties alleviates the tendency toward failure. For the past decade, the fracture analysis of the FGM's has been a key topic in solid mechanics. Erdogan in his paper [1] discussed the problem of crack growth in FGM's due to fatigue, creep and stress corrosion cracking, and fracture instability. He concluded that use of FGM as interfacial zones would reduce the magnitude of the residual and thermal stresses and therefore greatly improve the bonding stress. Another merit of the FGM interfacial zone is that it eliminates the crack tip stress oscillations for the classical interface crack model. It has been found that the crack tip singular field in FGM's has the same form as that in homogeneous media [2,3]. Thus many important parameters such as the stress intensity factor (SIF), energy release relate, crack opening displacement (COD), etc., which were developed in linear fracture mechanics for homogeneous materials can be applied directly to nonhomogeneous materials or FGM's. However, it is a nontrivial task to calculate these parameters for FGM's with arbitrarily varying properties because the associate elastic boundary value problem (BVP) is difficult to solve. So far, several ways have been developed to solve this problem, which we summarize as follows:

(1) Assume that the material properties of FGM's vary in a prescribed manner such that the associate BVP can be solved analytically. Two models have been presented: FGM's with properties varying in exponential manner and in power manner. The former one was fully developed by many investigators, especially by Erdogan and co-workers (therefore we refer to it as Erdogan's model). Here we do not intend to list all papers on this model, but only remind of some works about fracture analysis of FGM interfacial zones interested in this paper [4–14]. This model was also applied in analysis of dynamic fracture [15], thermal stress [16,17], viscoelastic fracture [18], etc. The second model is generally used in mode III fracture analysis [19–22].

(2) Model FGM's as piece-wise multilayered media (we refer to it as the PWML model) [23,24]. That is, divide FGM's into multiple layers in the gradient direction. The material properties in each layer are assumed constant. Indeed, this model was used early in modeling nonhomogeneous media such as earth [25].

(3) The perturbation approach may be applied for FGM's with properties varying slightly [26,27].

The first way is limited by the fact that the functions describing variation of material properties and suitable to find the analytical BVP solution are scarce. The second way cannot eliminate discontinuity of the material properties although it can reduce this discontinuity. Another problem in the second way is that no one has checked its convergence—how many layers should FGM's be divided into for a given manner in which the material properties vary? In this paper, we develop a new analytical model for the FGM's with continuously varying properties and solve the problem of a crack in a functionally graded interfacial zone between two dissimilar materials subjected to a static antiplane shearing load. The method is based on the fact that an arbitrary curve can be approached by a continuous broken line. Therefore we model the FGM interfacial zone as a multilayered medium with the linearly varying elastic properties in each sublayer and continuous on the subinterfaces. The Fourier integral transform technique and singular integral equation method are employed to solve the mixed boundary-value problem. The stress intensity factors are presented. We discuss the convergence of the new model, and compare it with Erdogan's model and the PWML model.

2 Problem Formulation

2.1 A New Multilayered Model for a Functionally Graded Interfacial Zone. Consider a Griffith crack on the interface between two bonded dissimilar elastic homogeneous half spaces, with the shear moduli μ, μ^*, which are loaded by remote antiplane shearing traction. Here the "interface" is not the "math-

[1]Corresponding author.
Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 19, 2001; final revision, Apr. 2, 2003. Associate Editor: B. M. Moran. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

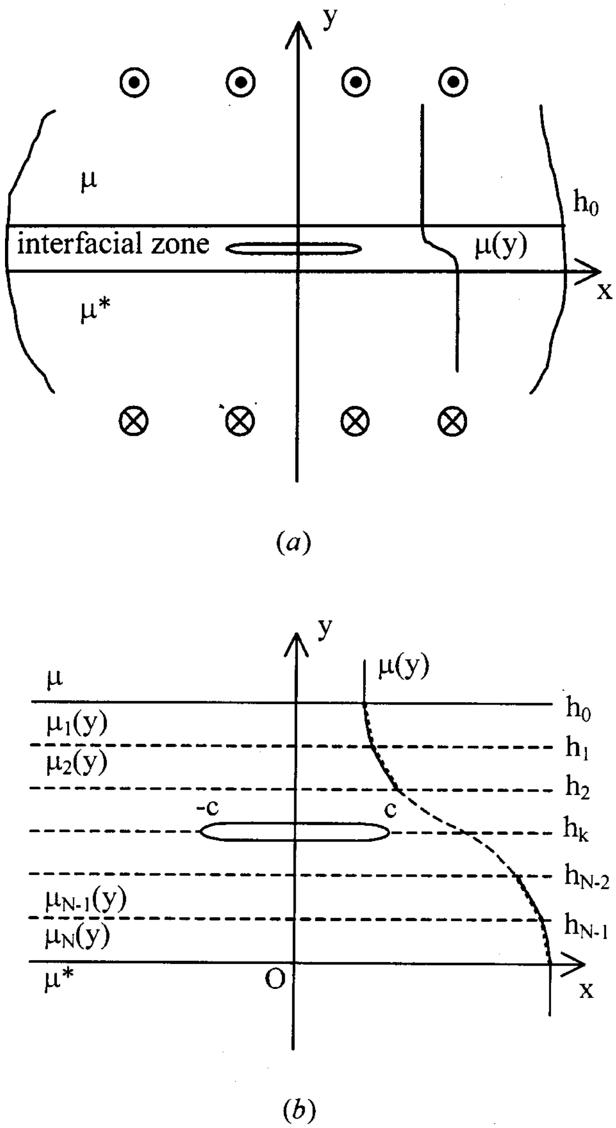


Fig. 1 Two dissimilar half spaces bonded through a FGM interfacial zone with a Griffith crack (*a*) and the new multilayered model of the FGM interface layer (*b*)

ematical interface'' with zero thickness, but an interface layer with finite thickness h_0, as shown in Fig. 1*a*. The crack is located in the interfacial region. This interface layer is made of the FGM's with shear modulus $\mu(y)$ continuously varying from μ to μ^*. Generally, $\mu(y)$ may be of an arbitrary form. However, considering the fact that an arbitrary curve can be approached by a continuous broken line, we develop a new multilayered model as shown in Fig. 1*b*. The interfacial zone $(h_0,0)$ is divided into N sublayers (h_{j-1},h_j) $(j=1,2,\ldots,N;h_N=0)$. The crack is on the kth interface (k may be any integer number from 0 to N). The shear modulus in each sublayer varies linearly with the form:

$$\mu(y)\approx\mu_j(y)=\overline{\mu_j}\cdot(a_j+b_jy),\quad h_j\leqslant y\leqslant h_{j-1},\quad j=1,2,\ldots,N,\tag{1}$$

where $\overline{\mu_j}$ is equal to the real values on the interfaces $y=h_j$, i.e., $\overline{\mu_j}=\mu_j(h_j)=\mu(h_j)$, which leads to

$$a_j=\frac{h_{j-1}-h_j\overline{\mu_{j-1}}/\overline{\mu_j}}{h_{j-1}-h_j},\quad b_j=\frac{\overline{\mu_{j-1}}/\overline{\mu_j}-1}{h_{j-1}-h_j}.\tag{2}$$

2.2 Transfer Matrix and Dual Integral Equations. The only nonzero displacement component is along the z axis, which, in each sublayer, satisfies the following wave equation:

$$\mu_j(y)\frac{\partial^2 w_j}{\partial x^2}+\frac{\partial}{\partial y}\left(\mu_j(y)\frac{\partial w_j}{\partial y}\right)=0,\quad j=1,2,\ldots,N.\tag{3}$$

Substituting Eqs. (1) and (2) into the above equation and applying Fourier integral transform with respect to x, we obtain

$$\frac{\partial^2\tilde{w}_j}{\partial y^2}+\frac{b_j}{a_j+b_jy}\frac{\partial\tilde{w}_j}{\partial y}-s^2\tilde{w}_j=0,\tag{4}$$

where ''~'' indicates the Fourier transform. The solution of Eq. (4) can be written as

$$\tilde{w}_j=C_{1j}(s)K_0(|s|\overline{y_j})+C_{2j}(s)I_0(|s|\overline{y_j}),\tag{5}$$

where C_{1j} and C_{2j} are unknown coefficients; $K_0(\ )$ and $I_0(\ )$ are modified Bessel functions; and $\overline{y_j}=b_j^{-1}(a_j+b_jy)$. Equation (5), together with the Fourier transform of stress component $\tilde{\tau}_{yzj}$, can be written as the matrix form:

$$\{S_j\}=[T_j(y)]\{C_j\},\quad j=1,2,\ldots,N,\tag{6}$$

where

$$\{S_j\}=[\tilde{w}_j,\tilde{\tau}_{yzj}]^T,\quad\{C_j\}=[C_{1j},C_{2j}]^T,\tag{7}$$

$$[T_j(y)]=\begin{bmatrix}K_0(|s|\overline{y_j}) & I_0(|s|\overline{y_j})\\ \mu_j(y)|s|K_0'(|s|\overline{y_j}) & \mu_j(y)|s|I_0'(|s|\overline{y_j})\end{bmatrix},\tag{8}$$

where the prime indicates the differentiation.

The solutions associated with the homogeneous half spaces are given by

$$\{S_0\}=[T_0(y)]\{B\}C_{10},\quad\{S_{N+1}\}=[T_{N+1}(y)]\{X\}C_{2,N+1},\tag{9}$$

where $\{B\}=\{1,0\}^T$, $\{X\}=\{0,1\}^T$, and

$$[T_0(y)]=\begin{bmatrix}e^{-|s|y} & e^{|s|y}\\ -\mu|s|e^{-|s|y} & \mu|s|e^{|s|y}\end{bmatrix},$$

$$[T_{N+1}(y)]=\begin{bmatrix}e^{-|s|y} & e^{|s|y}\\ -\mu^*|s|e^{-|s|y} & \mu^*|s|e^{|s|y}\end{bmatrix}.$$

The subscripts ''0'' and ''$N+1$'' correspond, respectively, to the upper and lower half spaces.

The shear stress and displacement are continuous on the interfaces except that the displacement involves a jump along the crack face on kth interface which is denoted as Δw_k. In the transformed domain, we have

$$\{S_j\}-\{S_{j+1}\}=\{\Delta S_k\}\delta_{kj},\quad y=h_j,\quad j=0,1,2\ldots,N,\tag{10}$$

where $\{\Delta S_k\}=\{\Delta\tilde{w}_k,0\}^T$, and δ_{kj} is Kronecker delta. The above equation is a recurrence relation which, on substitution of Eqs. (8) and (9), may yield the expression of $\{C_j\}$ in terms of $\{\Delta S_k\}$,

$$\{C_j\}=([\overline{L_{jk}}]+[\overline{K_{jk}}]H(j-k-1))\{\Delta S_k\},\quad j=1,2,\ldots,N,\tag{11}$$

where $H(\ )$ is the Heaviside function, and

$$[W_{j+1}]=[T_j(h_j)]^{-1}[T_{j+1}(h_j)];$$

$$[\overline{W}]=\{B\}[1,0]-[\overline{W_{N+1}}]\{X\}[1,0];$$

$$[L_k]=[\overline{W_k}][T_k(h_k)]^{-1};\quad[\overline{W_j}]=[W_2]\cdots[W_j],\quad j>1;$$

$$[\overline{W_1}]=[\mathrm{I}];$$

$$[\overline{L_{jk}}]=[\overline{W_j}]^{-1}[\overline{E_k}];\quad[\overline{K_{jk}}]=-[\overline{W_j}]^{-1}[L_k];$$

$$[\overline{E_k}]=\{B\}[1,0][\overline{W}]^{-1}[L_k].$$

Substituting Eq. (11) into Eq. (8) and taking the inverse Fourier transform, we have

$$\{w_j,\tau_{yzj}\}^T=\frac{1}{2\pi}\int_{-\infty}^{\infty}[M_{jk}]\{\Delta S_k\}e^{-isx}ds,\quad j=1,2,\ldots,N, \tag{12}$$

where we have denoted

$$[M_{jk}]=[T_j(y)]([\overline{L}_{jk}]+[\overline{K}_{jk}]H(j-k-1)), \tag{13}$$

which is nothing but the transfer matrix of the multiple layered medium with an interface crack. Writing this matrix as

$$[M_{jk}]=\begin{bmatrix} * & * \\ m_{jk} & * \end{bmatrix}, \tag{14}$$

then we have

$$\tau_{yzj}=\frac{1}{2\pi}\int_{-\infty}^{\infty}m_{jk}(s)\Delta\tilde{w}_k(s)e^{-isx}ds,\quad j=1,2,\ldots,N. \tag{15}$$

Suppose the stress caused by the remote loading in the medium without crack is $\tau_{yz}^L(x,y)$. Then the free traction condition on the crack faces yields

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{m}_k(s)\Delta\tilde{w}_k(s)e^{-isx}ds=-\tau_{yz}^L(x,h_k),\quad |x|<c, \tag{16}$$

where

$$\overline{m}_k=m_{jk}|_{y=h_k}\ \ (j=k\ \text{ or }\ k-1).$$

The single-valued condition for the displacement gives

$$\int_{-\infty}^{\infty}\Delta\tilde{w}_k(s)e^{-isx}ds=0,\quad |x|>c. \tag{17}$$

Equations (16) and (17) are the dual integral equations of the problem.

2.3 Cauchy Singular Equation. Introduce the dislocation density function of the crack,

$$\phi_k(x)=\frac{\partial}{\partial x}[\Delta w_k(x)],\quad |x|<c. \tag{18}$$

It is not difficult to prove, by considering the differential properties of the Fourier transform, that

$$\Delta\overline{w}_k=is^{-1}\int_{-c}^{c}\phi_k(u)e^{isu}du, \tag{19}$$

which, when substituted into Eqs. (16) and (17), yields

$$\frac{i}{2\pi}\int_{-\infty}^{\infty}s^{-1}\overline{m}_k(s)\int_{-c}^{c}\phi_k(u)e^{is(u-x)}duds=-\tau_{yz}^L(x,h_k),\qquad |x|<c, \tag{20}$$

$$\int_{-c}^{c}\phi_k(u)du=0. \tag{21}$$

Considering the asymptotic behavior of $K_0(\ )$ and $I_0(\ )$ for large arguments [28], one may prove

$$\lim_{s\to\pm\infty}s^{-1}\overline{m}_k(s)=\mp\overline{\mu}_k/2. \tag{22}$$

Therefore special care must be taken in interchanging the two integrations in Eq. (20). However, if we consider the following relation:

$$\int_{-\infty}^{\infty}\mathrm{sgn}(s)e^{is(u-x)}ds=\frac{2i}{u-x}, \tag{23}$$

and denote

$$P_k(u,x)=-\frac{1}{\pi}\int_0^{\infty}\left[s^{-1}\overline{m}_k(s)+\frac{1}{2}\overline{\mu}_k\right]\sin[s(u-x)]ds, \tag{24}$$

we can transform Eq. (20) into a Cauchy singular integral equation,

$$\frac{\overline{\mu}_k}{2\pi}\int_{-c}^{c}\frac{\phi_k(u)}{u-x}du+\int_{-c}^{c}\phi_k(u)P_k(u,x)du=-\tau_{yz}^L(x,h_k),\qquad |x|<c. \tag{25}$$

The method of Erdogan and Gupta [29] can be employed to solve Eqs. (25) and (21) numerically. To this end, we set

$$\phi_k(u)=\frac{f(u)}{\sqrt{1-u^2/c^2}}, \tag{26}$$

then Eqs. (25) and (21) reduce to [29]

$$\begin{cases}\dfrac{\overline{\mu}_k}{2M}\displaystyle\sum_{i=1}^{M}\left[\frac{f(c\eta_i)}{\eta_i-\xi_j}+\pi cP_k(c\eta_i,c\xi_j)f(c\eta_i)\right]=-\tau_{yz}^L(c\xi_j,h_k)\\[2ex] \dfrac{\pi}{M}\displaystyle\sum_{i=1}^{M}f(c\eta_i)=0\end{cases}, \tag{27}$$

where $\eta_i=\cos(\pi(2i-1)/2M)$, $\xi_j=\cos(\pi j/M)$, $j=1\ldots M-1$, and M is the number of the discrete points of $f(c\eta)$ in $(-c,c)$.

3 Stress Intensity Factors

In the present paper, special attention is devoted to the stress intensity factors (SIF's) which are defined as

$$K_{\mathrm{III}}^{\pm}=\lim_{x\to\pm c^{\pm}}\sqrt{2|x-c|}\,\tau_{yz}(x,h_k). \tag{28}$$

The principal part of the stress $\tau_{yz}(x,h_k)$ at the crack tips is

$$\tau_{yz}(x,h_k)\approx\frac{\overline{\mu}_k}{2\pi}\int_{-c}^{c}\frac{\phi_k(u)}{u-x}du,\quad |u|>c. \tag{29}$$

Following the procedure of analysis in Ref. [30], we have

$$K_{\mathrm{III}}^{\pm}=\frac{\overline{\mu}_k}{2}\sqrt{c}f(\pm c), \tag{30}$$

from which the SIF's can be computed.

4 Numerical Results And Discussion

Numerical results for some particular examples are presented in this section. As the first example, we consider a FGM interface layer with a Griffith crack subjected to uniform load τ_0 and lying on the interface between the layer and one component (see Fig. 1, but the crack is on the x axis). The shear modulus is assumed varying exponentially,

$$\mu(y)=\mu^*\exp(\beta y), \tag{31}$$

where β is determined by $\mu(h_0)=\mu$. The same problem was studied by Ozturk and Erdogan [6]. Here we use the present model to solve this problem. The interface layer is divided in N sublayers with the same thickness. The convergence of the numerical solution to Eq. (25) was proved by Erdogan and Gupta [29]. However, to ensure enough accuracy of the results and to avoid too much CPU time, we should choose proper values of M for the present particular problem. To this end, we calculate the SIF's for smaller and larger values of μ^*/μ and c/h_0 by choosing different values of M. The interfacial zone is divided into six sublayers, i.e., $N=6$. The results are shown in Table 1 where the SIF's are normalized by $\tau_0\sqrt{c}$. It is seen that smaller M can ensure

Table 1 Normalized SIF's for selected values of *M*, for exponential variation of the shear modulus of the FGM interface layer

	$\mu^*/\mu=1/22$		$\mu^*/\mu=22$	
M	$c/h_0=1$	$c/h_0=10$	$c/h_0=1$	$c/h_0=10$
20	0.814045	0.708540	1.387912	2.688155
30	0.814164	0.731126	1.387802	2.502428
40	0.814179	0.736229	1.387790	2.489127
50	0.814182	0.739543	1.387787	2.474036
60	0.814183	0.739271	1.387786	2.476729

enough accuracy for smaller μ^*/μ and/or c/h_0, while larger M is required for larger μ^*/μ and/or c/h_0. The following calculation is generally carried out by choosing $M=40$ or 60.

Now we have to answer another question—how many sublayers should the interfacial zone be divided into so that enough accurate results can be obtained? To answer this question, we calculate the SIF's for $\mu^*/\mu=1/22$ and 22 with $c/h_0=10$ by selecting different values of N, and list the results in Table 2 ($M=60$ is chosen in calculation). Erdogan's model gives $K_{\mathrm{III}}/\tau_0\sqrt{c}=0.737$ for $\mu^*/\mu=1/22$, and 2.5072 for $\mu^*/\mu=22$ (see Ozturk and Erdogan [6]). It is noted that $N=6$ or 8 can yield results approximating to Erdogan's. As a comparison, we calculate the same examples using the PWML model, the solution of which can be obtained by replacing $K_0(z)$ and $I_0(z)$ with, respectively, $\exp(-z)$ and $\exp(z)$ in above formulation. The interfacial zone is divided in N sublayers with the same thickness. As in Nozaki and Shindo [31], the shear modulus in each sublayer is taken to be the value of the midline of the sublayer, but those in the sublayers adjacent to the crack plane are assigned the real value at the crack plane in order to eliminate the discontinuity of the shear modulus across the crack plane. The results are also listed in Table 2. One may find that the present new model is more precise than the PWML model for the same N. Further comparison between these three models is shown in Fig. 2 which illustrates variations of SIF's with c/h_0 for $\mu^*/\mu=22$. The solid line is for Erdogan's model, the scattered crosses for the present new model with $N=6$, the scattered dots for the PWML model with $N=16$, and the dashed line for the PWML model with $N=6$. The former three are close to each other, while the last one deviates from them as c/h_0 increases. That is to say, a bigger N is necessary for the PWML model than for the present new model to reach the same precision, especially for larger values of c/h_0.

As mentioned before, the present new model allows for arbitrary variation of material properties of the interfacial zone. As an example, we consider a FGM interface layer with the shear modulus varying in the form of

$$\mu(y)=a+b\cos(\pi y/h_0), \tag{32}$$

where $a=(\mu+\mu^*)/2$ and $b=-(\mu-\mu^*)/2$. The crack is assumed on the midline of the interface layer. The normalized SIF's for some selected values of N with $\mu^*/\mu=22$ (or equally 1/22) and $c/h_0=1$ and 10 are listed in Table 3 ($M=60$ is chosen in calculation). The results of the same problem solved by the PWML model are also presented in the table. The detailed comparison is shown in Fig. 3, where the scattered crosses, dots, and the dashed line indicate the same quantities as in Fig. 2. Again we can see that the present new model is more efficient than the PWML model.

Table 2 Normalized SIF's for selected values of *N* with comparison of the present model and the PWML model, for exponential variation of the shear modulus of the FGM interface layer

	$\mu^*/\mu=1/22$		$\mu^*/\mu=22$	
N	Present model	PWML model	Present model	PWML model
2	0.736915	0.750788	2.387198	2.688155
4	0.738881	0.743316	2.461473	2.502428
6	0.739271	0.741420	2.476729	2.427718
8	0.739405	0.740654	2.482122	2.454279
10	0.739466	0.740270	2.484583	2.468313

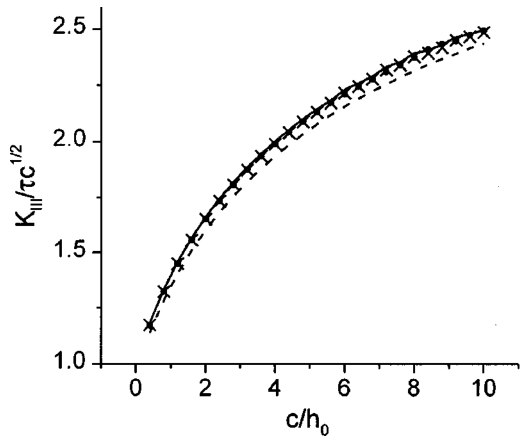


Fig. 2 Variation of SIF's with c/h_0 for $\mu^*/\mu=22$, with comparison between Erdogan's model, the PWML model, and the present new model, for exponential variation of the shear modulus of the FGM interface layer

5 Concluding Remarks

A new multilayered model for fracture analysis of a functionally graded interfacial zone with a Griffith crack under the antiplane shearing load is developed in the present paper. To check the efficiency of the new model, we calculate the SIF's for some

Table 3 Normalized SIF's for selected values of *N* with comparison of the present model and the PWML model, for cosine variation of the shear modulus of the FGM interface layer

	$c/h_0=1$		$c/h_0=10$	
N	Present model	PWML model	Present model	PWML model
2	1.307680	1.153604	2.060753	1.865835
4	1.374017	1.291829	2.114980	2.047207
6	1.388163	1.343830	2.126041	2.094627
8	1.393386	1.366477	2.130123	2.112845
10	1.395866	1.378069	2.132065	2.121659

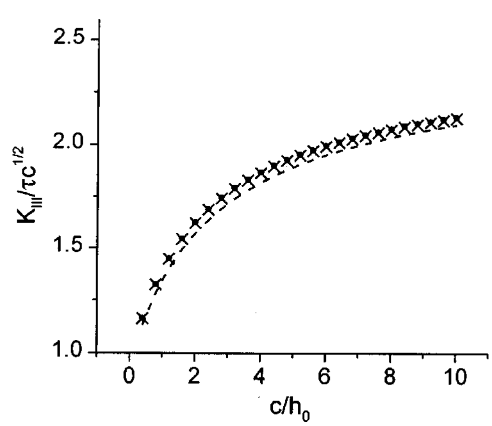


Fig. 3 Variation of SIF's with c/h_0 for $\mu^*/\mu=22$, with comparison between the PWML model and the present new model, for cosine variation of the shear modulus of the FGM interface layer

examples and compare the results with those of Erdogan's model and the PWML model. We have found the following advantages of the new model over the other two models:

(1) The present new model converges faster than the PWML model. For the two examples we considered, the new model converges with six to eight sublayers, while the PWML model requires about twice this number to reach the same precision.

(2) The present new model involves no discontinuities of the material properties, and therefore is expected to be applied in the plane problem of interface cracks to eliminate the stress oscillation at the crack tip. However, for the PWML model, special attention has to be paid to the two sublayers adjacent to the cracks.

(3) Compared with Erdogan's model, the present one allows for arbitrary variation of the material properties.

Acknowledgments

Y.-S. Wang is grateful to the Alexander von Humboldt Foundation for the support to the initiation of the present work.

References

[1] Erdogan, F., 1995, ''Fracture-Mechanics of Functionally Graded Materials,'' Composites Eng., **5**, pp. 753–770.
[2] Jin, Z.-H., and Noda, N., 1994, ''Crack-tip Singular Fields in Nonhomogeneous Materials,'' ASME J. Appl. Mech., **61**, pp. 738–740.
[3] Jin, Z. H., and Batra, R. C., 1996, ''Some Basic Fracture Mechanics Concepts in Functionally Graded Materials,'' J. Mech. Phys. Solids, **44**, pp. 1221–1235.
[4] Delale, F., and Erdogan, F., 1988, ''On the Mechanical Modeling of the Interfacial Region in Bonded Half-Plane,'' ASME J. Appl. Mech., **55**, pp. 317–324.
[5] Erdogan, F., Kaya, A. C., and Joseph, P. F., 1991, ''The Mode-III Crack Problem in Bonded Materials With a Nonhomogeneous Interfacial Zone,'' ASME J. Appl. Mech., **58**, pp. 419–427.
[6] Ozturk, M., and Erdogan, F., 1993, ''Antiplane Shear Crack Problem in Bonded Materials With a Graded Interfacial Zone,'' Int. J. Eng. Sci., **31**, pp. 1641–1657.
[7] Ozturk, M., and Erdogan, F., 1995, ''An Axisymmetrical Crack in Bonded Materials With a Nonhomogeneous Interfacial Zone Under Torsion,'' ASME J. Appl. Mech., **62**, pp. 116–125.
[8] Ozturk, M., and Erdogan, F., 1996, ''Axisymmetric Crack Problem in Bonded Materials With a Graded Interfacial Region,'' Int. J. Solids Struct., **33**, pp. 193–219.
[9] Fildis, H., and Yahsi, O. S., 1996, ''The Axisymmetric Crack Problem in a Non-Homogeneous Interfacial Region Between Homogeneous Half-Spaces,'' Int. J. Fract., **78**, pp. 139–164.
[10] Fildis, H., and Yahsi, O. S., 1997, ''The Mode III Axisymmetric Crack Problem in a Non-Homogeneous Interfacial Region Between Homogeneous Half-Spaces,'' Int. J. Fract., **85**, pp. 35–45.
[11] Choi, H. J., Lee, K. Y., and Jin, T. E., 1998, ''Collinear Cracks in a Layered Half-Plane With a Graded Nonhomogeneous Interfacial Zone-Part I: Mechanical Response,'' Int. J. Fract., **94**, pp. 103–122.
[12] Choi, H. J., Jin, T. E., and Lee, K. Y., 1998, ''Collinear Cracks in a Layered Half-Plane With a Graded Nonhomogeneous Interfacial Zone-Part II: Thermal Shock Response,'' Int. J. Fract., **94**, pp. 123–135.
[13] Shbeeb, N. I., and Binienda, W. K., 1999, ''Analysis of an Interface Crack for a Functionally Graded Strip Sandwiched Between Two Homogeneous Layers of Finite Thickness,'' Eng. Fract. Mech., **64**, pp. 693–720.
[14] Shbeeb, N. I., Binienda, W. K., and Kreider, K., 2000, ''Analysis of the Driving Force for a Generally Oriented Crack in a Functionally Graded Strip Sandwiched Between Two Homogeneous Half Planes,'' Int. J. Fract., **104**, pp. 23–50.
[15] BaBaei, R., and Lukasiewicz, S. S., 1998, ''Fracture in Functionally Gradient Materials Subjected to the Time-Dependent Anti-Plane Shear Load,'' Z. Angew. Math. Mech., **78**, pp. 383–390.
[16] Noda, N., and Jin, Z. H., 1993, ''Thermal Stress Intensity Factors for a Crack in a Strip of a Functionally Gradient Material,'' Int. J. Solids Struct., **30**, pp. 1039–1056.
[17] Erdogan, F., and Wu, B. H., 1996, ''Crack Problems in FGM Layers Under Thermal Stresses,'' J. Therm. Stresses, **19**, pp. 237–265.
[18] Paulino, G. H., and Jin, Z. H., 2001, ''Viscoelastic Functionally Graded Materials Subjected to Antiplane Shear Fracture,'' ASME J. Appl. Mech., **68**, pp. 284–293.
[19] Craster, R. V., and Atkinson, C., 1994, ''Mixed Boundary Value Problems in Non-Homogeneous Elastic Materials,'' Q. J. Math., **47**, pp. 183–206.
[20] Wang, X. Y., Wang, D., and Zou, Z. Z., 1996, ''On the Griffith Crack in a Nonhomogeneous Interlayer of Adjoining Two Different Elastic Materials,'' Int. J. Fract., **79**, pp. R51–R56.
[21] Dhaliwal, R. S., Saxena, H. S., and He, W. H., 1992, ''Stress Intensity Factor for the Cylindrical Interface Crack Between Nonhomogeneous Coaxial Finite Elastic Cylinders,'' Eng. Fract. Mech., **43**, pp. 1039–1051.
[22] Han, X. L., and Duo, W., 1996, ''The Crack Problem of a Fiber-Matrix Composite With a Nonhomogeneous Interfacial Zone Under Torsional Loading. 1. A Cylindrical Crack in the Interfacial Zone,'' Eng. Fract. Mech., **54**, pp. 63–69.
[23] Wang, B. L., Han, J. C., and Du, S. Y., 1999, ''Dynamic Response for Functionally Graded Materials With Penny-Shaped Cracks,'' Acta Mechanica Solida Sin., **12**, pp. 106–113.
[24] Itou, S., and Shima, Y., 1999, ''Stress Intensity Factors Around a Cylindrical Crack in an Interfacial Zone in Composite Materials,'' Int. J. Solids Struct., **36**, pp. 697–709.
[25] Ewing, W. M., Jardetzky, W. S., and Press, F., 1957, *Elastic Waves in Layered Media*, McGraw-Hill, New York.
[26] Gao, H., 1991, ''Fracture Analysis of Non-Homogeneous Materials via a Moduli-Perturbation Approach,'' Int. J. Solids Struct., **27**, pp. 1663–1682.
[27] Erguven, M. E., and Gross, D., 1999, ''On the Penny-Shaped Crack in Inhomogeneous Elastic Materials Under Normal Extension,'' Int. J. Solids Struct., **36**, pp. 1869–1882.
[28] Abramowitz, M., and Stegun, I. A., 1965, *Handbook of Mathematical Functions*, Dover, New York.
[29] Erdogan, F., and Gupta, G. D., 1972, ''On the Numerical Solution of Singular Integral Equations,'' Quart. Appl. Math., **29**, pp. 525–534.
[30] Wang, Y. S., and Wang, D., 1996, ''Scattering of Elastic Waves by a Rigid Cylindrical Inclusion Partially Debonded From its Surrounding Matrix-I. SH Case,'' Int. J. Solids Struct., **33**, pp. 2789–2815.
[31] Nozaki, H., and Shindo, Y., 1998, ''Effect of Interface Layers on Elastic Wave Propagation in a Fiber-Reinforced Metal-Matrix Composite,'' Int. J. Eng. Sci., **36**, pp. 383–394.

C. K. Soh
Professor

Y. Liu
Research Scholar

Y. Yang
Teaching Fellow

Y. Dong
Research Fellow

School of Civil & Environmental Engineering,
Nanyang Technological University,
Singapore, 639798

A Displacement Equivalence-Based Damage Model for Brittle Materials—Part I: Theory

In this paper, a displacement equivalence-based damage model for brittle materials is proposed. A new damage deactivation criterion, which depends on both the stress and strain states of the materials, is adopted. Based on the concept of effective stress, the virtual undamaged configuration is introduced, and the assumption of displacement equivalence is proposed to correlate the damaged and the virtual undamaged configurations. Then, an additional crack-opening-induced normal deformation is introduced, and the three-dimensional (3D) effect of these opened cracks is also considered. The evolution rule of damage is deduced using the Onsager relations, which also ensure that the second law of thermodynamics is satisfied. [DOI: 10.1115/1.1599914]

1 Introduction

Brittle materials such as concrete and ceramics are widely used in engineering structures. Hence, the constitutive models of the brittle materials are continuously studied by researchers. Damage theories, based on thermodynamics and with damage variables introduced as the internal state variables to describe the material defects, have been used by many researchers to model the behaviors of the materials. With much progress achieved in this area [1], the damage-mechanics-based models have been able to accurately account for the microcracking and softening behavior of materials [2].

There are three steps in analyzing an engineering problem using damage mechanics. The first step is to define the damage variables and to find out how the defects affect the materials' macrobehaviors. The second step is to describe the evolution rule of damage variables, which shows how the stress and strain lead to new damage in the materials. These first two steps essentially establish the constitutive relationship of materials with defects, i.e., they established the relations among damage, stress, and strain of the materials. The third step is to use the established constitutive relationship in the structural analysis to derive the result regarding when the macrodefects (cracks) will appear or when the structures will fail.

The definition of the damage includes choosing a proper variable (with or without apparent physical meaning) to describe the defects of the materials, and to find out how these defects affect the macrobehaviors of the materials. The damage variables could be scalar (single or multiple), vector, second-order, fourth-order, or even eighth-order tensor [3,4]. The influences of the defects in the continuum can be studied via three approaches: the effective stress-based methods, the pure phenomenological methods, and the representative-volume-element- (RVE) based methods.

The effective-stress-based methods assumed that there is a virtual undamaged configuration corresponding to the real damaged configuration. The virtual undamaged configuration is usually defined using the concept of the effective area and effective stress, and in this case, the damage variables usually have obvious physical meaning, e.g., a decrease in the effective area to resist load. It is also assumed that there is a certain equivalent relation between these two configurations. Thus, the stress-strain relationship of the damaged materials can be obtained from the corresponding virtual undamaged material via certain preassumed equivalence rules, such as the principle of strain equivalence [5], principle of energy equivalence [6], energy correlation hypothesis [7], and the assumption of displacement equivalence as proposed in this paper. This will be discussed in detail in Sec. 3.

In the pure phenomenological methods, the expression of certain free energy or strain energy is usually provided directly [8], from which the stress-strain relationship is derived. In this case, the damage variables may not have obvious physical meanings. The RVE-based methods are also known as the micromechanical methods, because in these methods, usually an RVE is introduced, and the micromechanical methods are then used to study the average behaviors of the RVE, which are viewed as the macrobehaviors of the continuum. This category includes the microplane method [4], the self-consistent method [9], the Mori-Tanaka method [10], and the generalized geometric definition [11,12].

Generally, the pure phenomenological definitions are consummated in theories, but it is not easy to specify the expression of the free energy. Although the micromechanical definitions could be regarded as the accurate definitions, usually they are too complicated for practical use. As indicated by Krajcinovic [13], a purely micromechanical theory may never replace a properly formulated macrophenomenological theory as a design tool. Compared with the micromechanics definitions, the effective-stress-based definitions are simple in formulation and thus are widely used in engineering analysis. But sometimes they are too simplistic in describing the complicated behaviors of the materials, so more defect-induced effects need to be taken into account. In this paper, the additional crack-opening-induced normal deformation and the three-dimensional (3D) effect of these opened cracks are thus introduced to take into account the misfit of crack face and the damage activation under compressive loading.

The evolution rule of damage can also be obtained by three methods: the experiment-based method, the micromechanical method, and the thermodynamics-based method. The first method includes curve fitting [14] and statistics based or stochastic method [9,15]. The experiment-based method is able to offer rational results for the tested cases, but is unable to be generalized to other more complicated cases. As for the second method, the micromechanical derivation of microcrack kinetic growth laws is currently achievable only for the case of originally homogenous linear isotropic elastic solids without microcrack interaction [16]. The thermodynamics-based method can be divided into three categories: the associated method, the nonassociated method, and the

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 21, 2000; final revision, December 1, 2002. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

direct method. The associated method [8] introduces a damage surface, which defines the reversible domain. The evolution of the damage should be normal to the damage surface (flux rule) and guarantee that the state variables of the material stay on the succeeding damage surface (consistent condition). The associated method can also be deduced equivalently from the principle of maximum dissipation [17], or the principle of minimum free energy [18]. The bounding surface method [2] is a generalization of the theory of bounding surface in plasticity (an associated plasticity with a refined hardening rule [19]) and thus is also a kind of associated method. In the nonassociated method [16,20], however, it is not necessary for the flux rule and consistent condition to be derived from the same surface. On the other hand, the direct method directly utilizes certain principles and relations in thermodynamics, for example, the Onsager relations [21], or specifies a dissipation function in the space of thermodynamics, $\phi(Y)$ (where Y is the conjugate variable of damage D), to deduce the evolution rule of damage by the relation [22],

$$\dot{D} = -\frac{\partial \phi}{\partial Y}. \tag{1}$$

The difficulty of the thermodynamics-based method lies in how to choose the dissipation potential and damage surface in the stress space or thermodynamics space. Usually, the experimental observation is resorted for this [20], while certain requirements in thermodynamics are considered, for example, the damage surface should be expressed by a homogeneous function of degree one [23].

In this paper, a displacement-equivalence-based damage model for brittle materials is proposed. The damage variable is defined using an effective-stress-based method. In the next section, a new damage deactivation criterion, which depends on both the stress and strain state of the materials, is discussed. The assumption of the displacement equivalence between the damaged configuration and the undamaged configuration is proposed in Section 3. Then, an additional crack opening induced normal deformation is introduced, and the 3D effect of these crack openings is also considered in Sec. 4. The evolution rule of damage is deduced using the Onsager relations in Sec. 5. The summary and discussion are finally presented in Sec. 6.

2 Damage Variable and Effective Damage

2.1 Basic Assumptions. In this paper, the virgin materials are assumed to be homogeneous in macroscale. The defects in the material due to loading are assumed to be penny-shaped microcracks or mesocracks [8]. It is also assumed that only elastic damage occurs in the materials, that is to say, there is no plastic flux, permanent deformation, or any other dissipation in the materials. A second-order damage tensor D is used to depict the damage in the materials. When the principal directions of stress and strain overlap with the principal directions of damage, the problem need only to be considered in the principal coordinate system.

2.2 Deactivation of Damage Effects. Deactivation of damage effects or unilateral effects refers to the phenomenon that certain cracks do not affect the behavior of the materials under certain circumstances. For example, the damage due to tension will not lead to an apparent decrease of elastic modulus when compressive loading is applied in the same direction because of crack closure [24], as shown in Fig. 1d. One of the approaches used to deal with the unilateral effects is to decompose the principal stress tensor into a positive part and a negative part [22],

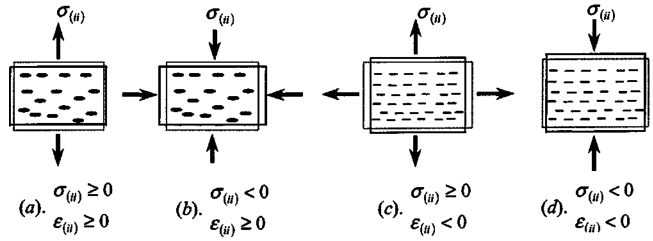


Fig. 1 Definition of effective damage tensor

$$[\sigma] = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} = \begin{bmatrix} \langle\sigma_1\rangle & 0 & 0 \\ 0 & \langle\sigma_2\rangle & 0 \\ 0 & 0 & \langle\sigma_3\rangle \end{bmatrix} - \begin{bmatrix} \langle-\sigma_1\rangle & 0 & 0 \\ 0 & \langle-\sigma_2\rangle & 0 \\ 0 & 0 & \langle-\sigma_3\rangle \end{bmatrix} \tag{2}$$

with the MacCauley brackets

$$\langle x \rangle = \begin{cases} x & \text{when } x > 0 \\ 0 & \text{when } x \leqslant 0 \end{cases}.$$

The Gibbs energy can be decomposed as

$$\varphi^* = \varphi_T^*(D_T, \langle\sigma\rangle) + \varphi_C^*(D_C, -\langle-\sigma\rangle). \tag{3}$$

Similar to the decomposition of the stress, there is also an approach to decompose the strain using projection operators [25].

Dragon and Halm [26] dealt with the unilateral effects in the principal directions of the second-order damage tensor and used the strain component as an indicator. If the normal strain component in the principal damage plane, assumed to have a normal vector N, is negative, that is,

$$N \cdot \epsilon \cdot N < 0, \tag{4}$$

then the crack in this plane is thought to be closed and have no effect on the free energy.

2.3 Definition of Effective Damage. In this paper, the effective damage $\tilde{D}$ is used to account for the opening and closing of cracks as shown in Fig. 1. We assumed that only in case Fig. 1d, where the stress and strain are all negative, will the unilateral effect be present. In the other cases (Figs. 1a, 1b, and 1c), the cracks will not be able to resist any loading, so the damage is "effective." In fact, Figs. 1a and 1b depict Eq. (4). As for Fig. 1c, although the cracks are closed, they cannot withstand any tension loading if the surfaces of the cracks are assumed to be smooth. Otherwise, a permanent deformation due to friction and interlock on the cracking surfaces will be present, and will contradict with our basic assumption stated in Sec. 2.1 (a detailed discussion is given in Appendix A). In fact, the crack surface of a real material will never be perfectly smooth, so this assumption only represents a necessary but not sufficient condition for the damage activation.

With this assumption, we define the effective principal damage components in the principal damage coordinate system as

$$\tilde{D}_i = \begin{cases} D_i & \text{when } \sigma_{(ii)} \geqslant 0 \text{ or } \epsilon_{(ii)} \geqslant 0 \\ 0 & \text{when } \sigma_{(ii)} < 0 \text{ and } \epsilon_{(ii)} < 0 \end{cases}, \tag{5}$$

where (ii) implies no summation convention on i. In the principal damage coordinate system, the effective damage tensor can also be expressed as

$$\tilde{D} = \text{diag}\{\tilde{D}_1, \tilde{D}_2, \tilde{D}_3\}. \tag{6}$$

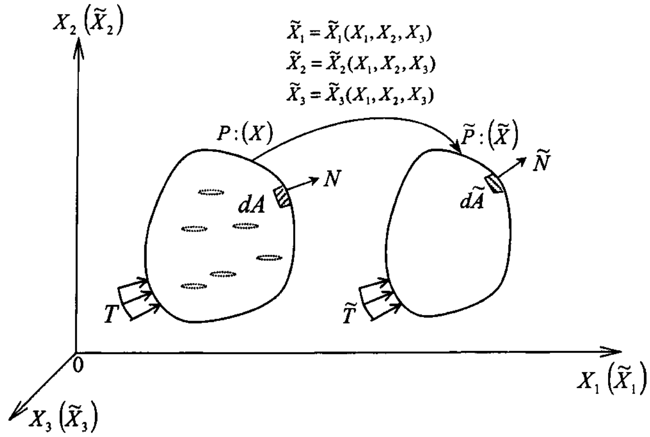


Fig. 2 Damaged and virtual undamaged configurations

3 Virtual Undamaged Configuration and Displacement Equivalence

3.1 Assumption of Displacement Equivalence. The effective stress and the equivalence hypotheses between damaged and virtual undamaged configurations are the two basic concepts in continuum damage mechanics [7]. For damaged material, the effect of the defects is assumed to be mainly on the decrease of the effective area for resisting the load. Thus, the virtual undamaged configuration is introduced, as shown in Fig. 2. A point in the damaged configuration, $P:(X)$, is mapped to a point in the virtual undamaged configuration, $\tilde{P}:(\tilde{X})$, and an area element in the damaged configuration, dA, with unit normal vector N is mapped to an area element in the virtual undamaged configuration, $d\tilde{A}$, with unit normal vector $\tilde{N}$. The distributive load on the surface of the damaged configuration, T, can be transplanted to the virtual undamaged configuration following the rule [27]

$$T\cdot dA=\tilde{T}\cdot d\tilde{A} \quad \text{or} \quad T_i dA|_{(X)}=\tilde{T}_i d\tilde{A}|_{(\tilde{X}(X))} \quad i=1,2,3. \tag{7}$$

Actually, Eq. (7) defines the effective stress, $\tilde{\sigma}$, in the virtual damaged configuration and further determines the relationship between the damage D and the mapping function $\tilde{X}=\tilde{X}(X)$ [27]:

$$\tilde{\sigma}^T=\sigma\cdot(I-D)^{-1}, \quad (I-D)=\det\left[\frac{\partial\tilde{X}}{\partial X}\right]\cdot\left[\frac{\partial\tilde{X}}{\partial X}\right]^{-T}, \tag{8}$$

where I is the unit identity matrix and the superscript $-T$ implies the transposition of the inverse matrix.

The two widely used principles of equivalence between damaged and undamaged configurations (strain and strain energy) are reviewed by Li [7], in which two energy equivalent principles are also proposed. These hypotheses focused either on geometrical intuition or on energy conservation (accumulated and dissipated). In this paper, an assumption of displacement equivalence is adopted and shown below that the equivalence of displacement (geometrical intuition) can also lead to the equivalence of work done by stress (energy conservation).

The *assumption of displacement equivalence* states that the displacement of the real damaged configuration is equal to that of the corresponding virtual undamaged configuration. In other words, any possible displacement that occurs in the virtual undamaged configuration $\delta\tilde{U}$ can be transplanted to the damaged configuration following the rule

$$\delta U=\delta\tilde{U} \quad \text{or} \quad \delta U_i(X)=\delta\tilde{U}_i(\tilde{X}(X)), \quad i=1,2,3. \tag{9}$$

From Eqs. (7) and (9), the following relation can be derived:

$$\int\int_{A(X)} T_i(X)\cdot\delta U_i(X)dA=\int\int_{\tilde{A}(X)}\tilde{T}_i(\tilde{X}(X))\delta\tilde{U}_i(\tilde{X}(X))d\tilde{A}$$

or

$$\int\int_A T\cdot\delta U dA=\int\int_{\tilde{A}}\tilde{T}\cdot\delta\tilde{U}d\tilde{A}. \tag{10}$$

The left part of Eq. (10) is the work done by the force applied on the damaged configuration, and the right part is the work done by the force applied on the virtual undamaged configuration.

According to the first law of thermodynamics, if there is no mechanical dissipation, thermal flux, or temperature variation in the deformation process, the strain energy accumulated in the two configurations will be equal, as shown in Eq. (10). If certain mechanical dissipation occurs in the process such as plasticity flux, the displacement δU can be decomposed into the reversible part, δU^E, and the irreversible part, δU^I. So the equivalence of δU^E leads to the equivalence of the strain energy accumulated in the two configurations, and the equivalence of δU^I leads to the equivalence of the energy dissipated in the two configurations. Because the energy dissipation due to the damage evolution is not included in the assumption of displacement equivalence, this assumption is different from the energy equivalent principles II proposed in Ref. [7], which requires that the total dissipative energy in the damaged configuration, including that due to damage evolution, equals the total dissipative energy in the virtual undamaged configuration. But when only elastic damage occurs in the damaged configuration, there will be no dissipative process in the corresponding virtual undamaged configuration. In this case, the energy equivalent principles II will not be satisfied.

3.2 Application of the Assumption of Displacement Equivalence. According to the basic assumptions stated in Sec. 2.1, only elastic damage occurs in the deformation process. So, we need only to assume that the elastic (or reversible) displacements of the two configurations are equivalent. We consider the element (with lateral length l) of the damaged configuration under three principal stresses. The stress-strain relationship can be expressed as

$$\epsilon=[S(D)]\cdot\sigma, \tag{11}$$

where $[S(D)]$ is the secant compliance matrix. Noting that

$$\begin{Bmatrix}\epsilon_1\\\epsilon_2\\\epsilon_3\end{Bmatrix}=[S_{ij}(D)]\begin{Bmatrix}\sigma_1\\\sigma_2\\\sigma_3\end{Bmatrix}=[S_{ij}(D)]\left(\begin{Bmatrix}\sigma_1\\0\\0\end{Bmatrix}+\begin{Bmatrix}0\\\sigma_2\\0\end{Bmatrix}+\begin{Bmatrix}0\\0\\\sigma_3\end{Bmatrix}\right)$$
$$=\begin{Bmatrix}\epsilon_1^1\\\epsilon_2^1\\\epsilon_3^1\end{Bmatrix}+\begin{Bmatrix}\epsilon_1^2\\\epsilon_2^2\\\epsilon_3^2\end{Bmatrix}+\begin{Bmatrix}\epsilon_1^3\\\epsilon_2^3\\\epsilon_3^3\end{Bmatrix}, \tag{12}$$

the strain of the element can be divided into three parts, with each part corresponding to a principal stress component. But this does not imply that the total deformation of the element equals the sum of deformations due to the three principal stresses applied separately, because the definition of the effective damage [Eq. (5)] depends on all of the state variables (D and σ).

Let us consider the part corresponding to σ_1 first. According to the concept of effective stress, the effective area for resisting the load, $\tilde{A}_1$, will decrease to

$$\tilde{A}_1=(1-\tilde{D}_1)A_1, \tag{13}$$

and the effective stress will increase to

$$\tilde{\sigma}_1=\sigma_1/(1-\tilde{D}_1). \tag{14}$$

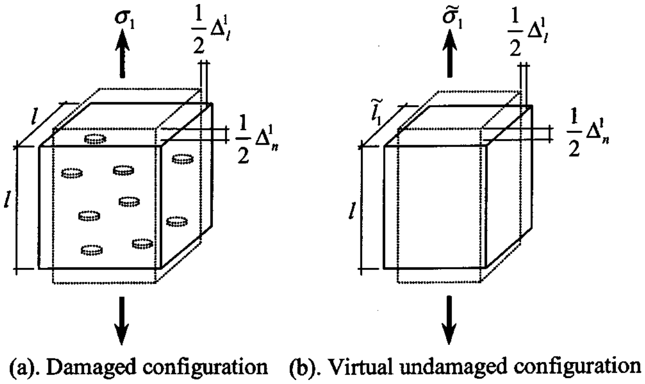


Fig. 3 Assumption of displacement equivalence

We can then define a virtual undamaged configuration with a cross sectional area $\tilde{A}_1$, under the loading $\tilde{\sigma}_1$, as shown in Fig. 3*b*. The displacement of the virtual undamaged configuration can be further deduced as

$$\Delta_n^1 = \frac{\tilde{\sigma}_1}{E} l = \frac{\sigma_1}{E(1-\tilde{D}_1)} l,$$
$$\Delta_l^1 = -\nu \frac{\tilde{\sigma}_1}{E} \tilde{l}_1 = -\nu \frac{\sigma_1}{E(1-\tilde{D}_1)} \tilde{l}_1 = -\nu \frac{\sigma_1}{E(1-\tilde{D}_1)} \sqrt{\tilde{A}_1}, \tag{15}$$

where Δ_n^1 and Δ_l^1 are the normal and lateral displacements due to $\tilde{\sigma}_1$, respectively, and E and ν are the elastic modulus and Poisson's ratio of the virgin material, respectively. According to the assumption of displacement equivalence, the strain of the element in the damaged configuration corresponding to σ_1 is

$$\epsilon_1^1 = \frac{\Delta_n^1}{l} = \frac{\sigma_1}{E(1-\tilde{D}_1)},$$
$$\epsilon_2^1 = \epsilon_3^1 = \frac{\Delta_l^1}{l} = -\nu \frac{\sigma_1}{E(1-\tilde{D}_1)} \frac{\sqrt{\tilde{A}_1}}{\sqrt{A_1}} = -\nu \frac{\sigma_1}{E\sqrt{1-\tilde{D}_1}}. \tag{16}$$

Similarly, we can obtain ϵ^2 and ϵ^3, and according to Eq. (12), the total strain of the element in the damaged configuration, ϵ, is

$$\epsilon = \epsilon^1 + \epsilon^2 + \epsilon^3. \tag{17}$$

The components of ϵ can be specified as

$$\begin{Bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{Bmatrix} = \frac{1}{E} \begin{bmatrix} 1 & -\nu\sqrt{1-\tilde{D}_1} & -\nu\sqrt{1-\tilde{D}_1} \\ -\nu\sqrt{1-\tilde{D}_2} & 1 & -\nu\sqrt{1-\tilde{D}_2} \\ -\nu\sqrt{1-\tilde{D}_3} & -\nu\sqrt{1-\tilde{D}_3} & 1 \end{bmatrix} \begin{Bmatrix} \tilde{\sigma}_1 \\ \tilde{\sigma}_2 \\ \tilde{\sigma}_3 \end{Bmatrix}, \tag{18}$$

where the effective stress components $\tilde{\sigma}_i = \sigma_i/(1-\tilde{D}_i)$.

3.3 Comparison of the Strain, Energy, and Displacement Equivalency. From Eq. (18), it is apparent that the difference between the assumption of displacement equivalence and the principle of strain equivalence lies in the fact that the former has an effect on Poisson's ratio while the latter does not. According to the assumption of displacement equivalence, the lateral deformation due to effective stress occurs only in the effective area $\tilde{A}$ in the virtual undamaged configuration. So the average lateral deformation in the nominal area, A in the damaged configuration, will decrease because $A \geqslant \tilde{A}$, and the lateral strain of the damaged configuration is thus smaller than that of the virtual undamaged configuration. Actually, the principle of strain equivalence defines a rule to transplant the strain component from the virtual undamaged configuration to the damaged configuration, i.e.,

$$\frac{\partial U_i}{\partial X_j}(X) = \frac{\partial \tilde{U}_i}{\partial \tilde{X}_j}(\tilde{X}(X)), \quad i,j=1,2,3. \tag{19}$$

However, according to the assumption of displacement equivalence, Eq. (9), we have

$$\frac{\partial U_i}{\partial X_j}(X) = \frac{\partial \tilde{U}_i}{\partial X_j}(\tilde{X}(X)) = \frac{\partial \tilde{U}_i}{\partial \tilde{X}_k}(\tilde{X}(X)) \frac{\partial \tilde{X}_k}{\partial X_j}, \quad i,j,k=1,2,3. \tag{20}$$

Equations (19) and (20) demonstrate the difference between the principle of strain equivalence and the assumption of the displacement equivalence.

From Eq. (20), the density of complementary strain energy in the damaged configuration can be derived, under the assumption of displacement equivalence, as

$$\begin{aligned} W^c(\sigma,D) &= \frac{1}{2}\sigma_{ij}\epsilon_{ij} = \frac{1}{2}\sigma_{ij}\frac{1}{2}\left(\frac{\partial U_i}{\partial X_j} + \frac{\partial U_j}{\partial X_i}\right) \\ &= \frac{1}{2}\sigma_{ij}\frac{\partial U_i}{\partial X_j} = \frac{1}{2}\sigma_{ij}\frac{\partial \tilde{U}_i}{\partial \tilde{X}_k}\frac{\partial \tilde{X}_k}{\partial X_j}, \\ & i,j,k=1,2,3. \end{aligned} \tag{21}$$

According to the principle of strain energy equivalence, the complementary strain energy in per unit volume of damaged material remains the same as that in the per unit volume of virtual undamaged material, i.e.,

$$W^c(\sigma,D) = \tilde{W}^c(\tilde{\sigma}) = \tfrac{1}{2}\tilde{\sigma}_{ij}\frac{\partial \tilde{U}_j}{\partial \tilde{X}_i} \tag{22}$$

[noting that in Eq. (22), $\tilde{\sigma}$ is not a symmetric tensor, so the conjugated variable of $\tilde{\sigma}_{ij}$ should be $\partial \tilde{U}_j/\partial \tilde{X}_i$]. According to Eq. (8), the density of complementary strain energy in the virtual undamaged configuration can be derived as

$$\tilde{W}^c(\tilde{\sigma}) = \tfrac{1}{2}\tilde{\sigma}_{ij}\frac{\partial \tilde{U}_j}{\partial \tilde{X}_i} = \tfrac{1}{2}\sigma_{jk}\frac{1}{\det[\partial \tilde{X}/\partial X]}\frac{\partial \tilde{X}_i}{\partial X_k}\cdot\frac{\partial \tilde{U}_j}{\partial \tilde{X}_i}$$
$$\xrightarrow{i\Rightarrow k,\ k\Rightarrow j,\ j\Rightarrow i} \tfrac{1}{2}\sigma_{ij}\frac{1}{\det[\partial \tilde{X}/\partial X]}\frac{\partial \tilde{U}_i}{\partial \tilde{X}_k}\frac{\partial \tilde{X}_k}{\partial X_j}. \tag{23}$$

Comparing Eq. (21) with Eq. (23), we can find that the assumption of displacement equivalence defines a relationship between the complementary strain energy density of damaged and virtual undamaged material as

$$W^c(\sigma,D) = \det\left[\frac{\partial \tilde{X}}{\partial X}\right]\cdot \tilde{W}^c(\tilde{\sigma}). \tag{24}$$

Noting that $d\tilde{V} = \det[\partial \tilde{X}/\partial X]\cdot dV$, where dV and $d\tilde{V}$ are the volumes of the elements in the damaged and virtual undamaged configurations, respectively. Hence, we can find that the difference between the assumption of displacement equivalence [Eq. (24)] and the principle of the strain energy equivalence [Eq. (22)] lies in the fact that the former considers the variation of volume from the damaged configuration to the virtual undamaged configuration, but the latter does not.

4 Additional Normal Strain Due to Crack Opening and 3D Effect of Cracks

4.1 Additional Normal Strain Due to Crack Opening. Extra strain has been used to consider the dilatancy effect in the brittle materials [28]. In this paper, it is attributed to the misfit of crack faces, and considered using additional normal strain due to cracking open. We assumed that the crack opening will lead to an additional normal strain as shown in Fig. 4, and the total normal strain will increase to

$$\epsilon_{(ii)}=\frac{\epsilon'_{(ii)}}{(1-\tilde{D}_i)^{\delta}}, \tag{25}$$

where δ is a model parameter and $\epsilon'_{(ii)}$ is the normal strain obtained from Eq. (18). From Eqs. (18) and (25), we obtain

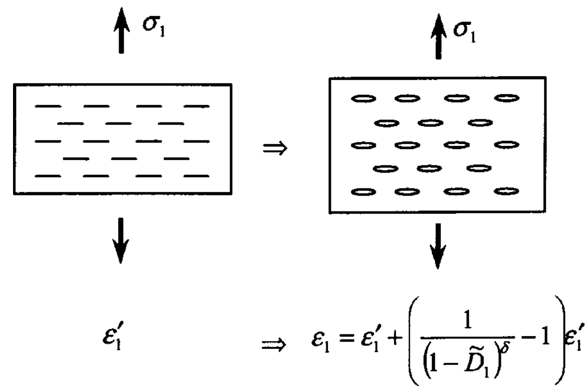


Fig. 4 Increase of normal strain caused by opening of cracks

$$\begin{Bmatrix}\epsilon_{11}\\ \epsilon_{22}\\ \epsilon_{33}\end{Bmatrix}=\frac{1}{E}\begin{bmatrix}(1-\tilde{D}_1)^{-1-\delta} & -\nu(1-\tilde{D}_1)^{-\delta}(1-\tilde{D}_2)^{-0.5} & -\nu(1-\tilde{D}_1)^{-\delta}(1-\tilde{D}_3)^{-0.5}\\ -\nu(1-\tilde{D}_1)^{-0.5}(1-\tilde{D}_2)^{-\delta} & (1-\tilde{D}_2)^{-1-\delta} & -\nu(1-\tilde{D}_2)^{-\delta}(1-\tilde{D}_3)^{-0.5}\\ -\nu(1-\tilde{D}_1)^{-0.5}(1-\tilde{D}_3)^{-\delta} & -\nu(1-\tilde{D}_2)^{-0.5}(1-\tilde{D}_3)^{-\delta} & (1-\tilde{D}_3)^{-1-\delta}\end{bmatrix}\begin{Bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\end{Bmatrix}. \tag{26}$$

We assumed that $\delta=0.5$, so as to ensure that the secant compliance matrix in Eq. (26) is symmetric.

4.2 3D Effect of Crack Opening. When uniaxial compression is applied as shown in Fig. 5, according to Eq. (5), $\tilde{D}_3=0$, and according to Eq. (26) the elastic modulus σ_3/ϵ_3 will not change. But it is well known that, in reality, there is degradation of the elasticity modulus in direction 3. The micromechanical analysis and experimental results [29] indicated that it is the frictional sliding on the preexisting slant flaws that leads to tension cracking and splitting in the direction of the compression loading (as illustrated in Fig. 5*a*). Thus, the degradation of the elasticity modulus under compression is due to a different mechanism from that under tension. Some authors used different damage variables for the tension loading and the compression loading, and the two damages hence evolved independently [30].

However, according to micromechanical analysis [29], the degradation of the materials' behavior in the direction of compression due to the microscale frictional sliding can also be measured or expressed by the damages $\tilde{D}_1$ and $\tilde{D}_2$, which are attributed to microscale tension cracking. Therefore, the cracks in brittle material, which are considered to be all penny-shaped, can also affect the elastic performance in the direction parallel to the cracking plane.

However, the complicated result of the micromechanical analysis may not be suitable for use in an effective-stress-based model, which is supposed to be simple and easy for engineering application, as indicated in Sec. 1. So, in this paper, we introduce a so-called 3D effect of the crack opening to account for this phenomenon, which assumes that there is a stress release zone around each preexisting flaw [31] and the tension crack will lead to an enlargement of the stress release zone. It thus leads to a decrease of the effective area in the direction of compression, as shown in Fig. 5*b*. To take into account this 3D effect, a new damage variable has been defined,

$$\omega_i=\begin{cases}(1-\tilde{D}_i) & \text{when } \sigma_i\geqslant 0 \text{ or } \epsilon_i\geqslant 0\\ (1-\tilde{D}_1)^{\beta}(1-\tilde{D}_2)^{\beta}(1-\tilde{D}_3)^{\beta} & \text{when } \sigma_i<0 \text{ and } \epsilon_i<0\end{cases}, \tag{27}$$

where β is a model parameter related to the stress release zone around each crack opening. Therefore, corresponding to the damaged configuration with the 3D effect, we define a new virtual undamaged configuration with cross section $\tilde{A}_i=A_i\omega_i$. Similar to the method used in Secs. 3.2 and 4.1, the new constitutive relationship can be expressed as

$$\begin{Bmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\end{Bmatrix}=\frac{1}{E}\begin{bmatrix}\omega_1^{-3/2} & -\nu\omega_1^{-1/2}\omega_2^{-1/2} & -\nu\omega_1^{-1/2}\omega_3^{-1/2}\\ -\nu\omega_1^{-1/2}\omega_2^{-1/2} & \omega_2^{-3/2} & -\nu\omega_2^{-1/2}\omega_3^{-1/2}\\ -\nu\omega_1^{-1/2}\omega_3^{-1/2} & -\nu\omega_2^{-1/2}\omega_3^{-1/2} & \omega_2^{-3/2}\end{bmatrix}\times\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\end{Bmatrix}, \tag{28}$$

and it is exactly the same as Eq. (26), if ω_i is replaced by $(1-\tilde{D}_i)$. The Gibbs free energy can be written as

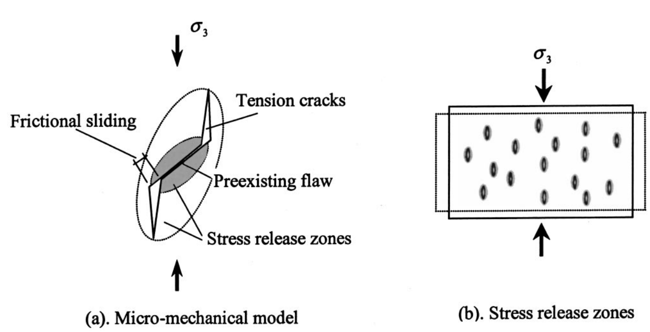


Fig. 5 Stress release zones under uniaxial compression

$$\varphi^* = \frac{1}{2E}\{\sigma_1 \quad \sigma_2 \quad \sigma_3\}$$

$$\cdot \begin{bmatrix} \omega_1^{-3/2} & -\nu\omega_1^{-1/2}\omega_2^{-1/2} & -\nu\omega_1^{-1/2}\omega_3^{-1/2} \\ -\nu\omega_1^{-1/2}\omega_2^{-1/2} & \omega_2^{-3/2} & -\nu\omega_2^{-1/2}\omega_3^{-1/2} \\ -\nu\omega_1^{-1/2}\omega_3^{-1/2} & -\nu\omega_2^{-1/2}\omega_3^{-1/2} & \omega_3^{-3/2} \end{bmatrix}$$

$$\cdot \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{Bmatrix} \tag{29}$$

It should be noted that the additional normal strain introduced in Eq. (25) and the 3D effect introduced in Eq. (27) are just assumptions based on experimental observation and other researcher's work. They may lack rigorous mechanical basis and may be viewed as compromises between the rigorous mechanical analysis and the engineering application.

5 Evolution of Damage

Although "thermodynamics alone does not provide all the necessary prerequisites for the formulation of a rational and successful continuum mode" [3], it is still the most convenient tool available to deduce a damage evolution rule, especially when the definition of the damage has been given, e.g., Eq. (28), and when the experimental results are still inadequate.

As an irreversible process, the evolution of damage must satisfy the second law of thermodynamics, which means that the dissipation ϕ must not be negative, i.e.,

$$\phi = \phi_D + \phi_P + \phi_T + \cdots \geqslant 0, \tag{30}$$

where the subscripts D, P, and T indicate the dissipation due to damage evolution, plasticity flux, and temperature variation, respectively. We assume that the dissipation due to damage can be uncoupled from the other two dissipation forms as the other two dissipations are not considered according to the basic assumptions stated in Sec. 2.1. Therefore, from continuum damage mechanics [20] we obtain

$$\phi_D = -Y:\dot{D} \geqslant 0, \tag{31}$$

where $\dot{D}$ is the rate of damage, and Y is the ratio of strain energy density release, which is defined as

$$Y = -\frac{\partial \varphi^*}{\partial D}, \tag{32}$$

where φ^* is the Gibbs free energy. According to Onsager relations, Eq. (31) will always be satisfied if we choose

$$\dot{D}_i = -F_{ij} Y_j, \tag{33}$$

where the matrix F is positively definitive. As indicated in Appendix B, matrix F can be specified as

$$F = \mathrm{diag}\{F_{11}, F_{22}, F_{33}\}. \tag{34}$$

So, Eq. (29) can serve as an evolution equation of the damage, and, from Eqs. (29), (32), and (33) we obtain

$$\dot{D}_i = F_{ij}\frac{\partial \varphi^*}{\partial D_j} = \frac{F_{ij}}{2}\sigma_{kl}\frac{\partial \epsilon_{kl}(D,\sigma)}{\partial D_j}. \tag{35}$$

It should be pointed out that the evolution rule, Eq. (33), can also be derived by first specifying a dissipation function in the thermodynamics space, i.e.,

$$\phi = \tfrac{1}{2}F_{ij}Y_iY_j, \tag{36}$$

and then applying the relation in Eq. (1).

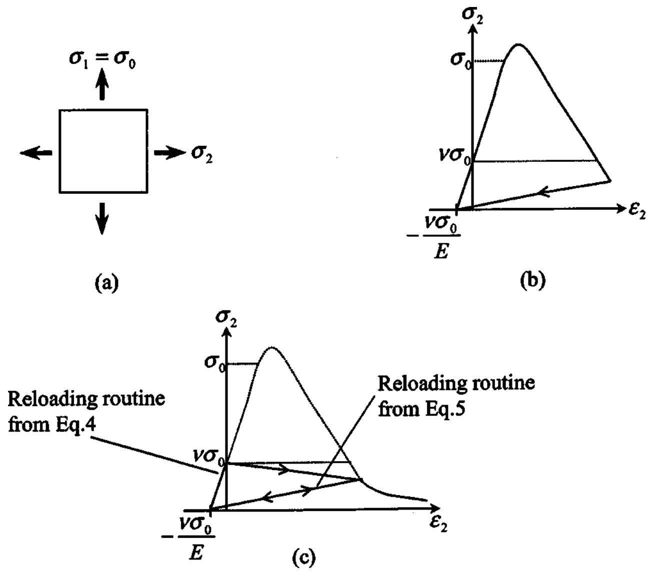


Fig. 6 Comparison of the damage deactivation criterion given by Eqs. (4) and (5)

6 Summary and Discussion

In this paper, a displacement equivalence-based damage model for brittle materials is proposed based on the concept of effective stress and virtual undamaged configuration. A new damage deactivation criterion, which depends on both the stress and strain states of the materials, is adopted. According to the discussion given in Appendix A, the new criterion can lead to more reasonable results in some ideal cases under the basic assumptions stated in Sec. 2.1. However, for the real engineering materials, the criterion is a conservative assessment of the ability of the material to withstand loading.

The assumption of displacement equivalence between the damaged configuration and the undamaged configuration is proposed. It has been proven that the equivalence of displacement (geometrical intuition) also leads to the equivalence of work done by stress (energy conservation). However, the principle of equivalent strain does not consider the damage effect on Poisson's ratio, and the principle of equivalent strain energy does not consider the geometrical change of the material. For example, it does not consider the decrease of the effective volume.

Also, an additional crack-opening-induced normal deformation is introduced to take into account the misfit of the crack face. The microscale mechanism of the 3D effect of the crack opening is also discussed and considered using the concept of stress release zones, from the macroscale viewpoint. The evolution rule of damage is deduced using the Onsager relations.

Acknowledgments

We would like to express our sincere gratitude to the reviewers for their invaluable comments and suggestions given to this paper.

Appendix A. Discussion on Damage Deactivation Criterion

The element shown in Fig. 6*a* is under a plane stress state. First, increase σ_1 until $\sigma_1 = \sigma_0$, and keep it constant (assume that no damage occurred in the process). Then, increase σ_2 under displacement control until the remaining strength is smaller than $\nu\sigma_0$. Finally, unload σ_2. According to the basic assumptions stated in Sec. 2.1, no permanent deformation will occur, as shown in Fig. 6*b*. Again, reload σ_2, and if we follow Eq. (4), D_2 will not be active until $\epsilon_2 > 0$, that is $\sigma_2 > \nu\sigma_0$ (as shown in Fig. 6*c*), and this is an unreasonable result. If we follow Eq. (5), D_2 will always be active, and the reload routine seems more reasonable.

But it should be noted that when $\epsilon_2 < 0$, according to Eq. (28), we must have $\sigma_2 < 2\nu f_t$, where f_t is the tensile strength of the material. For brittle materials such as concrete, $f_t \approx 0.1 f_c$ and $\nu \approx 0.2$; hence the loading is less than $0.04 f_c$. Thus we can say that the difference between Eqs. (4) and (5) is small.

In the actual deformation process of brittle materials, the cracking surface would assist in withstanding some tensile load because of the friction and interlock on the surface, making Eq. (5) a conservative assessment [while Eq. (4) may be an overassessment]. In fact, the change from an activation or deactivation state of the damage to the other state is not so abrupt, but gradual. So using this proposed criterion, the discontinuity of the stress-strain response seems inevitable when the unilateral condition takes place. As pointed out by Caboche [32], no theory can ensure that the continuity of the stress-strain response, the symmetry of elastic stiffness and the complete anisotropy will all be simultaneously satisfied.

Appendix B. Diagonalization of Matrix *F*

Consider a brittle material under uniaxial compression in the direction 3. According to Eq. (33), the evolution equation is

$$\begin{Bmatrix} \dot{D}_1 \\ \dot{D}_2 \\ \dot{D}_3 \end{Bmatrix} = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \\ y_3 \end{Bmatrix}. \tag{B1}$$

From Eqs. (27) and (29) we know that the Gibbs free energy $\varphi^* = \varphi^*(D_1, D_2)$, so

$$y_1 = \frac{\partial \varphi^*}{\partial D_1} \neq 0, \quad y_2 = \frac{\partial \varphi^*}{\partial D_2} \neq 0, \quad y_3 = \frac{\partial \varphi^*}{\partial D_3} = 0,$$

$$\dot{D}_3 = F_{31} y_1 + F_{32} y_2 = 0. \tag{B2}$$

Because of the symmetry of directions 1 and 2, one can obtain

$$y_1 = y_2, \quad F_{31} = F_{32}. \tag{B3}$$

Hence, it must be $F_{31} \equiv 0$, $F_{32} \equiv 0$. If we change the direction of load to be along direction 1 and direction 2, respectively, we can obtain

$$F_{ij} \equiv 0 \quad \text{when } i \neq j. \tag{B4}$$

References

[1] Krajcinovic, D., 1989, ''Damage Mechanics,'' Mech. Mater., **8**, pp. 117–197.

[2] Li, Q. B., and Ansari, F., 1999, ''Mechanics of Damage and Constitutive Relationships for Concrete,'' J. Eng. Mech. Div., **125**, pp. 1–10.

[3] Krajcinovic, D., 1996, *Damage Mechanics*, North-Holland Series in Applied Mathematics and Mechanics, Elsevier, New York.

[4] Carol, I., Bazant, Z. P., and Prat, C. P., 1991, ''Geometric Damage Tensor Based on Micro-Plane Method,'' J. Eng. Mech. Div., **117**, pp. 2429–2448.

[5] Lemaitre, J., 1985, ''A Continuous Damage Mechanics Model for Ductile Fracture,'' J. Eng. Mat. Tech., **107**, pp. 83–89.

[6] Sidoroff, F., 1981, ''Description of Anisotropic Damage Application to Elasticity,'' *IUTAM Colloquium, Physical Nonlinearities in Structural Analysis*, J. Hult and J. Lemaitre, eds., Springer-Verlag, Berlin, pp. 237–244.

[7] Li, Q. M., 2000, ''Energy Correlation Between a Damaged Macroscopic Continuum and its Sub-Scale,'' Int. J. Solids Struct., **37**, pp. 4539–4556.

[8] Krajcinovic, D., and Fonseka, G. U., 1981, ''The Continuous Damage Theory of Brittle Materials—Part 1: General Theory,'' ASME J. Appl. Mech., **48**, pp. 809–815.

[9] Ju, J. W., and Lee, X., 1991, ''Micro-Mechanical Damage Models for Brittle Solids I: Tensile Loading,'' J. Eng. Mech. Div., **117**, pp. 1495–1514.

[10] Tohgo, K., and Chou, T. W., 1996, ''Incremental Theory of Particulate-Reinforced Composites Including Debonding Damage,'' JSME Int. J., Ser. A, **39**, pp. 389–397.

[11] Ladeveze, P., 1983, ''Sur une theorie de l'endomagement anisotrope,'' Int. Report No. 34, Laboratoire de Mecanique et Technologie, Cachan, France (in French).

[12] Fichant, S., Pijaudier-Cabot, G., and La Borderie, C., 1997, ''Continuum Damages Modeling: Approximation of Crack Induced Anisotropy,'' Mech. Res. Commun., **24**, pp. 109–114.

[13] Krajcinovic, D., 1985, ''Constitutive Theories for Solids With Defective Microstructure,'' *Damage Mechanics and Continuum Modeling*, N. Stubbs and D. Krajcinovic, eds., ASCE, Reston, VA, pp. 39–56.

[14] Loland, K. E., 1980, ''Continuous Damage Model for Load-Response Estimation of Concrete,'' Cem. Concr. Res., **10**, pp. 392–492.

[15] Dong, L. L., Xie, H. P., and Zhao, P., 1995, ''Experimental Research on Complete Damage Process of Concrete Under Compression,'' J. Experimental Mech. (China), **10**, pp. 95–102.

[16] Ju, J. W., 1989, ''On Energy-Based Coupled Elastoplastic Damage Theories: Constitutive Modeling and Computational Aspects,'' Int. J. Solids Struct., **25**, pp. 803–833.

[17] Govindjee, S., Kay, G. J., and Simo, J., 1995, ''Anisotropic Modelling and Numerical Simulation of Brittle Damage in Concrete,'' Int. J. Numer. Methods Eng., **38**, pp. 3611–3633.

[18] Li, Q. M., 1999, ''Dissipative Flow Model Based on Dissipative Surface and Irreversible Thermodynamics,'' Archive Appl. Mech., Springer-Verlag, Berlin, pp. 379–392.

[19] Bazant, Z. P., and Prat, P. C., 1988, ''Micro-Plane Model for Brittle-Plastic Material: I Theory, II Verification,'' J. Eng. Mech. Div., **114**, pp. 1672–1700.

[20] Lemaitre, J., 1996, A Course on Damage Mechanics, Springer, Holland.

[21] Valanis, K. C., 1992, ''A Local and Non-Local Damage Theory: A Brief Review,'' *Damage Mechanics and Localization*, ASME, New York, **142**, pp. 145–152.

[22] Yu, T. Q., and Qian, J. C., 1993, *Damage Theory and its Application*, National Defense Industry Press of China, Beijing (in Chinese).

[23] Hansen, N. R., and Schreyer, H. L., 1994, ''A Thermodynamically Consistent Framework for Theories of Elastoplasticity Coupled With Damage,'' Int. J. Solids Struct., **31**, pp. 359–389.

[24] Van Mier, J. G. M., 1986, ''Fracture of Concrete Under Complex Stress,'' Heron, **31**, TNO-Delft, The Netherlands.

[25] Ortiz, M., 1985, ''A Constitutive Theory for the Inelastic Behavior of Concrete,'' Mech. Mater., **4**, pp. 67–93.

[26] Dragon, A., and Halm, D., 1998, ''A Meso-Crack Damage and Friction Coupled Model for Brittle Materials.'' *Damage Mechanics in Engineering Materials*, G. Z. Voyiadjis, J. W. Ju, and J. L. Chaboche, eds., Elsevier, New York, pp. 321–336.

[27] Murakami, S., and Ohno, M., 1980, ''A Continuum Theory of Creep and Creep Damage,'' *Proc. 3rd IUTAM Symp. on Creep in Structures*, A. R. S. Ponter and D. R. Hayhurst, eds., Springer-Verlag, Berlin, pp. 422–443.

[28] Fafitis, A., and Won, Y. H., 1992, ''A Multiaxial Stochastic Constitutive Law for Concrete: Part I—Theoretical Development and Part II—Comparison With Experimental Data,'' ASME J. Appl. Mech., **59**, pp. 283–294.

[29] Nemat-Nasser, N., and Hori, M., 1992, Micromechanics: Overall Properties of Heterogeneous Materials, North-Holland Series in Applied Mathematics and Mechanics, Elsevier, New York.

[30] Ramtani, S., 1990, ''Contribution a la Modelisation du Comportement Multiaxial du Beton Endommage avec Description du Caractere Unilateral,'' Thesis, Univ. Paris VI (in French).

[31] Frantziskonis, G., and Desai, C. S., 1987, ''Constitutive Model With Strain Softening,'' Int. J. Solids Struct., **23**, pp. 733–750.

[32] Caboche, J. L., 1992, ''On the Description of Damage Induced Anisotropy and Active/Passive Damage Effect,'' *Damage Mechanics in Engineering Materials*, J. W. Ju, D. Krajcinovic, and H. L. Schreyer, eds., ASME, New York, pp. 153–166.

Y. Liu
Research Scholar

C. K. Soh
Professor

Y. Dong
Research Fellow

Y. Yang
Teaching Fellow

School of Civil & Environmental Engineering,
Nanyang Technological University,
Singapore, 639798

A Displacement Equivalence-Based Damage Model for Brittle Materials—Part II: Verification

This paper focuses on the application of the damage model for brittle materials proposed in the preceding paper on theory. The evolution rule of damage derived using the Onsager relations is specified according to the experimental observation. The loading-unloading condition is discussed and the tangent modulus is derived. Then, the determination of the modal parameters is presented. For verification, the proposed model is applied to concrete under uniaxial and biaxial loading, and the numerical results are compared with those of other researchers and with the experimental results. The results are generally in good agreement and the proposed model is considered worthy for further research work. [DOI: 10.1115/1.1599915]

1 Introduction

In the preceding paper [1] on theory, a displacement equivalence-based damage model for brittle materials is proposed. A new damage deactivation criterion, which depends on both the stress and strain states of the materials, is adopted. Based on the concept of effective stress, the virtual undamaged configurations are introduced, and the assumption of the displacement equivalence is proposed to correlate the damaged and virtual undamaged configurations. Then, an additional crack-opening-induced normal deformation is introduced to take into account the misfit of crack faces. The three-dimensional (3D) effect of these opened cracks is also considered. The evolution rule of damage is deduced using the Onsager relations.

This paper focuses on the application of the proposed displacement equivalence based damage model. The damage evolution rules are discussed in detail in Sec. 2 for the uniaxial and multiaxial loading cases. The loading-unloading condition and the tangent modulus are also presented. The model parameters are discussed in Sec. 3. In Sec. 4, the proposed model is applied to concrete under uniaxial and biaxial loading, and the results are compared with those of the other researchers and with the published experimental results. Finally, the limitation of the proposed model and the further work are discussed in Sec. 5.

2 Evolution of Damage

In the preceding paper, the effective principal damage components are defined, as expressed in Eq. (1), to consider the activation and deactivation of damage in the principal damage coordinate system:

$$\tilde{D}_i=\begin{cases} D_i & \text{when } \sigma_{(ii)}\geqslant 0 \text{ or } \epsilon_{(ii)}\geqslant 0 \\ 0 & \text{when } \sigma_{(ii)}<0 \text{ and } \epsilon_{(ii)}<0 \end{cases}. \tag{1}$$

Then the secant compliance matrix is derived, as expressed in Eq. (2), from the assumption of displacement equivalence and the consideration of the additional normal strain and the 3D effect of the crack opening:

$$\begin{Bmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\end{Bmatrix}=\frac{1}{E}\begin{bmatrix}\omega_1^{-3/2} & -\nu\omega_1^{-1/2}\omega_2^{-1/2} & -\nu\omega_1^{-1/2}\omega_3^{-1/2}\\ -\nu\omega_1^{-1/2}\omega_2^{-1/2} & \omega_2^{-3/2} & -\nu\omega_2^{-1/2}\omega_3^{-1/2}\\ -\nu\omega_1^{-1/2}\omega_3^{-1/2} & -\nu\omega_2^{-1/2}\omega_3^{-1/2} & \omega_2^{-3/2}\end{bmatrix}\times\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\end{Bmatrix}=[S_{ij}]\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\end{Bmatrix}, \tag{2}$$

where S is the secant compliance matrix, and ω is the damage variable taking into account the 3D effect of the crack opening under compressive loading, which is defined as

$$\omega_i=\begin{cases}(1-\tilde{D}_i) & \text{when } \sigma_i\geqslant 0 \text{ or } \epsilon_i\geqslant 0\\ (1-\tilde{D}_1)^\beta(1-\tilde{D}_2)^\beta(1-\tilde{D}_3)^\beta & \text{when } \sigma_i<0 \text{ and } \epsilon_i<0\end{cases}. \tag{3}$$

Also a rate-dependent evolution rule of damage is deduced using the direct method:

$$\dot{D}_i=F_{ij}\frac{\partial\varphi^*}{\partial D_j}=\frac{F_{ij}}{2}\sigma_{kl}\frac{\partial\epsilon_{kl}(D,\sigma)}{\partial D_j}, \tag{4}$$

$$F_{ij}=\begin{cases}F_{(ii)} & \text{when } i=j\\ 0 & \text{when } i\neq j\end{cases}$$

In this section, the rate-independent form of the damage evolution rule is embodied for the uniaxial and multiaxial loading cases, respectively.

2.1 Uniaxial Tension. In the case of uniaxial tension, with $\sigma_1>0$, $\sigma_2=\sigma_3=0$, from Eq. (1) we know that $\tilde{D}_1=D_1>0$, $\tilde{D}_2=\tilde{D}_3=0$, and from Eq. (3) we know that

$$\omega_1=(1-D_1),\quad \omega_2=\omega_3=1. \tag{5}$$

Substituting Eq. (5) into Eq. (2), we obtain

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 21, 2000; final revision, December 1, 2001. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$\left\{\begin{array}{c}\epsilon_1\\ \epsilon_2\\ \epsilon_3\end{array}\right\}=\frac{1}{E}\begin{bmatrix}\frac{1}{\sqrt{(1-D_1)^3}} & \frac{-\nu}{\sqrt{(1-D_1)}} & \frac{-\nu}{\sqrt{(1-D_1)}}\\ \frac{-\nu}{\sqrt{(1-D_1)}} & 1 & -\nu\\ \frac{-\nu}{\sqrt{(1-D_1)}} & -\nu & 1\end{bmatrix}\left\{\begin{array}{c}\sigma_1\\ 0\\ 0\end{array}\right\}. \tag{6}$$

So, Eq. (4) becomes

$$\dot{D}_1=\frac{F_{11}}{2}\sigma_{kl}\frac{\partial\epsilon_{kl}(D,\sigma)}{\partial D_1}=\frac{3F_{11}\sigma_1^2}{4E}(1-D_1)^{-2.5},\quad \dot{D}_2=\dot{D}_3=0. \tag{7}$$

According to the experimental observation by Dong et al. [2], it is apparent that the ratio of the damage, $\dot{D}$, will become very small when $D\rightarrow 1$. We can ensure this tendency by choosing F_{11} to be

$$F_{11}=F_1(1-D_1)^{\gamma_t}, \tag{8}$$

where F_1 is a positive constant and γ_t is a model parameter used to depict the tendency of the strain, especially in the softening phase. Thus, the damage evolution equation for uniaxial tension can be written as

$$\dot{D}_1=\frac{3F_1\sigma_1^2}{4E}(1-D_1)^{\gamma_t-2.5}. \tag{9}$$

Assuming that $\dot{\epsilon}_1$ is kept constant when loading, we can obtain the analytical expression of the stress-strain relationship. Equation (9) can be rewritten as

$$dD_1=\dot{D}_1dt=\frac{1}{(d\epsilon_1/dt)}\dot{D}_1d\epsilon_1=\alpha_t\frac{\sigma_1^2}{E}(1-D_1)^{\gamma_t-2.5}d\epsilon_1, \tag{10}$$

where $\alpha_t=3F_1/4\dot{\epsilon}_1$ is a model parameter. Rewrite Eq. (10) as

$$(1-D_1)^{-\gamma_t-0.5}dD_1=\alpha_t E\epsilon_1^2 d\epsilon_1 \tag{11}$$

and integrate Eq. (11) to obtain the relationship between damage and strain,

$$C_t-(0.5-\gamma_t)^{-1}(1-D_1)^{(0.5-\gamma_t)}=\tfrac{1}{3}\alpha_t E\epsilon_1^3, \tag{12}$$

where C_t is an integration constant.

2.2 Uniaxial Compression. In the case of uniaxial compression with $\sigma_1=\sigma_2=0$, $\sigma_3<0$, from Eq. (1) we know that $\tilde{D}_1=D_1=\tilde{D}_2=D_2=D>0$, $\tilde{D}_3=0$, and from Eq. (3) we know that

$$\omega_1=(1-D_1)=\omega_2=1-D_2,\quad \omega_3=(1-D_1)^{\beta}(1-D_2)^{\beta}. \tag{13}$$

Substituting Eq. (13) into Eq. (2), we obtain

$$\left\{\begin{array}{c}\epsilon_1\\ \epsilon_2\\ \epsilon_3\end{array}\right\}=\frac{1}{E}\begin{bmatrix}\frac{1}{\sqrt{(1-D_1)^3}} & \frac{-\nu}{\sqrt{(1-D_1)(1-D_2)}} & \frac{-\nu}{\sqrt{(1-D_1)^{1+2\beta}}}\\ \frac{-\nu}{\sqrt{(1-D_1)(1-D_2)}} & \frac{1}{\sqrt{(1-D_2)^3}} & \frac{-\nu}{\sqrt{(1-D_2)^{1+2\beta}}}\\ \frac{-\nu}{\sqrt{(1-D_1)^{1+2\beta}}} & \frac{-\nu}{\sqrt{(1-D_2)^{1+2\beta}}} & \frac{1}{\sqrt{(1-D_1)^{3\beta}(1-D_2)^{3\beta}}}\end{bmatrix}\left\{\begin{array}{c}0\\ 0\\ \sigma_3\end{array}\right\}. \tag{14}$$

From Eq. (14), we derive

$$\begin{aligned}\dot{D}_1&=\frac{F}{2}\sigma_{kl}\frac{\partial\epsilon_{kl}(D_1,D_2,\sigma)}{\partial D_1}=\frac{3F\beta\sigma_3^2}{4E}(1-D)^{-3\beta-1}\\ &=\dot{D}_2=\dot{D},\quad \dot{D}_3=0.\end{aligned} \tag{15}$$

We can also choose the positive function F as

$$F=F'(1-D)^{\gamma_c} \tag{16}$$

Assuming that $\dot{\epsilon}_3$ is kept constant when loading, Eq. (15) can be rewritten as

$$(1-D)^{1-3\beta-\gamma_c}dD=-\alpha_c\beta E\epsilon_3^2 d\epsilon_3, \tag{17}$$

where $\alpha_c=3F'/4\dot{\epsilon}_3$ is also a model parameter. Integrating Eq. (17), we can derive the relationship between damage and strain as

$$C_c-(2-3\beta-\gamma_c)^{-1}(1-D)^{2-3\beta-\gamma_c}=-\tfrac{1}{3}\alpha_c\beta E\epsilon_3^3, \tag{18}$$

where C_c is an integration constant.

2.3 Multiaxial Loading. For the multi-axial loading case, it is assumed that the damage evolution equation can be expressed using the model parameters introduced for the uniaxial loading cases. Therefore, the damage evolution equation is deduced from Eqs. (4), (8), (10), (16), and (17) as

$$dD_i=\sum_{j=1}^{3}\frac{2}{3}\alpha_j\sigma_j(1-D_i)^{\gamma_j}\frac{\partial\epsilon_j}{\partial D_i}d\epsilon_j. \tag{19}$$

Here, the model parameters α_j and γ_j are either α_t or α_c, and either γ_t or γ_c, respectively, depending on either $\sigma_j>0$ or $\sigma_j<0$. $\partial\epsilon_j/\partial D_i$ can be deduced from the constitutive equation (2) as

$$\begin{aligned}\frac{\partial\epsilon_j}{\partial D_i}=\frac{1}{E}\sum_{k=1}^{3}\Bigg[&-\frac{3}{2}\omega_j^{-5/2}\frac{\partial\omega_k}{\partial D_i}\delta_{kj}\\ &+\frac{\nu}{2}\left(\omega_j^{-3/2}\omega_k^{-1/2}\frac{\partial\omega_j}{\partial D_i}+\omega_j^{-1/2}\omega_k^{-3/2}\frac{\partial\omega_k}{\partial D_i}\right)(1-\delta_{kj})\Bigg]\sigma_k,\end{aligned} \tag{20}$$

where δ_{kj} is the Kroneker symbol and

$$\frac{\partial \omega_i}{\partial D_j}=\begin{cases} -\beta\omega_i(1-D_j)^{-1} \quad \text{when } (i\neq j) \quad \text{and} \quad (\sigma_j>0 \text{ or } \epsilon_j>0) \quad \text{and} \quad (\sigma_i<0 \quad \text{and} \quad \epsilon_i<0) \\ -1 \quad \text{when } (i=j) \quad \text{and} \quad (\sigma_j>0 \text{ or } \epsilon_j>0) \\ 0 \quad \text{when all others.} \end{cases} \tag{21}$$

2.4 Loading-Unloading Condition. It is apparent from the damage evolution equations for the uniaxial loading case, i.e., Eqs. (10) and (17), that it will be convenient to express the loading-unloading condition using the tensile strain. Hence, for uniaxial tension, the loading-unloading condition can be expressed as

$$dD_1=\begin{cases} \text{Eq. (10)} \quad \text{when } \epsilon_1=\max\{\epsilon_t^{cr},\epsilon_1^{\max}\} \\ 0 \quad \text{when } \epsilon_1<\max\{\epsilon_t^{cr},\epsilon_1^{\max}\} \end{cases}. \tag{22}$$

where ϵ_t^{cr} is the damage threshold under uniaxial tension, and $\epsilon_1^{\max}$ is the maximum tensile strain in the whole loading history. For uniaxial compression, it can be expressed as

$$dD_1=\begin{cases} \text{Eq. (17)} \quad \text{when } \epsilon_1=\max\{\epsilon_c^{cr},\epsilon_1^{\max}\} \\ 0 \quad \text{when } \epsilon_1<\max\{\epsilon_c^{cr},\epsilon_1^{\max}\} \end{cases}, \tag{23}$$

where ϵ_c^{cr} is the damage threshold under uniaxial compression, and $\epsilon_1^{\max}$ is the maximum lateral tensile strain in the loading history. The damage threshold, ϵ_t^{cr} and ϵ_c^{cr}, will be discussed in the next section.

For the multiaxial loading case, the loading-unloading condition can also be written by combining those under axial loading. To do this, we rewrite Eq. (19) as

$$dD_i=\sum_{j=1}^{3} dD_i^j, \tag{24a}$$

$$dD_i^j=\frac{2}{3}\alpha_j\sigma_j(1-D_i)^{\gamma_j}\frac{\partial \epsilon_j}{\partial D_i}d\epsilon_j, \tag{24b}$$

where dD_i^j is the increment of damage due to stress σ_j and thus

$$dD_i^j=\begin{cases} \text{Eq. (24}b\text{)} \quad \text{when } \epsilon_i=\max\{\epsilon_i^{cr},\epsilon_i^{\max}\} \\ 0 \quad \text{when } \epsilon_i<\max\{\epsilon_i^{cr},\epsilon_i^{\max}\} \end{cases}, \tag{25}$$

where ϵ_i^{cr} is ϵ_t^{cr} when $\sigma_j>0$, or ϵ_c^{cr} when $\sigma_j<0$, and $\epsilon_i^{\max}$ is the maximum tensile strain ϵ_i in the loading history. The evolution equation in this section can also be deduced using the nonassociated method, as shown in Appendix A.

2.5 Tangent Modulus. The tangent modulus is useful for the numerical analysis. From Eq. (2), one can obtain

$$d\epsilon=\frac{\partial \epsilon}{\partial D}dD+Sd\sigma, \tag{26}$$

where $d\epsilon=\{d\epsilon_1\, d\epsilon_2\, d\epsilon_3\}^T$, $d\sigma=\{d\sigma_2\, d\sigma_2\, d\sigma_3\}^T$, $dD=\{dD_1\, dD_2\, dD_3\}^T$, and the matrix $[\partial\epsilon_j/\partial D_i]$ is given in Eq. (20). Rewrite Eq. (24) as

$$dD=[A_{ij}]\cdot d\epsilon, \tag{27a}$$

$$A_{ij}=\tfrac{2}{3}\alpha_j\sigma_j(1-D_i)^{\gamma_j}\frac{\partial\epsilon_j}{\partial D_i}. \tag{27b}$$

From Eqs. (26) and (27), the tangent modulus can be derived as

$$d\sigma=S^{-1}\left(I-\frac{\partial\epsilon}{\partial D}\cdot A\right)\cdot d\epsilon, \tag{28}$$

where I is the unit identity matrix.

3 Determination of Model Parameters

There are seven model parameters in Eqs. (2), (12), and (18), namely, β, α_t, α_c, γ_t, γ_c, C_t, and C_c. C_t and C_c are used to describe the thresholds of damage. When the tensile strain just reaches the thresholds of damage, it is assumed to be ϵ_t^{cr} in uniaxial tension and ϵ_c^{cr} in uniaxial compression. Setting $D_1=0$ and $D=0$ in Eqs. (12) and (18), respectively, we can obtain

$$C_t=(\tfrac{1}{2}-\gamma_t)^{-1}+\tfrac{1}{3}\alpha_t E(\epsilon_t^{cr})^3, \tag{29}$$

$$C_c=(2-3\beta-\gamma_c)^{-1}-\tfrac{1}{3}\alpha_c\beta E(-\epsilon_c^{cr}/\nu)^3. \tag{30}$$

It is also apparent from Eqs. (12) and (18) that it is γ_c and γ_t that determine the tendency of the strain while damage increases. For example in Eq. (12), $\epsilon_1\to\infty$, when $D_1\to 1$ if $\gamma_t>0.5$; and if $\gamma_t<0.5$, ϵ_1 is limited when $D_1\to 1$. So in practice, γ_t and γ_c are used to describe the behavior of concrete in the softening phase.

When uniaxial compression is applied, the ratio of the lateral strain to the normal one is

$$\frac{\epsilon_1}{\epsilon_3}=\frac{\epsilon_2}{\epsilon_3}=-(1-D)^{1/2-2\beta}=-\nu_c. \tag{31}$$

We notice that if $\beta<0.25$, the Poisson ratio under uniaxial compression, ν_c, will increase when the damage D develops; thus, β can be used to consider the dilatancy effect. α_t, α_c, C_t and C_c are used to fit the performance parameters of material. From Eqs. (6) and (12), we can derive the damage value when the material reaches its maximum tensile strength (See Appendix B) as

$$D_1^*=1-\left[\frac{4.5C_t}{4.5(0.5-\gamma_t)^{-1}+1}\right]^{1/(0.5-\gamma_t)}. \tag{32}$$

When the material reaches its maximum compressive strength, the damage will be

$$D^*=1-\left[\frac{9\beta C_c}{9\beta(2-3\beta-\gamma_c)^{-1}+1}\right]^{1/(2-3\beta-\gamma_c)}. \tag{33}$$

The expressions of the uniaxial tension and compression strength and the corresponding strains (f_t, f_c, ϵ_t, and ϵ_c) can thus be obtained from Eq. (2):

$$\epsilon_t=\left[\frac{3}{\alpha_t E}\,\frac{(0.5-\gamma_t)}{5-\gamma_t}C_t\right]^{1/3}, \tag{34}$$

$$f_t=E\epsilon_t(1-D_1^*)^{3/2}, \tag{35}$$

$$\epsilon_c=\left[\frac{3}{\alpha_c\beta E}\,\frac{2-3\beta-\gamma_c}{2+6\beta-\gamma_c}C_c\right]^{1/3}, \tag{36}$$

$$f_c=E\epsilon_c(1-D^*)^{3\beta}. \tag{37}$$

Therefore, the constants α_t, α_c, C_t, and C_c can be resolved as

Table 1 Material parameters of the six grades of concrete

No.	E (GPa)	ν	f_c (MPa)	ϵ_c	f_t (MPa)	ϵ_t
1	17.2	0.2	40.0	0.0029	4.14	0.00045
2	21.4	0.2	50.3	0.0030	5.24	0.00045
3	27.6	0.2	73.8	0.0036	7.59	0.00054
4	40.7	0.16	47.23	0.0019	3.10	0.00012
5	60.0	0.15	71.09	0.0022	3.96	0.00016
6	64.1	0.15	107.29	0.0020	5.63	0.00021

$$C_t=\left[(0.5-\gamma_t)^{-1}+\frac{2}{9}\right]\left(\frac{f_t}{E\epsilon_t}\right)^{(1-2\gamma_t)/3}, \qquad (38)$$

$$\alpha_t=\frac{3E^2}{f_t^3}\frac{0.5-\gamma_t}{5-\gamma_t}C_t\left[\frac{4.5C_t}{4.5(0.5-\gamma_t)^{-1}+1}\right]^{9/(1-2\gamma_t)}, \qquad (39)$$

$$C_c=\left[(2-3\beta-\gamma_c)^{-1}+\frac{1}{9\beta}\right]\left(\frac{f_c}{E\epsilon_c}\right)^{(2-3\beta-\gamma_c)/3\beta}, \qquad (40)$$

$$\alpha_c=\frac{3E^2}{f_c^3\beta}\frac{2-3\beta-\gamma_c}{2+6\beta-\gamma_c}C_c\left[\frac{9\beta C_c}{9\beta(2-3\beta-\gamma_c)^{-1}+1}\right]^{9\beta/(2-3\beta-\gamma_c)}. \qquad (41)$$

With these formulas, the stress-strain relationships under uniaxial tension and compression can then be written as

$$\sigma_1=\begin{cases}E(1-D_1)^{3/2}\epsilon_1=E\epsilon_1[(\tfrac{1}{2}-\gamma_t)(C_t-\tfrac{1}{3}\alpha_t E\epsilon_1^3)]^{3/(1-2\gamma_t)} & \text{when } \epsilon_1>\epsilon_t^{cr}\\ E\epsilon_1 & \text{when } \epsilon_1\leq\epsilon_t^{cr}\end{cases} \qquad (42)$$

and

$$\sigma_3=\begin{cases}E(1-D)^{3\beta}\epsilon_3=E\epsilon_3[(2-3\beta-\gamma_c)(C_c+\tfrac{1}{3}\alpha_c\beta E\epsilon_3^3)]^{3\beta/(2-3\beta-\gamma_c)} & \text{when } \epsilon_1>\epsilon_c^{cr}\\ E\epsilon_3 & \text{when } \epsilon_1\leq\epsilon_c^{cr}\end{cases} \qquad (43)$$

4 Numerical Examples

As a preliminary study, the proposed model is applied to concrete under uniaxial and biaxial loading, and the numerical results are compared with those of the other researchers and with the experimental results published by Fonseka and Krajcinovic [3], Li and Ansari [4], Dong and Xie [5], and Tasuji et al. [6].

4.1 Stress-Strain Response of Concrete Under Uniaxial Loading. The six grades of concrete used by Fonseka and Krajcinovic [3] and Li and Ansari [4] are considered. The material parameters are listed in Table 1. The stress-strain curves obtained are compared with their experimental results as depicted in Figs. 1–4. The curves depicting stress versus Poisson's ratio under

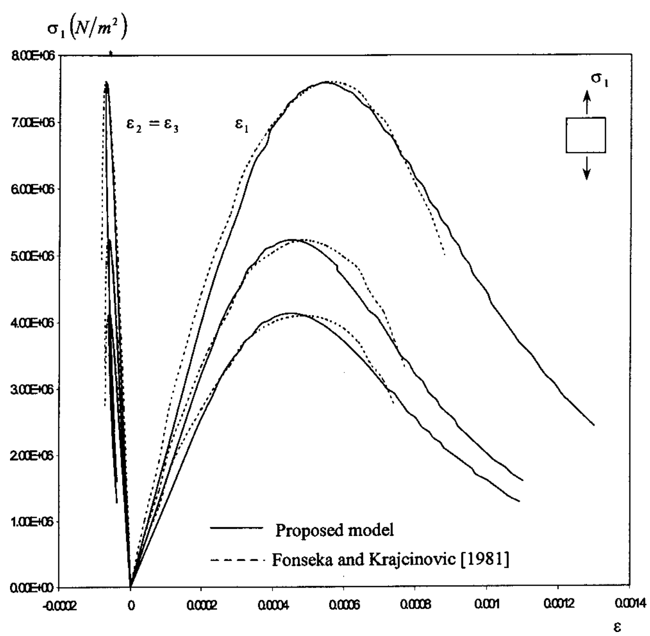


Fig. 1 Stress-strain response of concrete under uniaxial tension

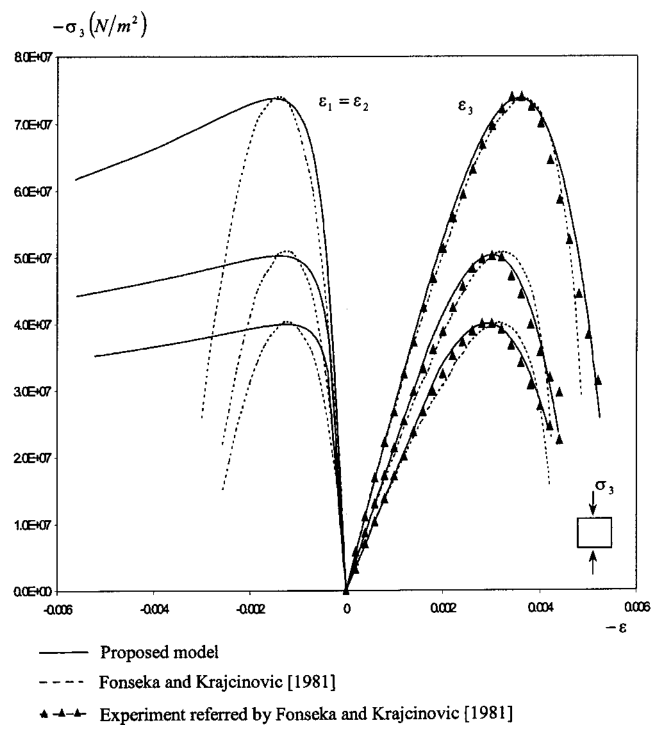


Fig. 2 Stress-strain response of concrete under uniaxial compression

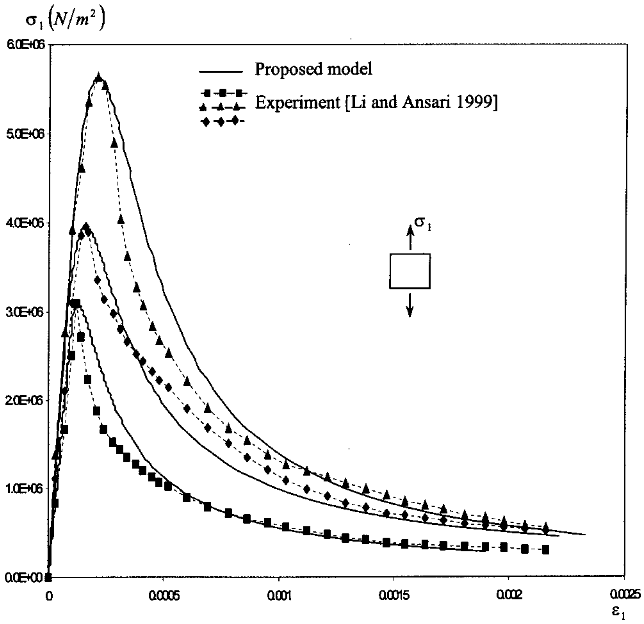


Fig. 3 Stress-strain response of high-strength concrete under uniaxial tension

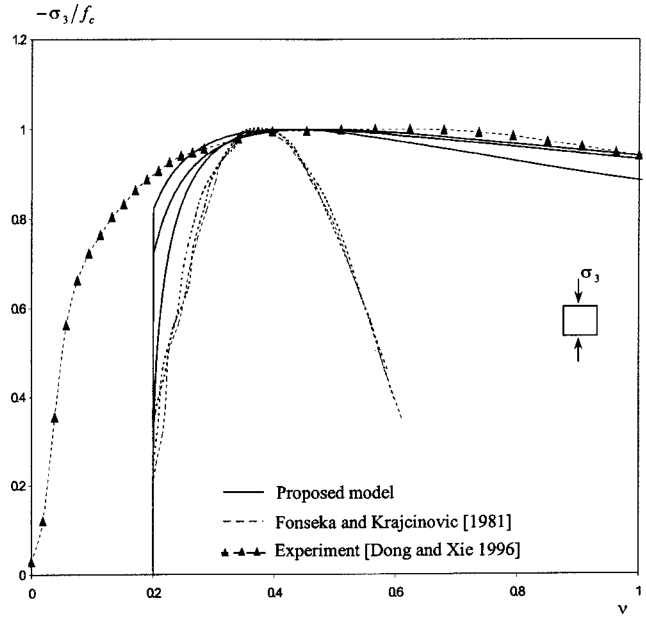


Fig. 5 Stress—Poisson's ratio response of concrete under uniaxial compression

uniaxial compression are also compared, in Fig. 5, with the work of Fonseka and Krajcinovic and a typical experimental result of Dong and Xie [5]. Figure 6 compares the volume-strain response of concrete under uniaxial compression obtained by Fonseka and Krajcinovic [3] with that obtained by the proposed model.

Figure 1 indicates that under uniaxial tension, the proposed model leads to a decrease of the lateral strain after the peak. This implies that unloading occurs in the lateral direction while loading progressed in the normal direction. Figures 2, 5, and 6 all show that the proposed model leads to more reasonable lateral strains and volume strains under uniaxial compression than the other researchers' results, as compared with the experimental observations. This is attributed to the fact that we have taken into account the additional normal strain caused by the crack opening, as given in Eq. (25), in the preceding paper on theory. Figures 3 and 4 show the application of the proposed model on high-strength concrete, and the results seem comparable with the experimental data.

4.2 Stress-Strain Response of Concrete Under Biaxial Loading. The stress-strain response of concrete under biaxial loading, for which the experimental data are available in Ref. [6], is also considered. The numerical and experimental results are compared in Figs. 7–9. Figure 10 shows the predicted biaxial ultimate strength envelope.

Figure 7 shows that under biaxial compression, the numerical result is comparable with the experimental results. However, Fig. 8 shows that the numerical-analysis-produced ultimate strengths

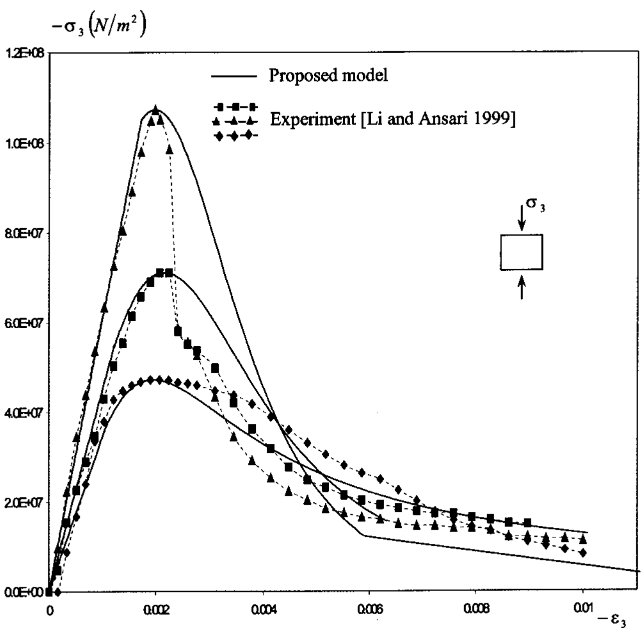


Fig. 4 Stress-strain response of high-strength concrete under uniaxial compression

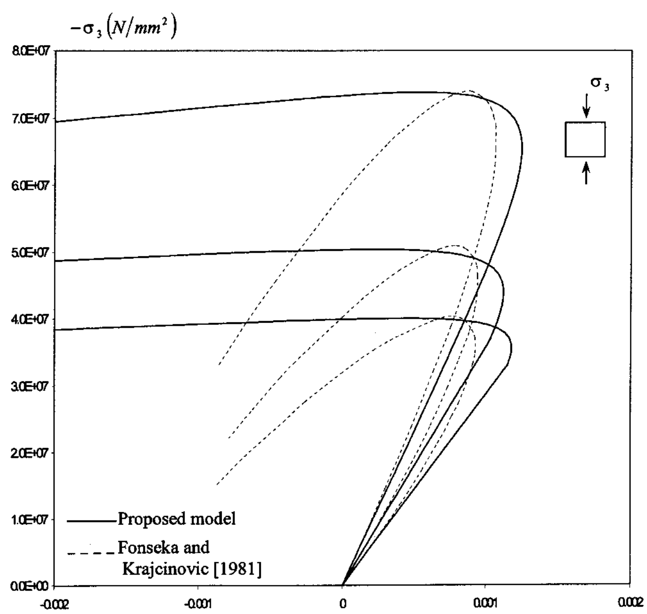


Fig. 6 Volume strain response of concrete under uniaxial compression

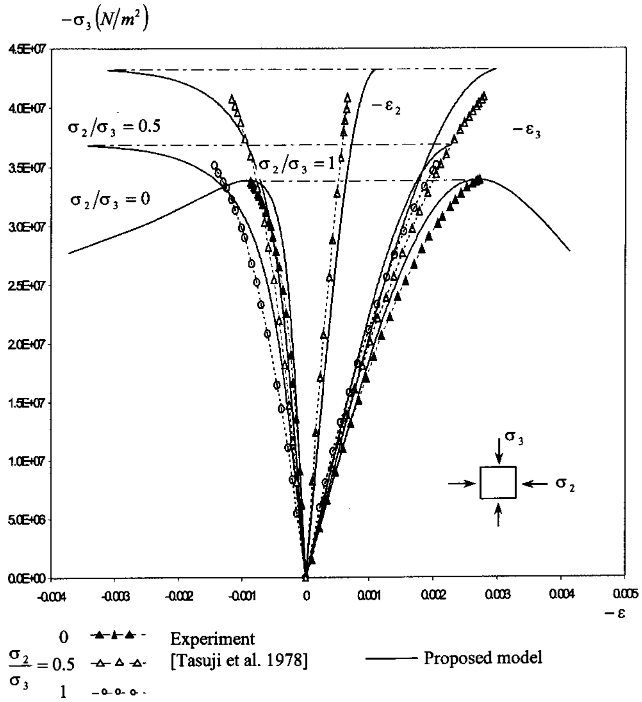


Fig. 7 Stress-strain response of concrete under biaxial compression

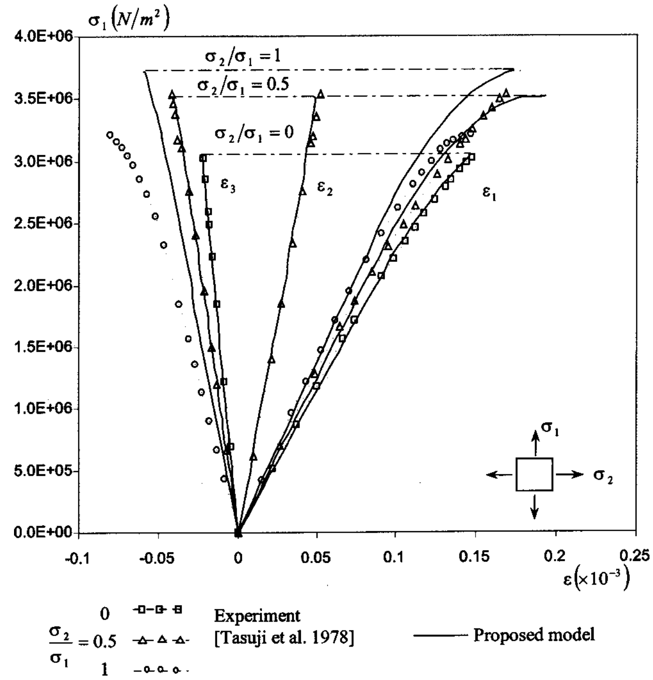


Fig. 9 Stress-strain response of concrete under biaxial tension

are smaller than that of the experiment under biaxial compression-tension. This may be attributed to the fact that in the experiment, friction exists on the surfaces under compression of the concrete specimen. This friction will have a restraining effect on the tension failure in the direction of σ_1. That is to say, the specimen is in fact not under perfect compression-tension loading, while ideal compression-tension loading is assumed in numerical simulation, and this effect of friction is not considered. Thus σ_1 will have a larger effect on the tension failure, which caused the difference between the two curves. This postulation is illustrated in Fig. 8 by cases $\sigma_1/\sigma_3 = -0.25$ and $\sigma_1/\sigma_3 = -0.05$. In the former case the effect of friction is smaller than that in the latter case, thus having a better agreement with the experimental result. Figure 9 shows that the numerical results under biaxial tension agree well with the experimental curves, except for the case of $\sigma_1/\sigma_2 = 1$, where the $\epsilon_3 \sim \sigma_1$ curves differ even during the elastic phase. The predicted biaxial ultimate strength envelope is shown in Fig. 10, and it also agrees with those reported by the other researchers [6–8].

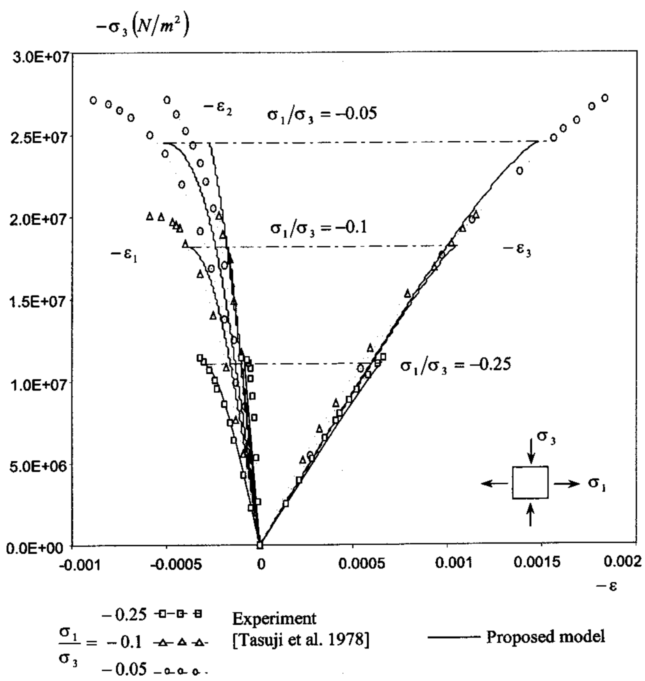


Fig. 8 Stress-strain response of concrete under biaxial compression tension

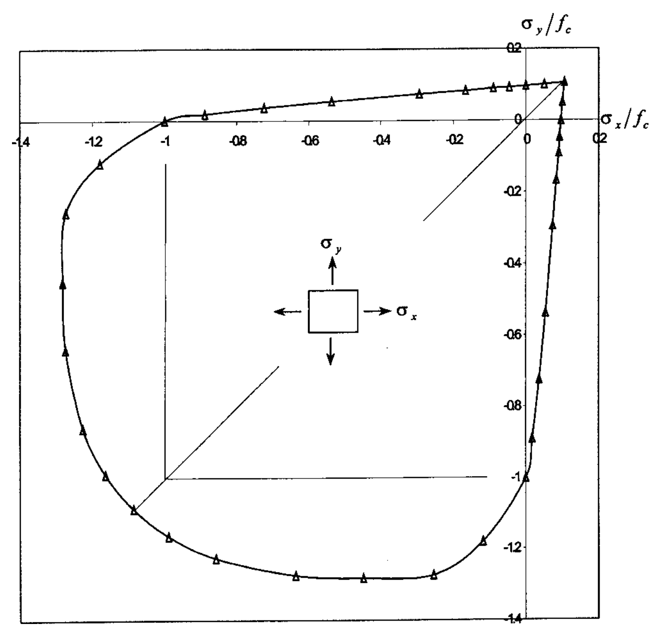


Fig. 10 The predicted biaxial ultimate strength envelope of concrete

5 Conclusions

This paper focuses on the application of the displacement equivalence-based damage model for brittle materials proposed in the preceding paper on theory [1]. The evolution rule of damage derived in the preceding paper using Onsager relations is applied to the cases of uniaxial and multiaxial loading. The loading-unloading condition is discussed and the tangent modulus is derived. The determination of the model parameters is then presented. The proposed model is applied to concrete under uniaxial and biaxial loading, and the numerical results are compared with those of other researchers and with the published experimental results. The results are generally in good agreement and the proposed model is considered to be worthy of further investigation.

However, the proposed model has been simplified because the material has been assumed to be totally brittle. This implies that there is only elastic damage, which leads to the degradation of the mechanical behaviors of the material, and no plasticity will occur in the material. Therefore it would be difficult to predict the response of brittle material under triaxial compression in which the plastic behavior is apparent, according to the experimental observation given in Ref. [4]. So in future work, coupling between the damage and plasticity (also under the assumption of displacement equivalence) should be considered.

In addition, another limitation of the proposed model is that the principal axis of damage should coincide with those of the stress and strain. To solve this problem, one could consider the shear stress and shear strain in the coordinate system of principal damage [9]. An alternative approach is to adopt the idea of rotating cracks in the smear crack model [10], i.e., to define a special damage evolution rule to force the principal axes of damage to coincide with those of the stress and strain.

Acknowledgments

We would like to express our sincere gratitude to the reviewers for their invaluable comments and suggestions given to this paper.

Appendix A: Derivation of Damage Evolution Rule Using the Nonassociated Method

First, decompose the ratio of strain energy density release

$$Y_i=\frac{\partial\varphi}{\partial D_i}=\frac{\partial\Sigma_{j=1}^{3}\varphi^j}{\partial D_i}=\sum_{j=1}^{3} Y_i^j,$$

$$Y_i^j=\frac{\partial\varphi^j}{\partial D_i}=\sigma_j\frac{\partial\epsilon_j}{\partial D_i}\ (\text{no summation for } j). \qquad (A1)$$

Then, introduce the nine flux potentials of damage, $g_i^j(Y,D)$, $i,j=1,2,3$:

$$g_i^j(Y,D)=\tfrac{1}{3}\alpha_j(1-D_i)^{\lambda_j}(Y_i^j)^2. \qquad (A2)$$

According to the normal flux rule, the increment of damage can be written as

$$dD_i^j=d\lambda_i^j\frac{\partial g_i^j}{\partial Y_i^j}=d\lambda_i^j\tfrac{2}{3}\alpha_j(1-D_i)^{\gamma_j}Y_i^j$$

$$=\tfrac{2}{3}\alpha_j(1-D_i)^{\gamma_j}\sigma_j\frac{\partial\epsilon_j}{\partial D_i}d\lambda_i^j,\quad i,j=1,2,3 \qquad (A3)$$

where $d\lambda_i^j$ are the damage consistency parameters which define the damage loading-unloading conditions according to the Kuhn-Tucker relations [11]

$$d\lambda_i^j\geqslant 0,\quad d\lambda_i^j(\epsilon_i-\max\{\epsilon_i^{cr},\epsilon_i^{\max}\})=0,\quad \epsilon_i-\max\{\epsilon_i^{cr},\epsilon_i^{\max}\}\leqslant 0. \qquad (A4)$$

So the damage evolution equation can be expressed as

$$dD_i=\sum_{j=1}^{3} dD_i^j=\sum_{j=1}^{3}\tfrac{2}{3}\alpha_j(1-D_i)^{\gamma_j}\sigma_j\frac{\partial\epsilon_j}{\partial D_i}d\lambda_i^j. \qquad (A5)$$

Equation (A5) is are the same as Eq. (25), as long as we set the damage consistency parameters to be

$$d\lambda_i^j=\begin{cases}|d\epsilon_j| & \text{when } \epsilon_i=\max\{\epsilon_i^{cr},\epsilon_i^{\max}\}\\ 0 & \text{when } \epsilon_i<\max\{\epsilon_i^{cr},\epsilon_i^{\max}\}\end{cases} \qquad (A6)$$

and reverse the sign of α_j, when $d\epsilon_j<0$.

Appendix B: Damage Value Corresponding to Tensile and Compressive Strength

For the case of uniaxial tension, from Eqs. (6) and (12), we can obtain

$$\sigma_1^3=\frac{3E^2}{\alpha}[C_t-(0.5-\gamma_t)^{-1}(1-D_1)^{0.5-\gamma_t}](1-D_1)^{4.5}, \qquad (B1)$$

$$\frac{d\sigma_1^3}{dD_1}=\frac{3E^2}{\alpha}(1-D_1)^{3.5}\{[4.5(0.5-\gamma_t)^{-1}+1]\times(1-D_1)^{1/2\gamma_t}-4.5C_t\}. \qquad (B2)$$

So when $D_1=1$ or

$$D_1=1-\left[\frac{4.5C_t}{4.5(0.5-\gamma_t)^{-1}+1}\right]^{1/(0.5-\gamma_t)},$$

σ_1^3 can reach its extreme value. $D_1=1$ represents a trivial solution and thus is discarded.

For the case of uniaxial compression, similar to the case of uniaxial tension, from Eqs. (14) and (18), we can obtain

$$\sigma_3^3=\frac{-3E^2}{\alpha\beta}[C_c-(2-3\beta-\gamma_c)^{-1}(1-D)^{2-3\beta-\gamma_c}](1-D)^{9\beta}, \qquad (B3)$$

$$\frac{d\sigma_3^3}{dD}=\frac{-3E^2}{\alpha\beta}(1-D)^{9\beta-1}[-9\beta C_c+9\beta(2-3\beta-\gamma_c)^{-1}\times(1-D_1)^{2-3\beta-\gamma_c}+(1-D)^{2-3\beta-\gamma_c}], \qquad (B4)$$

So when $D=1$ or

$$D=1-\left[\frac{9\beta C_c}{9\beta(2-3\beta-\gamma_c)^{-1}+1}\right]^{1/(2-3\beta-\gamma_c)},$$

σ_3^3 can reach its extreme value. $D=1$ represents a trivial solution and is also discarded.

References

[1] Soh, C. K., Liu, Y., Yang, Y., and Dong, Y., 2003, ''A Displacement Equivalence Based Damage Model for Brittle Materials—Part I: Theory,'' ASME J. Appl. Mech., **70**, pp. 1–7.
[2] Dong, L. L., Xie, H. P., and Zhao, P., 1995, ''Experimental Research on Complete Damage Process of Concrete Under Compression,'' J. Exp. Mech. (China), **10**, pp. 95–102 (in Chinese).
[3] Fonseka, G. U., and Krajcinovic, D., 1981, ''The Continuous Damage Theory of Brittle Materials- Part 2: Uniaxial and Plane Response Modes,'' ASME J. Appl. Mech., **48**, pp. 816–824.
[4] Li, Q. B., and Ansari, F., 1999, ''Mechanics of Damage and Constitutive Relationships for Concrete,'' J. Eng. Mech. Div., **125**, pp. 1–10.
[5] Dong, L. L., Xie, H. P., and Li, S. P., 1996, ''Continuum Damage Mechanics Constitutive Model of Concrete Under Compression,'' J. Eng. Mech. (China), **13**, pp. 44–53 (in Chinese).
[6] Tasuji, M. E., Slate, F. O., and Nilson, A. H., 1978, ''Stress-Strain Response and Fracture of Concrete in Biaxial Loading,'' ACI J., **7**, (July), pp. 306–312.
[7] Kupfer, H., and Hilsdorf, K., 1969, ''Behavior of Concrete Under Biaxial Stresses,'' ACI J., **66**, pp. 656–666.

[8] Liu, T. C. Y., Tony, C. Y., Nilson, A. H., and Slate, F. O., 1972, ''Stress-Strain Response and Fracture of Concrete in Uniaxial and Biaxial Compression,'' ACI J., **69**, pp. 291–295.
[9] Liu, Y., 2003, ''Computational Experiment of Reinforced Concrete Structural Elements Using Damage Mechanics,'' Ph.D. thesis, Nanyang Technological University, Singapore.
[10] Rots, J. G., 1998, ''Computational Modeling of Concrete Failure,'' thesis, Delft Univ. of Tech., Delft, the Netherlands.
[11] Ju, J. W., 1989, ''On Energy-Based Coupled Elastoplastic Damage Theories: Constitutive Modeling and Computational Aspects,'' Int. J. Solids Struct., **25**, pp. 803–833.

Stroh-Like Complex Variable Formalism for the Bending Theory of Anisotropic Plates

C. Hwu
Institute of Aeronautics and Astronautics,
National Cheng Kung University,
Tainan, Taiwan, Republic of China

Based upon the knowledge of the Stroh formalism and the Lekhnitskii formalism for two-dimensional anisotropic elasticity as well as the complex variable formalism developed by Lekhnitskii for plate bending problems, in this paper a Stroh-like formalism for the bending theory of anisotropic plates is established. The key feature that makes the Stroh formalism more attractive than the Lekhnitskii formalism is that the former possesses the eigenrelation that relates the eigenmodes of stress functions and displacements to the material properties. To retain this special feature, the associated eigenrelation and orthogonality relation have also been obtained for the present formalism. By intentional rearrangement, this new formalism and its associated relations look almost the same as those for the two-dimensional problems. Therefore, almost all the techniques developed for the two-dimensional problems can now be applied to the plate bending problems. Thus, many unsolved plate bending problems can now be solved if their corresponding two-dimensional problems have been solved successfully. To illustrate this benefit, two simple examples are shown in this paper. They are anisotropic plates containing elliptic holes or inclusions subjected to out-of-plane bending moments. The results are simple, exact and general. Note that the anisotropic plates treated in this paper consider only the homogeneous anisotropic plates. If a composite laminate is considered, it should be a symmetric laminate to avoid the coupling between stretching and bending behaviors.
[DOI: 10.1115/1.1600474]

1 Introduction

For two-dimensional linear anisotropic elasticity, there are two major complex variable formalisms in the literature. One is the Lekhnitskii formalism [1,2], which starts with the equilibrated stress functions followed by constitutive laws, strain-displacement relations, and compatibility equations; the other is the Stroh formalism [3,4], which starts with the compatible displacements followed by strain-displacement relations, constitutive laws, and equilibrium equations. A special feature of the Stroh formalism possesses an eigenrelation that relates the eigenmodes of stress functions and displacements to the material properties. Recently the Stroh formalism becomes more attractive than the Lekhnitskii formalism, especially when Ting [5] emphasized the Stroh formalism. In order to get benefits from both formalisms, there are also some works discussing the relations between these two formalisms, e.g., Refs. [6–11].

Besides the two-dimensional problems, around 60 years ago Lekhnitskii also developed a complex variable formalism for the plate bending problems [12]. After that, some researchers devoted their efforts to the development and application of the complex variable method on the laminates with the bending extension coupling such as Refs. [13–20]. In addition, the extension of the Stroh formalism to the coupled stretching-bending analysis of the composite laminates has been attempted by Lu [21], enhanced by Lu and Mahrenholtz [22] and modified by Cheng and Reddy [23]. Although there exist several different formalisms, due to the complexity, the resemblance between these formalisms and the Stroh formalism is not perfect enough to employ most of the key features of the Stroh formalism. The main difference of the existing formalisms with the Stroh formalism comes from the eigenrelation, which plays an important role for the Stroh formalism and will become a vital drawback if it cannot perfectly match that of the Stroh formalism.

Because the analogy between the in-plane problems and the plate bending problems has been observed long time ago, e.g., in Ref. [24], in this paper we turn back to the simple pure bending problems and try to develop a fully Stroh-like formalism in order to borrow all the techniques developed for the two-dimensional problems. By carefully reviewing Lekhnitskii formalism for both the two-dimensional and plate bending problems and catching the spirit of Stroh formalism for two-dimensional problems, we develop a Stroh-like formalism for the bending theory of anisotropic plates. Moreover, the explicit expressions of the material eigenvectors and fundamental matrix are also obtained in this paper. Again, it should be emphasized that this first attempt is valid only for homogeneous anisotropic plates which can be applied to the symmetric laminates not the general composite laminates. Better than those general formalisms developed in the literature, this new formalism looks very like the Stroh formalism for two-dimensional linear anisotropic elasticity. Hence, almost all the mathematical techniques developed for two-dimensional problems can lend to the plate bending problems. By simple analogy, many problems that cannot be solved previously now have the possibility to be solved even without a detailed derivation if their counterparts in two-dimensional problems have been solved.

Due to the success of the present formalism for the pure bending analysis, we have gone further to study how to improve the complex variable formalism for the coupled stretching-bending analysis in order to achieve a fully Stroh-like formalism. By comparing the difference of the present formalism with those presented in the literature such as Lu and Mahrenholtz [22] and Cheng and Reddy [23], we have found an alternative approach for developing the more Stroh-like formalism for the coupled stretching-bending analysis of general composite laminates [25].

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 3, 2002; final revision, Oct. 7, 2002. Associate Editor: J. R. Barber. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 Classical Bending Theory of Anisotropic Plates

In order to have a complete picture of the new formalism developed in this paper, the classical bending theory of anisotropic plates is briefly reviewed in this section [26,27]. In classical plate theory, the following assumptions, generally known as *Kirchhoff hypotheses*, are usually made [28]: (a) the material of the plate is elastic, homogeneous and isotropic; (b) the plate is initially flat; (c) the thickness of the plate is small compared to its other dimensions; (d) the deflections are small compared to the plate thickness; (e) the slopes of the deflected middle surface are small compared to unity; (f) the deformations are such that straight lines, initially normal to the middle surface, remain straight lines and normal to the middle surface, i.e., the deformations due to transverse shear will be neglected; (g) the stresses normal to the middle surface are of a negligible order of magnitude. For an anisotropic plate, all these assumptions remain valid except that the material of the plate is anisotropic instead of isotropic. Based upon these assumptions, the plate displacements u, v, and w in the x, y, and z directions can be expressed as

$$\begin{aligned} u(x,y,z)&=u_0(x,y)-z\frac{\partial w(x,y)}{\partial x},\\ v(x,y,z)&=v_0(x,y)-z\frac{\partial w(x,y)}{\partial y},\\ w(x,y,z)&=w_0(x,y), \end{aligned} \tag{1}$$

where u_0, v_0, and w_0 are, respectively, the middle surface displacements in the x, y, and z directions. If small deformations are considered, the strains of the plates can be written in terms of the middle surface displacements as follows:

$$\begin{aligned} \varepsilon_x&=\frac{\partial u}{\partial x}=\frac{\partial u_0}{\partial x}-z\frac{\partial^2 w}{\partial x^2},\\ \varepsilon_y&=\frac{\partial v}{\partial y}=\frac{\partial v_0}{\partial y}-z\frac{\partial^2 w}{\partial y^2},\\ \gamma_{xy}&=\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}=\frac{\partial u_0}{\partial y}+\frac{\partial v_0}{\partial x}-2z\frac{\partial^2 w}{\partial x\partial y}, \end{aligned} \tag{2a}$$

or, in matrix notation,

$$\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}_0+z\boldsymbol{\kappa}, \tag{2b}$$

where $\boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon}_0$, and $\boldsymbol{\kappa}$ denote, respectively, the strain vector, midsurface strain vector, and plate curvature vector, which are defined as

$$\boldsymbol{\varepsilon}=\begin{Bmatrix}\varepsilon_x\\ \varepsilon_y\\ \gamma_{xy}\end{Bmatrix}, \tag{3a}$$

$$\boldsymbol{\varepsilon}_0=\begin{Bmatrix}\varepsilon_x^0\\ \varepsilon_y^0\\ \gamma_{xy}^0\end{Bmatrix}=\begin{Bmatrix}\dfrac{\partial u_0}{\partial x}\\ \dfrac{\partial v_0}{\partial y}\\ \dfrac{\partial u_0}{\partial y}+\dfrac{\partial v_0}{\partial x}\end{Bmatrix}, \tag{3b}$$

$$\boldsymbol{\kappa}=\begin{Bmatrix}\kappa_x\\ \kappa_y\\ \kappa_{xy}\end{Bmatrix}=-\begin{Bmatrix}\dfrac{\partial^2 w}{\partial x^2}\\ \dfrac{\partial^2 w}{\partial y^2}\\ 2\dfrac{\partial^2 w}{\partial x\partial y}\end{Bmatrix}. \tag{3c}$$

Disregarding σ_z and assuming that the anisotropic materials have one plane of elastic symmetry located at the middle surface of the plate (or say, the material is monoclinic not the most general anisotropic), the stress-strain relationship of the anisotropic materials may then be separated into the following two parts:

$$\begin{Bmatrix}\sigma_x\\ \sigma_y\\ \tau_{xy}\end{Bmatrix}=\begin{bmatrix}\hat{C}_{11} & \hat{C}_{12} & \hat{C}_{16}\\ \hat{C}_{12} & \hat{C}_{22} & \hat{C}_{26}\\ \hat{C}_{16} & \hat{C}_{26} & \hat{C}_{66}\end{bmatrix}\begin{Bmatrix}\varepsilon_x\\ \varepsilon_y\\ \gamma_{xy}\end{Bmatrix},$$

$$\begin{Bmatrix}\tau_{yz}\\ \tau_{xz}\end{Bmatrix}=\begin{bmatrix}\hat{C}_{44} & \hat{C}_{45}\\ \hat{C}_{45} & \hat{C}_{55}\end{bmatrix}\begin{Bmatrix}\gamma_{yz}\\ \gamma_{xz}\end{Bmatrix}, \tag{4a}$$

or in matrix notation

$$\boldsymbol{\sigma}=\hat{\mathbf{C}}\boldsymbol{\varepsilon},\quad \boldsymbol{\sigma}_t=\hat{\mathbf{C}}_t\boldsymbol{\varepsilon}_t. \tag{4b}$$

In Eq. (4), $\hat{C}_{ij}$ is the reduced elastic stiffness defined as

$$\hat{C}_{ij}=C_{ij}-\frac{C_{i3}C_{3j}}{C_{33}}=\hat{C}_{ji},\quad i,j=1,2,4,5,6 \tag{5}$$

where C_{ij} is the elastic stiffness of the materials, which is also a contracted notation of the fourth-order elastic tensor C_{ijkl}. In engineering applications, one always likes to express the elastic stiffness in terms of engineering constants such as Young's moduli E, Poisson's ratios ν, and shear moduli G. These constants are usually measured in simple tests such as uniaxial tension or pure shear tests, which are all performed with a known load or stress. Thus, the components of the compliance matrix $\mathbf{S}$ (the inverse of the reduced elastic stiffness matrix $\hat{\mathbf{C}}^{-1}$) relating stresses and strains as $\boldsymbol{\varepsilon}=\mathbf{S}\boldsymbol{\sigma}$ are determined more directly than those of the stiffness matrix. For a general anisotropic material, the compliance matrix components in terms of the engineering constants are

$$\mathbf{S}=\hat{\mathbf{C}}^{-1}=\begin{bmatrix}\dfrac{1}{E_1} & \dfrac{-\nu_{21}}{E_2} & \dfrac{\eta_{1,12}}{G_{12}}\\ \dfrac{-\nu_{12}}{E_1} & \dfrac{1}{E_2} & \dfrac{\eta_{2,12}}{G_{12}}\\ \dfrac{\eta_{12,1}}{E_1} & \dfrac{\eta_{12,2}}{E_2} & \dfrac{1}{G_{12}}\end{bmatrix}, \tag{6}$$

where E_1 and E_2 are the Young's moduli in the x_1 and x_2 directions, respectively; ν_{ij} is Poisson's ratio for transverse strain in the x_j direction when stressed in the x_i direction, that is, $\nu_{ij}=-\varepsilon_j/\varepsilon_i$ for $\sigma_i=\sigma$ and all other stresses are zero; G_{12} is the shear modulus in the x_1x_2 plane; $\eta_{i,ij}$ is the coefficient of mutual influence of the first kind that characterizes stretching in the x_i direction caused by shear in the x_ix_j plane, that is, $\eta_{i,ij}=\varepsilon_i/\gamma_{ij}$ for $\tau_{ij}=\tau$ and all other stresses are zero; $\eta_{ij,i}$ is the coefficient of mutual influence of the second kind that characterizes shearing in the x_ix_j plane caused by a normal stress in the x_i direction, that is, $\eta_{ij,i}=\gamma_{ij}/\varepsilon_i$ for $\sigma_i=\sigma$, and all other stresses are zero. It is also known that the compliance (or stiffness) matrix is symmetric. Thus, $S_{ij}=S_{ji}$, or

$$\frac{\nu_{21}}{E_2}=\frac{\nu_{12}}{E_1},\quad \frac{\eta_{1,12}}{G_{12}}=\frac{\eta_{12,1}}{E_1},\quad \frac{\eta_{2,12}}{G_{12}}=\frac{\eta_{12,2}}{E_2}. \tag{7}$$

Note that the inverse of the elastic compliance matrix S_{ij} is the reduced elastic stiffness matrix $\hat{C}_{ij}$ instead of the elastic stiffness matrix C_{ij}.

Substituting Eq. (2b) into the stress-strain relation [first equation in (4b)], the stresses in the plates can also be written in terms of the middle surface strains $\boldsymbol{\varepsilon}_0$ and plate curvatures $\boldsymbol{\kappa}$ as

$$\begin{Bmatrix}\sigma_x\\ \sigma_y\\ \tau_{xy}\end{Bmatrix}=\boldsymbol{\sigma}=\hat{\mathbf{C}}(\boldsymbol{\varepsilon}_0+z\boldsymbol{\kappa}). \tag{8}$$

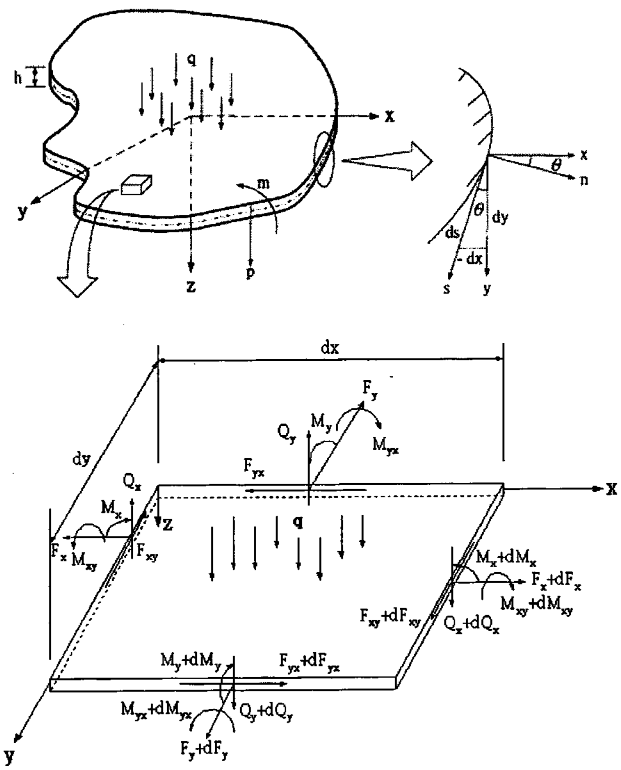


Fig. 1 Plate geometry, resultant forces, and moments

Since the thickness of the plate is considered to be small compared to its other dimensions, in classical plate theory an integral equivalent system of forces and moments acting on the plate cross section is used instead of spreading the stress distribution across the plate thickness. By integration of the stresses through the plate thickness, the resultant forces **F** and moments **M** acting on a plate cross section are defined as follows (Fig. 1):

$$\mathbf{F}=\begin{Bmatrix} F_x \\ F_y \\ F_{xy} \end{Bmatrix}=\int_{-h/2}^{h/2}\begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix}dz=\int_{-h/2}^{h/2}\boldsymbol{\sigma}\,dz,$$
$$\mathbf{M}=\begin{Bmatrix} M_x \\ M_y \\ M_{xy} \end{Bmatrix}=\int_{-h/2}^{h/2}\begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix}z\,dz=\int_{-h/2}^{h/2}\boldsymbol{\sigma}z\,dz, \tag{9}$$

where h is the thickness of the plate. Substituting Eq. (8) into (9), the resultant forces **F** and moments **M** can be written in terms of the middle surface strains $\boldsymbol{\varepsilon}_0$ and plate curvatures $\boldsymbol{\kappa}$ as

$$\mathbf{F}=h\hat{\mathbf{C}}\boldsymbol{\varepsilon}_0, \tag{10a}$$

$$\mathbf{M}=\mathbf{D}\boldsymbol{\kappa}, \tag{10b}$$

where **D** is the bending stiffness matrix that is related to $\hat{\mathbf{C}}$ by

$$\mathbf{D}=\frac{h^3}{12}\hat{\mathbf{C}}. \tag{11}$$

The results of Eq. (10) show that no coupling exists between the bending and extension of a plate, which is obviously due to the assumption of symmetry with respect to the middle surface of the plate. Hence, in the classical plate theory the bending problems are usually discussed separately with the in-plane problems. In the following we will then disregard the extensional forces and their associated deformations. When the plate is orthotropic and the directions of x and y axes coincide with the principal directions of elasticity, according to Eqs. (6) and (11) we have

$$D_{11}=\frac{E_1h^3}{12(1-\nu_{12}\nu_{21})},\quad D_{22}=\frac{E_2h^3}{12(1-\nu_{12}\nu_{21})},$$
$$D_{12}=\frac{\nu_{12}E_2h^3}{12(1-\nu_{12}\nu_{21})}, \tag{12}$$
$$D_{66}=\frac{G_{12}h^3}{12},\quad D_{16}=D_{26}=0.$$

In the case of an isotropic plate, we have

$$D_{11}=D_{22}=D_{12}+2D_{66}=\frac{Eh^3}{12(1-\nu^2)}=D,$$
$$D_{12}=\nu D,\quad D_{66}=\frac{1-\nu}{2}D,\quad D_{16}=D_{26}=0. \tag{13}$$

With the displacement fields, strain-displacement relations and constitutive laws given in Eqs. (1), (2), and (10) [or Eqs. (4) or (8)], to complete the structural analysis we need plate equilibrium equations. Since the equilibrium equations concern only the balance of forces acting upon the structures, it should be independent of the material types. For thin plates they are often developed by integration of the usual equilibrium equations of elasticity with respect to the coordinate z in thickness direction. By neglecting the body forces and the tractions on the top and bottom surfaces of the plate except the lateral load $q(x,y)$, the force and moment equilibrium equations of the plate can then be derived as

$$\frac{\partial Q_x}{\partial x}+\frac{\partial Q_y}{\partial y}+q=0, \tag{14a}$$

$$\frac{\partial M_x}{\partial x}+\frac{\partial M_{xy}}{\partial y}-Q_x=0, \tag{14b}$$

$$\frac{\partial M_{xy}}{\partial x}+\frac{\partial M_y}{\partial y}-Q_y=0, \tag{14c}$$

where Q_x and Q_y are the transverse shear forces defined as (Fig. 1)

$$\mathbf{Q}=\begin{Bmatrix} Q_x \\ Q_y \end{Bmatrix}=\int_{-h/2}^{h/2}\begin{Bmatrix} \tau_{xz} \\ \tau_{yz} \end{Bmatrix}dz. \tag{15}$$

Substitution of Eqs. (14b) and (14c) into Eq. (14a) will further lead the force equilibrium equation in the z direction to

$$\frac{\partial^2 M_x}{\partial x^2}+2\frac{\partial^2 M_{xy}}{\partial x\partial y}+\frac{\partial^2 M_y}{\partial y^2}+q=0. \tag{16}$$

With the moment-curvature relation (10b), the definition of curvature vector (3c), and the equilibrium equations (14b) and (14c), the moments and transverse shear forces can be expressed in terms of the lateral deflection as

$$M_x=-\left(D_{11}\frac{\partial^2 w}{\partial x^2}+D_{12}\frac{\partial^2 w}{\partial y^2}+2D_{16}\frac{\partial^2 w}{\partial x\partial y}\right),$$
$$M_y=-\left(D_{12}\frac{\partial^2 w}{\partial x^2}+D_{22}\frac{\partial^2 w}{\partial y^2}+2D_{26}\frac{\partial^2 w}{\partial x\partial y}\right), \tag{17a}$$
$$M_{xy}=-\left(D_{16}\frac{\partial^2 w}{\partial x^2}+D_{26}\frac{\partial^2 w}{\partial y^2}+2D_{66}\frac{\partial^2 w}{\partial x\partial y}\right),$$

$$Q_x=-\left[D_{11}\frac{\partial^3 w}{\partial x^3}+3D_{16}\frac{\partial^3 w}{\partial x^2\partial y}+(D_{12}+2D_{66})\frac{\partial^3 w}{\partial x\partial y^2}+D_{26}\frac{\partial^3 w}{\partial y^3}\right],$$

$$Q_y=-\left[D_{16}\frac{\partial^3 w}{\partial x^3}+(D_{12}+2D_{66})\frac{\partial^3 w}{\partial x^2\partial y}+3D_{26}\frac{\partial^3 w}{\partial x\partial y^2}+D_{22}\frac{\partial^3 w}{\partial y^3}\right]. \tag{17b}$$

Substituting Eq. (17*a*) into (16), the equilibrium equation can also be expressed in terms of the lateral deflection as

$$D_{11}\frac{\partial^4 w}{\partial x^4}+4D_{16}\frac{\partial^4 w}{\partial x^3\partial y}+2(D_{12}+2D_{66})\frac{\partial^4 w}{\partial x^2\partial y^2}+4D_{26}\frac{\partial^4 w}{\partial x\partial y^3}+D_{22}\frac{\partial^4 w}{\partial y^4}=q. \tag{18}$$

Equation (18) is the governing differential equation for deflection of anisotropic thin plates. To determine w through this equation, appropriate boundary conditions of the considered problems should be set properly. For a fourth-order differential equation, only two boundary conditions are required at each edge. These may be a given deflection and slope, or force and moment, or some combinations. Mathematically, they can be written as

$$\frac{\partial w}{\partial n}=\frac{\partial w^*}{\partial n}\ \text{ or }\ M_n=M_n^*\ \text{ or }\ M_n=k_m\frac{\partial w}{\partial n},$$

and (19)

$$w=w^*\ \text{ or }\ V_n=V_n^*\ \text{ or }\ V_n=k_v w,$$

where V_n is the well known *Kirchhoff force* of classical plate theory, or called the *effective transverse shear force* defined by

$$V_n=Q_n+\frac{\partial M_{nt}}{\partial t}. \tag{20}$$

The subscripts n and t denote, respectively, the directions normal and tangent to the boundary. The asterisk denotes the prescribed value. k_m and k_v are given spring constants. With the definition given in Eq. (20) and the expressions (17), we can also express the effective transverse shear force in terms of the lateral deflection as

$$V_x=Q_x+\frac{\partial M_{xy}}{\partial y}=-\left[D_{11}\frac{\partial^3 w}{\partial x^3}+4D_{16}\frac{\partial^3 w}{\partial x^2\partial y}+(D_{12}+4D_{66})\frac{\partial^3 w}{\partial x\partial y^2}+2D_{26}\frac{\partial^3 w}{\partial y^3}\right],$$

$$V_y=Q_y+\frac{\partial M_{xy}}{\partial x}=-\left[2D_{16}\frac{\partial^3 w}{\partial x^3}+(D_{12}+4D_{66})\frac{\partial^3 w}{\partial x^2\partial y}+4D_{26}\frac{\partial^3 w}{\partial x\partial y^2}+D_{22}\frac{\partial^3 w}{\partial y^3}\right]. \tag{21}$$

If θ denotes the angle between the normal n and x axis (Fig. 1), the values in the n-t coordinate can be calculated from the values in the x-y coordinate according to the following transformation laws:

$$\frac{\partial w}{\partial n}=\cos\theta\frac{\partial w}{\partial x}+\sin\theta\frac{\partial w}{\partial y}, \tag{22a}$$

$$M_n=\cos^2\theta M_x+\sin^2\theta M_y+2\sin\theta\cos\theta M_{xy},$$

$$M_t=\sin^2\theta M_x+\cos^2\theta M_y-2\sin\theta\cos\theta M_{xy}, \tag{22b}$$

$$M_{nt}=\sin\theta\cos\theta(M_y-M_x)+(\cos^2\theta-\sin^2\theta)M_{xy},$$

$$Q_n=\cos\theta Q_x+\sin\theta Q_y,$$

$$Q_t=-\sin\theta Q_x+\cos\theta Q_y. \tag{22c}$$

Among all the possible boundary conditions, the commonly used conditions such as simply supported, clamped, free, loaded, predisplaced and elastic supported edge boundary conditions can then be expressed as

$$\text{simply supported edge:}\quad w=0,\quad M_n=0; \tag{23a}$$

$$\text{clamped edge:}\quad w=0,\quad \frac{\partial w}{\partial n}=0; \tag{23b}$$

$$\text{free edge:}\quad V_n=0,\quad M_n=0; \tag{23c}$$

$$\text{loaded edge:}\quad V_n=p,\quad M_n=m; \tag{23d}$$

$$\text{predisplaced edge:}\quad w=w^*,\quad \frac{\partial w}{\partial n}=\alpha^*; \tag{23e}$$

$$\text{elastic supported edge:}\quad V_n=k_v w,\quad M_n=k_m\frac{\partial w}{\partial n}. \tag{23f}$$

3 Lekhnitskii's Complex Variable Formalism

To solve the governing differential equation (18) together with the boundary conditions (23) for deflection of anisotropic thin plate, Lekhnitskii [2] rewrote Eq. (18) symbolically with the use of four linear differential operators of the first order:

$$d_1d_2d_3d_4w=q. \tag{24a}$$

d_k $(k=1,2,3,4)$ designates the operation

$$d_k=\frac{\partial}{\partial y}-\mu_k\frac{\partial}{\partial x}, \tag{24b}$$

where μ_k are the roots of the characteristics equation

$$D_{22}\mu^4+4D_{26}\mu^3+2(D_{12}+2D_{66})\mu^2+4D_{16}\mu+D_{11}=0. \tag{24c}$$

It has been proved [12] that Eq. (24*c*) has no real roots for any elastic homogeneous material. Since the coefficients of the fourth-order equation for μ are real, there are two pairs of complex conjugates for μ. If we let

$$\operatorname{Im}\mu_k>0,\quad \mu_{k+2}=\overline{\mu_k},\quad k=1,2 \tag{25}$$

and assume that μ are distinct, the general solution for deflection w can be expressed as [2]

$$w=w_0+2\operatorname{Re}\{w_1(z_1)+w_2(z_2)\}, \tag{26a}$$

where Re and Im stand for the real and imaginary parts, respectively, and the overbar denotes the complex conjugate; w_0 is a particular solution of Eq. (18) whose form depends on the load distribution q on a plate surface; $w_1(z_1)$ and $w_2(z_2)$ are arbitrary analytical functions of complex variables

$$z_1=x+\mu_1 y\quad \text{and}\quad z_2=x+\mu_2 y. \tag{26b}$$

On the basis of Eqs. (17) and (21), general expressions for the moments and shear forces can be obtained as (for the case $\mu_1\neq\mu_2$) [2]:

$$M_x=M_x^0-2\operatorname{Re}\{\mu_1g_1w_1''(z_1)+\mu_2g_2w_2''(z_2)\}, \tag{27a}$$

$$M_y=M_y^0-2\operatorname{Re}\{h_1w_1''(z_1)+h_2w_2''(z_2)\}, \tag{27b}$$

$$M_{xy}=M_{xy}^0-2\operatorname{Re}\{r_1w_1''(z_1)+r_2w_2''(z_2)\}, \tag{27c}$$

$$Q_x=Q_x^0-2\operatorname{Re}\{\mu_1s_1w_1'''(z_1)+\mu_2s_2w_2'''(z_2)\}, \tag{27d}$$

$$Q_y = Q_y^0 + 2\,\mathrm{Re}\{s_1 w_1'''(z_1) + s_2 w_2'''(z_2)\}, \tag{27e}$$

$$V_x = V_x^0 + 2\,\mathrm{Re}\{h_1 \mu_1^2 w_1'''(z_1) + h_2 \mu_2^2 w_2'''(z_2)\}, \tag{27f}$$

$$V_y = V_y^0 + 2\,\mathrm{Re}\{g_1 w_1'''(z_1) + g_2 w_2'''(z_2)\}. \tag{27g}$$

Here $M_x^0, \ldots, Q_x^0, \ldots, V_y^0$ are the moments, transverse forces, and effective transverse forces corresponding to the particular solution of deflection w_0, which can be found by the relations given in Eqs. (17) and (21). The prime denotes differentiation with respect to the function argument z_k. The coefficients g_k, h_k, r_k, s_k, $k=1,2$, are defined as

$$g_k = \frac{D_{11}}{\mu_k} + D_{12}\mu_k + 2D_{16}, \quad h_k = D_{12} + D_{22}\mu_k^2 + 2D_{26}\mu_k,$$
$$r_k = D_{16} + D_{26}\mu_k^2 + 2D_{66}\mu_k, \tag{28}$$

$$s_k = \frac{D_{11}}{\mu_k} + 3D_{16} + (D_{12} + 2D_{66})\mu_k + D_{26}\mu_k^2, \quad k=1,2.$$

Through the use of the characteristic equation (24*c*), it can be proved that these coefficients possess the following relations:

$$s_k - r_k = g_k, \quad s_k + r_k = -h_k \mu_k, \quad k=1,2. \tag{29}$$

With the general solutions given in Eqs. (26) and (27), to get the complete solutions the only functions remained to be found are the complex functions $w_1(z_1)$ and $w_2(z_2)$, which should be determined through the satisfaction of boundary conditions. If the plate is subjected to bending only by forces and moments distributed along the edge and no transverse load is applied on the top or bottom surfaces (i.e., $q=0$), the boundary conditions of loaded and predisplaced edges shown in Eq. (23) can be replaced by [2]

$$2\,\mathrm{Re}\{g_1 w_1'(z_1) + g_2 w_2'(z_2)\} = -\int_0^s (m\,dy + f\,dx) - cx + c_1,$$
$$2\,\mathrm{Re}\{h_1 w_1'(z_1) + h_2 w_2'(z_2)\} = -\int_0^s (-m\,dx + f\,dy) + cy + c_2, \tag{30a}$$

and

$$2\,\mathrm{Re}\{w_1'(z_1) + w_2'(z_2)\} = -\frac{\partial w^*}{\partial s}\sin\theta + \alpha^*\cos\theta,$$
$$2\,\mathrm{Re}\{\mu_1 w_1'(z_1) + \mu_2 w_2'(z_2)\} = \frac{\partial w^*}{\partial s}\cos\theta + \alpha^*\sin\theta, \tag{30b}$$

where

$$f = \int_0^s p\,ds, \tag{30c}$$

and s is the arc length measured along a curved boundary; θ is the angle between the normal n and x axis (see Fig. 1).

Although Eqs. (30*a*)–(30*c*) are given in Lekhnitskii's book [2], only a reference written in Russian (Lekhnitskii [12]) is cited and no detailed explanation is given. Since their associated physical meaning is important for the following development, we now try to relate Eqs. (30*a*) and (30*b*) with the original boundary conditions (23*d*) and (23*e*).

Consider the loaded boundary condition for which $V_n = p$ and $M_n = m$ along the edges, as shown in Eq. (23*d*). At the first glance, it is hard to imagine why Lekhnitskii used Eq. (30*a*) to replace (23*d*). Looking at the definition of effective transverse shear force (20), the transformation law (22) and the complex variable expressions (27), we found that it is really complicated to describe the loaded conditions for curved boundaries by the usual way. For elasticity problems, the loaded conditions are usually called traction boundary conditions in which the tractions t_i are prescribed along the boundaries. Using Cauchy's formula $t_i = \sigma_{ij} n_j$, and introducing the stress functions ϕ_i such that $\sigma_{i1} = -\phi_{i,2}$ and $\sigma_{i2} = \phi_{i,1}$ for two-dimensional problems, a very simple relation $t_i = \partial\phi_i/\partial s$ has been found. This simple relation is very helpful for the description of traction conditions especially along the curved boundaries. Therefore, we now try to express the loaded boundary conditions by modifying the concept of traction boundary conditions used in two-dimensional (2D) elasticity.

By following the concept of 2D elasticity and considering the non-lateral load condition ($q=0$), the force and moment equilibrium equations of the plate shown in Eqs. (14) can be rewritten as

$$\frac{\partial^2(2M_{xy} - H_{xy} - H_{xy}^*)}{\partial x \partial y} = 0, \quad \frac{\partial M_x}{\partial x} + \frac{\partial H_{xy}}{\partial y} = 0,$$
$$\frac{\partial H_{xy}^*}{\partial x} + \frac{\partial M_y}{\partial y} = 0, \tag{31a}$$

where

$$\frac{\partial H_{xy}}{\partial y} = \frac{\partial M_{xy}}{\partial y} - Q_x, \quad \frac{\partial H_{xy}^*}{\partial x} = \frac{\partial M_{xy}}{\partial x} - Q_y. \tag{31b}$$

The second and third moment equilibrium equations (31*a*) will be satisfied automatically if we introduce the stress functions ϕ_1 and ϕ_2 such that

$$\frac{\partial \phi_1}{\partial x} = -M_y, \quad \frac{\partial \phi_1}{\partial y} = H_{xy}^*,$$
$$\frac{\partial \phi_2}{\partial y} = -M_x, \quad \frac{\partial \phi_2}{\partial x} = H_{xy}. \tag{32}$$

Similar to the surface traction $t_i\,(=\sigma_{ij}n_j)$, we now define the surface moment M_i by

$$M_1 = M_x n_1 + H_{xy} n_2, \quad M_2 = H_{xy}^* n_1 + M_y n_2. \tag{33}$$

From Fig. 1, we see that

$$n_1 = \cos\theta = \partial y/\partial s, \quad n_2 = \sin\theta = -\partial x/\partial s. \tag{34}$$

Substituting the relations given in Eqs. (32) and (34) into (33), we get

$$M_1 = -\frac{\partial \phi_2}{\partial s}, \quad M_2 = \frac{\partial \phi_1}{\partial s}. \tag{35}$$

Like the surface traction, the definitions given in Eq. (33) are the components of surface moment vector in the *x* and *y* directions. To find the component of surface moment in the direction normal to the boundary, we may use the transformation law for vectors, i.e., $M_1\cos\theta + M_2\sin\theta$. With the equations given in Eqs. (33), (34), and the first relations in Eqs. (31*a*) and (22*b*), it can be proved that

$$M_1\cos\theta + M_2\sin\theta = \cos^2\theta M_x + \sin^2\theta M_y + 2\sin\theta\cos\theta M_{xy}$$
$$= M_n. \tag{36}$$

Substituting Eq. (35) into (36), we can express M_n in terms of the stress functions as

$$M_n = \frac{\partial \phi_1}{\partial s}\sin\theta - \frac{\partial \phi_2}{\partial s}\cos\theta. \tag{37}$$

Similarly, we can prove that

$$V_n = \frac{\partial}{\partial s}(-M_1\sin\theta + M_2\cos\theta) = \frac{\partial}{\partial s}\left(\frac{\partial \phi_1}{\partial s}\cos\theta + \frac{\partial \phi_2}{\partial s}\sin\theta\right). \tag{38}$$

Using the relations obtained in Eqs. (37) and (38), and knowing that $dy = \cos\theta ds$, $dx = -\sin\theta ds$, we now prove that

$$\int_0^s (M_n dy + \tilde{V}_n dx) = -\phi_2(s) + \phi_2(0),$$

$$\int_0^s (-M_n dx + \tilde{V}_n dy) = \phi_1(s) - \phi_1(0), \tag{39a}$$

where

$$\tilde{V}_n = \int_0^s V_n ds. \tag{39b}$$

This result (39) together with the definition (32) and the expressions (27*a*) and (27*b*) can then help us to prove that the loaded boundary condition (23*d*) may be replaced by Eq. (30*a*). From this derivation, we know that (30*a*) stands for the prescribed conditions of the stress functions ϕ_1 and ϕ_2 which are defined in Eq. (32).

After proving the replacement of loaded boundary condition, we consider the pre-displaced boundary condition for which $w = w^*$ and $\partial w/\partial n = \alpha^*$ along the edges, as shown in Eq. (23*e*). Since $w = w^*$ along the edges, it leads to $\partial w/\partial s = \partial w^*/\partial s$ along the edges. Hence, the predisplaced boundary condition may be replaced by $\partial w/\partial s = \partial w^*/\partial s$ and $\partial w/\partial n = \alpha^*$, which means that both of the slopes in the tangential and normal directions of the boundary are prescribed. According to the transformation law of vectors, this condition can be further replaced by

$$\frac{\partial w}{\partial x} = \alpha^* \cos\theta - \frac{\partial w^*}{\partial s}\sin\theta, \quad \frac{\partial w}{\partial y} = \alpha^* \sin\theta + \frac{\partial w^*}{\partial s}\cos\theta. \tag{40}$$

Substituting Eq. (26) with $q=0$ into (40), we obtain Eq. (30*b*). From this derivation, we know that Eq. (30*b*) stands for the prescribed conditions of slopes in the x and y directions.

4 Stroh-Like Complex Variable Formalism

The complex variable formalism shown in the preceding section was developed by Lekhnitskii about 60 years ago. A very similar but more popular formalism also developed by Lekhnitskii is for two-dimensional linear anisotropic elasticity [1]. Another important complex variable formalism is called the Stroh formalism, which is now applied mainly to two-dimensional problems. The Stroh formalism can be traced to the work of Eshelby et al. [29]. However, it was named after Stroh because he made a major contribution about the establishment of a material eigenrelation, which let the Stroh formalism become more elegant and powerful than the Lekhnitskii formalism, and laid the foundations for researchers who followed him [3,4]. Recently, an important book about anisotropic elasticity was written by Ting [5], who reorganized the Stroh formalism and developed many important properties that were then used to solve many elasticity problems. During these few years, there are also some works discussing the relations between the Lekhnitskii formalism and Stroh formalism [6–11].

Although there are many contributions related to Lekhnitskii and Stroh formalisms, most of them are for two-dimensional problems. It is rare to see any contribution using Lekhnitskii's complex variable formalism described in the preceding section to solve the plate bending problems. The main reason for this low usage is possibly due to its mathematical difficulties. To remedy this, one may consider its counterpart of the Stroh formalism. In this paper, we try to reorganize the Lekhnitskii formalism into a Stroh-like formalism and hope that the merits of the Stroh formalism will make the solving of plate bending problems become easier.

In order to reorganize the Lekhnitskii formalism into a Stroh-like formalism, one should first know the Stroh formalism for two-dimensional anisotropic elasticity. For the convenience of the readers, the Stroh formalism is summarized in Appendix A. The key difference that makes the Stroh formalism more attractive than the Lekhnitskii formalism is that the former is presented in matrix form instead of scalar form and its associated eigenrelation and orthogonality relation provide many useful identities that make the complex variable mathematical manipulation much easier. Therefore, to be successful in forming a Stroh-like formalism for plate bending problems, the selection of the proper stress function vector and displacement vector is important. To be convenient for problem solving, this selection should be consistent with the expressions of boundary conditions. By viewing the boundary conditions of loaded and predisplaced edges shown in Eq. (30) and the derivation presented in Eqs. (31)–(40), we now introduce the stress function vector $\boldsymbol{\phi}$ and the slope vector $\boldsymbol{\alpha}$ as

$$\boldsymbol{\phi} = \begin{Bmatrix} \phi_1 \\ \phi_2 \end{Bmatrix} = \begin{Bmatrix} -\int M_y dx \\ -\int M_x dy \end{Bmatrix}, \quad \boldsymbol{\alpha} = \begin{Bmatrix} \alpha_1 \\ \alpha_2 \end{Bmatrix} = \begin{Bmatrix} -\dfrac{\partial w}{\partial y} \\ \dfrac{\partial w}{\partial x} \end{Bmatrix}. \tag{41}$$

It looks odd to put the minus sign and have the order of (y,x) instead of (x,y) in introducing new vectors (41). However, in order that our final solution form can match with the Stroh formalism for two-dimensional problems, this selection is necessary.

Employing these two new vectors, the general solutions for the plate bending problems given in Eqs. (26) and (27) can now be expressed in the following Stroh-like form,

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + 2\,\mathrm{Re}\{\mathbf{B}\mathbf{w}'(z)\}, \quad \boldsymbol{\phi} = \boldsymbol{\phi}_0 + 2\,\mathrm{Re}\{\mathbf{A}\mathbf{w}'(z)\}, \tag{42a}$$

where $\boldsymbol{\alpha}_0$ and $\boldsymbol{\phi}_0$ are the particular solutions related to the lateral load distribution q; $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{w}$ are defined as follows:

$$\mathbf{A} = [\mathbf{a}_1 \ \ \mathbf{a}_2], \quad \mathbf{B} = [\mathbf{b}_1 \ \ \mathbf{b}_2], \quad \mathbf{w}(z) = \begin{Bmatrix} w_1(z_1) \\ w_2(z_2) \end{Bmatrix}, \tag{42b}$$

where

$$\mathbf{a}_k = \begin{Bmatrix} h_k \\ g_k \end{Bmatrix}, \quad \mathbf{b}_k = \begin{Bmatrix} -\mu_k \\ 1 \end{Bmatrix}, \quad k = 1,2. \tag{42c}$$

The loaded and predisplaced boundary conditions (23*d*) and (23*e*) are replaced by Eqs. (30*a*) and (30*b*), which can also be written in matrix form as

$$\boldsymbol{\phi} = \begin{Bmatrix} \int_0^s (-m dx + f dy) + cy + c_2 \\ -\int_0^s (m dy + f dx) - cx + c_1 \end{Bmatrix}, \quad \boldsymbol{\alpha} = \begin{Bmatrix} -\dfrac{\partial w^*}{\partial s}\cos\theta - \alpha^* \sin\theta \\ -\dfrac{\partial w^*}{\partial s}\sin\theta + \alpha^* \cos\theta \end{Bmatrix}. \tag{43}$$

With the expressions (43), the commonly encountered boundary conditions (23*b*) and (23*c*) can now be written as

$$\text{clamped edge:} \quad \boldsymbol{\alpha} = \mathbf{0}; \qquad \text{free edge:} \quad \boldsymbol{\phi} = \mathbf{0}. \tag{44}$$

As to the mixed boundary conditions like the simply supported edge (23*a*) and the elastic supported edge (23*f*), no simple vector form expressions can be obtained. The same situation occurs for two-dimensional problems, which are usually solved by using the component expressions. In this sense, relations (37) and (38) may be utilized to express the moment and effective transverse shear force, while the deflection and slope may be expressed with the assistance of Eqs. (22*a*) and (26). For example, the boundary condition of simply supported edges can be written as

$$\text{simply supported edge:} \quad w = \frac{\partial \phi_1}{\partial s}\sin\theta - \frac{\partial \phi_2}{\partial s}\cos\theta = 0. \tag{45}$$

5 Eigenrelation

The key feature that makes the Stroh formalism more attractive than the Lekhnitskii formalism is that it possesses the eigenrelation that relates the eigenmodes of stress functions and displacements to the material properties. Without the eigenrelation, one cannot feel the benefit of the Stroh formalism. Therefore, it is important for us to establish the eigenrelation for the Stroh-like formalism developed in the preceding section. Otherwise, the formalism is just a skeleton without any spirit inside its body. In order to establish the eigenrelation for the Stroh-like formalism of plate bending problems, it is better for us to know the eigenrelation for the Stroh formalism of two-dimensional problems and its corresponding characteristic equation in Lekhnitskii formalism (see Appendixes A and B).

In general, the two-dimensional problems considered in anisotropic plates include not only in-plane but also antiplane problems and the problems where in-plane and antiplane deformations couple each other. For this general case, the characteristic equation associated with anisotropic materials is a sixth-order algebraic equation. In this paper, we only consider anisotropic plates having one plane of elastic symmetry located at the middle surface, i.e., the monoclinic plates. For this case, the in-plane and antiplane problems will decouple. The characteristic equation for the in-plane problems is a fourth-order algebraic equation, while that for the antiplane problems is a second-order equation. For the bending problems considered in this paper, a fourth-order characteristic equation is also obtained in Eq. (24*c*). Therefore, it is better for us to perform the comparison through in-plane problems for monoclinic plates not through general two-dimensional problems for anisotropic plates.

By comparing Eqs. (24*c*) with (*B*12), we find that these two characteristic equations will be equivalent if we make the following replacements:

$$\begin{aligned}&\hat{S}_{11}\leftrightarrow D_{22},\quad \hat{S}_{22}\leftrightarrow D_{11},\quad \hat{S}_{12}\leftrightarrow D_{12},\\ &\hat{S}_{16}\leftrightarrow -2D_{26},\quad \hat{S}_{26}\leftrightarrow -2D_{16},\quad \hat{S}_{66}\leftrightarrow 4D_{66}.\end{aligned} \tag{46}$$

A simple rule for the above replacement is that when we want to change $\hat{S}$ into D the subscripts 1 and 2 should be interchanged and the symbol D with subscript 6 should be multiplied by -2. According to this rule, we observe that the eigenvectors for the in-plane problems given in Eq. (*B*13) with g_k and h_k defined by Eq. (*B*11) can be successfully transformed to the eigenvectors for the plate bending problems given in Eq. (42*c*) with g_k and h_k defined by Eq. (28). By referring to the eigenrelation for the in-plane problems, the eigenrelation for the plate bending problems can be written by using exactly the same form as Eqs. (*A*11)–(*A*13), i.e.,

$$\mathbf{N}\boldsymbol{\xi}=\mu\boldsymbol{\xi}, \tag{47a}$$

where

$$\mathbf{N}=\begin{bmatrix}\mathbf{N}_1 & \mathbf{N}_2\\ \mathbf{N}_3 & \mathbf{N}_1^T\end{bmatrix},\qquad \boldsymbol{\xi}=\begin{Bmatrix}\mathbf{a}\\ \mathbf{b}\end{Bmatrix}, \tag{47b}$$

and

$$\mathbf{N}_1=-\mathbf{T}^{-1}\mathbf{R}^T,\quad \mathbf{N}_2=\mathbf{T}^{-1}=\mathbf{N}_2^T,\quad \mathbf{N}_3=\mathbf{R}\mathbf{T}^{-1}\mathbf{R}^T-\mathbf{Q}=\mathbf{N}_3^T. \tag{47c}$$

The three 2×2 real matrices $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{T}$ defined in Eq. (*A*18) for the in-plane problems are written in terms of the elastic constants C_{ij}, which are the inverse of the reduced elastic compliances $\hat{S}_{ij}$. Therefore, to construct $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{T}$ for the plate bending problems, we need to use the inverse of D_{ij}. According to the definition of bending stiffness (11) and the inversion relation (6), the inverse of D_{ij} can be written as

$$\mathbf{D}^{-1}=\begin{bmatrix}D_{11}^* & D_{12}^* & D_{16}^*\\ D_{12}^* & D_{22}^* & D_{26}^*\\ D_{16}^* & D_{26}^* & D_{66}^*\end{bmatrix}=\frac{12}{h^3}\hat{\mathbf{C}}^{-1}=\frac{12}{h^3}\mathbf{S}. \tag{48}$$

To match with the rule (46), we reorganize the inverse relation $\mathbf{D}\mathbf{D}^{-1}=\mathbf{I}$ into

$$\begin{bmatrix}D_{22} & D_{12} & -2D_{26}\\ D_{12} & D_{11} & -2D_{16}\\ -2D_{26} & -2D_{16} & 4D_{66}\end{bmatrix}\times\begin{bmatrix}D_{22}^* & D_{12}^* & -\frac{1}{2}D_{26}^*\\ D_{12}^* & D_{11}^* & -\frac{1}{2}D_{16}^*\\ -\frac{1}{2}D_{26}^* & -\frac{1}{2}D_{16}^* & \frac{1}{4}D_{66}^*\end{bmatrix}=\mathbf{I}. \tag{49}$$

With this relation and the definition (*A*18) for the in-plane problems, we know that $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{T}$ for the plate bending problems should be defined as follows:

$$\mathbf{Q}=\begin{bmatrix}D_{22}^* & -\frac{1}{2}D_{26}^*\\ -\frac{1}{2}D_{26}^* & \frac{1}{4}D_{66}^*\end{bmatrix},\quad \mathbf{R}=\begin{bmatrix}-\frac{1}{2}D_{26}^* & D_{12}^*\\ \frac{1}{4}D_{66}^* & -\frac{1}{2}D_{16}^*\end{bmatrix},$$
$$\mathbf{T}=\begin{bmatrix}\frac{1}{4}D_{66}^* & -\frac{1}{2}D_{16}^*\\ -\frac{1}{2}D_{16}^* & D_{11}^*\end{bmatrix}. \tag{50}$$

In order to check whether the eigenvalues and eigenvectors of the eigenrelation (47) with $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{T}$ defined in Eq. (50) are equivalent to those obtained in Eqs. (24*c*) and (42*c*), we first try to get the explicit expressions of $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_3$. By following the steps described in Ting's book [5], we obtain (see Appendix C)

$$\mathbf{N}_1=\begin{bmatrix}\dfrac{-2D_{26}}{D_{22}} & -1\\ \dfrac{D_{12}}{D_{22}} & 0\end{bmatrix},$$
$$\mathbf{N}_2=\begin{bmatrix}4D_{66}-\dfrac{4D_{26}^2}{D_{22}} & -2D_{16}+\dfrac{2D_{12}D_{26}}{D_{22}}\\ -2D_{16}+\dfrac{2D_{12}D_{26}}{D_{22}} & D_{11}-\dfrac{D_{12}^2}{D_{22}}\end{bmatrix}, \tag{51}$$
$$\mathbf{N}_3=\begin{bmatrix}\dfrac{-1}{D_{22}} & 0\\ 0 & 0\end{bmatrix}.$$

Substituting Eq. (51) into Eqs. (47*a*) and (47*b*), we can find that $\|\mathbf{N}-\mu\mathbf{I}\|=0$ is equivalent to Eq. (24*c*), and its associated eigenvector is the one written in Eq. (42*c*) where g_k and h_k are defined in Eq. (28).

6 Orthogonality Relation

The eigenrelation shown in Eq. (47) has exactly the same form as that for two-dimensional problems. The only difference between plate bending problems and two-dimensional problems is the definition of $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{T}$, which is Eq. (*A*14) for general two-dimensional problems, is Eq. (*A*18) for in-plane problems, and is Eq. (50) for plate bending problems. With this understanding, it is expected to have the same orthogonatlity relation as that for two-dimensional problems. That is (see Appendix D),

$$\begin{bmatrix} \mathbf{B}^T & \mathbf{A}^T \\ \overline{\mathbf{B}}^T & \overline{\mathbf{A}}^T \end{bmatrix} \begin{bmatrix} \mathbf{A} & \overline{\mathbf{A}} \\ \mathbf{B} & \overline{\mathbf{B}} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{52}$$

which is also the *orthogonality relation* for the plate bending problems. The two 4×4 matrices on the left side of Eq. (52) are the inverse of each other, and hence their products commute, i.e.,

$$\begin{bmatrix} \mathbf{A} & \overline{\mathbf{A}} \\ \mathbf{B} & \overline{\mathbf{B}} \end{bmatrix} \begin{bmatrix} \mathbf{B}^T & \mathbf{A}^T \\ \overline{\mathbf{B}}^T & \overline{\mathbf{A}}^T \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{53a}$$

or

$$\begin{aligned} &\mathbf{AB}^T + \overline{\mathbf{AB}}^T = \mathbf{I} = \mathbf{BA}^T + \overline{\mathbf{BA}}^T, \\ &\mathbf{AA}^T + \overline{\mathbf{AA}}^T = \mathbf{0} = \mathbf{BB}^T + \overline{\mathbf{BB}}^T. \end{aligned} \tag{53b}$$

From the relations obtained in Eq. (53*b*), it can be observed that the following three matrices $\mathbf{S}_1$, $\mathbf{H}_2$, and $\mathbf{L}_3$ are real, which appear often in the final solutions to two-dimensional anisotropic elasticity problems [5]:

$$\mathbf{S}_1 = i(2\mathbf{AB}^T - \mathbf{I}), \quad \mathbf{H}_2 = 2i\mathbf{AA}^T, \quad \mathbf{L}_3 = -2i\mathbf{BB}^T. \tag{54}$$

Usually the eigenvectors defined in Eq. (42*c*) may vary up to an arbitrary multiplier. However, if the orthogonality relation (52) or (53) is used in applications, normalization of eigenvectors are needed because the unit matrix $\mathbf{I}$ is employed in Eq. (52) or (53). By multiplying the eigenvectors $\mathbf{a}_k$ and $\mathbf{b}_k$ given in Eq. (42*c*) with a normalization factor c_k, and using the orthogonal relation (52) or (53), one may obtain the normalized eigenvectors as

$$\mathbf{a}_k = c_k \begin{Bmatrix} h_k \\ g_k \end{Bmatrix}, \quad \mathbf{b}_k = c_k \begin{Bmatrix} -\mu_k \\ 1 \end{Bmatrix}, \tag{55a}$$

where

$$c_k^2 = \frac{1}{2(g_k - \mu_k h_k)}, \quad k = 1,2. \tag{55b}$$

Since the eigenrelation (47) and the orthogonality relation (52) or (53) have exactly the same form as those for two-dimensional problems, all the identities developed based upon these two relations in the Stroh formalism for two-dimensional anisotropic elasticity should all be valid for the plate bending problems discussed in this paper. One may refer to Ting [5,30] for various useful identities.

The normalization given in Eq. (55) means that each component of $\mathbf{w}(z)$ defined in Eq. (42*b*) has a factor change and Eq. (26*a*) should be modified as

$$w = w_0 + 2\,\mathrm{Re}\{c_1 w_1(z_1) + c_2 w_2(z_2)\} \tag{56}$$

where c_1 and c_2 are the normalization factors given in Eq. (55*b*).

7 Examples

Before dealing with the examples shown in this section, we like to emphasize that the main purpose of this paper is trying to develop a fully Stroh-like formalism for the bending analysis of anisotropic plates and hope that the same concept can be extended to the more complicated problems like the coupled stretching-bending problems. Because even the coupled stretching-bending problems have been dealt with in the literature such as (Lu and Mahrenholtz [22]), repetition of the example solving exercise especially for the more simple cases discussed in this section is meaningless. The meaningful insight of the following illustrations is that the solution process as well as the final solution form are exactly the same as those of the solution of two-dimensional problems by the Stroh formalism. Therefore, after viewing the examples one should have more confidence in what we said, "By simple analogy of our present formalism with the Stroh formalism for two-dimensional problems, the pure bending problems for the homogeneous anisotropic plates can be solved directly even without detailed derivation if their counterparts in two-dimensional problems have been solved previously, or vice versa." In this sense, the examples given below are solved briefly and most of the critical steps (such as the choice of the unknown complex functions, the solving of the unknown coefficients, and the conversion to the real form solutions) are dealt with by referring to their counterparts of the two-dimensional problems and no detailed explanations are provided in this paper. For those who are interested in the detailed explanations of all these critical steps, please refer to Ting's book [5] for anisotropic elasticity or my previous related works [31–36].

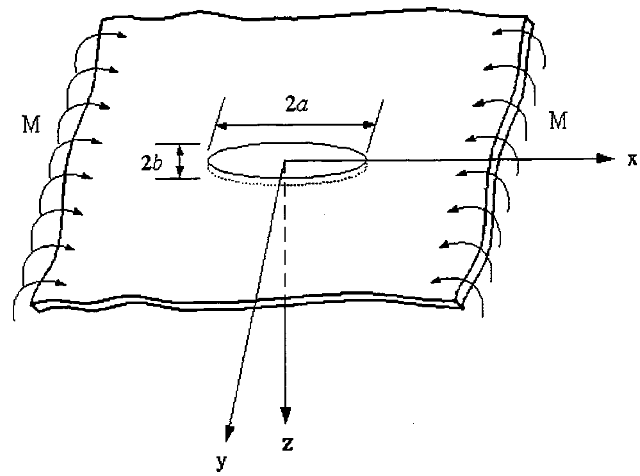


Fig. 2 An anisotropic plate weakens by an elliptical hole subjected to out-of-plane bending moments

7.1 Example 1: An Anisotropic Plate Weakened by an Elliptical Hole Subjected to Out-of-Plane Bending Moments

An anisotropic plate weaken by an elliptical hole is deflected under uniformly distributed moments M as shown in Fig. 2. There is no load around the edge of the hole. The diameter of the hole is considered to be small in comparison with the length of the plate sides, and the hole is situated far from the edges. Thus, the plate is considered as infinite. The contour of the hole is represented by

$$x = a\cos\psi, \quad y = b\sin\psi, \tag{57}$$

where $2a$ and $2b$ are the major and minor axes of the ellipse and ψ is a real parameter. By using the stress function vector $\boldsymbol{\phi}$ and slope vector $\boldsymbol{\alpha}$ introduced in Eq. (41), the boundary conditions of this problem can be expressed as

$$\begin{aligned} &\boldsymbol{\phi} = -My\mathbf{i}_2, \quad \text{at infinity}, \\ &\boldsymbol{\phi} = \mathbf{0}, \quad \text{around the hole boundary}, \end{aligned} \tag{58a}$$

where

$$\mathbf{i}_2 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}. \tag{58b}$$

Note that the first relation in Eq. (58*a*) is obtained by using $M_x = M$, $M_y = M_{xy} = 0$ at infinity and the definitions given in Eq. (32).

Since no lateral load is applied on the plate, the particular solution $\boldsymbol{\phi}_0$ and $\boldsymbol{\alpha}_0$ of Eq. (42*a*) is set to be zero. In order to satisfy the boundary condition at infinity, part of the homogenous solution $\mathbf{w}'(z)$ proportional to $\mathbf{z}$ is separated and denoted by $\boldsymbol{\phi}^\infty$ and $\boldsymbol{\alpha}^\infty$, which is then added to the general solution (42*a*) as

$$\boldsymbol{\phi} = \boldsymbol{\phi}^\infty + 2\,\mathrm{Re}\{\mathbf{Aw}'(z)\}, \quad \boldsymbol{\alpha} = \boldsymbol{\alpha}^\infty + 2\,\mathrm{Re}\{\mathbf{Bw}'(z)\}, \tag{59a}$$

where

$$\boldsymbol{\phi}^{\infty}=-My\mathbf{i}_2,\quad \boldsymbol{\alpha}^{\infty}=M(\mathbf{d}_2 y-\mathbf{d}_1 x), \tag{59b}$$

and

$$\mathbf{d}_1=\begin{Bmatrix}-D_{16}^*/2\\ D_{11}^*\end{Bmatrix},\quad \mathbf{d}_2=\begin{Bmatrix}D_{12}^*\\ -D_{16}^*/2\end{Bmatrix}. \tag{59c}$$

In Eqs. (59b) and (59c), $\boldsymbol{\phi}^{\infty}$ is selected to be match with the boundary condition at infinity, while $\boldsymbol{\alpha}^{\infty}$ is calculated by integrating the curvature vector $\boldsymbol{\kappa}$ from Eq. (10b) with $M_x=M$, $M_y=M_{xy}=0$ and using the symbol given in Eq. (48) for the inverse of the bending stiffness. Note that the constant terms of $\boldsymbol{\alpha}$ and $\boldsymbol{\phi}$ have been neglected since they have no contribution to the internal stresses of the plates.

To satisfy the free edge condition around the hole boundary shown in the second relation in Eq. (58a), by referring to the solutions of the corresponding two-dimensional problems [32–34], we select

$$\mathbf{w}'(z)=\langle \varsigma_k^{-1}\rangle\mathbf{k},\quad \varsigma_k=\frac{z_k+\sqrt{z_k^2-a^2-\mu_k^2 b^2}}{a-i\mu_k b} \tag{60}$$

where the angular bracket $\langle\ \rangle$ stands for the diagonal matrix, i.e., $\langle\varsigma_k^{-1}\rangle=\mathrm{diag}[\varsigma_1^{-1},\varsigma_2^{-1},\varsigma_3^{-1}]$, and $\mathbf{k}$ is the unknown coefficient vector to be determined through the satisfaction of the boundary condition. Substituting Eqs. (57) and (26b) into (60), we obtain

$$\mathbf{w}'(z)=e^{-i\psi}\mathbf{k}\ \text{ along the hole boundary.} \tag{61}$$

Substituting the first relation of Eq. (59b) and Eq. (61) into the first relation of (59a) with x and y given by Eq. (57), the free edge boundary condition, the second relation of Eq. (58a) now leads to

$$\mathbf{k}=\frac{ibM}{2}\mathbf{A}^{-1}\mathbf{i}_2. \tag{62}$$

By combining Eqs. (59), (60), and (62), the final explicit solution of the present problem can be expressed as

$$\begin{aligned}\boldsymbol{\phi}&=\boldsymbol{\phi}^{\infty}-bM\,\mathrm{Im}\{\mathbf{A}\langle\varsigma_k^{-1}\rangle\mathbf{A}^{-1}\}\mathbf{i}_2,\\ \boldsymbol{\alpha}&=\boldsymbol{\alpha}^{\infty}-bM\,\mathrm{Im}\{\mathbf{B}\langle\varsigma_k^{-1}\rangle\mathbf{A}^{-1}\}\mathbf{i}_2.\end{aligned} \tag{63}$$

With the explicit solution found in Eq. (63), the deflection, bending moments and transverse shear forces of the plate can then be obtained by using the relations given in Eqs. (41) and (32), or by using Eqs. (26) and (27) with w found by Eqs. (60), (62), and (56). Like the two-dimensional problems, by using some identities the moments around the hole boundary can be obtained in a real explicit form. A detailed derivation and the related numerical calculation and discussions as well as the reduction to crack problems can be found in Ref. [36].

7.2 Example 2: An Anisotropic Plate Embedded With a Rigid Elliptical Inclusion Subjected to Out-of-Plane Bending Moments. In this example, we are dealing with the predisplaced boundary conditions since the edge of rigid inclusion cannot undergo deformation. Like Eq. (58), the boundary condition of this problem can be expressed as

$$\begin{aligned}\boldsymbol{\phi}&=-My\mathbf{i}_2\quad \text{at infinity},\\ \boldsymbol{\alpha}&=\mathbf{0}\quad \text{around the hole boundary.}\end{aligned} \tag{64}$$

By the approach similar to example 1 and referring to its corresponding two-dimensional problems [31,35], we obtain the explicit solution for this example as

$$\begin{aligned}\boldsymbol{\phi}&=\boldsymbol{\phi}^{\infty}+M\,\mathrm{Re}\{\mathbf{A}\langle\varsigma_k^{-1}\rangle\mathbf{B}^{-1}(a\mathbf{d}_1-ib\mathbf{d}_2)\},\\ \boldsymbol{\alpha}&=\boldsymbol{\alpha}^{\infty}+M\,\mathrm{Re}\{\mathbf{B}\langle\varsigma_k^{-1}\rangle\mathbf{B}^{-1}(a\mathbf{d}_1-ib\mathbf{d}_2)\}.\end{aligned} \tag{65}$$

Detailed comparison and discussion of this result are also given in Ref. [36], from which we know that our solution can be reduced to the solution found in the literature for the special case of orthotropic plate embedded with a rigid circular core.

8 Conclusions

A Stroh-like complex variable formalism for the bending theory of anisotropic plates is established in this paper by taking the advantages of the Stroh-Lekhnitskii connection. Its associated eigenrelation and orthogonality relation are also obtained. Almost all the relations have been purposely arranged to have the same form as those of the corresponding two-dimensional formalism. The results show that the only difference from the two-dimensional formalism is that the eigenvector matrices **A** and **B** of the general solutions shown in Eq. (42) have been interchanged. Due to the similarity, almost all the mathematical techniques developed for two-dimensional problems can be adopted for the plate bending problems. By simple analogy, many problems that cannot be solved previously now have the possibility to be solved analytically. For the purpose of illustration, two examples are solved analytically. One is an anisotropic plate weakened by an elliptical hole subjected to out-of-plane bending moments, the other is an anisotropic plate embedded with a rigid elliptical inclusion subjected to out-of-plane bending moments. The former is a traction boundary value problem, while the latter is a displacement boundary value problem. The solutions show that they possess almost the same forms as those obtained for the corresponding two-dimensional problems. This gives us a hint that most of the pure bending problems for the homogeneous anisotropic plates can be solved even without detailed derivation if their counterparts in two-dimensional problems have been solved previously.

The results of this paper stimulated us to develop a fully Stroh-like formalism for the coupled stretching-bending analysis for general composites. For the readers who are interested in the further development, please refer to our most recent work [25].

Acknowledgments

The author would like to thank his Ph.D. student Mr. M. C. Hsieh for performing a numerical check on some relations proved in this paper, and thanks the National Science Council for support through Grant NSC 89-2212-E-006-127.

Appendix A: Stroh Formalism for Two-Dimensional Anisotropic Elasticity

In a fixed rectangular coordinate system x_i, $i=1,2,3$, let μ_i, σ_{ij}, ε_{ij} be, respectively, the displacement, stress and strain. The strain-displacement equations, the stress-strain laws, and the equations of equilibrium for anisotropic elasticity are

$$\varepsilon_{ij}=\tfrac{1}{2}(u_{i,j}+u_{j,i}), \tag{A1a}$$

$$\sigma_{ij}=C_{ijks}\varepsilon_{ks}, \tag{A1b}$$

$$\sigma_{ij,j}=0, \tag{A1c}$$

where repeated indices imply summation, a comma stands for differentiation, and C_{ijks} are elastic constants, which are assumed to be fully symmetric and positive definite. Consider a two-dimensional deformation in which u_i, $i=1,2,3$, depend on x_1 and x_2 only, the general solution to Eq. (A1) can be written as

$$\mathbf{u}=\sum_{k=1}^{6}\mathbf{a}_k f_k(z_k),\quad z_k=x_1+\mu_k x_2, \tag{A2}$$

in which f_k, $k=1,2,\ldots,6$ are arbitrary functions of their arguments and μ_k and $\mathbf{a}_k$ are the eigenvalues and eigenvectors of the following eigenrelation:

$$\{\mathbf{Q}+\mu(\mathbf{R}+\mathbf{R}^T)+\mu^2\mathbf{T}\}\mathbf{a}=\mathbf{0}. \tag{A3}$$

In Eq. (A3) the superscript T stands for the transpose and **Q**, **R**, **T** are the 3×3 real matrices given by

$$Q_{ik}=C_{i1k1},\quad R_{ik}=C_{i1k2},\quad T_{ik}=C_{i2k2}, \tag{A4}$$

In Eq. (A2), we assume that the eigenvalues μ_k, $k=1,2,\ldots,6$, are distinct and its associated eigenvectors $\mathbf{a}_k$, $k=1,2,\ldots,6$, are independent each other. Since μ_k cannot be real if the strain energy is positive [29], μ_k occurs as three pairs of complex conjugates. We let $\mu_{k+3}=\bar{\mu}_k$, $\mathrm{Im}(\mu_k)>0$, $k=1,2,3$, where an overbar denotes the complex conjugate and Im stands for the imaginary part. We then have $\mathbf{a}_{k+3}=\bar{\mathbf{a}}_k$, $k=1,2,3$. For the displacement $\mathbf{u}$ to be real, we let $f_{k+3}=\bar{f}_k$, $k=1,2,3$, and Eq. (A2) becomes

$$\mathbf{u}=2\,\mathrm{Re}\left\{\sum_{k=1}^{3}\mathbf{a}_k f_k(z_k)\right\}, \tag{A5}$$

in which Re stands for the real part. Introducing the vector

$$\mathbf{b}=(\mathbf{R}^T+\mu\mathbf{T})\mathbf{a}=-\frac{1}{\mu}(\mathbf{Q}+\mu\mathbf{R})\mathbf{a}, \tag{A6}$$

where the second equality comes from Eq. (A3), the stresses σ_{ij} obtained by substituting Eq. (A2) into (A1a) and (A1b) can be written as

$$\sigma_{i1}=-\phi_{i,2}, \quad \sigma_{i2}=\phi_{i,1}, \tag{A7}$$

where $\boldsymbol{\phi}$ is the stress function

$$\boldsymbol{\phi}=2\,\mathrm{Re}\left\{\sum_{k=1}^{3}\mathbf{b}_k f_k(z_k)\right\}. \tag{A8}$$

If we introduce a 3×1 column vector $\mathbf{f}(z)$ and two 3×3 complex matrices $\mathbf{A}$ and $\mathbf{B}$ by

$$\begin{gathered}\mathbf{f}(z)=\{f_1(z_1)f_2(z_2)f_3(z_3)\}^T,\\ \mathbf{A}=[\mathbf{a}_1\ \ \mathbf{a}_2\ \ \mathbf{a}_3], \quad \mathbf{B}=[\mathbf{b}_1\ \ \mathbf{b}_2\ \ \mathbf{b}_3],\end{gathered} \tag{A9}$$

Eqs. (A5) and (A8) can be written as

$$\mathbf{u}=2\,\mathrm{Re}\{\mathbf{Af}(z)\}, \quad \boldsymbol{\phi}=2\,\mathrm{Re}\{\mathbf{Bf}(z)\}. \tag{A10}$$

Equation (A6) can be reconstructed into the following standard eigenrelation

$$\mathbf{N}\boldsymbol{\xi}=\mu\boldsymbol{\xi}, \tag{A11}$$

where

$$\mathbf{N}=\begin{bmatrix}\mathbf{N}_1 & \mathbf{N}_2\\ \mathbf{N}_3 & \mathbf{N}_1^T\end{bmatrix}, \tag{A12a}$$

$$\boldsymbol{\xi}=\begin{Bmatrix}\mathbf{a}\\ \mathbf{b}\end{Bmatrix}, \tag{A12b}$$

and

$$\mathbf{N}_1=-\mathbf{T}^{-1}\mathbf{R}^T, \quad \mathbf{N}_2=\mathbf{T}^{-1}=\mathbf{N}_2^T, \quad \mathbf{N}_3=\mathbf{R}\mathbf{T}^{-1}\mathbf{R}^T-\mathbf{Q}=\mathbf{N}_3^T. \tag{A13}$$

Note that for the convenience of readers' usage, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{T}$ defined in Eq. (A4) are usually written as

$$\begin{gathered}\mathbf{Q}=\begin{bmatrix}C_{11} & C_{16} & C_{15}\\ C_{16} & C_{66} & C_{56}\\ C_{15} & C_{56} & C_{55}\end{bmatrix}, \quad \mathbf{R}=\begin{bmatrix}C_{16} & C_{12} & C_{14}\\ C_{66} & C_{26} & C_{46}\\ C_{56} & C_{25} & C_{45}\end{bmatrix},\\ \mathbf{T}=\begin{bmatrix}C_{66} & C_{26} & C_{46}\\ C_{26} & C_{22} & C_{24}\\ C_{46} & C_{24} & C_{44}\end{bmatrix},\end{gathered} \tag{A14}$$

where C_{ij} are the contracted notation of elastic tensor C_{ijkl} [5,26].

In-Plane Problem. Consider the anisotropic materials having one plane of elastic symmetry, i.e., the monoclinic materials. In this case, the in-plane and anti-plane problems will decouple. If only the in-plane problems are considered, the general solution (A9) and (A10) and its associated eigenrelations (A11)–(A14) can be reduced to the following. The general solution is

$$\mathbf{u}=2\,\mathrm{Re}\{\mathbf{Af}(z)\}, \quad \boldsymbol{\phi}=2\,\mathrm{Re}\{\mathbf{Bf}(z)\}, \tag{A15}$$

where

$$\begin{gathered}\mathbf{f}(z)=\{f_1(z_1)f_2(z_2)\}^T,\\ \mathbf{A}=[\mathbf{a}_1\ \ \mathbf{a}_2], \quad \mathbf{B}=[\mathbf{b}_1\ \ \mathbf{b}_2].\end{gathered} \tag{A16}$$

The eigenrelation is

$$\mathbf{N}\xi=\mu\boldsymbol{\xi}, \tag{A17}$$

where $\boldsymbol{\xi}$ is a 4×1 column vector defined in Eq. (A12b), and $\mathbf{N}$ is a 4×4 matrix defined in Eq. (A12a) of which $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_3$ are three 2×2 matrices defined in Eqs. (A13). $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{T}$ matrices defined in Eqs. (A14) are then reduced to

$$\mathbf{Q}=\begin{bmatrix}C_{11} & C_{16}\\ C_{16} & C_{66}\end{bmatrix}, \quad \mathbf{R}=\begin{bmatrix}C_{16} & C_{12}\\ C_{66} & C_{26}\end{bmatrix}, \quad \mathbf{T}=\begin{bmatrix}C_{66} & C_{26}\\ C_{26} & C_{22}\end{bmatrix}. \tag{A18}$$

Appendix B: Lekhnitskii Formalism for Two-Dimensional Anisotropic Elasticity

Consider the basic equations given in Eq. (A1), and assume that the stresses are independent of x_3. The equilibrium equation will be satisfied automatically if we set

$$\begin{gathered}\sigma_{11}=\chi_{,22}, \quad \sigma_{22}=\chi_{,11}, \quad \sigma_{12}=-\chi_{,12}, \quad \sigma_{32}=-\psi_{,1},\\ \sigma_{31}=-\psi_{,2}.\end{gathered} \tag{B1}$$

Substituting Eqs. (B1) into the stress-strain relation, we may obtain the expressions of strains in terms of χ and ψ. According to the strain-displacement equations, integration of strains may lead to the expressions of displacements. Moreover, since three displacement components come from five strain components, two compatibility conditions are needed. The satisfaction of these two compatibility conditions will then lead to the following sixth-order characteristic equation [1]

$$\ell_2(\mu)\ell_4(\mu)-[\ell_3(\mu)]^2=0, \tag{B2}$$

where

$$\begin{gathered}\ell_2(\mu)=\hat{S}_{55}\mu^2-2\hat{S}_{45}\mu+\hat{S}_{44},\\ \ell_3(\mu)=\hat{S}_{15}\mu^3-(\hat{S}_{14}+\hat{S}_{56})\mu^2+(\hat{S}_{25}+\hat{S}_{46})\mu-\hat{S}_{24},\\ \ell_4(\mu)=\hat{S}_{11}\mu^4-2\hat{S}_{16}\mu^3+(2\hat{S}_{12}+\hat{S}_{66})\mu^2-2\hat{S}_{26}\mu+\hat{S}_{22},\end{gathered} \tag{B3}$$

and $\hat{S}_{ij}$ are the reduced elastic compliances defined by

$$\hat{S}_{ij}=S_{ij}-S_{i3}S_{3j}/S_{33}. \tag{B4}$$

It has been proved that the roots of the characteristic equation (B2) are exactly the same as those obtained in Eq. (A3) or (A11) [7–11]. If we assume that μ_1, μ_2 and μ_3 are distinct, the general solutions of stresses and displacements derived according to the approach described above can then be expressed as

$$\begin{gathered}\sigma_{11}=2\,\mathrm{Re}\{\mu_1^2f_1'(z_1)+\mu_2^2f_2'(z_2)+\mu_3^2\lambda_3f_3'(z_3)\},\\ \sigma_{22}=2\,\mathrm{Re}\{f_1'(z_1)+f_2'(z_2)+\lambda_3f_3'(z_3)\},\\ \sigma_{12}=-2\,\mathrm{Re}\{\mu_1f_1'(z_1)+\mu_2f_2'(z_2)+\mu_3\lambda_3f_3'(z_3)\},\\ \sigma_{13}=2\,\mathrm{Re}\{\mu_1\lambda_1f_1'(z_1)+\mu_2\lambda_2f_2'(z_2)+\mu_3f_3'(z_3)\},\\ \sigma_{23}=2\,\mathrm{Re}\{\lambda_1f_1'(z_1)+\lambda_2f_2'(z_2)+f_3'(z_3)\},\end{gathered} \tag{B5}$$

and

$$u_1=2\,\mathrm{Re}\sum_{k=1}^{3} h_k f_k(z_k),$$

$$u_2=2\,\mathrm{Re}\sum_{k=1}^{3} g_k f_k(z_k), \tag{B6}$$

$$u_3=2\,\mathrm{Re}\sum_{k=1}^{3} e_k f_k(z_k),$$

where $f_k(z_k)$, $k=1,2,3$, are three holomorphic functions of complex variables z_k $(=x_1+\mu_k x_2)$, which will be determined by the boundary conditions; λ_k, h_k, g_k, and e_k are defined as

$$\lambda_1=-\ell_3(\mu_1)/\ell_2(\mu_1),\quad \lambda_2=-\ell_3(\mu_2)/\ell_2(\mu_2),\quad \lambda_3=-\ell_3(\mu_3)/\ell_4(\mu_3), \tag{B7}$$

and

$$\begin{aligned}
h_k&=\hat S_{11}\mu_k^2+\hat S_{12}-\hat S_{16}\mu_k+\lambda_k(\hat S_{15}\mu_k-\hat S_{14}),\\
g_k&=\hat S_{12}\mu_k+\frac{\hat S_{22}}{\mu_k}-\hat S_{26}+\lambda_k\left(\hat S_{25}-\frac{\hat S_{24}}{\mu_k}\right),\\
e_k&=\hat S_{14}\mu_k+\frac{\hat S_{24}}{\mu_k}-\hat S_{46}+\lambda_k\left(\hat S_{45}-\frac{\hat S_{44}}{\mu_k}\right),\quad k=1,2,\\
h_3&=\lambda_3(\hat S_{11}\mu_3^2+\hat S_{12}-\hat S_{16}\mu_3)+\hat S_{15}\mu_3-\hat S_{14},\\
g_3&=\lambda_3\left(\hat S_{12}\mu_3+\frac{\hat S_{22}}{\mu_3}-\hat S_{26}\right)+\hat S_{25}-\frac{\hat S_{24}}{\mu_3},\\
e_3&=\lambda_3\left(\hat S_{14}\mu_3+\frac{\hat S_{24}}{\mu_3}-\hat S_{46}\right)+\hat S_{45}-\frac{\hat S_{44}}{\mu_3}.
\end{aligned} \tag{B8}$$

By viewing the component form solutions $(B5)$–$(B8)$ and the matrix form solutions $(A9)$–$(A10)$ or $(A5)$–$(A8)$, it can be seen that

$$\mathbf{a}_k=\begin{Bmatrix} h_k\\ g_k\\ e_k\end{Bmatrix},\quad \mathbf{b}_k=\begin{Bmatrix} -\mu_k\\ 1\\ -\lambda_k\end{Bmatrix},\quad k=1,2,\quad \mathbf{a}_3=\begin{Bmatrix} h_3\\ g_3\\ e_3\end{Bmatrix},\quad \mathbf{b}_3=\begin{Bmatrix} -\mu_3\lambda_3\\ \lambda_3\\ -1\end{Bmatrix}. \tag{B9}$$

In-Plane Problem. Similar to Stroh formalism, for in-plane problems of monoclinic materials the general solutions $(B5)$–$(B8)$ can be reduced to

$$\begin{aligned}
\sigma_{11}&=2\,\mathrm{Re}\{\mu_1^2 f_1'(z_1)+\mu_2^2 f_2'(z_2)\},\\
\sigma_{22}&=2\,\mathrm{Re}\{f_1'(z_1)+f_2'(z_2)\},\\
\sigma_{12}&=-2\,\mathrm{Re}\{\mu_1 f_1'(z_1)+\mu_2 f_2'(z_2)\},\\
u_1&=2\,\mathrm{Re}\sum_{k=1}^{2} h_k f_k(z_k),\\
u_2&=2\,\mathrm{Re}\sum_{k=1}^{2} g_k f_k(z_k),
\end{aligned} \tag{B10}$$

where

$$h_k=\hat S_{11}\mu_k^2+\hat S_{12}-\hat S_{16}\mu_k,\qquad g_k=\hat S_{12}\mu_k+\frac{\hat S_{22}}{\mu_k}-\hat S_{26}. \tag{B11}$$

The associated characteristic equations $(B2)$–$(B4)$ can also be reduced to

$$\hat S_{11}\mu^4-2\hat S_{16}\mu^3+(2\hat S_{12}+\hat S_{66})\mu^2-2\hat S_{26}\mu+\hat S_{22}=0. \tag{B12}$$

From Eq. $(B9)$, the eigenvectors of Stroh formalism can also be reduced to

$$\mathbf{a}_k=\begin{Bmatrix} h_k\\ g_k\end{Bmatrix},\quad \mathbf{b}_k=\begin{Bmatrix} -\mu_k\\ 1\end{Bmatrix},\quad k=1,2. \tag{B13}$$

Appendix C: The Explicit Expressions of $\mathbf{N}_1$, $\mathbf{N}_2$ and $\mathbf{N}_3$ for the Plate Bending Problems

By following the steps described in Ting's book [5], we rearrange Eq. (49) into

$$\left[\begin{array}{cc:cc}
D_{22}^* & -\frac{1}{2}D_{26}^* & -\frac{1}{2}D_{26}^* & D_{12}^*\\
-\frac{1}{2}D_{26}^* & \frac{1}{4}D_{66}^* & \frac{1}{4}D_{66}^* & -\frac{1}{2}D_{16}^*\\
\hdashline
-\frac{1}{2}D_{26}^* & \frac{1}{4}D_{66}^* & \frac{1}{4}D_{66}^* & -\frac{1}{2}D_{16}^*\\
D_{12}^* & -\frac{1}{2}D_{16}^* & -\frac{1}{2}D_{16}^* & D_{11}^*
\end{array}\right]
\times\left[\begin{array}{cc:cc}
D_{22} & 0 & -2D_{26} & D_{12}\\
0 & 0 & 0 & 0\\
\hdashline
-2D_{26} & 0 & 4D_{66} & -2D_{16}\\
D_{12} & 0 & -2D_{16} & D_{11}
\end{array}\right]
=\left[\begin{array}{cc:cc}
1 & 0 & 0 & 0\\
0 & 0 & 1 & 0\\
\hdashline
0 & 0 & 1 & 0\\
0 & 0 & 0 & 1
\end{array}\right] \tag{C1}$$

Equation $(C1)$ can also be written in matrix form as

$$\begin{bmatrix}\mathbf{Q} & \mathbf{R}\\ \mathbf{R}^T & \mathbf{T}\end{bmatrix}\begin{bmatrix}\mathbf{Q}_2^* & \mathbf{R}_2^*\\ \mathbf{R}_2^{*T} & \mathbf{T}^*\end{bmatrix}=\begin{bmatrix}\mathbf{I}_1 & \mathbf{I}_{12}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}. \tag{C2}$$

Employing the relation

$$\begin{bmatrix}\mathbf{I} & \mathbf{N}_1^{\mathrm{T}}\\ \mathbf{0} & \mathbf{N}_2\end{bmatrix}\begin{bmatrix}\mathbf{Q} & \mathbf{R}\\ \mathbf{R}^T & \mathbf{T}\end{bmatrix}=\begin{bmatrix}-\mathbf{N}_3 & \mathbf{0}\\ -\mathbf{N}_1 & \mathbf{I}\end{bmatrix}, \tag{C3}$$

which can be verified by using the definition $(47c)$, Eq. $(C2)$ becomes

$$\begin{bmatrix}-\mathbf{N}_3 & \mathbf{0}\\ -\mathbf{N}_1 & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{Q}_2^* & \mathbf{R}_2^*\\ \mathbf{R}_2^{*\mathrm{T}} & \mathbf{T}^*\end{bmatrix}=\begin{bmatrix}\mathbf{I} & \mathbf{N}_1^{\mathrm{T}}\\ \mathbf{0} & \mathbf{N}_2\end{bmatrix}\begin{bmatrix}\mathbf{I}_1 & \mathbf{I}_{12}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \tag{C4}$$

or

$$-\mathbf{N}_3\mathbf{Q}_2^*=\mathbf{I}_1,\quad -\mathbf{N}_3\mathbf{R}_2^*=\mathbf{I}_{12}+\mathbf{N}_1^T,\quad -\mathbf{N}_1\mathbf{Q}_2^*+\mathbf{R}_2^{*T}=\mathbf{0},\quad -\mathbf{N}_1\mathbf{R}_2^*+\mathbf{T}^*=\mathbf{N}_2. \tag{C5}$$

By knowing the structures of $\mathbf{N}_1$ and $\mathbf{N}_3$, which are [30]

$$\mathbf{N}_1=\begin{bmatrix} * & -1\\ * & 0\end{bmatrix},\quad \mathbf{N}_3=\begin{bmatrix} * & 0\\ 0 & 0\end{bmatrix}, \tag{C6}$$

and making use of Eq. $(C5)$, the explicit expressions of $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_3$ can then be obtained as those shown in Eq. (51).

Appendix D: Proof of the Orthogonality Relation, Eq. (52)

Because the form of the eigenrelation for the plate bending problems is exactly the same as that for the two-dimensional problems, in order to prove its orthogonality relation we just extract the proof provided in Ting's book [5]. Since the matrix $\mathbf{N}$ of the eigenrelation (47) is not symmetric, its eigenvector $\boldsymbol{\xi}$ is generally called right eigenvector. Its associated left eigenvector $\boldsymbol{\eta}$ satisfies the following eigenrelation:

$$\mathbf{N}^T\boldsymbol{\eta}=\mu\,\boldsymbol{\eta}. \tag{D1}$$

Introducing the constant matrix

$$\mathbf{J}=\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{bmatrix}, \tag{D2}$$

it can be shown that $\mathbf{J}=\mathbf{J}^T=\mathbf{J}^{-1}$ and $\mathbf{JN}$ is symmetric, i.e., $\mathbf{JN}=(\mathbf{JN})^T=\mathbf{N}^T\mathbf{J}$. Premultiplying both sides of Eq. (47*a*) by $\mathbf{J}$ and using $\mathbf{JN}=\mathbf{N}^T\mathbf{J}$, we obtain

$$\mathbf{N}^T(\mathbf{J}\boldsymbol{\xi})=\mu(\mathbf{J}\boldsymbol{\xi}). \tag{D3}$$

By comparing Eqs. (*D*1) and (*D*3), the left eigenvector $\boldsymbol{\eta}$ can therefore be assumed to have the form

$$\boldsymbol{\eta}=\mathbf{J}\boldsymbol{\xi}=\begin{Bmatrix}\mathbf{b}\\ \mathbf{a}\end{Bmatrix}. \tag{D4}$$

To prove that the left and right eigenvectors associated with different eigenvalues are orthogonal to each other, we consider two distinct eigenvalues μ_i and μ_j. Their associated eigenrelations can be written as

$$\mathbf{N}\boldsymbol{\xi}_i=\mu_i\boldsymbol{\xi}_i, \tag{D5a}$$

$$\mathbf{N}^T\boldsymbol{\eta}_j=\mu_j\boldsymbol{\eta}_j. \tag{D5b}$$

Premultiplying both sides of Eq. (*D*5*a*) by $\boldsymbol{\eta}_j^T$ and premultiplying both sides of Eq. (*D*5*b*) by $\boldsymbol{\xi}_i^T$, we have

$$\boldsymbol{\eta}_j^T\mathbf{N}\boldsymbol{\xi}_i=\mu_i\boldsymbol{\eta}_j^T\boldsymbol{\xi}_i, \tag{D6a}$$

$$\boldsymbol{\xi}_i^T\mathbf{N}^T\boldsymbol{\eta}_j=\mu_j\boldsymbol{\xi}_i^T\boldsymbol{\eta}_j. \tag{D6b}$$

Transposing both sides of Eq. (*D*6*b*) and subtracting its results from Eq. (*D*6*a*), we obtain

$$(\mu_i-\mu_j)\boldsymbol{\eta}_j^T\boldsymbol{\xi}_i=0. \tag{D7}$$

Hence, if $\mu_i\neq\mu_j$, $\boldsymbol{\eta}_j^T\boldsymbol{\xi}_i=0$. The right eigenvector $\boldsymbol{\xi}$ given by Eq. (47), and hence the left eigenvector $\boldsymbol{\eta}$ by Eq. (*D*4) is unique up to an arbitrary multiplier. Assuming that μ_k, $k=1,2,3,4$ are distinct, we may then normalize $\boldsymbol{\xi}_\kappa$ by

$$\boldsymbol{\eta}_j^T\boldsymbol{\xi}_i=\delta_{ij}. \tag{D8}$$

In view of $\mu_{k+2}=\overline{\mu}_k$, $\mathbf{a}_{k+2}=\overline{\mathbf{a}}_k$, $\mathbf{b}_{k+2}=\overline{\mathbf{b}}_k$, $k=1,2$ and the definitions in the first two relations in Eq. (42*b*), the second relation in Eq. (47*b*), and relation (*D*4), the orthorgonality relation (*D*8) can therefore be written as that shown in Eq. (52).

References

[1] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Body*, Holden-Day, Inc., San Francisco.
[2] Lekhnitskii, S. G., 1968, *Anisotropic Plates*, Gordon and Breach, New York.
[3] Stroh, A. N., 1958, "Dislocations and Cracks in Anisotropic Elasticity," Philos. Mag., **7**, pp. 625–646.
[4] Stroh, A. N., 1962, "Steady State Problems in Anisotropic Elasticity," J. Math. Phys., **41**, pp. 77–103.
[5] Ting, T. C. T., 1996, *Anisotropic Elasticity—Theory and Applications*, Oxford, New York.
[6] Suo, Z., 1990, "Singularities, Interfaces and Cracks in Dissimilar Anisotropic Media," Proc. R. Soc. London, Ser. A, **427**, pp. 331–358.
[7] Hwu, C., 1996, "Correspondence Relations Between Anisotropic and Isotropic," Chin. J. Mech., **12**, pp. 483–493.
[8] Barnett, D. M., and Kirchner, H. O. K., 1997, "A Proof of the Equivalence of the Stroh and Lekhnitskii Sextic Equations for Plane Anisotropic Elastostatics," Philos. Mag. A, **76**, pp. 231–239.
[9] Ting, T. C. T., 1999, "A Modified Lekhnitskii Formalism a la Stroh for Anisotropic Elasticity and Classifications of the 6×6 Matrix **N**," Proc. R. Soc. London, Ser. A, **455**, pp. 69–89.
[10] Yin, W. L., 2000, "Deconstructing Plane Anisotropic Elasticity, Part I: The Latent Structure of Lekhnitskii's Formalism," Int. J. Solids Struct., **37**, pp. 5257–5276.
[11] Yin, W. L., 2000, "Deconstructing Plane Anisotropic Elasticity, Part II: Stroh's Formalism Sans Frills," Int. J. Solids Struct., **37**, pp. 5277–5296.
[12] Lekhnitskii, S. G., 1938, "Some Problems Related to the Theory of Bending of Thin Plates," Prikl. Mat. Mekh., **II**, p. 187.
[13] Becker, W., 1992, "Closed-Form Analytical Solutions for a Griffith Crack in a Non-Symmetric Laminate Plate," Compos. Struct., **21**, pp. 49–55.
[14] Becker, W., 1995, "Concentrated Forces and Moments on Laminates With Bending Extension Coupling," Compos. Struct., **30**, pp. 1–11.
[15] Zakharov, D. D., 1995, "Formulation of Boundary-Value Problems of Statics for Thin Elastic Asymmetrically-Laminated Anisotropic Plates and Solution Using Functions of a Complex Variable," J. Appl. Math. Mech., **59**, pp. 615–623.
[16] Zakharov, D. D., 1996, "Complex Potential Techniques in the Boundary Problems for Thin Anisotropic Asymmetric Laminates," Z. Angew. Math. Mech., **76**, pp. 709–710.
[17] Senthil, S. V., and Batra, R. C., 1999, "Analytical Solutions for Rectangular Thick Laminated Plates Subjected to Arbitrary Boundary Conditions," AIAA J., **37**, pp. 1464–1473.
[18] Zakharov, D. D., and Becker, W., 2000, "Unsymmetric Composite Laminate With a Discontinuity of the In-Plane Displacement or of the Slope," Acta Mech., **144**, pp. 127–135.
[19] Yuan, F. G., and Yang, S., 2000, "Asymptotic Crack-Tip Fields in an Anisotropic Plate Subjected to Bending, Twisting Moments and Transverse Shear Loads," Compos. Sci. Technol., **60**, pp. 2489–2502.
[20] Wang, M., and Zhao, Y., 2001, "Stroh Formalism and Its Generalization in Two-Dimensional Anisotropic Elasticity," Mech. Eng., **23**, pp. 7–30.
[21] Lu, P., 1994, "Stroh Type Formalism for Unsymmetric Laminated Plate," Mech. Res. Commun., **21**, pp. 249–254.
[22] Lu, P., and Mahrenholtz, O., 1994, "Extension of the Stroh Formalism to an Analysis of Bending of Anisotropic Elastic Plates," J. Mech. Phys. Solids, **42**, pp. 1725–1741.
[23] Cheng, Z. Q., and Reddy, J. N., 2002, "Octet Formalism for Kirchhoff Anisotropic Plates," Proc. R. Soc. London, Ser. A, **458**, pp. 1499–1517.
[24] Zastrow, U., 1985, "On the Complete System of Fundamental Solutions for Anisotropic Slices and Slabs: A Comparison by Use of the Slab Analogy," J. Elast., **15**, pp. 293–318.
[25] Hwu, C., 2003, "Stroh-Like Formalism for the Coupled Stretching-Bending Analysis of Composite Laminates," Int. J. Solids Struct., **40**(13/14), pp. 3681–3705.
[26] Jones, R. M., 1974, *Mechanics of Composite Materials*, Scripta, Washington, D.C.
[27] Vinson, J. R., and Sierakowski, R. L., 1986, *The Behavior of Structures Composed of Composite Materials*, Martinus Nijhoff, Dordrecht.
[28] Szilard, R., 1974, *Theory and Analysis of Plates—Classical and Numerical Methods*, Prentice Hall, Inc., New Jersey.
[29] Eshelby, J. D., Read, W. T., and Shockley, W., 1953, "Anisotropic Elasticity With Applications to Dislocation Theory," Acta Metall., **1**, pp. 251–259.
[30] Ting, T. C. T., 1988, "Some Identities and the Structures of $\mathbf{N}_i$ the Stroh Formalism of Anisotropic Elasticity," Q. Appl. Math., **46**, pp. 109–120.
[31] Hwu, C., and Ting, T. C. T., 1989, "Two-Dimensional Problems of the Anisotropic Elastic Solid With an Elliptic Inclusion," Q. J. Mech. Appl. Math., **42**, Pt. 4, pp. 553–572.
[32] Hwu, C., 1990, "Thermal Stresses in an Anisotropic Plate Disturbed by an Insulated Elliptic Hole or Crack," ASME J. Appl. Mech., **57**, No. 4, pp. 916–922.
[33] Hwu, C., 1992, "Polygonal Holes in Anisotropic Media," Int. J. Solids Struct., **29**, pp. 2369–2384.
[34] Hwu, C., and Yen, W. J., 1991, "Green's Functions of Two-Dimensional Anisotropic Plates Containing an Elliptic Hole," Int. J. Solids Struct., **27**, pp. 1705–1719.
[35] Hwu, C., and Yen, W. J., 1993, "On the Anisotropic Elastic Inclusions in Plane Elastostatics," ASME J. Appl. Mech., **60**, pp. 626–632.
[36] Hsieh, M. C., and Hwu, C., 2002, "Anisotropic Elastic Plates With Holes/Cracks/Inclusions Subjected to Out-of-Plane Bending Moments," Int. J. Solids Struct., **39**(19), pp. 4905–4925.

Response Power Spectrum of Multi-Degree-of-Freedom Nonlinear Systems by a Galerkin Technique

G. Failla
Doctor of Philosophy,
Università degli Studi di Palermo, DISeG, Viale delle Scienze 90128, Palermo, Italy

P. D. Spanos
Fellow, ASME
L. B. Ryon Chair in Engineering, Rice University, P.O. Box 1892, Houston, Texas 77251, USA

M. Di Paola
Professor,
Università degli Studi di Palermo, DISeG, Viale delle Scienze 90128, Palermo, Italy

This paper deals with the estimation of spectral properties of randomly excited multi-degree-of-freedom (MDOF) nonlinear vibrating systems. Each component of the vector of the stationary system response is expanded into a trigonometric Fourier series over an adequately long interval T. The unknown Fourier coefficients of individual samples of the response process are treated by harmonic balance, which leads to a set of nonlinear equations that are solved by Newton's method. For polynomial nonlinearities of cubic order, exact solutions are developed to compute the Fourier coefficients of the nonlinear terms, including those involved in the Jacobian matrix associated with the implementation of Newton's method. The proposed technique is also applicable for arbitrary nonlinearities via a cubicization procedure over the interval T. Upon determining the Fourier coefficients, estimates of the response power spectral density matrix are constructed by averaging their squared moduli over the samples ensemble. Examples of application prove the reliability of the technique by comparison with digital simulation data.

[DOI: 10.1115/1.1599916]

Introduction

In the investigation of randomly excited multi-degree-of-freedom (MDOF) dynamic systems the power spectral density matrix of the stationary response provides critical information for design and reliability analysis purposes. It is known that the frequency content of a linear system response is related to the frequency content of the excitation through a linear transformation that involves the system transfer function matrix. The frequency content of a nonlinear system response, however, does not lend itself to a straightforward mathematical representation, since a single-frequency excitation may generally lead to a multifrequency response.

A general approach to capture the effects of nonlinear behavior may be based on the so-called Volterra series expansion [1]. The central concept of this approach is to expand the response process in a time-domain series where terms of order n are n-fold convolution integrals, which involve n excitation times of the input. The corresponding frequency-domain representation, therefore, exhibits multifrequency response transfer functions. For polynomial nonlinearities, impulse response functions of any order n are expressed in terms of the first-order response functions, which are known from linear analysis. The system response power spectral density matrix can be then computed by multifold integrals in the frequency domain, as long as the Volterra series is truncated to finite order. In this context and to handle arbitrary nonlinearities, the so-called quadratization and cubicization procedures have been developed [2,3]. They involve replacing the original nonlinearities by equivalent polynomials of second or third order to be solved by the Volterra series method. The applicability of these methods, however, is hampered by the significant computational cost involved, especially for statistical cubicization [3,4].

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 16, 2001; final revision, June 5, 2002. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

As an improvement to the standard linearization solution [5], alternative strategies have been developed by introducing in the equivalent linear system damping and stiffness elements depending on random parameters [6–8]. Note, however, that these parameters relate to a van der Pol transformation adopted for the system response, which is taken to be narrow band. This method, then, is best suited for nonlinear systems with light damping subjected to wide-band excitations [9,10].

Recently, the authors have pursued an alternative formulation of a numerical technique, known in the literature as a "Galerkin technique" [11,12], which resorts to basic concepts of spectral analysis and signal processing to estimate the response power spectrum of nonlinear single-degree-of-freedom (SDOF) systems. In this context, the stationary response process has been expanded in an N-order Fourier series, over an adequately long interval T, and the unknown Fourier coefficients of individual samples of the response process have been determined by harmonic balance. To solve the set of nonlinear equations obtained from harmonic balance, the authors have developed an efficient solution scheme based on Newton's method, where exact solutions for the Fourier coefficients of the nonlinear terms have been used. In this manner, a significant reduction of computational effort has been achieved as compared to the previous formulation of the technique, where an extensive use of the fast Fourier transform (FFT) technique has been made [11]. Then, estimates of the response spectrum have been constructed by averaging the square modulus of the computed Fourier coefficients over a number of samples.

This paper presents a formulation of the technique that is applicable for both SDOF and MDOF systems. Numerical results for a Rayleigh oscillator and a four-degree-of-freedom system featuring flow-induced nonlinearities demonstrate its reliability.

Spectral Representation of the Response Process

Preliminary Background. Consider a scalar, real-valued stochastic process, $y(t)$, $-\infty<t<\infty$, stationary in the wide sense, with mean value μ_y. According to the spectral representation theory [13,14], $y(t)$ can be represented as

$$y(t)=\mu_y+\int_0^\infty \cos(\omega t)du(\omega)+\int_0^\infty \sin(\omega t)dv(\omega), \quad (1)$$

the integrals being defined in a mean-square sense. In Eq. (1) $u(\omega)$ and $v(\omega)$ are stochastic processes whose increments, $du(\omega)$ and $dv(\omega)$, have the properties

$$E[du(\omega)]=E[dv(\omega)]=0, \quad \omega\geqslant 0 \quad (2a)$$

$$E[du(\omega)du(\omega')]=E[dv(\omega)dv(\omega')]=0, \quad \omega,\omega'\geqslant 0, \quad \omega\neq\omega' \quad (2b)$$

$$E[du(\omega)dv(\omega')]=0, \quad \omega,\omega'\geqslant 0 \quad (2c)$$

$$E[|du(\omega)|^2]=E[|dv(\omega)|^2]=2S_{yy}(\omega)d\omega, \quad \omega\geqslant 0 \quad (2d)$$

with $S_{yy}(\omega)$ being the two-sided power spectral density of $y(t)$ and $E[\cdot]$ denoting the operator of mathematical expectation. Equation (1) may be interpreted as the limit in a mean-square sense of the trigonometric Fourier series

$$y(t)=\mu_y+\lim_{T\to\infty}\sum_{k=1}^{\infty} U_k\cos(\omega_k t)+V_k\sin(\omega_k t),$$

$$\omega_k=k\Delta\omega, \quad T=2\pi/\Delta\omega, \quad (3)$$

where $U_k=u(\omega_{k+1})-u(\omega_k)$, $V_k=v(\omega_{k+1})-v(\omega_k)$ and $\Delta\omega=\omega_{k+1}-\omega_k$. Therefore, properties similar to Eqs. (2) hold asymptotically, as $T\to\infty$, for the sequences of random coefficients $\{U_k\}$ and $\{V_k\}$,

$$E[U_k]=E[V_k]=0 \quad \text{for any } k, \quad (4a)$$

$$E[U_kU_j]=E[V_kV_j]=0 \quad \text{for any } k\neq j, \quad (4b)$$

$$E[U_kV_j]=0 \quad \text{for any } k,j, \quad (4c)$$

$$E[U_k^2]=E[V_k^2]=2S_{yy}(\omega_k)\Delta\omega \quad \text{for any } k. \quad (4d)$$

For numerical applications, an approximation of $y(t)$ can be obtained by truncating Eq. (3) to a certain order N, yielding

$$y^N(t)=\mu_y+\sum_{k=1}^{N} U_k\cos(\omega_k t)+V_k\sin(\omega_k t), \quad (5)$$

where the parameters T and N must be selected adequately large.

Application to the Nonlinear System Response. Next, consider the d-degree-of-freedom system

$$\mathbf{M}\ddot{\mathbf{X}}+\mathbf{C}\dot{\mathbf{X}}+\mathbf{K}\mathbf{X}+\mathbf{G}(\mathbf{X},\dot{\mathbf{X}})=\mathbf{F}(t), \quad (6)$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ denote $d\times d$ mass, damping, and stiffness matrices, respectively; $\mathbf{G}(\mathbf{X},\dot{\mathbf{X}})$ is an arbitrary nonlinear $d\times 1$ vector function of the variables $\mathbf{X}^T=[x_1,\ldots,x_d]$ and $\dot{\mathbf{X}}^T=[\dot{x}_1,\ldots,\dot{x}_d]$; and $\mathbf{F}(t)^T=[f_1(t),\ldots,f_d(t)]$ is a $d\times 1$ vector stationary random process. The symbol $(\ )^T$ denotes vectorial transposition. Assuming that the system response reaches stationarity, both the input and the output processes may be represented by a truncated Fourier series as in Eq. (5). That is,

$$f_i^N(t)=\mu_{f_i}+\sum_{k=1}^{N} A_k^{(i)}\cos(\omega_k t)+B_k^{(i)}\sin(\omega_k t), \quad i=1,2,\ldots,d \quad (7)$$

and

$$x_i^N(t)=\mu_{x_i}+\sum_{k=1}^{N} U_k^{(i)}\cos(\omega_k t)+V_k^{(i)}\sin(\omega_k t), \quad i=1,2,\ldots,d. \quad (8)$$

Substituting Eqs. (7) and (8) for the ith component of $\mathbf{F}(t)$ and $\mathbf{X}(t)$ in Eq. (6) leads to

$$\mathbf{M}\ddot{\mathbf{X}}^N+\mathbf{C}\dot{\mathbf{X}}^N+\mathbf{K}\mathbf{X}^N+\mathbf{G}(\mathbf{X}^N,\dot{\mathbf{X}}^N)=\mathbf{F}^N(t). \quad (9)$$

In the following, a T-periodic solution of Eq. (9), if it exists, will be referred to as an N-order Galerkin approximation of the response process $\mathbf{X}(t)$ [11].

Consider next an arbitrary realization of the excitation processes

$$\hat{f}_i^N(t)=\mu_{f_i}+\sum_{k=1}^{N} \hat{A}_k^{(i)}\cos(\omega_k t)+\hat{B}_k^{(i)}\sin(\omega_k t), \quad i=1,2,\ldots,d. \quad (10)$$

The caret is introduced to distinguish the random process, $f_i^N(t)$, from its generic sample function, $\hat{f}_i^N(t)$. To obtain $\hat{f}_i^N(t)$ in the form given by Eq. (10), digital simulation techniques may be used [15], or a standard FFT technique may be applied on experimental data. To determine the N-order Galerkin approximation of the corresponding realization of the response process, written as

$$\hat{x}_i^N(t)=\mu_{x_i}+\sum_{k=1}^{N} \hat{U}_k^{(i)}\cos(\omega_k t)+\hat{V}_k^{(i)}\sin(\omega_k t), \quad i=1,2,\ldots,d \quad (11)$$

the unknown mean value, μ_{x_i}, and the sequences of unknown Fourier coefficients, $\{\hat{U}_k^{(i)}\}$ and $\{\hat{V}_k^{(i)}\}$, must be evaluated for each degree of freedom. For this, expanding the nonlinear term $\mathbf{G}(\hat{\mathbf{X}}^N,\dot{\hat{\mathbf{X}}}^N)$ in Eq. (9) in a Fourier series, and using harmonic balance lead to the set of $(2N+1)d$ nonlinear equations

$$\mathbf{J}_k(\hat{\boldsymbol{\alpha}})=(-\omega_k^2\mathbf{M}+\mathbf{K})\hat{\mathbf{U}}_k+\omega_k\mathbf{C}\hat{\mathbf{V}}_k+\frac{2}{T}\int_0^T \mathbf{G}(\hat{\boldsymbol{\alpha}},t)\cos(\omega_k t)dt-\hat{\mathbf{A}}_k=\mathbf{0}, \quad (12)$$

$$\mathbf{J}_{N+k}(\hat{\boldsymbol{\alpha}})=(-\omega_k^2\mathbf{M}+\mathbf{K})\hat{\mathbf{V}}_k-\omega_k\mathbf{C}\hat{\mathbf{U}}_k+\frac{2}{T}\int_0^T \mathbf{G}(\hat{\boldsymbol{\alpha}},t)\sin(\omega_k t)dt-\hat{\mathbf{B}}_k=\mathbf{0}, \quad (13)$$

$$\mathbf{J}_\mu(\hat{\boldsymbol{\alpha}})=\boldsymbol{\mu}_X+\frac{1}{T}\int_0^T \mathbf{G}(\hat{\boldsymbol{\alpha}},t)dt-\boldsymbol{\mu}_F=\mathbf{0}, \quad (14)$$

where $k=1,2,\ldots,N$; the $d\times 1$ vectors $\hat{\mathbf{A}}_k$, $\hat{\mathbf{B}}_k$, and $\boldsymbol{\mu}_F$ are given by the equations

$$\hat{\mathbf{A}}_k^T=[\hat{A}_k^{(1)},\hat{A}_k^{(2)},\ldots,\hat{A}_k^{(d)}], \quad (15a)$$

$$\hat{\mathbf{B}}_k^T=[\hat{B}_k^{(1)},\hat{B}_k^{(2)},\ldots,\hat{B}_k^{(d)}], \quad (15b)$$

$$\boldsymbol{\mu}_F^T=[\mu_{f_1},\mu_{f_2},\ldots,\mu_{f_d}]. \quad (15c)$$

That is, $\hat{\mathbf{A}}_k$ and $\hat{\mathbf{B}}_k$ collect the Fourier coefficients of the kth harmonic of the input in each degree of freedom. Similarly, $\hat{\mathbf{U}}_k$, $\hat{\mathbf{V}}_k$, and $\boldsymbol{\mu}_X$ are written in the form

$$\hat{\mathbf{U}}_k^T=[\hat{U}_k^{(1)},\hat{U}_k^{(2)},\ldots,\hat{U}_k^{(d)}], \quad (16a)$$

$$\hat{\mathbf{V}}_k^T=[\hat{V}_k^{(1)},\hat{V}_k^{(2)},\ldots,\hat{V}_k^{(d)}], \quad (16b)$$

$$\boldsymbol{\mu}_X^T=[\mu_{x_1},\mu_{x_2},\ldots,\mu_{x_d}]. \quad (16c)$$

Further, the $[(2N+1)d]\times 1$ vector $\hat{\boldsymbol{\alpha}}$ collects all the unknowns

$$\hat{\boldsymbol{\alpha}}^T=[\hat{\mathbf{U}}_1^T,\ldots,\hat{\mathbf{U}}_N^T,\hat{\mathbf{V}}_1^T,\ldots,\hat{\mathbf{V}}_N^T,\boldsymbol{\mu}_X^T]. \quad (17)$$

For simplicity, the system of Eqs. (12)–(14) is recast in the form

$$\mathbf{J}(\hat{\boldsymbol{\alpha}})=\mathbf{0}, \quad (18)$$

where $\mathbf{J}(\hat{\boldsymbol{\alpha}})$ is a $[(2N+1)d]\times 1$ vector, given by the equation

$$\mathbf{J}^T(\hat{\boldsymbol{\alpha}})=[\mathbf{J}_1^T(\hat{\boldsymbol{\alpha}}),\ldots,\mathbf{J}_N^T(\hat{\boldsymbol{\alpha}}),\mathbf{J}_{N+1}^T(\hat{\boldsymbol{\alpha}}),\ldots,\mathbf{J}_{2N}^T(\hat{\boldsymbol{\alpha}}),\mathbf{J}_\mu^T(\hat{\boldsymbol{\alpha}})]. \quad (19)$$

Applying Newton's method to solve Eq. (18), at the nth iteration of the solution scheme one writes

$$\mathbf{J}(\hat{\boldsymbol{\alpha}}^{(n)})+\mathbf{\Delta J}(\hat{\boldsymbol{\alpha}}^{(n)})\cdot(\hat{\boldsymbol{\alpha}}^{(n+1)}-\hat{\boldsymbol{\alpha}}^{(n)})=\mathbf{0}. \tag{20}$$

Clearly, the entries of the vector $\mathbf{J}(\hat{\boldsymbol{\alpha}}^{(n)})$, and of the Jacobian matrix $\mathbf{\Delta J}(\hat{\boldsymbol{\alpha}}^{(n)})$ in Eq. (20) must be computed. For this, in Ref. [11] the use of the FFT technique has been suggested. In this context, given the Fourier coefficients at the nth step, first the FFT should be used d times to determine the time series $\hat{\mathbf{X}}^N(s\Delta t)$ at discrete time $t=s\Delta t$, over the time interval T. Then, it should be used d times to compute the Fourier coefficients of the time series of the nonlinear functions $\mathbf{G}(\hat{\boldsymbol{\alpha}}^{(n)},s\Delta t)$ in $\mathbf{J}(\hat{\boldsymbol{\alpha}}^{(n)})$, and $(2N+1)d^2$ times to compute the Fourier coefficients of the time series of the nonlinear functions in the Jacobian matrix, $\partial\mathbf{G}(\hat{\boldsymbol{\alpha}}^{(n)},s\Delta t)/\partial\hat{\alpha}_j$, where $j=1,2,\ldots,(2N+1)d$. Further, to capture supplementary harmonics introduced by the nonlinearities and to avoid aliasing, additional computational effort is required to generate the time series $\hat{\mathbf{X}}^N(s\Delta t)$, since a suitable number of zeros to the Fourier coefficients of $\hat{\mathbf{X}}^N(t)$ should be added at the beginning of each step [11]. Clearly, such a procedure becomes computationally costly as the dimension N of the problem increases.

Efficient Solution for Systems With Cubic or Equivalent Cubic Nonlinearity

In this section it is shown that, if the system nonlinearities can be put in a cubic polynomial form, all the integrals in Eq. (20) can be computed as closed-form functions of the unknowns $\hat{\boldsymbol{\alpha}}^{(n)}$. In this manner, the implementation of Newton's method is made very efficient and the drawbacks inherent in the use of the FFT technique are eliminated.

Introduce the $2d\times1$ state vector

$$\hat{\mathbf{Z}}^T=[\hat{\mathbf{X}}^T,\dot{\hat{\mathbf{X}}}^T]=[\hat{x}_1,\ldots,\hat{x}_d,\dot{\hat{x}}_1,\ldots,\dot{\hat{x}}_d], \tag{21}$$

and assume that

$$\mathbf{G}(\hat{\mathbf{Z}})=\mathbf{G}^c(\hat{\mathbf{Z}})=\mathbf{C}_0+\mathbf{C}_1\hat{\mathbf{Z}}+\mathbf{C}_2\hat{\mathbf{Z}}^{[2]}+\mathbf{C}_3\hat{\mathbf{Z}}^{[3]}, \tag{22}$$

where $\hat{\mathbf{Z}}^{[p]}$, for $p=2,3$, is a vector given by the lexicographic listing of the homogeneous p-forms, $\hat{z}_1^{p_1}\hat{z}_2^{p_2}\cdots\hat{z}_{2d}^{p_{2d}}$, with $\Sigma p_m=p$. That is, $\hat{\mathbf{Z}}^{[p]}$ is a $q\times1$ vector, where

$$q=\binom{2d+p-1}{p}.$$

Further, $\mathbf{C}_0$ is a $d\times1$ vector and $\mathbf{C}_p$ are time-invariant $d\times q$ matrices, for $p=1,2,3$. Note that cubic polynomial functions are considered in Eq. (22) in order to represent both symmetric and nonsymmetric nonlinear effects.

The ith component of the state vector $\hat{\mathbf{Z}}$, according to Eq. (11), can be represented as

$$\hat{z}_i^N=\mu_{z_i}+\sum_{k=1}^{N}\hat{U}_k^{z_i}\cos(\omega_k t)+\hat{V}_k^{z_i}\sin(\omega_k t). \tag{23}$$

Clearly, due to Eq. (21) it is seen that

$$\mu_{z_i}=0,\quad \hat{U}_k^{z_i}=\omega_k\hat{V}_k^{z_{i-d}},\quad \hat{V}_k^{z_i}=-\omega_k\hat{U}_k^{z_{i-d}},\quad i>d \tag{24}$$

for any $k=1,2,\ldots,N$. Consider next Eqs. (12)–(14). Substituting Eq. (22) for $\mathbf{G}(\hat{\boldsymbol{\alpha}},t)$, integrals of the form

$$Q_c=\int_0^T\hat{z}_i^N\hat{z}_j^N\cos(\omega_k t)dt, \tag{25a}$$

$$Q_s=\int_0^T\hat{z}_i^N\hat{z}_j^N\sin(\omega_k t)dt, \tag{25b}$$

$$P_c=\int_0^T\hat{z}_i^N\hat{z}_j^N\hat{z}_l^N\cos(\omega_k t)dt, \tag{26a}$$

$$P_s=\int_0^T\hat{z}_i^N\hat{z}_j^N\hat{z}_l^N\sin(\omega_k t)dt, \tag{26b}$$

must be computed for $k=1,2,\ldots,N$ and $i,j,l=1,2,\ldots,2d$. Obviously, the integrands in Eqs. (25) and (26) are given as products of harmonics. Thus, exact expressions can be readily found by using orthogonality properties of trigonometric functions over the period T. Accordingly, it is shown in the Appendix that the entries of both the vector $\mathbf{J}(\hat{\boldsymbol{\alpha}})$ and the matrix $\mathbf{\Delta J}(\hat{\boldsymbol{\alpha}})$ in Eq. (20) can be determined exactly.

For arbitrary nonlinearities, the proposed solution scheme can be still applied to an equivalent cubic polynomial whose coefficients are determined by minimizing a mean-square error over the period T. Specifically, introduce for each component $g_i(\hat{\mathbf{Z}})$ of the vector $\mathbf{G}(\hat{\mathbf{Z}})$ the error measure

$$\varepsilon_i=g_i(\hat{\mathbf{Z}})-C_0^{(i)}-\mathbf{C}_1^{(i)T}\hat{\mathbf{Z}}-\mathbf{C}_2^{(i)T}\hat{\mathbf{Z}}^{[2]}-\mathbf{C}_3^{(i)T}\hat{\mathbf{Z}}^{[3]}, \tag{27}$$

where $C_0^{(i)}$ is a constant parameter, and $\mathbf{C}_p^{(i)}$ is a $q\times1$ vector. Then, at each iteration of Newton's method the following minimization problem must be solved:

$$\min_{c_0^{(i)},\mathbf{C}_1^{(i)},\mathbf{C}_2^{(i)},\mathbf{C}_3^{(i)}}\frac{1}{T}\int_0^T\varepsilon_i^2dt. \tag{28}$$

From Eq. (28) the linear system of equations

$$\begin{bmatrix}\mathbf{A}_{11}^{(i)}&\mathbf{A}_{12}^{(i)}&\mathbf{A}_{13}^{(i)}&\mathbf{A}_{14}^{(i)}\\\mathbf{A}_{21}^{(i)}&\mathbf{A}_{22}^{(i)}&\mathbf{A}_{23}^{(i)}&\mathbf{A}_{24}^{(i)}\\\mathbf{A}_{31}^{(i)}&\mathbf{A}_{32}^{(i)}&\mathbf{A}_{33}^{(i)}&\mathbf{A}_{34}^{(i)}\\\mathbf{A}_{41}^{(i)}&\mathbf{A}_{42}^{(i)}&\mathbf{A}_{43}^{(i)}&\mathbf{A}_{44}^{(i)}\end{bmatrix}\begin{bmatrix}C_0^{(i)}\\\mathbf{C}_1^{(i)}\\\mathbf{C}_2^{(i)}\\\mathbf{C}_3^{(i)}\end{bmatrix}=\begin{bmatrix}\mathbf{B}_1^{(i)}\\\mathbf{B}_2^{(i)}\\\mathbf{B}_3^{(i)}\\\mathbf{B}_4^{(i)}\end{bmatrix}, \tag{29}$$

is derived, where the system matrix is symmetric and $\mathbf{A}_{jk}^{(i)}$ are block matrices given by

$$\mathbf{A}_{jk}^{(i)}=\frac{1}{T}\int_0^T\hat{\mathbf{Z}}^{[j-1]}\hat{\mathbf{Z}}^{[k-1]^T}dt, \tag{30}$$

and $\mathbf{B}_j^{(i)}$ are column vectors given by

$$\mathbf{B}_j^{(i)}=\frac{1}{T}\int_0^T\hat{\mathbf{Z}}^{[j-1]}g_i(\hat{\mathbf{Z}})dt, \tag{31}$$

for $j,k=1,\ldots,4$, being $\hat{\mathbf{Z}}^{[0]}=1$. Clearly, exact solutions can be used to compute the integrals defined by Eq. (30), and only the integrals defined by Eq. (31) should be in general computed numerically.

Irrespective of the nature of the system nonlinearities, polynomial or otherwise, from repetitive application of the preceding technique over various realizations of the input process, an approximate solution for the two-sided power spectral density matrix of the zero-mean response can be found using the equation

$$\mathbf{S}_{XX}(\omega_k)\approx\frac{E[\mathbf{U}_k\mathbf{U}_k^T+\mathbf{V}_k\mathbf{V}_k^T]}{4\Delta\omega},\quad k=1,2,\ldots,N. \tag{32}$$

Numerical Examples

Rayleigh Oscillator. To assess the accuracy of the method, consider first the so-called Rayleigh oscillator, in the form

$$\ddot{x}+2\zeta\omega_0(-1+\varepsilon\dot{x}^2)\dot{x}+\omega_0^2x=f(t), \tag{33}$$

with

$$\omega_0=4.0,\quad \zeta=0.05,\quad \varepsilon=4.0. \tag{34}$$

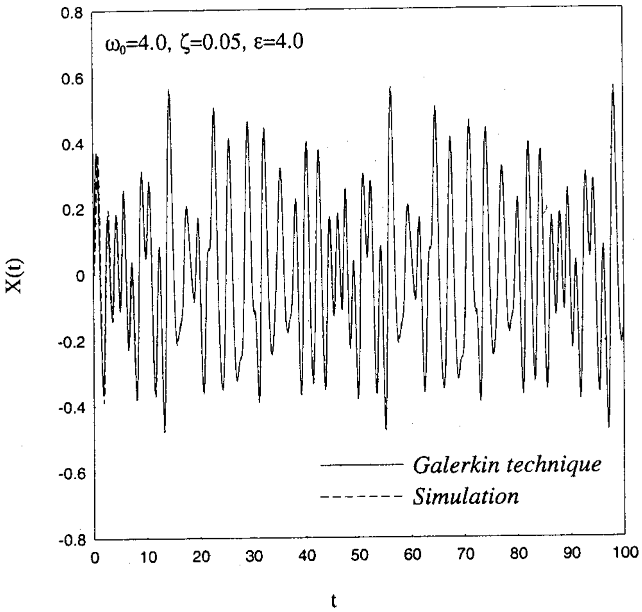


Fig. 1 Rayleigh oscillator, time trajectories of the displacement response

The input excitation $f(t)$ is a Gaussian, zero-mean process with two-sided power spectral density

$$S_{ff}(\omega)=\frac{\left(\dfrac{\omega}{\omega_s}\right)^2}{\left\{\left[1-\left(\dfrac{\omega}{\omega_s}\right)^2\right]^2+\left(2z_g\dfrac{\omega}{\omega_s}\right)^2\right\}^2},\quad \omega_s=2.0,\quad z_g=0.4. \tag{35}$$

The approximate time trajectory of the displacement $x(t)$, Eq. (8), compared to a Monte Carlo record [15], for $\Delta\omega=0.15$ and $N=50$, is shown in Fig. 1. To perform the digital simulation, a standard Runge-Kutta algorithm has been used, with a time step $\Delta t=0.01$. It is seen that, as the system reaches stationarity, the proposed technique reproduces well the response of the oscillator. Figure 2 shows the power spectral density of the response. The new technique and the simulation, both performed over 100 different samples of the excitation process, have yielded virtually identical results. The proposed technique, however, has proved more efficient, as approximately half a dozen iterations at most have been required to converge to the approximate Fourier coefficients of the stationary response, with a relative tolerance equal to 1×10^{-3}.

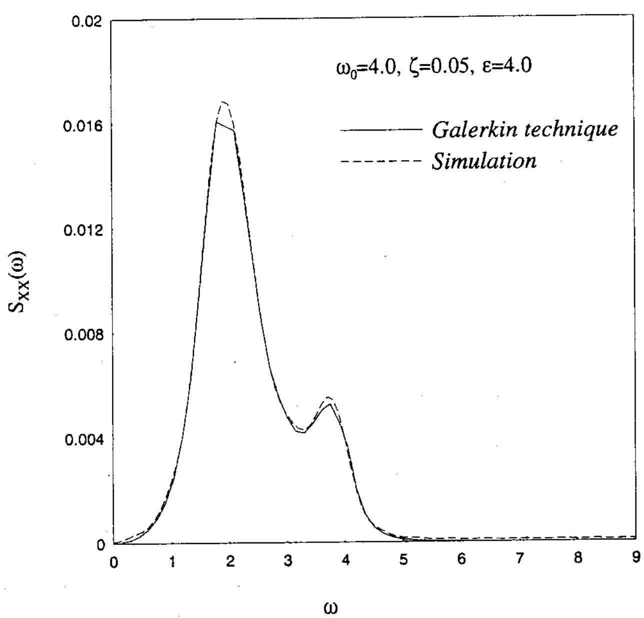


Fig. 2 Rayleigh oscillator, power spectral density of the displacement response

Flow-Induced Nonlinearity. To assess the usefulness of the procedure when dealing with an arbitrary nonlinearity, the four-degree-of-freedom system shown in Fig. 3 is now considered. The motion of this system is governed by Eq. (6), where

$$\mathbf{G}(\mathbf{X},\dot{\mathbf{X}})=\mathbf{G}^c(\dot{\mathbf{X}})+\mathbf{G}^f(\dot{\mathbf{X}}), \tag{36}$$

with

$$G_i^c(\dot{x}_{i-1},\dot{x}_i)=\beta(\dot{x}_i-\dot{x}_{i-1})^3,\quad i=1,\ldots,4 \tag{37a}$$

and

$$G_i^f(\dot{x}_i)=\gamma|a+\dot{x}_i|(a+\dot{x}_i),\quad i=1,\ldots,4. \tag{37b}$$

Note that Eq. (37*b*) is commonly used in offshore engineering for modeling flow-induced drag forces. The symbol $\mathbf{F}(t)$ denotes a vector of zero-mean, Gaussian processes whose *i*th component is given by the equation

$$f_i(t)=\int_{-\infty}^{\infty}h_f(\tau)w(t-\tau)d\tau, \tag{38}$$

where $w(t)$ is a white noise with unit power spectral density, that is,

$$E[w(t)w(t+\tau)]=2\pi\delta(\tau), \tag{39}$$

and $\delta(\tau)$ denotes the Dirac delta function. Further, the filter unit-impulse response function, $h_f(\tau)$, is associated with the transfer function

$$H_f(\omega)=\frac{1}{\sqrt{2(\omega_b-\omega_a)}},\quad \omega_a\leqslant|\omega|\leqslant\omega_b \tag{40a}$$

$$H_f(\omega)=0,\quad \text{everywhere else.} \tag{40b}$$

To apply the proposed technique as formulated in the preceding section, at each iteration of Newton's method the damping terms, Eq. (37*b*), are recast in the equivalent cubic polynomial form,

$$\gamma|a+\dot{x}_i|(a+\dot{x}_i)\approx c_0^{(i)}+c_1^{(i)}\dot{x}_i+c_2^{(i)}\dot{x}_i^2+c_3^{(i)}\dot{x}_i^3, \tag{41}$$

for $i=1,\ldots,4$. Note that, in this case, the minimization procedure (28) leads to the set of equations

$$\begin{bmatrix} a_{11}^{(i)} & a_{12}^{(i)} & a_{13}^{(i)} & a_{14}^{(i)} \\ - & a_{22}^{(i)} & a_{23}^{(i)} & a_{24}^{(i)} \\ - & - & a_{33}^{(i)} & a_{34}^{(i)} \\ - & - & - & a_{44}^{(i)} \end{bmatrix}\begin{bmatrix} c_0^{(i)} \\ c_1^{(i)} \\ c_2^{(i)} \\ c_3^{(i)} \end{bmatrix}=\begin{bmatrix} b_1^{(i)} \\ b_2^{(i)} \\ b_3^{(i)} \\ b_4^{(i)} \end{bmatrix}, \tag{42}$$

where

$$a_{jk}^{(i)}=\frac{1}{T}\int_0^T\dot{x}_i^{j+k-2}dt,\quad j,k=1,\ldots,4 \tag{43}$$

and

$$b_j^{(i)}=\frac{1}{T}\int_0^T\gamma|a+\dot{x}_i|(a+\dot{x}_i)\dot{x}_i^{j-1}dt,\quad j=1,\ldots,4. \tag{44}$$

A standard trapezoidal approximation is applied to compute the integrals (44), while exact solutions can be found for Eq. (43).

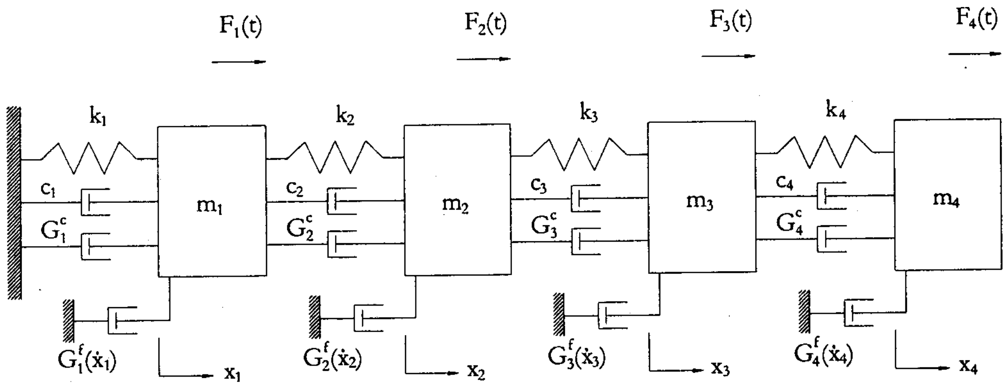


Fig. 3 Four-degree-of-freedom system with cubic damping and flow-induced nonlinearity

For the numerical computations, it is assumed that $\omega_a = 4.0$ and $\omega_b = 8.0$. Further, $\beta = 0.05$, $\gamma = 0.5$, $a = 0.3$, and the following values are selected for the entries of the matrices **M**, **C**, and **K**:

$$m_i = 1.0, \quad i = 1, \ldots, 4 \tag{45a}$$

$$c_i = 0.05, \quad i = 1, \ldots, 4 \tag{45b}$$

$$k_i = 1.0, \quad i = 2,3,4 \tag{45c}$$

and $k_1 = 2.0$. Figure 4 shows the trajectory $x_3(t)$ determined by the proposed technique and by a Monte Carlo simulation, for $\Delta\omega = 0.15$ and $N = 60$. It is seen that, after the transient initial discrepancies, an excellent agreement exists between the proposed technique results and the simulated data, as the system response reaches stationarity. The simulated records have been obtained by a standard Runge-Kutta algorithm with a time step $\Delta t = 0.01$. Figure 5 shows the power spectral density of the zero-mean response $x_4(t)$, as determined by the proposed technique and by the Monte Carlo study. It is seen that the two approaches yield virtually identical results. However, the proposed technique has been found more efficient than the digital simulation, even accounting for the cubicization procedure, Eq. (42). Similar results, in terms of reliability and efficiency, have been found for the rest of the system coordinates; they are excluded for brevity.

Concluding Remarks

A computationally efficient Galerkin or a harmonic balance technique has been used to estimate spectral properties of randomly excited MDOF nonlinear systems. By expanding the stationary system response in a suitable Fourier series, it has been shown that single realizations of the response process can be readily generated by solving a set of nonlinear algebraic equations derived by harmonic balance. Specifically, an expeditious solution scheme based on Newton's method has been developed by using exact solutions for the Fourier coefficients of the nonlinear terms, which are available for polynomial nonlinearities up to cubic order. Further, a simple cubicization procedure has been developed to approximate arbitrary nonlinearities by cubic polynomials. In this manner, the method has been made more efficient than a previous formulation requiring computationally costly use of the FFT technique to determine the Fourier coefficients of the nonlinear terms.

Numerical examples have shown the accuracy of the estimates both in terms of the response time trajectories and the power spectral density matrix. From a computational point of view, the technique has considerable advantages as compared to digital simulation, where the computational costs may become considerable due to long transient parts of the trajectories, which cannot be used to determine the statistical behavior of the stationary re-

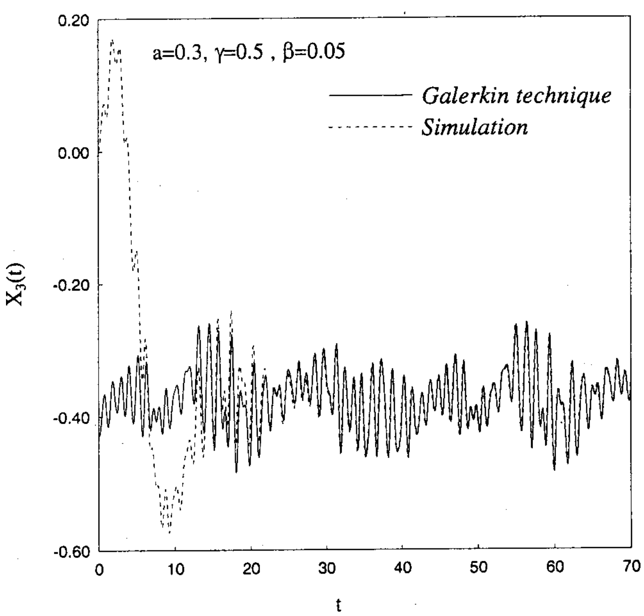


Fig. 4 Four-degree-of-freedom system, time trajectories of the displacement response of mass 3

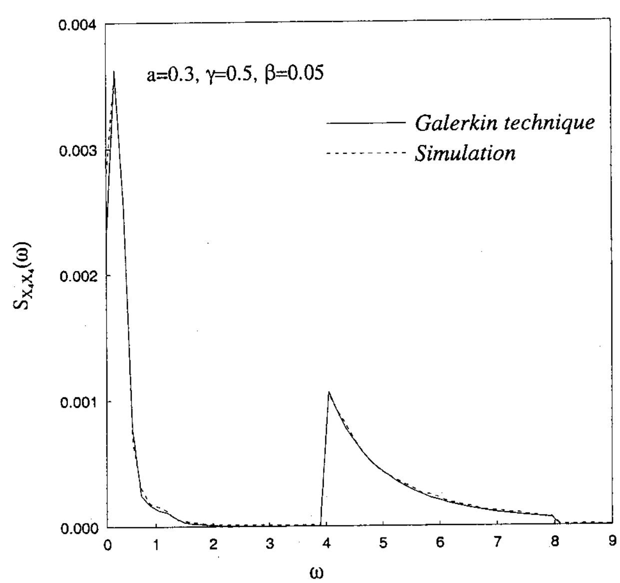


Fig. 5 Four-degree-of-freedom system, power spectral density of the zero-mean displacement response of mass 4

sponse process. Further, note that when dealing with nonlinear problems, the time step must be adequately small to eliminate numerical instabilities in the numerical integration algorithm, as well as aliasing when the FFT technique is applied to the discrete response record.

Irrespective of computational advantages, it is pointed out that Eq. (18), which is derived by the proposed technique, constitutes a nonlinear relationship between input/output parameters. Thus, it can be treated as a constraint in conjunction with reliability and optimization calculations of randomly excited dynamic systems [16]. Finally, it must be noted that the proposed technique has been found reliable for systems with nonlinear damping, for which a rapid convergence of Newton's method is achieved, regardless of the initial guess for the unknown Fourier coefficients of the response [12]. However, the applicability of the proposed technique for systems with nonlinear stiffness must be further investigated to account for jump phenomena, and presence of additional harmonics in the system response [17]. In this regard, existing analytical solutions [18,19] for specific nonlinear systems can be used as benchmarks.

Acknowledgment

The support of this work by a grant from the USA Office of Naval Research is gratefully acknowledged.

Appendix

This appendix addresses the evaluation of the integrals of the nonlinear terms involved in the solution of Eq. (18) by Newton's method. Specifically it is shown that, when the system nonlinearity assumes the cubic polynomial form (22), the integrals related to the vector $\mathbf{J}(\hat{\boldsymbol{\alpha}})$, given by

$$I_c=\int_0^T \mathbf{G}(\hat{\boldsymbol{\alpha}},t)\cos(\omega_k t)dt, \tag{A1}$$

$$I_s=\int_0^T \mathbf{G}(\hat{\boldsymbol{\alpha}},t)\sin(\omega_k t)dt, \tag{A2}$$

$$I=\int_0^T \mathbf{G}(\hat{\boldsymbol{\alpha}},t)dt, \tag{A3}$$

as well as the integrals related to the Jacobian matrix $\boldsymbol{\Delta}\mathbf{J}(\hat{\boldsymbol{\alpha}})$,

$$J_c=\int_0^T \frac{\partial \mathbf{G}(\hat{\boldsymbol{\alpha}},t)}{\partial \hat{\alpha}_j}\cos(\omega_k t)dt, \tag{A4}$$

$$J_s=\int_0^T \frac{\partial \mathbf{G}(\hat{\boldsymbol{\alpha}},t)}{\partial \hat{\alpha}_j}\sin(\omega_k t)dt, \tag{A5}$$

$$J=\int_0^T \frac{\partial \mathbf{G}(\hat{\boldsymbol{\alpha}},t)}{\partial \hat{\alpha}_j}dt, \tag{A6}$$

for $j=1,2,\ldots,(2N+1)d$, can be determined in a closed form.

First, consider Eqs. ($A1$)–($A3$). As long as Eq. (22) holds, the terms in Eq. (25) can be found by determining the integrals

$$Q_{ss}=\int_0^T \hat{z}_i \sin(\omega_r t)\sin(\omega_s t)dt, \tag{A7}$$

$$Q_{sc}=\int_0^T \hat{z}_i \sin(\omega_r t)\cos(\omega_s t)dt, \tag{A8}$$

$$Q_{cc}=\int_0^T \hat{z}_i \cos(\omega_r t)\cos(\omega_s t)dt, \quad r,s=1,2,\ldots,N \tag{A9}$$

for any $i=1,2,\ldots,2d$ [the superscript N in $\hat{z}_i$, which is given as Eq. (23), here is omitted for brevity]. Recognize that such integrals are also required in evaluating the entries of the Jacobian matrix, that is, the terms related to $\hat{\mathbf{Z}}^{[2]}$ in Eqs. ($A4$) and ($A5$), for j=1,2, . . . ,2N×d.Rudimentary algebraleadsto

$$Q_{ss}=\mu_{z_i}\frac{\pi}{\Delta\omega}\delta_{rs}+\frac{\pi}{2\Delta\omega}(\hat{U}^{z_i}_{|r-s|}-\hat{U}^{z_i}_{r+s}), \tag{A10}$$

$$Q_{sc}=\frac{\pi}{2\Delta\omega}\left(\frac{|r-s|}{r-s}\hat{V}^{z_i}_{|r-s|}+\hat{V}^{z_i}_{r+s}\right), \tag{A11}$$

and

$$Q_{cc}=\mu_{z_i}\frac{\pi}{\Delta\omega}\delta_{rs}+\frac{\pi}{2\Delta\omega}(\hat{U}^{z_i}_{|r-s|}+\hat{U}^{z_i}_{r+s}). \tag{A12}$$

The symbol δ_{rs} denotes the well-known Kronecker delta. To determine the terms involved in Eq. (26), the integrals

$$P_{ss}=\int_0^T \hat{z}_i\hat{z}_j \sin(\omega_r t)\sin(\omega_s t)dt, \tag{A13}$$

$$P_{sc}=\int_0^T \hat{z}_i\hat{z}_j \sin(\omega_r t)\cos(\omega_s t)dt, \tag{A14}$$

$$P_{cc}=\int_0^T \hat{z}_i\hat{z}_j \cos(\omega_r t)\cos(\omega_s t)dt, \quad r,s=1,2,\ldots,N \tag{A15}$$

must be computed for any $i,j=1,2,\ldots,2d$. Again, recognize that such integrals are also involved in the evaluation of the entries of the Jacobian matrix, that is, the terms related to $\hat{\mathbf{Z}}^{[3]}$ in Eqs. ($A4$) and ($A5$). Again, rudimentary algebra leads to

$$P_{ss}=\frac{\pi}{2\Delta\omega}[\mu_{z_i}(\hat{U}^{z_j}_{|r-s|}-\hat{U}^{z_j}_{r+s})+\mu_{z_j}(\hat{U}^{z_i}_{|r-s|}-\hat{U}^{z_i}_{r+s})]$$
$$+\frac{\pi}{\Delta\omega}\mu_{z_i}\mu_{z_j}\delta_{rs}+\hat{I}_1+\hat{I}_2, \tag{A16}$$

where

$$\hat{I}_1=-\sum_{k=1}^{N-n_s}\frac{\pi}{4\Delta\omega}[\hat{U}^{z_i}_k\cdot\hat{U}^{z_j}_{k+n_s}+\hat{U}^{z_j}_k\cdot\hat{U}^{z_i}_{k+n_s}+\hat{V}^{z_i}_k\cdot\hat{V}^{z_j}_{k+n_s}$$
$$+\hat{V}^{z_j}_k\cdot\hat{V}^{z_i}_{k+n_s}]+\sum_{k=1}^{n_s-1}\frac{\pi}{8\Delta\omega}[-\hat{U}^{z_i}_k\cdot\hat{U}^{z_j}_{n_s-k}-\hat{U}^{z_j}_k\cdot\hat{U}^{z_i}_{n_s-k}$$
$$+\hat{V}^{z_i}_k\cdot\hat{V}^{z_j}_{n_s-k}+\hat{V}^{z_j}_k\cdot\hat{V}^{z_i}_{n_s-k}] \tag{A17}$$

and

$$\hat{I}_2=\sum_{k=1}^{N-n_d}\frac{\pi}{4\Delta\omega}[\hat{U}^{z_i}_k\cdot\hat{U}^{z_j}_{k+n_d}+\hat{U}^{z_j}_k\cdot\hat{U}^{z_i}_{k+n_d}+\hat{V}^{z_i}_k\cdot\hat{V}^{z_j}_{k+n_d}$$
$$+\hat{V}^{z_j}_k\cdot\hat{V}^{z_i}_{k+n_d}]+\sum_{k=1}^{n_d-1}\frac{\pi}{8\Delta\omega}[\hat{U}^{z_i}_k\cdot\hat{U}^{z_j}_{n_d-k}$$
$$+\hat{U}^{z_j}_k\cdot\hat{U}^{z_i}_{n_d-k}-\hat{V}^{z_i}_k\cdot\hat{V}^{z_j}_{n_d-k}-\hat{V}^{z_j}_k\cdot\hat{V}^{z_i}_{n_d-k}], \tag{A18}$$

with $n_s=r+s$, $n_d=|r-s|$;

$$P_{sc}=\frac{\pi}{2\Delta\omega}\left[\mu_{z_i}\left(\frac{|r-s|}{r-s}\hat{V}^{z_j}_{|r-s|}+\hat{V}^{z_j}_{r+s}\right)\right.$$
$$\left.+\mu_{z_j}\left(\frac{|r-s|}{r-s}\hat{V}^{z_i}_{|r-s|}+\hat{V}^{z_i}_{r+s}\right)\right]+\hat{J}_1+\hat{J}_2, \tag{A19}$$

where

$$\hat{J}_1=\sum_{k=1}^{N-n_s}\frac{\pi}{4\Delta\omega}[\hat{U}_k^{z_j}\cdot\hat{V}_{k+n_s}^{z_i}+\hat{U}_k^{z_i}\cdot\hat{V}_{k+n_s}^{z_j}-\hat{V}_k^{z_j}\cdot\hat{U}_{k+n_s}^{z_i}-\hat{V}_k^{z_i}\cdot\hat{U}_{k+n_s}^{z_j}]+\sum_{k=1}^{n_s-1}\frac{\pi}{8\Delta\omega}[\hat{U}_k^{z_i}\cdot\hat{V}_{n_s-k}^{z_j}+\hat{U}_k^{z_j}\cdot\hat{V}_{n_s-k}^{z_i}+\hat{V}_k^{z_j}\cdot\hat{U}_{n_s-k}^{z_i}+\hat{V}_k^{z_i}\cdot\hat{U}_{n_s-k}^{z_j}], \tag{A20}$$

$$\hat{J}_2=\sum_{k=1}^{N-n_d}\frac{\pi}{4\Delta\omega}\frac{|r-s|}{r-s}[\hat{U}_k^{z_j}\cdot\hat{V}_{k+n_d}^{z_i}+\hat{U}_k^{z_i}\cdot\hat{V}_{k+n_d}^{z_j}-\hat{V}_k^{z_j}\cdot\hat{U}_{k+n_d}^{z_i}-\hat{V}_k^{z_i}\cdot\hat{U}_{k+n_d}^{z_j}]+\sum_{k=1}^{n_d-1}\frac{\pi}{8\Delta\omega}\frac{|r-s|}{r-s}[\hat{U}_k^{z_j}\cdot\hat{V}_{n_d-k}^{z_i}+\hat{U}_k^{z_i}\cdot\hat{V}_{n_d-k}^{z_j}+\hat{V}_k^{z_j}\cdot\hat{U}_{n_d-k}^{z_i}+\hat{V}_k^{z_i}\cdot\hat{U}_{n_d-k}^{z_j}]; \tag{A21}$$

$$P_{cc}=\frac{\pi}{2\Delta\omega}[\mu_{z_i}(\hat{U}_{|r-s|}^{z_j}+\hat{U}_{r+s}^{z_j})+\mu_{z_j}(\hat{U}_{|r-s|}^{z_i}+\hat{U}_{r+s}^{z_i})]+\frac{\pi}{\Delta\omega}\mu_{z_i}\mu_{z_j}\delta_{rs}+\tilde{J}_1+\tilde{J}_2, \tag{A22}$$

where

$$\tilde{J}_1=\sum_{k=1}^{N-n_s}\frac{\pi}{4\Delta\omega}[\hat{U}_k^{z_i}\cdot\hat{U}_{k+n_s}^{z_j}+\hat{U}_k^{z_j}\cdot\hat{U}_{k+n_s}^{z_i}+\hat{V}_k^{z_i}\cdot\hat{V}_{k+n_s}^{z_j}+\hat{V}_k^{z_j}\cdot\hat{V}_{k+n_s}^{z_i}]+\sum_{k=1}^{n_s-1}\frac{\pi}{8\Delta\omega}[\hat{U}_k^{z_i}\cdot\hat{U}_{n_s-k}^{z_j}+\hat{U}_k^{z_j}\cdot\hat{U}_{n_s-k}^{z_i}-\hat{V}_k^{z_i}\cdot\hat{V}_{n_s-k}^{z_j}-\hat{V}_k^{z_j}\cdot\hat{V}_{n_s-k}^{z_i}] \tag{A23}$$

and

$$\tilde{J}_2=\sum_{k=1}^{N-n_d}\frac{\pi}{4\Delta\omega}[\hat{U}_k^{z_i}\cdot\hat{U}_{k+n_d}^{z_j}+\hat{U}_k^{z_j}\cdot\hat{U}_{k+n_d}^{z_i}+\hat{V}_k^{z_i}\cdot\hat{V}_{k+n_d}^{z_j}+\hat{V}_k^{z_j}\cdot\hat{V}_{k+n_d}^{z_i}]+\sum_{k=1}^{n_d-1}\frac{\pi}{8\Delta\omega}[\hat{U}_k^{z_i}\cdot\hat{U}_{n_d-k}^{z_j}+\hat{U}_k^{z_j}\cdot\hat{U}_{n_d-k}^{z_i}-\hat{V}_k^{z_i}\cdot\hat{V}_{n_d-k}^{z_j}-\hat{V}_k^{z_j}\cdot\hat{V}_{n_d-k}^{z_i}]. \tag{A24}$$

References

[1] Schetzen, M., 1980, *The Volterra and Wiener Theories of Nonlinear Systems*, Wiley, New York.

[2] Spanos, P. D., and Donley, M. G., 1992, ''Non-Linear Multi-Degree-of-Freedom System Random Vibration by Equivalent Statistical Quadratization,'' Int. J. Non-Linear Mech., **27**, pp. 735–748.

[3] Tognarelli, M. A., Zhao, J., Rao, K. B., and Kareem, A., 1997, ''Equivalent Statistical Quadratization and Cubicization for Nonlinear Systems,'' J. Eng. Mech., **123**, pp. 512–523.

[4] Li, X. M., 1998, *Stochastic Response of Offshore Platforms*, Computational Mechanics Publications, Southampton, U.K.

[5] Roberts, J. B., and Spanos, P. D., 1990, *Random Vibration and Statistical Linearization*, Wiley, New York.

[6] Iwan, W. D., and Spanos, P. D., 1978, ''Response Envelope Statistics for Nonlinear Oscillators With Random Excitation,'' J. Appl. Mech., **45**, pp. 170–174.

[7] Bellizzi, S., and Bouc, R., 1999, ''Analysis of Multi-Degree of Freedom Strongly Non-Linear Mechanical Systems With Random Input, Part I,'' Probab. Eng. Mech., **14**, pp. 229–244.

[8] Bellizzi, S., and Bouc, R., 1999, ''Analysis of Multi-Degree of Freedom Strongly Non-Linear Mechanical Systems With Random Input, Part II,'' Probab. Eng. Mech., **14**, pp. 245–256.

[9] Bouc, R., and Defilippi, M., 1997, ''Multimodal Nonlinear Spectral Response of a Beam With Impact Under Random Input,'' Probab. Eng. Mech., **12**, pp. 163–170.

[10] Bellizzi, S., Bouc, R., Defilippi, M., and Guihot, P., 1998, ''Response Spectral Densities and Identification of a Randomly Excited Non-Linear Squeeze Film Oscillator,'' Mech. Syst. Signal Process., **12**, pp. 693–711.

[11] Bouc, R., and Defilippi, M., 1987, ''A Galerkin Multiharmonic Procedure for Nonlinear Multidimensional Random Vibration,'' Int. J. Eng. Sci., **25**, pp. 723–733.

[12] Spanos, P. D., Di Paola, M., and Failla G., 2002, ''A Galerkin Approach for Power Spectrum Determination of Nonlinear Oscillators,'' Meccanica, International Journal of the Italian Association of Theoretical and Applied Mechanics, **37**, pp. 51–65.

[13] Papoulis, A., 1965, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York.

[14] Priestley, M. B., 1981, *Spectral Analysis and Time Series*, Academic, New York.

[15] Spanos, P. D., and Zeldin, B. A., 1998, ''Monte Carlo Treatment of Random Fields: A Broad Perspective,'' Appl. Mech. Rev., **51**, pp. 219–237.

[16] Madsen, H. O., Krenk, S., and Lind, N. C., 1986, *Methods of Structural Safety*, Prentice-Hall, Englewood Cliffs, New Jersey.

[17] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, Wiley, New York.

[18] Dimentberg, M. F., Hou, Z., and Noori, M., 1995, ''Spectral Density of a Non-Linear Single-Degree-of-Freedom System's Response to a White-Noise Random Excitation: A Unique Case of an Exact Solution,'' Int. J. Non-Linear Mech., **30**, pp. 673–676.

[19] Dimentberg, M. F., and Haenisch, H. G., 1998, ''Pseudo-Linear Vibro-Impact System With a Secondary Structure: Response to a White-Noise Excitation,'' J. Appl. Mech., **65**, pp. 772–774.

H. Luş
Assistant Professor,
Dept. of Civil Eng., Boğaziçi University, Bebek,
34342, Istanbul, Turkey

M. De Angelis
Assistant Professor,
Dip. Ing. Strutt. e Geotecnica, Univ. di Roma "La Sapienza," Rome, Italy

R. Betti
Professor,
Dept. of Civil Eng. and Eng. Mechanics,
Columbia University, New York, NY 10027

A New Approach for Reduced Order Modeling of Mechanical Systems Using Vibration Measurements

This study investigates the possibilities of obtaining reduced order mass-damping-stiffness models of mechanical systems using state space realizations identified via dynamic tests. It is shown that even when the system is insufficiently instrumented with sensors and actuators, it is still possible to create physically meaningful reduced order mass-damping-stiffness models that incorporate measured and unmeasured degrees of freedom. It is further discussed that certain assumptions, such as having a diagonal mass matrix or having classical damping in the system, allow one to develop alternative reduced order representations with different physical interpretations. The theoretical presentation is supplemented by a numerical example that illustrates the applications of the formulations developed herein. [DOI: 10.1115/1.1602482]

1 Introduction

A well-known technique employed in modeling the dynamics of mechanical systems is the use of second-order matrix differential equations. In such a formulation, the coefficient matrices contain the physical mass, damping, and stiffness parameters of the system, which in turn affect the modal vibrational parameters such as the natural frequencies and mode shapes. The construction of a mass-damping-stiffness model based on material properties and the system's geometry, as done in finite element analysis, is a relatively straightforward procedure and is widely employed in predicting the response of a structure to prescribed inputs (forward analysis). The identification of such a model from the measured dynamic response, on the other hand, has proven to be a tough challenge, and it is often referred to as the "(linear) inverse vibration problem."

This problem has been addressed by various scholars in the past as evidenced by the works of Agbabian et al. [1], Mottershead and Friswell [2], Berman [3], Baruch [4], Beck and Katafygiotis [5], Alvin et al. [6,7], Tseng et al. [8,9], and Balmès [10]. Recently the authors have presented a solution to this problem based on identified state space realizations ([11,12]). This solution has proven to be more flexible and general than the previously available solutions of the problem utilizing state space realizations, and it has been used effectively in estimating the physical parameters of various structural models ([13]). On the other hand, even though the requirements on the number of available sensors and actuators for a full order identification has been improved with the aforementioned solution, the question of obtaining reduced order models in the absence of full instrumentation has not been fully investigated yet. A noteworthy exception to this claim is the study by Alvin et al. [7] in which the authors have provided a methodology that utilizes undamped second-order frequencies and mass normalized mode shapes to construct reduced order mass, damping, and stiffness matrices. The current study therefore is concerned with developing formulations that address the problem of insufficient instrumentation, and attempts at providing new methods based on identified state space models that utilize both damped and undamped modal information.

The following sections are devoted to the review of the full order modeling problem, effects of insufficient instrumentation, possible reduced order modeling schemes, and alternative formulations that can be developed by considering frequently employed assumptions such as having a diagonal mass matrix and/or a classically damped system. The proposed methodology/solution consists of three well-defined phases. First, a first-order model of the system is determined using the recorded input-output data. Once a first-order model has been determined, the next step is to construct the transformation matrix that relates the arbitrary coordinates of the identified state space model to the set of modal coordinates that are derived via the symmetric eigenvalue problem formulation. The construction of such a transformation matrix has been shown to be possible if there is (at least) one co-located sensor-actuator pair ([11,12]). With the use of this transformation, it will be shown to be possible to evaluate some partitions of the complex eigenvector matrix of the symmetric eigenvalue problem. Utilizing such information, the last step consists of constructing reduced order mass-damping-stiffness models based on the identified complex eigenvalues and partitions of the complex eigenvector matrix. It will be shown in the latter sections that such reduced order models are physically meaningful, and furthermore, that they can be analytically related to the full order matrices. The last section is devoted to the presentation of a numerical example which illustrates the applicability of the solution and the formulations developed in this study.

2 Statement of the Problem

Consider an N degree of freedom viscously damped linear structural system, subjected to r external excitations. The equations of motion for such a system can be expressed as

$$\mathcal{M}\ddot{\mathbf{q}}(t)+\mathcal{L}\dot{\mathbf{q}}(t)+\mathcal{K}\mathbf{q}(t)=\mathcal{B}\mathbf{u}(t), \tag{1}$$

where $\mathbf{q}(t)$ indicates the vector of the generalized nodal displacements, with $(\dot{\ })$ and $(\ddot{\ })$ representing respectively the first- and second-order derivatives with respect to time. The vector $\mathbf{u}(t)$, of dimension $r\times 1$, is the input vector containing the r external ex-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 5, 2002; final revision, Jan. 28, 2003. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

citations acting on the system, with $\mathcal{B} \in \mathrm{Re}^{N\times r}$ being the input matrix that relates the inputs to the DOF's. The matrices $\mathcal{M} \in \mathrm{Re}^{N\times N}$, $\mathcal{L} \in \mathrm{Re}^{N\times N}$, and $\mathcal{K} \in \mathrm{Re}^{N\times N}$ are the symmetric positive definite mass, damping, and stiffness matrices, respectively. Let us assume that only m output time histories of the structural response are available, so that the measurement vector $\mathbf{y}(t)$, of dimensions $m\times 1$, can be written as

$$\mathbf{y}(t)=[(\mathcal{C}_p\mathbf{q}(t))^T \;\; (\mathcal{C}_v\dot{\mathbf{q}}(t))^T \;\; (\mathcal{C}_a\ddot{\mathbf{q}}(t))^T]^T, \tag{2}$$

where the matrices $\mathcal{C}_p$, $\mathcal{C}_v$, and $\mathcal{C}_a$ relate the measurements to positions, velocities, and accelerations, respectively, and the superscript $(\;)^T$ denotes the transpose.

While the cases of a complete set of sensors ($m=N$ ([6])) and of a complete set of actuators ($r=N$ ([8,9])) have been previously addressed, the "more general" case of a sufficient number of sensors and actuators ($m+r=N+1$) with one co-located sensor-actuator pair has only recently been studied by the authors ([11,12]). In these recent studies the basic assumption is that, at each degree of freedom of the system, there is either a sensor or an actuator with at least one degree of freedom having a co-located sensor-actuator pair. In the present study, this assumption is removed and the analysis focuses on the case where, still considering a co-located sensor-actuator pair, a sufficient number of sensors and actuators is not available ($m+r<N+1$) so that there will be degrees of freedom with neither a sensor nor an actuator. This is a common scenario in real life applications where only limited testing and measuring equipment is available. However, even with these limitations, some dynamic characteristics of the structural system can be retrieved and a "reduced" second-order model of the "larger" structural system can still be obtained.

3 Transformation to a First-Order Modal Model

A well-known fact from control theory is that it is possible (and, in some cases, convenient) to transform the system of second-order differential equations of motion into a system of first-order differential equations by introducing a state vector $\mathbf{z}(t)=[\mathbf{q}(t)^T\dot{\mathbf{q}}(t)^T]^T$. As discussed in the works of Luş [11] and De Angelis et al. [12], the equations of motion (1) and the output equations (2) can be conveniently rewritten as

$$\begin{bmatrix}\mathcal{L} & \mathcal{M}\\ \mathcal{M} & \mathbf{0}\end{bmatrix}\dot{\mathbf{z}}(t)+\begin{bmatrix}\mathcal{K} & \mathbf{0}\\ \mathbf{0} & -\mathcal{M}\end{bmatrix}\mathbf{z}(t)=\begin{bmatrix}\mathcal{B}\\ \mathbf{0}\end{bmatrix}\mathbf{u}(t), \tag{3a}$$

$$\mathbf{y}(t)=[\mathcal{C}_p \;\; \mathbf{0}]\mathbf{z}(t), \tag{3b}$$

where, for ease of exposition, we have considered only position measurements in the output equation (3b). It should be pointed out, however, that the following results are valid for any type of measurements (positions, velocities, or accelerations), as discussed in Refs. [11], [12]. The advantage of rewriting Eq. (1) into Eqs. (3) is that now the associated eigenvalue problem in the state space formulation preserves its symmetry, and this yields a great advantage in posing the identification problem, as will be shown in the following formulations. By indicating with $\psi_{N\times 2N}=[\psi_1\psi_2\cdots\psi_{2N}]$ the matrix containing the eigenvectors of the complex eigenvalue problem

$$(\lambda_i^2\mathcal{M}+\lambda_i\mathcal{L}+\mathcal{K})\psi_i=0 \tag{4}$$

and with $\mathbf{\Lambda}_{2N\times 2N}$ the diagonal matrix containing all the complex eigenvalues λ_i $(i=1,2,\ldots,2N)$, it is possible to rewrite Eqs. (3) in a modal form. Since the eigenvectors ψ_i $(i=1,2,\ldots,2N)$ can be arbitrarily scaled, the scaling choice considered in this study is such that (see Sestieri and Ibrahim [14], Balmès [10])

$$\begin{bmatrix}\psi\\ \psi\mathbf{\Lambda}\end{bmatrix}^T\begin{bmatrix}\mathcal{L} & \mathcal{M}\\ \mathcal{M} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\psi\\ \psi\mathbf{\Lambda}\end{bmatrix}=\mathbf{I}, \tag{5a}$$

$$\begin{bmatrix}\psi\\ \psi\mathbf{\Lambda}\end{bmatrix}^T\begin{bmatrix}\mathcal{K} & \mathbf{0}\\ \mathbf{0} & -\mathcal{M}\end{bmatrix}\begin{bmatrix}\psi\\ \psi\mathbf{\Lambda}\end{bmatrix}=-\mathbf{\Lambda}, \tag{5b}$$

and with this scaling, for a *proportionally damped system*, the real and imaginary parts of the components of these complex eigenvectors are equal in magnitude. By using the transformation $\mathbf{z}(t)=[\psi^T(\psi\mathbf{\Lambda})^T]^T\boldsymbol{\zeta}(t)$, and pre-multiplying Eq. (3a) by $[\psi^T(\psi\mathbf{\Lambda})^T]$, the equations in modal coordinates can be written as

$$\dot{\boldsymbol{\zeta}}(t)=\mathbf{\Lambda}\boldsymbol{\zeta}(t)+\psi^T\mathcal{B}\mathbf{u}(t), \tag{6a}$$

$$\mathbf{y}(t)=\mathcal{C}_p\psi\boldsymbol{\zeta}(t). \tag{6b}$$

4 Determination of a "Reduced" First-Order Model of the System

Here we assume that a state space realization (in some arbitrary basis) of the dynamic system under investigation has been obtained using general input/output data. Such a realization can be expressed as

$$\begin{aligned}\dot{\mathbf{x}}(t)&=\mathbf{A}_C\mathbf{x}(t)+\mathbf{B}_C\mathbf{u}(t),\\ \mathbf{y}(t)&=\mathbf{C}_C\mathbf{x}(t)+\mathbf{D}_C\mathbf{u}(t),\end{aligned} \tag{7}$$

where $\mathbf{A}_C\in\mathrm{Re}^{2N\times 2N}$, $\mathbf{B}_C\in\mathrm{Re}^{2N\times r}$, $\mathbf{C}_C\in\mathrm{Re}^{m\times 2N}$, and $\mathbf{D}_C\in\mathrm{Re}^{m\times r}$ are the continuous time system matrices. In this study, an ERA/OKID based approach, as discussed by Juang and co-workers [15,16] and Luş and co-workers [17,18], is considered for the identification of the discrete time system matrices of a state space realization, and these discrete time matrices are converted to their continuous time counterparts using the zero-order hold assumption. Here we assume that the external excitation is sufficiently rich so that all the vibrational modes of the structure are adequately excited. The advantages of using the ERA/OKID based approach are: (1) no data manipulation (integration or differentiation) is needed, and (2) it has proven to be quite effective in accurately identifying the dynamic characteristics of complex systems using very limited sets of sensors and actuators, even in the presence of noise (Luş [11]). Hence it is very appropriate for the purpose of this study, which is to analyze the case when the number of sensors and actuators available is much smaller than the number of degrees of freedom of the structure.

By considering the transformation $\mathbf{x}(t)=\varphi\theta(t)$, the continuous time system of Eqs. (7) can also be written in modal coordinates as

$$\dot{\theta}(t)=\mathbf{\Lambda}'\theta(t)+\varphi^{-1}\mathbf{B}_C\mathbf{u}(t), \tag{8a}$$

$$\mathbf{y}(t)=\mathbf{C}_C\varphi\theta(t), \tag{8b}$$

where the matrix $\mathbf{\Lambda}'$ contains the $2N$ continuous time eigenvalues of the identified state space model, and φ, of size $2N\times 2N$, is the matrix of the corresponding eigenvectors. The matrix $\mathbf{D}_C$ has been omitted in Eq. (8b) because it is independent of coordinate transformations, and its presence or absence does not in any way alter the development of subsequent results. At this point, it is noteworthy that, since the dimensions of $(\varphi^{-1}\mathbf{B}_C)$ are $2N\times r$ and those of $(\mathbf{C}_C\varphi)$ are $m\times 2N$, the modal model represented in Eq. (8b) can "only" be used for sensors and actuators placed at locations specified by the experimental setup used in dynamic testing. This limitation will be overcome by the proposed approach, since it will allow us to expand the input/output mapping to include "new" sensor/actuator locations.

4.1 Identifiable Partitions of the Complex Eigenvector Matrix. If the first-order system of Eqs. (7) was identified using data that actually came from the second-order model of Eq. (1), the models represented by Eqs. (6) and (8) are different models of the same system, with the same set of eigenvalues. Therefore there must be a transformation matrix $\mathcal{T}$ that relates these two representations, so that:

$$\mathcal{T}^{-1}\mathbf{\Lambda}'\mathcal{T}=\mathbf{\Lambda}, \tag{9a}$$

$$\mathcal{T}^{-1}\varphi^{-1}\mathbf{B}_C=\psi^T\mathcal{B}, \tag{9b}$$

$$\mathbf{C}_C \varphi \mathcal{T} = \mathcal{C}_p \psi. \tag{9c}$$

The matrix $\mathbf{\Lambda}$, which belongs to the state space model of Eqs. (6), is always diagonal since it is obtained via a symmetric eigenvalue problem. On the other hand, the matrix $\mathbf{\Lambda}'$ of Eqs. (8) comes from an asymmetric eigenvalue problem, and it is known that diagonalization is not always guaranteed in an asymmetric eigenvalue problem. In this case, however, since the asymmetric eigenvalue problem is in fact derived from the symmetric eigenvalue problem, there will always exist a set of eigenvectors that will yield $\mathbf{\Lambda}' = \mathbf{\Lambda}$ (see Appendix A for a proof of this statement). Consequently, the transformation matrix $\mathcal{T}$ is also a diagonal matrix, denoted as $\mathcal{T} = \mathrm{diag}(t_1, t_2, \ldots, t_{2N})$. Furthermore, it is assumed that the structure is properly constrained (which generically is the case in a modal testing situation) so that there are no rigid body modes. In addition, the input and output matrices ($\mathcal{B}$ and $\mathcal{C}_p$, respectively) of the finite element model, which contain information about the actuator and sensor locations, are assumed to be known.

The identification of the transformation matrix is contingent upon the existence of (at least) one co-located sensor actuator pair. To briefly summarize the procedure presented by Luş [11] and De Angelis et al. [12], let us assume that there is a co-located sensor-actuator pair at the generic jth degree of freedom (DOF). The co-location requirement may be written as

$$\begin{aligned} &\mathcal{C}_p(\text{row corresponding to the } j\text{th DOF},:)\psi \\ &\quad = (\psi^T \mathcal{B}(:,\text{column corresponding to the } j\text{th DOF}))^T. \end{aligned} \tag{10}$$

For ease of presentation, we shall resort to the notation $\mathbf{M}(j,:)$ to denote the row of a generic matrix $\mathbf{M}$ that corresponds to the jth DOF, and we shall denote by $\mathbf{M}(:,j)$ the column of that matrix corresponding to the jth DOF. Hence, the co-location requirement is given by $\mathcal{C}_p(j,:)\psi = (\psi^T \mathcal{B}(:,j))^T$, or, with the use of the transformation equations,

$$\mathbf{C}_C(j,:)\varphi \mathcal{T}^2 = (\varphi^{-1}\mathbf{B}_C(:,j))^T, \tag{11}$$

which yields $2N$ equations for the $2N$ unknowns in the diagonal matrix $\mathcal{T}$. It should be noted that the co-location requirement does not need to be satisfied in a strict sense. For example, in the case of a rigid body, such a requirement may be rephrased in alternative forms by properly combining the inputs and outputs. In any case, it is evident that the total number of available sensors and actuators has no bearing on the identifiability of the transformation matrix, rather it is in the determination of the eigenvector matrix ψ that the limitation imposed by the insufficient set of sensors/actuators appears. Some components of the matrix ψ can be determined using the information contained in the input and output matrices. When there is a sensor at the kth DOF, then the kth row of the matrix ψ can be evaluated via Eq. (9c), which may be written as

$$\psi(k,:) = \mathbf{C}_C(k,:)\varphi \mathcal{T}. \tag{12}$$

On the other hand, if there is an actuator located at the kth DOF, then the kth row of the matrix ψ can be obtained using Eq. (9b), i.e.,

$$\psi(k,:) = (\mathcal{T}^{-1}\varphi^{-1}\mathbf{B}_C(:,k))^T. \tag{13}$$

Clearly, this argument can be applied to determine "only" $n = m + r - 1$ rows of the eigenvector matrix ψ, assuming that there is only one co-located sensor-actuator pair ($n = m + r - n_c$ if there are n_c co-located sensor-actuator pairs). It is important to notice that, in correspondence of a degree of freedom with either a sensor or an actuator, the entire row (corresponding to that degree of freedom) of the eigenvector matrix ψ can be evaluated. This is equivalent to saying that, at each of these degrees of freedom, the corresponding components of all the $2N$ complex vibrational modes are determined. For the other degrees of freedom that are not instrumented with either a sensor or an actuator, however, the corresponding components of the complex eigenvector matrix ψ cannot be evaluated since it will not be possible to set up Eq. (12) or Eq. (13). For a detailed discussion of the aforementioned methodology, the reader is referred to the presentations in the works of Luş [11] and De Angelis et al. [12].

Once all the identifiable rows of the matrix ψ have been computed using either Eq. (12) or Eq. (13), it is convenient to rearrange them by moving all the known rows at the top of the matrix ψ while the unknown rows are moved down. This is equivalent to rearranging the vector of the degrees of freedom in "known" and "unknown" DOF's. The eigenvector matrix ψ can then be represented as

$$\psi = \begin{bmatrix} \hat{\psi}_{1,1} & \hat{\psi}_{1,2} & \cdots & \hat{\psi}_{1,2N} \\ \hat{\psi}_{2,1} & \hat{\psi}_{2,2} & \cdots & \hat{\psi}_{2,2N} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\psi}_{n,1} & \hat{\psi}_{n,2} & \cdots & \hat{\psi}_{n,2N} \\ \bar{\psi}_{n+1,1} & \bar{\psi}_{n+1,2} & \cdots & \bar{\psi}_{n+1,2N} \\ \vdots & \vdots & \vdots & \vdots \\ \bar{\psi}_{N,1} & \bar{\psi}_{N,2} & \cdots & \bar{\psi}_{N,2N} \end{bmatrix}, \tag{14}$$

where $\hat{\psi}_{i,j}$ denotes the "known" component of the jth mode at the ith degree of freedom, while $\bar{\psi}_{k,j}$ denotes the "unknown" component of the jth mode at the kth degree of freedom.

4.2 Expanding the Input-Output Mapping. Having determined the n rows of the matrix ψ (denoted by $\hat{\psi}$), it is now possible to construct a new state space model, which can also predict the system's response at *actuator* locations for excitations applied at *sensor* locations. This new system is an improvement on the initial first-order system, for which the actuator and sensor locations were fixed and limited to the initial test configuration.

If a certain degree of freedom, e.g., the jth DOF, has an actuator placed on it, then the contribution of the excitation to the state equation is through the term $\psi^T\mathcal{B}(:,j)$. Analogously, if a sensor is placed on the jth DOF, then the state vector is related to that output through the term $\mathcal{C}_p(j,:)\psi$. By assuming that sensors are placed at the DOF's where the actuators are placed and that actuators are placed at the DOF's with sensors, a hypothetical "expanded set of $(m + r - n_c)$ co-located sensor-actuator pairs" can be created. For each of such pairs, the co-location requirement is given by $\mathcal{C}_p(j,:) = \mathcal{B}(:,j)^T$, for $j = 1$ *to* $(m + r - n_c)$. This would allow us to create new input and output matrices, namely $\hat{\mathcal{B}}$ and $\hat{\mathcal{C}}_p$, that can be used in a new state equation, which can be written as

$$\dot{\boldsymbol{\zeta}}(t) = \mathbf{\Lambda}\boldsymbol{\zeta}(t) + \hat{\psi}^T\hat{\mathcal{B}}\hat{\mathbf{u}}(t), \tag{15a}$$

$$\hat{\mathbf{y}}(t) = \hat{\mathcal{C}}_p\hat{\psi}\boldsymbol{\zeta}(t), \tag{15b}$$

where the quantities with (ˆ) are related to the n degrees of freedom with either a sensor or an actuator. It is important to note that the new input vector $\hat{\mathbf{u}}(t)$ and the new output vector $\hat{\mathbf{y}}(t)$ now contain information about all the n active degrees of freedom, providing a more general input/output mapping. This is to say that (1) it is now possible to predict the output at any of the DOF's that did not have a sensor but had an actuator placed on them, and (2) if it so happens that the system is subjected to a new input applied at any of the aforementioned n active DOF's, it is also possible to accurately predict the response at any of those locations. This operation can be seen as the time domain equivalent of building the transfer function matrix of the system using the symmetry property of such a matrix. If, for example, one had identified the components of the transfer function matrix that relates the ith input to the jth output, then it is also possible to obtain immediately the component relating a hypothetical output at the ith DOF to a hypothetical input at the jth DOF.

5 Retrieving the Mass, Damping, and Stiffness Matrices of the "Reduced Order Model"

5.1 The General Case. Having determined the n rows of the matrix ψ (denoted by $\hat{\psi}$), it is now possible to determine a compact form of the mass, damping, and stiffness matrices related to the reduced model. To this end, let us consider the general expressions of the mass, damping, and stiffness matrices of the larger system which are obtained by imposing the orthogonality conditions given in Eqs. (5). As shown in Refs. [10] and [12], these matrices can be expressed as

$$\mathcal{M}=(\psi\Lambda\psi^T)^{-1}, \tag{16a}$$

$$\mathcal{L}=-\mathcal{M}\psi\Lambda^2\psi^T\mathcal{M}, \tag{16b}$$

$$\mathcal{K}=-(\psi\Lambda^{-1}\psi^T)^{-1}. \tag{16c}$$

Since the mass and stiffness matrices have similar structures, let us first analyze the reduced form of the mass and stiffness matrices. Using the subdivision of the matrix ψ presented in Eq. (14), it is possible to express the matrices $\mathcal{M}$ and $\mathcal{K}$ in partitioned forms as

$$\mathcal{M}=\begin{bmatrix}\hat{\psi}\Lambda\hat{\psi}^T & \hat{\psi}\Lambda\overline{\psi}^T\\ \overline{\psi}\Lambda\hat{\psi}^T & \overline{\psi}\Lambda\overline{\psi}^T\end{bmatrix}^{-1};\quad \mathcal{K}=-\begin{bmatrix}\hat{\psi}\Lambda^{-1}\hat{\psi}^T & \hat{\psi}\Lambda^{-1}\overline{\psi}^T\\ \overline{\psi}\Lambda^{-1}\hat{\psi}^T & \overline{\psi}\Lambda^{-1}\overline{\psi}^T\end{bmatrix}^{-1}, \tag{17}$$

where $\hat{\psi}$, of dimension $n\times 2N$, and $\overline{\psi}$, of dimension $(N-n)\times 2N$, are the submatrices of ψ corresponding to known and unknown degrees of freedom, respectively. By employing the inverse of only the known portion of $\mathcal{M}^{-1}$, it is possible to obtain a "reduced" order mass matrix of the structural system, $\hat{\mathcal{M}}$, of dimension $n\times n$, as

$$\hat{\mathcal{M}}=[\hat{\psi}\Lambda\hat{\psi}^T]^{-1}. \tag{18}$$

Similarly, a reduced order $n\times n$ stiffness matrix of the system, $\hat{\mathcal{K}}$, can be obtained as

$$\hat{\mathcal{K}}=-[\hat{\psi}\Lambda^{-1}\hat{\psi}^T]^{-1}. \tag{19}$$

These two matrices are symmetric and are related to the general $N\times N$ mass and stiffness matrices of the structural system through a static condensation relationship. In fact, from Eqs. (18) and (19), both matrices are presented as the inverse of a symmetric matrix which itself is a partition of a larger matrix. If we now rewrite the partitioned form of $\mathcal{K}$ as

$$\mathcal{K}=\begin{bmatrix}\mathcal{K}_{nn} & \mathcal{K}_{nu}\\ \mathcal{K}_{un} & \mathcal{K}_{uu}\end{bmatrix}=-\begin{bmatrix}\hat{\psi}\Lambda^{-1}\hat{\psi}^T & \hat{\psi}\Lambda^{-1}\overline{\psi}^T\\ \overline{\psi}\Lambda^{-1}\hat{\psi}^T & \overline{\psi}\Lambda^{-1}\overline{\psi}^T\end{bmatrix}^{-1} \tag{20}$$

then what we have denoted as $\hat{\mathcal{K}}=-[\hat{\psi}\Lambda^{-1}\hat{\psi}^T]^{-1}$ is in fact

$$\hat{\mathcal{K}}=[\mathcal{K}_{nn}-\mathcal{K}_{nu}\mathcal{K}_{uu}^{-1}\mathcal{K}_{un}]. \tag{21}$$

It is important to note that the expression in Eq. (21) is identical to an expression we would have got if we had considered the static condensation of the matrix $\mathcal{K}$ by employing the DOF's with either an actuator or a sensor as the "independent" DOF's, and the DOF's with neither an actuator nor a sensor as "dependent" (condensed) DOF's. This expression of $\hat{\mathcal{K}}$ is also identical to the reduced stiffness matrix one would obtain using the Guyan reduction ([19]). Because of the similarity in the structure of $\mathcal{M}$ and $\mathcal{K}$, the first argument holds true also for the mass matrix, so that the reduced mass matrix corresponds to

$$\hat{\mathcal{M}}=[\mathcal{M}_{nn}-\mathcal{M}_{nu}\mathcal{M}_{uu}^{-1}\mathcal{M}_{un}] \tag{22}$$

with the partitions defined analogously. However, since this reduced mass matrix is obtained only with the partitions of the larger mass matrix and does not involve any contribution from the stiffness matrix, it is in general quite different from the reduced mass matrix one would get using the Guyan reduction. In fact, these reduced order matrices $\hat{\mathcal{K}}$ and $\hat{\mathcal{M}}$ are also known as the Schur complements of the partitions $\mathcal{K}_{uu}$ and $\mathcal{M}_{uu}$, respectively.

With regard to the damping matrix, its reduced form can be obtained from Eqs. (16), by accounting only for the known partitions of ψ and $\mathcal{M}$. This reduced order damping matrix $\hat{\mathcal{L}}$, of dimensions $n\times n$, is symmetric and can be expressed as

$$\hat{\mathcal{L}}=-\hat{\mathcal{M}}\hat{\psi}\Lambda^2\hat{\psi}^T\hat{\mathcal{M}}, \tag{23}$$

where the information about all the N DOF's of the "larger" system have been included through $\hat{\mathcal{M}}$, $\hat{\psi}$, and Λ.

5.2 Block Diagonal Mass Matrix Case. If the system under consideration has a block diagonal mass matrix, it is possible to provide new interpretations to the results of the previous section. Let us first note that a block diagonal mass matrix may be partitioned as

$$\mathcal{M}=\begin{bmatrix}\mathcal{M}_{nn} & \mathbf{0}\\ \mathbf{0} & \mathcal{M}_{uu}\end{bmatrix}$$

and that for this case Eq. (22) leads to

$$\hat{\mathcal{M}}=[\hat{\psi}\Lambda\hat{\psi}^T]^{-1}=\mathcal{M}_{nn}. \tag{24}$$

Therefore, for a system with a block diagonal mass matrix, the subpartition of the mass matrix related to the instrumented DOF's can be directly and exactly evaluated using the available complex modal information. Similarly, since in this case the expression for the damping matrix can be written as

$$\mathcal{L}=\begin{bmatrix}\mathcal{L}_{nn} & \mathcal{L}_{nu}\\ \mathcal{L}_{un} & \mathcal{L}_{uu}\end{bmatrix}=-\begin{bmatrix}\mathcal{M}_{nn} & \mathbf{0}\\ \mathbf{0} & \mathcal{M}_{uu}\end{bmatrix}\begin{bmatrix}\hat{\psi}\Lambda^2\hat{\psi}^T & \hat{\psi}\Lambda^2\overline{\psi}^T\\ \overline{\psi}\Lambda^{-2}\hat{\psi}^T & \overline{\psi}\Lambda^2\overline{\psi}^T\end{bmatrix}\begin{bmatrix}\mathcal{M}_{nn} & \mathbf{0}\\ \mathbf{0} & \mathcal{M}_{uu}\end{bmatrix}, \tag{25}$$

the partition $\mathcal{L}_{nn}$ can also be directly and exactly evaluated as

$$\mathcal{L}_{nn}=-\mathcal{M}_{nn}\hat{\psi}\Lambda^2\hat{\psi}^T\mathcal{M}_{nn}. \tag{26}$$

The partitions of the stiffness matrix, however, are not so straightforward to evaluate, and only approximate results may be obtained. In order to clarify this comment, let us go back to Eq. (5b) and rewrite it as

$$\psi^T\mathcal{K}\psi-\Lambda\psi^T\mathcal{M}\psi\Lambda=-\Lambda. \tag{27}$$

Premultiplying Eq. (27) with $\psi\Lambda$ and postmultiplying it with $\Lambda\psi^T$ leads, with the help of the relations in Eqs. (16), to

$$\mathcal{K}=-\mathcal{M}\psi\Lambda^3\psi^T\mathcal{M}+\mathcal{L}\mathcal{M}^{-1}\mathcal{L} \tag{28}$$

and, using the partitions of the mass and damping matrices, Eq. (28) may be rewritten as

$$\mathcal{K}_{nn}=\mathcal{L}_{nn}\mathcal{M}_{nn}^{-1}\mathcal{L}_{nn}+\mathcal{L}_{nu}\mathcal{M}_{uu}^{-1}\mathcal{L}_{un}-\mathcal{M}_{nn}\hat{\psi}\Lambda^3\hat{\psi}^T\mathcal{M}_{nn}. \tag{29}$$

Clearly this expression can not be evaluated exactly since $\mathcal{L}_{nu}$, $\mathcal{M}_{uu}$, and $\mathcal{L}_{un}$ are unknown quantities. The "closeness" of this estimation naturally depends on the contribution of the unknown term, and unfortunately it is not trivial to quantify due to its sole dependence on unattainable parameters of the system.

5.3 Diagonal Mass Matrix Case With Mass Normalized Normal Modes. The discussion until now has focused on the complex eigenvectors and eigenvalues which are a direct product of the state space formulation. The normal modal parameters, on the other hand, are also widely used in modal analysis, and in this section we present a new methodology that utilizes such information to construct partitions of the mass, damping, and stiffness matrices.

The eigenvalue problem for the so called normal modal parameters is given by

$$\mathcal{M}\phi\boldsymbol{\Omega}^2=\mathcal{K}\phi, \tag{30}$$

where $\boldsymbol{\Omega}^2=\mathrm{diag}(\Omega_1^2,\Omega_2^2,\ldots,\Omega_N^2)$ is a diagonal matrix containing the squares of the undamped natural frequencies, and $\phi_{N\times N}$ is the eigenvector matrix whose columns are the normal (undamped) eigenvectors of the system. Here we assume that the eigenvectors are mass normalized, i.e., that they are scaled such that

$$\phi^T\mathcal{M}\phi=\mathbf{I},\quad \phi^T\mathcal{K}\phi=\boldsymbol{\Omega}^2. \tag{31}$$

The determination of these normal modal parameters from experiments is possible for classically damped systems; the reader is referred to the work of Alvin et al. [7] for a brief discussion of candidate procedures. Here we employ the same assumptions as Alvin et al. [7] regarding the availability of these modal parameters, and assume that (i) the system is classically damped, so that the undamped eigenvalues and modal damping percentages can be easily evaluated from the continuous time poles of the identified state space model (as discussed, for example, by Luş [11]), and that (ii) the mass normalized undamped mode shapes are known at the sensor locations.

Using these normal modal parameters, the mass, damping, and stiffness matrices can be constructed using

$$\mathcal{M}=\phi^{-T}\phi^{-1},\quad \mathcal{L}=\phi^{-T}\mathcal{E}\phi^{-1},\quad \mathcal{K}=\phi^{-T}\boldsymbol{\Omega}^2\phi^{-1}, \tag{32}$$

where $\mathcal{E}$ is the damping matrix in modal coordinates (diagonal for a classically damped system). If, on the other hand, the system is not fully instrumented, then it will not be possible to evaluate the whole eigenvector matrix ϕ, and hence Eqs. (32) will not be applicable. It will still be possible, however, to employ alternate expressions to estimate partitions of the mass, damping, and stiffness matrices, provided that the system has a diagonal mass matrix. To investigate this claim, let us start by partitioning the eigenvector matrix such that

$$\phi=\begin{bmatrix}\hat{\phi}\\ \bar{\phi}\end{bmatrix}, \tag{33}$$

where $\hat{\phi}_{n\times N}$ is the partition of ϕ that can be determined, and $\bar{\phi}_{(N-n)\times N}$ denotes the partition that can not be evaluated due to insufficient instrumentation. Furthermore, the singular value decomposition of ϕ is given by

$$\phi=\begin{bmatrix}\hat{\phi}\\ \bar{\phi}\end{bmatrix}=\mathbf{USV}^T=\begin{bmatrix}\hat{\mathbf{U}}\mathbf{SV}^T\\ \bar{\mathbf{U}}\mathbf{SV}^T\end{bmatrix}, \tag{34}$$

where $\mathbf{U}_{N\times N}$ and $\mathbf{V}_{N\times N}$ are real unitary matrices containing the left and right singular vectors, and $\mathbf{S}_{N\times N}$ is a diagonal matrix containing the singular values. Using this decomposition, the inverse of ϕ can be expressed as

$$\phi^{-1}=\mathbf{VS}^{-1}\mathbf{U}^T=[\mathbf{VS}^{-1}\hat{\mathbf{U}}^T\ \ \mathbf{VS}^{-1}\bar{\mathbf{U}}^T] \tag{35}$$

and, using Eqs. (32), the expressions for the mass, damping, and stiffness matrices can be written as

$$\mathcal{M}=\begin{bmatrix}\mathcal{M}_{nn} & \mathcal{M}_{nu}\\ \mathcal{M}_{un} & \mathcal{M}_{uu}\end{bmatrix}=\begin{bmatrix}\hat{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathbf{VS}\hat{\mathbf{U}}^T & \hat{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathbf{VS}\bar{\mathbf{U}}^T\\ \bar{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathbf{VS}\hat{\mathbf{U}}^T & \bar{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathbf{VS}\bar{\mathbf{U}}^T\end{bmatrix}, \tag{36a}$$

$$\mathcal{L}=\begin{bmatrix}\mathcal{L}_{nn} & \mathcal{L}_{nu}\\ \mathcal{L}_{un} & \mathcal{L}_{uu}\end{bmatrix}=\begin{bmatrix}\hat{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathcal{E}\mathbf{VS}\hat{\mathbf{U}}^T & \hat{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathcal{E}\mathbf{VS}\bar{\mathbf{U}}^T\\ \bar{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathcal{E}\mathbf{VS}\hat{\mathbf{U}}^T & \bar{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\mathcal{E}\mathbf{VS}\bar{\mathbf{U}}^T\end{bmatrix}, \tag{36b}$$

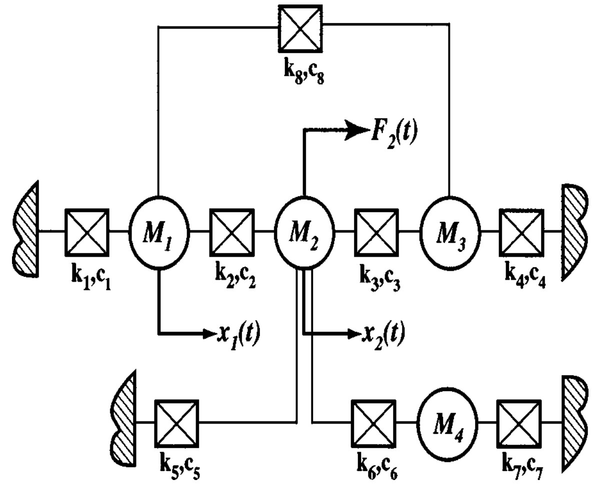


Fig. 1 4 DOF lumped mass system

$$\mathcal{K}=\begin{bmatrix}\mathcal{K}_{nn} & \mathcal{K}_{nu}\\ \mathcal{K}_{un} & \mathcal{K}_{uu}\end{bmatrix}=\begin{bmatrix}\hat{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\boldsymbol{\Omega}^2\mathbf{VS}\hat{\mathbf{U}}^T & \hat{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\boldsymbol{\Omega}^2\mathbf{VS}\bar{\mathbf{U}}^T\\ \bar{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\boldsymbol{\Omega}^2\mathbf{VS}\hat{\mathbf{U}}^T & \bar{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^T\boldsymbol{\Omega}^2\mathbf{VS}\bar{\mathbf{U}}^T\end{bmatrix}. \tag{36c}$$

It should be emphasized that, since the whole eigenvector matrix ϕ is not available, it is in general not possible to determine the matrices $\hat{\mathbf{U}}$, $\bar{\mathbf{U}}$, $\mathbf{S}$, and $\mathbf{V}$. It can be shown, however, that if the system has a diagonal mass matrix, then the partition $\mathbf{VS}^{-1}\hat{\mathbf{U}}^T$ is identically equal to the Moore-Penrose pseudoinverse of $\hat{\phi}$, i.e.,

$$\mathbf{VS}^{-1}\hat{\mathbf{U}}^T=\hat{\phi}^{\dagger}, \tag{37}$$

where $\hat{\phi}^{\dagger}$ denotes the Moore-Penrose pseudoinverse (see Appendix B for the proof of the statement). Therefore the partitions of the mass, damping, and stiffness matrices corresponding to the known DOF's can be exactly evaluated using

$$\mathcal{M}_{nn}=(\hat{\phi}^{\dagger})^T(\hat{\phi}^{\dagger}),\quad \mathcal{L}_{nn}=(\hat{\phi}^{\dagger})^T\mathcal{E}(\hat{\phi}^{\dagger}),$$
$$\mathcal{K}_{nn}=(\hat{\phi}^{\dagger})^T\boldsymbol{\Omega}^2(\hat{\phi}^{\dagger}). \tag{38}$$

It should be noted that these expressions complement the results presented in the work of Alvin et al. [7], wherein the authors had provided expressions that yielded reduced order matrices equivalent to ones that would be obtained via Guyan reduction.

As a final note, it should be mentioned that reduced order physical matrices, obtained via the procedures described herein, do not completely reflect the full dynamics of the system, i.e., the response of the measured/excited DOF's as predicted by the reduced order system will not be equal to the true response of the respective DOF's of the full order model. This problem is common to all reduction schemes, whether they start from an identified model or an analytical model, and is mainly due to the fact that the reduced order matrices lead to a different eigenvalue problem than that of the full order matrices, and that modal properties of the full order system are not preserved in this new eigenvalue problem. Therefore it is best to use the information obtained via the proposed schemes for obtaining reduced order models in investigating the partitions of the full order matrices, e.g., in health monitoring or model updating. In order to accurately predict the structural response, one should use the state space model represented by the system in Eqs. (15).

6 Numerical Examples

6.1 4-DOF System. In order to present the various stages of the proposed approach, let us consider a brief numerical example. The system we study is the 4 DOF lumped mass model shown in Fig. 1. The dynamic data to be used in the identification is ob-

Table 1 Actual and identified continuous time poles of the model. Re( ) and Im( ) refer, respectively, to the real and imaginary components of the poles.

Actual		Identified	
$\mathrm{Re}(\lambda_i)$	$\mathrm{Im}(\lambda_i)$	$\mathrm{Re}(\lambda_i)$	$\mathrm{Im}(\lambda_i)$
−0.077	−1.244	−0.077	−1.244
−0.077	+1.244	−0.077	+1.244
−0.170	−1.838	−0.170	−1.838
−0.170	+1.838	−0.170	+1.838
−0.263	−2.278	−0.263	−2.278
−0.263	+2.278	−0.263	+2.278
−0.281	−2.353	−0.281	−2.353
−0.281	+2.353	−0.281	+2.353

tained by subjecting this system to a white noise input at the second DOF, and displacements are assumed to be measured at the first and the second DOF's. It should be strongly emphasized, however, that the methodology is also applicable with velocity and/or acceleration measurements. The reader is referred to the works of Luş [11] and De Angelis et al. [12] for a thorough discussion of this issue.

The system properties are chosen as follows: $M_1=0.8$, $M_2=2.0$, $M_3=1.2$, $M_4=0.6$, $k_i=1.0$ for $i=1$, 4, 5, 7, $k_i=2.0$ for $i=2$, 3, 6, 8, and $c_i=0.1\times k_i$ for all i. Note that this choice of the viscous damping coefficients leads to a classically damped system, thereby allowing us to discuss all the formulations developed in this study. With these parameters, the mass, damping, and stiffness matrices can be constructed as

$$\mathcal{M}=\begin{bmatrix} 0.8 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 0.6 \end{bmatrix};$$

$$\mathcal{L}=\begin{bmatrix} 0.4 & -0.1 & -0.1 & 0 \\ -0.1 & 0.5 & -0.1 & -0.1 \\ -0.1 & -0.1 & 0.4 & 0 \\ 0 & -0.1 & 0 & 0.3 \end{bmatrix};$$

$$\mathcal{K}=\begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 5 & -1 & -1 \\ -1 & -1 & 4 & 0 \\ 0 & -1 & 0 & 3 \end{bmatrix}.$$

Note that the partitions of the stiffness matrix defined in Eq. (20) are given for this system and instrumentation set up by

$$\mathcal{K}_{nn}=\begin{bmatrix} 4 & -1 \\ -1 & 5 \end{bmatrix}; \quad \mathcal{K}_{nu}=\begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix},$$

$$\mathcal{K}_{un}=\begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}; \quad \mathcal{K}_{uu}=\begin{bmatrix} -4 & 0 \\ 0 & 3 \end{bmatrix}.$$

The partitions of the mass and the damping matrices are defined analogously.

The initial stage in the proposed methodology is the identification of a first-order model of the system. Using the OKID/ERA approach ([16,18]), a single input–two output system is easily identified. The success of this identification may be easily judged by comparing the actual and identified values of the continuous time poles, which are presented in Table 1.

Clearly, the OKID/ERA algorithm has been extremely successful in realizing an initial state space model even with two outputs and a single input, as evidenced by the exact agreement between the actual and identified values of the poles. Also note that with the rich input employed in the test, it was in fact possible to identify all the poles of the system. Starting with this initial model, the next step is to find the matrix $\mathcal{T}$ that allows us to transform the equations in Eqs. (8) to the modal coordinates of the symmetric eigenvalue problem given by Eqs. (6). Since the system has a co-located sensor-actuator pair located at the second DOF, it is indeed possible to identify the transformation matrix and execute the transformation. Furthermore, using the equations in the transformed modal coordinates, the partitions of the complex eigenvector matrix corresponding to the instrumented DOF's (the first and the second DOF's for this example) may be evaluated. In fact, in this case we get $\hat{\psi}=[\hat{\psi}_1\hat{\psi}_1^*\hat{\psi}_2\hat{\psi}_2^*\hat{\psi}_3\hat{\psi}_3^*\hat{\psi}_4\hat{\psi}_4^*]$, where the superscript (*) denotes complex conjugate, with

$$\hat{\psi}_1=\begin{bmatrix} 0.1613+\jmath 0.1613 \\ 0.2515+\jmath 0.2515 \end{bmatrix}; \quad \hat{\psi}_2=\begin{bmatrix} 0.1054+\jmath 0.1054 \\ -0.1288-\jmath 0.1288 \end{bmatrix},$$

$$\hat{\psi}_3=\begin{bmatrix} 0.1394+\jmath 0.1394 \\ 0.0560+\jmath 0.0560 \end{bmatrix}; \quad \hat{\psi}_4=\begin{bmatrix} -0.3026-\jmath 0.3026 \\ 0.0608+\jmath 0.0608 \end{bmatrix}.$$

Note that the real and imaginary parts of the components of these eigenvectors are equal in magnitude; this phenomenon is due to the particular scaling choice expressed in Eqs. (5) and it is indicative of the fact that the system is classically damped. At this point, we have all the information we need to evaluate the reduced order matrices. Starting with Eqs. (18) and (19), we have

$$\hat{\mathcal{M}}=[\hat{\psi}\Lambda\hat{\psi}^T]^{-1}=\begin{bmatrix} 0.8 & 0.0 \\ 0.0 & 2.0 \end{bmatrix}, \tag{39a}$$

$$\hat{\mathcal{K}}=-[\hat{\psi}\Lambda^{-1}\hat{\psi}^T]^{-1}=\begin{bmatrix} 3.75 & -1.25 \\ -1.25 & 4.12 \end{bmatrix}. \tag{39b}$$

Note that, since the mass matrix was diagonal, the identified reduced order mass matrix $\hat{\mathcal{M}}$ is nothing but the 2×2 partition corresponding to the first and the second DOF's (i.e., $\mathcal{M}_{nn}$), and this result was to be expected due to Eq. (24). The identified reduced order stiffness matrix $\hat{\mathcal{K}}$, on the other hand, is exactly equal to the matrix one would get by statically condensing the third and the fourth DOF's (i.e., Guyan reduction); the result of such a reduction would be given by the formula $\mathcal{K}_{nn}-\mathcal{K}_{nu}\mathcal{K}_{uu}^{-1}\mathcal{K}_{un}$ with the partitions presented above. Analogously, a reduced order damping matrix can be obtained via Eq. (23) as

$$\hat{\mathcal{L}}=-\hat{\mathcal{M}}\hat{\psi}\Lambda^2\hat{\psi}^T\hat{\mathcal{M}}=\begin{bmatrix} 0.4 & -0.1 \\ -0.1 & 0.5 \end{bmatrix}.$$

Once again, owing to the diagonal nature of the mass matrix, this reduced order damping matrix is nothing but the 2×2 partition corresponding to the first and the second DOF's, i.e., $\hat{\mathcal{L}}=\mathcal{L}_{nn}$, as was expressed in Eq. (26).

An alternate expression one could obtain for a reduced order stiffness matrix would be with the use of Eq. (29). In fact, using this equation, one could get an estimate of the partition $\mathcal{K}_{nn}$ by ignoring the contribution of the unknown term $\mathcal{L}_{nu}\mathcal{M}_{uu}^{-1}\mathcal{L}_{un}$, which in this case leads to

$$\mathcal{K}_{nn}^{est}=\begin{bmatrix} 3.99 & -1.01 \\ -1.01 & 4.98 \end{bmatrix}.$$

Even though this estimate is in error, it is interesting to note that the error is in fact quite small for the system under consideration. In general, the ''closeness'' of this estimate will naturally depend on the nature of the neglected term as discussed in the previous sections.

As a final note, let us consider the estimates we could obtain using normal modal parameters. With the instrumentation at hand, it would be possible (see Alvin et al. [7] and Bernal and Gunes [20]) to get the partition $\hat{\phi}$ of the mass normalized eigenvectors as

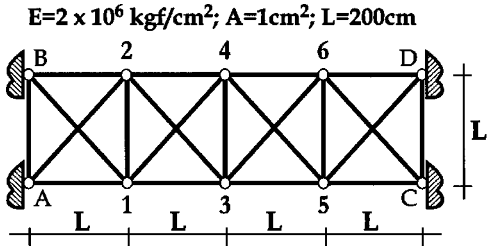


Fig. 2 Example truss structure

$$\hat{\phi}=\begin{bmatrix}-0.3598 & 0.2858 & -0.4207 & 0.9284\\ -0.5610 & -0.3492 & -0.1691 & -0.1865\end{bmatrix}.$$

Since the second-order eigenvalues and modal damping ratios can be evaluated directly from the continuous time poles of the identified first-order system (for an explanation of the procedure, see, for example, the study by Luş et al. [18]), the second-order matrices $\mathcal{E}$ and $\mathbf{\Omega}^2$ may be easily constructed as

$$\mathbf{\Omega}^2=\begin{bmatrix}1.553 & 0 & 0 & 0\\ 0 & 3.407 & 0 & 0\\ 0 & 0 & 5.257 & 0\\ 0 & 0 & 0 & 5.616\end{bmatrix};$$

$$\mathcal{E}=\begin{bmatrix}0.1553 & 0 & 0 & 0\\ 0 & 0.3407 & 0 & 0\\ 0 & 0 & 0.5257 & 0\\ 0 & 0 & 0 & 0.5616\end{bmatrix}.$$

Hence, using Eqs. (38), the partitions $\mathcal{M}_{nn}$, $\mathcal{L}_{nn}$, and $\mathcal{K}_{nn}$ can be evaluated as

$$\mathcal{M}_{nn}=\begin{bmatrix}0.8 & 0.0\\ 0.0 & 2.0\end{bmatrix};\quad \mathcal{L}_{nn}=\begin{bmatrix}0.4 & -0.1\\ -0.1 & 0.5\end{bmatrix};$$

$$\mathcal{K}_{nn}=\begin{bmatrix}4.0 & -1.0\\ -1.0 & 5.0\end{bmatrix}$$

and clearly these identified matrices are identically equal to their respective actual values.

6.2 Applications for Health Monitoring. To briefly investigate the applicability of the proposed technique to structural health monitoring, let us now consider a second structural system, which is represented by the truss shown in Fig. 2. The material and geometric properties of the various elements are reported in the figure while the values of the lumped masses are: $m_1=m_6=100\,\mathrm{K\,gm}$ and $m_2=m_3=m_4=m_5=200\,\mathrm{K\,gm}$. For this example, it is assumed that the system is classically damped. Such a truss is subjected to three applied random forces, two of which are acting in the horizontal direction (at nodes 1 and 5) and one is applied along the vertical direction (at node 4).

For the case of a limited set of sensors, let us assume that the displacements along the horizontal direction have been measured at nodes 2, 3, and 4 while the vertical displacements have been measured only at nodes 1, 2, 3, and 4. It is clear that, in this case, the co-located sensor-actuator pair is the one corresponding to the horizontal direction at node 4. To monitor the structural health of such a truss system, we consider two configurations: (1) the "undamaged" configuration, in which the structure is assumed to be as previously described, and (2) the "damaged" configuration, in which a structural element (i.e., 6-C) has been removed to simulate the effect of damage.

Using this input-output sets, the identified first-order undamaged and damaged models still show the contribution of 12 vibrational modes, as shown in Table 2. The damping and frequency values differ for the undamaged and the damaged structure due to the structural change. Analogous changes appear in the identified reduced (9×9) stiffness matrix for the two cases, as presented in Table 3. By comparing these two matrices, it is possible to identify the elements that have suffered structural damage: here it is clear that all the stiffness coefficients related to node 6 (presented in bold face) are affected by the change. This provides a useful indication about the location of the damaged area.

Table 2 Comparison of identified damping factors and natural frequencies for damaged and undamaged model for the truss system

Mode	Undamaged		Damaged	
	ξ_i(%)	f_i (Hz.)	ξ_i(%)	f_i (Hz.)
1	2.3243	0.4709	3.2772	0.3972
2	3.4069	0.9513	4.1605	0.8354
3	3.5979	1.0212	3.6028	1.0148
4	4.3500	1.2862	4.4715	1.2537
5	5.1973	1.5739	5.2079	1.5718
6	5.7328	1.7525	5.7328	1.7525
7	6.6063	2.0408	7.0593	1.9999
8	6.6756	2.0930	6.7820	2.0913
9	7.5120	2.3369	7.9273	2.1667
10	7.6604	2.3853	7.6316	2.3705
11	8.6475	2.7058	8.6960	2.6528
12	9.2078	2.8871	9.2128	2.8646

With regard to the mass and damping matrices, the identified reduced mass matrix is identical to that obtained from static condensation of the mass matrix of the entire system while the reduced damping matrix comes from the elimination of appropriate rows and columns (in this case, the 10th, 11th and 12th) from the general damping matrix.

7 Conclusions

In this study, the authors have derived various formulations regarding the construction of reduced order mass-damping-stiffness models from identified state space realizations. In particular, it has been shown that

(1) For nonclassically damped systems, it is possible to retrieve statically condensed versions of the mass and stiffness matrices. Such a condensation treats the instrumented DOF's (instrumentation in the form of acceleration, velocity, and/or position sensors and force actuators) as the "independent" DOF's, while the remaining DOF's are treated as the "dependent" DOF's and condensed.

(2) If the system has a block diagonal mass matrix, it is possible to exactly construct the partitions of the full order mass and damping matrices corresponding to the instrumented DOF's. Furthermore, it is also possible to obtain an estimate of the partition of the full order stiffness matrix corresponding to the instrumented DOF's, although this estimate is not exact.

(3) Provided that the partition of the mass normalized eigenvectors corresponding to the instrumented DOF's is available, it is possible to exactly identify the partitions of the mass, the damping, and the stiffness matrices corresponding to those instrumented DOF's for classically damped systems with diagonal mass matrices.

The theoretical presentation has been supplemented with a numerical example that illustrates the applicability of the proposed methodology and various issues regarding the formulations developed in this study.

It is anticipated that the methodology presented herein may be applied to various problems in mechanics including finite element model updating and health monitoring of systems with insufficient information. These issues and other investigations such as the effects of noise perturbations on the identified parameters are the

Table 3 Comparison reduced stiffness matrices for undamaged and damaged truss system

	2.707	0.000	0.000	0.000	−1.000	0.000	−0.354	−0.354	0.000
	0.000	1.707	0.000	−1.000	0.000	0.000	−0.354	−0.354	0.000
	0.000	0.000	2.707	0.000	−0.354	0.354	−1.000	0.000	0.000
	0.000	−1.000	0.000	1.707	0.354	−0.354	0.000	0.000	0.000
$\hat{\mathcal{K}}^U = 10^4\times$	−1.000	0.000	−0.354	0.354	***2.549***	***−0.158***	***−0.065***	***−0.065***	−1.000
	0.000	0.000	0.354	−0.354	***−0.158***	***1.549***	***−0.065***	***−1.065***	0.000
	−0.354	−0.354	−1.000	0.000	***−0.065***	***−0.065***	***2.226***	***0.111***	−0.354
	−0.354	−0.354	0.000	0.000	***−0.065***	***−1.065***	***0.111***	***1.596***	0.354
	0.000	0.000	0.000	0.000	−1.000	0.000	−0.354	0.354	2.707
	2.707	0.000	0.000	−0.000	−1.000	0.000	−0.354	−0.354	0.000
	0.000	1.707	−0.000	−1.000	−0.000	−0.000	−0.354	−0.354	0.000
	0.000	−0.000	2.707	0.000	−0.354	0.354	−1.000	−0.000	0.000
	−0.000	−1.000	0.000	1.707	0.354	−0.354	0.000	−0.000	−0.000
$\hat{\mathcal{K}}^D = 10^4\times$	−1.000	−0.000	−0.354	0.354	***2.528***	***−0.179***	***−0.000***	***−0.087***	−1.000
	0.000	−0.000	0.354	−0.354	***−0.179***	***1.528***	***−0.000***	***−1.087***	0.000
	−0.354	−0.354	−1.000	0.000	***−0.000***	***−0.000***	***2.030***	***0.177***	−0.354
	−0.354	−0.354	−0.000	−0.000	***−0.087***	***−1.087***	***0.177***	***1.574***	0.354
	0.000	0.000	0.000	−0.000	−1.000	0.000	−0.354	0.354	2.707

subjects of current research and beyond the scope and intent of this study, and hence they will be addressed and reported in future work.

Acknowledgments

This research has been sponsored through a research grant by the National Science Foundation (CMS-9457305), whose support has been greatly appreciated.

Appendix A: Proof Regarding the Diagonalization in the Asymmetric Eigenvalue Problem

Consider the real symmetric generalized eigenvalue problem given by

$$\mathbf{A}\boldsymbol{\Phi}\boldsymbol{\Theta}=\mathbf{B}\boldsymbol{\Phi}, \tag{40}$$

where $\mathbf{A}$ and $\mathbf{B}$ are two full rank symmetric matrices, $\boldsymbol{\Phi}$ is the eigenvector matrix, and $\boldsymbol{\Theta}$ is the eigenvalue matrix. It is known that, even in the case of repeated roots, there exists a unique set of eigenvectors such that

$$\boldsymbol{\Phi}^T\mathbf{A}\boldsymbol{\Phi}=\mathbf{I}, \tag{41a}$$

$$\boldsymbol{\Phi}^T\mathbf{B}\boldsymbol{\Phi}=\boldsymbol{\Theta}, \tag{41b}$$

and that $\boldsymbol{\Theta}$ is a diagonal matrix. Now consider a reformulation of this problem as

$$\boldsymbol{\Psi}\hat{\boldsymbol{\Theta}}=\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\Psi}, \tag{42}$$

where $\hat{\boldsymbol{\Theta}}$ is the corresponding eigenvalue matrix, and $\boldsymbol{\Psi}$ is the corresponding eigenvector matrix. Since the product of two symmetric matrices is not necessarily symmetric, Eq. (42) is in general an asymmetric eigenvalue problem, and diagonalizability in an asymmetric eigenvalue problem is not guaranteed. Since, however, in this case the asymmetric eigenvalue problem of Eq. (42) is in fact derived from the symmetric eigenvalue problem of Eqs. (40) and (41), there indeed exist sets of eigenvectors that will yield a diagonal $\hat{\boldsymbol{\Theta}}$. In fact, assume that $\boldsymbol{\Psi}=\boldsymbol{\Phi}\mathbf{T}$, such that

$$\begin{aligned}\boldsymbol{\Psi}^{-1}\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\Psi}&=(\boldsymbol{\Phi}\mathbf{T})^{-1}\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\Phi}\mathbf{T}\\ &=\mathbf{T}^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\Phi}\boldsymbol{\Phi}^T\mathbf{B}\boldsymbol{\Phi}\mathbf{T} \quad \text{via Eq. } (41a)\\ &=\mathbf{T}^{-1}\boldsymbol{\Theta}\mathbf{T} \quad \text{via Eq. } (41b).\end{aligned}$$

Therefore, for any diagonal transformation matrix $\mathbf{T}$, the eigenvector matrix $\boldsymbol{\Psi}=\boldsymbol{\Phi}\mathbf{T}$ yields $\boldsymbol{\Psi}^{-1}\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\Psi}=\hat{\boldsymbol{\Theta}}=\boldsymbol{\Theta}$, and hence the existence of $\boldsymbol{\Phi}$ implies the diagonalizability of the asymmetric eigenvalue problem in Eq. (42).

Appendix B: Proof That the Moore-Penrose Pseudoinverse of the Known Partition of the Mass Normalized Eigenvector Matrix is Equal to a Partition of the Inverse of the Mass Normalized Eigenvector Matrix

The mass normalized eigenvector matrix $\phi_{N\times N}$ can be partitioned such that

$$\phi=\left[\frac{\hat{\phi}}{\bar{\phi}}\right]=\mathbf{U}\mathbf{S}\mathbf{V}^T=\left[\frac{\hat{\mathbf{U}}\mathbf{S}\mathbf{V}^T}{\bar{\mathbf{U}}\mathbf{S}\mathbf{V}^T}\right], \tag{43}$$

where $\mathbf{U}_{N\times N}$ and $\mathbf{V}_{N\times N}$ are real unitary matrices containing the left and right singular vectors, and $\mathbf{S}_{N\times N}$ is a diagonal matrix containing the singular values. The Moore-Penrose pseudoinverse of the partition $\hat{\phi}$, which is denoted here as $\hat{\phi}^\dagger$, must satisfy the following four conditions:

$$(1)\quad \hat{\phi}\hat{\phi}^\dagger\hat{\phi}=\hat{\phi}, \tag{44a}$$

$$(2)\quad \hat{\phi}^\dagger\hat{\phi}\hat{\phi}^\dagger=\hat{\phi}^\dagger, \tag{44b}$$

$$(3)\quad \hat{\phi}\hat{\phi}^\dagger=(\hat{\phi}\hat{\phi}^\dagger)^T, \tag{44c}$$

$$(4)\quad \hat{\phi}^\dagger\hat{\phi}=(\hat{\phi}^\dagger\hat{\phi})^T. \tag{44d}$$

Let us assume that $\hat{\phi}^\dagger=\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^T$, and check if this solution satisfies the necessary conditions. Keeping in mind that $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and that $\hat{\phi}=\hat{\mathbf{U}}\mathbf{S}\mathbf{V}^T$, we get

$$(1)\quad \hat{\phi}\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^T\hat{\phi}=\hat{\mathbf{U}}\mathbf{S}\mathbf{V}^T=\hat{\phi}, \tag{45a}$$

$$(2)\quad \mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^T\hat{\phi}\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^T=\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^T, \tag{45b}$$

$$(3)\quad \hat{\phi}\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^T=\hat{\mathbf{U}}\hat{\mathbf{U}}^T=(\hat{\phi}\mathbf{V}\mathbf{S}^{-1})^T, \tag{45c}$$

$$(4)\quad \hat{\phi}^{\dagger}\hat{\phi}=\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^{T}\hat{\mathbf{U}}\mathbf{S}\mathbf{V}^{T};\quad (\hat{\phi}^{\dagger}\hat{\phi})^{T}=\mathbf{V}\mathbf{S}\hat{\mathbf{U}}^{T}\hat{\mathbf{U}}\mathbf{S}^{-1}\mathbf{V}^{T}. \tag{45d}$$

Note that the proposed solution $\hat{\phi}^{\dagger}=\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^{T}$ satisfies the first three conditions, but it does not satisfy the fourth condition *unless* $\hat{\mathbf{U}}^{T}\hat{\mathbf{U}}$ is a diagonal matrix, and so in general this solution does not hold. In the case of a diagonal mass matrix, however, the fourth condition will also be satisfied. To show the validity of this claim, let us write the expression for the inverse of mass matrix in terms of the normal eigenvectors to yield

$$\mathcal{M}^{-1}=\phi\phi^{T}=\mathbf{U}\mathbf{S}\mathbf{V}^{T}\mathbf{V}\mathbf{S}\mathbf{U}^{T}=\mathbf{U}\mathbf{S}^{2}\mathbf{U}^{T}. \tag{46}$$

Therefore the columns of $\mathbf{U}$ are the eigenvectors for the eigenvalue problem

$$\mathcal{M}^{-1}\mathbf{U}=\mathbf{U}\mathbf{S}^{2} \tag{47}$$

and $\mathbf{U}$ does not have a prescribed structure for a general mass matrix. For a diagonal mass matrix, on the other hand, $\mathcal{M}^{-1}$ is diagonal, and so is $\mathbf{U}$; in fact, in such a case,

$$\mathbf{U}(i,j)=\pm\delta_{ij} \tag{48}$$

where $\mathbf{U}(i,j)$ refers to the element on the ith row and jth column of $\mathbf{U}$, and δ_{ij} is the Kronecker delta. In this case therefore the product $\hat{\mathbf{U}}^{T}\hat{\mathbf{U}}$ becomes

$$\hat{\mathbf{U}}^{T}\hat{\mathbf{U}}=\begin{bmatrix}\mathbf{I}_{n\times n} & \mathbf{0}_{n\times(N-n)}\\ \mathbf{0}_{(N-n)\times n} & \mathbf{0}_{(N-n)\times(N-n)}\end{bmatrix} \tag{49}$$

and hence diagonal, such that condition (4) is also satisfied. In conclusion, for a system with a diagonal mass matrix, the proposed solution $\hat{\phi}^{\dagger}=\mathbf{V}\mathbf{S}^{-1}\hat{\mathbf{U}}^{T}$ indeed satisfies all the necessary conditions and hence is the true solution.

References

[1] Agbabian, M. S., Masri, S. F., Miller, R. K., and Caughey, T. K., 1991, ''System Identification Approach to Detection of Structural Changes,'' J. Eng. Mech., **117**, pp. 370–390.

[2] Mottershead, J. E., and Friswell, M. I., 1993, ''Model Updating in Structural Dynamics: A Survey,'' J. Sound Vib., **165**, pp. 347–375.

[3] Berman, A., 1979, ''Mass Matrix Correction Using an Incomplete Set of Measured Modes,'' AIAA J., **17**, pp. 1147–1148.

[4] Baruch, M., 1982, ''Optimal Correction of Mass and Stiffness Matrices Using Measured Modes,'' AIAA J., **20**, pp. 1623–1626.

[5] Beck, J. L., and Katafygiotis, L. S., 1998, ''Updating Models and Their Uncertainties. I: Bayesian Statistical Framework,'' J. Eng. Mech., **124**, pp. 455–461.

[6] Alvin, K. F., and Park, K. C., 1994, ''Second-Order Structural Identification Procedure Via State-Space-Based System Identification,'' AIAA J., **32**, pp. 397–406.

[7] Alvin, K. F., Peterson, L. D., and Park, K. C., 1995, ''Method for Determining Minimum-Order Mass and Stiffness Matrices From Modal Test Data,'' AIAA J., **33**, pp. 128–135.

[8] Tseng, D. H., Longman, R. W., and Juang, J. N., 1994, ''Identification of Gyroscopic and Nongyroscopic Second Order Mechanical Systems Including Repeated Root Problems,'' Adv. Astronaut. Sci., **87**, pp. 145–165.

[9] Tseng, D. H., Longman, R. W., and Juang, J. N., 1994, ''Identification of the Structure of the Damping Matrix in Second Order Mechanical Systems,'' Adv. Astronaut. Sci., **87**, pp. 166–190.

[10] Balmès, E., 1997, ''New Results on the Identification of Normal Modes From Experimental Complex Modes,'' Mech. Syst. Signal Process., **11**, pp. 229–243.

[11] Luş, H., 2001, ''Control Theory Based System Identification,'' Ph.D. thesis, Columbia University, New York.

[12] DeAngelis, M., Luş, H., Betti, R., and Longman, R. W., 2002, ''Extracting Physical Parameters of Mechanical Models From Identified State Space Representations,'' ASME J. Appl. Mech., **69**, pp. 617–625.

[13] Luş, H., Betti, R., Yu, J., and DeAngelis, M., ''Investigation of a System Identification Methodology in the Context of the ASCE Benchmark Problem,'' J. Eng. Mech., to appear.

[14] Sestieri, A., and Ibrahim, S. R., 1994, ''Analysis of Errors and Approximations in the Use of Modal Coordinates,'' J. Sound Vib., **177**, pp. 145–157.

[15] Juang, J. N., Cooper, J. E., and Wright, J. R., 1988, ''An Eigensystem Realization Algorithm Using Data Correlations (ERA/DC) for Modal Parameter Identification,'' Control Theory and Adv. Technol., **4**(1), pp. 5–14.

[16] Juang, J. N., Phan, M., Horta, L. G., and Longman, R. W., 1993, ''Identification of Observer/Kalman Filter Markov Parameters: Theory and Experiments,'' J. Guid. Control Dyn., **16**, pp. 320–329.

[17] Luş, H., Betti, R., and Longman, R. W., 1999, ''Identification of Linear Structural Systems Using Earthquake-Induced Vibration Data,'' Earthquake Eng. Struct. Dyn., **28**, pp. 1449–1467.

[18] Luş, H., Betti, R., and Longman, R. W., 2002, ''Obtaining Refined First Order Predictive Models of Linear Structural Systems,'' Earthquake Eng. Struct. Dyn., **31**, pp. 1413–1440.

[19] Guyan, Robert J., 1965, ''Reduction of Stiffness and Mass Matrices,'' AIAA J., **3**, p. 380.

[20] Bernal, D., and Gunes, B., 2000, ''Extraction of Second Order System Matrices From State Space Realizations,'' 14th ASCE Engineering Mechanics Conference.

Linear Multi-Degree-of-Freedom System Stochastic Response by Using the Harmonic Wavelet Transform

P. Tratskas[1]
e-mail: petros@rice.edu

P. D. Spanos
Fellow ASME,
L.B. Ryon Chair in Engineering
e-mail: spanos@rice.edu

Rice University,
6100 S. Main,
Houston, TX 77005, U.S.A.

The wavelet transform is used to capture localized features in either the time domain or the frequency domain of the response of a multi-degree-of-freedom linear system subject to a nonstationary stochastic excitation. The family of the harmonic wavelets is used due to the convenient spectral characteristics of its basis functions. A wavelet-based system representation is derived by converting the system frequency response matrix into a time-frequency wavelet "tensor." Excitation-response relationships are obtained for the wavelet-based representation which involve linear system theory, spectral representation of the excitation and of the response vectors, and the wavelet transfer tensor of the system. Numericalresults demonstrate the usefulness of the developed analytical procedure.
[DOI: 10.1115/1.1601252]

Introduction

The wavelet transform provides a combined time-scale representation of signals ([1–4]). Its localization properties are quite useful for representing random fields, and for analyzing nonstationary signals ([5–7]). Problems of vibration analysis have been approached by a wavelet-Galerkin scheme ([8]), and by wavelet-finite element schemes ([9]). Pertinent applications involve analysis of beams ([10]), base isolation systems ([11]), and plate structures ([9]). Studies related to treatment of linear systems have either used a numerical approach to derive the response, or have relied on restrictive assumptions regarding the nature of the excitation and the spectral properties of the wavelet function used ([12–14]).

This paper develops a general wavelet-based system representation, and excitation-response relationships in the wavelet domain. For this purpose, traditional concepts such as the system impulse response matrix and the frequency response matrix are treated in a wavelet basis context. Further, excitation-response relationships are derived by relying on linear system theory, and the scheme that generates the harmonic wavelet coefficients. Finally, the wavelet representation and the evolutionary spectrum of the stochastic response are obtained. The usefulness of the proposed method is demonstrated by applying it to a two-degree-of-freedom (2DOF) system subject to filtered white noise.

The Wavelet Transform

The wavelet transform assigns to an arbitrary function $f(t)$ another function of two variables α and b via the equation

$$WT_f(\alpha,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}f(t)w^*\left(\frac{t-b}{a}\right)dt. \tag{1}$$

This equation defines the continuous wavelet transform. The function $w(t)$ is the mother wavelet function; the symbol α represents the scale, and the symbol b denotes the time positioning of $w(t)$, while "*" denotes complex conjugation.

Modifying the continuous scheme, a discrete wavelet scheme can be introduced. Several discretization schemes have been employed in previous studies. The so-called dyadic wavelet transform is a widely used orthogonal scheme where the scale parameter is sampled on a dyadic grid. Specifically, $a=2^j$ and $b=k2^j$, where $j\in N$ and $k\in Z$. The symbols j and k denote the scale and time position respectively in the discrete wavelet domain. Therefore Eq. (1) is rewritten as ([3])

$$DWT_f(j,k)=2^{-j/2}\int_{-\infty}^{\infty}f(t)w^*(2^{-j}t-k)dt. \tag{2}$$

The harmonic wavelet transform is used in this study. The choice of the harmonic wavelets is made due to two appealing features that they exhibit. First, they are defined by a simple explicit formula, and second they have an exactly boxlike representation in the frequency domain ([5]). These properties lead to tractable solutions for spectral estimation and for system response determination. Specifically, the mother wavelet of the harmonic wavelet transform has a spectrum given by the equation

$$W(\omega)=\frac{1}{2\pi},\quad 2\pi\leqslant\omega\leqslant4\pi\quad\text{and}$$
$$W(\omega)=0,\quad\text{elsewhere.} \tag{3}$$

By considering the inverse Fourier transform of the above function, the harmonic wavelet representation in the time domain is derived. Specifically,

$$w(t)=\frac{e^{i4\pi t}-e^{i2\pi t}}{i2\pi t};\quad i=\sqrt{-1}. \tag{4}$$

Two kinds of discrete harmonic transform have been introduced. These are the dyadic, and the general transforms ([15]). The difference between the two harmonic schemes is in the way that they divide the time-frequency plane. An improved scheme has also been presented, which enhances the time resolution of the general scheme ([16]). In the dyadic harmonic wavelet scheme, a wavelet that belongs to scale j and time position k is given by the formula

[1]Research Engineer, Shell Energy Company, London, UK.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 26, 2002; final revision, Dec. 4, 2002. Associate Editor: A.A. Ferri. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$w_{j,k}(t)=w(2^j t-k)=\frac{e^{i4\pi(2^j t-k)}-e^{i2\pi(2^j t-k)}}{i2\pi(2^j t-k)}, \tag{5}$$

and its Fourier transform is

$$W_{j,k}(\omega)=\frac{1}{2\pi}2^{-j}e^{-i\omega k/2^j}, \quad 2^j 2\pi\leqslant\omega\leqslant 2^j 4\pi \quad \text{and}$$

$$W_{j,k}(\omega)=0, \quad \text{elsewhere} \tag{6}$$

For a discrete signal of length $N=2^{\mathrm{n}}$, the scale index j runs from 0 to $n-2$. At each scale j there are 2^{j} wavelets positioned at different time instants by the index k. The wavelet coefficients are determined by the equation ([5])

$$a(j,k)=2^j\int_{-\infty}^{\infty}f(t)w^*(2^j t-k)dt. \tag{7}$$

By using the relationship between the harmonic wavelet $w(t)$ and its Fourier transform $W(\omega)$ described by the formula

$$w^*(2^j t-k)=\int_{2\pi 2^j}^{4\pi 2^j}W_{j,k}(\omega)e^{i\omega t}d\omega, \tag{8}$$

Eq. (7) may be rewritten in the form

$$a(j,k)=\int_{2\pi 2^j}^{4\pi 2^j}F(\omega)e^{i\omega k/2^j}d\omega, \tag{9}$$

where $F(\omega)$ is the Fourier transform of $f(t)$. The computation of the wavelet coefficients may be then performed in the discrete format by expressing $F(\omega)$ in a Fourier series. Specifically,

$$F_{2^j+s}=2\pi F[2\pi(2^j+s)]. \tag{10}$$

Thus the harmonic wavelet coefficients may be determined by using the equation

$$a(j,k)=\sum_{s=0}^{2^{j-1}}F_{2^j+s}e^{i2\pi 2k/2^j}, \tag{11}$$

where the factor $\Delta\omega$ has been canceled by the factor $1/2\pi$ from Eq. (10). Therefore, to compute the wavelet coefficients, octave blocks of the sequence F are processed by the inverse Fourier transform.

Harmonic wavelets of the dyadic scheme that correspond to different scales, and also wavelets of the same scale but at different time positions are orthogonal to each other. That is,

$$\int_{-\infty}^{\infty}w(2^j t+k)w^*(2^r t+s)dt=0 \quad \forall \quad j,k,r,s, \tag{12}$$

except when $j=r$ and $s=k$.

The general harmonic wavelet transform divides the frequency axis in a different manner. Two indices are used to define the scale at this scheme. A wavelet of scale (m, n) and position k is defined in the frequency domain as

$$W_{m,n,k}(\omega)=\frac{1}{(n-m)2\pi}e^{-i\omega k/(n-m)}, \quad m2\pi\leqslant\omega\leqslant n2\pi \quad \text{and}$$

$$W_{m,n,k}(\omega)=0, \quad \text{elsewhere.} \tag{13}$$

Its time domain representation is given by the equation

$$w_{m,n,k}(t)=w\left(t-\frac{k}{n-m}\right)$$

$$=\frac{e^{in2\pi\{+-[k/(n-m)]\}}-e^{-im2\pi\{t-[k/(n-m)]\}}}{i2\pi(n-m)\left(t-\frac{k}{n-m}\right)}, \tag{14}$$

where m and n are positive real numbers. This is the expression for a harmonic wavelet centered at time $k/(n-m)$ and frequency $(m+n)\pi$ with bandwidth $(n-m)/(2\pi)$. The general harmonic scheme introduces an orthogonal transform as well. The complex general harmonic wavelet coefficients of a function $f(t)$ are determined by the equation

$$a[(m,n),k]=(n-m)\int_{-\infty}^{\infty}f(t)w^*\left(t-\frac{k}{n-m}\right)dt. \tag{15}$$

An improved version of the harmonic wavelet transform introduced by Newland ([16]) enhances the time resolution in the wavelet map for a given frequency resolution by filtering the mother wavelet function in the frequency domain using a Hanning filter. Specifically,

$$\hat{W}(\omega)=\frac{1}{(n-m)2\pi}\left[1-\cos\left(\frac{\omega-m2\pi}{n-m}\right)\right]. \tag{16}$$

For every scale, this representation yields wavelet coefficients corresponding to points spaced T units apart, where T is the sampling period. The improved harmonic scheme will be called henceforth the filtered harmonic scheme.

Wavelet Representation of Non-Stationary Excitations

The wavelet representation of a function $f(t)$ expressed in a discrete form with length $N=2^{\mathrm{n}}$ by using the dyadic scheme of the harmonic wavelet transform ([17]) is given by the equation

$$f(t)=a_0+\sum_{j=0}^{n-2}\sum_{k=0}^{2^{j-1}}[a_{j,k}w(2^j t-k)+a^*_{j,k}w^*(2^j t-k)]. \tag{17}$$

For a real and zero mean $f(t)$, $\alpha(0)=0$ and $\alpha^*_{\mathrm{j,k}}$ is the complex conjugate of $\alpha_{\mathrm{j,k}}$. The function $f(t)$ thus can be represented by the formula

$$f(t)=2\cdot\mathrm{Re}\sum_{j=0}^{n-2}\sum_{k=0}^{2^{j-1}}[a_{j,k}w(2^j t-k)]. \tag{18}$$

As shown by Priestley [18], the spectral representation of a nonstationary process may assume the form

$$f(t)=\int_{-\infty}^{\infty}\phi_t(\omega)dZ(\omega), \tag{19}$$

where $\varphi_{\mathrm{t}}(\omega)$ is an oscillatory function, and $Z(\omega)$ is a stochastic process with orthogonal increments. This oscillatory function can have the form of an amplitude-modulated trigonometric function,

$$\phi_t(\omega)=A_t(\omega)e^{i\theta(\omega)t}, \tag{20}$$

where $\theta(\omega)=\omega_0$ is the frequency at which $\varphi_{\mathrm{t}}(\omega)$ has its maximum fast Fourier transform (FFT) magnitude. The function $A_{\mathrm{t}}(\omega)$ acts as an envelope to the trigonometric function. The non-normalized evolutionary spectrum of the nonstationary process described by Eq. (18) is traditionally given by the formula

$$dH_i(\omega)=|A_t(\omega)|^2E[|dZ(\omega)|^2], \tag{21}$$

where E is the operator of mathematical expectation. In conjunction with the time-frequency localization attributes of wavelets, it has been suggested by Spanos and Tratskas [19] that representations (18) and (19) can be interpreted in a wavelet coefficient based local spectrum of a nonstationary process context. Specifically, the non-normalized local spectrum of a nonstationary process at scale j and time position k is given in that study by the formula

$$H(j,k)=E[|a(j,k)|^2]. \tag{22}$$

The normalized local spectrum of a nonstationary process is also defined in that study as the product of the non-normalized local spectrum multiplied by the energy of the wavelet function at the specific scale. That is,

$$S(j,k)=\frac{E[|a(j,k)|^2]}{2^j}. \tag{23}$$

This equation defines the local spectrum in specific time and frequency regions. In real units, these regions for the dyadic scheme involve the sampling period T, and the total number of data points per sample record N of $f(t)$. Specifically,

$$\frac{2^j 2\pi}{NT}\leqslant\omega\leqslant\frac{2^j 4\pi}{NT}; \quad \text{and}$$

$$\frac{NTk}{2^j}\leqslant t\leqslant\frac{NT(k+1)}{\cdot 2^j}. \tag{24}$$

If the filtered scheme is used, the local spectrum of the stationary process is determined by the formula

$$S((m,n),k)=\frac{E[|a((m,n),k)|^2]}{n-m}. \tag{25}$$

The applicable time-frequency regions are

$$\frac{m2\pi}{NT}\leqslant\omega\leqslant\frac{n2\pi}{NT}, \quad \text{with } n-m=10 \quad \text{and}$$

$$rT\leqslant t\leqslant(r+1)T, \quad \text{with } r=1{:}N-1. \tag{26}$$

Note that cross-wavelet coefficients can be used to derive the cross-spectrum between two stochastic processes ([19]). Such processes can be the excitation and the response of a dynamic system. The cross-wavelet coefficients (CWT's) between the excitation x_1 and the response f_r to a multi-degree-of-freedom (MDOF) system are given at scale j and time position k by the formula ([20])

$$E[CWTx_l f_r(j,k)]=E[WTf_r^*(j,k)WTx_l(j,k)]. \tag{27}$$

The cross spectrum between the excitation f_r and the response x_1 is then determined by the equation

$$CS(j,k)=\frac{E[|CWT(j,k)|^2]}{2^j}. \tag{28}$$

Wavelet Representation of a MDOF Linear System

A time-frequency representation of a multi-degree-of-freedom (MDOF) system is derived in this section. For clarity, the procedure is first demonstrated for single-degree-of-freedom systems. The equation that describes the response of a SDOF dynamic system is given by the formula

$$m\dot{x}+c\dot{x}+kx=f(t), \tag{29}$$

where m represents the mass of the system, c denotes the damping, k is the stiffness parameter, and f denotes the excitation. Although the symbol k has been already used earlier to denote wavelet time positions, it is also used in this section of the paper to denote, consistently with a standard vibrations notation, the system stiffness. The solution to the above equation can be derived either in the time domain or in the frequency domain. In the time domain the response is given as a combination of the homogeneous and a particular solution. The homogeneous solution can be expressed as

$$x_h(t)=\sum_{j=1}^{2} C_j e^{\lambda_j(t-t_0)}, \tag{30}$$

where C_j are the appropriate constants, and λ_j are the roots of the characteristic equation of the system. A particular solution is found by convolving the excitation $f(t)$ with the impulse response of the system $h(t)$. That is,

$$x_p(t)=\int_0^\infty h(t-\tau)f(\tau)d\tau. \tag{31}$$

The impulse response function for a lightly damped system is given by the equation

$$h(t)=\frac{1}{\omega_d}e^{-\zeta\omega_0 t}\sin(\omega_d t). \tag{32}$$

In this equation, $\omega_0=\sqrt{k/m}$ denotes the natural frequency, $\zeta=c/2\sqrt{km}$ denotes the damping ratio, and $\omega_d=\omega_0\sqrt{(1-\zeta^2)}$ denotes the damped natural frequency of the system. Well-known modifications of Eq. (32) apply for $\zeta\geqslant 1$.

In the frequency domain, the system is described by the frequency response function. This function is given by the equation

$$H(\omega)=\frac{1}{-m\omega^2+ic\omega+k}. \tag{33}$$

The system response is obtained in the frequency domain by using the Fourier representation $F(\omega)$ of the excitation and the frequency response function $H(\omega)$. Specifically,

$$X(\omega)=H(\omega)F(\omega). \tag{34}$$

The dyadic harmonic wavelet scheme generates its coefficients in two steps. First the discrete time sequence of the signal is converted to a frequency sequence through the Fourier transform. Second, octave bands of the frequency sequence are processed by the inverse Fourier transform to generate the wavelet coefficients at various scales ([5]).

In the case of a dynamic system representation, the frequency response function is a frequency sequence, when expressed in a discrete format. The Fourier coefficients of this sequence can be directly implemented into the second step of the harmonic wavelet scheme. After processing octave bands of the frequency response sequence, one obtains the harmonic wavelet coefficients of the system by inverse Fourier transforming. These coefficients provide a time-frequency representation of the system that originates from its frequency representation. The relationship between the frequency response function $H(\omega)$ and the harmonic wavelet coefficients $T(j,k)$ that represent the system at scale j and position k is derived by resorting to Eq. (9). That is,

$$T(j,k)=\int_{2\pi 2^j}^{4\pi 2^j} H(\omega)e^{i\omega k/2^j}d\omega. \tag{35}$$

The frequency response function $H(\omega)$ can be expressed as a discrete sequence in the form

$$H_{2^j+s}=2\pi H[\omega=2\pi(2^j+s)]. \tag{36}$$

Equation (35) can be then written in a discrete format. That is,

$$T(j,k)=\sum_{s=0}^{2^j-1} H_{2^j+s}e^{i2\pi sk/2^j}. \tag{37}$$

The representation of the system based on the wavelet coefficients extracted from its frequency response function as described by Eq. (35) or Eq. (37) will be called the "wavelet transfer function."

Obviously, the Fourier coefficients of the frequency response function that are used to provide information in the frequency domain can be distributed in octave blocks to provide time information, as well. The resolution of the wavelet map can be increased by using the filtered scheme of the harmonic wavelet transform which provides the best possible resolution in time, given the number of data points available per sample. The frequency resolution increases when the length of the frequency response sequence increases.

For multi-degree-of-freedom systems, the equation of motion can be written in the form

$$\underline{\underline{M}}\ddot{\underline{x}}+\underline{\underline{C}}\dot{\underline{x}}+\underline{\underline{K}}\underline{x}=\underline{F}, \tag{38}$$

where $\underline{\underline{M}}$ is the system mass matrix, $\underline{\underline{C}}$ is the system damping matrix, and $\underline{\underline{K}}$ is the system stiffness matrix; the symbol $\underline{F}$ denotes the excitation vector, and $\underset{\sim}{x}$ denotes the response vector. The system frequency response matrix is given by the equation

$$\underset{\approx}{H}(\omega)=[-\omega^2\underset{\approx}{M}+i\omega\underset{\approx}{C}+\underset{\approx}{K}]^{-1}. \tag{39}$$

One may derive the wavelet representation of a multi-degree-of-freedom system by following two different approaches. The first approach involves the use of the frequency response matrix, and the wavelet method presented in this section for SDOF systems. Specifically, each element of the frequency response matrix is treated separately to derive a wavelet representation by using Eq. (35). Further, assembling the wavelet representations for each element of the frequency response matrix, the "wavelet transfer tensor" of the MDOF system is formed. For a linear system with n excitations and m responses, the frequency response matrix, and the wavelet transfer tensor can be written as

$$\underset{\approx}{H}(\omega)=\begin{bmatrix} H_{11}(\omega) & H_{12}(\omega) & \dots & H_{1n}(\omega) \\ H_{21}(\omega) & \cdot & & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & & \cdot \\ H_{m1}(\omega) & \cdot & \dots & H_{mn}(\omega) \end{bmatrix}$$

and

$$\underset{\approx}{T}(j,k)=\begin{bmatrix} T_{11}(j,k) & T_{12}(j,k) & \dots & T_{1n}(j,k) \\ T_{21}(j,k) & \cdot & & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & & \cdot \\ T_{m1}(j,k) & \cdot & \dots & T_{mn}(j,k) \end{bmatrix}, \tag{40}$$

where $T_{lr}(j,k)$, $l=1,\dots,m$ and $r=1,\dots,n$ signify the system wavelet coefficient at scale j and time position k corresponding to the $H_{lr}(\omega)$ element of the frequency response matrix. The equation that describes this relationship is

$$T_{l,r}(j,k)=\int_{2\pi 2^j}^{4\pi 2^j} H_{lr}(\omega)e^{i\omega k/2^j}d\omega. \tag{41}$$

This approach is quite simple to implement since it assumes parallel computations to derive the wavelet coefficients corresponding to each element of the frequency response matrix.

The second approach relies on modal analysis, and on representing the system response in terms of its eigenmodes. The response of a MDOF system $x(t)$ is expressed via the eigenmodes of the system $(\underset{\sim}{\phi}^{(1)},\underset{\sim}{\phi}^{(2)},\dots,\underset{\sim}{\phi}^{(\mathrm{d})})$ in the form

$$\underset{\sim}{x}(t)=\sum_{q=1}^{d}\underset{\sim}{\phi}^{(j)}\eta_q(t). \tag{42}$$

Under commonly used approximations for the damping matrix C, the system response can be represented by a set of d uncoupled equations with each equation describing the motion in a particular mode of vibration. That is,

$$\ddot{\eta}_q(t)+2\zeta_q\omega_q\dot{\eta}_q(t)+\omega_q^2\eta_q(t)=g_q(t), \tag{43}$$

where ω_{q}, ζ_{q}, and $g_{\mathrm{q}}=(\underset{\sim}{\phi}^{(\mathrm{q})})^{\mathrm{T}}\underset{\sim}{F}/(\underset{\sim}{\phi}^{(\mathrm{q})})^{(\mathrm{T})}\underset{\approx}{M}\underset{\sim}{\phi}^{(\mathrm{q})}$, are the natural frequency, damping ratio, and participation factor of the qth mode, respectively. This approach decomposes the MDOF system to an ensemble of SDOF, oscillatory Eqs. (43), for which the wavelet transfer function can be readily determined by using Eq. (35). In this manner the wavelet representation of the MDOF system consists of the wavelet transfer functions assigned to the eigenmodes of the system.

Excitation-Response Relationships in the Wavelet Domain

A harmonic wavelet map provides a time-frequency representation with each scale related to a nonoverlapping frequency band. Therefore reconstructing a signal by using the wavelet coefficients that correspond to one scale only, would result in a monochromatic signal in the limit case of a narrow frequency band. Signal reconstruction based on the wavelet coefficients that correspond to a time interval and extend to all scales approaches in the limit case, the wavelet representation of a pulse signal. The pulse signals and the monochromatic sinusoids are the signals used in the generation of the impulse, and the frequency response functions, respectively. Therefore the harmonic wavelet representation includes elements of both the time and the frequency domains characteristic functions of a system. However, the harmonic wavelet transform cannot, due to the uncertainty principle, represent exactly neither a pulse nor a monochromatic signal.

By wavelet transforming Eq. (29) that describes a SDOF system, integrating by parts and using the chain rule of differentiation, one obtains the following equation ([21]):

$$m\frac{\partial}{\partial k^2}\mathrm{WT}_x(j,k)+c\frac{\partial}{\partial k}\mathrm{WT}_x(j,k)+\kappa\mathrm{WT}_x(j,k)=\mathrm{WT}_f(j,k), \tag{44}$$

where $\mathrm{WT_x}$ and $\mathrm{WT_f}$ denote the wavelet transforms of the response, and of the excitation, respectively.

To derive the wavelet representation of the response $\mathrm{WT_x}$ each scale j is considered separately. The wavelet coefficients of the excitation $\mathrm{WT_f}(j,.)$ that correspond to that scale form a vector with elements at various time positions. Using linear theory, $\mathrm{WT_x}(j,.)$ are derived by convolving the wavelet coefficients of the excitation with the impulse response of the corresponding to scale j subsystem. This subsystem describes the dynamic behavior of the original system, Eq. (44), in the frequency band associated with scale j and involves the energy of the frequency response function in a limited frequency region.

The impulse response of the subsystem $h_{\mathrm{j}}(k)$ is obtained in a discrete form by applying the inverse Fourier transform on the frequency response sequence H, Eq. (36), corresponding to the frequency range associated with scale j. That is,

$$h_j(k)=\sum_{s=0}^{2^j-1}H_{2^j+s}e^{i2\pi sk/2^i}. \tag{45}$$

Convolution is then considered in a discrete format between the impulse response of the subsystem $h_{\mathrm{j}}(k)$, and the wavelet coefficients of the excitation $\mathrm{WT_f}(j,.)$ to derive the wavelet coefficients of the response $\mathrm{WT_x}(j,.)$. Specifically,

$$\mathrm{WT}_x(j,k)=\sum_{n=0}^{2^j-1}h_j(n)\mathrm{WT}_f(j,k-n). \tag{46}$$

This calculation is repeated for every scale to derive the wavelet coefficients of the response.

By comparing the impulse response function of the subsystem $h_{\mathrm{j}}(k)$, Eq. (45), to the coefficients of the system wavelet transfer function $T(j,k)$ as described by Eq. (37), it is concluded that the two representations are identical. Both are derived by applying the inverse Fourier transform in blocks of the frequency response sequence of the system. Therefore Eq. (46) can be written in the form

$$\mathrm{WT}_x(j,k)=\sum_{z=0}^{z^j-1}T(j,z)\mathrm{WT}_f(j,k-z). \tag{47}$$

Equation (47) presents an excitation-response relationship given explicitly in the time-frequency domain. In the case of stochastic excitations the expected value of the response wavelet coefficients is given by the formula

$$E[\mathrm{WT}_x(j,k)] = \sum_{z=0}^{2^j-1} T(j,z)E[\mathrm{WT}_f(j,k-z)], \tag{48}$$

where the operator of mathematical expectation is applied on ensembles of system excitations and responses. Applying Eq. (48) to Eq. (23), the local spectrum at scale j and time position k of the response process is determined for the dyadic scheme. That is,

$$S(j,k) = \frac{E\left[\left(\sum_{z=0}^{2^j-1} T(j,z)\mathrm{WT}_f(j,k-z)\right)^2\right]}{2^j}. \tag{49}$$

Equation (49) determines the evolutionary spectrum of the response at scale j and time position k of the dyadic harmonic scheme in terms of the wavelet coefficients of the excitation process and of the system representation. For the filtered harmonic scheme, the response spectrum is expressed in the form

$$S((m,n),r) = \frac{E\left[\left(\sum_{z=0}^{n-m-1} T((m,n),z)\mathrm{WT}_f((m,n),r-z)\right)^2\right]}{n-m}. \tag{50}$$

In the case of a MDOF system, two approaches for deriving the wavelet representation of the response are possible based on the two different ways of representing the system in the wavelet domain, Eqs. (40) and (43). The first approach uses the frequency response matrix to derive the wavelet transfer tensor, Eqs. (40) and (41). The dyadic harmonic coefficients of the response process x_1 corresponding to the excitation process f_r are derived by the equation

$$\mathrm{WT}x_{tr}(j,k) = \sum_{z=0}^{2^j-1} T_{lr}(j,z)\mathrm{WT}f_r(j,k-z). \tag{51}$$

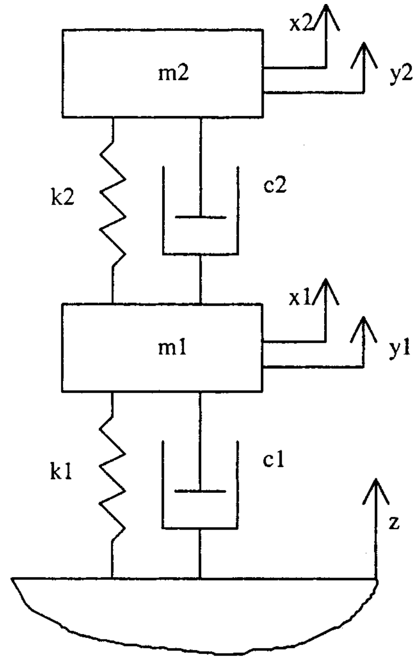


Fig. 1 Stochastic response of a 2DOF linear system under base excitation

The wavelet coefficients of the output due to each excitation are then added to provide the wavelet coefficients of the response $\mathrm{WT}x_1$. That is,

$$\mathrm{WT}x_l(j,k) = \sum_{r=1}^{m} \mathrm{WT}x_{lr}(j,k). \tag{52}$$

The preceding formula involves the superposition principle in the wavelet domain.

As an alternative, modal analysis can be employed. By wavelet transforming both sides of Eq. (42), which defines the response of

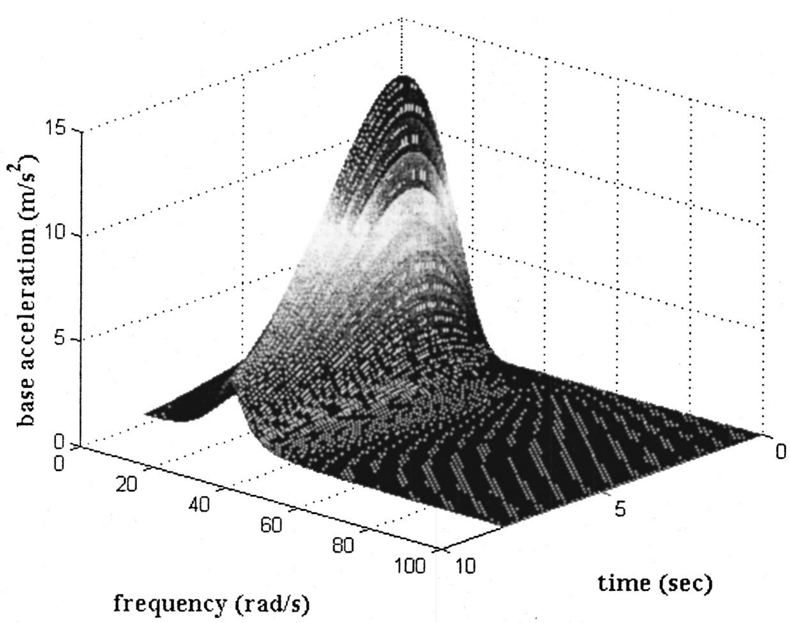


Fig. 2 Evolutionary spectrum of the base acceleration

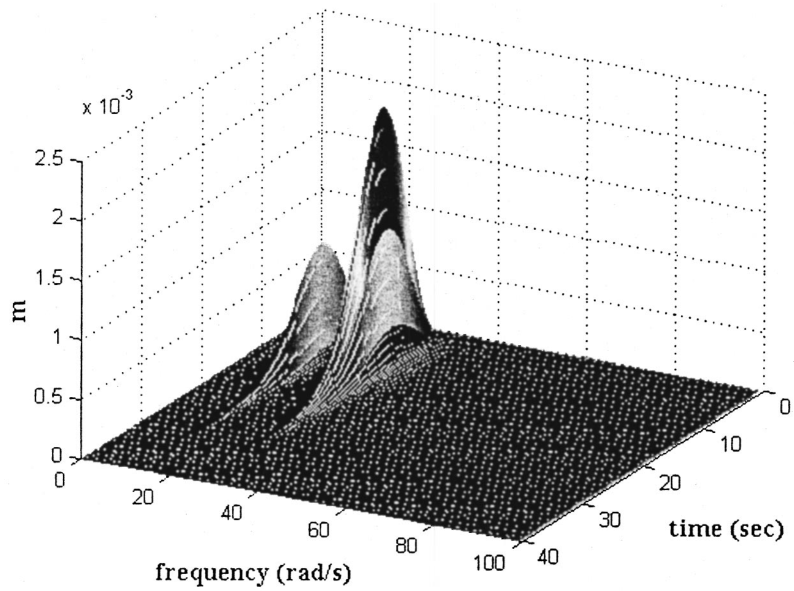


Fig. 3 Spectrum of the response x_1 by using the wavelet method

the system in the time domain, one may obtain the wavelet transform of the response with respect to the wavelet representation of the functions $\eta_q(t)$,

$$\mathrm{WT}\underline{x}(j,k)=\sum_{q=1}^{d}\underline{\phi}^{(q)}\mathrm{WT}\eta_q(j,k). \tag{53}$$

The wavelet coefficients corresponding to functions $\eta_q(t)$ are then obtained by wavelet transforming and Eq. (43), and using Eq. (47). This leads to the solution of the equation

$$\frac{\partial^2}{\partial k^2}\mathrm{WT}\eta_q(j,k)+2\zeta_q\omega_q\frac{\partial}{\partial k}\mathrm{WT}\eta_q(j,k)+\omega_q^2\mathrm{WT}\eta_q(j,k)$$
$$=\mathrm{WT}g_q(j,k). \tag{54}$$

2DOF System Response to Stochastic Excitation

In this section the usefulness of applying the wavelet-based method for deriving the response of MDOF linear systems to non-stationary excitations is assessed. For this purpose the 2DOF linear system shown in Fig. 1 is considered subject to base excitation. The absolute displacements of the two masses are denoted by y_1 and y_2, while x_1 and x_2 are the relative displacements with respect to the base displacement z; $\ddot{z}$ denotes the base acceleration. The equations of motion of the system are

$$m_1\ddot{x}_1+(c_1+c_2)\dot{x}_1-c_2\dot{x}_2+(k_1+k_2)x_1-k_2x_2=-m_1\ddot{z}, \tag{55}$$

and $$m_2\ddot{x}_2+c_2\dot{x}_2-c_2\dot{x}_1+k_2x_2-k_2x_1=-m_2\ddot{z}. \tag{56}$$

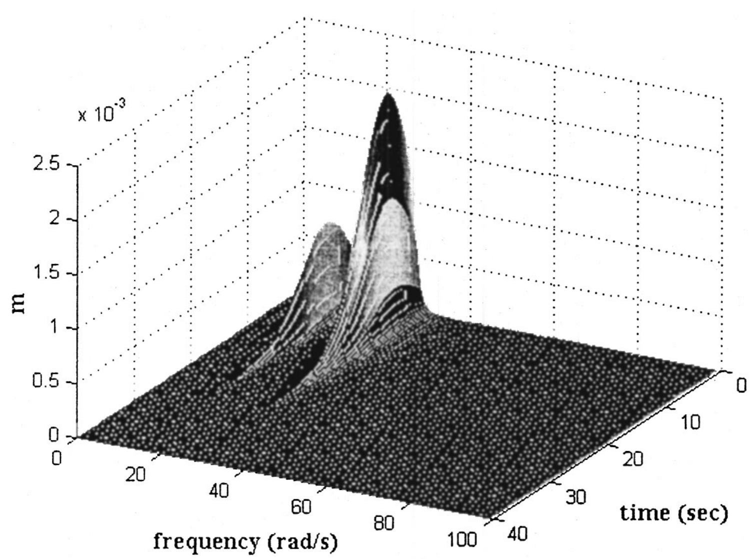


Fig. 4 Spectrum of the response x_1 by using Monte Carlo simulation

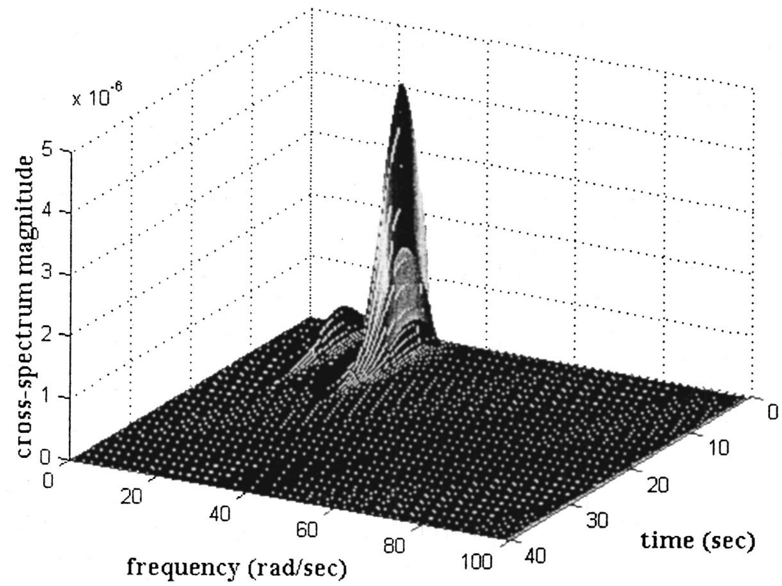


Fig. 5 Cross-spectrum magnitude between F_1 and x_1 derived by the wavelet-based method

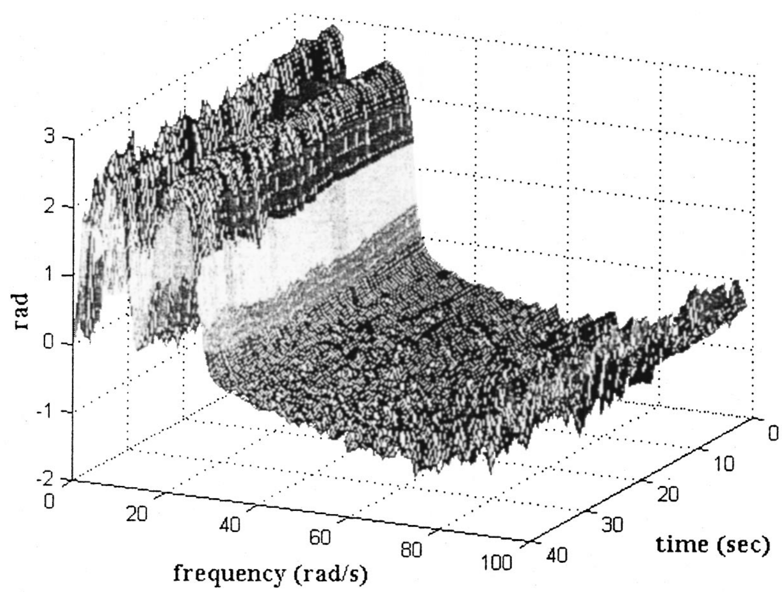


Fig. 6 Cross-spectrum phase between F_1 and x_1 derived by the wavelet-based method

The parameters m_1 and m_2 denote the mass of the two bodies, c_1 and c_2 are the damping coefficients, and k_1 and k_2 denote the stiffness coefficients. In context with the notation of Eq. (38), the mass, stiffness, and damping matrices, and the excitation and response vectors are

$$\underline{\underline{M}}=\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad \underline{\underline{C}}=\begin{bmatrix} c_1+c_2 & -c_2 \\ -c_2 & c_2 \end{bmatrix}, \quad \underline{\underline{K}}=\begin{bmatrix} k_1+k_2 & -k_2 \\ -k_2 & k_2 \end{bmatrix}, \tag{57}$$

$$\underline{F}=\begin{bmatrix} -m_1\ddot{z} \\ -m_2\ddot{z} \end{bmatrix}, \quad \underline{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{58}$$

Using the frequency response matrix, Eq. (39), which in this case is a 2×2 matrix, one can derive the wavelet transfer tensor of the system, Eqs. (40) and (41). By employing Eq. (48), the wavelet representation of the response is determined. Further, Eq. (50) provides the evolutionary response spectrum based on the filtered harmonic scheme.

A numerical example is considered. The following values for the system parameters are chosen: $m_1=12$ kg, $m_2=5$ kg, $k_1=4000$ N/m, $k_2=2000$ N/m, $c_1=8$ N/m/s, $c_2=2$ N/m/s. A filtered white noise is used as the excitation. A second-order filter with natural frequency equal to $\omega_o=18$ rad/s, and damping ratio equal to $\zeta=0.2$ is used to generate the excitation process. Those values are chosen such that the excitation spectrum includes energy in the band between the two natural frequencies of the system which are 13.7 and 26.7 rad/s. The evolutionary spectrum of the base acceleration is shown in Fig. 2.

Through filtering, 300 records of the excitation process are generated. The spectra of the system response are derived by two methods. First the wavelet method of determining the wavelet coefficients of the response process and the corresponding values of its evolutionary spectrum as captured by Eq. (46) and Eq. (23) is employed. Second, Monte Carlo simulation is used to derive the spectrum of the system response ([19]).

Pertinent results are shown in Figs. 3 and 4. It is seen that the evolutionary spectrum generated by the wavelet-based method is virtually indistinguishable from the one generated by Monte Carlo simulation. It accurately locates the magnitude and position of the response peaks in the time-frequency plane. Besides providing an explicit excitation-response relationship in the time-frequency domain, the proposed wavelet-based method is computationally advantageous compared to the Monte Carlo simulation. This is due to the efficiency of the convolution only, which is required to determine the expected value of the response wavelet coefficients, versus the numerical integrations and the spectral processing required by the Monte Carlo method. Similar results have been obtained for the response variable $x_2(t)$; they are not shown here for brevity.

In Figs. 5 and 6 the cross-spectrum magnitude between the excitation process F_1 and the response process x_1 are presented as derived by both the Monte Carlo simulation and the wavelet-based analytical approach using Eq. (28); they are practically indistinguishable from the relevant Monte Carlo simulation results which are not shown for brevity.

Concluding Remarks

The wavelet transform has been applied to the problem of determining the response of a MDOF linear system to nonstationary stochastic excitation. A time-frequency representation of the frequency response matrix has been developed involving the harmonic wavelet scheme. Further, excitation-response relationships in the harmonic representation scheme have been derived. Specifically, it has been shown that the wavelet coefficients of the response can be determined by convolving the wavelet coefficients of the excitation with the wavelet representation of the system at each scale. Numerical results pertaining to the response of a 2DOF linear system subject to filtered white noise excitation have shown that the proposed method captures quite reliably the variation of the frequency temporal content of the system response.

Acknowledgment

The support of this work from a grant from the Basic Engineering Sciences Program of the Department of Energy is gratefully acknowledged.

Nomenclature

A = modulating envelope
CWT = cross-wavelet transform
CS = cross-wavelet spectrum
DWT = discrete wavelet transform
E = mathematical expectation
F = Fourier transform of signal f
$\underline{F}$ = excitation vector
H = evolutionary spectrum of a non-stationary process
$\underline{H}$ = frequency response matrix
$\underline{K}$ = stiffness matrix
$\underline{M}$ = mass matrix
N = number of sample points
S = evolutionary spectrum derived by the wavelet method
T = wavelet transfer function
$\underline{T}$ = wavelet transfer tensor
W = Fourier transform of the wavelet function
$\hat{W}$ = Fourier transform of the filtered harmonic wavelet
WT = wavelet transform
X = Fourier transform of the response
Z = stochastic process with orthogonal increments
b = wavelet time position
c = damping coefficient
f = excitation signal
h = impulse response function
j = DWT scale
k = DWT time position
m = mass
t = time
w = wavelet function
x = response
α = wavelet scale
α = harmonic wavelet coefficient
ζ = damping ratio
λ = eigenvalue
ω = frequency
ω_o = natural frequency
ω_d = damped natural frequency

References

[1] Mallat, S., 1989, ''A Theory for Multi-Resolution Signal Decomposition: The Wavelet Representation,'' IEEE Pattern Anal. and Machine Intell., **11**, pp. 674–693.
[2] Mallat, S., 1989, ''Multi-Resolution Approximations and Wavelets,'' Trans. Am. Math. Soc., **315**, pp. 69–88.
[3] Daubechies, I., 1992, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA.
[4] Meyer, Y., 1992, *Wavelets and Operators*, Cambridge University Press, New York.
[5] Newland, D. E., 1993, *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*, Longman Scientific & Technical, London.
[6] Spanos, P. D., and Zeldin, B. A., 1997, ''Wavelet Concepts on Computational Stochastic Mechanics,'' Journal of Probabilistic Engineering Mechanics, **12**, pp. 244–255.
[7] Carmona, R., Hwang, W. J., and Torresani, B., 1998, *Practical Time-Frequency Analysis*, Academic Press, New York.
[8] Diaz, A., 1999, ''A Wavelet-Galerkin Scheme for the Analysis of Large-Scale Problems on Simple Domains,'' Int. J. Numer. Methods Eng., **44**, pp. 1599–1616.
[9] Youche, Z., Jizeng, W., and Xiaojing, Z., 1998, ''Applications of Wavelet FEM to Bending of Beam and Plate Structures,'' Appl. Math. Mech., **19**, pp. 745–755.
[10] Mei, H., and Agrawal, O. P., 1998, ''Wavelet Based Model for Stochastic Analysis of Beam Structures,'' AIAA J., **36**, pp. 465–470.
[11] Basu, B., and Gupta, V. K., 1999, ''Wavelet-Based Analysis of a Non-Stationary Response of a Slipping Foundation,'' J. Sound Vib., **222**, pp. 547–563.
[12] Basu, B., and Gupta, V. K., 1997, ''Non-Stationary Seismic Response of MDOF Systems by Wavelet Transform,'' Earthquake Eng. Struct. Dyn., **26**, pp. 1243–1258.
[13] Basu, B., and Gupta, V. K., 2000, ''Stochastic Seismic Response of Single-Degree-of-Freedom Systems through Wavelets,'' Eng. Struct., **22**, pp. 1714–1722.
[14] Iyama, J., and Kuwamura, H., 1999, ''Applications of Wavelets to Analysis and Simulation of Earthquake Records,'' Earthquake Eng. Struct. Dyn., **28**, pp. 255–272.
[15] Newland, D. E., 1994, ''Wavelet Analysis of Vibration, Part I: Theory,'' J. Vibr. Acoust., **116**, pp. 409–416.
[16] Newland, D. E., 1999, ''Ridge and Phase Identification in the Frequency Analysis of Transient Signals by Harmonic Wavelets,'' J. Vibr. Acoust., **121**, pp. 149–155.
[17] Newland, D. E., 1994, ''Wavelet Analysis of Vibration, Part II: Wavelet Maps,'' J. Vibr. Acoust., **116**, pp. 417–425.
[18] Priestley, M. B., 1988, *Nonlinear and Non-Stationary Time Series Analysis*, Academic Press, New York.
[19] Spanos, P. D., and Tratskas, P., 2001, ''Power spectrum estimation of stochastic processes by using the wavelet transform,'' *Proceedings of the 8th International Conference on Structural Safety and Reliability*, Newport, CA.
[20] Kyprianou, A., and Staszewski, W. J., 1999, ''On the Cross-Wavelet Analysis of Duffing Oscillator,'' J. Sound Vib., **228**, pp. 199–210.
[21] Basu, B., and Gupta, V. K., 1998, ''Seismic Response of SDOF Systems by Wavelet Modeling of Non-Stationary Processes,'' J. Eng. Mech., **124**, pp. 1142–1150.

K. Yagasaki
Department of Mechanical and Systems Engineering,
Gifu University,
Gifu, Gifu 501-1193, Japan

Nonintegrability of an Infinite-Degree-of-Freedom Model for Unforced and Undamped, Straight Beams

We study a mathematical model for unforced and undamped, initially straight beams. This system is governed by an integro-partial differential equation, and its energy is conserved: It is an infinite-degree-of-freedom Hamiltonian system. We can derive "exact" finite-degree-of-freedom mode truncations for it. Using the differential Galois theory for Hamiltonian systems, we prove that any two or more modal truncations for the model are nonintegrable in the following sense: The Hamiltonian systems do not have the same number of "meromorphic" first complex integrals which are independent and in involution, as the number of their degrees of freedom, when they are regarded as Hamiltonian systems with complex time and coordinates. This also means the nonintegrability of the infinite-degree-of-freedom model for the beams. We present numerical simulation results and observe that chaotic motions occur as in typical nonintegrable Hamiltonian systems.
[DOI: 10.1115/1.1602483]

1 Introduction

In this paper we study a mathematical model for an unforced and undamped, initially straight beam ([1–10]), such as shown in Fig. 1,

$$\ddot{u}+u''''-\left[\Gamma+\kappa\int_0^1 (u')^2 d\zeta\right]u''=0, \tag{1}$$

where $u=u(t,z)$ represents the transverse displacement, and the overdot and prime represent partial differentiation with respect to t and z, respectively. The expression in the bracket of Eq. (1) approximately gives the extensive force, where Γ and κ are constants and the former especially represents the initial tension. See, e.g., Refs. [1], [2], [5], [10] for the derivation of Eq. (1). For simplicity of computation, we choose hinged ends as the boundary conditions:

$$u(t,0)=u(t,1)=0, \quad u''(t,0)=u''(t,1)=0. \tag{2}$$

Equation (1) represents an infinite-degree-of-freedom Hamiltonian system (see Appendix A).

If Γ is negative and its absolute value is sufficiently large, then Eq. (1) can be used to analyze vibrations of a buckled beam. In this situation, extending a global perturbation technique called Melnikov's method ([11,12]), Holmes and Marsden [13] showed that chaotic dynamics occur when the beam is subjected to linear damping and periodic, transverse excitation. Such chaotic vibrations in related experimental models were also observed by Tseng and Dugundji [14] and subsequently studied by Moon and Holmes [15]. Furthermore, it was proven recently by Melnikov-type techniques ([16,17]) that very complicated behaviors occur even in the absence of the damping and periodic excitation ([18]): chaos and fast diffusion for which a mechanism is very similar to Arnold diffusion ([19,20]) occurs. Numerical observations of these behaviors were provided in Ref. [21].

On the other hand, for the case in which the beam is unbuckled or straight (i.e., Γ is positive or its absolute value is sufficiently small), free and forced periodic vibrations were analytically and experimentally investigated in an early stage of this subject ([1–9]). In particular, Tseng and Dugundji [7] analytically and experimentally studied harmonic, subharmonic, superharmonic, and ultrasubharmonic motions under damping and periodic forcing. Ultrasubharmonic vibrations were also analyzed in Refs. [8], [9]. See Ref. [10] for more details and references on early researches. Moreover, it was shown theoretically in Ref. [22] and experimentally in Ref. [23] that chaotic vibrations may occur when the straight beam is subjected to two-frequency quasiperiodic forcing.

The objective of this paper is to prove the nonintegrability of the infinite-degree-of-freedom model (1) for unforced and undamped, initially straight beams. To this end, we use the differential Galois theory for Hamiltonian systems ([24,25]) and its application to a special class of two-degree-of-freedom Hamiltonians ([25–27]). Here "nonintegrability" means that the Hamiltonian system does not have the same number of "meromorphic" first complex integrals which are independent and in involution, as the number of its degrees of freedom, when it is regarded as a Hamiltonian system with complex time and coordinates. When a Hamiltonian system is nonintegrable in this sense, we can often find chaotic motions in the system although theoretical explanations are seldom given (exceptions were presented in Refs. [25], [28], [29]). We also remark that the presence of chaotic motions implies the nonintegrability of the system ([30]).

The outline of this paper is as follows. In Sec. 2 we derive "exact" finite-degree-of-freedom modal truncations for Eq. (1). In Secs. 3 and 4, we review the differential Galois theory for Hamiltonian systems and its application to a class of two-degree-of-freedom systems with potentials. In Sec. 5 we apply the general results to the finite-degree-of-freedom modal truncations and show that any truncations of Eq. (1) with two or more modes are nonintegrable. This also implies the nonintegrability of the infinite-degree-of-freedom model (1). In Sec. 6 we give numeri-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 4, 2002; final revision, Feb. 20, 2003. Associate Editor: O. O'Reilly. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

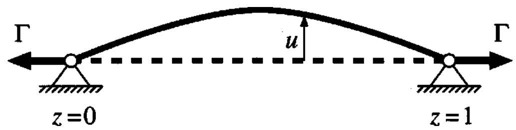


Fig. 1 Transverse deformation of an initially straight beam

cal simulation results and observe that chaotic motions arise as in typical nonintegrable Hamiltonian systems. Finally, we conclude with a summary and some comments in Sec. 7.

2 "Exact" Finite-Degree-of-Freedom Modal Truncations

The eigenvalues λ_j and eigenfunctions u_j, $j=1,2,\ldots$, for the linearization of Eq. (1) about the trivial solution $u=0$ form a countable set:

$$\lambda_j=\pm i(j\pi)\sqrt{\Gamma+(j\pi)^2},\quad u_j(z)=\sin j\pi z,\quad j=1,2,\ldots. \tag{3}$$

For $N>0$ an integer, let

$$u(z,t)=\sum_{l=1}^{N} a_l(t)\sin j_l\pi z, \tag{4}$$

where j_l, $l=1,\ldots,N$, are distinct, positive integers. Note that Eq. (4) satisfies the boundary conditions (2). Substitution of Eq. (4) into Eq. (1) yields

$$\sum_{l=1}^{N}\left\{\ddot{a}_l+(j_l\pi)^2\left[(j_l\pi)^2+\Gamma+\frac{\kappa}{2}\sum_{r=1}^{N}(j_r\pi)^2a_r^2\right]a_l\right\}\sin j_l\pi z=0, \tag{5}$$

from which we obtain

$$\ddot{a}_l+(j_l\pi)^2\left[(j_l\pi)^2+\Gamma+\frac{\kappa}{2}\sum_{r=1}^{N}(j_r\pi)^2a_r^2\right]a_l=0,\quad l=1,\ldots,N. \tag{6}$$

Conversely, if $a_l(t)$, $l=1,\ldots,N$, satisfy Eq. (6), then Eq. (4) gives an *exact* solution of Eq. (1).

Scaling the time variable $t\mapsto\omega_0 t$ with $\omega_0=\pi\sqrt{\Gamma+\pi^2}$ and the coordinate variables as $a_l\mapsto(\pi\sqrt{\kappa/2(\Gamma+\pi^2)})a_l$ in Eq. (6), we have

$$\ddot{a}_l+\omega_l^2a_l+j_l^2\left(\sum_{r=1}^{N}j_r^2a_r^2\right)a_l=0,\quad l=1,\ldots,N, \tag{7}$$

where $\omega_l=j_l\sqrt{(\Gamma+(j_l\pi)^2)/(\Gamma+\pi^2)}$. Finally, setting $x_l=a_l$ and $y_l=\dot{a}_l$, we obtain

$$\dot{x}_l=y_l,\quad \dot{y}_l=-\omega_l^2x_l-j_l^2\left(\sum_{r=1}^{N}j_r^2x_r^2\right)x_l,\quad l=1,\ldots,N, \tag{8}$$

which is an N-degree-of-freedom Hamiltonian system with a Hamiltonian function

$$H_N(x,y)=\frac{1}{2}\sum_{l=1}^{N}\omega_l^2x_l^2+\frac{1}{4}\left(\sum_{l=1}^{N}j_l^2x_l^2\right)^2+\frac{1}{2}\sum_{l=1}^{N}y_l^2, \tag{9}$$

where $x=(x_1,\ldots,x_N)$ and $y=(y_1,\ldots,y_N)\in\mathbb{R}^N$.

3 Differential Galois Theory for Hamiltonian Systems

In this and the next section we outline the differential Galois theory of Ref. [24] for Hamiltonian systems and its application to two-degree-of-freedom systems with potentials of a special form. See Refs. [24–27] for more details on the theory and results.

Consider an N-degree-of-freedom canonical Hamiltonian system

$$\dot{z}=J_N\mathrm{D}_zH(z),\quad z\in\mathbb{R}^{2N}, \tag{10}$$

where $H:\mathbb{R}^{2N}\to\mathbb{R}$ is analytic and J_N is the $2N\times 2N$ canonical symplectic matrix

$$J_N=\begin{pmatrix}0 & \mathrm{id}^N\\ -\mathrm{id}^N & 0\end{pmatrix} \tag{11}$$

with id^N the $N\times N$ identity matrix. We regard the Hamiltonian system (10) as a restriction to $\mathbb{R}^{2N}$ of a complex analytic Hamiltonian system (with complex time) defined in a domain $D\subset\mathbb{C}^{2N}$.

We call a function $f:D\to\mathbb{C}$ a *first integral* of Eq. (10) if

$$\{f,H\}\equiv\mathrm{D}_zf(z)\cdot J_N\mathrm{D}_zH(z)=\mathrm{D}_zf(z)\cdot\dot{z}=\frac{\mathrm{d}}{\mathrm{d}t}[f(z(t))]=0, \tag{12}$$

i.e., f is constant along any trajectory of Eq. (10). Here the bracket is the canonical *Poisson bracket* and "$\cdot$" represents the inner product. We say that first integrals $f_1,\ldots,f_k$ $(2\leqslant k\leqslant N)$ are *in involution* (on a set) if the Poisson brackets $\{f_i,f_j\}\equiv\mathrm{D}_zf_i\cdot J_N\mathrm{D}_zf_j=0$ (on the set) for any $i\neq j$, and that they are *independent* (on a set) if $\mathrm{D}_zf_1,\ldots,\mathrm{D}_zf_k$ are independent (on the set). The Hamiltonian system (10) is said to be (meromorphic) *integrable* in a set $U\subset D$ if there are N (meromorphic) first integrals which are in involution and independent on an open dense subset of U. If it is not (meromorphic) integrable in U, it is said to be (meromorphic) nonintegrable in U. When the Hamiltonian system (10) is integrable, we can obtain general solutions using N first integrals even if they are not complex analytic ([31]).

Let C be an integral curve of the Hamiltonian system (10) and denote by $z_C(t)$ the trajectory of C. The *variational equation* (VE) along C is given by

$$\dot{\xi}=J_N\mathrm{D}_z^2H(z_C(t))\xi. \tag{13}$$

If there are $k(<N)$ first integrals $f_1,\ldots,f_k$ which are in involution and independent in a neighborhood of C, then $\xi=J_N\mathrm{D}_zf_i(z_C(t))$, $i=1,\ldots,k$, are solutions to Eq. (13). Using this fact, we can reduce Eq. (13) to a $2(N-k)$-dimensional linear differential equation called the *normal variational equation* (NVE) along C when k such first integrals exist ([24,25]). Since the Hamiltonian function is a first integral itself, we always have a $2(N-1)$-dimensional NVE along C.

In the context of Eq. (13), we now give background information on the differential Galois theory for linear differential equations ([32–34]). See Appendix B for some basic algebraic terminologies. We first note that the coefficients of Eq. (13) can be generally considered to be in the field K of meromorphic functions in D. Let $\xi_1,\ldots,\xi_{2N}$ be a fundamental system of solutions to Eq. (13). A set of rational functions of $\xi_1,\ldots,\xi_{2N}$ with coefficients in K, which we denote by $L=K(\xi_1,\ldots,\xi_{2N})$, is also a field. We refer to L as the *Picard-Vessiot extension* of K for Eq. (13), and define the *Galois group* of Eq. (13), $G=\mathrm{Gal}_K(L)=\mathrm{Gal}(L/K)$, as the group of all (differential) automorphisms of L such that K is left invariant. From a fundamental result in differential Galois theory ([32–34]), Eq. (13) is integrable in the meaning that its general solution can be expressed by a combination of integrals, exponentials of integrals, and algebraic functions of elements of K if and only if the identity component of G, which we denote by G^0, is solvable. In particular, if G^0 is abelian, then Eq. (13) is integrable.

In such a situation, extending a result of Ziglin [35], Morales-Ruiz and Ramis [24,25] proved that if there are N meromorphic first integrals of Eq. (10) which are in involution and independent over a neighborhood of C (not necessarily on C itself), then the identity component of the Galois group of the VE (13) is abelian. Moreover, if k of the first integrals are (functionally) independent on C itself, then the identity component of the Galois group of the associated $2(N-k)$-dimensional NVE is abelian. Thus, if the identity component of the Galois group of the VE or NVE is not abelian, then the Hamiltonian system (10) is nonintegrable in the sense that there do not exist N meromorphic first integrals which are in involution and independent.

The Galois group G can generally be represented as a group of nonsingular matrices with constant coefficients. So we only have to study the matrix representation in order to determine whether G is abelian (or whether it is solvable). At present the task is often routine to some extent. See, e.g., Refs. [25], [33], [34], [36] and references therein for more details. In the next section we outline one of such results in Refs. [25], [36], which plays an essential role in our analysis of the beam equation (1) and its finite-degree-of-freedom mode truncation (8).

4 Two-Degree-of-Freedom Systems With Potentials

Now consider a two-degree-of-freedom canonical Hamiltonian system with a potential $V(x_1,x_2)$,

$$\dot{x}_j=y_j, \quad \dot{y}_j=-\frac{\partial V}{\partial x_j}(x_1,x_2), \quad j=1,2. \tag{14}$$

The Hamiltonian function for Eq. (14) is

$$H=\frac{1}{2}(y_1^2+y_2^2)+V(x_1,x_2). \tag{15}$$

We also assume that the potential V is an analytic function in $\mathbb{R}^2$ and has the form

$$V(x_1,x_2)=\varphi(x_1)-\frac{1}{2}\alpha(x_1)x_2^2+\mathcal{O}(x_2^3) \tag{16}$$

where φ and α are analytic functions. When $N=2$, Eq. (8) has the form of Eq. (14) with

$$\varphi(x_1)=\frac{1}{2}\omega_1x_1^2+\frac{1}{4}j_1^4x_1^4, \quad \alpha(x_1)=-(\omega_2^2+j_1^2j_2^2x_1^2). \tag{17}$$

The (x_1,y_1) plane is invariant under the flow of Eq. (14) and the equations of motion restricted to it are

$$\dot{x}_1=y_1, \quad \dot{y}_1=-\frac{\mathrm{d}\varphi}{\mathrm{d}x_1}(x_1). \tag{18}$$

We assume that on the invariant plane there exists a continuous family of integral curves C_h parametrized by the Hamiltonian energy h.

Let $(x_1^h(t),y_1^h(t))$ be the trajectory of Eq. (18) for C_h. We also denote by C_h the integral curve corresponding to $(x_1^h(t),y_1^h(t))$ in the (x_1,y_2,x_2,y_2) space. The VE of Eq. (14) along C_h is

$$\dot{\xi}_1=\xi_3, \quad \dot{\xi}_2=\xi_4, \quad \dot{\xi}_3=-\frac{\mathrm{d}^2\varphi}{\mathrm{d}x_1^2}(x_1^h(t))\xi_1, \quad \dot{\xi}_4=\alpha(x_1^h(t))\xi_2,$$

or equivalently

$$\ddot{\xi}_1+\frac{\mathrm{d}^2\varphi}{\mathrm{d}x_1^2}(x_1^h(t))\xi_1=0, \quad \ddot{\xi}_2-\alpha(x_1^h(t))\xi_2=0. \tag{19}$$

The first equation of Eq. (19) has a solution $\xi_1=y_2^h(t)$ due to the fact that the Hamiltonian function is a first integral. The NVE of Eq. (14) along C_h is

$$\dot{\eta}_1=\eta_2, \quad \dot{\eta}_2=\alpha(x_1^h(t))\eta_1,$$

or equivalently

$$\ddot{\eta}_1-\alpha(x_1^h(t))\eta_1=0, \tag{20}$$

which is a nontrivial part of the VE (19). Morales-Ruiz and Simó [25,27] showed that, for several cases of φ and α including Eq. (17), the NVE (20) can be transformed into the Weierstrass form of the Lamé equation ([37]):

$$\ddot{\eta}_1-(A\mathcal{P}(t)+B)\eta_1=0, \tag{21}$$

where $\mathcal{P}(t)$ is the Weierstrass elliptic function, which is a solution of

$$(\dot{z})^2=4z^3-g_2z-g_3, \tag{22}$$

where g_2 and g_3 are constants such that $27g_3^2-g_2^3\neq 0$, and A and B are complex numbers.

Suppose that the NVE (20) is of the type (21). Denote $\hat{\alpha}(t,h)=\alpha(x_1^h(t))$. Noting that $\mathcal{P}$, A, and B depend on h, we see that

$$\hat{\alpha}(t,h)=A(h)\mathcal{P}(t,h)+B(h) \tag{23}$$

must hold. Since $\mathcal{P}(t)$ satisfies Eq. (22), we have

$$\begin{aligned}\dot{\hat{\alpha}}^2(t,h)=&\frac{4}{A(h)}\hat{\alpha}^3(t,h)-\frac{12B(h)}{A(h)}\hat{\alpha}^2(t,h)\\&+\left(\frac{12B^2(h)}{A(h)}-g_2A(h)\right)\hat{\alpha}(t,h)\\&+\left(-\frac{4B^3(h)}{A(h)}+g_2A(h)B(h)-g_3A^2(h)\right),\end{aligned} \tag{24}$$

where we used Eq. (23). Let $\hat{x}_1(\alpha_0)$ be a (possibly multivalued) function such that $\alpha(\hat{x}_1(\alpha_0))=\alpha_0$. Since $x_1^h(t)$ is a solution of

$$\frac{1}{2}\dot{x}_1^2+\varphi(x_1)=h, \tag{25}$$

we also have

$$\dot{\hat{\alpha}}^2(t,h)=2[\alpha'(\hat{x}_1(\hat{\alpha}(t,h)))]^2[h-\varphi(\hat{x}_1(\hat{\alpha}(t,h)))]. \tag{26}$$

Hence by Eqs. (24) and (26) we can write

$$\dot{\hat{\alpha}}^2=P_1(\hat{\alpha})+hP_2(\hat{\alpha}), \tag{27}$$

where $P_j(\alpha)$, $j=1,2$, are cubic polynomials of α, i.e.,

$$P_j(\alpha)=a_j\alpha^3+b_j\alpha^2+c_j\alpha+d_j, \quad j=1,2, \tag{28}$$

with a_j, b_j, c_j, and d_j, $j=1,2$, constants. Moreover, comparing Eqs. (26) and (27), we have

$$\varphi(\hat{x}_1(\alpha))=-\frac{P_1(\alpha)}{P_2(\alpha)}, \quad [\alpha'(x_1)]^2=\frac{1}{2}P_2(\alpha(x_1)). \tag{29}$$

Conversely, if there are cubic polynomials $P_j(\alpha)$, $j=1,2$, satisfying Eq. (29), then the NVE (20) is of the Lamé type (21).

In this situation, using a result on the integrability of the Lamé equation (21) ([25,36]) and the differential Galois theory of Refs. [24], [25] described in Sec. 3, Morales-Ruiz and Simó [25,27] proved that the Hamiltonian system (14) is nonintegrable if $a_2\neq 0$ or if none of the following conditions holds:

(i) $a_1=4/n(n+1)$ for some $n\in\mathbb{N}$;

(ii) $a_1=16/(4m^2-1)$ for some $m\in\mathbb{N}$, $b_2=0$ and either

(iia) $m=1$ and $b_1=0$;

(iib) $m=2$ and $c_2=16a_1c_1+3b_1^2=0$;

(iic) $m=3$, $16a_1d_2+11b_1c_2=1024a_1^2d_1+704a_1b_1c_1+45b_1^3=0$;

or

(iid) $m>3$, $b_1=0$ and either

(iid1) $m=1$, 2, 4 or 5 mod 6 and $c_1=c_2=0$

or

(iid2) m is odd and $d_1=d_2=0$;

(iii) $a_1=4/n(n+1)$ with $n+(1/2)\in(1/3)\mathbb{Z}\cup(1/4)\mathbb{Z}\cup(1/5)\mathbb{Z}\setminus\mathbb{Z}$, $b_2=0$ and either

(iiia) $c_2=b_1^2-3a_1c_1=0$;

or

(iiib) $c_2b_1-3a_1d_2=2b_1^3-9a_1b_1c_1+27a_1^2d_1=0$.

The condition (iid) was obtained based on a conjecture which was confirmed by a numerical computation ([25,27]). However, concerning condition (ii), we only require the first equality, $a_1 = 16/(4m^2-1)$, later, and our main result, given in the next section, does not depend on the conjecture.

5 Nonintegrability of the Straight Beam Model

We return to the ''exact'' finite-degree-of-freedom model (8) for the straight beam. We first set $N=2$ and apply the theoretical result stated in Sec. 4. From Eqs. (17) and (29) we obtain

$$P_1(\alpha)=\frac{2}{\beta_2}(\alpha+\omega_2^2)^2[\beta_1(\alpha+\omega_2^2)-2\beta_2\omega_1^2],$$
$$P_2(\alpha)=-8\beta_1\beta_2(\alpha+\omega_2^2), \tag{30}$$

where $\beta_l=j_l^2$. Hence

$$a_1=\frac{2\beta_1}{\beta_2}, \quad b_1=-4\omega_1^2+\frac{6\beta_1}{\beta_2}\omega_2^2,$$
$$c_1=\left(-8\omega_1^2+\frac{6\beta_1}{\beta_2}\omega_2^2\right)\omega_2^2, \quad d_1=\left(-4\omega_1^2+\frac{2\beta_1}{\beta_2}\omega_2^2\right)\omega_2^4, \tag{31}$$
$$a_2=b_2=0, \quad c_2=-8\beta_1\beta_2, \quad d_2=-8\beta_1\beta_2\omega_2^2.$$

Moreover, comparing Eqs. (24) and (27), we have

$$g_2=\frac{4}{3}(3\beta_1^2h+\omega_1^4), \quad g_3=\frac{4\omega_1^2}{27}(9\beta_1h+2\omega_1^4), \tag{32}$$

so that

$$27g_3^2-g_2^3=-16\beta_1^4h^2(4\beta_1^2h+\omega_1^4), \tag{33}$$

which is nonzero unless $h=0$.

First, we easily see that condition (iii) does not hold since $c_2 \neq 0$ and

$$c_2b_1-3a_1d_2=-32\beta_1\beta_2\omega_1^2\neq 0,$$
$$2b_1^3-9a_1b_1c_1+27a_1^2d_1=-128\omega_1^6\neq 0. \tag{34}$$

Next, suppose that condition (ii) holds. Then we have $a_1/2 = j_1^2/j_2^2 = 2^3/(4m^2-1)$ so that we can write $j_1=2^2p$ for some $p \in \mathbb{N}$. Hence we obtain $j_2^2=2p^2(4m^2-1)$, which means that $4m^2-1$ is even. So, condition (ii) does not hold. Finally, suppose that condition (i) holds and

$$\frac{j_1^2}{j_2^2}=\frac{2}{n(n+1)} \tag{35}$$

for some $n\in\mathbb{N}$. In fact, there are infinitely many pairs (j_1,j_2) such that Eq. (35) holds for some $n\in\mathbb{N}$ (see Appendix A of Ref. [18] for a proof of this fact when $j_1=1$). However, unless $j_1 = j_2$, i.e., $n=1$, there does not exist $n'\in\mathbb{N}$ such that

$$\frac{j_2^2}{j_1^2}=\frac{n(n+1)}{2}=\frac{2}{n'(n'+1)}. \tag{36}$$

Replacing $l=1$ with $l=2$ and applying the above arguments, we easily prove the nonintegrability of Eq. (8) with $N=2$ near a trajectory $(x_1,y_1,x_2,y_2)=(0,0,x_2^h(t),y_2^h(t))$ where $(x_2^h(t),y_2^h(t))$ is a solution of

$$\dot{x}_2=y_2, \quad \dot{y}_2=-\omega_2^2x_2-j_2^4x_2^3. \tag{37}$$

In conclusion, the Hamiltonian system (8) is nonintegrable when $N=2$.

We turn to the case of $N>2$. We also denote by Γ_h the integral curve corresponding to $(x_1^h(t),y_1^h(t))$ in the full $2N$-dimensional phase space. The NVE of Eq. (8) along Γ_h is

$$\ddot{\eta}_l+(\omega_{l+1}^2+j_1^2j_{l+1}^2[x_1^h(t)]^2)\eta_l=0, \quad l=1,\ldots,N-1, \tag{38}$$

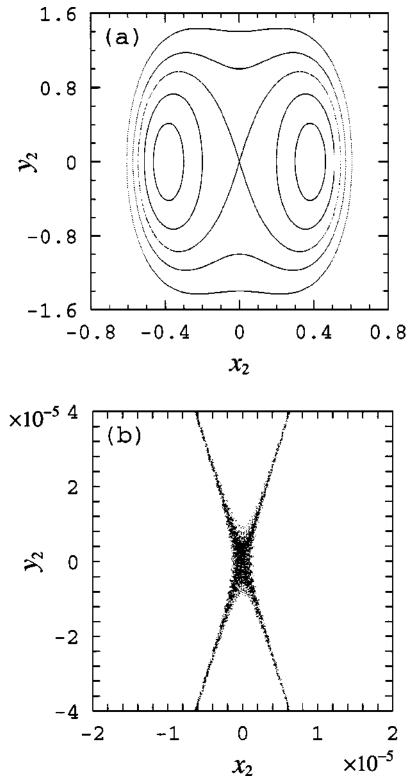


Fig. 2 Orbits of the Poincaré map of Eq. (8) with $N=2$ for $j_1=1$, $j_2=2$, $\omega_1=1$, $\omega_2=3.2$, and $H=11$. (b) is an enlargement of (a) near the origin.

which consists of $N-1$ independent Lamé equations. Hence if all the Galois groups for the Lamé equations are not abelian, then the Hamiltonian system (8) is nonintegrable. From the above discussion we see that this is the case. Thus we have proven the nonintegrability of Eq. (8) for $N>2$.

Finally we consider the infinite-degree-of-freedom system (1). Let Λ be a set of all solutions to Eq. (1) given by

$$u(t,z)=\sum_{l=1}^{\infty}a_l(t)\sin j_l\pi z, \quad \sum_{l=1}^{\infty}|a_l(t)|^2<\infty. \tag{39}$$

Note that $u(t,z)\in\Lambda$ always satisfies the boundary conditions (2), and that conversely every ''square integrable'' solution $u(t,z)$ of Eq. (1) satisfying Eq. (2) is present in Λ with some integer sequence $j_1, j_2, \ldots$, due to a fundamental result on Fourier series (e.g., Ref. [37]). Suppose that there are the necessary (i.e., infinite) number of first integrals for obtaining general solutions to Eq. (1) under the boundary condition (2). Then one can express the solution (39) by these first integrals and hence Eq. (4) by some of them. However, this is not the case because of the nonintegrability of Eq. (8) for any $N\geqslant 2$. Therefore we see that the infinite-degree-of-freedom system (1) is nonintegrable.

6 Numerical Simulations

To see whether the nonintegrability also implies the occurrence of chaotic motions in the Hamiltonian system (8), numerical simulations for $N=2$ were performed using a Fortran code called ''DOP853'' ([38]). The code is based on the explicit Runge-Kutta method of order 8 by Dormand and Prince [39], and a fifth-order error estimator with third-order correction is utilized. It also has a dense output of order 7. See Ref. [38] for more details on the method. A tolerance of 10^{-12} was chosen in the computations so that very precise results could be obtained.

Figure 2 shows orbits of the Poincaré map for $j_1=1$, $j_2=2$,

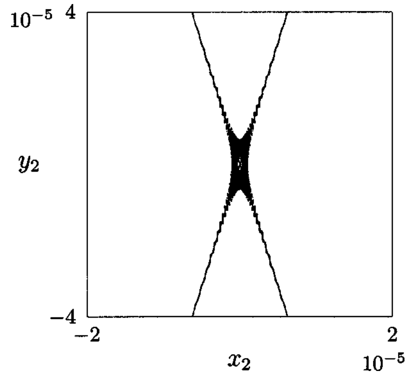


Fig. 3 Numerically computed stable and unstable manifolds of a periodic orbit for the Poincaré map of Eq. (8) with $N=2$. The same parameter values as in Fig. 2 are taken.

$\omega_1=1$, $\omega_2=3.2$, and $H=11$. Here we take as a Poincaré section the three-dimensional hyperplane $\{(x,y)\in\mathbb{R}^2\times\mathbb{R}^2 | x_1=0, y_1>0\}$. To obtain a point at which a computed trajectory intersects the Poincaré section, an interval $[t_{n-1},t_n]$ of numerical integration such that $x_1(t_{n-1})<0$ and $x_1(t_n)\geqslant 0$ was searched and the method of bisection was used for the interval with a tolerance of $|x_1|<10^{-12}$. Figure 2(b), which was obtained by performing 20,000 iterations of the Poincaré map, suggests that chaotic motions occur although they exist only in a very narrow region of the phase space.

It is easy to see that the periodic orbit $(x_1,y_1,x_2,y_2)=(x_1^h(t),y_1^h(t),0,0)$ corresponding to the origin in Fig. 2 is unstable. Actually, the NVE for the periodic orbit is

$$\ddot{\eta}_2+(\omega_2^2+j_1^2 j_2^2[x_1^h(t)]^2)\eta_2=0, \tag{40}$$

which is a special case of Hill's equation (see e.g., Ref. [10]), where $x_1^h(t)$ is a solution of

$$\ddot{x}_1+\omega_1^2 x_1+j_1^4 x_1^3=0. \tag{41}$$

In general, Hill's equation has wide parameter regions of instability, and so does Eq. (40). Such an unstable periodic orbit has stable and unstable manifolds, which are supposed to intersect transversely. Their transversal intersection is responsible for the occurrence of chaotic motions by the Smale-Birkhoff homoclinic theorem ([11,12]).

Figure 3 shows numerically computed stable and unstable manifolds of the periodic orbit for the Poincaré map of Eq. (8). To draw Fig. 3, a software called "DYNAMICS" ([40]) with the assistance of the C version of DOP853 ([38]) was used, and 5000 points on small segments with lengths of about 6×10^{-6} in the stable and unstable directions were iterated by the Poincaré map up to 95 times. Since the computation was performed with very small errors ($\sim10^{-12}$), the 95th iterate of the Poincaré map could be computed with very high accuracy. See Refs. [41], [42] for the details on the method. We see that the stable and unstable manifolds intersect transversely and the occurrence of chaos in Eq. (8) indicated by Fig. 2(b) is not an artifact of numerical errors. Thus we observe numerically that chaotic orbits exist in the two modal truncation and hence in the original beam equation (1).

7 Concluding Remarks

In this paper, we have studied the infinite-degree-of-freedom model (1) for unforced and undamped, initially straight beams and showed that it is nonintegrable in a specific sense. We first derived "exact" finite-degree-of-freedom modal truncations for the model, and proved that any $N(>1)$-modal truncations are nonintegrable in the sense that there do not exist N meromorphic first integrals which are independent and in involution. We also showed that this implies the nonintegrability of the infinite-degree-of-freedom model (1). Moreover, we provided numerical simulation results for a two-mode truncation system and observed that chaotic motions occur as in typical nonintegrable Hamiltonian systems.

Here we mention the nonintegrability of Eq. (1) when $-\Gamma\in(\pi^2,4\pi^2)$, i.e., when the beam is buckled and only the first mode is unstable. In this case, the "exact" finite-degree-of-freedom system (8) becomes

$$\dot{x}_1=y_2,\quad \dot{y}_1=x_1-\left(\sum_{r=1}^{N} j_r^2 x_r^2\right)x_1,$$
$$\dot{x}_l=y_l,\quad \dot{y}_l=-\omega_l^2 x_l-j_l^2\left(\sum_{r=1}^{N} j_r^2 x_r^2\right)x_l,\quad l=2,\ldots,N, \tag{42}$$

where we set $j_1=1$, scaled the coordinate variables as $a_l\mapsto(\pi\sqrt{-\kappa/2(\Gamma+\pi^2)})$, and replaced ω_0 by $\omega_0=\pi\sqrt{-(\Gamma+\pi^2)}$ and $\omega_l=j_l\sqrt{-((j_l\pi)^2+\Gamma)/(\Gamma+\pi^2)}$, $l=2,\ldots,N$. For $N=2$, we compute the coefficients of $P_j(\alpha)$, $j=1,2$, as

$$b_1=4+\frac{6\beta_1}{\beta_2}\omega_2^2,\quad c_1=\left(8+\frac{6\beta_1}{\beta_2}\omega_2^2\right)\omega_2^2,$$
$$d_1=\left(4+\frac{2\beta_1}{\beta_2}\omega_2^2\right)\omega_2^4 \tag{43}$$

while the other coefficients are the same as ones in Eq. (31). Repeating the arguments given in Sec. 5, we can show that the finite-degree-of-freedom system (42) with $N\geqslant2$ and the infinite-degree-of-freedom system (1) with $-\Gamma\in(\pi^2,4\pi^2)$ are nonintegrable. In particular, the two-mode truncation is nonintegrable near integral curves on the (x_1,y_1) plane including a pair of homoclinic orbits if

$$j_2^2\neq\frac{n(n+1)}{2}\quad\text{for any}\quad n\in\mathbb{N}, \tag{44}$$

which was also proven by a Melnikov-type technique of Ref. [16] in Ref. [18] to give a criterion for the existence of chaos in the system. Thus we expect that there is a close relation between nonintegrability and chaotic dynamics of Hamiltonian systems. In fact, general results for such a relation in two-degree-of-freedom systems were given in Refs. [25], [28], [29].

There has been much work on chaotic vibrations of "forced" and "damped" beams (e.g., Refs. [11], [13], [15], [22], [23]). However, in these researches, it is assumed that "unforced" and "undamped" beams are integrable, and one analyzes the forced and damped beams. So we have the following question: What novel behavior arises in forced and damped beams from the nonintegrability of unforced and undamped ones? This will be a challenging problem. A partial answer is given elsewhere ([45]).

Since our discovery of chaos we have often seen that the dynamics of simple systems are not so simple as we thought. The result given here also suggests that we should take the lesson to heart once more.

Acknowledgments

The author thanks the anonymous referees and the Associate Editor for helpful comments and suggestions, which have improved this work.

Appendix A: Description of Eq. (1) as a Hamiltonian System in the Abstract Framework

In the geometric theory of mechanics ([43,44]) we can regard Eq. (1) as an infinite-degree-of-freedom Hamiltonian system.

Let $\mathcal{X}=\mathcal{H}^1[0,1]\times\mathcal{L}^2[0,1]$, where $\mathcal{H}^1[0,1]$ is the space of functions on the interval $[0,1]$ which are square integrable along with their first derivatives, and $\mathcal{L}^2[0,1]$ is the space of square integrable functions on $[0,1]$. The inner product for $w_j=(u_j,v_j)\in\mathcal{X}$, $j=1,2$, is given by

$$\langle w_1,w_2\rangle=\int_0^1 w_1(z)\cdot w_2(z)\mathrm{d}z, \tag{A1}$$

where "$\cdot$" represents a usual inner product in $\mathbb{R}^2$. We equip the functional space $\mathcal{X}$ with a symplectic form

$$\Omega(w_1,w_2)=\int_0^1[u_1v_2-v_1u_2]\mathrm{d}z \tag{A2}$$

and a Hamiltonian function

$$H(w)=\frac{1}{2}\int_0^1[v^2+\Gamma(u')^2+(u'')^2]\mathrm{d}z+\frac{\kappa}{4}\left[\int_0^1(u')^2\mathrm{d}z\right]^2, \tag{A3}$$

where $w=(u,v)$, $w_j=(u_j,v_j)\in\mathcal{X}$, $j=1,2$. So we can write Eq. (1) as

$$\dot{w}=X_H(w), \tag{A4}$$

where $w=(u,v)\in\mathcal{X}$ and X_H represents a vector field such that $\Omega(X_H(w_1),w_2)=\langle \mathrm{D}_wH(w_1),w_2\rangle$. Here D_wH represents the *Gâteaux derivative* defined by

$$\langle \mathrm{D}_wH(w),h\rangle=\lim_{t\to 0}\frac{1}{t}[H(w+th)-H(w)] \tag{A5}$$

for any $h\in\mathcal{X}$. Note that $H(\phi_t(w))=H(w)$, i.e., the Hamiltonian energy is conserved under the flow ϕ_t, where ϕ_t is the flow generated by Eq. (A4), i.e., by Eq. (1).

Appendix B: Some Fundamental Concepts From Algebra

In this appendix we gather some necessary concepts from algebra for convenience of the reader. See, e.g., Ref. [46] for more details except some of terminologies in the last paragraph, which come from introductory concepts of linear algebraic groups.

A set S with a law of composition "$\circ$" is called a *semigroup* if the associative law holds, i.e., for any a, b, $c\in S$ one has $(a\circ b)\circ c=a\circ(b\circ c)$. A *group* G is a semigroup such that it has a *unit element* e (i.e., $a\circ e=a$ for all $a\in G$) and an *inverse element* a' for each $a\in G$ (i.e., $a\circ a'=e$). We easily see that $e\circ a=a$ and $a'\circ a=e$. The group G is said to be *commutative* or *abelian* if the commutative law also holds (i.e., $a\circ b=b\circ a$ for any a, $b\in G$). A *subgroup* H of G is a subset of G such that H itself is a group with the same law of composition. In the following we drop the symbol "$\circ$" and write the law of composition like multiplication unless we treat two or more laws of composition.

Let G be a group and let H be a subgroup of G. A *left coset* of H in G is a subset of G of type $aH=\{ax|x\in H\}$ for some $a\in G$, and a *right coset* of H in G is a subset of G of type Ha for some $a\in G$. The subgroup H is said to be *normal* if $xH=Hx$, which is often rewritten as $xHx^{-1}=H$, for all $x\in G$. If H is normal, then a left coset of H is equal to a right coset so that we do not need to distinguish between them. Suppose that H is normal and let G' be a set of cosets of H. If xH, $yH\in G'$, then $xHyH=xyH\in G'$. We easily see that G' is a group with this product as a law of composition. In particular, H is a unit element and $x^{-1}H$ is an inverse element of $xH\in G'$. We write G' as G/H and call it the *factor group* of G by H.

A sequence of subgroups of a group G,

$$G=G_0\supset G_1\supset\ldots\supset G_m, \tag{B1}$$

is called a *tower*. The tower is said to be *normal* if G_{i+1} is normal in G_i for $i=0,\ldots,m-1$. It is said to be *abelian* if it is normal and if each factor group G_i/G_{i+1} is abelian. The group G is said to be *solvable* if it has an abelian tower such that the last element is the trivial subgroup $G_m=\{e\}$. Obviously, G is solvable if it is abelian since $G\supset\{e\}$ is an abelian tower.

We now introduce some more general concepts. A set M is called a *monoid* if it has an associative law of composition and a unit element. A group G is a monoid such that every element has an inverse. A *ring* is a set R with two laws of composition called addition $(a+b)$ and multiplication (ab) if it is a commutative group with respect to the addition and a semigroup with respect to the multiplication and the distributive law holds, i.e., $a(b+c)=ab+ac$ and $(a+b)+c=ac+bc$, where a, b, $c\in R$. The unit element of the addition is also referred to as the *zero* element and denoted by 0. A ring F is said to be a *field* if the multiplication has a unit element denoted by 1 and every nonzero element has an inverse element. For example, the set $\mathbb{Q}$ of all rational numbers is a field with usual addition and multiplication.

Let M, M' be monoids. A *homomorphism* of M into M' is a mapping $f:M\to M'$ such that $f(xy)=f(x)f(y)$ for all x, $y\in M$. A homomorphism $f:M\to M'$ is called an *isomorphism* if there is a homomorphism $g:M'\to M$ such that composite mappings $f\circ g$ and $g\circ f$ are the identity mappings in M' and M, respectively. When $M=M'$, a homomorphism and isomorphism are called an *endomorphism* and *automorphism*, respectively.

Let X be a topological space. Then X is said to be *irreducible* if X cannot be written as the union of two proper, nonempty, closed subsets. A subspace Y of X is called irreducible if it is irreducible as a topological space. Every irreducible subspace of X is contained in a maximal one, which is called an *irreducible component* of X. For $\mathbb{C}^n$, where n is a positive integer, one can define a topology in which sets consisting of zeros of polynomials with n unknown variables and their intersections are closed. This topology is called the *Zariski topology*. Since the Galois group G can be represented as a group of, say $n\times n$, nonsingular matrices as stated in the last paragraph of Sec. 3, we introduce the Zariski topology in G, which is subset of $\mathbb{C}^{n^2}$. An irreducible component of G containing the unit element e, which is unique, is called the *identity component* of G. See Refs. [32–34] and references therein for more details.

References

[1] Woinowsky-Krieger, S., 1950, "The Effect of an Axial Force on the Vibration of Hinged Bars," ASME J. Appl. Mech., **17**, pp. 35–36.
[2] Burgreen, D., 1951, "Free Vibrations of a Pin-Ended Column With Constant Distance Between Pin Ends," ASME J. Appl. Mech., **18**, pp. 135–139.
[3] Eisley, J. G., 1964, "Nonlinear Vibration of Beams and Rectangular Plates," Z. Angew. Math. Phys., **15**, pp. 167–175.
[4] Srinivasan, A. V., 1966, "Nonlinear Vibrations of Beams and Plates," Int. J. Non-Linear Mech., **1**, pp. 179–191.
[5] Ray, J. D., and Bert, C. W., 1969, "Nonlinear Vibrations of a Beam With Pinned Ends," ASME J. Eng. Ind., **91**, pp. 997–1004.
[6] Bennet, J. A., and Eisley, J. G., 1970, "A Multiple Degree-of-Freedom Approach to Nonlinear Beam Vibrations," AIAA J., **8**, pp. 734–739.
[7] Tseng, W. Y., and Dugundji, J., 1970, "Nonlinear Vibrations of a Beam Under Harmonic Excitation," ASME J. Appl. Mech., **37**, pp. 292–297.
[8] Bennett, J. A., and Rinkel, R. L., 1972, "Ultraharmonic Vibrations of Nonlinear Beams," AIAA J., **10**, pp. 715–716.
[9] Bennett, J. A., 1973, "Ultraharmonic Motion of a Viscously Damped Nonlinear Beams," AIAA J., **11**, pp. 710–715.
[10] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, John Wiley and Sons, New York.
[11] Guckenheimer, J., and Holmes, P. J., 1983, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York.
[12] Wiggins, S., 1990, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, New York.
[13] Holmes, P. J., and Marsden, J. E., 1981, "A Partial Differential Equation With Infinitely Many Periodic Orbits: Chaotic Oscillations of a Forced Beam," Arch. Ration. Mech. Anal., **76**, pp. 135–165.
[14] Tseng, W. Y., and Dugundji, J., 1971, "Nonlinear Vibrations of a Buckled Beam Under Harmonic Excitation," ASME J. Appl. Mech., **38**, pp. 467–476.
[15] Moon, F. C., and Holmes, P. J., 1979, "A Magnetoelastic Strange Attractor," J. Sound Vib., **65**, pp. 275–296.
[16] Yagasaki, K., 2000, "Horseshoes in Two-Degree-of-Freedom Hamiltonian Systems With Saddle-Centers," Arch. Ration. Mech. Anal., **154**(2000), pp. 275–296.
[17] Yagasaki, K., 2003, "Homoclinic and Heteroclinic Orbits to Invariant Tori in

Multi-Degree-of-Freedom Hamiltonian Systems With Saddle-Centers,'' submitted for publication.
[18] Yagasaki, K., 2001, ''Homoclinic and Heteroclinic Behavior in an Infinite-Degree-of-Freedom Hamiltonian System: Chaotic Free Vibrations of an Undamped, Buckled Beam,'' Phys. Lett. A, **285**, pp. 55–62.
[19] Arnold, V. I., 1964, ''Instability of Dynamical Systems With Many Degrees of Freedom,'' Sov. Math. Dokl., **5**, pp. 581–585.
[20] Lochak, P., 1999, ''Arnold Diffusion: A Compendium of Remarks and Questions,'' *Hamiltonian Systems with Three or More Degrees of Freedom*, edited by C. Simó, Kluwer, Dordrecht, pp. 168–183.
[21] Yagasaki, K., 2002, ''Numerical Evidence of Fast Diffusion in a Three-Degree-of-Freedom Hamiltonian System With a Saddle-Center,'' Phys. Lett. A, **301**, pp. 45–52.
[22] Yagasaki, K., 1992, ''Chaotic Dynamics of a Quasiperiodically Forced Beam,'' ASME J. Appl. Mech., **59**, pp. 161–167.
[23] Yagasaki, K., 1995, ''Bifurcations and Chaos in a Quasi-Periodically Forced Beam: Theory, Simulation, and Experiment,'' J. Sound Vib., **183**, pp. 1–31.
[24] Morales-Ruiz, J. J., and Ramis, J. P., 2001, ''Galoisian Obstructions to Integrability of Hamiltonian Systems,'' Methods Appl. Anal., **8**, pp. 33–96.
[25] Morales-Ruiz, J. J., 1999, *Differential Galois Theory and Non-Integrability of Hamiltonian Systems*, Birkhäuser, Basel.
[26] Morales-Ruiz, J. J., 2001, ''Meromorphic Nonintegrability of Hamiltonian Systems,'' Rep. Math. Phys., **48**, pp. 183–194.
[27] Morales-Ruiz, J. J., and Simó, C., 1996, ''Non-Integrability Criteria for Hamiltonians in the Case of Lamé Normal Variational Equations,'' J. Diff. Eqns., **129**, pp. 111–135; Corrigendum: 1998, **144**, pp. 477–478.
[28] Morales-Ruiz, J. J., and Peris, J. M., 1999, ''On a Galoisian Approach to the Splitting of Separatrices,'' Ann. Fac. Sci. Toulouse VI. Sér., Math., **8**, pp. 125–141.
[29] Yagasaki, K., 2003, ''Galoisian Obstructions to Integrability and Melnikov Criteria for Chaos in Two-Degree-of-Freedom Hamiltonian Systems With Saddle-Centers,'' Nonlinearity, to appear.
[30] Moser, J., 1973, *Stable and Random Motions in Dynamical Systems*, Princeton University Press, Princeton, NJ.
[31] Arnold, V. I., 1989, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer-Verlag, New York.
[32] Kaplansky, I., 1976, *An Introduction to Differential Algebra*, 2nd ed., Hermann, Paris.
[33] Singer, M. F., 1989, ''An Outline of Differential Galois Theory,'' *Computer Algebra and Differential Equations*, edited by E. Tournier, Academic Press, London, pp. 3–57.
[34] Van der Put, M., and Singer, M. F., 2003, *Galois Theory of Linear Differential Equations*, Springer-Verlag, New York.
[35] Ziglin, S. L., 1982, ''Branching of Solutions and Non-Existence of First Integrals in Hamiltonian Mechanics, I,'' Funct. Anal. Appl., **16**, pp. 181–189.
[36] Morales-Ruiz, J. J., and Simó, C., 1994, ''Picard-Vessiot Theory and Ziglin's Theorem,'' J. Diff. Eqns., **104**, pp. 140–162.
[37] Whittaker, E. T., and Watson, G. N., 1927, *A Course of Modern Analysis*, Cambridge University Press, Cambridge.
[38] Hairer, E., Nørsett, S. P., and Wanner, G., 1993, *Solving Ordinary Differential Equations I*, 2nd ed., Springer-Verlag, Berlin.
[39] Dormand, J. R., and Prince, P. J., 1989, ''Practical Runge-Kutta Processes,'' SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput., **10**, pp. 977–989.
[40] Nusse, E. H., and Yorke, J. A., 1997, *Dynamics: Numerical Explorations*, 2nd ed., Springer-Verlag, New York.
[41] Yagasaki, K., 2003, ''Numerical Analysis for Local and Global Bifurcations of Periodic Orbits in Autonomous Differential Equations,'' in preparation.
[42] Yagasaki, K., 2003, HomMap: *A Package of AUTO and Dynamics Drivers for Homoclinic Bifurcation Analysis for Periodic Orbits of Maps and ODEs, Version 2.0*, Gifu University, Gifu, in preparation.
[43] Abraham, R., and Marsden, J. E., 1978, *Foundations of Mechanics*, 2nd ed., Addison-Wesley, Redwood City, CA.
[44] Marsden, J. E., and Ratiu, T., 1999, *Introduction to Mechanics and Symmetry*, 2nd ed., Springer-Verlag, New York.
[45] Yagasaki, K., 2003, ''Homoclinic and Heteroclinic Motions for Resonant Periodic Orbits in Forced, Two-Degree-of-Freedom Systems,'' in preparation.
[46] Lang, S., 1990, *Undergraduate Algebra*, 2nd ed., Springer-Verlag, New York.

L. Kong
Graduate Student

R. G. Parker
Corresponding Author, Mem.ASME
e-mail: parker.242@osu.edu

Department of Mechanical Engineering,
The Ohio State University,
206 W 18th Avenue,
Columbus, OH 43210
(614)-688-3922

Equilibrium and Belt-Pulley Vibration Coupling in Serpentine Belt Drives

Serpentine belt drives with spring-loaded tensioners are now widely used in the automotive industry. Experimental measurements show that linear system vibration coupling exists between the pulley rotations and the transverse span deflections. Former models that treat the belt as a string and neglect the belt bending stiffness cannot explain this coupling phenomenon. In this paper, a new serpentine belt system model incorporating the belt bending stiffness is established. The finite belt bending stiffness causes nontrivial transverse span equilibria, in contrast to string models with straight span equilibria. Nontrivial span equilibria cause linear span-pulley coupling, and the degree of coupling is determined by the equilibrium curvatures. A computational method based on boundary value problem solvers is developed to obtain the numerically exact solution of the nonlinear equilibrium equations. An approximate analytical solution of closed-form is also obtained for the case of small bending stiffness. Based on these solutions, the effects of design variables on the equilibrium deflections and span-pulley coupling are investigated.
[DOI: 10.1115/1.1598477]

Introduction

Serpentine belt drives with multiribbed belts of small section heights are commonly used in the automobile industry to drive accessories such as the alternator, air condition, power steering pump, and so on. A prominent characteristic of such systems is the introduction of a spring-loaded tensioner assembly, which greatly improves the system dynamic performance by automatically compensating for tension changes as accessories are activated or deactivated.

Considerable research has been done on the vibration of serpentine belt systems since their wide application about two decades ago. Most studies consider only the pulley rotational motions with the belt acting only as a longitudinal stiffness [1–4]. In contrast, Beikmann et al. [5–7] treated the belt as a continuum string and developed a prototypical model consisting of three pulleys and a tensioner. This model captures linear coupling only between the tensioner rotation and the transverse vibration of the two spans adjacent to the tensioner, see Fig. 1. This coupling results from tensioner rotation moving the boundary points of the two adjacent spans. For spans away from the tensioner bounded by fixed pulleys, the belt is modeled as an axially moving string whose boundary points have no transverse deflections, thereby decoupling the span vibrations from pulley rotations. Parker [8] used similar modeling to analyze vibration of a general n-pulley system. Beikmann's experiments [6,7], however, show a degree of linear coupling between pulley rotations and the span transverse vibrations between fixed pulleys, which can be excited despite the fact that the motions of the two end pulleys are only along the axis of the span treated as a string. Observations on current automobiles also demonstrate strong, apparently linear coupling where pulley rotations excite undesirable transverse vibration of the adjacent spans.

In this paper, general equations of motion that incorporate belt bending stiffness are formulated using Hamilton's principle for a three-pulley system. The equations show that vibration coupling exists between spans away from the tensioner and the pulley rotations. The mechanism of the coupling depends upon the equilibrium curvature of the spans. Correspondingly, a coupling indicator $\Gamma = \Sigma \Gamma_i$ determined by the equilibrium state is defined to measure the magnitude of vibration coupling of the whole system, where Γ_i denotes the coupling indicator for individual spans.

The main work of this paper is to determine the span equilibrium deflections from the set of nonlinear equations. The equilibrium equations are differential-integral equations coupled with several algebraic equations. A key step is the application of ordinary differential equation (ODE) conversion techniques to reformulate the governing equations into a standard boundary value problem (BVP) form that can be readily accepted by generalpurpose BVP solver codes [9]. Taking advantage of the reliability and high performance of modern BVP solver codes, almost exact numerical results are found with little programming effort. For the practically important case of small dimensionless bending stiffness, singular perturbation techniques are used to derive an approximate closed-form solution.

Based on these numerical and analytical solutions, the effects of major design variables on the equilibrium deflections and coupling indicators are investigated. The analytical solution explicitly reveals the qualitative and quantitative impact of design variables

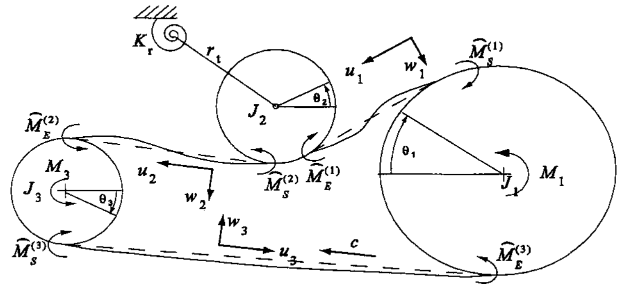


Fig. 1 A prototypical three-pulley serpentine belt drive system

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 17, 2002; final revision, Sept. 27, 2002. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

on the span equilibria and magnitude of the vibration coupling. A closed-form approximation for the coupling indicator captures this information in a simple expression.

The equilibrium solution presented here is essential for subsequent dynamic analysis of coupled belt-pulley serpentine drive vibration [10].

System Model

The prototypical system of Fig. 1 was first used by Beikmann et al. [5–7] to study the vibration of serpentine belt systems. It contains the essential components in automotive serpentine drives: a driving pulley (pulley 1), a driven pulley (pulley 3), a belt, and a spring-loaded tensioner assembly. The tensioner assembly consists of a tensioner arm spring-loaded at its pivot with an idler pulley (pulley 2) in contact with the belt. The accessory driving torques $M_1(t)$ and $M_3(t)$ are present, but $M_2(t)=0$ for the idler pulley. (See Nomenclature for definitions of the symbols.)

The dynamic motions are the pulley rotations $\theta_i(t)$, $i=1,2,3$, the tensioner arm rotation $\theta_t(t)$, and the transverse $(w_i(x_i,t))$ and longitudinal $(u_i(x_i,t))$ displacements of each belt span. The spans are modeled as continua translating with constant speed c. Each span is subjected to constant moments at its ends arising from the bending of the belt around the pulleys. No microslip or gross slip is considered at the belt-pulley interface, which is taken to be a single contact point that does not vary with belt motion [11,12].

All motions are measured relative to a reference state. The reference state corresponds to the equilibrium for a stationary belt with no bending stiffness, that is, the equilibrium for the system with the belt modeled as a string. Steady accessory torques are present in the reference state, so different spans may have different reference state tensions. Beikmann et al. [6] presented a method to calculate the reference equilibrium. Measuring displacements from this reference state clearly shows the effect of belt bending stiffness.

Consider the reference equilibrium when all accessory torques are zero. In this case, all span tensions are P_0, which is used in the subsequent nondimensionalization procedure.

The equations of motions are derived from Hamilton's principle. The kinetic energy T is

$$T=\frac{1}{2}\sum_{i=1}^{3} J_i\left(\frac{c}{r_i}+\dot{\theta}_i\right)^2+\frac{1}{2}J_t\dot{\theta}_t^2+\sum_{i=1}^{3}\int_0^{l_i}\frac{1}{2}m[(w_{i,t}-cw_{i,x})^2+(u_{i,t}-cu_{i,x}-c)^2]dx_i, \tag{1}$$

where $J_t=J_{arm}+m_2r_t^2$, J_{arm} is the rotational inertia of the tensioner arm, m_2 is the mass of the tensioner pulley, and θ_2 is the *absolute* rotation angle of the tensioner idler pulley (as opposed to rotation relative to the tensioner arm). The potential energy V is

$$V=\frac{1}{2}k_r(\theta_t+\theta_{tr})^2+\sum_{i=1}^{3}\int_0^{l_i}\left[\frac{EA}{2}\left(\frac{P_i}{EA}+u_{i,x}+\frac{1}{2}w_{i,x}^2\right)^2+\frac{1}{2}EIw_{i,xx}^2\right]dx_i, \tag{2}$$

where θ_{tr} is the rotation angle of the tensioner arm in the reference state (Fig. 2). The virtual work is

$$\delta W=\sum_{i=1}^{3}\hat{M}_S^{(i)}\delta w_{i,x}(0,t)+\sum_{i=1}^{3}\hat{M}_E^{(i)}\delta w_{i,x}(l_i,t)-\sum_{i=1}^{3}M_i\delta\theta_i, \tag{3}$$

where $\hat{M}_S^{(i)}$ is the end moment on the ith span start point and $\hat{M}_E^{(i)}$ is the end moment on the ith span end point. Beam theory applied to the belt at the pulley contact point requires

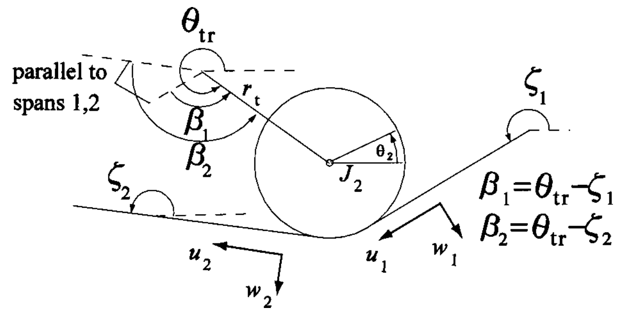


Fig. 2 Detail of tensioner region and pulley 2, defining alignment angles

$$\hat{M}_S^{(i)}=EI/r_S,\quad \hat{M}_E^{(i)}=EI/r_E, \tag{4}$$

where r_S and r_E are the radii of the pulleys that bound a span [11,12].

The kinematic constraints are obtained from Figs. 1 and 2,

$$w_1(0,t)=0,\quad w_1(l_1,t)=r_t\theta_t\cos\beta_1, \tag{5}$$

$$u_1(0,t)=-r_1\theta_1,\quad u_1(l_1,t)=-r_2\theta_2-r_t\theta_t\sin\beta_1, \tag{6}$$

$$w_2(0,t)=r_t\theta_t\cos\beta_2,\quad w_2(l_2,t)=0, \tag{7}$$

$$u_2(0,t)=-r_2\theta_2-r_t\theta_t\sin\beta_2,\quad u_2(l_2,t)=-r_3\theta_3, \tag{8}$$

$$w_3(0,t)=0,\quad w_3(l_3,t)=0, \tag{9}$$

$$u_3(0,t)=-r_3\theta_3,\quad u_3(l_3,t)=-r_1\theta_1, \tag{10}$$

where $\beta_{1,2}=\theta_{tr}-\zeta_{1,2}$ are the orientation angles of the tensioner arm relative to the two spans adjacent to the tensioner. For the details of the tensioner alignment, see Fig. 2. The positive direction of w_i is always to the inside of the belt loop, the positive direction of u_i is counterclockwise, and the positive direction of belt travel c is clockwise; θ_{tr}, θ_t are positive counterclockwise; θ_i is positive if its rotation is in the direction of belt travel; ζ_i is the angle from due east to the *outside* of the ith span; the pulleys and spans are numbered sequentially in the counterclockwise direction (Figs. 1 and 2). The definition of the tensioner orientation using $\beta_{1,2}$ differs from prior research [5–7,13]; this new definition seems easier to understand, although the two approaches are equivalent.

Application of Hamilton's principle yields the equations of motion. The field equations for the spans are

$$m(w_{i,tt}-2cw_{i,xt}+c^2w_{i,xx})-\left\{\left[EA\left(u_{i,x}+\frac{1}{2}w_{i,x}^2\right)+P_i\right]w_{i,x}\right\}_{,x}+EIw_{i,xxxx}=0,\quad i=1,2,3, \tag{11}$$

$$m(u_{i,tt}-2cu_{i,xt}+c^2u_{i,xx})-\left[EA\left(u_{i,x}+\frac{1}{2}w_{i,x}^2\right)+P_i\right]_{,x}=0,\quad i=1,2,3. \tag{12}$$

For practical serpentine drives, the longitudinal stiffness EA is much greater than the transverse stiffness from belt tension and bending. As a consequence, longitudinal waves propagate much more rapidly than transverse waves, and one may adopt the *quasistatic assumption* [14]. Under this assumption, the inertia terms are neglected in the field equations for longitudinal motion, and the dynamic tension $EA(u_{i,x}+1/2w_{i,x}^2)$ becomes uniform throughout the span,

$$\tilde{P}_i(t)=EA\left(u_{i,x}+\frac{1}{2}w_{i,x}^2\right)=\frac{EA}{l_i}\left[u_i(l_i,t)-u_i(0,t)+\int_0^{l_i}\frac{1}{2}w_{i,x}^2(x_i,t)dx_i\right],\quad i=1,2,3. \tag{13}$$

The transverse vibration equations become

$$m(w_{i,tt}-2cw_{i,xt}+c^2w_{i,xx})-[(P_i+\tilde{P}_i)w_{i,x}]_{,x}$$
$$+EIw_{i,xxxx}=0,\quad i=1,2,3, \tag{14}$$

$$w_1(0,t)=0,\quad EIw_{1,xx}(0,t)=\frac{EI}{r_1},$$

$$w_1(l_1,t)=r_t\theta_t\cos\beta_1,\quad EIw_{1,xx}(l_1,t)=-\frac{EI}{r_2}, \tag{15}$$

$$w_2(0,t)=r_t\theta_t\cos\beta_2,\quad EIw_{2,xx}(0,t)=-\frac{EI}{r_2},$$

$$w_2(l_2,t)=0,\quad EIw_{2,xx}(l_2,t)=\frac{EI}{r_3}, \tag{16}$$

$$w_3(0,t)=0,\quad EIw_{3,xx}(0,t)=\frac{EI}{r_3},\quad w_3(l_3,t)=0,$$

$$EIw_{3,xx}(l_3,t)=\frac{EI}{r_1}. \tag{17}$$

The equations for the tensioner and pulleys are

$$J_t\ddot{\theta}_t+k_r\theta_t+[mcw_{1,t}(l_1)+(P_1-mc^2+\tilde{P}_1)w_{1,x}(l_1)$$
$$+EIw_{1,xxx}(l_1)]r_t\cos\beta_1+(mc^2-\tilde{P}_1)r_t\sin\beta_1$$
$$-[mcw_{2,t}(0)+(P_2-mc^2+\tilde{P}_2)w_{2,x}(0)$$
$$+EIw_{2,xxx}(0)]r_t\cos\beta_2-(mc^2-\tilde{P}_2)r_t\sin\beta_2=0, \tag{18}$$

$$J_1\ddot{\theta}_1+\tilde{P}_1r_1-\tilde{P}_3r_1=0, \tag{19}$$

$$J_2\ddot{\theta}_2-\tilde{P}_1r_2+\tilde{P}_2r_2=0, \tag{20}$$

$$J_3\ddot{\theta}_3-\tilde{P}_2r_3+\tilde{P}_3r_3=0, \tag{21}$$

where

$$\tilde{P}_1=\frac{EA}{l_1}\left(r_2\theta_2-r_1\theta_1-r_t\theta_t\sin\beta_1+\int_0^{l_1}\frac{1}{2}w_{1,x}^2dx\right), \tag{22}$$

$$\tilde{P}_2=\frac{EA}{l_2}\left(r_3\theta_3-r_2\theta_2+r_t\theta_t\sin\beta_2+\int_0^{l_2}\frac{1}{2}w_{2,x}^2dx\right), \tag{23}$$

$$\tilde{P}_3=\frac{EA}{l_3}\left(r_1\theta_1-r_3\theta_3+\int_0^{l_3}\frac{1}{2}w_{3,x}^2dx\right). \tag{24}$$

Note that the reference state equilibrium equations $-P_1r_t\sin\beta_1+P_2r_t\sin\beta_2+k_r\theta_{tr}=0$, $P_1r_1-P_3r_1+M_1=0$, $-P_1r_2+P_2r_2=0$, and $-P_2r_3+P_3r_3+M_3=0$ have been used to simplify Eqs. (18)–(21).

The following nondimensional variables are defined:

$$\hat{x}_i=\frac{x_i}{l_i},\quad \hat{w}_i=\frac{w_i}{l_i},\quad l=\frac{l_1+l_2+l_3}{3},\quad \hat{t}=t\sqrt{\frac{P_0}{ml^2}},$$
$$\hat{P}_i=\frac{P_i}{P_0},\quad \hat{T}_i=\frac{\tilde{P}_i}{P_0}, \tag{25}$$
$$\varepsilon^2=\frac{EI}{P_0l^2},\quad s=c\sqrt{\frac{m}{P_0}},\quad k_s=\frac{k_r}{P_0r_t},\quad \gamma=\frac{EA}{P_0},$$

where P_0 is the initial tension of the string model at rest with no accessory torques, as mentioned previously. l is the characteristic length taken as the average span length.

Eliminating time derivative terms and dropping the hat on dimensionless variables yields the nondimensional equilibrium equations

$$\varepsilon^2\left(\frac{l}{l_i}\right)^2w_{i,xxxx}-[P_i-s^2+T_i]w_{i,xx}=0,\quad 0<x<1,\quad i=1,2,3, \tag{26}$$

$$w_1(0)=0,\quad w_{1,xx}(0)=\frac{l_1}{r_1},$$

$$w_1(1)=\frac{r_t}{l_1}\cos\beta_1\theta_t,\quad w_{1,xx}(1)=-\frac{l_1}{r_2}, \tag{27}$$

$$w_2(0)=\frac{r_t}{l_2}\cos\beta_2\theta_t,\quad w_{2,xx}(0)=-\frac{l_2}{r_2},$$

$$w_2(1)=0,\quad w_{2,xx}(1)=\frac{l_2}{r_3}, \tag{28}$$

$$w_3(0)=0,\quad w_{3,xx}(0)=\frac{l_3}{r_3},\quad w_3(1)=0,\quad w_{3,xx}(1)=\frac{l_3}{r_1}, \tag{29}$$

$$\left[-(P_1-s^2+T_1)w_{1,x}(1)+\varepsilon^2\left(\frac{l}{l_1}\right)^2w_{1,xxx}(1)\right]\cos\beta_1,$$
$$-\left[-(P_2-s^2+T_2)w_{2,x}(0)+\varepsilon^2\left(\frac{l}{l_2}\right)^2w_{2,xxx}(0)\right]\cos\beta_2,$$
$$+(T_1-s^2)\sin\beta_1-(T_2-s^2)\sin\beta_2-k_s\theta_t=0, \tag{30}$$

$$T_1-T_3=0, \tag{31}$$

$$-T_1+T_2=0, \tag{32}$$

$$-T_2+T_3=0, \tag{33}$$

with the dimensionless tensions

$$T_1=\frac{\tilde{P}_1}{P_0}=\gamma\left(-\frac{r_2}{l_1}\theta_2+\frac{r_t}{l_1}\theta_1-\frac{r_1}{l_1}\theta_t\sin\beta_1+\int_0^1\frac{1}{2}w_{1,x}^2dx\right), \tag{34}$$

$$T_2=\frac{\tilde{P}_2}{P_0}=\gamma\left(-\frac{r_3}{l_2}\theta_3+\frac{r_2}{l_2}\theta_2+\frac{r_t}{l_2}\theta_t\sin\beta_2+\int_0^1\frac{1}{2}w_{2,x}^2dx\right), \tag{35}$$

$$T_3=\frac{\tilde{P}_3}{P_0}=\gamma\left(-\frac{r_1}{l_3}\theta_1+\frac{r_3}{l_3}\theta_3+\int_0^1\frac{1}{2}w_{3,x}^2dx\right). \tag{36}$$

Equations (26)–(36) are a mixed differential-integral-algebraic system. The unknowns are θ_t, w_i, and θ_i. The design variables that control the equilibrium span deflections are the initial span tensions P_i, bending stiffness ε^2, speed s, tensioner spring stiffness k_s, longitudinal belt stiffness γ, and drive geometry.

From the pulley Eqs. (31)–(33),

$$T_1=T_2=T_3=T. \tag{37}$$

From Eqs. (34)–(36),

$$\frac{T_1}{\gamma}\frac{l_1}{l}=-\frac{r_2}{l}\theta_2+\frac{r_1}{l}\theta_1-\frac{r_t}{l}\theta_t\sin\beta_1+\frac{l_1}{l}\int_0^1\frac{1}{2}w_{1,x}^2dx, \tag{38}$$

$$\frac{T_2}{\gamma}\frac{l_2}{l}=-\frac{r_3}{l}\theta_3+\frac{r_2}{l}\theta_2+\frac{r_t}{l}\theta_t\sin\beta_2+\frac{l_2}{l}\int_0^1\frac{1}{2}w_{2,x}^2dx, \tag{39}$$

$$\frac{T_3}{\gamma}\frac{l_3}{l}=-\frac{r_1}{l}\theta_1+\frac{r_3}{l}\theta_3+\frac{l_3}{l}\int_0^1\frac{1}{2}w_{3,x}^2dx. \tag{40}$$

Addition of Eqs. (38), (39), and (40) and substitution of the relation (37) yield

$$\left(\frac{l_1+l_2+l_3}{l}\right)\frac{T}{\gamma}-\frac{l_1}{l}\int_0^1\frac{1}{2}w_{1,x}^2dx-\frac{l_2}{l}\int_0^1\frac{1}{2}w_{2,x}^2dx$$

$$-\frac{l_3}{l}\int_0^1\frac{1}{2}w_{3,x}^2dx+\frac{r_t}{l}(\sin\beta_1-\sin\beta_2)\theta_t=0. \tag{41}$$

Defining T as the new unknown variable and substituting Eq. (37) into Eqs. (26) and (30) gives

$$\varepsilon^2\left(\frac{l}{l_i}\right)^2 w_{1,xxxx}-(P_i-s^2+T)w_{1,xx}=0,\quad 0<x<1,\quad i=1,2,3, \tag{42}$$

$$\left[-(P_1-s^2+T)w_{1,x}(1)+\varepsilon^2\left(\frac{l}{l_1}\right)^2 w_{1,xxx}(1)\right]\cos\beta_1$$
$$-\left[-(P_2-s^2+T)w_{2,x}(0)+\varepsilon^2\left(\frac{l}{l_2}\right)^2 w_{2,xxx}(0)\right]\cos\beta_2 \tag{43}$$
$$+(T-s^2)\sin\beta_1-(T-s^2)\sin\beta_2-k_s\theta_t=0.$$

Equations (41)–(43) and boundary conditions (27)–(29) define a simplified nondimensional system that is equivalent to the original system (26)–(36). Compared with the original equilibrium equations, the variables θ_t and T replace θ_t and θ_i, $i=1,2,3$.

Numerical Solution

The above system consists of boundary value problem (BVP) Eqs. (42) coupled with algebraic Eqs. (41) and (43). The boundary conditions (27) and (28) for the two spans adjacent to the tensioner w_1 and w_2 are nontrivial and coupled with the tensioner rotation θ_t. The algebraic Eq. (41) contains integral terms involving the w_i. Furthermore, all equations are nonlinear. The combination of these characteristics initially makes it seem difficult to formulate an accurate numerical solution. By applying ODE conversion techniques, however, the above system can be transformed into a standard form defined on the interval [0,1] [9]. This formulation can be accepted by general-purpose BVP solvers yielding convenient and highly accurate solutions with minimal programming.

The standard form required for most BVP solvers is

$$y'(x)=f(x,y(x)),\quad a<x<b, \tag{44}$$

$$g(y(a),y(b))=0, \tag{45}$$

where f, y, and g are n-dimensional vectors and f and g may be nonlinear [9]. This standard form cannot contain integral terms or algebraic equations as are present in the current system.

To adapt the belt drive equilibrium equations to standard form, the following three conversion techniques are used:

- Define the constants T and θ_t as functions $T=T(x)$, $\theta_t=\theta_t(x)$ governed by
$$\frac{dT(x)}{dx}=0,\quad \frac{d\theta_t(x)}{dx}=0,\quad 0<x<1. \tag{46}$$
- For the integral terms in the algebraic equation, define $I_i(x)=\int_0^x 1/2w_{i,\sigma}^2 d\sigma$, which gives
$$\frac{dI_i(x)}{dx}=\frac{1}{2}w_{i,x}^2,\quad I_i(0)=0,\quad i=1,2,3. \tag{47}$$
$I_i(1)$ is then equivalent to the original integral terms $\int_0^1 1/2w_{i,x}^2dx$ in Eq. (41).
- With T and θ_t defined as functions of x as in Eq. (46) and the definition of $I_i(x)$ in Eq. (47), the algebraic Eqs. (41) and (43) are treated as boundary conditions as seen in Eqs. (54) and (55) below. This conveniently draws the discrete variable Eqs. (41) and (43) into the continuum BVP formulation.

This process yields the following differential equations:

$$T_{,x}=0,\quad \theta_{t,x}=0,\quad 0<x<1, \tag{48}$$

$$w_{i,xxxx}-\frac{1}{\varepsilon^2}\left(\frac{l_i}{l}\right)^2(P_i-s^2+T)w_{i,xx}=0,\quad I_{i,x}=\frac{1}{2}w_{i,x}^2,$$
$$i=1,2,3,\quad 0<x<1, \tag{49}$$

with boundary conditions

$$I_1(0)=0,\quad I_2(0)=0,\quad I_3(0)=0, \tag{50}$$

$$w_1(0)=0,\quad w_1(1)=\frac{r_t}{l_1}\cos\beta_1\theta_t(1),$$
$$w_{1,xx}(0)=\frac{l_1}{r_1},\quad w_{1,xx}(1)=-\frac{l_1}{r_2}, \tag{51}$$

$$w_2(0)=\frac{r_t}{l_2}\cos\beta_2\theta_t(1),\quad w_2(1)=0,$$
$$w_{2,xx}(0)=-\frac{l_2}{r_2},\quad w_{2,xx}(1)=\frac{l_2}{r_3}, \tag{52}$$

$$w_3(0)=0,\quad w_3(1)=0,\quad w_{3,xx}(0)=\frac{l_3}{r_3},\quad w_{3,xx}(1)=\frac{l_3}{r_1}, \tag{53}$$

$$\frac{l_1+l_2+l_3}{l}\frac{1}{\gamma}T(1)-\frac{l_1}{l}I_1(1)-\frac{l_2}{l}I_2(1)-\frac{l_3}{l}I_3(1)$$
$$+\frac{r_t}{l}(\sin\beta_1-\sin\beta_2)\theta_t(1)=0, \tag{54}$$

$$\left[-(P_1-s^2+T(1))w_{1,x}(1)+\varepsilon^2\left(\frac{l}{l_1}\right)^2 w_{1,xxx}(1)\right]\cos\beta_1$$
$$-\left[-(P_2-s^2+T(1))w_{2,x}(0)+\varepsilon^2\left(\frac{l}{l_2}\right)^2 w_{2,xxx}(0)\right]\cos\beta_2$$
$$+(T(1)-s^2)\sin\beta_1-(T(1)-s^2)\sin\beta_2-k_s\theta_t(1)=0. \tag{55}$$

Notice that the 17 boundary conditions (50)–(55) equal the total order of the eight differential Eqs. (48)–(49). Equations (48)–(55) involving higher derivatives can be readily reduced to standard first-order form (44)–(45) with the definitions $y_1(x)=w_1(x)$, $y_2(x)=w_{1,x}(x)$, $y_1'(x)=y_2(x)$, and so on. Also note that the problem is cast entirely on the interval $x\in[0,1]$ even though the problem involves multiple spans of different lengths. This standard form is readily implemented in BVP solver software. Here, the solver BVP4C in the Matlab software is adopted [15].

This approach for solving the equilibrium problem of coupled continuous-discrete systems has several advantages:

- It is easy to implement in readily available professional software once the problem is cast in standard form. This minimizes software development needs and setup time.
- Second, with the high quality and robustness of general-purpose codes, the numerical results can be excellent. For example, in this study the relative tolerance RelTol=0.001 was used for the BVP4C calculations, which is a high criterion for numerical computation [15]. Because there is no spatial discretization, the final results can be viewed as numerically exact.
- Finally, the method can be extended to other nonlinear, continuous-discrete BVP systems. By applying similar con-

Table 1 Physical properties of the prototypical system

	Pulley 1	Pulley 2	Pulley 3	Tensioner
Center	(0.5525 m, 0.0556 m)	(0.3477 m, 0.05715 m)	(0,0)	(0.2508 m, 0.0635 m)
Radius	0.0889 m	0.0452 m	0.02697 m	0.097 m
Other				
Physical properties		Belt modulus: $EA = 120{,}000$ N Span lengths: $l_1 = 0.1548$ m, $l_2 = 0.3449$ m, $l_3 = 0.5518$ m Tensioner spring stiffness: $k_r = 116.4$ N-m/rad Tensioner rotation at reference state: $\theta_{tr} = 0.1688$ rad		
Calculated values		Belt tension at reference state: $P_0 = 300$ N Tensioner alignment angles: $\beta_1 = 135.79$ deg, $\beta_2 = 178.74$ deg		

version techniques, these systems can be transformed into a standard BVP system. The algebraic equations associated with the discrete variables (e.g., tensioner rotations) typically serve as boundary conditions.

Numerical Results and Discussion

In this section, numerical equilibrium results are presented for a prototypical three-pulley serpentine belt system. The physical properties shown in Table 1 are drawn from Refs. [5–7]. Equilibrium deflections are plotted in three-dimensional perspective for span 1 adjacent to the tensoner and span 3 away from the tensioner. Tensioner rotation are evident from these plots using $w_1(1) = r_t/l_1 \cos\beta_1\theta_t$.

Figure 3 shows that nondimensional bending stiffness ε^2 strongly influences the equilibrium deflections. When bending stiffness increases, the equilibrium deflection increases markedly. When ε^2 is small, there are boundary layers in the equilibrium deflections. As ε^2 grows larger, the boundary layers become less pronounced and the equilibrium deflections become bigger and smoother. Approximating the bending stiffness of a poly-ribbed belt typical of vehicle applications by $EI = (m-1)2.867 \times 10^{-3}\, N \cdot m^2$ (where m is the number of ribs), reasonable values of ε fall in the range $0.01 \leqslant \varepsilon \leqslant 0.12$.

Figure 4 illustrates the influences of tensions P_i on the equilibrium deflections. The unequal ranges of these tensions shown in Fig. 4 result from calculating span tension variations as the steady accessory torques M_1, M_3 ($M_3 = -(r_3/r_1)M_1$) are varied across $0 \leqslant M_1 \leqslant 37.34$ N·m. Notice that the variation range of the tensions of the two spans adjacent to the tensioner $P_1 = P_2$ is much smaller than that of the third span P_3. This occurs because, when the accessory torques change, the tensioner assembly compensates for the tension loss or gain to stabilize the tensions of the two spans adjacent to it [6]. The span equilibria increase considerably with decreasing tension.

Compared with ε^2 and P_i, the speed s has smaller influence on the equilibrium deflections as seen in Fig. 5. Small equilibrium changes with speed seem contradictory to physical intuition because increasing the speed means decreasing the effective tension. This phenomenon can be explained by the mechanism of the tensioner, whose main purpose is automatic tension loss compensation to keep variations of the *tractive* tension of the belt ($P_i - s^2 + T$) small when the speed is changed [6].

The tensioner spring stiffness k_s has small influence on the equilibrium deflections over a wide range of variation (Fig. 6). Although for span 1 the equilibrium deflection appears to change significantly when k_s increases, these changes are predominantly from rotation of the whole span (which is caused by tensioner rotation and the span boundary condition $w_1(1) = (r_t/l_1)\cos\beta_1\theta_t$). Subsequent results using the coupling indicator confirm this. Figure 7 shows that the longitudinal belt stiffness γ has very small influence on the equilibrium deflections.

Linearization of the general *dynamic* equations about the equilibrium configuration shows that the degree of pulley-span vibration coupling is determined mainly by the equilibrium curvature of each span. For example, the linearized equation for span 3 is

$$\left(\frac{l_3}{l}\right)^2 w_{3,tt} - 2s\frac{l_3}{l}w_{3,xt} - \overline{P_3}w_{3,xx} + \varepsilon^2\left(\frac{l}{l_3}\right)^2 w_{3,xxxx} - \gamma\left(-\frac{r_1}{l_3}\theta_1 + \frac{r_3}{l_3}\theta_3 + \int_0^1 w_{3,x}w_{3,x}^* dx\right)w_{3,xx}^* = 0, \tag{56}$$

where $\overline{P_3} = P_3 - s^2 + P_3^*$, $P_3^* = \gamma[-r_1/l_3\theta_1^* + r_3/l_3\theta_3^* + \int_0^1 1/2(w_{3,x}^*)^2 dx]$, the asterisk denotes the equilibrium configuration, and w_3, θ_1, and θ_3 are small vibrations about equilibrium

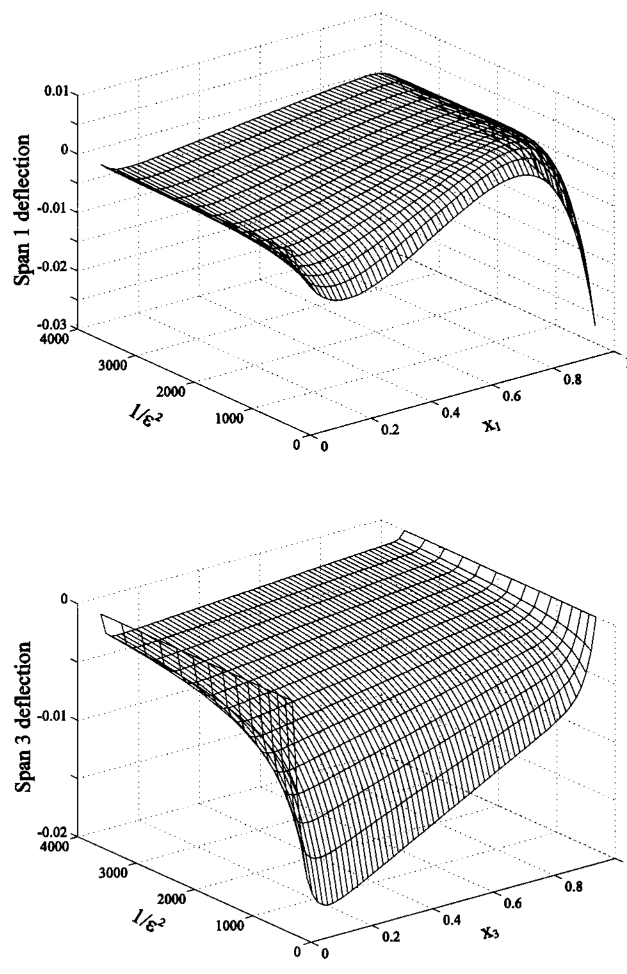


Fig. 3 Equilibrium deflections of spans 1 and 3 for varying belt bending stiffness: $s=0$, $k_s=4$, $\gamma=400$, $P_1=P_2=P_3=1$

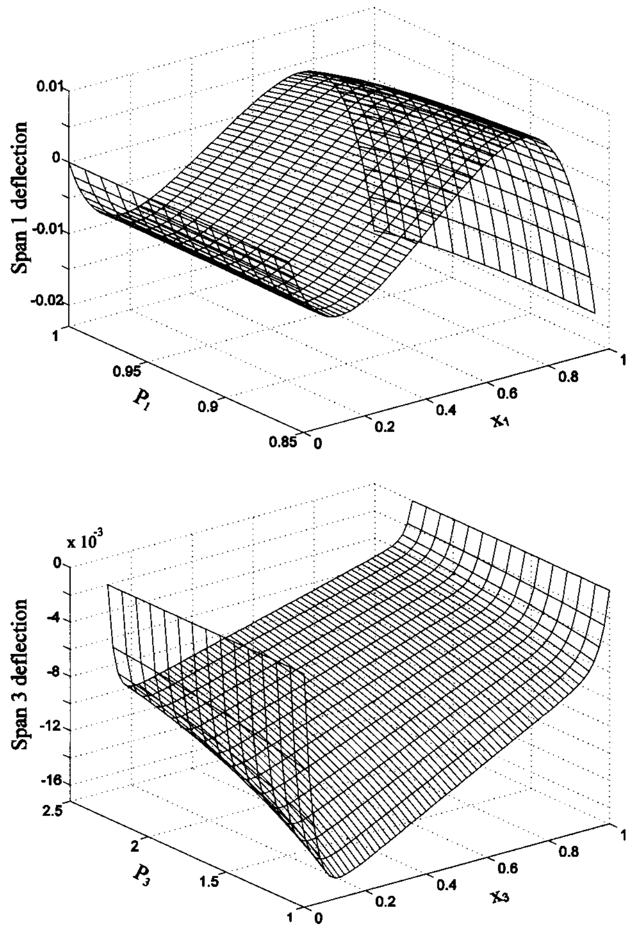


Fig. 4 Equilibrium deflections of spans 1 and 3 for varying span tensions: ε=0.05, s=0, k_s=4, γ=400

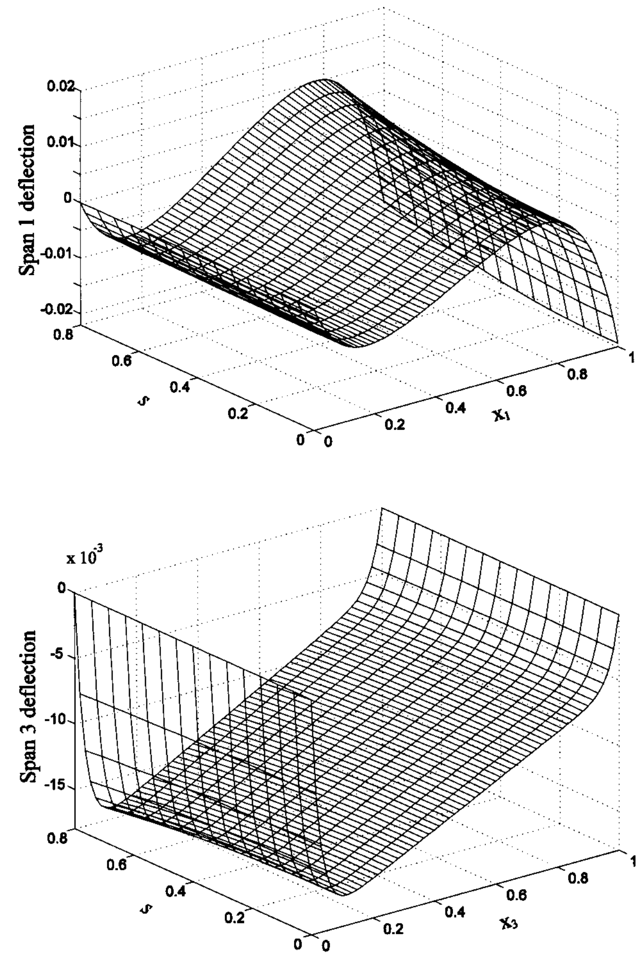


Fig. 5 Equilibrium deflections of spans 1 and 3 for varying speed: ε=0.05, k_s=4, γ=400, $P_1=P_2=P_3=1$

(for this equation only). Looking at the θ_1 and θ_3 terms coupling the span and pulley vibrations, the equilibrium curvature $w^*_{3,xx}$ governs the magnitude of coupling. Returning to the notation of $w_i(x)$ representing equilibrium deflections, one can define the equilibrium parameter $\Gamma_i=\int_0^1 w_{i,xx}^2 dx$ as the *coupling indicator* for each span and the sum $\Gamma=\Sigma_{i=1}^3\Gamma_i=\Sigma_{i=1}^3\int_0^1 w_{i,xx}^2 dx$ as the coupling indicator for the whole system. Increasing Γ indicates increasing coupling between the rotations of the pulleys and transverse motions of the belt.

Figures 8 and 9 show that belt bending stiffness ε^2 and span tensions P_i strongly affect coupling. Generally, large ε^2 and small tensions result in a more beam-like belt with relatively large equilibrium span deflections and strong pulley-span vibration coupling. In such cases, bending stiffness cannot be ignored.

Belt speed s weakly affects coupling for properly designed systems, but the effect of speed can rise if the tensioner is not properly designed. These results are shown in Figs. 10 and 11 where η=0.78 corresponds to a well designed system (as used for all other results in this paper) and η=0 is a poorly designed system. η is the tensioner effectiveness, and it indicates the ability of the tensioner to maintain constant *tractive* belt tension in response to changes in belt speed [6]. Mathematically, $\eta=\partial T/\partial(s^2)$ at $s=0$, $\varepsilon=0$ (an analytical approximation is derived subsequently). η is close to unity for well designed systems, while small η indicates poor tensioner design. Figure 11 shows the variation of tension for ranges of ε and s^2 for the values η=0.78 and η=0 (the change in η is induced by adjusting the tensioner orientation angles $\beta_{1,2}$). For the well-designed system η=0.78, the surface remains nearly planar and $T\approx\eta s^2$ over the entire region. For η=0, the surface is clearly curved except for $\varepsilon\approx0$. In this case, $T\approx\eta s^2$ is a poor approximation for $\varepsilon>0$; η is no longer effective to describe the change of tension with respect to speed changes. Physically, the reason behind this phenomenon is that when tensioner effectiveness is small, the system cannot maintain constant tractive tension through tensioner rotation as speed increases. Increased speed then decreases the tractive tension and, because of the nonzero bending stiffness, enlarges the belt span deflections (string model equilibria always have *zero* span deflections and are unaffected by decreased tension). The increased span deflections in turn cause additional tension due to the increased span lengths. Thus, when bending stiffness is considered, η cannot correctly describe the relationship between tension and speed for poorly designed systems.

The plots of Γ versus tensioner spring stiffness k_s and longitudinal belt stiffness γ are not shown because coupling indicator is a weak function of these quantities. When $\varepsilon=0.015$, $s=0$, $P_1=P_2=P_3=1$, the value of Γ varies less than 3% over the ranges of k_s and γ corresponding to that in Figs. 6 and 7.

Approximate Closed-Form Solution

When the belt bending stiffness is small, which means $\varepsilon\ll1$ (this is the usual case for practical serpentine belt systems), an approximate solution can be obtained using singular perturbation. This closed-form solution shows explicit dependence of the equilibria on system parameters. It also leads to a simple equation that reveals quantitative relations between the coupling indicator and the key design variables.

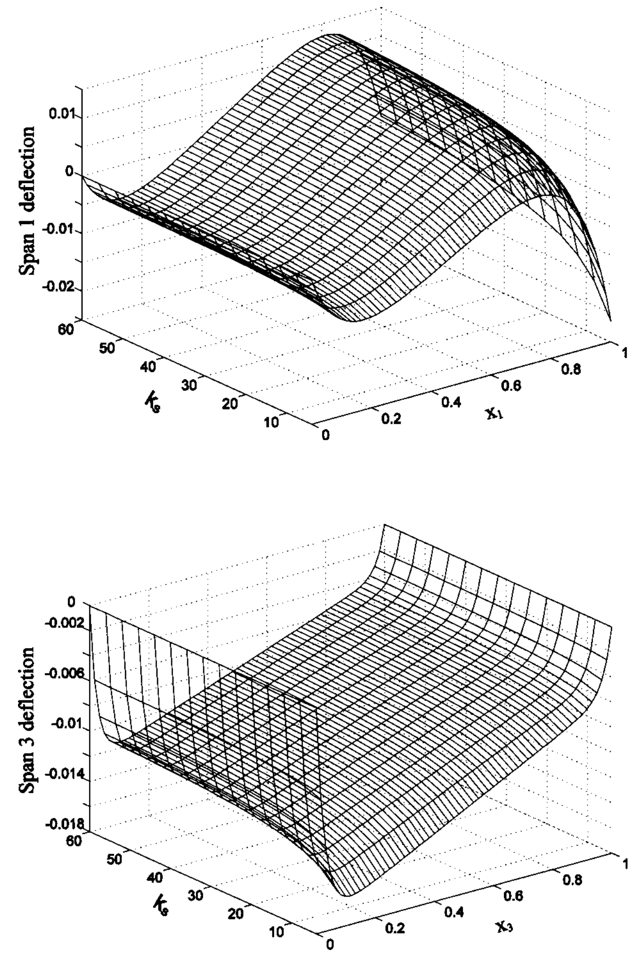


Fig. 6 Equilibrium deflections of spans 1 and 3 for varying tensioner spring stiffness: ε=0.05, s=0, γ=400, $P_1=P_2=P_3$ =1

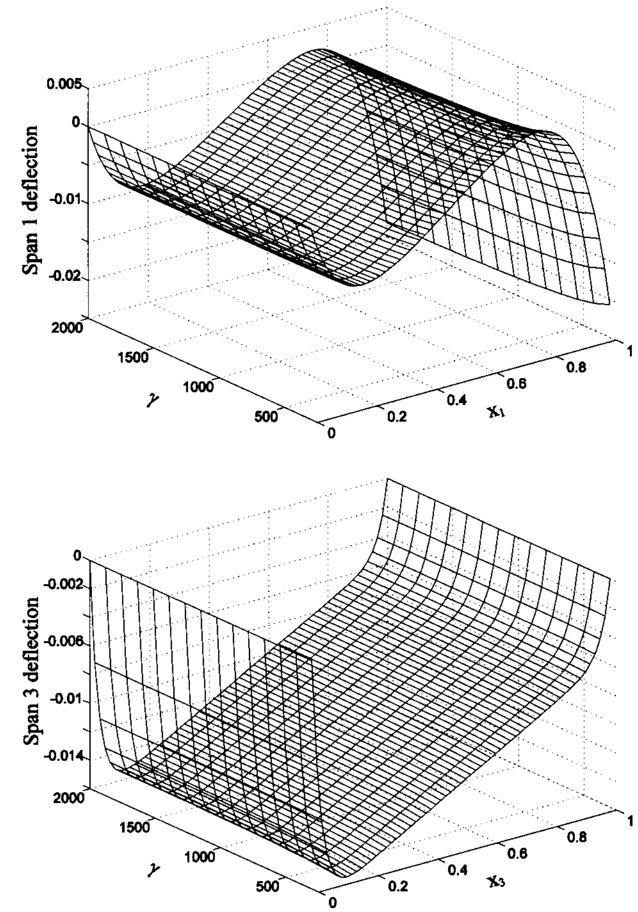


Fig. 7 Equilibrium deflections of spans 1 and 3 for varying longitudinal belt stiffness: ε=0.05, s=0, k_s=4, $P_1=P_2=P_3=1$

First, let us investigate the linearized equations of the equilibrium system (42)–(44). When nonlinear terms are neglected, one has the linear system

$$\varepsilon^2\left(\frac{l}{l_i}\right)^2 w_{i,xxxx}-(P_i-s^2)w_{i,xx}=0,\quad i=1,2,3,\quad 0<x<1, \tag{57}$$

$$\begin{aligned}&\left[-(P_1-s^2)w_{1,x}(1)+\varepsilon^2\left(\frac{l}{l_1}\right)^2 w_{1,xxx}(1)\right]\cos\beta_1\\ &\quad-\left[-(P_2-s^2)w_{2,x}(0)+\varepsilon^2\left(\frac{l}{l_2}\right)^2 w_{2,xxx}(0)\right]\cos\beta_2\\ &\quad+(T-s^2)\sin\beta_1-(T-s^2)\sin\beta_2-k_s\theta_t=0,\end{aligned} \tag{58}$$

$$\frac{l_1+l_2+l_3}{l}\frac{T}{\gamma}+\frac{r_t}{l}(\sin\beta_1-\sin\beta_2)\theta_t=0. \tag{59}$$

Boundary conditions for the w_i are still Eqs. (27)–(29). While readily solvable, the linear system approximation is unsatisfactory as seen in Figs. 12 and 13. Thus the equilibrium problem is nonlinear in essence, and the linear system does not give an effective approximation.

Returning to the nonlinear model, the variables are represented as

$$T=T^{(0)}+\varepsilon T^{(1)}+\cdots, \tag{60}$$

$$\theta_t=\theta_t^{(0)}+\varepsilon\theta_t^{(1)}+\cdots, \tag{61}$$

$$w_i=w_i^c+\cdots, \tag{62}$$

where w_i^c is the leading composite combination of the inner and outer solution of the ith span. Substitution of Eqs. (60)–(62) into Eq. (42) yields equations for the spans.

First consider span 1, which has the equation

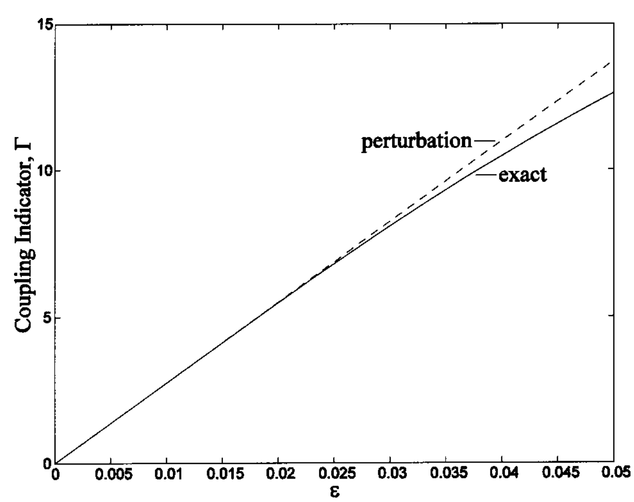


Fig. 8 System coupling indicator Γ for varying belt bending stiffness: s=0, k_s=4, γ=400, $P_1=P_2=P_3=1$

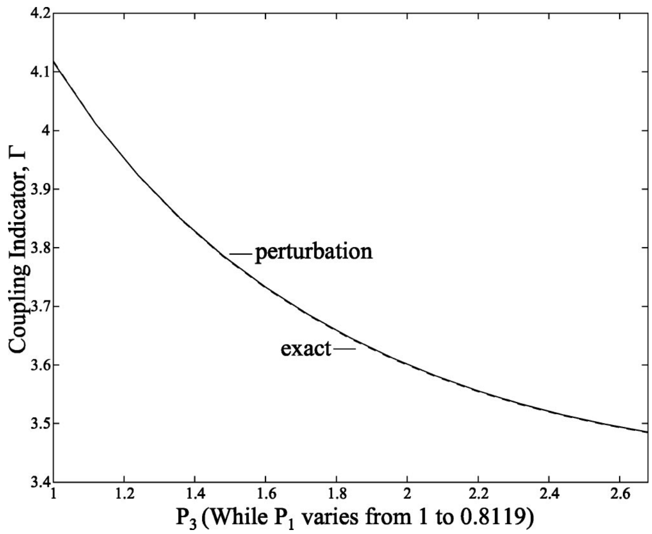


Fig. 9 System coupling indicator Γ for varying span tensions: ε=0.015, s=0, k_s=4, γ=400

$$\varepsilon^2\left(\frac{l}{l_1}\right)^2 w^c_{1,xxxx}-(P_1-s^2+T^{(0)}+\varepsilon T^{(1)})w^c_{1,xx}=0,\quad 0<x<1. \tag{63}$$

The outer expansion has the form

$$w_1^O=y_0(x)+\varepsilon y_1(x)+\varepsilon^2 y_2(x)+\cdots. \tag{64}$$

Substitution of Eq. (64) into Eq. (63) gives

$$y_{0,xx}=0, \tag{65}$$

$$(P_1-s^2+T^{(0)})y_{1,xx}=-T^{(1)}y_{0,xx}, \tag{66}$$

$$(P_1-s^2+T^{(0)})y_{2,xx}=\left(\frac{l}{l_1}\right)^2 y_{0,xxxx}-T^{(1)}y_{1,xx}. \tag{67}$$

Solution of these problems in sequence yields

$$w_1^O=D_0+D_1x+\varepsilon(D_2+D_3x)+\varepsilon^2(D_4+D_5x), \tag{68}$$

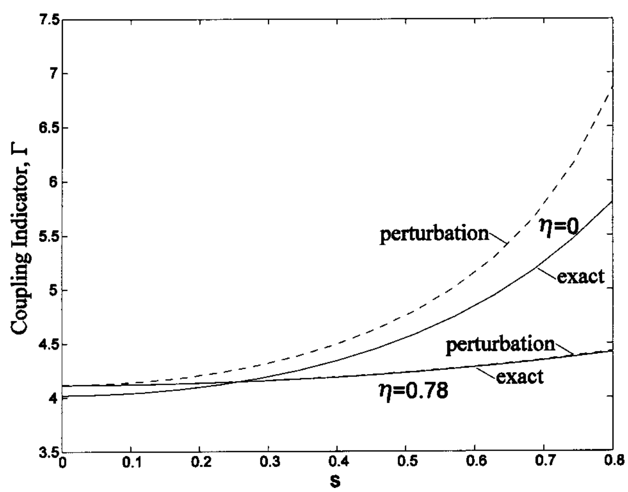


Fig. 10 System coupling indicator Γ for varying speed: ε=0.015, k_s=4, γ=400, $P_1=P_2=P_3=1$. η=0.78 corresponds to tensioner orientation β_1=135.79 deg, β_2=178.74 deg in Table 1; while η=0 corresponds to tensioner orientation β_1=68.53 deg, β_2=111.47 deg.

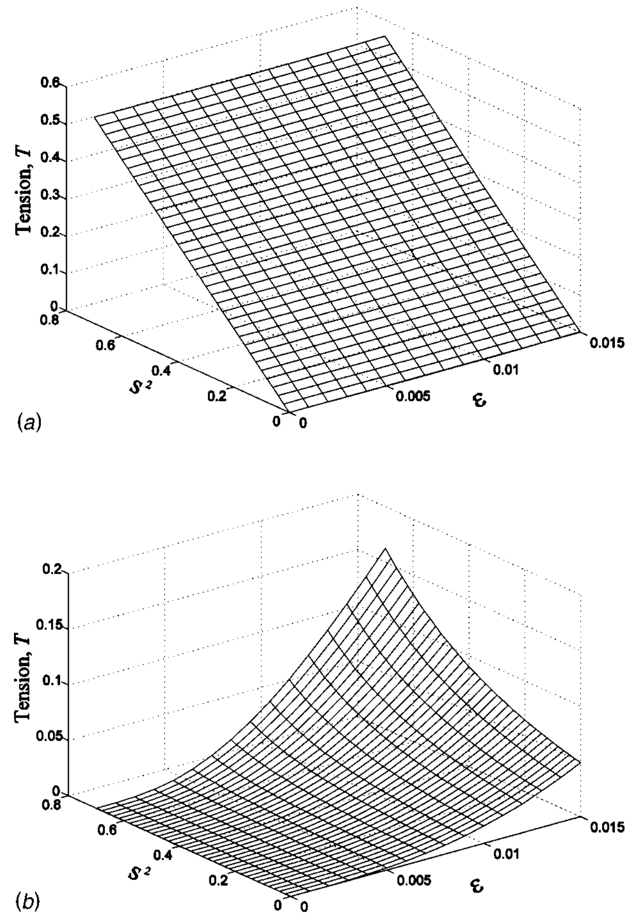


Fig. 11 Tension variation with different tensioner orientation: k_s=4, γ=400, $P_1=P_2=P_3=1$. (*a*) β_1=135.79 deg, β_2=178.74 deg, η=0.78. (*b*) β_1=68.53 deg, β_2=111.47 deg, η=0.

where $D_0\sim D_5$ are arbitrary constants. This outer expansion is not required to satisfy any boundary conditions. It must be matched with two boundary layer expansions, one valid near $x=0$ and the other valid near $x=1$.

To find the inner expansion near $x=0$, one introduces the stretching transformation

$$\xi=\frac{x}{\varepsilon^\lambda},\quad \lambda>0. \tag{69}$$

Denoting the inner expansion near $x=0$ by the superscript i, Eq. (63) becomes

$$\varepsilon^{2-4\lambda}\left(\frac{l}{l_1}\right)^2\frac{d^4w_1^i}{d\xi^4}-(P_1-s^2+T^{(0)}+\varepsilon T^{(1)})\varepsilon^{-2\lambda}\frac{d^2w_1^i}{d\xi^2}=0. \tag{70}$$

As $\varepsilon\to 0$, the distinguished limit corresponds to $\lambda=1$. Then, w_1^i is governed by

$$\left(\frac{l}{l_1}\right)^2\frac{d^4w_1^i}{d\xi^4}-(P_1-s^2+T^{(0)}+\varepsilon T^{(1)})\frac{d^2w_1^i}{d\xi^2}=0,$$
$$w_1^i(0)=0,\quad \frac{d^2w_1^i}{d\xi^2}(0)=\varepsilon^2\left(\frac{l_1}{r_1}\right). \tag{71}$$

With the inner expansion $w_1^i=W_0(\xi)+\varepsilon W_1(\xi)+\varepsilon^2W_2(\xi)$, Eq. (71) gives

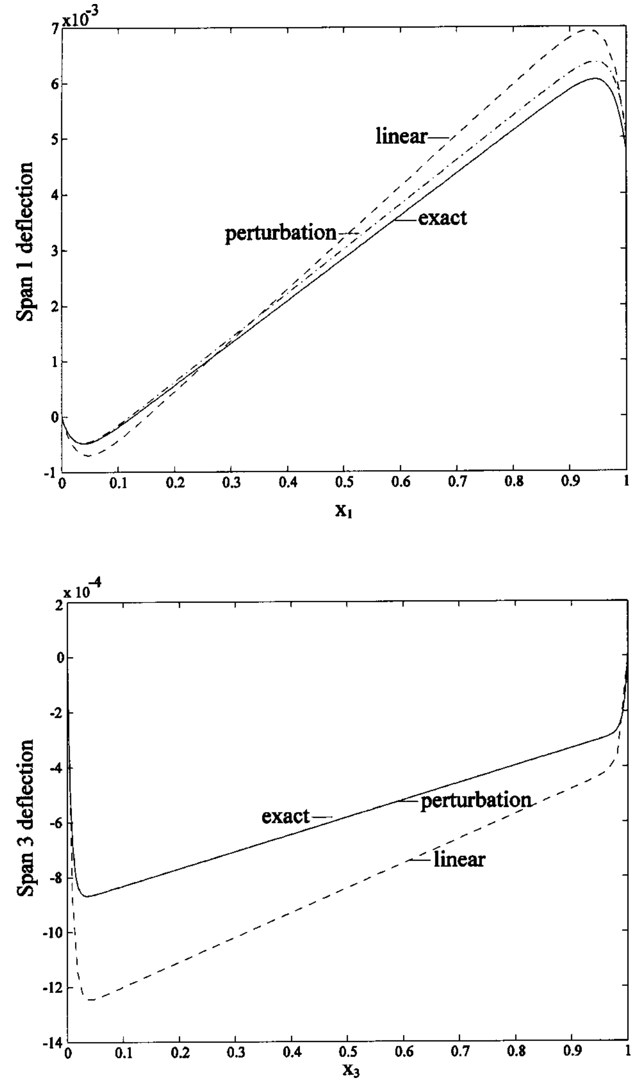


Fig. 12 Equilibrium deflections of the first and third spans: $\varepsilon=0.01$, $s=0.6$, $k_s=4$, $\gamma=400$, $P_1=P_2=P_3=1$

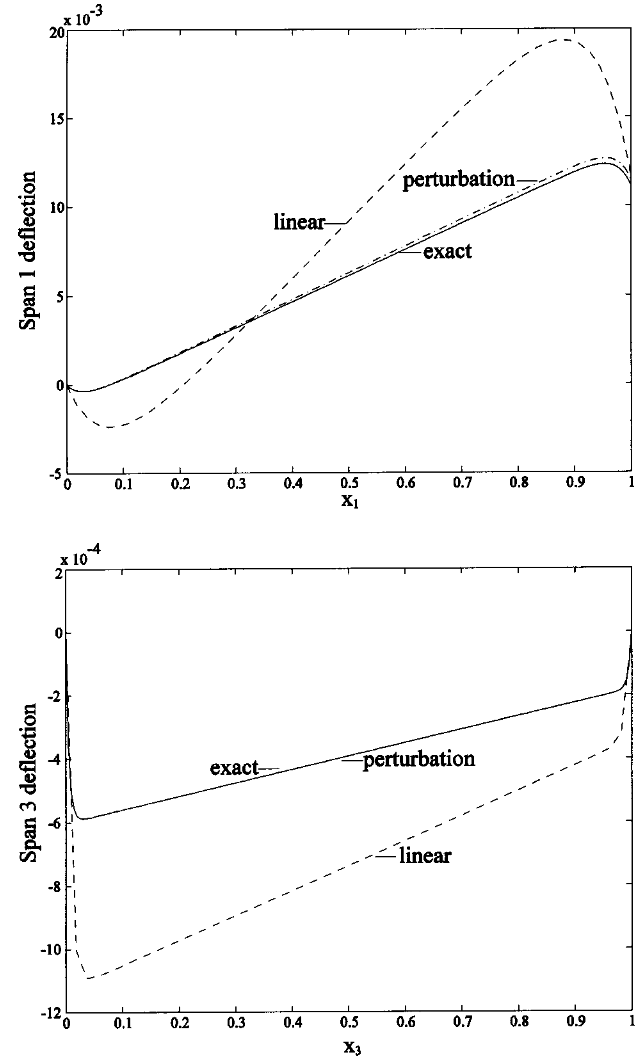


Fig. 13 Equilibrium deflections of the first and third spans: $\varepsilon=0.01$, $s=0.9$, $k_s=4$, $\gamma=400$, $P_1=P_2=0.9395$, $P_3=1.5395$

$$\left(\frac{l}{l_1}\right)^2\frac{d^4W_0}{d\xi^4}-(P_1-s^2+T^{(0)})\frac{d^2W_0}{d\xi^2}=0,$$

$$W_0(0)=0 \quad W_{0,\xi\xi}(0)=0, \tag{72}$$

$$\left(\frac{l}{l_1}\right)^2\frac{d^4W_1}{d\xi^4}-(P_1-s^2+T^{(0)})\frac{d^2W_1}{d\xi^2}=T^{(1)}\frac{d^2W_0}{d\xi^2},$$

$$W_1(0)=0 \quad W_{1,\xi\xi}(0)=0, \tag{73}$$

$$\left(\frac{l}{l_1}\right)^2\frac{d^4W_2}{d\xi^4}-(P_1-s^2+T^{(0)})\frac{d^2W_2}{d\xi^2}=T^{(1)}\frac{d^2W_1}{d\xi^2},$$

$$W_2(0)=0 \quad W_{2,\xi\xi}(0)=\frac{l_1}{r_1}. \tag{74}$$

The general solution of Eq. (72) is

$$W_0(\xi)=a_0+b_0\xi+c_0e^{-\xi(l_1/l)\sqrt{P_1-s^2+T^{(0)}}}+d_0e^{\xi(l_1/l)\sqrt{P_1-s^2+T^{(0)}}}. \tag{75}$$

The constant d_0 must be zero; otherwise, $W_0(\xi)$ would grow exponentially with ξ, making it unmatchable with the outer expansion. The boundary conditions in Eq. (72) lead to $W_0(\xi)=b_0\xi$. Subsequent solution of Eqs. (73) and (74) for W_1 and W_2 yields the inner expansion

$$w_1^i(\xi)=b_0\xi+\varepsilon b_1\xi+\varepsilon^2\left[b_2\xi-\left(\frac{l}{l_1}\right)^2\left(\frac{l_1}{r_1}\right)\frac{1}{P_1-s^2+T^{(0)}}(1-e^{-\xi(l_1/l)\sqrt{P_1-s^2+T^{(0)}}})\right], \tag{76}$$

where exponential growth terms have been eliminated.

With use of Eq. (69) and $\lambda=1$, the inner expansion of Eq. (68) is

$$(w_1^O)^i=D_0+\varepsilon(D_1\xi+D_2)+\varepsilon^2(D_3\xi+D_4). \tag{77}$$

The outer expansion of Eq. (76) as $\varepsilon\to 0$ is

$$(w_1^i)^O=b_0\xi+\varepsilon b_1\xi+\varepsilon^2\left[b_2\xi-\left(\frac{l}{l_1}\right)^2\left(\frac{l_1}{r_1}\right)\frac{1}{P_1-s^2+T^{(0)}}\right]. \tag{78}$$

The matching principle $(w_1^O)^i=(w_1^i)^O$ yields

$$D_0=0, \quad b_0=0, \quad b_1=D_1, \quad D_3=b_2, \quad D_2=0,$$

$$D_4=-\left(\frac{l}{l_1}\right)^2\left(\frac{l_1}{r_1}\right)\frac{1}{P_1-s^2+T^{(0)}}. \tag{79}$$

For the boundary layer near $x=1$, the stretching transformation is $\zeta=(1-x)/\varepsilon$. Denoting the inner expansion of Eq. (63) near $x=1$ by the superscript I, the governing equation as $\varepsilon\to 0$ is

$$\left(\frac{l}{l_1}\right)^2\frac{d^4 w_1^I}{d\zeta^4}-(P_1-s^2+T^{(0)}+\varepsilon T^{(1)})\frac{d^2 w_1^I}{d\zeta^2}=0. \tag{80}$$

Application of similar procedures as for w_1^i to w_1^I yields

$$w_1^I=\frac{r_t}{l_1}\cos\beta_1\theta_t^{(0)}+B_0\zeta+\varepsilon\left(B_1\zeta+\frac{r_t}{l_1}\cos\beta_1\theta_t^{(1)}\right)+\varepsilon^2\left[B_2\zeta+\left(\frac{l}{l_1}\right)^2\left(\frac{l_1}{r_2}\right)\frac{1}{P_1-s^2+T^{(0)}}\left(1-e^{-\zeta(l_1/l)\sqrt{P_1-s^2+T^{(0)}}}\right)\right], \tag{81}$$

where B_0, B_1, and B_2 are arbitrary constants.

The inner expansion near $x=1$ of the outer solution (68) is

$$(w_1^O)^I=D_0+D_1+\varepsilon[-D_1\zeta+D_2+D_3]+\varepsilon^2[-D_3\zeta+D_4+D_5]. \tag{82}$$

The matching condition $(w_1^O)^I=(w_1^I)^O$ leads to

$$D_0+D_1=\frac{r_t}{l_1}\cos\beta_1\theta_t^{(0)},\quad -D_1=B_1,$$
$$D_2+D_3=\frac{r_3}{l_1}\cos\beta_1\theta_t^{(1)},$$
$$-D_3=B_2,\quad B_0=0,\quad D_4+D_5=\left(\frac{l}{l_1}\right)^2\left(\frac{l_1}{r_2}\right)\frac{1}{P_1-s^2+T^{(0)}}. \tag{83}$$

Solution of Eqs. (83) and (79) gives

$$D_0=0,\quad D_1=\frac{r_t}{l_1}\cos\beta_1\theta_t^{(0)},\quad D_2=0,\quad D_3=\frac{r_t}{l_1}\cos\beta_1\theta_t^{(1)},$$
$$D_4=-\left(\frac{l}{l_1}\right)^2\left(\frac{l_1}{r_1}\right)\frac{1}{P_1-s^2+T^{(0)}},$$
$$D_5=\left(\frac{l}{l_1}\right)^2\frac{1}{P_1-s^2+T^{(0)}}\left(\frac{l_1}{r_1}+\frac{l_1}{r_2}\right),$$
$$b_0=0,\quad b_1=\frac{r_t}{l_1}\cos\beta_1\theta_t^{(0)}, \tag{84}$$
$$b_2=\frac{r_t}{l_1}\cos\beta_1\theta_t^{(1)},\quad B_0=0,\quad B_1=-\frac{r_t}{l_1}\cos\beta_1\theta_t^{(0)},$$
$$B_2=-\frac{r_t}{l_1}\cos\beta_1\theta_t^{(1)}.$$

The composite expansion for span 1 is

$$w_1\sim w_1^c=w_1^O+w_1^i+w_1^I-(w_1^i)^O-(w_1^I)^O$$
$$=\varepsilon^2\left(\frac{l}{l_1}\right)^2\frac{1}{P_1-s^2+T^{(0)}}\left[-\frac{l_1}{r_1}+\left(\frac{l_1}{r_1}+\frac{l_1}{r_2}\right)x+\frac{l_1}{r_1}e^{-(x/\varepsilon)(l_1/l)\sqrt{P_1-s^2+T^{(0)}}}-\frac{l_1}{r_2}e^{[(x-1)/\varepsilon](l_1/l)\sqrt{P_1-s^2+T^{(0)}}}\right]+\frac{r_t}{l_1}\cos\beta_1(\theta_t^{(0)}+\varepsilon\theta_t^{(1)})x. \tag{85}$$

Similar perturbation processes for w_2 and w_3 yield

$$w_2\sim w_2^c=\varepsilon^2\left(\frac{l}{l_2}\right)^2\frac{1}{P_2-s^2+T^{(0)}}\left[\frac{l_2}{r_2}+\left(-\frac{l_2}{r_2}-\frac{l_2}{r_3}\right)x-\frac{l_2}{r_2}e^{-(x/\varepsilon)(l_2/l)\sqrt{P_2-s^2+T^{(0)}}}+\frac{l_2}{r_3}e^{[(x-1)/\varepsilon](l_2/l)\sqrt{P_2-s^2+T^{(0)}}}\right]-\frac{r_t}{l_2}\cos\beta_2(\theta_t^{(0)}+\varepsilon\theta_t^{(1)})(x-1), \tag{86}$$

$$w_3\sim w_3^c=\varepsilon^2\left(\frac{l}{l_3}\right)^2\frac{1}{P_3-s^2+T^{(0)}}\left[-\frac{l_3}{r_3}+\left(\frac{l_3}{r_3}-\frac{l_3}{r_1}\right)x+\frac{l_3}{r_3}e^{-(x/\varepsilon)(l_3/l)\sqrt{P_3-s^2+T^{(0)}}}+\frac{l_3}{r_1}e^{[(x-1)/\varepsilon](l_3/l)\sqrt{P_3-s^2+T^{(0)}}}\right]. \tag{87}$$

Substitution of Eqs. (60), (61), (85), and (86) into the tensioner arm Eq. (44) gives the following conditions corresponding to the ε^0 and ε^1 terms, respectively:

$$(-s^2+T^{(0)})(\sin\beta_1-\sin\beta_2)-\left[k_s+\frac{r_t}{l_1}\cos^2\beta_1(P_1-s^2+T^{(0)})+\frac{r_t}{l_2}\cos^2\beta_2(P_2-s^2+T^{(0)})\right]\theta_t^{(0)}=0, \tag{88}$$

$$T^{(1)}(\sin\beta_1-\sin\beta_2)-\left[k_s+\frac{r_t}{l_1}\cos^2\beta_1(P_1-s^2+T^{(0)})+\frac{r_t}{l_2}\cos^2\beta_2(P_2-s^2+T^{(0)})\right]\theta_t^{(1)}-\left(\frac{r_t}{l_1}\cos^2\beta_1+\frac{r_t}{l_2}\cos^2\beta_2\right)T^{(1)}\theta_t^{(0)}=0. \tag{89}$$

Substitution of Eqs. (60), (61), and (85)–(87) into Eq. (41) yields the ε^0 and ε^1 conditions

$$\frac{1}{\gamma}\frac{l_1+l_2+l_3}{l}T^{(0)}-\frac{1}{2}\left[\frac{l_1}{l}\left(\frac{r_t}{l_1}\cos\beta_1\right)^2+\frac{l_2}{l}\left(\frac{r_t}{l_2}\cos\beta_2\right)^2\right]\times(\theta_t^{(0)})^2+\frac{r_t}{l}(\sin\beta_1-\sin\beta_2)\theta_t^{(0)}=0, \tag{90}$$

$$\frac{1}{\gamma}\frac{l_1+l_2+l_3}{l}T^{(1)}-\left[\frac{l_1}{l}\left(\frac{r_t}{l_1}\cos\beta_1\right)^2+\frac{l_2}{l}\left(\frac{r_t}{l_2}\cos\beta_2\right)^2\right]\times\theta_t^{(0)}\theta_t^{(1)}+\frac{r_t}{l}(\sin\beta_1-\sin\beta_2)\theta_t^{(1)}=0. \tag{91}$$

Generally, $T^{(0)}$ and $\theta_t^{(0)}$ must be solved numerically from the nonlinear algebraic Eqs. (88) and (90). After solving for $T^{(0)}$ and $\theta_t^{(0)}$, higher-order terms can be calculated successively, like the $T^{(1)}$ and $\theta_t^{(1)}$ are obtained by solution of Eqs. (89) and (91) (which give $T^{(1)}=\theta_t^{(1)}=0$). These values complete the leading-order composite span solutions of Eqs. (85), (86), and (87). Representative agreement between analytical and numerical solutions are shown in Figs. 8–10, 12–13, and listed in Table 2.

Note if the bending stiffness vanishes, then the analytical results reduce to those for the string model. For example, according to Eq. (87), the deflection in span 3 is zero throughout the span

Table 2 Comparison of approximate analytical and numerical solutions. Case 1 parameters are those used in Fig. 12, and Case 2 parameters are those used in Fig. 13.

	Exact T	Approximate T	Exact θ_t	Approximate θ_t
Case 1	0.2828	0.2836	−0.0106	−0.0113
Case 2	0.6408	0.6417	−0.0247	−0.0255

when ε is zero; also see Eqs. (85) and (86). When the bending stiffness approaches zero, no matter how small it is, boundary layers always exist at the ends of each span, although their thickness and height are very small, as seen in Figs. 3 and 12.

Substitution of Eqs. (85)–(87) into the definitions of system and span coupling indicator gives

$$\Gamma=\sum_{i=1}^{3}\Gamma_i\sim\sum_{i=1}^{3}\frac{1}{2}\frac{\varepsilon}{\sqrt{P_i-s^2+T^{(0)}}}\frac{l}{l_i}\left[\left(\frac{l_i}{r_{si}}\right)^2+\left(\frac{l_i}{r_{ei}}\right)^2\right], \tag{92}$$

where r_{si} and r_{ei} are the radii of the two pulleys at the ends of the ith span. Analytical and numerical solutions agree well as shown in Figs. 8, 9, and 10. For general serpentine belt systems having n spans, generalization of Eq. (92) gives

$$\Gamma\sim\frac{1}{2}\sum_{i=1}^{n}\frac{\varepsilon}{\sqrt{P_i-s^2+T^{(0)}}}\frac{l}{l_i}\left[\left(\frac{l_i}{r_{si}}\right)^2+\left(\frac{l_i}{r_{ei}}\right)^2\right]. \tag{93}$$

Note that from Eq. (93), when the bending stiffness is zero ($\varepsilon=0$), then $\Gamma=0$. This corresponds to the string model where there is no equilibrium curvature.

From Eqs. (60) and (61), the leading-order tension and tensioner rotation approximations ($T^{(0)}$ and $\theta_t^{(0)}$) are independent of the belt bending stiffness ε^2, which means that $T^{(0)}$ and $\theta_t^{(0)}$ are determined by the string model of the serpentine belt drives. For string models, former researchers [6,16] have indicated that

$$T^{(0)}\approx\eta s^2,\qquad \eta=\frac{dT}{d(s^2)}\quad\text{at } s^2=0, \tag{94}$$

where η is the tensioner effectiveness, as mentioned before. Equations (88) and (90) lead to

$$\eta=\frac{1}{\dfrac{l_1+l_2+l_3}{\gamma}\left[\dfrac{P_1\left(\dfrac{1}{l_1}\cos^2\beta_1+\dfrac{1}{l_2}\cos^2\beta_2\right)+k_s}{(\sin\beta_1-\sin\beta_2)^2}\right]+1} \tag{95}$$

which corresponds to that defined in Ref. [6].

Further comparison with numerical solutions shows that the above approximation is good for properly designed system, but the approximation becomes poor if η is away from unity (Fig. 10). In such cases, $T^{(0)}\approx\eta s^2$ is no longer the dominant part of T due to the poor speed compensation ability of the tensioner, as mentioned previously; bending stiffness significantly impacts tension. For good approximation for $\eta\approx 0$, more terms need to be incorporated in Eqs. (60) and (61) (with no change in Eq. (62) for approximation through ε^4). This leads to additional equations like Eqs. (89) and (91). Approximation through ε^4 gives excellent agreement with Fig. 11*b* where $\eta=0$.

Considering well-designed systems, substitution of Eq. (94) into Eq. (93) leads to

$$\Gamma\sim\frac{1}{2}\sum_{i=1}^{n}\frac{\varepsilon}{\sqrt{P_i-(1-\eta)s^2}}\frac{l}{l_i}\left[\left(\frac{l_i}{r_{si}}\right)^2+\left(\frac{l_i}{r_{ei}}\right)^2\right]. \tag{96}$$

The simple relation (96) confirms the coupling indicator's dependence on bending stiffness, tensions, and speed as shown by the numerical solution. The effect of speed is small due to the small coefficient $1-\eta$ preceding it; the influence of speed would rise as η gets smaller. The effects of tensioner spring stiffness k_s and longitudinal belt stiffness γ are very small because their impact is felt only through the tensioner effectiveness η.

Equations (93) and (96) reveal that the span-pulley vibration coupling is determined primarily by three factors: the system geometry (especially the ratios of the pulley radii to the span lengths), the bending stiffness, and the span tensions. Large Γ is necessary for strong vibration coupling between the spans and pulleys; reduced Γ prevents pulley rotations from generating undesirable span vibration. By tuning the system design variables, the degree of coupling Γ can be adjusted. The most effective methods to reduce this coupling are: (1) decreasing the belt bending stiffness, (2) increasing the span tensions, which are determined by the initial tensioner torque and the accessory torques exerted on the pulleys, and (3) increasing the ratio between the pulley radii and the span lengths. Higher bearing loads negatively balance the benefits of higher span tensions. The increased bearing loads occur at all pulleys, even though the troublesome coupling is typically concentrated at a single span. Increasing a pulley radius can be an effective, low cost solution to reduce a practical span vibration problem where pulley rotations drive large span vibrations. This solution is localized to address only the problem span vibration.

Summary and Conclusions

A model of a serpentine belt system including belt bending stiffness is established and an equilibrium analysis is performed. A numerically exact solution is presented to determine the span and tensioner equilibria. This requires a novel transformation of the governing equations to a standard ODE form readily accepted by general-purpose BVP solver codes. A closed-form analytical solution is also developed for the case of small bending stiffness using singular perturbation. A coupling indicator (Γ) is defined based on the equilibria to quantify the undesirable vibration coupling between pulleys and spans. The perturbation solution analytically exposes the effects of the design variables on the equilibria and coupling in terms of simple expressions. The major conclusions include:

1. Belt bending stiffness and span tensions strongly influence the equilibrium deflections and pulley-span coupling.
2. Speed has a much smaller effect on this coupling due to the automatic tension compensation ability of the tensioner assembly when properly designed.
3. The effects of tensioner spring stiffness and longitudinal belt stiffness on equilibrium are small over a wide range of variation.
4. The vibration coupling between the pulleys and spans is determined mainly by the equilibrium span curvatures. A defined *coupling indicator* captures the magnitude of coupling for each span.
5. System geometry, especially the pulley radius to span length ratios, significantly affects equilibria and coupling. This suggests that an effective, low cost solution to troubleshoot coupling problems is increasing the radii of pulleys that bound the problem span.

Acknowledgments

The authors thank Mark IV Automotive/Dayco Corporation and the National Science Foundation for support of this research.

Nomenclature

EA = longitudinal belt stiffness modulus
EI = bending stiffness
J_i = rotation inertia of pulley i
M_i = applied static torque on accessory i
P_i = belt tension of span i in the reference state

P_0 = static tension in the reference state without accessory torques
c = steady state belt speed
k_r = tensioner spring stiffness
l_i = length of belt span i
m = belt mass per unit length
r_i = radius of pulley i
u_i = longitudinal displacement of span i
w_i = transverse displacement of span i
θ_i = rotation of pulley i
θ_t = rotation of the tensioner
θ_{tr} = tensioner arm rotation in the reference state
$\zeta_1, \zeta_2, \beta_1, \beta_2$ = alignment angles of tensioner (see Fig. 2)
$\hat{M}_S^{(i)}, \hat{M}_E^{(i)}$ = end moments of span i

References

[1] Hawker, L. E., 1991, "A Vibration Analysis of Automotive Serpentine Accessory Drives Systems," Ph.D. dissertation, University of Windsor, Ontario, Canada.
[2] Barker, C. R., Oliver, L. R., and Breig, W. F., 1991, "Dynamic Analysis of Belt Drive Tension Forces During Rapid Engine Acceleration," SAE Paper No. 910687.
[3] Hwang, S. J., Perkins, N. C., Ulsoy, A. G., and Meckstroth, R., 1994, "Rotational Response and Slip Prediction of Serpentine Belt Drives Systems," ASME J. Vibr. Acoust., **116**, pp. 71–78.
[4] Leamy, M. J., and Perkins, N. C., 1998, "Nonlinear Periodic Response of Engine Accessory Drives With Dry Friction Tensioners," ASME J. Vibr. Acoust., **120**, pp. 909–916.
[5] Beikmann, R. S., 1992, "Static and Dynamic Behavior of Serpentine Belt Drive Systems: Theory and Experiments," Ph.D. dissertation, The University of Michigan, Ann Arbor, MI.
[6] Beikmann, R. S., Perkins, N. C., and Ulsoy, A. G., 1996, "Design and Analysis of Automotive Serpentine Belt Drive Systems for Steady State Performance," ASME J. Mech. Des., **119**, pp. 162–168.
[7] Beikmann, R. S., Perkins, N. C., and Ulsoy, A. G., 1996, "Free Vibration of Serpentine Belt Drive Systems," ASME J. Vibr. Acoust., **118**, pp. 406–413.
[8] Parker, R. G., 2003, "Efficient Eigensolution, Dynamic Response, and Eigensentivity of Serpentine Belt Drives," J. Sound Vib., in press.
[9] Ascher, U., and Russell, R., 1981, "Reformulation of Boundary Value Problems into 'Standard' Form," SIAM (Soc. Ind. Appl. Math.) Rev., **23**, pp. 238–254.
[10] Kong, L., and Parker, R. G., 2003, "Coupled Belt-Pulley Vibration in Serpentine Drives With Belt Bending Stiffness," ASME J. Appl. Mech., in press.
[11] Wang, K. W., and Mote, C. D., Jr., 1986, "Vibration Coupling Analysis of Band/Wheel Mechanical Systems," J. Sound Vib., **109**, pp. 237–258.
[12] Mote, C. D., Jr., and Wu, W. Z., 1985, "Vibration Coupling in Continuous Belt and Band Systems," J. Sound Vib., **102**, pp. 1–9.
[13] Zhang, L., and Zu, J. W., 1999, "Modal Analysis of Serpentine Belt Drive Systems," J. Sound Vib., **222**(2), pp. 259–279.
[14] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, John Wiley and Sons, New York, pp. 447–455.
[15] Shampine, L. F., Kierzenka, J., and Reichelt, M. W., 2000, "Solving Boundary Value Problems for Ordinary Differential Equations in Matlab With Bvp4c," available at ftp://ftp.mathworks.com/pub/doc/papers/bvp/
[16] Mote, C. D., Jr., 1965, "A Study of Band Saw Vibrations," J. Franklin Inst., **279**, pp. 430–444.

The Application of Lagrange Equations to Mechanical Systems With Mass Explicitly Dependent on Position

C. P. Pesce
Mem. ASME
Department of Mechanical Engineering,
Escola Politécnica,
University of São Paulo,
Brazil
e-mail: ceppesce@usp.br

The derivation and application of the Lagrange equations of motion to systems with mass varying explicitly with position are discussed. Two perspectives can be followed: systems with a material type of source, attached to particles continuously gaining or losing mass, and systems for which the variation of mass is of a nonlinear control volume type, mass trespassing a control surface. This is the case if, for some theoretical or practical reason, a partition into subsystems is considered. An important class of problems in which the extended Lagrange equations turn to be useful emerges from "hydromechanics," whenever a finite number of generalized coordinates can be used, under the concept of the added mass tensor. A particular and interesting one is addressed in the present paper: the classical hydrodynamic impact of a rigid body against a liquid free surface.
[DOI: 10.1115/1.1601249]

1 Introduction

Variable-mass systems have been the focus of a large number of problems in classical mechanics. As early as in 1857, Cayley [1] discussed the problem of a chain being coiled up at a table. Levi-Civita [2], in 1928, treated the motion of a variable mass point body, in the two-body problem, introducing an extended form of Newton's law. Such a form, however, is only valid if mass is gained or lost at null velocity with respect to an inertial frame. A renewed interest in this subject emerged with the "rocket problem," now included in many textbooks. Another special class of problems, which has deserved consideration, relates to "tethered satellite systems." Similar to the coiling chain problem such a class concerns the deploying or the retrieving of a cable from or into a body moving along it, see, e.g., Crellin et al. [3,4]. Specific problems also related to cable systems, such as the lifting-crane problem, Cveticanin [5], have been studied recently. The textile industry is another source of variable-mass systems problems in mechanics, Cveticanin [6–9]. All those applied research activities gave rise to the need of new theoretical investigations, as those conducted in the 1990s by Cveticanin, [10–12] and even earlier, in the 1980s, by Ge [13,14].

The purpose of the present paper is to discuss some theoretical aspects involved in the variable-mass systems dynamics, usually hidden behind many derivations. Particular emphasis is given to the case of systems where the variation of mass is an explicit function of position. The discussion is extended to "hydromechanics," here meant as a class of problems involving potential flows around bodies, whenever a finite number of generalized coordinates can be used as a proper representation of the motion, under the concept of the added mass tensor. As added mass usually depends on position of the body, explicitly, this may render incorrect the application of the usual Euler-Lagrange equations of motion to any isolated subsystem. A particular and interesting one is addressed in the present paper: the classical hydrodynamic impact of a rigid body against the water free surface; see, e.g., Korobkin and Pukhnachov [15], Cooker and Peregrine [16], and Molin, Cointe, and Fontaine [17].

2 Some Theoretical Considerations

Kinetic energy, $T=T(q_j;\dot{q}_j;t)$, is, by definition, at least a bilinear form in generalized velocities, $\dot{q}_j$, and also, in many cases, a function of the generalized coordinates, q_j. However, physically speaking, there is a rather large distinction between the kinetic energy being an *explicit* function of the kinematic state of the system (generalized coordinates and velocities), or else, an *implicit* function of those, through a possible dependence of mass, in the form $m_i=m_i(q_j;\dot{q}_j;t)$. Whenever this is the case, the mechanical system does not obey the usual form of the classical Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right)-\frac{\partial T}{\partial q_j}=Q_j,$$

unless nonconservative generalized forces associated to fluxes of mass are already considered included in Q_j, as those usually referred to as Metchersky's reactive forces, see, e.g., Cveticanin [10]. Otherwise, the derived equations of motion will take an erroneous form. It could be argued that the usual Euler-Lagrange (or simply Lagrange) equations would be suitable to the case for which $m=m(t)$, mass varying solely as an explicit function of time. This is true, however, only if mass is gained or lost at zero velocity with respect to an inertial frame of reference, in a "continuous impact manner," as assumed by Agostinelli [18] (p. 257), who concluded being the Lagrange equations invariant for holonomic variable-mass systems. In that particular work, Levi-Civita's special form of momentum equation $d\mathbf{p}/dt=\mathbf{f}$ was used, with no reference to any reactive force, proportional to the velocity of the particle that is being expelled from or incorporated to the system.

The reason for these subtle distinctions, concerning how mass changes, if as an *explicit* or an *implicit* function of time, will be shown next. The answer hides behind the derivation of the most general form of the Lagrange equations, as presented, e.g., in Cveticanin [10]. We point out that, in systems with mass explicitly dependent on position, a naive application of the usual Lagrange equations, without any special consideration on generalized

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, April 27, 2001; final revision, February 20, 2003. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

forces, leads to equations of motions which lack terms of the form $(1/2)(\partial m/\partial q)\dot{q}^2$. For instance, a nonproper application of Lagrange equation to the hydrodynamic impact problem leads to an erroneous term of the form $(1/2)(dM_a/dZ)W^2$, being M_a the instantaneous added mass, Z the penetration of the impacting body into the initially quiescent free surface, and W the downward vertical velocity of the body.

We now proceed, deriving the "extended" Lagrange equations in Secs. 3 and 4. In Sec. 5, we present two didactic examples: (i) the reel problem; (ii) the free-surface impact problem.

3 The Classical Lagrange Equations

Consider a system of N particles of mass m_i. Let P_i be the corresponding position in a given inertial frame of reference and $\mathbf{p}_i = m_i \mathbf{v}_i$ the momentum. By extending Levi-Civita's form of Newton's law to cases when mass is gained or lost with no null velocity, D'Alembert's principle can be written

$$\sum_i \left(\frac{d\mathbf{p}_i}{dt} - \mathbf{F}_i \right) \cdot \delta P_i = \mathbf{0}, \tag{1}$$

where

$$\mathbf{F}_i = \mathbf{f}_i + \mathbf{h}_i, \tag{2}$$

$\mathbf{f}_i$ being the sum of all active forces acting on P_i, and

$$\mathbf{h}_i = \dot{m}_i \mathbf{v}_{oi} \tag{3}$$

is a reactive force, proportional to the rate of variation of mass with respect to time and to the velocity $\mathbf{v}_{oi}$ of the expelled (or gained) mass. Note that the reactive force known as Metchersky's force, in the Russian technical literature, is usually written as function of relative velocities, in the form

$$\mathbf{\Phi}_i = \dot{m}_i(\mathbf{v}_{oi} - \mathbf{v}_i) = \mathbf{h}_i - \dot{m}_i \mathbf{v}_i. \tag{4}$$

Under this latter interpretation, the extended D'Alembert's Principle, Eq. (1), would be equivalently written (see, e.g., Cveticanin [10])

$$\sum_i \left(m_i \frac{d\mathbf{v}_i}{dt} - (\mathbf{f}_i + \mathbf{\Phi}_i) \right) \cdot \delta P_i = \mathbf{0}. \tag{1'}$$

Consider virtual displacements δP_i, and a M set of generalized coordinates q_j (for simplicity, the system is considered holonomic) such that

$$\delta P_i = \sum_j \frac{\partial P_i}{\partial q_j} \cdot \delta q_j. \tag{5}$$

The velocities $\mathbf{v}_i = \mathbf{v}_i(q_j; \dot{q}_j; t)$; $j = 1, \ldots, M$ are, as usual, considered as functions of generalized coordinates and derivatives, as well of time t. The following common and straightforwardly derivable kinematic relations, which can be found in any good textbook in classical mechanics (see, e.g., Targ [19], p. 508), will be used as well:

$$\frac{\partial \mathbf{v}_i}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial P_i}{\partial q_j} \right), \tag{6}$$

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} = \frac{\partial P_i}{\partial q_j}, \tag{7}$$

and

$$\frac{d\mathbf{v}_i}{dt} \cdot \frac{\partial P_i}{\partial q_j} = \frac{d}{dt}\left(\frac{1}{2} \frac{\partial \mathbf{v}_i^2}{\partial \dot{q}_j} \right) - \frac{\partial}{\partial q_j}\left(\frac{1}{2} \mathbf{v}_i^2 \right). \tag{8}$$

We also define the generalized, nonconservative force, which already includes the reactive force $\mathbf{h}_i = \dot{m}_i \mathbf{v}_{oi}$, as

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial P_i}{\partial q_j} = \sum_i (\mathbf{f}_i + \mathbf{h}_i) \cdot \frac{\partial P_i}{\partial q_j}. \tag{9}$$

3.1 The Simplest Case of Systems of Particles with Constant Mass. Usually, for systems of constant mass, the kinetic energy $T_i = 1/2 m_i \mathbf{v}_i^2$ of a given particle P_i is, apart from the mass m_i, identified in both terms of Eq. (8), such that

$$m_i \frac{d\mathbf{v}_i}{dt} \cdot \frac{\partial P_i}{\partial q_j} = \frac{d}{dt}\left(\frac{1}{2} \frac{\partial m_i \mathbf{v}_i^2}{\partial \dot{q}_j} \right) - \frac{\partial}{\partial q_j}\left(\frac{1}{2} m_i \mathbf{v}_i^2 \right) = \frac{d}{dt}\left(\frac{\partial T_i}{\partial \dot{q}_j} \right) - \frac{\partial T_i}{\partial q_j}. \tag{10}$$

Observing that, in this simplest case, $dm_i/dt = 0$, such that $d\mathbf{p}_i/dt = m_i(d\mathbf{v}_i/dt)$, substituting Eqs. (5) and (10) in the extended D'Alembert principle (1), and observing that the generalized forces Q_j reduce, from Eq. (9), to the usual form,

$$Q_j = \sum_i \mathbf{f}_i \cdot \frac{\partial P_i}{\partial q_j}, \tag{11}$$

one easily obtains the usual Lagrange equations

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j; \quad j = 1, \ldots, M, \tag{12}$$

for a system where all particles have invariant mass.

3.2 Systems of Particles with Mass as Explicit Function of Time $m_i = m_i(t)$. Before entering the more general case, where $m_i = m_i(q_j; \dot{q}_j; t)$, it is also instructive to consider the case where mass is solely an explicit function of time, $m_i = m_i(t)$. D'Alembert's principle reads

$$\sum_i \left(\frac{d\mathbf{p}_i}{dt} - (\mathbf{f}_i + \mathbf{h}_i) \right) \cdot \delta P_i = \sum_j \sum_i \left(m_i \frac{d\mathbf{v}_i}{dt} + \frac{dm_i}{dt} \mathbf{v}_i - (\mathbf{f}_i + \mathbf{h}_i) \right) \cdot \frac{\partial P_i}{\partial q_j} \delta q_j = \mathbf{0}. \tag{13}$$

Integrating by parts, the first term transforms as follows:

$$\begin{aligned} m_i \frac{d\mathbf{v}_i}{dt} \cdot \frac{\partial P_i}{\partial q_j} &= \frac{d}{dt}\left(\frac{1}{2} m_i \frac{\partial \mathbf{v}_i^2}{\partial \dot{q}_j} \right) - \frac{dm_i}{dt}\left(\frac{1}{2} \frac{\partial \mathbf{v}_i^2}{\partial \dot{q}_j} \right) - \frac{\partial}{\partial q_j}\left(\frac{1}{2} m_i \mathbf{v}_i^2 \right) \\ &= \frac{d}{dt}\left(\frac{1}{2} \frac{\partial m_i \mathbf{v}_i^2}{\partial \dot{q}_j} \right) - \frac{dm_i}{dt}\left\{ \frac{\partial}{\partial \dot{q}_j}\left[\frac{\partial}{\partial m_i}\left(\frac{1}{2} m_i \mathbf{v}_i^2 \right) \right] \right\} \\ &\quad - \frac{\partial}{\partial q_j}\left(\frac{1}{2} m_i \mathbf{v}_i^2 \right) \\ &= \frac{d}{dt}\left(\frac{\partial T_i}{\partial \dot{q}_j} \right) - \frac{dm_i}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\frac{\partial T_i}{\partial m_i} \right) \right] - \frac{\partial T_i}{\partial q_j}. \end{aligned} \tag{14}$$

Observing Eq. (7), the second term in Eq. (13) transforms as

$$\frac{dm_i}{dt} \mathbf{v}_i \cdot \frac{\partial P_i}{\partial q_j} = \frac{dm_i}{dt} \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} = \frac{1}{2} \frac{dm_i}{dt} \frac{\partial \mathbf{v}_i^2}{\partial \dot{q}_j} = \frac{dm_i}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\frac{\partial T_i}{\partial m_i} \right) \right]. \tag{15}$$

This latter expression is the most general (and concise) form for the parcel that depends on the variation of mass in the momentum time derivative. Note that this form is exactly the opposite of the second term appearing on the right-hand side of Eq. (14).

They cancel each other when Eqs. (14) and (15) are substituted into Eq. (13), leading to a equation of motion which has the same form as Eq. (12), with the generalized forces given by Eq. (9). This is a very subtle step which can explain why a system of particles with variable mass, but given solely as an explicit function of time, $m_i = m_i(t)$, obey the same form of Euler-Lagrange equations that govern a system of particles of invariant mass. This is essentially Agostinelli's [15] result (p. 257), now having the generalized forces extended according to Eq. (9), by including the reactive forces defined by Eq. (3).

4 The Lagrange Equations for Systems of Particles with Variable Mass as Function of Time, Generalized Coordinates and Velocities, $m_i = m_i(q_j; \dot{q}_j; t)$

As before, use is made of the extended D'Alembert principle, in the form (13) besides Eqs. (6)–(9). In this general case the first term in Eq. (13) is given as

$$
\begin{aligned}
m_i \frac{\partial \mathbf{v}_i}{dt} \cdot \frac{\partial P_i}{\partial q_j} &= \frac{d}{dt}\left(\frac{1}{2} m_i \frac{\partial \mathbf{v}_i^2}{\partial \dot{q}_j}\right) - \frac{1}{2}\frac{dm_i}{dt}\left(\frac{\partial \mathbf{v}_i^2}{\partial \dot{q}_j}\right) - \frac{\partial}{\partial q_j}\left(\frac{1}{2} m_i \mathbf{v}_i^2\right) \\
&\quad + \frac{1}{2}\frac{\partial m_i}{\partial q_j}(\mathbf{v}_i^2) \\
&= \frac{d}{dt}\left(\frac{1}{2}\frac{\partial m_i \mathbf{v}_i^2}{\partial \dot{q}_j}\right) - \frac{1}{2}\frac{d}{dt}\left(\frac{\partial m_i}{\partial \dot{q}_j}\right)\mathbf{v}_i^2 - \frac{1}{2}\frac{dm_i}{dt}\left(\frac{\partial \mathbf{v}_i^2}{\partial \dot{q}_j}\right) \\
&\quad - \frac{\partial}{\partial q_j}\left(\frac{1}{2} m_i \mathbf{v}_i^2\right) + \frac{1}{2}\frac{\partial m_i}{\partial q_j}(\mathbf{v}_i^2) \\
&= \frac{d}{dt}\left(\frac{\partial T_i}{\partial \dot{q}_j}\right) - \frac{1}{2}\frac{d}{dt}\left(\frac{\partial m_i}{\partial \dot{q}_j}\right)\mathbf{v}_i^2 - \frac{dm_i}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\frac{\partial T_i}{\partial m_i}\right)\right] \\
&\quad - \frac{\partial T_i}{\partial q_j} + \frac{1}{2}\frac{\partial m_i}{\partial q_j}(\mathbf{v}_i^2).
\end{aligned} \tag{16}
$$

This is the most general form for the parcel that depends on the acceleration in the momentum time derivative. Taking both most general forms, Eqs. (15) and (16), and substituting in Eq. (13), with the generalized forces given by Eq. (9), one finally obtains[1]

$$
\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j + \sum_i \left\{\frac{1}{2}\frac{d}{dt}\left(\frac{\partial m_i}{\partial \dot{q}_j}(\mathbf{v}_i)^2\right) - \frac{1}{2}\frac{\partial m_i}{\partial q_j}(\mathbf{v}_i)^2\right\}; \quad j = 1, \ldots, M. \tag{17}
$$

These are the dynamic equations for a system of particles with variable mass in the form $m_i = m_i(q_j; \dot{q}_j; t)$. The last two terms can be interpreted as additional parcels of momentum time rate, caused by the changes in mass with position and velocities. Alternatively, they could be interpreted as additional (nonconservative) ''equivalent generalized forces'' that take into account the variation of mass of each particle in the system. By properly defining a nonconservative generalized force $\hat{Q}_j$, Eq. (17) can finally be written in the simplest (and usual) form,

$$
\begin{aligned}
&\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = \hat{Q}_j; \quad j = 1, \ldots, M, \\
&\hat{Q}_j = \sum_i (\mathbf{f}_i + \dot{m}_i \mathbf{v}_{oi}) \cdot \frac{\partial P_i}{\partial q_j} + \sum_i \left\{\frac{1}{2}\frac{d}{dt}\left(\frac{\partial m_i}{\partial \dot{q}_j}(\mathbf{v}_i)^2\right) - \frac{1}{2}\frac{\partial m_i}{\partial q_j}(\mathbf{v}_i)^2\right\}.
\end{aligned} \tag{18}
$$

Moreover, if the active forces $\mathbf{f}_i$ are split into

$$
\mathbf{f}_i = \mathbf{f}_i^c + \mathbf{f}_i^{nc}, \tag{19}
$$

$\mathbf{f}_i^c$ being conservative and $\mathbf{f}_i^{nc}$ nonconservative parcels, respectively, such that

$$
\sum_i \mathbf{f}_i^c \cdot \delta P_i = -\sum_j \frac{\partial V}{\partial q_j}\delta q_j, \tag{20}
$$

V being the usual potential-energy function, Eq. (18) can be written in the convenient form,

$$
\begin{aligned}
&\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \hat{Q}_j^{nc}; \quad j = 1, \ldots, M, \\
&\hat{Q}_j^{nc} = \sum_i (\mathbf{f}_i^{nc} + \dot{m}_i \mathbf{v}_{oi}) \cdot \frac{\partial P_i}{\partial q_j} + \sum_i \left\{\frac{1}{2}\frac{d}{dt}\left(\frac{\partial m_i}{\partial \dot{q}_j}(\mathbf{v}_i)^2\right) - \frac{1}{2}\frac{\partial m_i}{\partial q_j}(\mathbf{v}_i)^2\right\},
\end{aligned} \tag{21}
$$

where $L = T - V$ is the Lagrangian function. As a matter of fact, problems in classical mechanics where mass is an explicit function of velocities are hard to conceive, to say the least.[2] Therefore, if just position dependence is considered, we obtain

$$
\begin{aligned}
&\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \hat{Q}_j^{nc}; \quad j = 1, \ldots, M, \\
&\hat{Q}_j^{nc} = \sum_i (\mathbf{f}_i^{nc} + \dot{m}_i \mathbf{v}_{oi}) \cdot \frac{\partial P_i}{\partial q_j} - \sum_i \left\{\frac{1}{2}\frac{\partial m_i}{\partial q_j}(\mathbf{v}_i)^2\right\}.
\end{aligned} \tag{22}
$$

Equation (22) can be verified to agree with the derivation provided by Cveticanin [10], for the practical case where mass is solely dependent on generalized coordinates (not in velocities). It must be observed that the first term appearing in Cveticanin's Eq. (8) is exactly our term given by Eq. (15). In the present derivation, the Mertchersky's reactive force has been split in the form $\Phi_i = \dot{m}_i(\mathbf{v}_{oi} - \mathbf{v}_i) = \mathbf{h}_i - \dot{m}_i \mathbf{v}_i$, such that the term given by Eq. (15) ends to be cancelled out, as observed before, turning the final Eq. (22) somewhat simpler in form.

5 Two Didactic and Illustrative Examples

5.1 A Very Simple Example in Mechanical Engineering: The Deploying of a Heavy Cable from a Reel. Consider the classical and very simple problem of a heavy cable being deployed from a reel by the action of gravity, as presented in Fig. 1. This *very well-known problem* has been elected with the only purpose of exemplifying how partition into subsystems might lead to an erroneous use of Euler-Lagrange equations. The reel has radius R and moment of inertia I_O, around the axis of rotation.

[1] Recall that $\mathbf{v}_i = \mathbf{v}_i(q_j; \dot{q}_j; t)$; $j = 1, \ldots, M$.

[2] Such cases are, however, the core of relativistic problems, see, e.g., Pars [20], Chap. XI, p. 190.

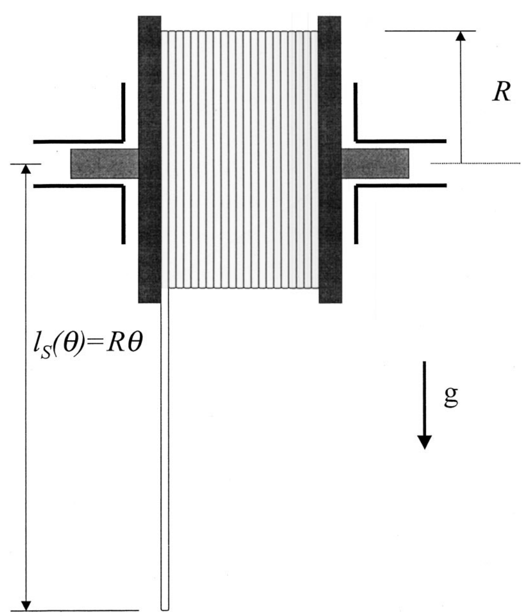


Fig. 1 Cable being deployed from a reel

Let μ be the mass per unit of length of the cable, supposed non-extensible and infinitely flexible. Without loss of generality, let θ be the generalized coordinate, measured from horizontal such that, at a given instant t, the cable suspended length is $l_S(\theta)=R\Theta$. Let also L be the total length of the cable such that $m=\mu L$ is the total cable's mass. For simplicity, we take the cable diameter to be very small compared to the reel's radius such that the winding pitch is also small and that all winding turns can be accommodated into a single winding layer. Let also $m_S(\theta)=\mu l_S(\theta)=\mu R\theta$ and $m_R(\theta)=m-m_S(\theta)=\mu(L-\theta R)$ be, respectively, the cable's suspended and wound masses.

Obviously, for this simple problem, the best and shortest way to apply Lagrange equation is to consider the whole (invariant mass) system. Kinetic energy is simply $T=1/2(I_O+mR^2)\dot{\theta}^2$. Accordingly, potential energy is $V=-1/2[m_S(\theta)gR\theta]=-1/2\mu gR^2\theta^2$. A straightforward application of the usual Lagrange equation $d/dt(\partial L/\partial\dot{\theta})-\partial L/\partial\theta=0$ to this invariant mass system leads to the correct equation of motion $(I_O+mR^2)\ddot{\theta}-\mu gR^2\theta=0$.

Suppose now that, for some practical reason, the analyst decides to take a subsystem composed by the reel and by the wound part of the cable, considering the suspended part of the cable as a second subsystem. Note that the suspended part of the cable can be considered as a material point gaining mass at rate $\dot{m}_S(\theta)=\mu R\dot{\theta}$, with velocity $\nu_o=R\dot{\theta}$. Let the active force be $f(\theta)=m_s(\theta)g-\tau(\theta)$, $\tau(\theta)$ being the traction at the upper section. Applying the extended Levi-Civita form of Newton's law to the suspended part, we easily obtain $(d/dt)[m_S(\theta)R\dot{\theta}]=m_s(\theta)g-\tau(\theta)+\dot{m}_S(\theta)R\dot{\theta}$. Hence the tension force applied by the wound part to the suspended part of the cable is simply $\tau(\theta)=\mu R\theta(g-R\ddot{\theta})$.

Now let $I_1=I_O+m_R(\theta)R^2=I_O+\mu R^2(L-R\theta)$ be the moment of inertia of the first (reel+wound part of the cable) subsystem, such that the corresponding kinetic energy is given by $T_1=1/2(I_R)\dot{\theta}^2=1/2[I_O+m_R(\theta)R^2]\dot{\theta}^2=1/2[I_O+\mu R^2(L-R\theta)]\dot{\theta}^2$. Note that mass exits the wound part with velocity $\nu_o=R\dot{\theta}$ at a rate $\dot{m}_R(\theta)=-\mu R\dot{\theta}$. If, *erroneously*, the Lagrange equation is applied to the first subsystem in the form (12), with $Q_\theta=[\tau(\theta)+\dot{m}_R(\theta)R\dot{\theta}]R$, such that $(d/dt)(\partial T_1/\partial\dot{\theta})-(\partial T_1/\partial\theta)=[\tau(\theta)+\dot{m}_R(\theta)R\dot{\theta}]R$, the following and obviously incorrect equation of motion is obtained: $(I_O+mR^2)\ddot{\theta}+(1/2)\mu R^3\dot{\theta}^2-\mu gR^2\theta=0$. Note the presence of an erroneous quadratic term in velocity, namely, $(1/2)\mu R^3\dot{\theta}^2$.

However, if the correct form of the Lagrange equation, given by Eq. (18), is applied to this variable mass sub-system, with $\hat{Q}_\theta=[\tau(\theta)+\dot{m}_R(\theta)R\dot{\theta}]R-(1/2)(dm_R/d\theta)R^2\dot{\theta}^2$, i.e., such that $(d/dt)(\partial T_1/\partial\dot{\theta})-(\partial T_1/\partial\theta)=[\tau(\theta)+\dot{m}_R(\theta)R\dot{\theta}]R-(1/2)\times(dm_R/d\theta)R^2\dot{\theta}^2$, the *correct* equation of motion, $(I_O+mR^2)\ddot{\theta}-\mu gR^2\theta=0$, previously derived when the whole system was considered, is readily recovered.

5.2 The Impact of a Rigid Body Against the Water Free Surface. It is usual practice to treat potential hydrodynamic problems involving motion of solid bodies within the frame of system dynamics. This is done whenever a finite number of generalized coordinates can be used as a proper representation for the motion of the whole fluid. We shall refer to this kind of approach as "hydromechanical." This is made possible through the introduction of the well-known concept of the added mass tensor; see, e.g., Newman [21] or Lamb [22], where a thorough analysis is presented on this subject. Particularly, art. 137 of Lamb's opus is dedicated to the application of the Lagrangian formalism, envisaging to address problems where the fluid kinetic energy, transferred from a moving body, depends on body's position. This is, for instance, the case of a body moving close to a solid barrier. In this particular example, fluid kinetic energy varies according to the distance of the body from the barrier, and such a variation can be represented through the added mass tensor.

This would also be the case for the classical hydrodynamic impact problem of a rigid body against a free surface. In fact, during the initial stage of impact, sprays (or jets) are formed, relieving a very high field pressure that is formed around the body, as shown in Fig. 2.

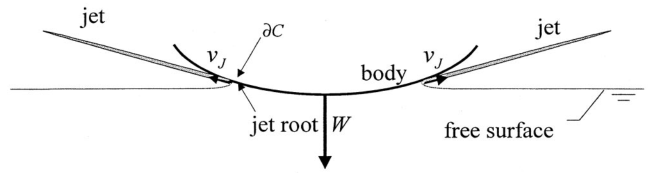


Fig. 2 The impact of a rigid body against a liquid free surface. Jets or sprays are formed. ∂c indicates the instantaneous position of jet's root, across which there is a flux of kinetic energy and mass.

Nonetheless, as common in hydrodynamics, the analyst would probably split the whole body of fluid into two subsystems: (i) the bulk of fluid, and (ii) the fluid inside the jets. If this procedure is followed, an exchanging flux of mass and energy clearly exists between these two subsystems, through their common frontiers, so that, even being the mass invariant for the whole material system, it is not for the subsystems.

Within the hydromechanical approach, the present impact problem can be consistently formulated under the Lagrangian formalism. We take, for simplicity, a purely symmetric and vertical impact case of a body being dropped against a free surface of a liquid, with infinite depth. The aim is to determine the impacting force on the body, given its initial kinematic state. We take the subsystem (i) (the bulk of liquid), recalling the added mass dependence on the position of the body. Let W be the impacting body velocity, Z the body's penetration into the water, measured from the initially quiescent free surface and $M_a^B=M_a^B(Z)$ the body's added mass, in the vertical direction. Consistently, the true added mass is defined in the bulk of the fluid subsystem (excluding the jets), at each instant of time, what is explicitly represented through the superscript^B for "bulk." This is accomplished by taking into account the so-called "wetted surface correction on added mass," caused by the marching of the jet root on the body's surface. This marching generates a geometric dependence of the kinetic energy on the generalized coordinate Z. Note that the generalized coordinate Z is an unknown function of the time t. At a given instant t, the kinetic energy *in the bulk of fluid* is then written

$$T=\frac{1}{2}M_a^BW^2,$$

$$M_a^B=M_a^B(Z),$$

$$Z=\int_{0^+}^{t}Wdt. \tag{23}$$

However, as already pointed out, considering just the bulk of fluid implies that a volume of control has been defined by cutting the jets out. An actual variation of mass should therefore be taken into account within the subsystem under consideration. The bulk of fluid loses mass and kinetic energy to the jets, through the jet roots ∂C.

The correct "Lagrangian formalism" approach is therefore to apply Eq. (18).[3] Note that the pressure force is applied along the interface formed by the bulk of fluid and the body; Fig. 2. $-F_z{}^B$ being the force applied on the bulk of fluid,[4] this equation reads

$$-F_z{}^B=-\frac{d}{dt}\left(\frac{\partial T}{\partial W}\right)+\frac{\partial T}{\partial Z}-\frac{1}{2}\frac{dM_a^B}{dZ}W^2-2\dot{m}\nu_J\sin\alpha. \tag{24}$$

[3]Buoyancy and gravitational effects are neglectable in the very starting of impact, when inertia forces are, by far, the dominant ones.

[4]The opposite sign of the force applied on the body by the fluid.

The fourth term corresponds to the reactive force, $\dot{m}$ being the flux of mass through the jets and ν_J the absolute velocity of the fluid particles at the jet root; α is the instantaneous angle of the jets with respect to the horizontal. This equation transforms as

$$-F_z{}^B = -\frac{d}{dt}\left(\frac{\partial T}{\partial W}\right) + \frac{\partial T}{\partial Z} - \frac{1}{2}\frac{dM_a^B}{dZ}W^2 - 2\dot{m}\nu_J \sin\alpha$$

$$= -\frac{d}{dt}(M_a^B W) + \frac{1}{2}W^2\frac{dM_a^B}{dZ} - \frac{1}{2}\frac{dM_a^B}{dZ}W^2 - 2\dot{m}\nu_J \sin\alpha,$$

so that

$$F_z^B = -\frac{d}{dt}(M_a^B W) + 2\bar{m}\nu_J \sin\alpha, \tag{25}$$

recovering an expected result, in accordance with Eq. (1).

The way to calculate the instantaneous added mass, the flux of mass, the velocity ν_J at the jet root, ∂C (as well as the velocity ν_R of the jet root itself), can be found, e.g., in Faltinsen and Zhao [23], and Cointe and Armand [24]. It is also crucial to mention a neat conclusion drawn, in two independent papers, by Korobkin and Pukhnachov [15] and by Molin et al. [17]: *half of the kinetic energy is transferred to the jets and half to the bulk of the fluid.*

As a matter of fact, the analysis by Molin et al. [17], after an asymptotic analysis by Cointe and Armand [24] on the particular and important case of a circular cylinder of radius R, proved that, $\varepsilon = \sqrt{Wt/R}$ being a small parameter—or, in other words, within a short scale of time—the thickness of the jet root is of order $\delta_J = O(\varepsilon^3 \pi R/4)$ and the velocity at the jet root is of order $\nu_J = O(\varepsilon^{-1}W)$. It then follows that mass flux through the jets is of order $\dot{m} = O(\varepsilon^2 \pi\rho RLW/2)$ and so, $\dot{m}\nu_J = O(\varepsilon\pi\rho RLW^2/2)$, L being the cylinder's length $(L/R \gg 1)$. Hence, the vertical force applied by the jets on the bulk of fluid is of order $O(\varepsilon\pi\rho RLW^2 \sin\alpha)$. *Contrarily*, the energy flux is of order $G = O(\pi\rho RLW^3)$ and $(d/dt)(M_a^B W) = O(\varepsilon^{-2}\pi\rho RLW^2)$. Therefore the impact force on the body, for this case, could be written

$$F_z{}^B = -\frac{d}{dt}(M_a^B W). \tag{26}$$

This derivation consistently recovers, up to third order, statements as given in Faltinsen and Zhao [23], such that $F_z = -(d/dt) \times (M_a^B W)$.

On the other hand, if (see, e.g., Wu [25]) the third and fourth terms appearing on the right-hand side of Eq. (24) were not considered at all, a *different assertive would be obtained*, according to which $F_z = -M_a^B(dW/dt) - (1/2)W(dM_a^B/dt)$.

This contradiction has been in fact recognized as a (apparent) controversy, stating that important discrepancies would be obtained in impact forces calculations, if either integrating pressures methods or energy approaches were used. It has been pointed out by a series of authors (e.g., Miloh [26]), even as late as in 1998; Wu [25]. The explanation for such an apparent controversy is actually the one observed by Molin et al. [17], here re-interpreted from the point of view of analytical mechanics.

It should be mentioned that a generalization of the dynamic equations to continuum systems is not straightforward as it could appear from this simple example. A rigorous treatment of Hamilton principles in continuum mechanics can be found in Seliger and Whitham [27]. However, to the present date, and to the author's knowledge, no theoretical extension has been made considering the case of continuum systems with variable mass as an explicit function of coordinates (and velocities).

6 Conclusions

A system of particles with mass varying explicitly with position (and velocity) does not obey the classical form of Lagrange equations of motions. A general derivation has been provided, recovering, e.g., Cveticanin's [10] result. Some simple examples were given to exemplify the use of such a general equation. Application to engineering problems is somewhat rare, however. In fact, it is rather difficult to conceive practical applications in which mass is an explicit function of position (and, even more rare, of velocities).

Apart from classical problems such as the ''rocket problem,'' ''tethered satellite systems,'' ''deployment of cables,'' etc., good examples can be extracted from potential flows around bodies. This is true whenever a finite number of generalized coordinates can be used as a proper representation for the position of the whole system. In such cases, partition into subsystems, together with proper definition of control volumes, is often mandatory. Real losses of mass associated to fluxes of energy through permeable surfaces are then likely to occur, rendering conceptually incorrect the application of the usual Lagrange equations of motion.

A particular example where this kind of treatment turns out to be puzzling is the classical problem concerning the hydrodynamic impact of a rigid body against a liquid surface. During the initial stage of impact, jets or sprays are formed along the intersection between the body and the free surface. Ignoring energy flux through the jets has been the cause of a (apparent) controversy, by which pressure integration and energy methods would lead to different expressions for the impacting force. According to Molin et al. [17], although flux of mass through the jets can be shown to be neglectable, the flux of kinetic energy is not. Half the kinetic energy coming from the impacting body is transferred to the bulk of the fluid and half is transferred to the jets. We believe that this simple derivation, from the point of view of Lagrange equations, may contribute with the reasoning of the cited authors, even emphasizing the need of a consistent definition of the added-mass quantity.

Acknowledgments

The author acknowledges a research grant, No. 304062/85, from CNPq—the Brazilian National Research Council and a grant from the Brazilian Navy, during a sabbatical year, 1999, spent at the Naval Architecture and Marine Engineering Department (NAME), University of Michigan. The author is especially grateful to Professor Armin Troesch, University of Michigan, for introducing him to the impact problem, indicating a bench of valuable references and for the time spent in very interesting discussions. Thanks to the whole (NAME) staff and faculty members, especially to Professor Michael Bernitsas, for their support and comprehension. Thanks also to Professors L.N.F. França, J.A.P. Aranha and C.G. Ragazzo, University of São Paulo, for their time reading the original manuscript, and to Dr. E. A. Tannuri, for interesting suggestions regarding engineering applications to simple systems in which the present analysis might be relevant. The author is also grateful to two of the reviewers, who patiently read the original manuscript, made a number of very enlightening criticisms and rich suggestions, and brought to this author's knowledge the works by Cayley, Crellin, Cveticanin, Agostinelli, and Ge.

References

[1] Cayley, A., 1857, ''On a Class of Dynamical Problems,'' Proc. R. Soc. London, **8**, pp. 506–511.

[2] Levi-Civita, T., 1928, ''Sul Moto di un Corpo de Massa Variabile,'' Rendiconti delle Sedute della Reale Accademia Nazionale dei Lincei,, **8**, pp. 329–333, Aggiunta alla nota, pp. 621–622.

[3] Crellin et al., 1995, ''Deployment and Retraction of a Continuous Tethered Satellite—the Equations Revisited,'' *Proceedings of the Fourth Int. Conference on Tethers in Space*, Hampton, VA, pp. 1415–1423.

[4] Crellin, et al., 1997, ''On Balance and Variational Formulations of the Equation of a Motion of a Body Deploying Along a Cable,'' J. Appl. Mech., **64**, pp. 369–374.

[5] Cveticanin, L., 1995, ''Dynamic Behavior of the Lifting Crane Mechanism,'' Mech. Mach. Theory, **30**, pp. 141–151.

[6] Cveticanin, L., 1984, ''Vibrations of a Textile Machine Rotor,'' J. Sound Vib., **97**, pp. 181–187.

[7] Cveticanin, L., 1989, ''Stability of a Clamped-free Rotor with Variable mass for the Case of Radial Rubbing,'' J. Sound Vib., **129**, pp. 489–499.

[8] Cveticanin, L., 1992, ''The Influence of the Reactive Force on a Nonlinear Oscillation with Variable Parameters,'' J. Vibr. Acoust., **114**, 578–580.
[9] Cveticanin, L., 1993, ''The Influence of the Reactive Force on the Motion of the Rotor on which the Band is Winding Up,'' J. Sound Vib., **167**, pp. 382–384.
[10] Cveticanin, L., 1993, ''Conservation Laws in Systems with Variable Mass,'' J. Appl. Mech., **60**, pp. 954–958.
[11] Cveticanin, L., 1994, ''Some Conservation Laws for Orbits Involving Variable Mass and Linear Damping,'' J. Guid. Control Dyn., **17**, pp. 209–211.
[12] Cveticanin, L., 2001, ''Dynamic Buckling of a Single-degree-of-freedom System with Variable Mass,'' Eur. J. Mech. A/Solids, **20**, pp. 661–672.
[13] Ge, Z.-M., 1982, ''Extended Kane's Equations for Nonholonomic Variable Mass System,'' J. Appl. Mech., **49**, pp. 429–431.
[14] Ge, Z.-M., 1984, ''The Equations of Motion of Nonlinear Nonholonomic Variable Mass System with Applications,'' J. Appl. Mech., **51**, pp. 435–437.
[15] Korobkin, A. A., and Pukhnachov, V. V., 1988, ''Initial Stage of Water Impact,'' Annu. Rev. Fluid Mech., **20**, pp. 159–185.
[16] Cooker, M. J., and Peregrine, D. H., 1995, ''Pressure-impulse Theory for Liquid Impact Problems,'' J. Fluid Mech., **297**, pp. 193–214.
[17] Molin, B., Cointe, R., and Fontaine, E., 1996, ''On Energy Arguments Applied to the Slamming Force,'' *11th Int. Workshop on Water Waves and Floating Bodies*, Hamburg.
[18] Agostinelli, C., 1936, ''Sui Sistemi Dinamici di Masse Variabili,'' Atti della Reale Accademia delle Scienze di Torino, Classe di Scienze Fisiche, Matemat. Naturali, **71**, I, pp. 254–272.
[19] Targ, S., 1976, *Theoretical Mechanics, A Short Course*, Mir Publishers, Moscow, English Translation of 2nd ed., 525 pp.
[20] Pars, L. A., 1965, *A Treatise on Analytical Dynamics* Heinemann, London, 641 pp.
[21] Newman, J. N., 1978, *Marine Hydrodynamics*, The MIT Press, 402 pp.
[22] Lamb, H., 1932, *Hydrodynamics*, Dover Publications, N.Y., 6th ed., 738 pp.
[23] Faltinsen, O. M., and Zhao, R., 1997, ''Water Entry of Ship Sections and Axi-symmetric Bodies,'' *Agard Ukraine Inst. on Hydromechanics Workshop on High Speed Body Motion in Water*.
[24] Cointe, R., and Armand, J. L., 1987, ''Hydrodynamic Impact Analysis of a Cylinder,'' ASME J. Offshore Mech. Arct. Eng. , **109**, pp. 237–243.
[25] Wu, G. X., 1998, ''Hydrodynamic Force on a Rigid Body During Impact with Liquid,'' J. Fluids Struct., **12**, pp. 549–559.
[26] Miloh, T., 1981, ''Wave Slam on a Sphere Penetrating a Free Surface,'' J. Eng. Math., **15**, pp. 221–240.
[27] Seliger, R. L., and Whitham, G. B., 1968, ''Variational Principles in Continuum Mechanics,'' Proc. R. Soc. London, Ser. A, **305**, pp. 1–25.

Approximate Model for a Viscoelastic Oscillator

Y. Ketema
Department of Aerospace Engineering and Mechanics,
University of Minnesota,
Minneapolis, MN 55455

An oscillator where the restoring force is furnished by a viscoelastic bar and therefore depends on the history of the motion is considered. The history-dependent force is characterized by a relaxation modulus and a relaxation time. Assuming that the relaxation time is small, an approximate model for the oscillator is derived. This model is then linearized for the study of small vibrations. It is shown that the viscoelastic force, in addition to viscous damping, effects an apparent decrease in mass that modifies the natural frequency of the linear oscillator. The temperature dependence of the relaxation time, and consequently the frequency shift, is studied. [DOI: 10.1115/1.1607355]

1 Introduction

The application of viscoelastic materials in vibration damping and control is a subject of increasing interest. Examples of such applications are found in vibration damping of flexible structures such as beams, plates, shells, etc., where the viscoelastic material is made to vibrate with the structural member, for example, in the form of a layer attached to a beam or a plate (see, e.g., [1–3]).

It is common to model the dynamics of viscoelastic materials, especially in linear problems, through the use of complex moduli. A more general model of viscoelastic materials is the material-with-memory model. In this model the stress at a point in the material not only depends on the current state of deformation but also on the history of the deformation. Such materials are termed *history type materials,* or *materials with memory,* [4]. Examples of history type materials include gum rubber, silicone gel, and high polymer solutions.

When the relative deformations of the recent past are more important than those further back in time in determining the history-dependent force, the material is said to have "fading memory." A rigorous treatise on the properties of such materials has been given by Coleman [5].

In the simplest case where the history-type model is used to describe viscoelastic materials, the stress in the material is characterized by two material parameters: the relaxation time γ, which determines the rate at which the influence of past states of strain on present stress diminishes with elapsed time, and the relaxation modulus G_0, which determines the overall strength of the history dependence of the stress. More generally, history-type materials may exhibit a spectrum of relaxation times that influence the behavior of the material simultaneously. In addition, the parameters γ and G_0 themselves may be sensitive to outside effects such as change in temperature. This temperature dependence may lead to undesirable changes in mechanical behavior in the case of uncontrolled temperature variations. However, it also affords the technologically important possibility of "tuning" the viscoelastic parameters to desired values through induced temperature changes in a specific viscoelastic element.

In a previous paper, [6], the properties of a linear oscillator subject to history-type restoring forces that have elastic and dissipative parts was studied. It was shown that such forces give the linear oscillator force-amplification and resonance properties that are very different from those of a conventional oscillator with viscous damping. Thus, for example, the oscillator possesses two resonance frequencies in the range of variation for the relaxation modulus. This property has since been used in the study of vibration absorbers with adaptable frequencies of operation (see [7,8]).

A one degree-of-freedom oscillator where both the elastic and history-dependent dissipative forces are nonlinear was considered in [9]. As in the case of the linear oscillator with history-dependent restoring forces, the resulting dynamical system may be studied in a three-dimensional phase space. The problem of the preservation of Hamiltonian orbits was addressed and the results were used in a new method of measurement of the relaxation time and relaxation modulus for the viscoelastic material.

In the present paper, an oscillator that is subject to history-type restoring forces is considered under the assumption that the relaxation time γ is small in comparison to a characteristic time of the motion of the oscillator. Based on this assumption an approximate model that gives the dissipative forces in an algebraic form, i.e., in terms of the displacement and velocity, is derived. In this model the resulting dynamical system has a two-dimensional phase space and is therefore simpler to analyze than the full history-dependent model. The approximate model studied in this paper is equivalent to what would be obtained in accordance to [10], though the method of approximation is slightly different.

The rest of this paper is organized as follows: In Section 2 the full history-dependent model is introduced. The approximate model, which allows for the study of the special case of small relaxation times, is studied in Section 3. Lastly, in Section 4, the dependence of the natural frequency of the oscillator on the relaxation time γ, the temperature dependence of γ, and the resulting temperature dependence of the natural frequency is studied.

2 A One-Degree-of-Freedom Oscillator With Memory

Consider a one-degree-of-freedom oscillator in which the restoring force is furnished by a viscoelastic bar whose motion is restricted to be in the direction of motion of the oscillating mass (Fig. 1).

To describe the dynamics of the viscoelatic bar, let $x=\chi(X,t)$ denote the position at the present time t of a particle of the viscoelastic bar which is at $X=\chi(X,0)$ in its undistorted natural state at time $t=0$. Then the history of the motion is represented by $\chi(X,t-s)$, $\forall s \geqslant 0$. If we now let $F \equiv \partial\chi/\partial X$ denote the deformation gradient, the relative deformation gradient history is given by $F(X,t-s)/F(X,t)$, $\forall s \geqslant 0$. The relative history is then characterized by

$$J_t(X,t-s)=\frac{[F(X,t-s)]^2}{[F(X,t)]^2}-1, \quad \forall s \geqslant 0. \tag{1}$$

The constitutive response function for determining the present value of the axial force $f(X,t)$ on the particle X in the viscoelastic bar is assumed to be of the finite-linear form

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 20, 2000; final revision, Apr. 24, 2003. Associate Editor: B. M. Moran. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

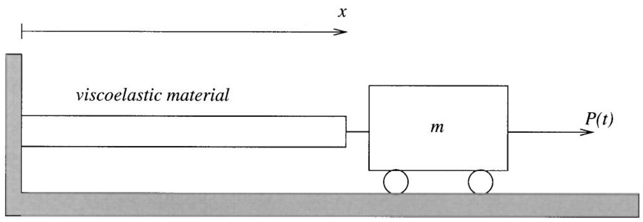


Fig. 1 Schematic diagram of a linear oscillator with history-dependent forces

$$f(X,t)=\overline{f^e}(F(X,t))+\int_0^\infty G(s)J_t(X,t-s)ds, \tag{2}$$

where $\overline{f^e}(\cdot)$ denotes the elastic response function and $G(\cdot)$ is the viscoelastic relaxation kernel for the material.

Assuming that the motion of the viscoelastic bar is homogeneous, we write

$$x=\chi(X,t)=X\tilde{x}(t). \tag{3}$$

Then, the deformation gradient $F(X,t)=\partial\chi(X,t)/\partial X=\tilde{x}(t)$ is, in fact, the homogeneous "stretch" of material filaments so that

$$\overline{f^e}(F(X,t))=\overline{f^e}(\tilde{x}(t)), \quad J_t(X,t-s)=\frac{\tilde{x}^2(t-s)-\tilde{x}^2(t)}{\tilde{x}^2(t)}. \tag{4}$$

We shall assume that $G(\cdot)$ is given by the exponentially decaying relaxation function

$$G(s)\equiv G_0e^{-s/\gamma}, \quad \forall s\geqslant 0, \tag{5}$$

where $G_0>0$ and the relaxation time $\gamma>0$. If we let L_0 denote the referential length of the viscoelastic bar, then the dynamical equation for the mass m is

$$m\ddot{\chi}(L_0,t)=p(t)-f(L_0,t), \tag{6}$$

which with (1), (3), and (4) can be rewritten as

$$mL_0\ddot{\tilde{x}}(t)=-\overline{f^e}(\tilde{x}(t))+\frac{G_0}{\gamma}\int_0^\infty e^{-s/\gamma}\frac{\tilde{x}^2(t-s)-\tilde{x}^2(t)}{\tilde{x}^2(t)}ds$$

$$+P\cos\Omega t, \tag{7}$$

where we have introduced the special forcing function $p(t)=P\cos\Omega t$.

It is convenient to rewrite (7) as a system of first-order ordinary differential equations, and to do this we define the auxiliary function

$$\tilde{\zeta}(t)=\int_0^\infty e^{-s/\gamma}\frac{\tilde{x}^2(t-s)-\tilde{x}^2(t)}{\tilde{x}^2(t)}ds. \tag{8}$$

Then, it readily follows that (7) has the equivalent form

$$\dot{\tilde{x}}(t)=\tilde{\eta}(t),$$

$$mL_0\dot{\tilde{\eta}}(t)=-\overline{f^e}(\tilde{x}(t))+\frac{G_0}{\gamma}\tilde{\zeta}(t)+P\cos\Omega t, \tag{9}$$

$$\dot{\tilde{\zeta}}(t)=-\left[\frac{1}{\gamma}+\frac{2\tilde{\eta}(t)}{\tilde{x}(t)}\right]\tilde{\zeta}(t)-\gamma\frac{2\tilde{\eta}(t)}{\tilde{x}(t)}.$$

Clearly, an equilibrium point $(\tilde{x}^*,\tilde{\eta}^*,\tilde{\zeta}^*)$ for (9) (i.e., when $P=0$) is given by $(\tilde{x}^*,0,0)$, where $\tilde{x}^*$ is any root of the equation $\overline{f^e}(\tilde{x}^*)=0$.

In order to bring the system (9) into a dimensionless form we first introduce τ through

$$\tau=\omega_0 t, \tag{10}$$

where ω_0 is the natural frequency of small oscillations about the equilibrium point $\tilde{x}^*$ of (9), and is given by

$$\omega_0=\sqrt{\frac{k}{mL_0}}, \tag{11}$$

with

$$k=\frac{d\overline{f^e}(\tilde{x}^*)}{d\tilde{x}}>0. \tag{12}$$

Now, let $\tilde{\xi}(t)$ such that $\tilde{x}(t)=\tilde{x}^*+\tilde{\xi}(t)$, and introduce

$$\xi(\tau)\equiv\tilde{\xi}\left(\frac{\tau}{\omega_0}\right), \quad \eta(\tau)\equiv\frac{1}{\omega_0}\tilde{\eta}\left(\frac{\tau}{\omega_0}\right), \quad \zeta(\tau)\equiv\tilde{\zeta}\left(\frac{\tau}{\omega_0}\right). \tag{13}$$

Then, (9) will take the form

$$\frac{d\xi(\tau)}{d\tau}=\eta(\tau),$$

$$\frac{d\eta(\tau)}{d\tau}=-f(\xi(\tau))+\Phi_0\zeta(\tau)+p\cos\omega\tau, \tag{14}$$

$$\frac{d\zeta(\tau)}{d\tau}=-\left[\frac{1}{\varepsilon\nu}+\frac{2\eta(\tau)}{\tilde{x}^*+\xi(\tau)}\right]\zeta(\tau)-\frac{2\eta(\tau)}{\tilde{x}^*+\xi(\tau)},$$

where

$$f(\xi)=\frac{\overline{f^e}(\tilde{x}^*+\xi)}{k}, \quad \Phi_0=\frac{G_0}{k}, \quad p=\frac{P}{k} \quad \varepsilon\nu=\gamma\omega_0 \tag{15}$$

and

$$\omega=\frac{\Omega}{\omega_0}. \tag{16}$$

We assume that $0\leqslant\varepsilon\ll 1$ and that ν is $\mathcal{O}(1)$ thus that the relaxation time is small. The other parameters p and Φ_0 are assumed to be $O(1)$. In what follows, we derive a second-order approximation of (15), i.e., one that is accurate to $\mathcal{O}(\varepsilon)$, but that is in the simpler form of a single second-order differential equation. The main purpose of this analysis is to describe the dynamics of the system in terms of the well-known behavior of the linear harmonic oscillator.

3 A Second-Order Approximation

As is customary in regular perturbation theory (see, e.g., [11]), we begin by expanding $\zeta(\tau)$ in an asymptotic series in powers of ε:

$$\zeta(\tau)=\zeta_0(\tau)+\varepsilon\zeta_1(\tau)+\varepsilon^2\zeta_2(\tau)+\mathcal{O}(\varepsilon^3). \tag{17}$$

Substituting this expression into the third equation of (14) and collecting terms of the same order in ε gives

$$\varepsilon^0: \quad \zeta_0(\tau)=0, \tag{18}$$

$$\varepsilon^1: \quad \zeta_1(\tau)=-2\nu\frac{\eta(\tau)}{\tilde{x}^*+\xi(\tau)}, \tag{19}$$

$$\varepsilon^2: \quad \dot{\zeta}_1(\tau)=-\frac{1}{\nu}\zeta_2(\tau)-2\frac{\eta(\tau)\zeta_1(\tau)}{\tilde{x}^*+\xi(\tau)}. \tag{20}$$

Now, using (19) in (20) and solving for $\zeta_2(\tau)$, we have

$$\zeta_2(\tau)=\frac{2\nu^2}{\tilde{x}^*+\xi(\tau)}\left(\dot{\eta}(\tau)+\frac{\eta(\tau)^2}{\tilde{x}^*+\xi(\tau)}\right). \tag{21}$$

We can therefore write

$$\zeta(\tau)=-2\varepsilon\nu\frac{\eta(\tau)}{\tilde{x}^*+\xi(\tau)}+\varepsilon^2\frac{2\nu^2}{\tilde{x}^*+\xi(\tau)}\left(\dot{\eta}(\tau)+\frac{\eta(\tau)^2}{\tilde{x}^*+\xi(\tau)}\right). \tag{22}$$

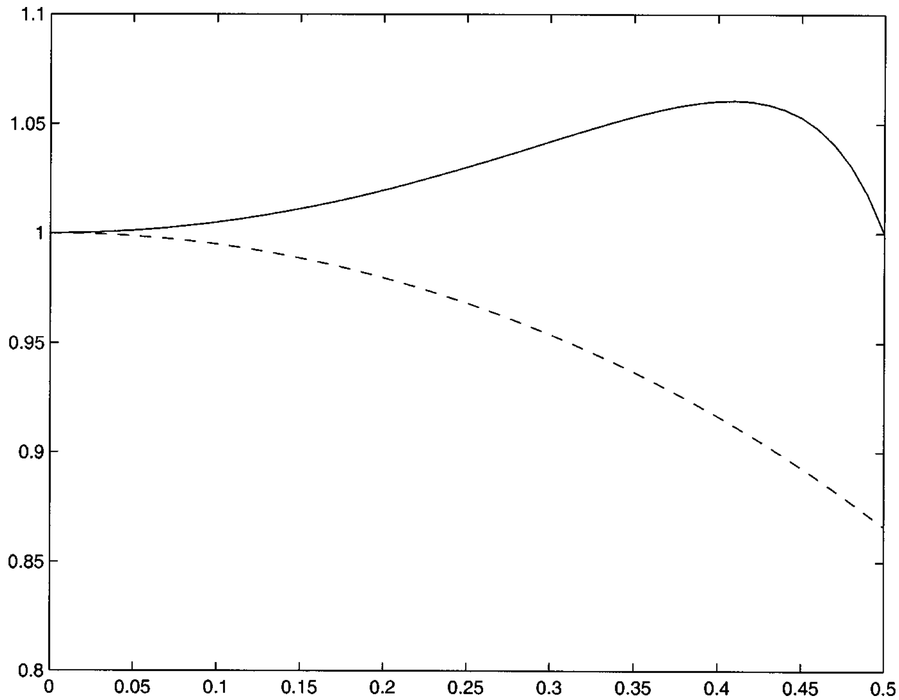


Fig. 2 Damped natural frequency of a viscoelastic oscillator as a function of the relaxation time: (*a*) second-order approximation (solid line), (*b*) first-order approximation, (dashed line). $\Phi_0=1$, $\tilde{x}^*=1$.

This last expression may be substituted into the second equation of (14), and results in

$$\mu(\xi(\tau))\ddot{\xi}(\tau)=-f(\xi(\tau))-2\varepsilon\nu\Phi_0\frac{\eta(\tau)}{\tilde{x}^*+\xi(\tau)}+2\varepsilon^2\nu^2\Phi_0\frac{\eta(\tau)^2}{(\tilde{x}^*+\xi(\tau))^2}+p\cos\omega t \qquad (23)$$

where

$$\mu(\xi(\tau))=1-\frac{2\varepsilon^2\nu^2\Phi_0}{\tilde{x}^*+\xi(\tau)}. \qquad (24)$$

3.1 The Linear Viscoelastic Oscillator. Assuming small vibrations about the equilibrium position $\xi=0$ one can linearize (23) to obtain the equation

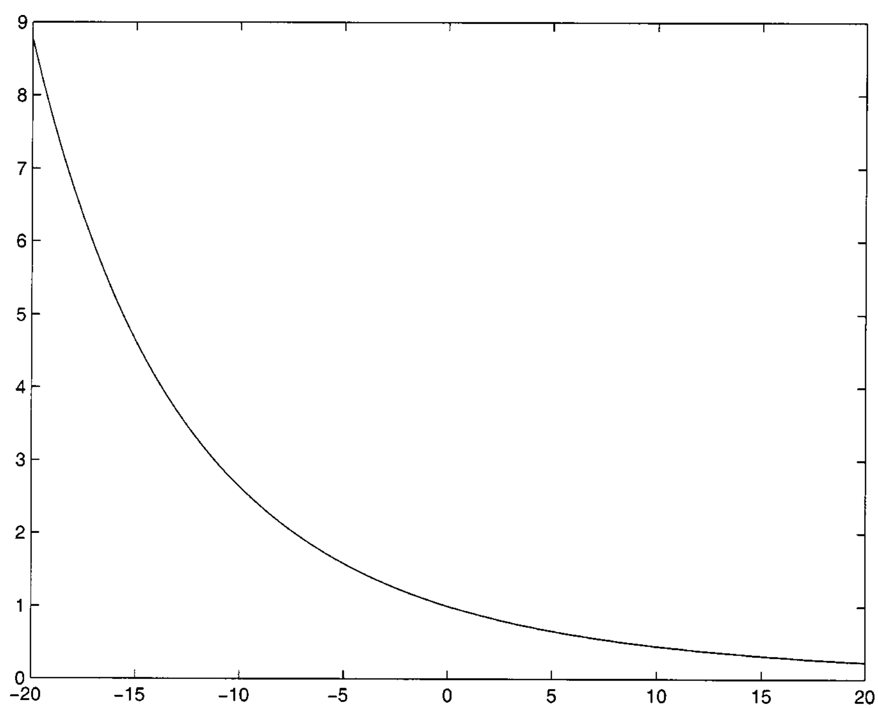


Fig. 3 The variation of the normalized relaxation time with temperature

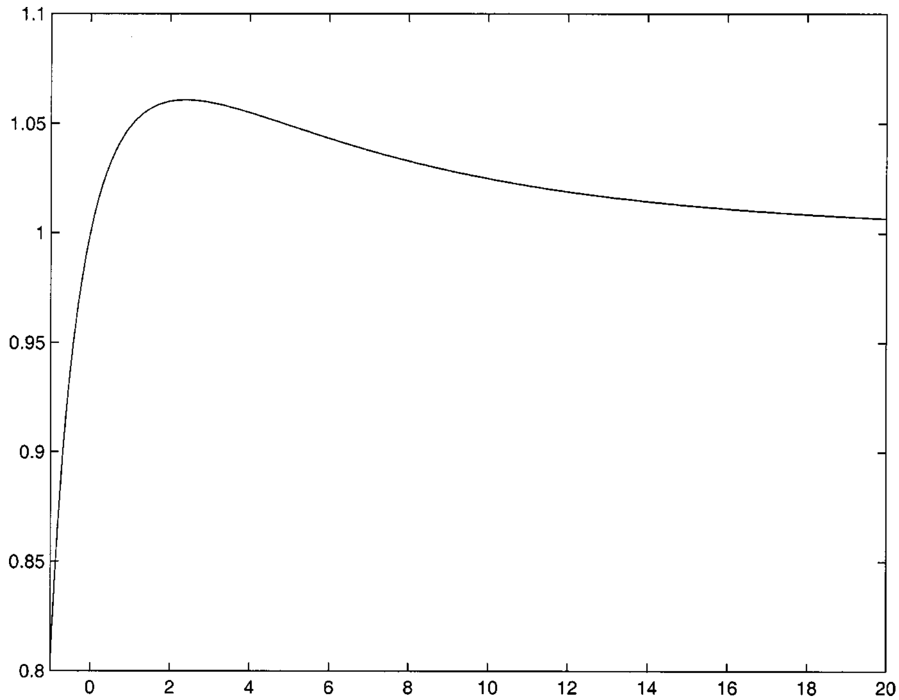


Fig. 4 The variation of the normalized damped natural frequency with temperature

$$\mu_0 \ddot{\xi}(\tau) = -\xi(\tau) - 2\varepsilon\nu \frac{\Phi_0}{\tilde{x}^*} \eta(\tau) + p \cos \omega t, \tag{25}$$

where

$$\mu_0 = 1 - 2\varepsilon^2 \nu^2 \frac{\Phi_0}{\tilde{x}^*}. \tag{26}$$

Note that (25) has the form of the equation of motion for an oscillator where the (scaled) mass is μ_0, i.e., a mass that is smaller than the physical mass of the oscillator. In the case where Φ_0 is large the "subtracted mass" in (24) and its effect on the dynamics may be of significance.

Consider, for example, the damped natural frequency for the linear oscillator, ω_d. This is given by

$$\omega_d = \sqrt{\frac{1}{\mu_0} - \frac{1}{\mu_0^2} \left(\frac{\varepsilon\nu\Phi_0}{\tilde{x}^*} \right)^2} \tag{27}$$

and therefore depends on μ_0. Figure 2 shows ω_d as a function of $\gamma\omega_0 = \varepsilon\nu$ (solid line). At $\gamma\omega_0 \approx 0.4$ the natural frequency has a maximum. This is where the tendency of the natural frequency to increase due to the decrease of the effective mass μ_0 is balanced by its tendency to decrease due to the increasing damping. For comparison, the damped natural frequency of a damped linear oscillator (with no second-order effect of the subtracted mass) is shown in the same figure (dashed line). For both cases it is assumed $\Phi_0 = \tilde{x}^* = 1$.

4 Concluding Remarks: Thermal Effects

Viscoelastic materials are important components in many dynamical applications (including structural applications) where their damping characteristics are put to use. It is therefore crucial to understand all aspects of the dynamics that they introduce to a mechanical system. In this paper, it has been shown that in the case of small relaxation times and large relaxation moduli, a significant shift in the damped natural frequency of the system may arise. The strong dependence of the relaxation time on temperature is a factor that makes this frequency shift of practical importance.

Consider, for example, polymers for which the William-Landal-Ferry formula (see, e.g., [12,13]) gives

$$\gamma(T) = \gamma_0 e^{\psi(T)}, \quad \psi(T) = \frac{-c_1(T - T_s)}{c_2 + T - T_s}, \tag{28}$$

where the material constants c_1 and c_2 are positive and depend on the reference temperature T_s, and where γ_0 is the value of γ at $T = T_s$. For rubbers c_1 and c_2 take on the values of approximately -8.86 and 101.6, respectively, the reference temperature T_s lying roughly in the range 200 K–300 K. Based on these values (with $T_s = 250$ K). Figure 3 shows the variation of the normalized relaxation time γ/γ_0 with temperature. It is evident from the figure that relatively small changes in temperature can result in significant changes in the normalized relaxation time.

To estimate the corresponding change in damped natural frequency, we begin by noting that for any given temperature difference $T - T_s$, the ratio of the scaled relaxation time to its corresponding value at the reference temperature T_s, i.e., the fraction ν/ν_0, can be calculated using (28) simply by replacing γ and γ_0 by $\varepsilon\nu$ and $\varepsilon\nu_0$, respectively. Then, by making use of (27) the damped natural frequency can be calculated for any $T - T_s$. Figure 4 shows ω_d as a function of $T - T_s$ with the assumption that $\varepsilon\nu = \varepsilon\nu_0 = 0.5$ at $T = T_s = 250$ K.

References

[1] Okazaki, A., Urata, Y., and Tatemichi, A., 1990, "Damping Properties of a Three Layered Shallow Spherical Shell With a Constrained Viscoelastic Layer," JSME Int. J., Ser. I, **33**(2), pp. 145–151.

[2] Gautham, B. P., and Ganesan, N., 1994, "Vibration and Damping Characteristics of Spherical Shells With a Viscoelastic Core," J. Sound Vib., **170**(**3**), pp. 289–301.

[3] Culkowski, P. M., and Reismann, H., 1971, "The Spherical Sandwich Shell Under Axisymmetric Static and Dynamic Loading," J. Sound Vib., **14**, pp. 229–240.

[4] Truesdell, C., and Noll, W., 1965, "The Non-Linear Field Theories of Mechanics," *Handbook of Physics*, III/3, Springer, New York.

[5] Coleman, B. D., 1964, "Thermodynamics of Materials With Memory," Arch. Ration. Mech. Anal., **17**(1), pp. 1–46.

[6] Fosdick, R. L., Ketema, Y., and Yu, J. H., 1998, "Vibration Damping Through the Use of Materials With Memory," Int. J. Solids Struct., **35**, pp. 403–420.

[7] Fosdick, R. L., and Ketema, Y., 1998, ''A Thermoviscoelastic Dynamic Vibration Absorber,'' J. Appl. Mech., **65**, pp. 17–24.
[8] Ketema, Y., 1998, ''A Viscoelastic Dynamic Vibration Absorber With Adaptable Suppression Band: A Feasibility Study,'' J. Sound Vib., **216**(1), pp. 133–145.
[9] Fosdick, R. J., Ketema, Y., and Yu, J. H., 1998, ''A Nonlinear Oscillator With History Dependent Forces,'' Int. J. Non-Linear Mech., **33**, pp. 447–459.
[10] Coleman, B. D., and Noll, W., 1960, ''An Approximation Theorem for Functionals, With Applications in Continuum Mechanics,'' Arch. Ration. Mech. Anal., **6**, pp. 355–370.
[11] Nayfeh, A. H., 1973, *Perturbation Methods*, John Wiley and Sons, New York.
[12] Ferry, J. D., 1970, *Viscoelastic Properties of Polymers*, 2nd Ed., John Wiley and Sons, New York.
[13] Moore, D. F., 1993, *Viscoelastic Machine Elements*, Butterworth-Heineman Ltd., Oxford.

Flow Control Using Rotating Cylinders: Effect of Gap

S. Mittal
Department of Aerospace Engineering,
Indian Institute of Technology,
Kanpur, UP 208 016,
India
e-mail: smittal@iitk.ac.in

The flow past a bluff body can be controlled significantly by placing small rotating cylinders at appropriate locations. Computational results for control of $Re = 10^4$ flow past a circular cylinder are presented. Two control cylinders of one-twentieth the diameter of the main cylinder rotate at a rate such that their tip speed is five times the free-stream speed of the flow. Computations are carried out for various values of the gap between the main and control cylinders. A stabilized finite element method is utilized to solve the incompressible Navier-Stokes equations in the primitive variables formulation. A gap value of one-tenth the diameter of the main cylinder is found to be close to the optimal value. Compared to the flow past an isolated cylinder a very significant reduction in the drag and unsteady forces is observed for the flow with control. [DOI: 10.1115/1.1601250]

1 Introduction

Various flow control techniques that result in reduction of drag and unsteady forces have been suggested and tested in the past. For example, the review article by Gad-el-Hak and Bushnell [1] discusses some of them. One such method employed for separation control is the moving-surface boundary layer control (MSBC). Rotating cylinder elements are employed to inject momentum into the already existing boundary layer.

Modi and his co-workers have applied MSBC to reduce drag and unsteady forces on bluff bodies, for example, Modi [2], Modi et al. [3,4], and Munshi et al. [5,6]. The bluff bodies that they have studied include flat plates at various angles of attack, rectangular prisms, D sections, and tractor-trailer truck configurations. In all the cases, for high rotation rates of the control cylinders ($U_c/U \geqslant 4$, where U_c is the tip speed of the control cylinder and U is the free-stream speed of the flow) result in a narrowing of the wake, delay of separation, and a significant reduction in the drag. The method has also been demonstrated to reduce flow induced vibrations. Another effort by Choi and Choi [7] utilizes a belt mounted on a cylinder that moves due to the shear stress acting on the wall. Computations carried out at Re=100 indicate that up to 11% reduction in drag is obtained for a belt installed between 112.5 deg and 135 deg from the stagnation point. Another computational study by Park et al. [8] utilized a pair of blowing and suction slots located at ±110 deg from the leading edge stagnation point. Complete suppression is achieved for the Re=60 flow.

The degree of flow control depends on various parameters. Some of them include the number of control cylinders, their diameter relative to the characteristic dimension of the bluff body of interest, their speed of rotation and the gap between the bluff body and the control cylinder(s). The rotating control cylinders generate circulation and inject momentum from the outer flow into the wake of the main cylinder. It is expected that, if the gap between the main and control cylinders is too large, even though the control cylinders rotating at high speeds may generate enough circulation, the effect on the main cylinders will be small. On the other hand, if the gap is too small, the control cylinders will not be able to generate significant circulation to achieve good flow control. Therefore the gap is an important parameter in obtaining the optimal performance of the flow control system.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, August 29, 2001; final revision, February 11, 2003. Associate Editor: T. E. Tezduyar. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Some finite element simulations for flow control using rotating cylinders have been presented by Mittal [9]. Computations were reported for two values of the gap (=0.01D and 0.075D) and for Re=100 and 10^4. It was observed that $U_c/U=5$ results in a steady flow for Re=100. At Re=10^4, even though the flow remains unsteady, the wake is highly organized and narrower compared to the one without control. In all the cases, a significant reduction in the overall drag coefficient and the unsteady aerodynamic forces acting on the body is observed. The effect of the gap is found to be more critical for the Re=10^4 flows compared to that at lower Reynolds numbers. It is found that the coefficient of power required for rotating the control cylinders is significant for low Reynolds numbers but small for relatively higher Reynolds number (10^4) flows.

Flow past a translating and rotating cylinder is one of the most vital components of the flow control technique being studied in the present work. A detailed investigation of the flow past an isolated spinning cylinder and a review of the work by other researchers was presented by Mittal [9]. Their results show that for, rapidly spinning cylinders, the end conditions and the aspect ratio of the cylinder have a significant effect on the flow. The end-condition corresponding to a no-slip wall results in loss of lift and increase in drag owing to certain centrifugal instabilities. The results from two-dimensional computations approach those in three dimensions when the aspect ratio of the cylinder is large. This observation is in line with that of Tokumaru and Dimotakis [10]. It was shown in an earlier study ([11]) that the two-dimensional (2D) flow is stable to 2D disturbances for large rotation rates. This was established by computing flows past an eccentrically spinning cylinder.

The present work is a sequel to that reported in our earlier work ([9]). Flow control past a circular cylinder using rotating control cylinders, of much smaller diameter, is studied numerically. The rotation rate is fixed at $U_c/U=5$ and the Reynolds number of the flow is 10^4. The effect of the gap between the main and control cylinders is investigated. The computations are restricted to two dimensions. These results are expected to simulate the situations where the aspect ratios of the main and control cylinders are large. Detailed results, including time histories of the aerodynamic coefficients for forces acting on the bodies and power required, are presented.

Stabilized finite element formulations are utilized to solve the viscous incompressible flow equations using primitive variables. The large-scale equation systems that result from the finite element discretization of the flow problem are solved using iterative solution techniques. The methods that are being used for the numerical simulations are well proven and have been used earlier to solve a variety of flow problems. Some of the applications can be

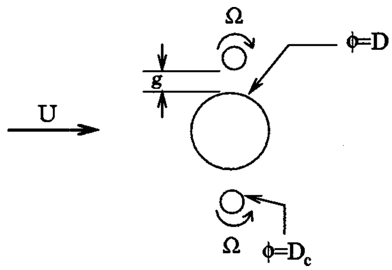


Fig. 1 Description of the relative location of the main and control cylinders

found in the articles by Mittal and Tezduyar [12], Mittal and Kumar [13], and Mittal et al. [14]. The method has also been employed by Mittal and Raghuvanshi [15] to investigate the effect of placing a small stationary cylinder close to the main cylinder. It was found that for certain locations of the control cylinder and for small Reynolds numbers the vortex shedding is completely suppressed.

The outline of the rest of the article is as follows. We begin by giving the problem description along with a schematic of the flow control device in Sec. 2. The SUPG (streamline-upwind/Petrov-Galerkin) and PSPG (pressure-stabilizing/Petrov-Galerkin) stabilization techniques are employed to stabilize our computations against spurious numerical oscillations and to enable us to use equal-order-interpolation velocity-pressure elements. Details on these techniques can be found in the article by Tezduyar et al. [16]. In Sec. 3 computational results for flows involving the main and control cylinders are presented and discussed. Finally, a few concluding remarks are made in Sec. 4.

2 Problem Description

Figure 1 shows the schematic of the typical arrangement of the setup in the present work. Two control cylinders of diameter D_c, each, are placed close to the main cylinder of diameter D. The control cylinders are placed close to the shoulder of the main cylinder. The line joining the centers of the three cylinders is normal to the free-stream flow. The upper cylinder rotates in the clockwise and the lower one in the anticlockwise direction. The rotation rates of both the cylinders is Ω. The nondimensional rotational rate is defined as $\alpha = U_c/U$ where U_c $(=D_c\Omega/2)$ is the tip speed of the rotating cylinders and U is the free-stream speed. All the results presented in this article are with $D/D_c=20$ and $\alpha=5$. The Reynolds number based on the diameter of the main cylinder (D), free-stream velocity and the kinematic viscosity of the fluid is 10^4. The gap between the main and control cylinders is g, i.e., the distance between the centers of the main and control cylinders is $g+D/2+D_c/2$. The objective of the present study is to investigate the effect of gap on the control effectiveness.

Power is needed to sustain both the translational and rotational motion of the main and control cylinders $(C_P=C_P^{Rot}+C_P^{Trans})$. The power coefficient for translatory motion, C_P^{Trans}, is simply the sum of the drag coefficients for the main and control cylinders, i.e., $C_P^{Trans}=C_{D1}+C_{D2}+C_{D3}$. The subscript 1 is for the main cylinder while 2 and 3 are for the two control cylinders. The power coefficient due to the rotary motion of the two control cylinders is $C_P^{Rot}=(C_{M1}+C_{M2})(D/D_c)(U_c/U)$. Here, C_M is the moment coefficient acting on the center of the cylinder. All the results presented in this article are with respect to nondimensional time $\tau=Ut/D$, where t is the actual time.

3 Numerical Simulations

The cylinders reside in a rectangular domain whose upstream and downstream boundaries are located at 8 and 30 cylinder diameters, respectively, from the center of the main cylinder. The upper and lower boundaries are placed at 8 diameters, each, from the center of the main cylinder. The no-slip condition is specified for the velocity on the surface of the cylinders and free-stream values are assigned for the velocity at the upstream boundary. At the downstream boundary, a Neumann-type boundary condition for the velocity is specified that corresponds to zero viscous stress vector. On the upper and lower boundaries, the component of velocity normal to and the component of stress vector along these boundaries is prescribed zero value.

All the values for the lift and drag coefficients and the Strouhal number, reported in this article, have been nondimensionalized with respect to the diameter of the main cylinder, D, and the free-stream speed, U. The quantities with suffix "1" refer to the main cylinder while the ones with "2" and "3" correspond to the upper and lower control cylinders, respectively.

The computations presented in this article have been carried out without the use of a turbulence model even though the Re is large enough for the wake to show significant turbulence. RANS (Reynolds averaged Navier-Stokes) equations is not a good option for the present problem as the flows we are dealing with are inherently unsteady and the turbulence model may interfere with the physics of the flow. Another option is to carry out large eddy simulation (LES) by employing a subgrid scale model, for example, a Smagorinsky turbulence model. It has been shown, recently, by Akin et al. [17], via test problems, that in most of the flow domain the numerical viscosity generated by the SUPG stabilization, in terms of its maximum value in the flow direction, is much larger than the eddy viscosity due to a Smagorinsky turbulence model. Similar observations have been made by Mittal and Moin [18] for large eddy simulation (LES) past a cylinder at Re $=3900$ using a higher order upwind-biased finite difference schemes. The mean velocity profiles from computations with and without the subgrid scale model did not shown any significant difference. However, the one-dimensional spectrum of velocity at a downstream location reveals that the numerical viscosity removes substantial energy from the high wave number regime.

3.1 The Finite Element Mesh. The flows that are being computed here involve interaction of wakes of the main and control cylinders. The disparity between the geometric scales of the cylinders is expected to result in vortical structures of widely varying length scales. The value of Re based on the diameter of the control cylinders and free-stream speed is 500. The flow for no rotation of the control cylinders is associated with vortex shedding from the main as well as control cylinders. Details of the flow for two values of gap can be seen in Mittal [9]. At high rotation rates of the control cylinders, a thin region of high speed fluid is expected to surround each one of them. The boundary/shear layers for such flows are extremely thin. The finite element mesh utilized in the present study is fine enough to resolve all the details of the flow. A structured finite-element mesh is used close to the three cylinders. An unstructured mesh is generated using the Delaunay technique in the rest of the domain via an automatic mesh generator. Mesh convergence studies have been carried out for these flow problems to assess the adequacy of the resolution for $g=0.01D$, $g=0.075D$, and $0.10D$. With approximately twice the number of grid points and one-half the time step, the two solutions do not show any appreciable difference except for a slight increase (less than 2%) in the amplitude of the unsteady aerodynamic coefficients with the refined mesh.

For example, for the $g=0.075D$ case, the computations begin with a mesh consisting of 22,221 nodes and 43,558 triangular elements and a time step of 0.02. In the structured part of the mesh around the main cylinder there are 600 elements in the circumferential and 8 in the radial direction. The radial thickness of the band of structured elements is $0.025D$. The band of elements around the control cylinders has a thickness of $0.0025D$. It has 120 elements in the circumferential and 4 in the radial direction. The radial thickness of the elements close to the cylinders is 5

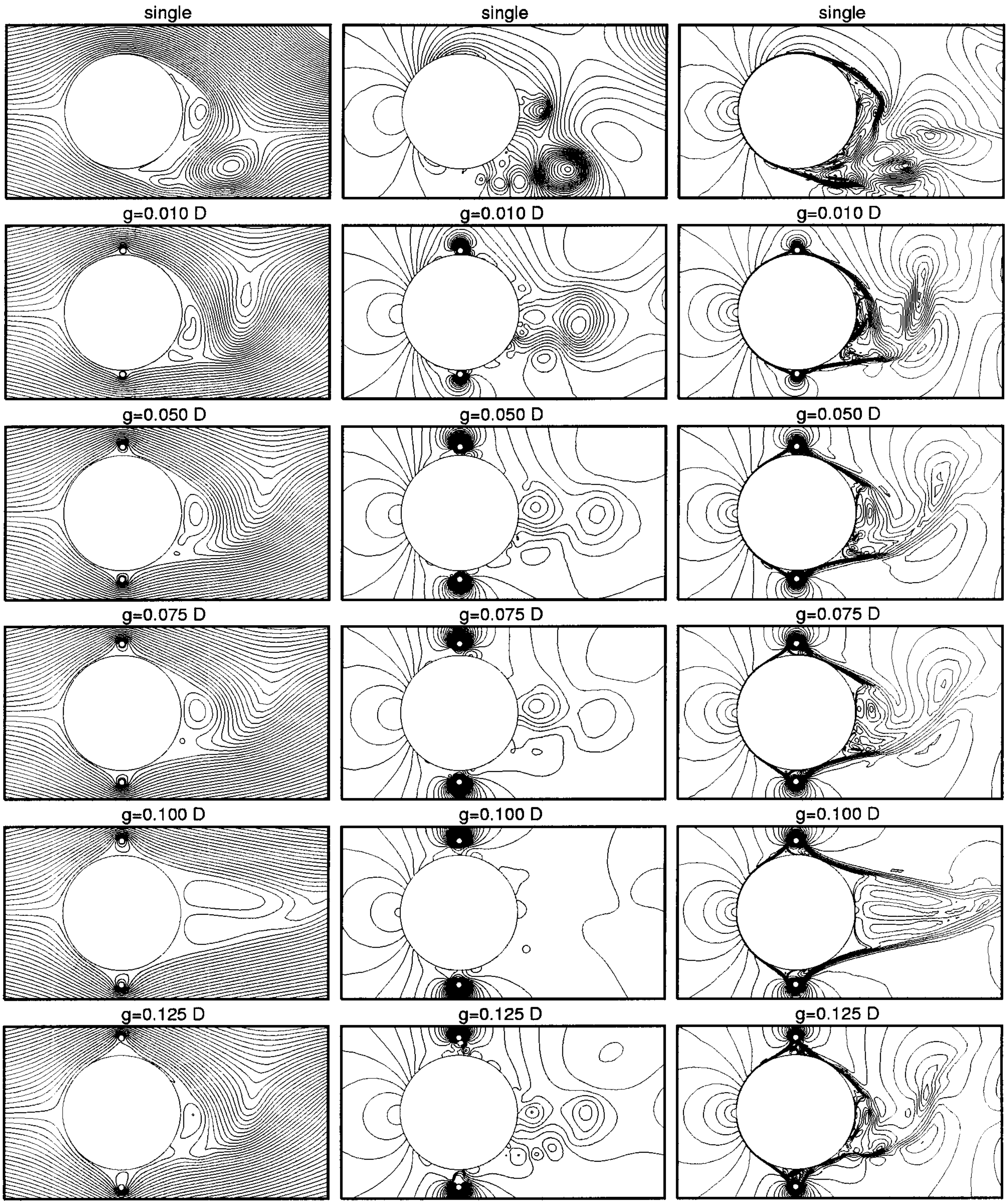


Fig. 2 Re=10^4 flow past main and control cylinders: stream function (left), pressure (middle), and magnitude of velocity (right) fields at a time instant corresponding to the peak value of the lift coefficient for the main cylinder

$\times 10^{-4}D$. To check the effect of the time step and spatial resolution, the solution is projected on a finer mesh with 27,029 nodes and 53,014 elements and computations are continued with a reduced time step of 0.01. The mesh with larger number of nodes is similar to the other one except for more refinement close to the control cylinders and the gap. In the regions close to the control cylinder it is, approximately, twice as refined as the earlier one. This mesh has 200 elements in the circumferential and 8 in the radial direction in the annulus consisting of the structured part of the mesh around the control cylinders. The first element thickness, in this case, is $5\times 10^{-5}D$. The solutions computed on the two meshes do not show any appreciable difference except for a slight increase (less than 2%) in the amplitude of the unsteady aerodynamic coefficients with the refined mesh.

The $g=0.10D$ case was initially computed with a mesh with 28,323 nodes and 55,602 elements. The time step used was 0.001. This mesh has 600 elements in the the circumferential and 8 in the radial direction. The radial thickness of the elements close to the cylinders is $1\times10^{-4}D$. The solution is projected on a finer mesh with 38,093 nodes and 75,142 elements and computations continued. Less than 1% difference was observed in the solutions from the two meshes.

The solution for the single cylinder (no control cylinders) was computed using a mesh with 47,011 nodes and 93,574 elements. The structured part of the mesh (an annulus of radial thickness $0.05D$) consists of 400 elements in the the circumferential and 40 in the radial direction. The radial thickness of the elements close to the cylinder is $1\times10^{-4}D$.

3.2 Flow Field for Various Values of Gap. Figure 2 shows the streamlines, pressure and velocity magnitude for various values of gap between the main and control cylinders. The pictures are taken at a time instant that corresponds to the peak value of the lift coefficient for the main cylinder. The solution for a single cylinder is also shown in the same figure. The flow past a single cylinder separates quite early resulting in a wake with large lateral width. This is expected for a subcritical flow. From Fig. 2, it is observed that the instantaneous flow shows an upward bias of the flow. This has been observed by other researchers in the past, as well. Behr [19] reports that for Re=2000 and larger, the vortex street oscillates about the centerline of the domain. The effect of this oscillation is also observed in the time histories of the lift and drag coefficient. The basic vortex shedding frequency is modulated with a lower secondary frequency that is, approximately, ten times the vortex shedding frequency. Our computations indicate that the time-averaged flow for several vortex shedding cycles results in an almost symmetric flow. The time-averaged drag coefficient is 1.8, approximately. This is on the higher side compared to the measurements from experiments. Computations were also carried out for Re=3900 flow. The time-averaged drag coefficient obtained for this flow is 1.79. This compares well with the results from Beaudean and Moin [20] who reported a value of 1.74 from their two-dimensional computations. Three-dimensional computations result in more realistic results. However, they require significantly larger computational resources. In the presence of control cylinders, the demand on computational resources is even higher. Therefore the present computations are restricted to two dimensions.

With the rotating control cylinders in place, the wake is much narrower and organized. In all the cases the control cylinders are associated with a set of closed streamlines near their surface. The rotation of the control cylinders causes the streamlines close to the stagnation streamline for the main cylinder to go over the control cylinders. The fluid that negotiates the control cylinders is given an increased momentum and is pumped into the wake of the main cylinder. The flow pictures suggest that $g=0.100D$ results in a wake with the lowest unsteadiness.

The velocity profiles for the isolated cylinder and for various gaps are shown in Fig. 3. The profiles on the upper surface of the control cylinder are quite similar in all the cases. The flow accelerates over the upper surface and achieves $u/U\sim6$ close to the

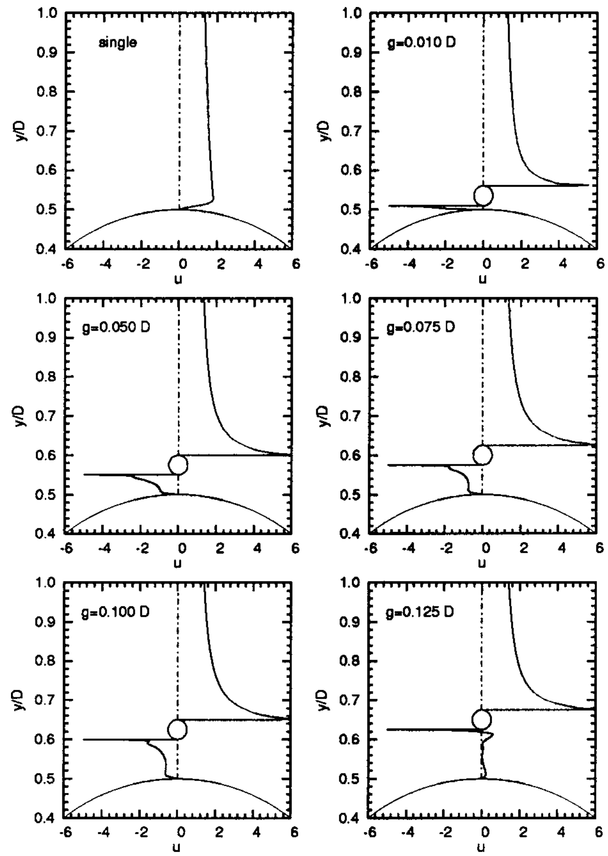


Fig. 3 Re=10^4 flow past main and control cylinders: variation of the *x* component of velocity in the gap region and close to the upper control cylinder at a time instant corresponding to the peak value of the lift coefficient for the main cylinder

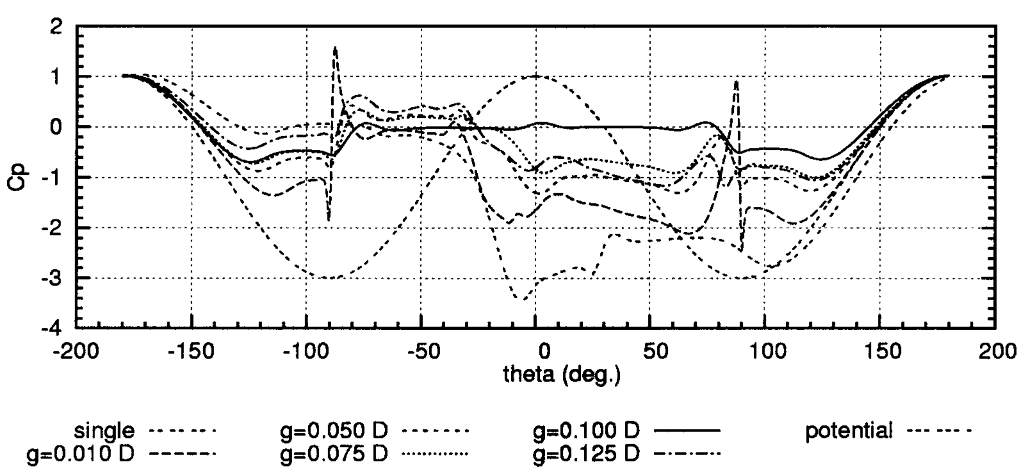


Fig. 4 Re=10^4 flow past main and control cylinders: variation of pressure coefficient on the surface of main cylinder at a time instant corresponding to the peak value of the lift coefficient

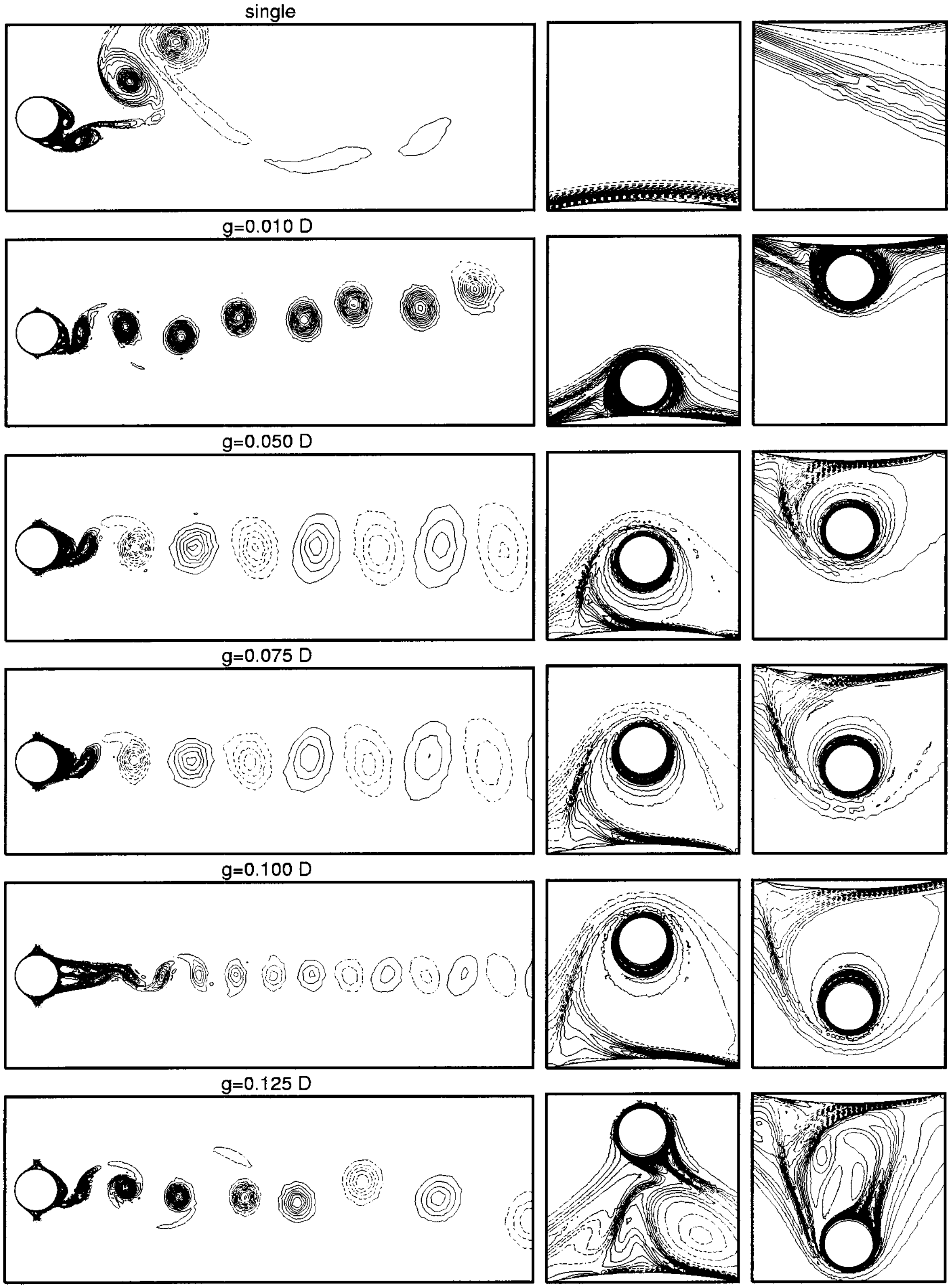


Fig. 5 Re=10^4 flow past main and control cylinders: the vorticity field and its close-up at a time instant corresponding to the peak value of the lift coefficient for the main cylinder. Clockwise vorticity is in broken lines while the counter clockwise vorticity is shown in solid lines.

control cylinder. The flow in the gap is quite interesting. Except when the gap is large ($g=0.125D$) the entire fluid in the gap region is dragged by the rotating control cylinder. All the fluid particles have a velocity in the direction opposite that of the free-stream flow. For $g=0.125D$, except for a region close to the control cylinders, the flow in the gap has a velocity in the free-stream direction.

Shown in Fig. 4 is the variation of the pressure coefficient on

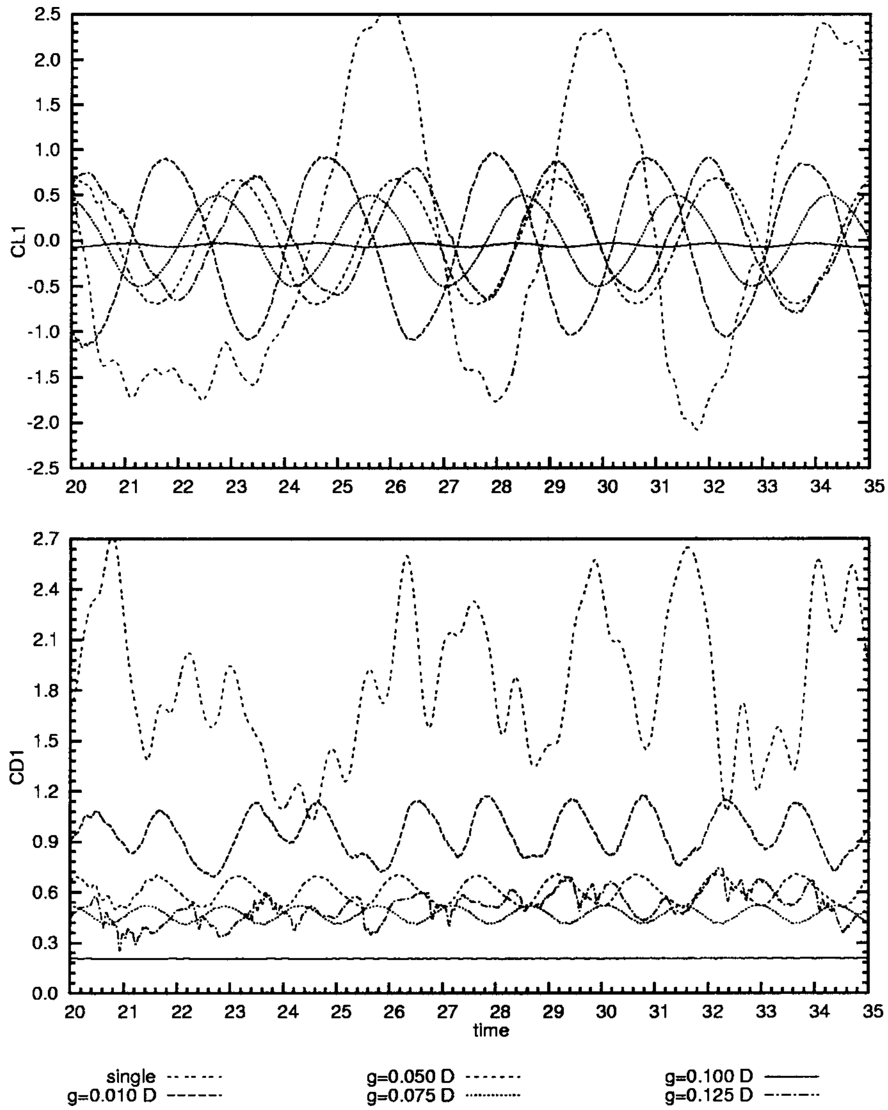


Fig. 6 Re=10^4 flow past main and control cylinders: time histories of the lift and drag coefficients for the main cylinder

the surface of main cylinder for various gaps. The variations for the isolated cylinder and that for the potential flow solution are also shown. The departure of the C_p plot from symmetry around $\theta=0$ deg is associated with a nonzero lift coefficient acting on the cylinder. The plots correspond to the time instant when the lift coefficient of the main cylinder achieves its peak value. Therefore, at this time instant, the C_p distribution is expected to exhibit the maximum asymmetry about $\theta=0$ deg.

In general, a lower base pressure results in a larger drag coefficient. The instantaneous C_p distribution for a single cylinder is quite unsymmetric about $\theta=0$ deg. The base suction coefficient $(-C_p)$ for the viscous case has a less negative value compared to that for the potential flow solution. Therefore it is associated with large mean drag and amplitude of lift coefficients. The $g=0.010D$ case restores the symmetry of the C_p plot to a significant extent and also increases the base pressure. Since the control cylinders are fairly close to the main cylinders, the variation in the pressure distribution on their surfaces caused by the rotation is felt by the main cylinder as well. This appears as sharp peaks and valleys near the shoulder of the main cylinder. As the gap is increased the peaks in the C_p variations due to the control cylinders reduce. In addition, the solution shows increasing symmetry about $\theta=0$ deg and a larger base pressure coefficient. Maximum base pressure is observed for $g=0.100D$. Beyond this value of gap, the base pressure starts decreasing again. The $g=0.100D$ solution is also the most symmetric one among all the cases. This suggests that the $g=0.100D$ case results in very low lift and drag coefficients.

The local peaks in the C_p plots for the main cylinder, shown in Fig. 4, can also be explained as follows. The rotation of the control cylinders cause the surrounding fluid to be dragged along with them. If the gap between the main and rotating control cylinders is too small, not all the fluid can pass through the gap. A part of this high speed moving fluid impinges on the main cylinder, slightly downstream of the gap, leading to a very high pressure and in certain cases, close to the stagnation value. This can also be observed very clearly in the plots for pressure field shown in Fig. 2. The high pressure on the main cylinder due to this effect reduces as the gap increases and the affected region moves further downstream. Beyond a certain value of g not all the fluid in the gap region moves in the direction of the control cylinders. Slightly away from the control cylinders, the flow in the gap, has a velocity in the free-stream direction. This results in a qualitative change in the nature of flow. The wake of the control cylinder that impinges on the main cylinder, perhaps, becomes unstable leading to unsteadiness.

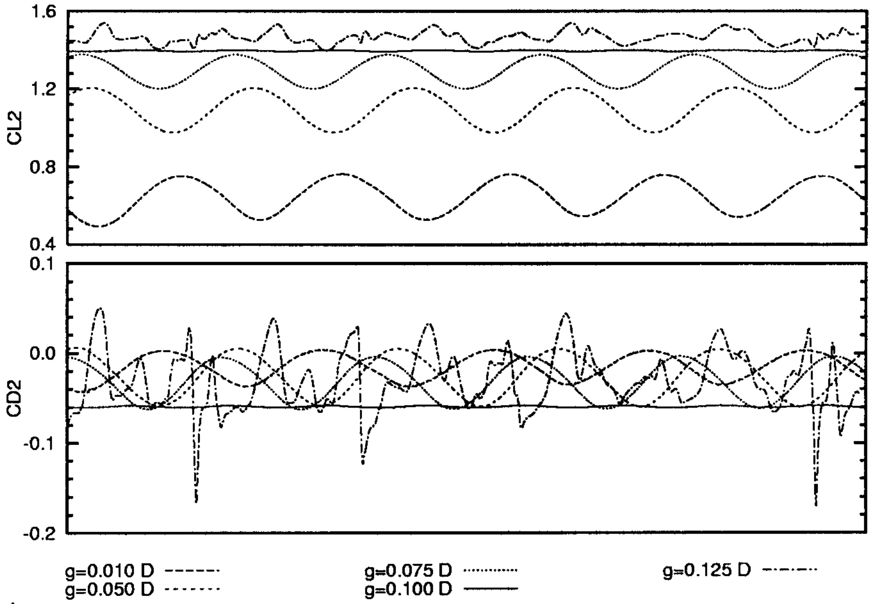


Fig. 7 Re=10^4 flow past main and control cylinders: time histories of the lift and drag coefficients for the upper control cylinder

The vorticity fields for the various cases are shown in Fig. 5. As was observed earlier, the flows with rotating control cylinders are more organized compared to that past an isolated cylinder. The solution for $g=0.100D$ exhibits the lowest level of unsteadiness. Among the cases studied, it appears to be the most optimal gap value. For the small gap case ($g=0.010D$) the upper control cylinder generates counterclockwise vorticity around itself. The boundary layer from the main cylinder is lifted off the surface and is dragged by the control cylinder around it. High speed flow is injected into the wake of the main cylinder by the control cylinders. The efficiency of the control cylinders increases with increase in gap. They drag lesser low momentum fluid from the main cylinder and deliver the high momentum fluid to the main cylinder closer to the shoulder region. However, beyond a certain gap, the effectiveness of the control cylinders is reduced. As can be observed from the streamline plots in Fig. 2, in all the cases except for $g=0.125D$, the gap region is occupied by fluid that goes around the control cylinder. This is suggested by the closed streamlines around the control cylinders. For $g=0.125D$, fluid from the flow past the main cylinder is able to pass through the gap. This observation is also supported by the velocity profiles shown in Fig. 3. This flow has a significant impact on the near wake of the rotating control cylinder that impinges on the main cylinder. In fact, the flow pictures seem to suggest that the wake of the control cylinders is unstable for this value of the gap. This view point is strengthened by the time histories of the force coefficients on the cylinders as shown in Figs. 6 and 7. As a result, the control cylinder is not as effective as for $g=0.100D$.

3.3 Aerodynamic Coefficients for Various Values of Gap. Figures 6 and 7 show the time histories of the lift and drag coefficients for the main and control cylinders for the various values of gap. Compared to flow past an isolated cylinder, flows with rotating control cylinders result in a significant reduction in both mean drag and amplitude of unsteady force coefficients. The case with $g=0.100D$ results in the least value of mean drag and amplitude of lift coefficient. Except for $g=0.125D$, the frequency of the time variation of the forces on the control cylinders matches that on the main cylinder. For $g=0.125D$, the wake from the control cylinders is unstable and time variations at higher frequencies are observed in the time histories of the force coefficients. The increase in gap results in an increase in the lift acting on the control cylinders. This is in line with the observation that the presence of main cylinder has a detrimental effect on the circulation generated by the rotating control cylinders.

An interesting observation from Fig. 7 is the generation of thrust by the control cylinders. A possible explanation is as follows. The oncoming flow, according to an observer placed on the control cylinder, is at a slight angle to the free-stream flow due to the presence of the main cylinder. Therefore the local lift vector is slightly tilted with respect to the free-stream direction and the effective drag coefficient (which is along the free-stream direction) gets a contribution from this tilted lift vector. For larger rotation rates, when the lift on the control cylinder is large, the drag component due to the tilt of the lift vector can become large thereby resulting in a negative drag coefficient. Of course, power is still needed to rotate the control cylinders to overcome the aerodynamic moment.

The variation of the Strouhal number, power requirements and rms value of the time variation of force coefficients are summarized in Figs. 8 and 9. The Strouhal number is based on the time

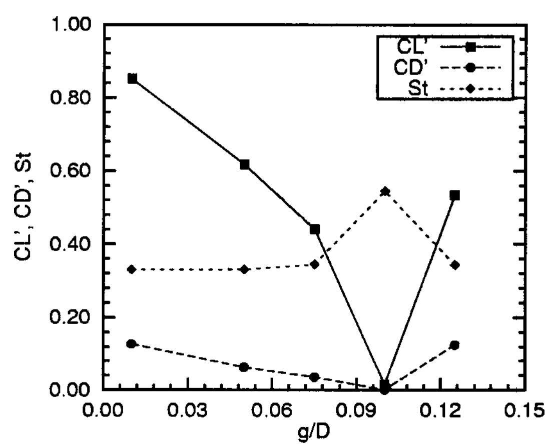


Fig. 8 Re=10^4 flow past main and control cylinders: variation with gap of the rms values of the unsteady force coefficients and Strouhal number

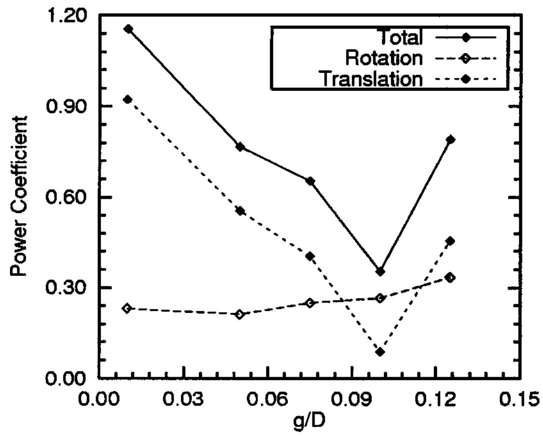


Fig. 9 Re=10^4 flow past main and control cylinders: variation with gap of the time averaged power coefficient

variation of the lift coefficient acting on the main cylinder. The best flow control is achieved for $g=0.100D$. The rotating control cylinders cause an increase in the Strouhal number compared to that for an isolated cylinder. From the figure it appears that, the smaller the rms value of the force coefficients, the larger the Strouhal number. A number of researchers have suggested the idea of ''universal'' Strouhal number where the characteristic length is the distance between the separation points of a bluff body (for example, Griffin (1981) [21], Roshko (1961) [22]). Flow control leads to narrowing of the wake and therefore if one uses the cylinder diameter as the characteristic length, a large value of the Strouhal number is expected. Additionally, the case with optimal gap is associated with the narrowest wake and should therefore exhibit the largest value of Strouhal number.

The total power coefficient, shown in Fig. 9, includes the contributions from drag on all three cylinders and moment acting on the two control cylinders. The power input required for the rotation of the two cylinders increases with gap. It is expected that beyond a certain gap, for each cylinder, the power input should reach a constant value that corresponds to that required to spin an isolated cylinder. Close to $g=0.100D$, the flow is very sensitive to the gap. The rms values of the aerodynamic coefficients and Strouhal number show a sharp variation with gap in this region. The power required for translation shows more variation with the gap as compared to the power needed to sustain the rotary motion. It is quite possible that for a certain value of gap close to $g=0.100D$ the flow is steady. More simulations with finer variations in the gap need to be carried out to investigate this possibility.

4 Concluding Remarks

Control of flow past a circular cylinder, using rotating cylinder elements, has been studied numerically. The Reynolds number is set to 10^4 and the control cylinders rotate at a rate such that their tip speed is five times the free-stream speed of the flow. The effect of the gap between the main and rotating cylinders has been investigated. Compared to the flow past an isolated cylinder the flow with rotating control cylinders results in more organized wake, lower unsteadiness and a significant reduction in the drag coefficient. This reduction in drag is caused by two effects as outlined below.

The control cylinders inject high speed fluid in the wake of the main cylinder. This causes the flow to re-attach resulting in a higher base pressure, compared to that on a single cylinder. The increased base pressure on the main cylinder results in lower drag. The rotary motion of the control cylinders is responsible for the generation of lift. The presence of the main cylinders tilts the lift vector slightly and results in thrust generation by the control cylinders. This causes a further reduction in the drag coefficient. It would be interesting to study the effect of the angular location of the control cylinders. Perhaps, locating them towards the windward side of the main cylinder may result in better control as it will increase the tilt of the lift vector leading to larger thrust.

The rotation of the control cylinders requires a power input. This power requirement increases slightly with increase in gap. However, the power to overcome the drag first reduces and then increases with gap. It is found that the gap is an important parameter in the design of such control strategies. A value of $g=0.100D$ has been found to be close to optimal. It results in a very significant reduction in power savings (close to 70%). Too small a value of gap limits the circulation that the rotating cylinders generate. Very large values of gap result in the unsteadiness of the near wake of the control cylinder, thereby reducing their effectiveness.

Such flow control is also quite effective in reducing flow-induced oscillations because it leads to a reduction in the unsteady forces as well. It is expected that the actual power saving will be smaller than the values reported here. Three-dimensional effects, including those arising from the end wall effects, may reduce the control effectiveness.

Acknowledgment

Partial support for this work has come from the Department of Science and Technology, India.

References

[1] Gad el Hak, M., and Bushnell, D. M., 1991, ''Separation Control: Review,'' ASME J. Fluids Eng., **113**, pp. 5–29.
[2] Modi, V. J., 1997, ''Moving Surface Boundary-Layer Control: A Review,'' J. Fluids Struct., **11**, pp. 627–663.
[3] Modi, V. J., Fernando, M. S. U. K., and Yokomizo, T., 1991, ''Moving Surface Boundary-Layer Control: Studies with Bluff Bodies and Applications,'' AIAA J., **29**, pp. 1400–1406.
[4] Modi, V. J., Shih, E., Ying, B., and Yokomizo, T., 1992, ''Drag Reduction of Bluff Bodies through Momentum Injection,'' J. Aircr., **29**, pp. 429–436.
[5] Munshi, S. R., Modi, V. J., and Yokomizo, T., 1997, ''Control of Fluid-Structure Interaction Instabilities through Momentum Injection,'' in *Proceedings of the Seventh Asian Congress of Fluid Mechanics*, pp. 335–338, Indian Institute of Technology Madras, Chennai, India, Allied Publishers Limited.
[6] Munshi, S. R., Modi, V. J., and Yokomizo, T., 1997, ''Aerodynamics and Dynamics of Rectangular Prisms with Momentum Injection,'' J. Fluids Struct., **11**, pp. 873–892.
[7] Choi, B., and Choi, H., 1992, ''Drag Reduction with a Sliding Wall in Flow over a Circular Cylinder,'' AIAA J., **38**, pp. 715–717.
[8] Park, D. S., Ladd, D. M., and Hendricks, E. W., 1994, ''Feedback Control of von Karman Vortex Shedding behind a Circular Cylinder at Low Reynolds Numbers,'' Phys. Fluids, **6**, pp. 2390–2405.
[9] Mittal, S., 2001, ''Control of Flow Past Bluff Bodies using Rotating Control Cylinders,'' J. Fluids Struct., **15**, pp. 291–326.
[10] Tokumaru, P. T., and Dimotakis, P. E., 1993, ''The Lift of a Cylinder Executing Rotary Motions in a Uniform Flow,'' J. Fluid Mech., **255**, pp. 1–10.
[11] Mittal, S., 2001, ''Flow Past Rotating Cylinders: Effect of Eccentricity,'' ASME J. Appl. Mech., **68**, pp. 543–552.
[12] Mittal, S., and Tezduyan, T. E., 1995, ''Parallel Finite Element Simulation of 3D Incompressible Flows: Fluid-Structure Interactions,'' Int. J. Numer. Methods Fluids, **21**, pp. 933–953.
[13] Mittal, S., and Kumar, V., 1999, ''Finite Element Study of Vortex-Induced Cross-Flow and In-Line Oscillations of a Circular Cylinder at Low Reynolds Numbers,'' Int. J. Numer. Methods Fluids, **31**, pp. 1087–1120.
[14] Mittal, S., Kumar, V., and Raghuvanshi, A., 1997, ''Unsteady Incompressible Flow Past Two Cylinders in Tandem and Staggered Arrangements,'' Int. J. Numer. Methods Fluids, **25**, pp. 1315–1344.
[15] Mittal, S., and Raghuvanshi, A., 2001, ''Control of Vortex Shedding Behind Circular Cylinder for Flow at Low Reynolds Numbers,'' Int. J. Numer. Methods Fluids, **35**, pp. 421–447.
[16] Tezduyar, T. E., Mittal, S., Ray, S. E., and Shih, R., 1992, ''Incompressible Flow Computations with Stabilized Bilinear and Linear Equal-Order-Interpolation Velocity-Pressure Elements,'' Comput. Methods Appl. Mech. Eng., **95**, pp. 221–242.
[17] Akin, J. E., Tezduyar, T. E., Ungor, M., and Mittal, S., 2003, ''Stabilization Parameters and Smaogorinsky Turbulence Model,'' J. Appl. Mech. **70**, pp. 2–9.
[18] Mittal, R., and Moin, P., 1997, ''Suitability of Upwind-Biased-Finite Differ-

ence Schemes for Large-Eddy Simulation of Turbulent Flows,'' AIAA J., **35**, pp. 1415.
[19] Behr, M., 1992, ''Stabilized Finite Element Methods for Incompressible Flows with Emphasis on Moving Boundaries and Interfaces,'' Ph.D. thesis, Department of Aerospace Engineering, University of Minnesota.
[20] Beaudan, P., and Moin, P., 1994, ''Numerical Experiments on the Flow Past a Circular Cylinder at Sub-Critical Reynolds Number,'' Technical Report TF-62, Stanford University, Stanford, CA 94035.
[21] Griffin, O. M., 1981, ''Universal Similarity in the Wakes of Stationary and Vibrating Bluff Structures,'' ASME J. Fluids Eng., **103**, pp. 52–58.
[22] Roshko, A., 1961, ''Experiments on the Flow Past a Circular Cylinder at Very High Reynolds Numbers,'' J. Fluid Mech., **10**, pp. 345–356.

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 2500 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 Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME International, Three Park Avenue, New York, NY 10016-5990, or to the 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.

Helical Collapse of a Whirling Elastic Rod Forced to Lie on a Cylinder

G. H. M. van der Heijden
Center for Nonlinear Dynamics, University College London, London WC1E 6BT, UK

W. B. Fraser
School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia

1 Introduction

Whirling rods in a constrained environment are encountered in a variety of industrial applications, e.g., rotating drill strings confined to narrow boreholes in oilwell drilling (see, e.g., [1,2] for recent references) and in textile yarn manufacturing processes such as two-for-one twisting where the yarn is constrained by a cylindrical guide surface, [3].

Most analytical studies on the buckling of drill strings have assumed continuous contact between the drill string and the borehole wall. In fact, in most cases a helical shape for the drill string is assumed, [2]. In previous work by one of us, [4], continuous contact was considered but no assumption on the shape of the rod was made. An interesting feature found was the collapse of the rod into a helical configuration at certain critical loads.

In this note we extend the analysis in [4] to allow for steady motions of a periodically driven rod (i.e., the configuration of the rod is stationary when viewed in a reference frame that rotates with a constant angular velocity). As we shall be restricting our attention to unshearable linearly elastic rods of inextensible centerline and uniform symmetrical cross section, we take this opportunity to present the formulation of the problem using the traditional notation of engineering structural mechanics, rather than the Cosserat formulation in [4]. We also explore the quasi-statical helical collapse of the rod more carefully as a function of the physical parameters. Lastly, we point out that the so-called balanced ply solutions in recent studies of yarn twisting [5] and DNA supercoiling, [6–8], are special solutions of a rod with cylindrical constraint and therefore occur naturally within our wider formulation.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Jan. 2, 2001, final revision, May 6, 2003. Associate Editor: N. Triantafyllidis.

2 General Formulation

Let $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ be the basis vectors of a right-handed orthonormal coordinate system rotating with respect to an inertial frame with constant angular velocity ω about $\mathbf{k}$ which is pointing along the axis of the cylinder. We also introduce the corresponding cylindrical coordinates (r,θ,z) with basis vectors $(\mathbf{e}_r,\mathbf{e}_\theta,\mathbf{e}_z)$ given by

$$\begin{aligned}\mathbf{e}_r&=\cos\theta\mathbf{i}+\sin\theta\mathbf{j},\\ \mathbf{e}_\theta&=-\sin\theta\mathbf{i}+\cos\theta\mathbf{j},\\ \mathbf{e}_z&=\mathbf{k}.\end{aligned}\tag{1}$$

The dimensionless equations for the rate of change of linear and angular momentum can then be written as (see [9] for details)

$$\Omega^2\{D^2\mathbf{R}+2\mathbf{k}\times D\mathbf{R}+\mathbf{k}\times(\mathbf{k}\times\mathbf{R})\}=(T\mathbf{R}'+\mathbf{V})'-F\mathbf{e}_r-\mu F\mathbf{e}_\theta,\tag{2}$$

$$\varepsilon^2\Omega^2(D\mathbf{H}+\mathbf{k}\times\mathbf{H})=(Q\mathbf{R}'+\mathbf{M})'+\mathbf{R}'\times\mathbf{V}-\varepsilon\mu F\mathbf{k},\tag{3}$$

where

$$\mathbf{H}\delta s=\frac{1}{2}\,\delta s\left\{\omega_t\mathbf{R}'+\frac{1}{2}[\mathbf{R}'\times(D\mathbf{R}'+\mathbf{k}\times\mathbf{R}')]\right\}$$

is the angular momentum vector (relative to the center of mass) of the rod element δs at position vector $\mathbf{R}(s,t)$. s denotes arclength, t time. T is the tension, $\mathbf{V}$ the shear force, Q the torque, $\mathbf{M}$ the bending moment, $D(\ )=\partial(\ )/\partial t$ with respect to the rotating reference frame, $(\ )'=\partial(\ )/\partial s$, $-F\mathbf{e}_r$ the force per unit length of rod that the cylinder exerts on the rod in the direction normal to the cylinder (with F positive when this force is pointing inward), and μ the coefficient of friction between the rod and the cylinder. Note that in the frictionless case the cylinder reaction force has only a component normal to the cylinder. $\omega_t\mathbf{R}'$ is the angular velocity of the rod element about the rod axis.

To these equations we add the following constitutive and constraint equations:

$$\mathbf{R}'\cdot\mathbf{R}'=1,\tag{4}$$

$$\mathbf{R}'\cdot\mathbf{V}=0,\tag{5}$$

$$\mathbf{M}=\mathbf{R}'\times\mathbf{R}'',\tag{6}$$

$$\mathbf{R}'\cdot\mathbf{e}_r=0,\tag{7}$$

expressing, respectively, inextensibility, unshearability, linear elasticity in bending, and the cylindrical constraint.

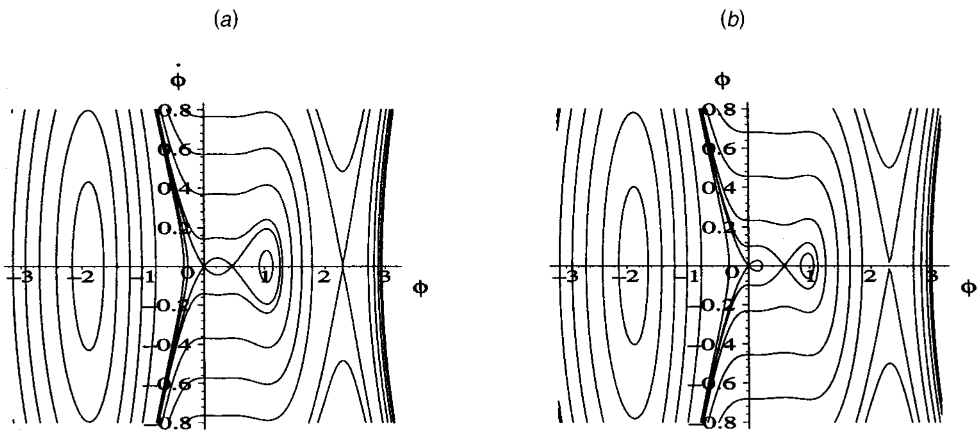


Fig. 1 Phase-plane diagrams for the equivalent oscillator (16) subject to (19) for $r=1$, $K=0.8$ and (*a*) $P=P_c=0.1683$, (*b*) $P=0.1322$. Notice the saddle connection between the origin and the nontrivial fixed point at $\phi=0.4660$ (26.70°) in (*a*).

The Eqs. (2)–(7) have been made dimensionless with respect to the bending stiffness B and mass density m of the rod, the time scale $1/\omega$ and an arbitrary characteristic length scale L (e.g., the radius of the yarn package in unwinding or the radius of the casing in case of an oil drill pipe). The dimensionless parameters Ω and ϵ are thus given by

$$\Omega^2=\frac{m\omega^2L^4}{B},\qquad \epsilon=\frac{a}{L},\tag{8}$$

where a is the radius of the rod.

In the case of a steel drill pipe of radius 0.065 m rotating at say 2 Hz in a casing of radius 0.45 m, $\Omega^2\approx0.2$ and $\varepsilon^2=0.02$ so that it may be reasonable to neglect the rotary inertia terms relative to the translational inertia terms. Dimensions quoted here are compatible with the ranges of physical parameters given in [10]. In the case of yarn twisting dynamics, $\varepsilon\approx10^{-3}$, [3,9].

3 Reduction of the Problem: The Equivalent Oscillator

We now consider the steady-state problem so that $D(\,)\equiv0$, and in the light of the above discussion we neglect the rotary inertia. We assume that the force and moment loads at the remote ends of the rod are applied in the direction of the axis of the cylinder, though not necessarily coaxial. We also ignore friction ($\mu=0$). Thus Eqs. (2) and (3) reduce to

$$(T\mathbf{R}'+\mathbf{V})'=(F-\Omega^2r)\mathbf{e}_r=:G\mathbf{e}_r,\tag{9}$$

$$(Q\mathbf{R}')'+\mathbf{M}'+\mathbf{R}'\times\mathbf{V}=\mathbf{0},\tag{10}$$

where G is the effective reaction force and r is now the constant (dimensionless) radius of the cylinder. These equations can be simplified in several steps as follows. First form the scalar product of $\mathbf{k}$ with (9) and integrate to arrive at

$$(T\mathbf{R}'+\mathbf{V})\cdot\mathbf{k}=P,\tag{11}$$

where P is the applied end force. Next take the scalar product of $\mathbf{R}'$ with (10) to get $Q'=0$, implying that the twisting moment Q is constant along the rod. An expression for the shear force is obtained by forming the vector product of $\mathbf{R}'$ with (10) and using (6):

$$\mathbf{V}=Q(\mathbf{R}'\times\mathbf{R}'')-[(\mathbf{R}''\cdot\mathbf{R}'')\mathbf{R}'+\mathbf{R}'''].\tag{12}$$

The formation of the scalar product of $\mathbf{R}'$ with (9), the use of $\mathbf{V}'\cdot\mathbf{R}'=-\mathbf{V}\cdot\mathbf{R}''$ (since $\mathbf{V}\cdot\mathbf{R}'=0$) and (12) followed by integration gives an expression for the tension

$$T=T_0-\frac{1}{2}(\mathbf{R}''\cdot\mathbf{R}''),\tag{13}$$

where T_0, the integration constant, is a reference tension (taken at one of the ends of the rod, for instance).

Because r is constant it follows from (1) that the position and (unit) tangent vectors can be written as

$$\begin{aligned}\mathbf{R}&=r\cos\theta\mathbf{i}+r\sin\theta\mathbf{j}+z\mathbf{k},\\ \mathbf{R}'&=r\theta'\mathbf{e}_\theta+z'\mathbf{k}=\sin\phi\mathbf{e}_\theta+\cos\phi\mathbf{k},\end{aligned}\tag{14}$$

where ϕ is the angle between the tangent of the rod and the axis of the cylinder. Substitution of (12), (13) and (14) into (11) gives

$$P=T_0\cos\phi-\frac{3}{2}\left[\frac{\sin^4\phi\cos\phi}{r^2}+\phi'^2\cos\phi\right]-(\cos\phi)''+\frac{Q\sin^3\phi}{r}.\tag{15}$$

This equation can finally be integrated to arrive at an equivalent oscillator for the angle ϕ which can be written as

$$\frac{1}{2}\phi'^2+V(\phi)=T_0,\tag{16}$$

with the "potential energy" $V(\phi)$ given by

$$V(\phi)=P\cos\phi+\frac{K\sin\phi}{r}-\frac{Q\sin\phi\cos\phi}{r}-\frac{\sin^4\phi}{2r^2},\tag{17}$$

where K is the applied end moment. This result agrees with the derivation in [4].

All physical quantities, such as P, T, $\mathbf{V}$, and $\mathbf{M}$ can now be expressed in terms of the angle ϕ and its derivatives. A final quantity that will be useful later is the effective reaction force G for which we have, from (9),

$$G=(T\mathbf{R}''+\mathbf{V}')\cdot\mathbf{e}_r.\tag{18}$$

4 Solutions and Helical Collapse

Fixed points of (16) are given by solutions of $V'(\phi)=0$ and correspond to helical solutions with pitch angle $\pi/2-\phi$ and axial wavelength λ given by $\lambda=2\pi r/|\tan\phi|$. By (14), the helix is right-handed if $0<\phi<\pi/2$ or $-\pi<\phi<-\pi/2$, and left-handed if $\pi/2<\phi<\pi$ or $-\pi/2<\phi<0$; it is in tension if $|\phi|<\pi/2$, and in compression if $\pi/2<|\phi|<\pi$.

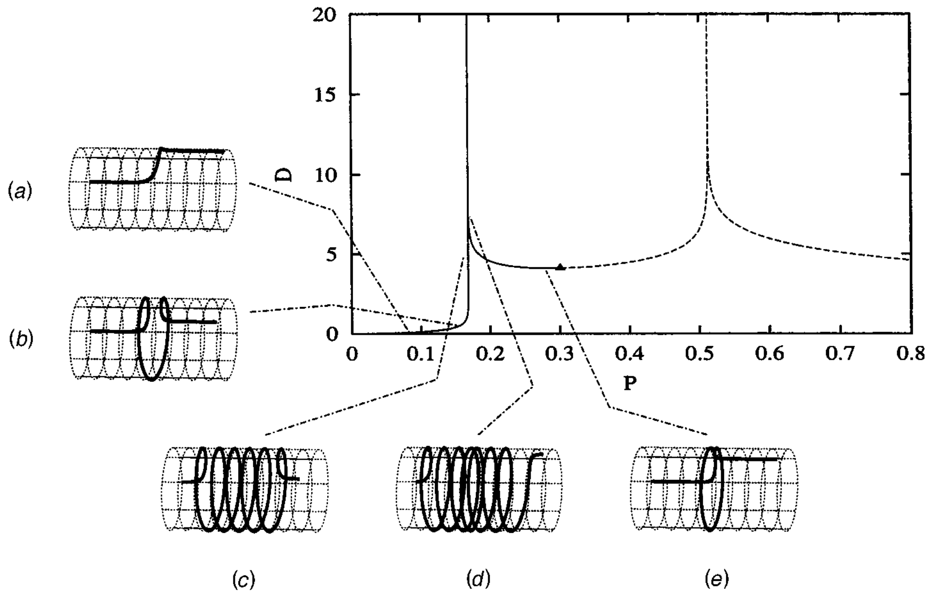


Fig. 2 Load-deflection characteristic and evolution, under varying load P, of the localized solution with initial $\phi>0$. There is a critical collapse load corresponding to a right-handed tensile helix at $P_c=0.1683$. The triangle indicates where the rod starts to go backwards on the cylinder over some section of rod. This is soon followed by self-intersection, so the dashed part of the curve, including the second critical load at $P=0.5123$, is nonphysical. D is the dimensionless end shortening. ($r=1$, $K=0.8$.)

We shall only consider values of the integration constants that admit the straight rod ($\phi\equiv 0$) as a solution. This means that we have to choose

$$Q=K. \tag{19}$$

The origin is then a saddle if $P>0$ (straight rod in tension) and a center if $P<0$ (straight rod in compression).

Figure 1 shows two phase portraits for the oscillator (16), subject to (19), taking $r=1$, $K=0.8$. Generically, the origin has two homoclinic orbits that connect the saddle to itself (as in Fig. 1(*b*)). These solutions correspond to asymptotically straight localized solutions. At special values of the parameters, however, a (heteroclinic) connection may be formed between the origin and a nontrivial saddle (as in Fig. 1(*a*)). Since this nontrivial saddle corresponds to a helix, these critical parameter values define loads at which the rod collapses into a helix. This is illustrated by the load-deflection diagram in Fig. 2 which shows the end shortening D, defined by

$$D=\int_{-\infty}^{\infty}(1-\cos\phi(s))\,ds, \tag{20}$$

as a function of the applied load for one of the homoclinic orbits to the origin in Fig. 1(*b*)) (the one with $\phi>0$). As the critical load $P_c=0.1683$ is approached the rod coils up and D diverges. At P_c the solution is a pure (infinite) helix.

In [4] it is shown that there are at most two critical collapse loads, one of which involves self-intersections. The physical collapse load P_c only exists for $0\leqslant rK\leqslant 0.9648$ and $1.9093\leqslant K/\sqrt{P}\leqslant 2$, or, equivalently,

$$1.9093\sqrt{P}\leqslant K\leqslant \min\{0.9648/r, 2\sqrt{P}\}. \tag{21}$$

The angle ϕ of the critical helix varies between $0\,(K=0)$ and $0.7320\,(K=0.9648)$. This means the helix is always right-handed and in tension. Noncritical helices of course are less constrained in their characteristics.

G is constant along the central helical part of the solution and changes rapidly in the transition to the straight terminal sections where G tends to zero. For a rod inside a cylinder, as in the drill string problem, wall contact will be maintained as long as $F>0$. Where F drops to zero the rod will lift off and the present model ceases to be valid. In terms of G the contact condition is $G>-m\omega^2 r$, so the critical liftoff level is set by the centrifugal force, and contact can always be preserved by using a sufficiently high driving frequency ω.

For a helical solution (18) and (16) yield relatively simple expressions for G and P:

$$G=\frac{\sin\phi(1-\cos\phi)(\sin\phi(1+\cos\phi)-rK)}{r^3}, \tag{22}$$

$$P=\frac{-rK(1-\cos\phi)+2\sin^2\phi(rK-\sin\phi\cos\phi)}{r^2\sin\phi}. \tag{23}$$

We consider three special cases:

1. the straight rod: $\phi=0$, $G=0$, P and K indeterminate.
2. the (multicovered) ring: $\phi=\pi/2$, $G=(1-rK)/r^3$, $P=K/r$. So G drops to zero when the applied axial moment K provides the bending moment required to hold the rod in a ring of radius r.
3. the free helix: $G=0$, $rK=\sin\phi(1+\cos\phi)$, $P=(\sin^2\phi)/r^2$. These are well-known relations for a Kirchhoff rod bent into a helix of radius r (see, e.g., [11]).

5 Discussion

Although our work does not give information on the stability of the solutions considered, the presence of the cylindrical constraint is expected to make a number of them stable and therefore observable in practice (this was the experience with constrained Euler buckling in [12]).

The analysis can be extended to the case of a rod of nonsymmetric cross section. However, since one has to keep track of the orientation of the cross section a director formulation as employed in [4] is required which complicates matters. It is also no longer possible to reduce the system of equations to a planar oscillator as in Section 3. Instead, one gets so-called spatial chaos with infinitely many localized solutions including multipulse ones. For more on this the reader is referred to [13].

The work described in this note is relevant for ply solutions as have recently attracted a good deal of interest in studies of yarn twisting, [5], and DNA supercoiling, [6–8]. In its simplest form a ply consists of two segments of rod in continuous contact along a straight line, the ply axis. Thus a ply is a special case of a rod (more precisely, a pair of rods) winding on a cylinder of radius equal to the radius of the rod, one segment providing the required pressure force G (now negative!) to a 180° rotated copy of itself. Indeed, (16) is a generalization of Eq. (2.41) in [6] for a balanced (i.e., unloaded) ply to the case of nonzero end loading. See [14] for more on this.

References

[1] Tucker, R. W., and Wang, C., 1999, "An Integrated Model for Drill-String Dynamics," J. Sound Vib., **224**, pp. 123–165.
[2] Huang, N. C., and Pattillo, P. D., 2000, "Helical Buckling of a Tube in an Inclined Wellbore," Int. J. Non-Linear Mech., **35**, pp. 911–923.
[3] Fraser, W. B., 1993, "On the Dynamics of the Two-for-One Twister," Proc. R. Soc. London, Ser. A, **447**, pp. 409–425.
[4] van der Heijden, G. H. M., 2001, "The Static Deformation of a Twisted Elastic Rod Constrained to Lie on a Cylinder," Proc. R. Soc. London, Ser. A, **457**, pp. 695–715.
[5] Fraser, W. B., and Stump, D. M., 1998, "The Equilibrium of the Convergence Point in Two-Strand Yarn Plying," Int. J. Solids Struct., **35**, pp. 285–298.
[6] Coleman, B. D., and Swigon, D., 2000, "Theory of Supercoiled Elastic Rings With Self-Contact and Its Application to DNA Plasmids," J. Elast., **60**, pp. 173–221.
[7] Stump, D. M., Fraser, W. B., and Gates, K. E., 1998, "The Writhing of Circular Cross-Section Rods: Undersea Cables to DNA Supercoils," Proc. R. Soc. London, Ser. A, **454**, pp. 2123–2156.
[8] Stump, D. M., and Fraser, W. B., 2000, "Multiple Solutions for Writhed Rods: Implications for DNA Supercoiling," Proc. R. Soc. London, Ser. A, **456**, pp. 455–467.
[9] Clark, J. D., Fraser, W. B., and Stump, D. M., 2001, "Modelling of Tension in Yarn Package Unwinding," J. Eng. Math., **40**, pp. 59–75.
[10] Jansen, J. D., 1991, "Non-Linear Rotor Dynamics as Applied to Oilwell Drill-string Vibrations," J. Sound Vib., **147**, 115–135.
[11] van der Heijden, G. H. M., and Thompson, J. M. T., 2000, "Helical and Localised Buckling in Twisted Rods: A Unified Analysis of the Symmetric Case," Nonlinear Dyn., **21**, pp. 71–99.
[12] Holmes, P., Domokos, G., Schmitt, J., and Szeberényi, I., 1999, "Constrained Euler Buckling: An Interplay of Computation and Analysis," Comput. Methods Appl. Mech. Eng., **170**, pp. 175–207.
[13] van der Heijden, G. H. M., Champneys, A. R., and Thompson, J. M. T., 2002, "Spatially Complex Localisation in Twisted Elastic Rods Constrained to a Cylinder," Int. J. Solids Struct., **39**, pp. 1863–1883.
[14] Thompson, J. M. T., van der Heijden, G. H. M., and Neukirch, S., 2002, "Supercoiling of DNA Plasmids: Mechanics of the Generalized Ply," Proc. R. Soc. London, Ser. A, **458**, pp. 959–985.

Nonlinear Free Flexural Vibration of Oval Rings

M. Ganapathi
Professor

B. P. Patel
Professor

D. P. Makhecha
Senior Research Fellow

Institute of Armament Technology, Deemed University, Girinagar, Pune 411 025, Maharastra, India

In this article, the nonlinear free vibration characteristics of isotropic oval rings are analyzed using a shear flexible cubic B-spline curved beam element. The amplitude-frequency relationships are estimated from the response history. the participation of various modes in the total response is highlighted.

[DOI: 10.1115/1.1604834]

1 Introduction

The nonlinear characteristics of the large-amplitude asymmetric flexural vibrations of the isotropic thin circular/oval rings have been given little attention in the literature, [1], compared to that of beam and plate elements. The notable contributions are the work of Evensen [2], Dowell [3], Sathyamoorthy and Pandalai [4], and Prathap and Pandalai [5]. In the work of Evensen [2] and Sathyamoorthy and Pandalai [4], the nature of nonlinear vibration behavior is predicted based on the inextensionality assumption, whereas such assumption is relaxed by Dowell [3] and Prathap and Pandalai [5]. All these analyses are based on Galerkin technique with two-mode approximation. It is concluded in the work of Chen and Babcock [6] that analytical methods with assumed displacement fields, without proper judgement, may even qualitatively predict different type of nonlinear behavior. Numerical methods such as finite element procedure may be preferred over the analytical methods in the sense that there is no need for an a priori assumption of mode shapes and such analysis for rings appears to be lacking in the literature. Furthermore, there is no information available in the existing literature on the participation of various asymmetric modes while vibrating the oval ring at large amplitudes in a particular mode.

In the present paper, a shear flexible curved beam element using cubic B-spline functions developed recently, [7], is employed to analyze the nonlinear free vibrations of isotropic oval rings. The dynamic responses are obtained using Newmark's integration procedure coupled with Newton-Raphson iterations. The amplitude-frequency relationships are estimated from the response history. The amount of participation of various modes in the total response is evaluated using modal expansion approach.

2 Formulation

A curved beam is considered with the coordinates x along the axis of the beam and z along the thickness direction. The tangential and normal displacements (u,w) at a point (x,z) are expressed in terms of midplane displacements u_o and w, and independent rotation θ of the normal in xz-plane, as

$$\begin{aligned} u(x,z,t) &= u_o(x,t)(1-z/r) - zw_{,x}(x,t) \\ &\quad + f(z)[w_{,x}(x,t)+\theta(x,t)]; \\ w(x,z,t) &= w(x,t), \end{aligned} \tag{1}$$

where t is the time and the radius of curvature r is taken as $r_o/[1+\zeta\cos(2x/r_o)]$ with ζ being ovality parameter; r_o is the radius of a circle whose perimeter is equal to that of oval ring. The function $f(z)$, higher order in nature, is defined as $[h/\pi\sin(\pi z/h)]$ and it avoids the use of any shear correction factor.

The strain-displacement relations, based on von Karman's assumption, are written as

$$\begin{aligned} \varepsilon_{\mathbf{x}} &= u_{o,x} + w/r - zw_{,xx} + f(z)(w_{,xx}+\theta_{,x}) \\ &\quad - zu_{o,x}/r + zu_o r_{,x}/r^2 + (1/2)w_{,x}^2; \\ \{\varepsilon_s\} &= f_{,z}(w_{,xx}+\theta_{,x}) \end{aligned} \tag{2}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 12, 2001; final revision, Apr. 25, 2003. Associate Editor: V. K. Kinra.

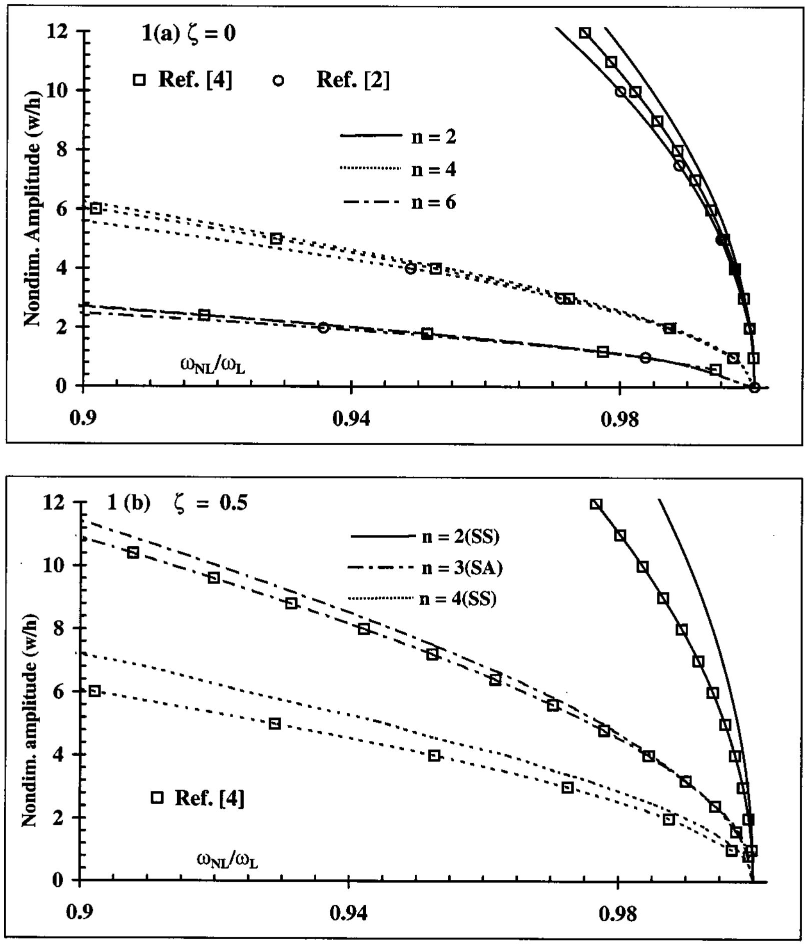


Fig. 1 Nonlinear amplitude frequency relationship for isotropic rings with $r_o/h=100$

where ε_x and ε_s are the in-plane normal and transverse shear strains. The subscript comma denotes the partial derivative with respect to the spatial coordinate succeeding it.

The strain energy of the ring can be expressed in terms of the field variable u_o, w, θ, and their derivatives. The kinetic energy includes the effect of in-plane and rotary inertia terms. The governing equations obtained using the Lagrange's equation of motion are solved based on finite element method, [7]. Using Eqs. (1) and (2) and following the procedure outlined in the work of Rajasekaran and Murray [8], the finite element equations thus derived are

$$[\mathbf{M}]\{\ddot{\delta}\}+[[\mathbf{K}]+(1/2)[\mathbf{N}_1]+(1/3)[\mathbf{N}_2]]\{\delta\}=\{\mathbf{0}\}. \quad (3)$$

Here, $[\mathbf{K}]$ and $[\mathbf{M}]$ are the linear stiffness and mass matrices; $[\mathbf{N}_1]$ and $[\mathbf{N}_2]$ are nonlinear stiffness matrices, linearly and quadrati-

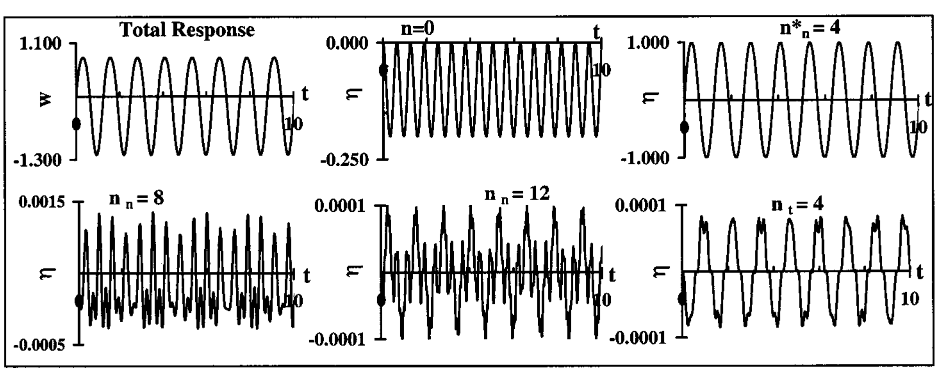


Fig. 2 Modal participation factors for circular ring [$\zeta=0$, $r_o/h=100$, $n^*_{\mathrm{excited}}=4$]

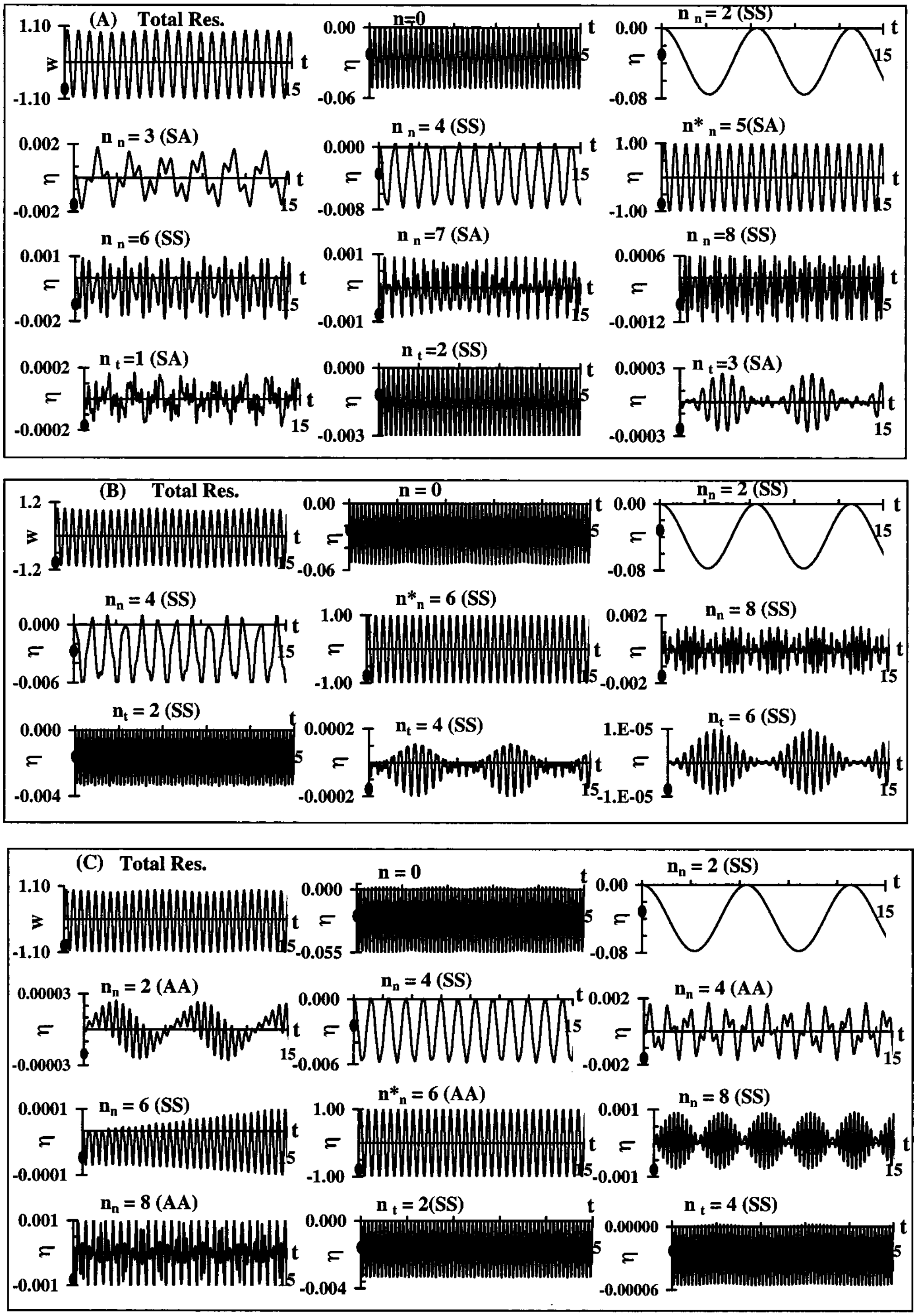


Fig. 3 Modal participation factors for oval ring [ζ=0.5, r_o/h=100]. (A) n^*_{excited}=5 (SA Mode); (B) n^*_{excited}=6 (SS Mode); (C) n^*_{excited}=6 (AA Mode).

cally dependent on the field variables; $\{\ddot{\delta}\}$ and $\{\delta\}$ are the acceleration and displacement vectors, respectively. The resulting-nonlinear Eq. (3) is solved for the dynamic response histories by varying the initial displacement vectors proportional to linear flexural modes.

3 Results and Discussion

Based on progressive mesh refinement, cubic B-spline section with $q=60$ is found to be adequate in modeling the full ring. The frequency values obtained from the present formulation for linear and nonlinear vibration of circular rings have been compared with the analytical solutions, [5,9], and excellent agreement was seen. For the sake of brevity, such comparisons are not presented here. The analysis of oval rings is carried out considering four different type of spatially fixed asymmetric modes such as (i) modes symmetric about both geometrical symmetry axes with even number of circumferential waves (n)–SS, (ii) modes antisymmetric about both axes with even n–AA (iii) modes of odd n symmetric about one and antisymmetric about other axis−SA and (iv) modes of odd n antisymmetric about one and symmetric about other axis−AS.

In Fig. 1, the nonlinear frequency ratios obtained here for the rings ($\zeta=0$ and 0.5) considering modes having different n are compared with those of the approximate analytical methods. Although the present results are fairly in close agreement with the available solutions, the difference in the results increases with the increase in the amplitudes of vibration. This discrepancy is attributed to the limited number of terms retained for displacement in the analytical models.

Next, the degree of participation of various natural modes (axisymmetric mode, $n=0$; asymmetric modes dominated with normal deflection, n_n; and asymmetric modes dominated with tangential deflection, n_t) in the total response, while exciting the particular asymmetric mode, is examined in terms of modal participation factors (η).

The responses of different modes are presented in Figs. 2 and 3 for the circular ($\zeta=0$) and oval rings ($\zeta=0.5$). It is observed from these figures that the amplitude of the total response of the ring in the outward direction is less than that of the inward direction. The participation of contractional double frequency axisymmetric mode in asymmetric vibrations is brought out here, instead of presuming such mode in the analytical methods. In addition to $n=0$ mode, the participation of modes having kn (k is an integer; n is the excited one) circumferential waves, is demonstrated in Fig. 2 for the circular ring. However, it is revealed from Fig. 3 that, for the oval case, the participation of asymmetric modes of different n highly depends on the type of excited asymmetric mode. The participation of all SS modes while exciting a particular SS mode, participation of all SS as well as AA modes while exciting a AA mode, and participation of all SS as well as SA modes while exciting a SA mode are brought out. The nature of the participating lower SS modes ($n<n_{\text{excited}}$) is of contractional type whereas higher participating SS modes ($n>n_{\text{excited}}$) have both contraction and expansion phases. For other participating modes (AA and SA), irrespective of n compared to n_{excited}, both contraction and expansion phases are involved in their responses. In general, the response frequencies of participating modes with $n<n_{\text{excited}}$ are lower whereas it is higher for the participating modes with $n>n_{\text{excited}}$ while comparing with that of excited mode. Furthermore, for the oval ring case, the contribution of lower asymmetric modes (n_n, n_t) is significant compared to that of the higher modes. However, the contribution of n_t modes is very less for circular ring. It is hoped that the present study is very useful for the researchers while accurately modeling and analyzing the closed noncircular structures.

References

[1] Sathyamoorthy, M., 1998, *Nonlinear Analysis of Structures*, Boca Raton, FL.
[2] Evensen, D. A., 1965, "A Theoretical and Experimental Study of the Nonlinear Flexural Vibrations of Thin Circular Rings," Paper No. NASA TR R-227.
[3] Dowell, E. H., 1967, "On the Nonlinear Flexural Vibrations of Rings," AIAA J., **5**, pp. 1508–1509.
[4] Sathyamoorthy, M., and Pandalai, K. A. V., 1971, "Nonlinear Flexural Vibrations of Oval Rings," J. Aeronaut. Soc. India, **23**, pp. 1–12.
[5] Prathap, G., and Pandalai, K. A. V., 1978, "The Role of Median Surface Curvature in Large Amplitude Flexural Vibrations of Thin Shells," J. Sound Vib., **60**, pp. 119–131.
[6] Chen, J. C., and Babcock, C. D., 1975, "Nonlinear Vibration of Cylindrical Shells," AIAA J., **13**, pp. 868–876.
[7] Patel, B. P., Ganapathi, M., and Saravanan, J., 1999, "Shear Flexible Field-Consistent Curved Spline Beam Element for Vibration Analysis," Int. J. Numer. Methods Eng., **46**, pp. 387–407.
[8] Rajasekaran, S., and Murray, D. W., 1973, "Incremental Finite Element Matrices," J. Struct. Div. ASCE, **99**, pp. 2423–2438.
[9] Blevins, R. D., 1979, *Formulas for Natural Frequency and Mode Shape*, Van Nostrand Reinhold, New York.

Coincidence of Boobnov-Galerkin and Closed-Form Solutions in an Applied Mechanics Problem

I. Elishakoff
Dept. of Mechanical Engineering, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431-0991, USA
e-mail: elishako@fau.edu

M. Zingales
Dip. di Ingegneria Strutturale e Geotecnica, Viale delle Scienze, Palermo I-90128, Italy

It is shown that the utilization of the Filonenko-Borodich set of functions, as the comparison functions in the Boobnov-Galerkin method leads to the result that coincides with the closed-form solution for the clamped-clamped uniform beam under uniformly distributed load. It is hoped that this remarkable, direct coincidence could be used in graduate courses and books on mechanics of solids [DOI: 10.1115/1.1598474]

1 Introduction

Since the germ of an idea in 1913 by Boobnov and the first paper by Galerkin [1] two years later, numerous investigators in various fields of engineering sciences adopted this technique showing its unusual potential [2,3]. The Boobnov-Galerkin method was proved to converge [4] for a large class of mechanical problems to the exact solutions ([5–9]); a pertinent description of the method and actual applications have been reported in Ref. [1]. Moreover, the equivalence of the Boobnov-Galerkin and Rayleigh-Ritz methods was shown by Singer [10] and several other investigators. Elishakoff and Lee [11] demonstrated that, for the uniform beams, simply supported at both ends, the Boobnov-Galerkin method and the Fourier series method lead to the identical solution. Note also that Galerkin [1] considered bending of uniform beams clamped at both ends using the set of functions

$$P_j(x)=1-(-1)^j\cos(2j\pi x/L), \tag{1}$$

where $P_j(x)$ are comparison functions, L=length of the beam, j is the serial number of the comparison function, and x=axial coordinate measured from the mid-span of the beam. He summed up the series resulting from his method and showed that the results coincided with the exact solution, obtainable by direct integration. In this paper we address ourselves to the interrelation of the Boobnov-Galerkin method and the exact solution in the beam deflection problems. Namely we show the coincidence of these two methods for clamped-clamped boundary conditions, using an alternative set of functions proposed by Filonenko-Borodich [12].

2 Clamped-Clamped Beam Under Uniformly Distributed Load

Let us consider a clamped-clamped uniform beam under static transverse load $q(x)$. The differential equation of the transverse deflection $w(x)$ of the beam reads

$$EId^4w(x)/dx^4=q(x) \tag{2}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 30, 2001; final revision, Nov. 8, 2002. Associate Editor: D. A. Siginer.

with E and I, respectively, the modulus of elasticity, and the moment of inertia of the beam cross section with respect to the neutral axis. The boundary conditions associated with the transverse displacement $w(x)$ read

$$w(x)=dw(x)/dx \quad \text{at } x=0, \quad x=L, \tag{3}$$

L being the length of the beam. For the uniform transverse load $q(x)=q_0$, displacement $w(L/2)$ in the mid-span is given by the well-known formula

$$w(L/2)=q_0L^4/384EI \tag{4}$$

obtained by direct integration of Eq. (2) with attendant boundary conditions in Eq. (3).

The Boobnov-Galerkin solution of Eq. (2) is achieved by choosing approximate solution $\tilde{w}_n(x)$ in the series form:

$$\tilde{w}_n(x)=\sum_{j=0}^{n} A_jP_j(x). \tag{5}$$

Herein the comparison function $P_j(x)$ is represented in the following form:

$$P_{2j}(x)=\cos(2j\pi x/L)-\cos[2(j+1)\pi x/L], \tag{6}$$

$$P_{2j+1}(x)=\cos[(2j+1)\pi x/L]-\cos[(2j+3)\pi x/L], \tag{7}$$

where distinction has been made between symmetric (Eq. (6)) and asymmetric (Eq. (7)) functions $P_j(x)$ with respect to the mid-span of the beam. The set of functions $P_j(x)$, $j=1,2,\ldots \operatorname{int}[n/2]$ constitutes a class of complete functions in the range $[0,L]$ (see Stepanov [13]) satisfying the boundary conditions in Eq. (3); int[•] indicates the integer part. This set was apparently first introduced by Filonenko-Borodich [12]. Moreover, functions $P_j(x)$, $j=1,2,\ldots$ are quasiorthogonal, that is, the following conditions hold:

$$\begin{aligned} &\int_0^L \frac{d^4P_j(x)}{dx^4}P_k(x)dx=-\frac{\pi^4}{2L^3}\begin{cases} j^4 & \text{if } j-k=2 \\ (j+2)^4 & \text{if } k-j=2 \end{cases}, \\ &\int_0^L \frac{d^4P_j(x)}{dx^4}P_k(x)dx=\begin{cases} \dfrac{\pi^4[j^4+(j+2)^4]}{2L^3} & \text{if } j=k \\ 0 & \text{if } j\neq k \end{cases}. \end{aligned} \tag{8}$$

Substitution of Eq. (5) into Eq. (2) yields an error $\varepsilon_n(x)$ given by

$$\varepsilon_n(x)=EI\sum_{j=1}^{n} A_j\frac{d^4P_j(x)}{dx^4}-q_0. \tag{9}$$

Requirement of orthogonality reads

$$(\varepsilon_n,P_j(x))=0; \quad j=1,2,\ldots n, \tag{10}$$

where the scalar product is represented as follows:

$$(\varepsilon_n,P_j(x))=\int_0^L \varepsilon_n(x)P_j(x)dx. \tag{11}$$

Equation (10) represents an algebraic system of n equation in the unknowns A_j, $j=1,2,\ldots n$ which reads, for even $P_{2j}(x)$ and odd comparison $P_{2j+1}(x)$ functions, respectively, the following sets of equations:

$$\begin{cases} A_0-A_2=q_0L^4/8EI\equiv K, \\ 2^4(A_2-A_4)-(A_0-A_2)=0, \\ \cdots \\ (j+1)^4(A_{2j}-A_{2j+2})-j^4(A_{2j-2}-A_{2j})=0, \\ \cdots \end{cases} \tag{12}$$

$$\begin{cases} A_1+3^4(A_1-A_3)=0, \\ 3^4A_3+5^4(A_3-A_5)=0, \\ \cdots \\ (j+1)^4A_{2j+1}+(j+3)^4(A_{2j+1}-A_{2j+3})=0, \\ \cdots \end{cases} \tag{13}$$

where Eq. (8) has been taken into account and integrals involving terms $q(x)$ have been evaluated yielding

$$\int_0^L q(x)P_j(x)dx=\begin{cases} q_0L & \text{if } j=0 \\ 0 & \text{if } j>0 \end{cases}. \tag{14}$$

Solution of the algebraic system in Eq. (13) is trivial, that is, odd-numbered constants A_{2j+1} are identically vanishing, and the approximate transverse displacement of the beam axis may be rewritten after some straightforward manipulations as follows:

$$\tilde{w}_n(x)=A_0+\sum_{j=0}^{n} (A_{2j+2}-A_{2j})\cos[2(j+1)\pi x/L]. \tag{15}$$

Solution of the algebraic systems in Eq. (12) reads

$$\begin{aligned} A_{2n}&=\frac{K}{(n+1)^4}=\frac{q_0L^4}{8\pi^4EI}\frac{1}{(n+1)^4}, \\ A_{2n-2}&=K/(n-1)^4+A_{2n}, \\ &\cdots \\ A_0&=K\sum_{j=0}^{n}\frac{1}{(j+1)^4}=\frac{q_0L^4}{8\pi^4EI}\sum_{j=0}^{n}\frac{1}{(j+1)^4} \end{aligned} \tag{16}$$

since $A_{2n+2}=0$ in Eq. (12) because only the first $2n$ terms have been retained. Equation (15) can be rewritten as

$$\tilde{w}_n(x)=K\left[\sum_{j=0}^{n}\frac{1}{(j+1)^4}-\sum_{j=1}^{n}\frac{\cos(2j\pi x/L)}{j^4}\right]. \tag{17}$$

Letting $n\to\infty$ in Eq. (15) leads to the expression, denoted by $\tilde{w}_\infty$ [14,15],

$$\tilde{w}_\infty(x)=K\left[\frac{\pi^4x^2}{3L^2}-\frac{4\pi^4x^3}{3L^3}+\frac{3\pi^4x^4}{4L^4}\right]; \quad 0\leqslant x\leqslant L. \tag{18}$$

Evaluation of mid-span transverse displacement $\tilde{w}_\infty(L/2)$, from Eq. (18), yields an expression which coincides with the well-known formula given in Eq. (4).

Bending moment $M(x)$, along the beam axis, can be evaluated by means of Eq. (18), bearing in mind the familiar relation

$$\frac{d^2w(x)}{dx^2}=-\frac{M(x)}{EI}. \tag{19}$$

Substituting expression for $\tilde{w}_\infty(x)$ into Eq. (19) results in

$$M(x)=-K\left(\frac{2\pi^4}{3L^2}-\frac{4\pi^4x}{L^3}+\frac{4\pi^4x^2}{L^4}\right). \tag{20}$$

Setting $x=0$, or $x=L$, yields the well-known expression for the bending moment at the end of clamped-clamped beam, namely $M(0)=M(L)=-q_0L^2/12$. Note, that some mathematical properties of Filonenko-Borodich functions defined in Eq. (6) and Eq. (7) were proved by Vilenkin [15].

References

[1] Galerkin, B. G., 1915, "Beams and Plates," Vestnik Inzhenerov, **1**, pp. 897–908, (see also *Collected Works*, USSR Academy of Sciences Publishing, 1952, pp. 182–183).

[2] Duncan, W. J., 1937, "Galerkin's Method in Mechanics and Differential Equations," Aeronautical Research Council Reports and Memoranda No. 1798.

[3] Duncan, W. J., 1938, "The Principles of the Galerkin Method," Aeronautical Research Report and Memoranda No. 1894.

[4] Mikhlin, S. G., 1964, *Variational Methods in Mathematical Physics*, Pergamon Press, Oxford.
[5] Hoff, N. J., 1956, *The Analysis of Structures*, Wiley, New York.
[6] Kantorovich, L. V., and Krylov, V. I., 1958, *Approximate Method for Higher Analysis*, Interscience (English translation), New York.
[7] Ishlinski, A. Ju. (Editor), 1975, *Advances of Mechanics of Deformable Continua* (dedicated to the 100th Anniversary of the Birth of B. G. Galerkin), Nauka, Moscow (in Russian).
[8] Strang, G., and Fix, G. J., 1973, *An Analysis of the Finite Element Method*, Prentice-Hall Englewood Cliffs, New Jersey.
[9] Leipholz, H. H. E., 1976, "Use of Galerkin's Method for Vibration Problems," Shock Vib. Dig., **8**, pp. 3–18.
[10] Singer, J., 1962, "On the Equivalence of Galerkin's and Rayleigh Ritz Methods," Journal of the Royal Aeronautical Society, **66**, p. 592.
[11] Elishakoff, I., and Lee, L. H. N., 1986, "On Equivalence of the Galerkin and Fourier Series Methods for one Class of Problems," J. Sound Vib., **109**, pp. 174–177.
[12] Filonenenko-Borodich, M. M., 1946, "On a Certain System of Functions and its Applications in the Theory of Elasticity," Prikl. Mat. Mekh., **10**, pp. 193–208 (in Russian).
[13] Stepanov, V. V., 1945, "On Some Complete Non-Orthogonal Systems," *Proceedings of the USSR Academy of Sciences*, Vol. XL VIII, No. 6 (in Russian).
[14] Gradshteyn, I. S., and Ryzhik I. M., 1980, *Tables of Integrals Series and Products*, Academic Press, New York.
[15] Vilenkin, N. Ya., 1952, "On Some Nearly Periodic Systems of Functions," PMM, **16**(3), pp. 812–814 (in Russian).

Thermal Stresses in an Infinite Slab Under an Arbitrary Thermal Shock

A. E. Segall
Associate Professor, Engineering Science and Mechanics, The Pennsylvania State University, University Park, PA 16802. Mem. ASME

Introduction

A number of studies on the thermal shock behavior of finite-thickness slabs with the temperature varying through the thickness have been conducted over the years by researchers such as Albrecht [1], Chen [2], and Nied [3], to name just a few. However, a review of these solutions reveals that the transient thermal loading is usually restricted to either a step or linear function of time whereas many thermal shocks are asymptotic in form (Vedula et al. [4] and Tu and Segall [5]). Consequently, the actual time and magnitude of the maximum stress may vary significantly from an overly conservative prediction based on step loading. From an engineering and design perspective, this shortcoming can be critical since reliability and failure predictions can become significantly skewed. Because of the need to realistically address the time-dependent nature of thermal shock, a study was undertaken to derive transient stress solutions for an infinite-slab subjected to an arbitrary thermal loading. The results presented herein show how generalized solutions are possible if the materials properties are assumed constant, the surface temperature history can be described by a polynomial, and the temperature gradients are restricted to the thickness coordinate.

Analytical Considerations

For a finite-thickness slab, the partial differential equation governing the transient temperature distribution is

$$\frac{\partial^2 T}{\partial x^2}=\frac{1}{D}\frac{\partial T}{\partial t} \tag{1}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 16, 2002, final revision, Mar. 31, 2003. Associate Editor: J. R. Barber.

where T is the temperature, t is time, x is the coordinate from the thermally loaded surface, and D is the thermal diffusivity that is assumed to be independent of temperature. The initial and boundary conditions of the problem are

$$T(x,0)=0 \tag{2a}$$

$$\frac{\partial T}{\partial x}(h,t)=0 \tag{2b}$$

$$T(0,t)=T_s(t) \tag{2c}$$

where h is the thickness of the slab and $T_s(t)$ is the surface thermal-loading that is an arbitrary function of time. If a unit step-loading, $T_s(0,t)$=constant is assumed, Eq. (1) can be solved into the following form as shown by Austin [6]:

$$T(x,t)=T_s\left[1-\sum_{k=1}^{\infty}\Psi_k e^{-b_k^2 t}\right] \tag{3}$$

where

$$\Psi_k=A_k\cos\left(Q_k\frac{x-h}{h}\right) \tag{4a}$$

$$A_k=\frac{4(-1)^{k+1}}{\pi(2k-1)} \tag{4b}$$

$$Q_k=\frac{\pi(2k-1)}{2} \tag{4c}$$

$$b_k^2=\frac{DQ_k^2}{h^2}. \tag{4d}$$

The resulting step response can then be used as a kernel with Duhamel's relationship (Fodor [7]) for a more generalized solution:

$$T(x,t)=H(0)\Phi(x,t)+\int_0^t\frac{\partial H(0,\tau)}{\partial\tau}\Phi(x,t-\tau)d\tau \tag{5}$$

to determine the temperature response, $T(x,t)$ to an arbitrary excitation $H(0,t)$. The front term, $H(0)$ represents the step-excitation at $x=0$ while $\Phi(x,t)$ is Eq. (5) with $T_s=1$. Generalization of the excitation $H(0,t)$ is possible via a polynomial containing both integral and half-order terms up to Nth order:

$$H(0,t)=a_0+a_1t^{1/2}+a_2t+a_3t^{3/2}+\ldots=\sum_{n=0}^{N}a_nt^{n/2} \tag{6}$$

where a_n represent polynomial coefficients determined using least squares methods. Given the form of Eqs. (3)–(6), the transient response to a system initially at a temperature of zero under a generalized polynomial excitation becomes

$$T(x,t)=a_0\delta_0(x,t)+\sum_{n=1}^{N}a_n\left[t^{n/2}-\frac{n}{2}\delta_n(x,t)\right] \tag{7}$$

where the response functions $\delta_n(x,t)$ for $n=0,1,2\ldots N$ are defined as

$$\delta_0(x,t)=\Phi(x,t) \tag{8a}$$

with the remaining half and integral-order terms defined as follows:

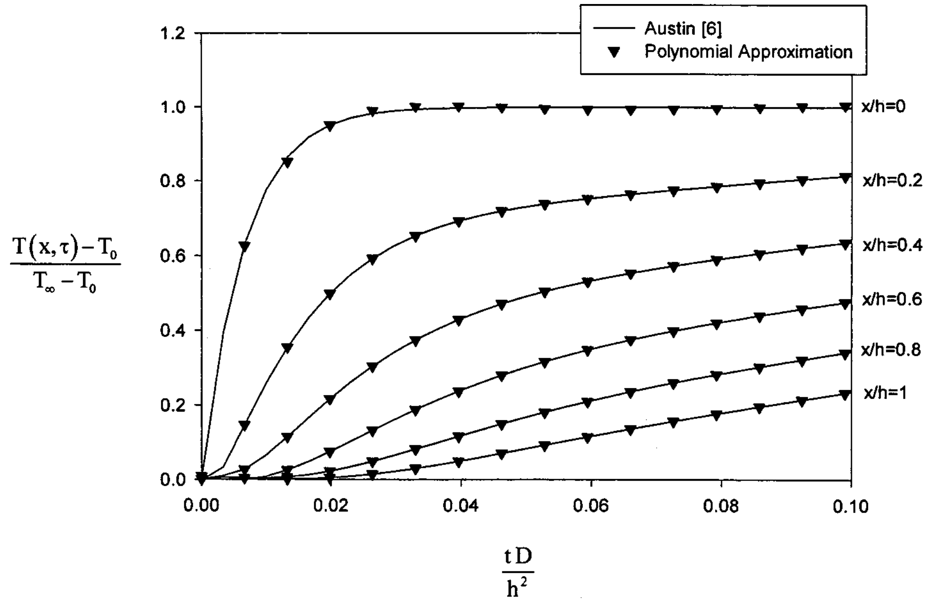


Fig. 1 Transient temperature distributions as a function of nondimensional time and location through the thickness of the slab

$$\delta_{2j-1}(x,t)=\frac{(-1)^{j+1}}{2^{j-1}}\zeta(j-1,2j-3)\sum_{k=1}^{\infty}\frac{\Psi_k\Omega_k}{b_k^{2(j-1)}}+\sum_{i=0}^{j-2}\left[(-1)^i\frac{t^{\frac{2(j-i)-3}{2}}}{2^i}\zeta(i,2j-3)\sum_{k=1}^{\infty}\frac{\Psi_k}{b_k^{2(i+1)}}\right] \quad (8b)$$

$$\delta_{2j}(x,t)=(-1)^j\chi(j-1,j-1)\sum_{k=1}^{\infty}\frac{\Psi_k e^{-b_k^2 t}}{b_k^{2j}}+\sum_{i=0}^{j-1}\left[(-1)^i t^{j-i-1}\chi(i,j-1)\sum_{k=1}^{\infty}\frac{\Psi_k}{b_k^{2(i+1)}}\right] \quad (8c)$$

with the subscript integer set as $j=1,2,3\ldots N$. It is important to note that for $j=1$, the second term in Eq. (8b) is equal to zero for the unlikely summation over the range of $i=0\rightarrow-1$.

The reoccurring functions $\chi(\eta,\lambda)$ and $\zeta(\eta,\lambda)$ found in Eqs. (8b) and (8c) are defined as follows:

$$\chi(\eta,\lambda)=\frac{\lambda!}{(\lambda-\eta)!} \quad (9a)$$

$$\zeta(\eta,\lambda)=\frac{2^{\eta}\Gamma(\lambda/2+1)}{\Gamma(\lambda/2+1-\eta)} \quad (9b)$$

where Γ is the Gamma function. The use of half-order terms by the polynomial results in an additional reoccurring function that includes complex arguments $i=\sqrt{-1}$:

$$\Omega_k(t)=\int_0^t\frac{-e^{(-b_k^2(t-\tau))}}{\sqrt{\tau}}=\frac{-i\sqrt{\pi}\,\mathrm{erf}(ib_k\sqrt{t})e^{-b_k^2 t}}{b_k}. \quad (10)$$

As an alternative to Eq. (10), the Laplace Transform identity of the convolution integral can be used such that

$$\Omega_k(s)=\mathcal{L}\left\{\int_0^t\tau^{-1/2}e^{-b_k^2(t-\tau)}d\tau\right\}=\sqrt{\frac{\pi}{s}}\frac{1}{(s+b_k^2)} \quad (11)$$

where s is the Laplace variable. Once in this form, the inversion can be approximated by using the ten-term Gavor-Stehfest inversion technique advocated by Woo [8]:

$$\Omega_k(t)\approx\frac{\mathrm{Ln}(2)}{t}\sum_{m=1}^{10}\varphi_m\left[\frac{\sqrt{\pi}}{\sqrt{s_m^*}(s_m^*+b_k^2)}\right] \quad (12)$$

where the modified Laplace variable s^* is

$$s_m^*=\frac{\mathrm{Ln}(2)}{t}m \quad (13a)$$

and the series coefficients φ_m, can be determined by the following relationship:

$$\varphi_m=(-1)^{m+p/2}\times\sum_{q}^{\min(m,p/2)}\frac{q^{p/2}(2q)!}{(p/2-q)!q!(q-1)!(m-q)!(2q-m)!}. \quad (13b)$$

In Eq. (13b), q represents the integer value of $(m+1)/2$ and p is the number of terms in the series that is set at ten for the current analysis. Using ten terms, Eq. (12) shows excellent agreement with numerical integration over many decades of the argument. It should also be noted that a first-order approximation developed by Segall [9] can be employed to integrate Eq. (10) for small values of time:

$$\int_0^t\tau^{-1/2}e^{-b_k^2(t-\tau)}d\tau\approx 2t^{1/2}e^{-b_k^2 t} \quad (14)$$

provided the condition $b_k^2 t>0.1$ is satisfied to keep the relative error below 3%.

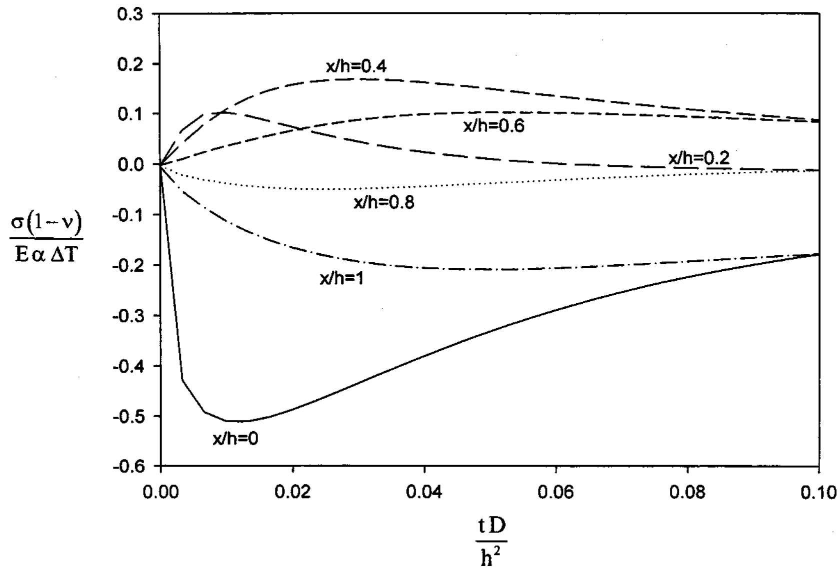


Fig. 2 Transient thermoelastic distributions as a function of nondimensional time and location through the thickness of the slab

Thermoelastic Stresses

Once the transient temperature-distributions are known, the stresses can be determined by the following relationship (Nied [3]):

$$\sigma_y(x,t)=\frac{\alpha E}{(1-\nu)}\left[\frac{4h-6x}{h^2}\int_0^h T(x,t)dx + \frac{12x-6h}{h^3}\int_0^h xT(x,t)dx-T(x,t)\right] \quad (15a)$$

where E is the elastic modulus, ν is Poisson's ratio, and α is the coefficient of thermal expansion. Substitution of the transient temperature distribution results in the following relationship for the prevailing thermal stress:

$$\sigma_y(x,t)=\frac{\alpha E}{(1-\nu)}\left[\frac{4h-6x}{h^2}\lambda 1+\frac{12x-6h}{h^3}\lambda 2-T(x,t)\right] \quad (16)$$

where

$$\lambda 1=a_0h\overline{\delta}_0+\sum_{n=1}^{N}a_nh\left[t^{n/2}-\frac{n}{2}\overline{\delta}_n(t)\right] \quad (17)$$

$$\lambda 2=a_0\frac{h^2}{2}\overline{\overline{\delta}}_0+\sum_{n=1}^{N}a_n\frac{h^2}{2}\left[t^{n/2}-\frac{n}{2}\overline{\overline{\delta}}_n(t)\right]. \quad (18)$$

The redefined terms $\overline{\delta}_n$ and $\overline{\overline{\delta}}_n$ represent the response functions given by Eq. (8) with Ψ_k in each infinite sum replaced by Eqs. (19a) and (19b), respectively

$$\overline{\Psi}_k=\frac{A_k}{Q_k}\sin(Q_k) \quad (19a)$$

$$\overline{\overline{\Psi}}_k=\frac{2A_k}{Q_k^2}[1-\cos(Q_k)] \quad (19b)$$

Results

For an asymptotic thermal loading of the form

$$T(0,t)=\Delta T(1-e^{-1/2t}) \quad (20)$$

with $\Delta T=1$, a six-term polynomial nearly perfectly fits the time-temperature history as shown by the $x/h=0$ curve in Fig. 1. For the current analysis, the six-term polynomial was found to be adequate with lower orders deviating from the imposed asymptotic loading. However, it should be noted that higher-order polynomials would not necessarily improve the quality or accuracy of the fit because curve instabilities or "wiggles" around the individual date points would also develop. Hence, the potential versatility of higher-order polynomials should be carefully weighed against their tendency to meander about the data points.

As also shown in Fig. 1, the predicted response for a thermally thick material ($h=0.05$ m and $D=0.085$ cm^2/s) shows excellent agreement with the closed form expression (Austin [6]) for various values of nondimensional time and depth away from the exposed surface. Figure 2 shows the resulting thermoelastic stresses for the same material ($E=207$ GPa and α=1.8E−6/°C) for the same nondimensional times and depths. For the current calculations, only the first ten terms in the infinite series (Eq. (3)) were required to produce reasonable results; calculations involving up to a 1000 terms produced no significant changes in the calculated thermal response and the resulting thermoelastic stresses.

Conclusions

A series solution was derived that allows the calculation of the thermal transients and the resulting thermal-stresses caused by an arbitrary surface loading on a finite-thickness slab. The arbitrary nature of the thermal loading was allowed through the use of a polynomial that employed integral and half-order terms. The method appears well suited for complicated thermal shocks provided the analysis is restricted to the time interval used to determine the polynomial and the thermophysical properties do not vary with temperature.

References

[1] Albrecht, W., 1969, "How Thickness and Materials Properties Influence Thermal Shock Stresses in Flat Plates and Cylinders," *ASME Paper G9-GT-107*.

[2] Chen, S. H., 1961, "One-Dimensional Heat Conduction With Arbitrary Heating Rate," J. Aerosp. Sci., pp. 336–337.

[3] Nied, H. F., 1987, ''Thermal Shock in an Edge-Cracked Plate Subjected to Uniform Surface Heating,'' Eng. Fract. Mech., **26**, pp. 239–246.
[4] Vedula, V. R., Green, D. J., Hellmann, J. R., and Segall, A. E., 1999, ''Test Methodology for the Thermal Shock Characterization of Ceramics,'' J. Mater. Sci., **33**, pp. 5427–5432.
[5] Tu, J. J., and Segall, A. E., 1997, ''Thermomechanical Analysis of a Complex, Refractory Tundish Flow Modifier During Preheating,'' *Proceedings of the Unified International Technical Conference on Refractories, 5th Biennial Worldwide Congress*, New Orleans, LA, Nov.
[6] Austin, J. B., 1932, ''Temperature Distribution in Solid Bodies During Heating or Cooling,'' J. Appl. Phys., **3**, pp. 179–184.
[7] Fodor, G., 1965, *Laplace Transforms in Engineering*, Akademiai Kiado, Budapest.
[8] Woo, K., 1980, ''TI-59 Inverts Laplace Transforms for Time-Domain Analysis,'' Electronics, Oct. 9, pp. 178–179.
[9] Segall, A. E., 2001, ''Relationships for the Approximation of Direct and Inverse Problems With Asymptotic Kernels,'' Inverse Prob. Eng., **9**, pp. 127–140.

Journal of Applied Mechanics

Discussion

Discussion: "Normal Indentation of Elastic Half-Space With a Rigid Frictionless Axisymmetric Punch" (Fu, G., and Chandra, A., 2002, ASME J. Appl. Mech., 69, pp. 142–147)

F. M. Borodich
L. M. Keer
Northwestern University, Evanston, IL 60208-3109

The authors presented an interesting consideration of an axisymmetric frictionless contact problem with the aid of mathematical software MATHEMATICA. Evidently, the use of modern analytical software gives a possibility to obtain new results, check known solutions, and correct possible misprints. However, some papers in the field should be added to their reference list.

In 1939 an analytical solution for a punch described by a monomial function of r of a positive even degree α was obtained by Shtaerman [1]. It is worth mentioning that after A. E. H. Love had obtained his solution, the problem for conical punch was also solved by Lur'e [2] in 1941. The problem for a punch described by a monomial function of r of an arbitrary real degree α was solved by Galin (see Chap. 2, paragraph 5 in Ref. [3]). Then this problem was also analyzed by Sneddon [4]. In 1957 the problem was analyzed by Segedin [5] for a punch whose shape is represented by a series (a polynomial function of r) with integer degrees α. For a punch described by a fractional power series of r, the problem was analyzed in Ref. [6]. The analysis in Ref. [6] was based on the Galin's solution ([3]). It was shown that the solution can be also used in the case when the punch is a transversally isotropic solid and the half space has homogeneous initial stresses. In particular, a formula similar to formula (17) was obtained.

References

[1] Shtaerman, I. Ya., 1939, "On the Hertz Theory of Local Deformations Resulting From the Pressure of Elastic Solids," Dokl. Akad. Nauk SSSR, **25**, pp. 360–362 (in Russian).

[2] Lur'e, A. I., 1941, "Some Contact Problems of the Theory of Elasticity," Prikl. Mat. Mekh., **5**, pp. 383–408 (in Russian); Lur'e, A. I., 1964, *Three-Dimensional Problems of the Theory of Elasticity*, Interscience Publishers, New York.

[3] Galin, L. A., 1953, *Contact Problems in the Theory of Elasticity*, Gostekhizdat, Moscow-Leningrad. (English translation by H. Moss, edited by I. N. Sneddon, North Carolina State College, Departments of Mathematics and Engineering Research, NSF Grant No. G16447, 1961).

[4] Sneddon, I. N., 1965, "The Relation Between Load and Penetration in the Axisymmetric Boussineq Problem for a Punch of Arbitrary Profile," Int. J. Eng. Sci., **3**, pp. 47–57.

[5] Segedin, C. M., 1957, "The Relation Between Load and Penetration for a Spherical Punch," Mathematika, **4**, pp. 156–161.

[6] Borodich, F. M., 1990, "Hertz Contact Problems for Elastic Anisotropic Half-Space With Initial Stress," Soviet Appl. Mechanics, **26**, pp. 126–132.

Closure to "Discussion of 'Normal Indentation of Elastic Half-Space With a Rigid Frictionless Axisymmetric Punch' " (2003, ASME J. Appl. Mech., 70, pp. 783)

A. Chandra
G. Fu

Mechanical Engineering Department, Iowa State University, Ames, IA 50011

Our work ([1]) was based on Green's solution ([2]). We thank the discussors for pointing out a different approach taken by Borodich [3] following the work of Galin [4]. At the time of publication, we were unaware of the work by Borodich. The usage of our derived solution is straightforward. With modern mathematical software, hypergeometric function will be like a regular elementary function and the final result is easy to be obtain. It can also be used to check analytical expressions for possible misprints.

As it is pointed out in the paper, the power of the "polynomial" can be any non-negative number, such as 0, 2, 1/12, e, π. With this solution, one can use multiple terms to define the punch shape instead of a monomial function of the punch radius.

We appreciate the fact that there exist numerous contributions to this field in the Russian literature, and our understanding of this work is mainly based on the books by Gladwell [5] and Sneddon [6]. Johnson [7] also mentioned the solutions by Shtaerman and Galin in his book.

References

[1] Fu, G., and Chandra, A., 2002, "Normal Indentation of Elastic Half-Space With a Rigid Frictionless Axisymmetric Punch," ASME J. Appl. Mech., **69**, pp. 142–147.

[2] Green, A. E., and Zerna, W., 1954, *Theoretical Elasticity*, Oxford University Press, London, Great Britain.

[3] Borodich, F. M., 1990, "Hertz Contact Problems for Elastic Anisotropic Half-Space With Initial Stress," Soviet Appl. Mechanics, **26**, pp. 126–132.

[4] Galin, L. A., 1953, *Contact Problems in the Theory of Elasticity*, Gostekhizdat, Moscow-Leningrad (English translation by H. Moss, edited by I. N. Sneddon, North Carolina State College, Departments of Mathematics and Engineering Research, NSF Grant No. G16447, 1961).

[5] Gladwell, G. M. L., 1980, *Contact Problems in the Classical Theory of Elasticity*, Sijthoff & Noordhoff, Alphen aan den Rijn, the Netherlands.

[6] Sneddon, I. N., 1966, *Mixed Boundary Value Problems in Potential Theory*, North-Holland Publishing Company, Amsterdam, Holland.

[7] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

Discussion: "Dynamic Condensation and Synthesis of Unsymmetric Structural Systems" (Rao, G. V., 2002, ASME J. Appl. Mech., 69, pp. 610–616)

Z.-Q. Qu[1]

Department of Civil Engineering, University of Arkansas, Fayetteville, AR 72701

The dynamic condensation method ([1]) was successfully extended by Rao [2] to handle the unsymmetric systems with damping. This method is very interesting and useful in the finite element modeling, vibration control, etc. However, one misunderstanding occurred when this approach was utilized in substructure synthesis.

As stated by the author in Sec. 4, the reduced order matrices $[M_R]$ and $[K_R]$ of each substructure in Eqs. (16) and (17) have the form

$$[M_R]=\begin{bmatrix}[0] & [M_{mmR}]\\ -[M_{mmR}] & -[C_{mmR}]\end{bmatrix},\quad [K_R]=\begin{bmatrix}[M_{mmR}] & [0]\\ [0] & [K_{mmR}]\end{bmatrix} \tag{1}$$

in which $[M_{mmR}]$, $[C_{mmR}]$, and $[K_{mmR}]$ are the reduced order mass, damping, and stiffness matrices of order $m\times m$.

Actually, if the reduced order matrices $[M_R]$ and $[K_R]$ are computed from Eqs. (16) and (17) as indicated by the author, these two matrices are generally fully populated and do not have the forms shown in Eq. (1). This will be explained in detail later. Hence one cannot simply convert these two matrices into the displacement space with the explicit forms of the reduced order matrices $[M_{mmR}]$, $[C_{mmR}]$ and $[K_{mmR}]$. If the matrices on the right-hand sides of Eq. (1) are known and those on the left-hand sides are unknowns, the relations shown in this equation are right. However, the problem is how we get the reduced matrices $[M_{mmR}]$, $[C_{mmR}]$, and $[K_{mmR}]$ before we have $[M_R]$ and $[K_R]$.

To simplify the discussion, consider a symmetric problem. After the simplification, the full order matrices $[\bar{M}]$ and $[\bar{K}]$ in Ref. [2] become

$$[\bar{M}]=-\begin{bmatrix}[0] & [M]\\ [M] & [C]\end{bmatrix},\quad [\bar{K}]=\begin{bmatrix}-[M] & [0]\\ [0] & [K]\end{bmatrix}, \tag{2}$$

if the eigenproblem in Sec. 3 rather than the dynamic equations of equilibrium in Sec. 2 is considered. The transformation matrices $[R]$ and $[S]$ are the same and indicated by $[R]$. The corresponding governing equation for the transformation matrix is given by

$$[R]=[\bar{K}_{ss}]^{-1}[([\bar{M}_{sm}]+[\bar{M}_{ss}][R])[M_R]^{-1}[K_R]-[\bar{K}_{sm}]] \tag{3}$$

and the initial approximation is

$$[R]^{(0)}=-[\bar{K}_{ss}]^{-1}[\bar{K}_{sm}]. \tag{4}$$

A very simple numerical example is given to show the form of reduced order matrices $[M_R]$ and $[K_R]$. In this example, the mass, damping, and stiffness matrices are

$$[M]=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\quad [C]=\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\quad [K]=\begin{bmatrix}2&-1&0\\-1&2&-1\\0&-1&1\end{bmatrix}. \tag{5}$$

Two cases that the first and the third degrees of freedom are, respectively, selected as the master degrees of freedom are considered. The resulted reduced order matrices $[M_R]$ and $[K_R]$ from the initial approximation and the first three iterations are listed in Table 1. The results show that reduced order matrices $[M_R]$ and $[K_R]$ obtained from the initial approximation, that is, Guyan condensation, have the forms given in Eq. (2). This conclusion can be proven simply. After two iterations, both reduced order matrices are fully populated. The further discussion on the dynamic condensation of viscously damped, symmetric models may be found in Refs. [3–6].

Table 1 Reduced order matrices $[M_R]$ and $[K_R]$ during iteration

Iteration	Case 1 $[M_R]$		Case 1 $[K_R]$		Case 2 $[M_R]$		Case 2 $[K_R]$	
0	0	−1	−1	0	0	−1	−1	0
	−1	−1	0	1	−1	−0.1111	0	0.3333
1	0	−3	−3	0	0.2469	−1.4815	−1.4815	0
	−3	−1	0	1	−1.4815	−0.1111	0	0.3333
2	−3.3333	−4.6667	−2.4444	0.5556	0.1963	−1.5716	−1.5042	−0.0311
	−4.6667	−1	0.5556	1.5556	−1.5716	−0.1962	−0.0311	0.3512
3	−3.0183	−5.1174	−3.0159	0.6003	0.2750	−1.6296	−1.6196	−0.04670
	−5.1174	−0.6281	0.6003	1.6537	−1.6296	−0.2148	−0.04670	0.3507

References

[1] Qu, Z.-Q., and Fu, Z.-F., 1998, "New Structural Dynamic Condensation Method for Finite Element Models," AIAA J., **36**, pp. 1320–1324.
[2] Rao, G. V., 2002, "Dynamic Condensation and Synthesis of Unsymmetric Structural Systems," ASME J. Appl. Mech., **69**, pp. 610–616.
[3] Rivera, M. A., Singh, M. P., and Suarez, L. E., 1999, "Dynamic Condensation Approach for Nonclassically Damped Structures," AIAA J., **37**, pp. 564–571.
[4] Qu, Z.-Q., and Selvam, R. P., 2000, "Dynamic Condensation Methods for Viscously Damped Models," *Proceedings of the 18th International Modal Analysis Conference & Exhibit* (San Antonio, Texas), Society for Experimental Mechanics, CT, USA, pp. 1752–1757.
[5] Qu, Z.-Q., and Chang, W., 2000, "Dynamic Condensation Method for Viscously Damped Vibration Systems in Engineering," Eng. Struct., **22**, pp. 1426–1432.
[6] Qu, Z.-Q., and Selvam, R. P., 2002, "Efficient Method for Dynamic Condensation of Nonclassically Damped Vibration Systems," AIAA J., **40**, pp. 368–375.

[1]Current address: 2409 Wynncrest Circle, 6205, Arlington, TX 76006.

Closure to "Discussion on 'Dynamic Condensation and Synthesis of Unsymmetrical Systems'" (2003, ASME J. Appl. Mech., 70, p. 784)

G. V. Rao
EMRC 607/900 Barton Centre, M. G. Road, Bangalore, Karnataka 560001, India

I appreciate Zu-Qing Qu for the interest shown and for making useful comments on the contents of my paper.

The reduced order matrices M_R and K_R retain the same form as given in Eq. (1) as the iterations progress. The computation results presented by Zu-Qing Qu seem to be incorrect.

With the help of the procedure in Sec. 3 of my paper, the computations for the first three iterations are carried out on the numerical example cited by Zu-Qing Qu. The results of the first three iterations are given in Table 1 hereunder.

In addition to the above, the eigenvalues are also extracted for the full system and the two cases of master selection. The converged eigenvalues are shown in Table 2 below.

Further to the above, I wish to add here that since the formulation in my paper finally falls into the category of unsymmetric matrices—particularly so in the case of the mass matrix in Eq. (2)—the assumption by Zu-Qing Qu that the transformation matrices [R] and [S] are the same even for a symmetric structure is not valid.

Table 1 M_R and K_R for three iterations

	Case 1				Case 2			
Iteration	M_R		K_R		M_R		K_R	
0	0 −1	1 −1	1 0	0 1	0 −1	1 0.1111	1 0	0 0.333
1	0 −3.0	1 −1	1 0	0 1	0 −1.481	1 −0.111	1 0	0 0.333
2	0 −3.0	1 −0.5165	1 0	0 1	0 −1.50	1 −0.167	1 0	0 0.333
3	0 −4.1111	1 −0.4444	1 0	0 1	0 −1.572	1 −0.161	1 0	0 0.333
10	0 −4.765	1 −0.516	1 0	0 1	0 −1.588	1 −0.172	1 0	0 0.333

Table 2 Eigenvalues

Mode no.	Full system	Case 1	Case 2
1	−0.0542+J 0.4549	−0.0507+J 0.4553	−0.05419+J0.4549
2	−0.0542−J 0.4549	−0.0507−J 0.4553	−0.05419−j 0.4549
3	−0.3336+J 1.2374		
4	−0.3336−J 1.2374		
5	−0.11225+J 1.6996		
6	−0.11225−J 1.6996		

