ASME
SETTING THE STANDARD

Journal of Applied Mechanics

Published Bimonthly by ASME

VOLUME 74 • NUMBER 1 • JANUARY 2007

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published bimonthly (Jan., Mar., May, July, Sept., Nov.) by The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300. **CHANGES OF ADDRESS** must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.

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.

Journal of Applied Mechanics Volume 74, Number 1 JANUARY 2007

TECHNICAL BRIEFS

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

Alloys Index, Aluminum Industry Abstracts, Applied Science & Technology Index, 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

Disappearance Conditions of Stress Singularities for Anisotropic Bimaterial Half-Plane Wedges Under Antiplane Shear

Chuan-I Liu[1]
Department of Structure Analysis,
Aerospace Industrial Development Corporation,
Taichung, Taiwan, ROC
e-mail: chuaniliu@gmail.com

Ching-Hwei Chue
Department of Mechanical Engineering,
National Cheng Kung University,
Tainan, Taiwan, ROC

Based on the anisotropic elasticity theory and Lekhnitskii's complex potential functions, the analytical eigenequations of anisotropic bimaterial half-plane wedges under antiplane shear are derived in brief forms. The boundary surfaces of half-plane wedges can be combinations of free and/or clamped edges. The disappearance conditions of stress singularities can be obtained directly from the derived eigenequations, which can be applied to improve the safety of the structures. The interesting phenomenon on the periodic appearance of the singularity orders is proposed and discussed as well.
[DOI: 10.1115/1.1989356]

1 Introduction

In elastic wedge structures, the stresses may be singular due to geometric and/or material discontinuities. This singularity stress is responsible for the initiation of delamination in composite structures. For the antiplane problem, Ma and Hour [1] used the Mellin transform to study the antiplane stress singularities of anisotropic wedges. They found that the order of stress singularity is real for a general anisotropic bimaterial wedge of the antiplane problem subjected to different boundary conditions (i.e., free-free, clamped-clamped, and free-clamped). The problems of cracking are also extensively investigated. For example, Ma and Hour [2] studied antiplane problems in anisotropic materials with an inclined crack terminating at a bimaterial interface. Shahani [3] investigated the anisotropic finite wedge under antiplane deformation. Pageau et al. [4] obtained the numerical antiplane singular stress field of anisotropic multimaterial wedges and junctions. Recently, Xie and Chaudhuri [5] proposed an eigenfunction approach for analyzing the three-dimensional asymptotic stress field of a symmetric bimaterial wedge under antiplane shear loading. A similar problem with an asymmetric bimaterial pie-shaped wedge has been analyzed by Chiu and Chaudhuri [6].

The disappearance of stress singularities is very important in engineering practice. It can be applied to improve the safety of the structures. Hu et al. [7] studied the disappearance conditions of stress singularities near the vertex of bonding edges in single-lap joints. Wu [8] designed the singularity-free bimaterial components. They were focused on isotropic material. For anisotropic, Chue and Liu [9] used the numerical contours to obtain the disappearance condition of stress singularities in composite laminates. However, the brief formulation for disappearance conditions of stress singularities will be more practical in an engineering application.

This paper uses the Lekhnitskii's formulations [10] to derive generalized eigenequations, which govern the antiplane stress singularity orders for the anisotropic bimaterial half-plane wedge $(\alpha, 180\ \text{deg}-\alpha)$ shown in Fig. 1. This structure usually appears in multilayer laminate with a chamfer angle on the edge. In addition, the material principal axes $\hat{x}_k$ axis $(k=1,2)$ are assumed to lie in the x-y plane and makes an angle η_k $(-90\ \text{deg} \leqslant \eta_k \leqslant 90\ \text{deg})$ with the positive x axis.

Two interesting points related to the stress singularity order will be discussed in detail. First, the disappearance condition of the stress singularity can be obtained directly from the derived eigenequations. It is very important and useful for designing the structures to improve the safety of the structures. Second, the period of the singularity orders varied by fiber orientation is proposed and discussed analytically in detail.

2 Formulations of a Two-Bonded Anisotropic Wedge in Antiplane Fields

2.1 General Formulation. Consider two anisotropic wedge occupied regions Ω_1 and Ω_2 with wedge angle α and 180 deg $-\alpha$, respectively, that they are perfectly bonded along the common edge (Fig. 1) and subjected to longitudinal shear loading. Note that the wedge angle α of region Ω_1 is kept in the interval $0\ \text{deg} \leqslant \alpha \leqslant 180\ \text{deg}$. Assume that the x-y plane is a symmetric plane and the inplane and antiplane stress fields are decoupled. For the antiplane problem, only the stress $\tau_{yz}^{(k)}$, $\tau_{xz}^{(k)}$, and displacement $w^{(k)}$ are considered, in which the superscript k denotes the region Ω_k. The strain components are related to the stresses by

$$\begin{aligned} \gamma_{yz}^{(k)} &= s_{44}^{(k)}\tau_{yz}^{(k)} + s_{45}^{(k)}\tau_{xz}^{(k)} \\ \gamma_{xz}^{(k)} &= s_{45}^{(k)}\tau_{yz}^{(k)} + s_{55}^{(k)}\tau_{xz}^{(k)} \end{aligned} \quad (k=1,2) \tag{1}$$

where $s_{ij}^{(k)}$ are the components of elastic compliance matrix referred to global xyz-coordinate system. Based on Lekhnitskii's formulations [10], the stresses and displacement corresponding to cylindrical coordinates (r,θ) can be expressed in terms of complex function $\phi_k(z_k)$ as follows:

$$\tau_{\theta z}^{(k)} = -2\ \text{Re}[(\cos\theta + \mu_k \sin\theta)\phi_k'(z_k)] \tag{2}$$

$$w^{(k)} = 2\ \text{Re}[t_k \phi_k(z_k)] \tag{3}$$

where

$$z_k = x + \mu_k y = r(\cos\theta + \mu_k \sin\theta) \equiv r\ \zeta_k(\theta) \tag{4}$$

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 1, 2004; final manuscript received February 13, 2005. Assoc. Editor: Z. Suo. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

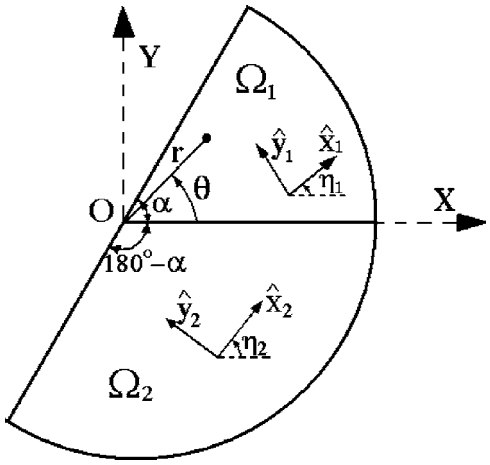


Fig. 1 Geometry of two dissimilar anisotropic materials for half-plane wedge

$$\mu_k = \frac{s_{45}^{(k)} + i[s_{44}^{(k)} s_{55}^{(k)} - (s_{45}^{(k)})^2]^{1/2}}{s_{55}^{(k)}} \tag{5}$$

and

$$t_k = s_{45}^{(k)} - s_{44}^{(k)}/\mu_k = i\sqrt{s_{44}^{(k)} s_{55}^{(k)} - (s_{45}^{(k)})^2}. \tag{6}$$

Thus, the anisotropic antiplane problems are reduced to the determination of complex function $\phi_k(z_k)$.

In order to evaluate the singular stress behavior at the apex, the complex stress potentials $\phi_k(z_k)$ are expanded in following forms:

$$\phi_k(z_k) = c_{1k} z_k^{\lambda} + c_{2k} \bar{z}_k^{\lambda} \quad (k = 1,2) \tag{7}$$

where c_{1k} and c_{2k} are unknown complex constants and λ is a complex eigenvalue to be determined. Substituting Eqs. (4) and (7) into Eq. (2), the stress can be expended in the following form:

$$\tau_{\theta z}^{(k)} = -(c_{1k}\lambda r^{\lambda-1}\zeta_k^{\lambda} + c_{2k}\bar{\lambda} r^{\bar{\lambda}-1}\zeta_k^{\bar{\lambda}} + \bar{c}_{1k}\bar{\lambda} r^{\bar{\lambda}-1}\bar{\zeta}_k^{\bar{\lambda}} + \bar{c}_{2k}\lambda r^{\lambda-1}\bar{\zeta}_k^{\lambda}) \quad (k = 1,2). \tag{8}$$

Based on the complex theory, Eq. (8) can be arranged as

$$\tau_{\theta z}^{(k)} = -2\,\mathrm{Re}[\lambda r^{\lambda-1}(c_{1k}\zeta_k^{\lambda} + \bar{c}_{2k}\bar{\zeta}_k^{\lambda})] \quad (k = 1,2). \tag{9}$$

Similarly, the displacement of Eq. (3) becomes

$$w^{(k)} = 2\,\mathrm{Re}[r^{\lambda}(c_{1k}t_k\zeta_k^{\lambda} + \bar{c}_{2k}\bar{t}_k\bar{\zeta}_k^{\lambda})] \quad (k = 1,2). \tag{10}$$

Here $(\lambda-1)$ with $-1<\mathrm{Re}(\lambda-1)<0$ is called the stress singularity order and smaller λ gives the stronger stress singularity.

The orthotropic materials with incline principal axes are used frequently in composite structures. Therefore, consider that the materials of wedges are orthotropic, in which the principal axes are denoted as $\hat{x}_k$, $\hat{y}_k$, and $\hat{z}_k$ for each region Ω_k $(k=1,2)$. The $\hat{z}_k$ axis coincides with the z axis and the $\hat{x}_k$ axis makes an angle η_k $(-90\ \mathrm{deg} \leqslant \eta_k \leqslant 90\ \mathrm{deg})$ with the positive x axis. If $\hat{s}_{44}^{(k)}$ and $\hat{s}_{55}^{(k)}$ are the elastic constants of orthotropic materials referred to $\hat{x}_k\hat{y}_k\hat{z}_k$ coordinates $(\hat{s}_{45}^{(k)}=0)$, the elastic constants referred to xyz coordinates are obtained by transformation

$$\begin{aligned} s_{44}^{(k)} &= \hat{s}_{44}^{(k)}\cos^2(-\eta_k) + \hat{s}_{55}^{(k)}\sin^2(-\eta_k) \\ s_{45}^{(k)} &= (\hat{s}_{44}^{(k)} - \hat{s}_{55}^{(k)})\cos(-\eta_k)\sin(-\eta_k) \quad (k = 1,2). \\ s_{55}^{(k)} &= \hat{s}_{55}^{(k)}\cos^2(-\eta_k) + \hat{s}_{44}^{(k)}\sin^2(-\eta_k) \end{aligned} \tag{11}$$

Substituting Eq. (11) into Eq. (5), μ_k leads to

$$\mu_k = \frac{-\sqrt{m_k}\cos\eta_k + i\sin\eta_k}{\sqrt{m_k}\sin\eta_k + i\cos\eta_k} \tag{12}$$

where $m_k = \hat{s}_{44}^{(k)}/\hat{s}_{55}^{(k)}$. For isotropic material, $m_k=1$.

Therefore, the parameter t_k defined in Eq. (6) can be reduced to

$$t_k = i\sqrt{\hat{s}_{44}^{(k)}\hat{s}_{55}^{(k)}}. \tag{13}$$

Based on Eqs. (4) and (12), the function $\zeta_k(\theta)$ is given by

$$\zeta_k(\theta) = \frac{[(1-m_k)\sin(\theta-\eta_k)\sin\eta_k + \cos\theta] + i\sqrt{m_k}\sin\theta}{m_k\sin^2\eta_k + \cos^2\eta_k}. \tag{14}$$

2.2 Essential Character of Function $\zeta_k(\theta)$. In this paper, most of the emphasis is placed on the study of the stress singularity order. The stress $\tau_{\theta z}^{(k)}$ of Eq. (9) and displacement $w^{(k)}$ of Eq. (10) applied on the boundary are used to derive the eigenequations. Consequently, the function $\zeta_k(\theta)$ will play an extremely important role in discussing the singularity behavior. A close examination on the nature of function $\zeta_k(\theta)$ is necessary.

Let the absolute value and principal argument of $\zeta_k(\theta)$ be denoted as $R_k(\theta)$ and $\Theta_k(\theta)$, respectively. They are given below

$$R_k(\theta) \equiv |\zeta_k(\theta)| = \sqrt{\frac{(1+m_k)+(1-m_k)\cos[2(\theta-\eta_k)]}{(1+m_k)+(1-m_k)\cos(2\eta_k)}} \tag{15}$$

$$\mathrm{Tan}[\Theta_k(\theta)] \equiv \mathrm{Tan}[\mathrm{Arg}\,\zeta_k(\theta)] = \frac{\sqrt{m_k}\sin\theta}{(1-m_k)\sin(\theta-\eta_k)\sin\eta_k + \cos\theta}. \tag{16}$$

The principal argument $\Theta_k(\theta)$ depends on m_k and η_k and is a function of θ. While the interval of angle θ is defined between 0 deg and 180 deg, then $0 \leqslant \Theta_k(\theta) \leqslant 180$ deg. The value of $\Theta_k(\theta)$ with $-180\ \mathrm{deg} \leqslant \theta \leqslant 0$ deg can be obtained from $\Theta_k(\theta)$ with $0\ \mathrm{deg} \leqslant \theta \leqslant 180$ deg.

Figures 2 and 3 plot the variations of the principal argument $\Theta_k(\theta)$ with η_k at different θ and m_k, respectively. From these figures or Eq. (16), some conclusions can be made as follows:

1. For all η_k, Fig. 3 shows that the argument $\Theta_k(\theta) \to 0$ deg as $\theta \to 0$ deg and $\Theta_k(\theta) \to 180$ deg as $\theta \to 180$ deg.
2. As the parameter m_k $(=\hat{s}_{44}^{(k)}/\hat{s}_{55}^{(k)})$ approaches infinity or zero, the difference between extreme values of $\Theta_k(\theta)$ becomes larger (i.e., $\cong 180$ deg). It means that the principal argument depends strongly on η_k.
3. While the material is degenerated to isotropic material (i.e., $m_k=1$), $\Theta_k(\theta)=\theta$ for all η_k.
4. The argument $\Theta_k(\theta)$ in Fig. 2 is a period function of η_k. Its period is 180 deg. However, only the half period $(\theta/2$ -90 deg$) \leqslant \eta_k \leqslant \theta/2$ is needed. For example, as $\theta=60$ deg, the argument $\Theta_k(\theta)$ in $-60\ \mathrm{deg} \leqslant \eta_k \leqslant 30$ deg is considered. The other values of $\Theta_k(\theta)$ can be deduced from this half period. Consequently, for given m_k, it can be proved that the principal argument $\Theta_k(\theta)$ of Eq. (16) remains unchanged when the fiber orientation η_k is replaced by $(\theta-\eta_k)$, i.e.,

$$\Theta_k(\theta)_{\eta_k} = \Theta_k(\theta)_{\theta-\eta_k}. \tag{17}$$

This conclusion is very important in discussing the repeated occurrence of the singularity order.

In the following section, the application of these relations to anisotropic bimaterial wedges will be discussed in detail.

2.3 Half-Plane Wedges $(\alpha, 180\ \mathrm{deg}-\alpha)$. Consider the case that the half-plane wedges perfectly bonded along the common edge (i.e., x axis in Fig. 1). Using Eqs. (9) and (10), the continuity conditions of the stress $[\tau_{\theta z}^{(1)}(r,0) = \tau_{\theta z}^{(2)}(r,0)]$ and displacement $[w^{(1)}(r,0) = w^{(2)}(r,0)]$ along the interface will result in the following relations:

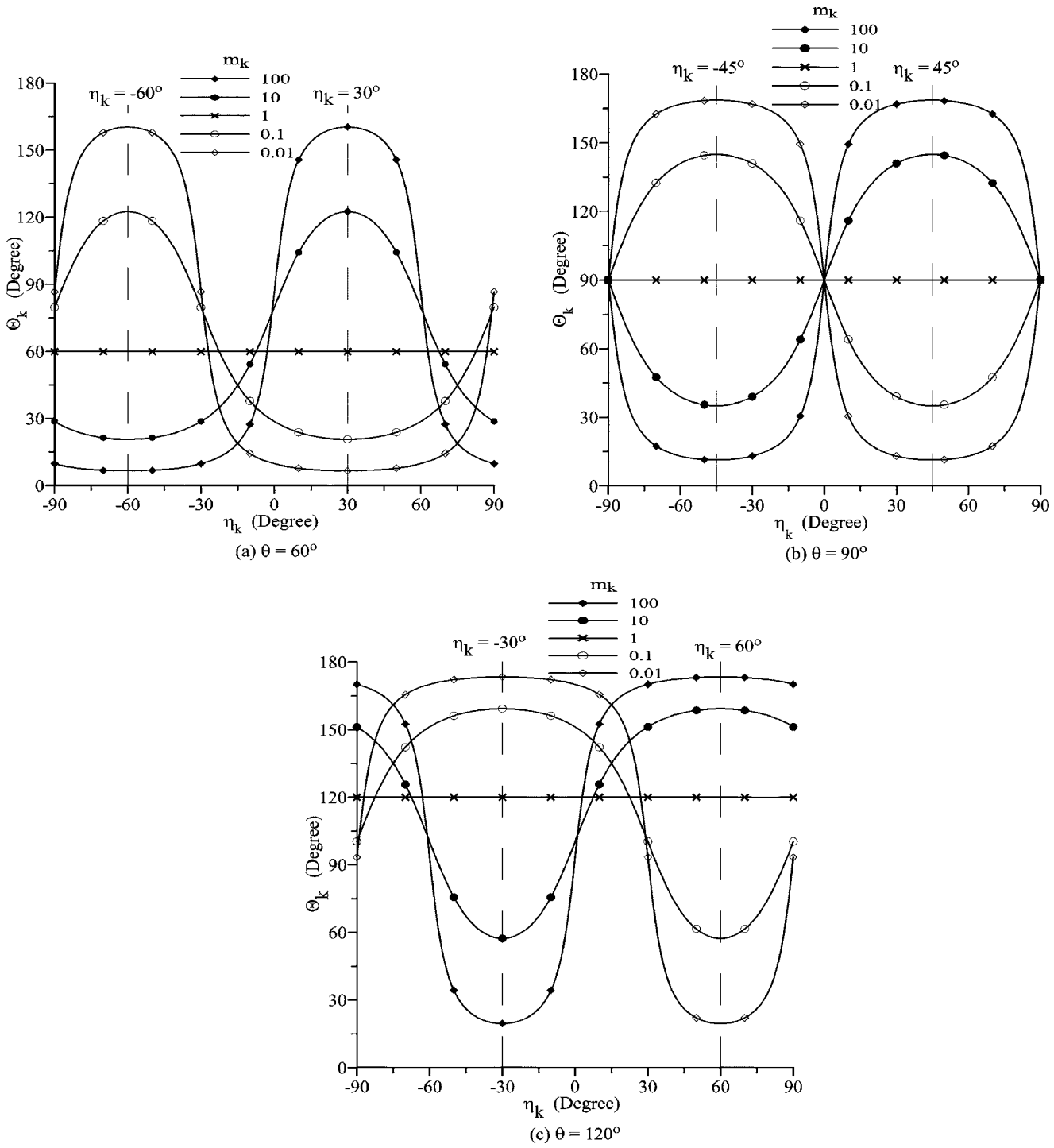


Fig. 2 The variations of $\Theta_k(\theta)$ with fiber orientation η_k when (a) θ=60 deg; (b) θ=90 deg; (c) θ=120 deg

$$c_{11} + \bar{c}_{21} - c_{12} - \bar{c}_{22} = 0 \tag{18}$$

$$c_{11} - \bar{c}_{21} - \omega\, c_{12} + \omega\, \bar{c}_{22} = 0. \tag{19}$$

Here the relations $\zeta_1(\theta=0)=\zeta_2(\theta=0)=1$ from Eq. (14) have been used. The parameter ω is the ratio of the elastic constants

$$\omega = \sqrt{\hat{s}_{44}^{(2)}\hat{s}_{55}^{(2)} / \hat{s}_{44}^{(1)}\hat{s}_{55}^{(1)}}. \tag{20}$$

While the materials of two wedges are the same with different fiber orientation, $\omega=1$.

The following four combinations of boundary conditions are considered:

2.3.1 Free-Free Boundary Conditions. Both of the edges are traction free, i.e., $\tau_{\theta z}^{(1)}(r, \theta=\alpha)=0$ and $\tau_{\theta z}^{(2)}(r, \theta=\alpha-\pi)=0$. Based on Eq. (9), they reduce to the relations, respectively,

$$c_{11}[\zeta_1(\alpha)]^\lambda + \bar{c}_{21}\overline{\zeta_1(\alpha)^\lambda} = 0 \tag{21}$$

$$c_{12}[\zeta_2(\alpha-\pi)]^\lambda + \bar{c}_{22}\overline{\zeta_2(\alpha-\pi)^\lambda} = 0. \tag{22}$$

Owing to Eqs. (15) and (16), the functions $\zeta_1(\alpha)$ and $\zeta_2(\alpha-\pi)$ can be expanded as follows:

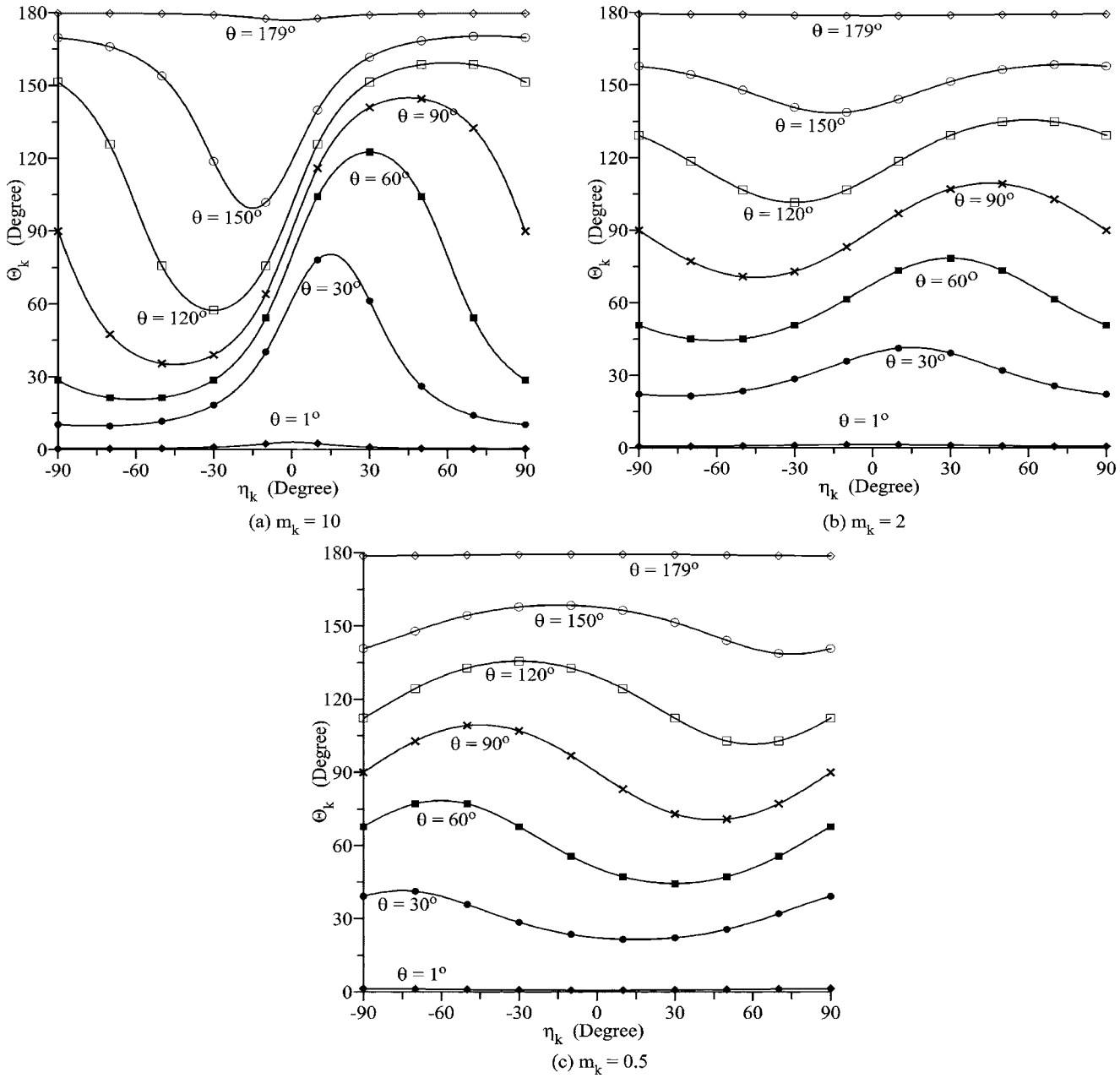


Fig. 3 The variations of $\Theta_k(\theta)$ with fiber orientation η_k when (a) m_k=10; (b) m_k=2; (c) m_k=0.5

$$\zeta_1(\alpha) = R_1(\alpha)e^{i\Theta_1(\alpha)} \tag{23}$$

$$\zeta_2(\alpha-\pi) = \cos(\alpha-\pi) + \mu_2 \sin(\alpha-\pi) = -\zeta_2(\alpha) = R_2(\alpha)e^{i[\Theta_2(\alpha)-\pi]}. \tag{24}$$

Equations (18), (19), (21), and (22) yield four linear homogeneous equations with unknowns c_{11}, $\bar{c}_{21}$, c_{12}, and $\bar{c}_{22}$. They can be rewritten in the following matrix form:

$$\begin{bmatrix} e^{i\lambda\Theta_1(\alpha)} & e^{-i\lambda\Theta_1(\alpha)} & 0 & 0 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -\omega & \omega \\ 0 & 0 & e^{i[\Theta_2(\alpha)-\pi]\lambda} & e^{-i[\Theta_2(\alpha)-\pi]\lambda} \end{bmatrix} \begin{bmatrix} c_{11} \\ \bar{c}_{21} \\ c_{12} \\ \bar{c}_{22} \end{bmatrix} = 0. \tag{25}$$

For nontrivial solutions of c_{ij} $(i,j=1,2)$, the determinant of the matrix must vanish. It reduces to the following eigenequation:

$$\tan\{\lambda[\Theta_2(\alpha)-\pi]\}\cot[\lambda\Theta_1(\alpha)] = \omega. \tag{26}$$

The stress singularity order $(\lambda-1)$ depends on $\Theta_k(\alpha)$ and ω. From the essential character of $\Theta_k(\theta)$ (i.e., Eq. (17)), the singularity order will not be changed if the fiber orientation (η_1, η_2) is replaced by $(\alpha-\eta_1, \alpha-\eta_2)$, $(\alpha-\eta_1, \eta_2)$, or $(\eta_1, \alpha-\eta_2)$. The numerical results of the singularity order $(\lambda-1)$ are listed in Table 1 when α=60 deg. The material properties are assumed that $\hat{s}_{44}^{(1)} = 2\hat{s}_{55}^{(1)}$, $\hat{s}_{44}^{(2)}, = 2\hat{s}_{55}^{(2)}$, and $\hat{s}_{44}^{(1)} = \hat{s}_{44}^{(2)}$. It is shown that the singularity order −0.099 246 at (η_1, η_2)=(15 deg, 80 deg) is equal to the singularity orders at $(\alpha-\eta_1, \alpha-\eta_2)$=(45 deg, −20 deg), $(\alpha-\eta_1, \eta_2)$ =(45 deg, 80 deg), and $(\eta_1, \alpha-\eta_2)$=(15 deg, −20 deg). The repeated occurrence of stress singularity orders can be observed. It is useful to select the multiple fiber orientation for structure design under an invariable singularity order.

In most engineering applications, both of the materials are usu-

Table 1 The antiplane stress singularity orders ($\lambda-1$) of a 60 deg–120 deg wedge with free-free edges. (N.S.: no singularity).

Region Ω_1	Region Ω_2	
η_1	$\eta_2=-20$ deg	$\eta_2=80$ deg
−90 deg	0.027914 (N.S.)	0.027914 (N.S.)
−75 deg	0.056733 (N.S.)	0.056733 (N.S.)
−60 deg	0.066612 (N.S.)	0.066612 (N.S.)
−45 deg	0.056733 (N.S.)	0.056733 (N.S.)
−30 deg	0.027914 (N.S.)	0.027914 (N.S.)
−15 deg	−0.015647	−0.015647
0 deg	−0.063162	−0.063162
15 deg	−0.099246	−0.099246
30 deg	−0.112454	−0.112454
45 deg	−0.099246	−0.099246
60 deg	−0.063162	−0.063162
75 deg	−0.015647	−0.015647
90 deg	0.027914 (N.S.)	0.027914 (N.S.)

ally the same but with different fiber orientations (i.e., $\hat{s}_{44}^{(1)}=\hat{s}_{44}^{(2)}$, $\hat{s}_{55}^{(1)}=\hat{s}_{55}^{(2)}$, $\eta_1\neq\eta_2$). It leads to $m_1=m_2$, $\omega=1$ but $\mu_1\neq\mu_2$. Substituting $\omega=1$ into Eq. (26), the eigenequation can be simplified as

$$\sin\{[\Theta_2(\alpha)-\Theta_1(\alpha)-\pi]\lambda\}=0. \tag{27}$$

The solution is given by

$$\lambda=\frac{n\pi}{\Theta_2(\alpha)-\Theta_1(\alpha)-\pi}\quad (n=\text{integer}). \tag{28}$$

It is important to examine the disappearance conditions of stress singularity. The values of $[\Theta_2(\alpha)-\Theta_1(\alpha)]$, ranging from $-\pi$ to π, can be obtained from Eq. (16) while m_1 (or m_2, since $m_1=m_2$), η_k, and α are given. The interesting values of $(\lambda-1)$ are in the interval $-1<\lambda-1<0$. Figure 4 plots the variations of $(\lambda-1)$ with $\Theta_2(\alpha)-\Theta_1(\alpha)$. As $\Theta_2(\alpha)<\Theta_1(\alpha)$, the stress field is singular. The strongest stress singularity occurs when $\Theta_2(\alpha)-\Theta_1(\alpha)=-\pi$ and the order is −0.5. Based on Fig. 2 or Eq. (16), the case of $\lambda-1=-0.5$ exists in a combination of $\eta_1=\alpha/2$ and $\eta_2=\alpha/2-\pi/2$ for $m_k\approx\infty$ or $\eta_1=\alpha/2-\pi/2$ and $\eta_2=\alpha/2$ for $m_k\approx0$. On the other hand, the disappearance conditions of stress singularity is derived as

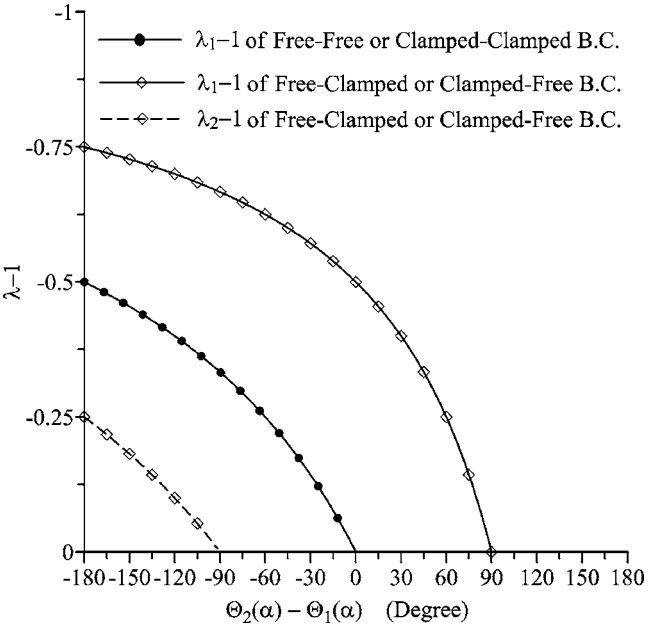


Fig. 4 The variations of stress singularity $\lambda-1$ with $\Theta_2-\Theta_1$ ($\omega=1$)

$$\Theta_2(\alpha)\geq\Theta_1(\alpha). \tag{29}$$

By using Eq. (16), the condition for no-singular stress of Eq. (29) can be expressed in the following form:

$$(1-m_1)\sin(\alpha-\eta_1)\sin\eta_1\geq(1-m_1)\sin(\alpha-\eta_2)\sin\eta_2. \tag{30}$$

This inequality will be discussed in three parts:

(1) If $m_1=m_2>1$, Eq. (30) can be simplified as follows:

$$\sin(\alpha-\eta_1-\eta_2)\sin(\eta_1-\eta_2)\leq0. \tag{31}$$

The results for all combinations of η_1 and η_2 are presented in Fig. 5(*a*). The horizontal and vertical axes plotted with dotted lines are denoted as η_1 and η_2, respectively. This plot can be divided into four regions and four lines ($\Gamma_1\sim\Gamma_4$) to be discussed. For region 1 (bonded by points A, B, O, and H), there are $0\text{ deg}<\eta_1-\eta_2<180$ deg, $0\text{ deg}<\eta_1+\eta_2<180$ deg, and $0\text{ deg}\leq\alpha\leq180$ deg. These lead to $\sin(\eta_1-\eta_2)>0$ and $-180\text{ deg}<\alpha-\eta_1-\eta_2<180$ deg. In order to create a no-singular stress field, the condition $\sin(\alpha-\eta_1-\eta_2)\leq0$ based on Eq. (31) is needed. Hence, the inequality is expanded as follows

$$0\leq\alpha\leq\eta_1+\eta_2\quad\text{for region 1 in Fig. (5(}a\text{))}. \tag{32}$$

It means that if the combination of η_1 and η_2 locates in region 1, the designer can cut the wedge angle according as Eq. (32) to improve the safety of structures. For example, if the material properties and principal orientations are as $m_1=m_2=2$, $\eta_1=60$ deg, and $\eta_2=20$ deg, the wedge angle of region Ω_1 need to be cut in the interval $0\text{ deg}\leq\alpha\leq80$ deg for avoiding the stress concentration. The order $(\lambda-1)=0.027\,914$ in Table 1 with $(\eta_1,\eta_2)=(90\text{ deg},80\text{ deg})$ is an example of a nonsingularity case. The case $(\eta_1,\eta_2)=(30\text{ deg},-20\text{ deg})$, which does not satisfy Eq. (32), causes singular stress field with order $(\lambda-1)=-0.112\,454$. Similarly, we can obtain the disappearance condition for the other three regions as

$$\eta_1+\eta_2\leq\alpha\leq180\text{ deg}\quad\text{for region 2 in Fig. (5(}a\text{))} \tag{33}$$

$$0\leq\alpha\leq\eta_1+\eta_2+180\text{ deg}\quad\text{for region 3 in Fig. (5(}a\text{))} \tag{34}$$

$$\eta_1+\eta_2+180\text{ deg}\leq\alpha\leq180\text{ deg}\quad\text{for region 4 in Fig. (5(}a\text{))}. \tag{35}$$

Since the condition of $\eta_1=\eta_2$ on the lines Γ_1 and Γ_3 reduces the wedge to a single-material half-plane, there is no singularity regardless of angle α in this case. The line Γ_2 denotes $\eta_1-\eta_2$ for $\eta_1<0$ and $\eta_2>0$. Taking these relations into Eq. (31), it reduces to $0\text{ deg}\leq\alpha\leq180$ deg. It means that the inequality of Eq. (31) is always satisfied for arbitrary α. For the case on line Γ_4 which denotes $\eta_1=-\eta_2$ for $\eta_1>0$ and $\eta_2<0$, Eq. (31) is reduced to $\sin\alpha\leq0$. Then α must to be 0 deg or 180 deg and the wedge becomes a single-material half-plane.

(2) If $m_1=m_2<1$, the disappearance conditions for nonsingular antiplane stress can be obtained in the same way as case (1) and are shown in Fig. 5(*b*).

(3) If $m_1=m_2=1$, Eq. (30) is always satisfied. Since $\omega=1$ has been assumed, the bimaterial wedges contain two identical isotropic materials. The stress field is nonsingular.

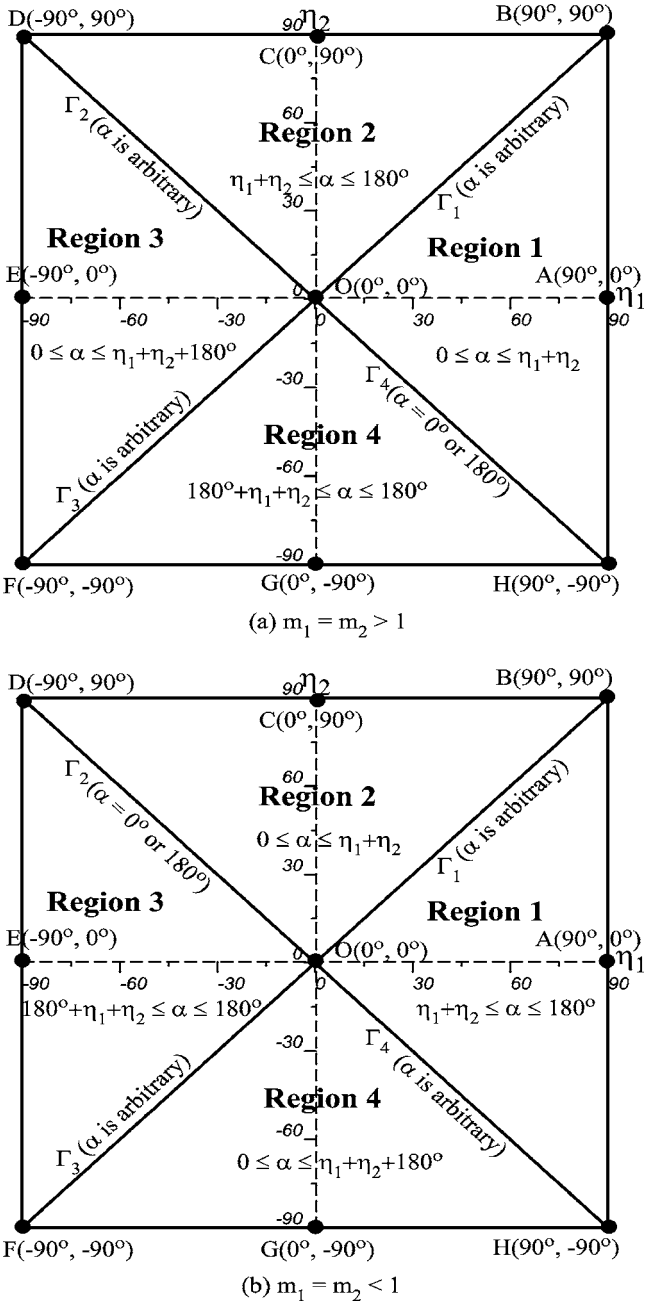


Fig. 5 The conditions of nonsingular antiplane stress field for half-plane wedges with free-free or clamped-clamped edges when (*a*) $m_1=m_2>1$ and (*b*) $m_1=m_2<1$

The eigenequation of Eq. (26) can also be used to study the problem of isotropic bimaterial half-plane. Assume that the materials are different, i.e., $\omega\neq 1$. Let $\hat{s}_{44}^{(1)}=\hat{s}_{55}^{(1)}=s^{(1)}$, $\hat{s}_{44}^{(2)}=\hat{s}_{55}^{(2)}=s^{(2)}$, and $m_1=m_2=1$. From Fig. 2 or Eq. (16), it gives $\Theta_1(\alpha)=\Theta_2(\alpha)=\alpha$. Taking these into Eq. (26), the eigenequation reduces as following:

$$\frac{\sin(\lambda\pi)}{\sin[\lambda(2\alpha-\pi)]}=-\frac{Q_{\text{iso}}-1}{Q_{\text{iso}}+1} \tag{36}$$

where $Q_{\text{iso}}=s^{(2)}/s^{(1)}$. It is exactly the same as Eq. (23) in Ma and Hour [1] if the wedge angles are $(\alpha, 180\ \text{deg}-\alpha)$.

2.3.2 Clamped-Clamped Boundary Conditions. Based on Eq. (10), the clamped-clamped boundary conditions (i.e., $w^{(1)}(r,\alpha)=0$ and $w^{(2)}(r,\alpha-\pi)=0$) reduce to the relations, respectively,

$$c_{11}\zeta_1^{\lambda}(\alpha)-\bar{c}_{21}\overline{\zeta_1(\alpha)}^{\lambda}=0 \tag{37}$$

$$c_{12}[\zeta_2(\alpha-\pi)]^{\lambda}-\bar{c}_{22}\overline{\zeta_2(\alpha-\pi)}^{\lambda}=0. \tag{38}$$

Combining Eqs. (18), (19), (37), and (38), a system of four homogeneous equations is formed. It can be further simplified to the following form:

$$\tan\{\lambda[\Theta_2(\alpha)-\pi]\}\cot[\lambda\Theta_1(\alpha)]=\omega^{-1}. \tag{39}$$

Following the same procedure of Sec. 2.3.1, the singularity order remains the same when the fiber orientations are changed from (η_1,η_2) to $(\alpha-\eta_1,\alpha-\eta_2)$, $(\alpha-\eta_1,\eta_2)$ or $(\eta_1,\alpha-\eta_2)$.

If the bimaterials are the same, i.e., $\omega=1$, the eigenequation of Eq. (39) is identical with Eq. (27). Consequently, Eq. (28) is the solution and Fig. 4 shows the singularity order graphically. The disappearance conditions for non-singular antiplane stress are shown in Fig. 5.

While the bonded wedges contain two different isotropic materials ($m_1=m_2=1, \omega\neq 1$), the eigenequation of Eq. (39) is rearranged as follows:

$$\frac{\sin(\lambda\pi)}{\sin[\lambda(2\alpha-\pi)]}=-\frac{Q_{\text{iso}}^{-1}-1}{Q_{\text{iso}}^{-1}+1}, \tag{40}$$

where $Q_{\text{iso}}=s^{(2)}/s^{(1)}$. Equation (40) is exactly the same as Eq. (38) in Ma and Hour [1] if the wedge angles are $(\alpha,\pi-\alpha)$. Comparing Eqs. (36) and (40), the singularity orders with free-free boundary surfaces are the same with clamped-clamped boundary surfaces after interchanging two isotropic materials. This conclusion is in agreement with Chiu and Chaudhuri [6].

2.3.3 Clamped-Free Boundary Conditions. The characteristic matrix is given by combining Eqs. (18), (19), (22), and (37). The eigenequation becomes

$$\tan\{\lambda[\Theta_2(\alpha)-\pi]\}\tan[\lambda\Theta_1(\alpha)]=-\omega. \tag{41}$$

It can be observed again that the singularity order remains the same when the directions of $\hat{x}_k$-axis locate at (η_1,η_2), $(\alpha-\eta_1,\alpha-\eta_2)$, $(\alpha-\eta_1,\eta_2)$, or $(\eta_1,\alpha-\eta_2)$.

While identical materials are used ($\omega=1$) with different fiber orientations, Eq. (41) reduces to

$$\cos\{[\Theta_2(\alpha)-\Theta_1(\alpha)-\pi]\lambda\}=0. \tag{42}$$

The root of Eq. (42) is

$$\lambda=\frac{(2n+1)\pi}{2[\Theta_2(\alpha)-\Theta_1(\alpha)-\pi]}\quad (n=\text{integer}). \tag{43}$$

Figure 4 plots the variations of $(\lambda-1)$ with $\Theta_2(\alpha)-\Theta_1(\alpha)$. The dotted line represents the second root of $(\lambda-1)$. The disappearance condition of stress singularity can be expressed as the following form:

$$\pi/2\leqslant\Theta_2(\alpha)-\Theta_1(\alpha)\leqslant\pi. \tag{44}$$

According to Figs. 2 and 3, nonsingular stress field can be reached by properly selecting the Ω_1 wedge angle α, elastic constant ratio m_k and fiber orientations (η_1,η_2). It shows while the parameter $m_k(=\hat{s}_{44}^{(k)}/\hat{s}_{55}^{(k)})$ is far from unity, the disappearance of stress singularity is possibly reached. For example, the case of $\alpha=60$ deg, $\eta_1=-60$ deg, $\eta_2=30$ deg, and $m_1=m_2=10$ leads to $\Theta_2-\Theta_1=122.58\ \text{deg}-20.69\ \text{deg}=101.89$ deg. Then nonsingular stress field is expected. Furthermore, if the materials have the same orientations of principal axes (i.e., $\Theta_2(\alpha)-\Theta_1(\alpha)=0$) which reduces one-material half-plane wedge, the stress singularity order $\lambda-1$ is equal to -0.5 regardless of $\hat{s}_{44}/\hat{s}_{55}$. The strongest singularity order is $\lambda-1=-0.75$ as $\Theta_2(\alpha)-\Theta_1(\alpha)=-\pi$. There are two roots for

$-\pi < \Theta_2(\alpha) - \Theta_1(\alpha) < -\pi/2$.

For a dissimilar isotropic wedge, the eigenequation Eq. (41) can be simplified as follows:

$$\frac{\cos(\lambda\pi)}{\cos[(2\alpha-\pi)\lambda]} = -\frac{Q_{\text{iso}}-1}{Q_{\text{iso}}+1}. \tag{45}$$

2.3.4 Free-Clamped Boundary Conditions. For free-clamped boundary, Eqs. (18), (19), (21), and (38) are used to form the characteristic matrix. The eigenequation is given by

$$\tan\{\lambda[\Theta_2(\alpha)-\pi]\}\tan[\lambda\Theta_1(\alpha)] = -\omega^{-1}. \tag{46}$$

The repeated occurrence of the stress singularity can also be observed. While the same materials are used ($\omega=1$) with different fiber orientations, the eigenequation leads to be identical with Eq. (42).

Consider that the materials 1 and 2 are two different isotropic materials ($\omega \neq 1$), that the eigenequation of Eq. (46) can be rewritten as follows:

$$\frac{\cos(\lambda\pi)}{\cos[(2\alpha-\pi)\lambda]} = -\frac{Q_{\text{iso}}^{-1}-1}{Q_{\text{iso}}^{-1}+1}, \tag{47}$$

which is exactly the same as Eq. (28) in Ma and Hour [1] if the wedge angles are $(\alpha, \pi-\alpha)$. Comparing Eqs. (45) and (47), the singularity orders with free-clamped boundary surfaces are the same with clamped-free boundary surfaces after interchanging two isotropic materials. It is coincident with Chiu and Chaudhuri [6].

3 Conclusions

This paper has derived the general closed-form solutions of eigenequations for half-plane wedges under antiplane shear. The boundary surfaces of half-plane wedges can be the combinations of free and/or clamped edges. The disappearance conditions of stress singularities are derived from the eigenequations, which can be applied to improve the safety of structures. Some conclusions can be made: (1) The antiplane stress singularity order remains unchanged when the fiber orientations are changed from (η_1, η_2) to $(\alpha-\eta_1, \alpha-\eta_2)$, $(\alpha-\eta_1, \eta_2)$, or $(\eta_1, \alpha-\eta_2)$; (2) As the material properties of regions 1 and 2 are identical regardless of fiber orientations, the stress singularity orders with free-free or clamped-clamped edges are equal. This is also true for clamped-free and free-clamped edges; and (3) The singularity orders with free-free boundary surfaces are the same with clamped-clamped boundary surfaces after interchanging two isotropic materials. This is also true for clamped-free and free-clamped boundary conditions.

Acknowledgment

The authors are grateful to the National Science Council of the ROC for financial support under Contract No. NSC89-2212-E006-119.

References

[1] Ma, C. C., and Hour, B. L., 1989, "Analysis of Dissimilar Anisotropic Wedges Subjected to Antiplane Shear Deformation," Int. J. Solids Struct., **25**, pp. 1295–1309.

[2] Ma, C. C., and Hour, B. L., 1990, "Antiplane Problems in Composite Anisotropic Materials With an Inclined Crack Terminating at a Bimaterial Interface," Int. J. Solids Struct., **26**, pp. 1387–1400.

[3] Shahani, A. R., 1999, "Analysis of an Anisotropic Finite Wedge Under Antiplane Deformation," J. Elast., **56**, pp. 17–32.

[4] Pageau, S. S., Joseph, P. F., and Biggers, S. B. Jr., 1995, "Finite Element Analysis of Singular Stress Fields in Anisotropic Materials Loaded in Antiplane Shear," Int. J. Numer. Methods Eng., **38**, pp. 81–97.

[5] Xie, M., and Chaudhuri, R. A., 2001, "Three-Dimensional Asymptotic Stress Field at the Front of a Bimaterial Wedge of Symmetric Geometry Under Antiplane Shear Loading," Compos. Struct., **54**, pp. 509–514.

[6] Chiu, Jack S. H., and Chaudhuri, R. A., 2002, "Three-Dimensional Asymptotic Stress Field at the Front of an Unsymmetric Bimaterial Pie-Shaped Wedge Under Antiplane Shear Loading," Compos. Struct., **58**, pp. 129–137.

[7] Hu, N., Wang, B., Sekine, H., Yao, Z., and Tan, G., 1998, "Shape-Optimum Design of a Bi-Material Single-Lap Joint," Compos. Struct., **41**, pp. 315–330.

[8] Wu, Z., 2004, "Design Free of Stress Singularities for Bi-Material Components," Compos. Struct., **65**, pp. 339–345.

[9] Chue, C. H., and Liu, C. I., 2002, "Disappearance of Free-Edge Stress Singularity in Composite Laminates," Compos. Struct., **56**, pp. 111–129.

[10] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Body*, Holden-Day, San Francisco.

On the Surface Stability of a Spherical Void Embedded in a Stressed Matrix

Jérôme Colin
Laboratoire de Métallurgie Physique,
UMR 6630 du CNRS,
Université de Poitiers,
BP 30179,
86962 Futuroscope Cedex, France
e-mail: jerome.colin@univ-poitiers.fr

The linear stability analysis of the shape of a spherical cavity embedded in an infinite-size matrix under stress has been performed when infinitesimal perturbation from sphericity of the rod is assumed to appear by surface diffusion. Developing the perturbation on a basis of complete spherical harmonics, the growth rate of each harmonic $Y_l^m(\theta,\varphi)$ *has been determined and the conditions for the development of the different fluctuations have been discussed as a function of the applied stress and the order l of the perturbation.* [DOI: 10.1115/1.2165244]

1 Introduction

The shrintage of voids and the sintering of particles are key phenomena in understanding the mechanical behavior of a number of materials. For example, the study of sapphire fibers tested in tension at low loading rate and elevated temperature (between 800 and 1500°C) shows that micron-sized pores embedded in the matrix modify the strength and lifetime, since cracks have been found to be generated from the surface of the pores [1]. Internal pore channels have been also produced in undoped sapphire matrix. Introducing sinusoidal perturbations of controlled amplitude on their lateral surface, the different regimes of the pore channel evolution by surface diffusion have then been investigated as a function of the imposed initial wavelength of the perturbation [2]. The modeling of void shrintage and sintering of spherical particles has been carried out from numerical and theoretical points of view assuming atoms are diffusing onto the surfaces [3,4]. The formation of cracks from spherical pores has been simulated as well as the formation of spherical cavities from cylindrical pore channels. The problem of the diffusion-controlled growth of a spherical particle in contact with its supersaturated melt has been studied by Mullins and Sekerka [5] and a critical radius of the precipitate above which the surface of the particle is unstable with respect to shape perturbation has been determined.

The stability of the free surface of a non-hydrostatically stressed solid has been first investigated by Asaro and Tiller [6], Grinfeld [7] and co-workers [8–11]. A critical wavelength has been determined above which the surface is unstable. Later, the effect of stress on the morphological evolution of different cylindrical structures such as pore channels, whiskers, or tubules has been investigated when the diffusion mechanism is surface diffusion [12–15]. It has been observed that the principal effect of stress is to accelerate the growth rate of the shape fluctuations and to reduce their periodicity. Under sufficiently high conditions of the stress, the development by diffusion of an ellipsoidal shape may be favored which can lead to the formation of crack-like structures [12]. A perturbation method has also been used to determine the stress near a smooth polygonal hole [16] and the morphology of holes under stress in anisotropic matrix has been then studied [17] as well as the morphology of growing spherical precipitates epitaxially stressed in a matrix when bulk diffusion is activated [18,19]. In the case of a spherical germ growing in its supersatured melt, the "memory" effect of stress generated by the composition dependence of the lattice parameter of the solid has been characterized [20].

In this paper, the stability of the shape of a spherical cavity embedded in an infinite-size matrix has been investigated with respect to infinitesimal fluctuation of the sphere expanded on a basis of spherical harmonics in the case where the matrix is submitted to a constant stress tensor at infinity (see Fig. 1 for axes). Following the method of Leo et al. [18,19], the elastic strain and stress tensors of relaxation have been determined to the first order in the fluctuation amplitude near the cavity and the diffusion equation ruling the time evolution of each harmonic has been solved. The possibility of development of an ellipsoidal shape and of more complicated morphologies of the pore has then been investigated.

2 Morphological Instability of a Cavity Embedded in an Infinite-Size Matrix Under Constant Stress

A spherical cavity of radius $\rho=r_0$ is considered in an infinite-size matrix of shear modulus μ and Poisson's ratio ν (see Fig. 1). In the following, the calculation has been carried out using spherical coordinates. To model a number of experiments where the materials are submitted to stress, it is assumed that at sufficiently large distance from the cavity in the matrix, the medium is under stress T^0

$$\lim_{|\vec{r}|\to\infty} T^0(\vec{r}) = \begin{pmatrix} T_0 & 0 & 0 \\ 0 & T_0 & 0 \\ 0 & 0 & T_0 \end{pmatrix} \tag{1}$$

where T_0 is a constant. This initial stress satisfying the Cauchy equation in the bulk

$$\nabla \cdot T^0 = 0 \tag{2}$$

must also fulfill the traction-free condition onto the surface of the cavity $\rho=r_0$

$$T^0 \tilde{n} = 0 \tag{3}$$

with $\tilde{n}$ the normal to the sphere pointing into the vacuum. The resulting strain tensor E^0 defined by

$$E_{ij}^0 = \frac{1}{2}\left(\frac{\partial u_i^0}{\partial x_j} + \frac{\partial u_j^0}{\partial x_i}\right) \tag{4}$$

with u^0 the displacement field, also satisfies Hooke's law

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 26, 2005; final manuscript received October 19, 2005. Review conducted by Z. Suo. Discussion on the paper should be addressed to the Editor, Prof Robert M. McMeeking, Journal of Applied Mechanics, 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.

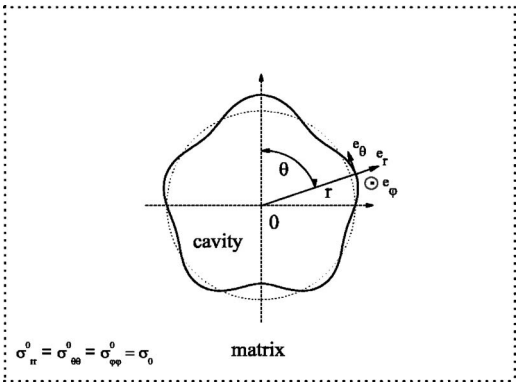


Fig. 1 The profile of an initially spherical cavity of radius r_0 embedded in an infinite-size matrix under stress is perturbed with the help of a spherical harmonic $Y_{l=5}^{m}(\theta,\varphi)$

$$T_{ij}^0 = C_{ijkl}E_{kl}^0 \tag{5}$$

where C_{ijkl} are the elastic constants. Introducing Eq. (5) in Eq. (2) leads to Navier's equation

$$\nabla^2\vec{u}^0 + \frac{1}{1-2\nu}\nabla(\nabla\cdot\vec{u}^0) = 0 \tag{6}$$

The general expression of the displacement is $\vec{u}^0(\vec{r}) = u_r^0(r)\vec{e}_r$, with $\vec{e}_r$ the radial unit vector and u_r^0 the radial displacement given by [21]

$$u_r^0(r) = \alpha_0 r + \frac{\beta_0}{r^2} \tag{7}$$

where α_0 and β_0 are constants that have determined with the help of Eqs. (1) and (3). The initial stress tensor satisfying to elasticity Eqs. (2) and (3) finally reads

$$T_{rr}^0(r) = T_0\left(1 - \frac{r_0^3}{r^3}\right) \qquad T_{\theta\theta}^0(r) = T_{\varphi\varphi}^0(r) = T_0\left(1 + \frac{r_0^3}{2r^3}\right). \tag{8}$$

It can be observed that this initial stress which is constant at infinity depends on r near the sphere [21]. The shape of the cavity is then assumed to change under the action of diffusing atoms. The set of equations describing the time evolution of the rod under stress has been first recalled in the case where atoms are diffusing onto the surface. The surface flux of atoms is defined with the help of the chemical potential [22–24]

$$\vec{J}_S = -\frac{D_S\zeta}{kT}\nabla_S\mu_c \tag{9}$$

where D_S is the surface diffusivity of atoms, k Boltzmann's constant, T the temperature, ζ the surface diffusivity of lattice sites, and μ_c the chemical potential. In the case of a single component matrix, the surface gradient of chemical potential is written as [10,24]

$$\nabla_S\mu_c = \Omega\nabla_S(\gamma\kappa + G^{\text{elas}}) \tag{10}$$

with Ω the atomic volume, γ the surface energy in the reference state, κ the curvature of the surface in the reference state defined for a normal pointing into the cavity, and G^{elas} the elastic energy density onto the surface. The motion of the surface is controlled by the accumulation of diffusing atoms [10]:

$$\frac{\partial\vec{r}}{\partial t} = -\Omega(\nabla_S\cdot\vec{J}_S)\vec{n} \tag{11}$$

The time evolution of the radial component of $\vec{r}$ is finally written as [10,23,24]

$$\frac{\partial\rho}{\partial t} = -\frac{D_S\Omega^2\zeta\gamma}{kT}(1+|\nabla\rho|^2)^{1/2}\nabla_S^2\left(\kappa + \frac{G^{\text{elas}}}{\gamma}\right) \tag{12}$$

The above system of equations has been solved in the case where a perturbation

$$\rho(\theta,t) = r_0 + \delta(t)Y_l^m(\theta,\varphi) \tag{13}$$

of the radius ρ of the cavity is assumed to appear and develops by diffusion, where δ is the amplitude of the fluctuation Y_l^m, l assuming integral values ($l=2,\ldots$). The case $l=1$ is not considered since it does not break the symmetry of the cavity. When the morphology of the cavity is modified, the solid undergoes elastic strain E^{rel} and stress T^{rel} of relaxation satisfying Hooke's law, $T_{ij}^{\text{rel}} = C_{ijkl}E_{kl}^{\text{rel}}$. This relaxation stress can be determined re-writing the mechanical equilibrium Eq. (3) onto the perturbed surface

$$(T^0 + T^{\text{rel}})\vec{n} = 0 \tag{14}$$

with

$$\vec{n} = \left(-1, \frac{\delta}{r_0}\frac{\partial Y_l^m}{\partial\theta}, \frac{\delta}{r_0\sin\theta}\frac{\partial Y_l^m}{\partial\varphi}\right)$$

the outward normal to the sphere. The total strain and stress in the matrix are then, respectively, defined as: $E_{ij}^{\text{tot}} = E_{ij}^0 + E_{ij}^{\text{rel}}$ and $T_{ij}^{\text{tot}} = T_{ij}^0 + T_{ij}^{\text{rel}}$, the resulting elasticity term in Eq. (10) being written as $G^{\text{elas}} = 1/2T_{ij}^{\text{tot}}E_{ij}^{\text{tot}}$. An elastic displacement can be derived from the strain of relaxation,

$$E_{ij}^{\text{rel}} = \frac{1}{2}\left(\frac{\partial u_i^{\text{rel}}}{\partial x_j} + \frac{\partial u_j^{\text{rel}}}{\partial x_i}\right) \tag{15}$$

The equilibrium Eq. (14) satisfied by the relaxation stress has been expanded to the first order in δ amplitude

$$T_{rr}^{\text{rel}}(r_0,\theta,\varphi) = -\left.\frac{\partial T_{rr}^0}{\partial r}\right|_{r=r_0}\delta Y_l^m(\theta,\varphi) + \Theta(\delta^2) \tag{16}$$

$$T_{r\theta}^{\text{rel}}(r_0,\theta,\varphi) = T_{\theta\theta}^0(r_0,\theta,\varphi)\frac{\delta}{r_0}\frac{\partial Y_l^m}{\partial\theta}(\theta,\varphi) + \Theta(\delta^2) \tag{17}$$

The calculation of the elastic relaxation has been performed to the first order in perturbation amplitude δ using the method developed by Leo et al. [18,19]. The elastic displacement of relaxation can be taken of the form

$$u_r^{\text{rel}}(r,\theta,\varphi) = f(r)Y_l^m(\theta,\varphi)$$

$$u_\theta^{\text{rel}}(r,\theta,\varphi) = g(r)\frac{\partial Y_l^m}{\partial\theta}(\theta,\varphi)$$

$$u_\varphi^{\text{rel}}(r,\theta,\varphi) = g(r)\frac{1}{\sin\theta}\frac{\partial Y_l^m}{\partial\varphi}(\theta,\varphi) \tag{18}$$

with

$$f(r) = \alpha_1 r^{-l-2} + \beta_1 r^{-l}, \quad g(r) = -\frac{1}{l+1}\alpha_1 r^{-l-2} + \varsigma\beta_1 r^{-l} \tag{19}$$

and

$$\varsigma = \frac{4-l-4\nu}{l(l+3-4\nu)} \tag{20}$$

The general expression of strain tensor can then be easily derived using Eq. (15). The stress is determined with the help of Hooke's law. Using Eqs. (16) and (17), the two constants α_1 and β_1 have been found to be

$$\alpha_1 = \frac{3\delta}{8\mu}r_0^{l+2}T_0\frac{(l+1)(l^2-l+2(1-\nu))}{l^2+l(1-2\nu)+1-\nu} \tag{21}$$

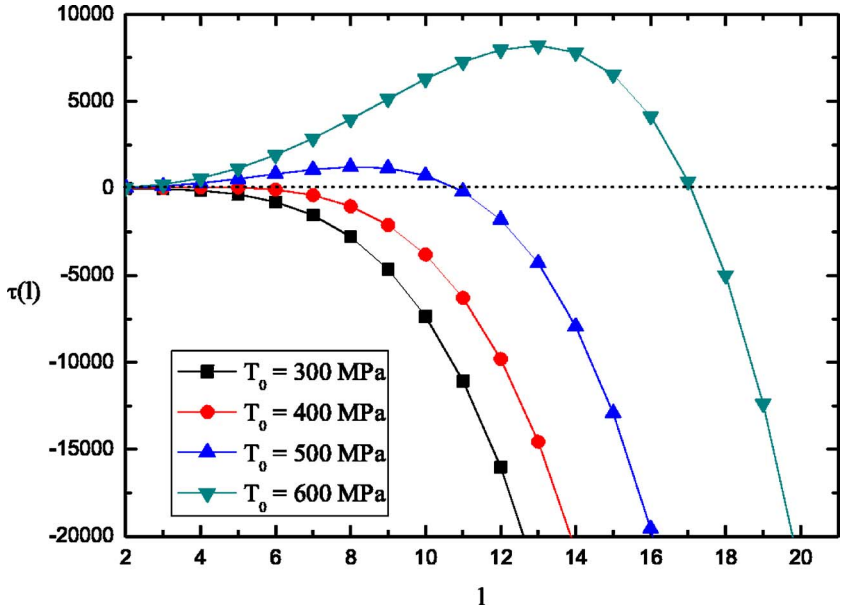


Fig. 2 Growth rate τ versus l for different values of the applied stress T_0

$$\beta_1 = -\frac{3\delta}{8\mu} r_0^l T_0 \frac{l(l-1)(l+3-4\nu)}{l^2+l(1-2\nu)+1-\nu} \tag{22}$$

The elasticity term G^{elas} has been expanded to the first order in perturbation amplitude δ

$$G^{\text{elas}} = \frac{9(1-\nu)}{8(1+\nu)}\frac{T_0^2}{\mu} + \frac{9(1-\nu)l(2l^2-3l+1)}{8[l^2+l(1-2\nu)+1-\nu]}\frac{T_0^2}{\mu r_0}\delta Y_l^m(\theta,\varphi) + \Theta(\delta^2) \tag{23}$$

and the evolution Eq. (12) of the surface has been found to be

$$\frac{1}{\delta}\frac{\partial\delta}{\partial t} = \frac{D_S\Omega^2\zeta\gamma}{r_0^4 kT}\left[-(l-1)l(l+1)(l+2) + \frac{T_0^2 r_0}{\mu\gamma}\frac{9(1-\nu)l^2(l+1)(2l^2-3l+1)}{8[l^2+l(1-2\nu)+1-\nu]}\right] \tag{24}$$

The time evolution of the lth order fluctuation is finally given by

$$\delta(t) = \delta_0 \exp\left[\frac{\tau}{\tau_0}t\right] \tag{25}$$

with $\tau_0 = kTr_0^4/D_S\Omega^2\zeta\gamma$, $\delta(0)=\delta_0$ and τ the dimensionless growth rate of the fluctuation defined by

$$\tau(l) = -(l-1)l(l+1)(l+2) + K\frac{9(1-\nu)l^2(l+1)(2l^2-3l+1)}{8[l^2+l(1-2\nu)+1-\nu]} \tag{26}$$

with the dimensionless control parameter $K = T_0^2 r_0/\mu\gamma$. The first capillarity term in the growth rate expression has already been calculated by Mullins [23]. This term is negative and favors the decay of the perturbation. The second term coming from elasticity is positive and favors the development of the fluctuation. In these conditions, the development of the lth order fluctuation is favorable when $\tau(l>1)\geqslant 0$. The growth rate τ of the fluctuation appearing on the surface of a cavity of initial radius $r_0=1$ μm has been plotted in Fig. 2 as a function of l for different values of T_0 and for the following physical parameters corresponding to an aluminum matrix $\gamma \approx 1$ J m^{-2}, $\mu=27$ GPa, and $\nu=0.33$. It can be observed that as the applied stress T_0 increases, the number of fluctuations destabilizing the shape of the cavity increases. It is assumed that a given void undergoes morphological change when at least the second order harmonic Y_2^m develops that is for $\tau(l=2)\geqslant 0$. In that case, the cavity develops an ellipsoidal shape which has already been observed in many cases such as cylindrical pore channels in a bi-axially stressed matrix [12] or spherical precipitates embedded in a matrix [25]. The condition $\tau(l=2)=0$ has been used to define a critical stress $T_c^{l=2}$ for the cavity:

$$T_c^{l=2} = \frac{4}{3\sqrt{3}}\sqrt{\frac{7-5\nu}{1-\nu}}\sqrt{\frac{\mu\gamma}{r_0}} \tag{27}$$

above which morphological change is expected to appear. This stress has been found to be $T_c^{l=2}=357.5$ MPa for a cavity of radius $r_0=1$ μm embedded in an aluminum matrix. This threshold seems to be coherent with those already determined for cylindrical [12] and spherical cavities embedded in matrix [26] even if the applied stresses differ from the initial stress considered in this work. As the stress is increased the ellipsoidal shape can be destabilizing as soon as spherical harmonics of higher order may appear. In that case, a morphological change consisting in the development of a number of harmonics takes place and the pore may grow crack tips [13]. For the lth order fluctuation, the corresponding applied stress T_c^l is defined by

$$\frac{T_c^l}{T_c^{l=2}} = \sqrt{\frac{3}{2}\frac{1-\nu}{7-5\nu}}\sqrt{\frac{(l+2)[l^2+l(1-2\nu)+1-\nu]}{(1-\nu)l(2l-1)}} \tag{28}$$

Relation (28) holds for cavities of any radius, for example in the case of aluminum ($\nu=0.33$) the critical stress above which the fluctuation Y_4^m may develop has been found to be $T_c^{l=4}=1.04\ T_c^{l=2}$. The critical stress above which Y_6^m may appear is $T_c^{l=6}=1.147\ T_c^{l=2}$. It can be observed that when the initial stress T_0 slightly increases from the value $T_c^{l=2}$, the development of a number of harmonics becomes favorable. A critical mode $l_c=\lfloor l_* \rfloor$ such that $\tau(l_*)=0$ can be then defined, where $\lfloor x \rfloor$ is the integer part of a number x. For $l\leqslant l_c$, the growth rate is positive and the corresponding fluctuation may develop. This critical mode l_c must satisfy the following relation

$$8l_*^3 + [8(3-2\nu) - 18K(1-\nu)]l_*^2 + [8(3-5\nu) + 9K(1-\nu)]l_* + 16(1-\nu) = 0 \tag{29}$$

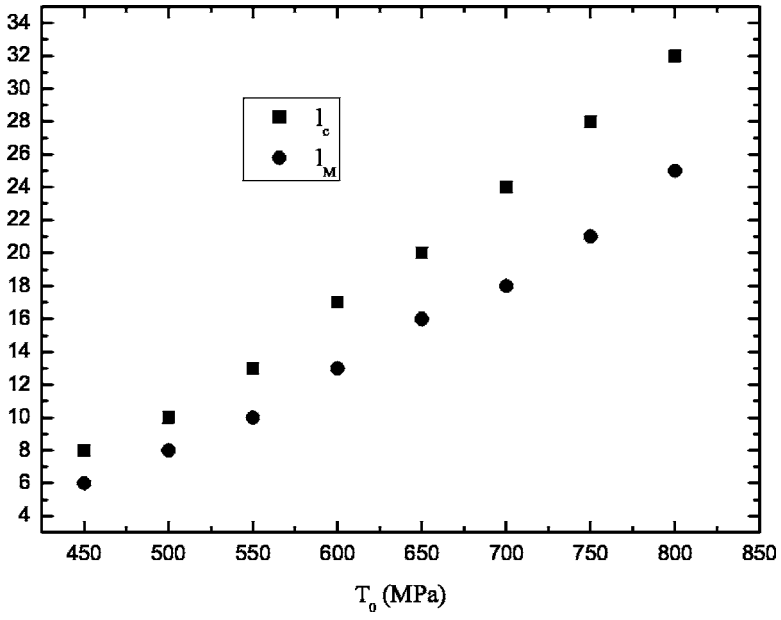


Fig. 3 Critical and most probable modes l_c and l_M versus applied stress T_0

Among all the possible modes, one is assumed to develop faster than the other. This mode l_M selected by surface diffusion is determined differentiating the growth rate with respect to l. The optimal mode l_M has been found to be $l_M \in [[l_*],[l_*]+1]$ with l_* satisfying the following relation

$$\begin{aligned} &-32l_*^7+2[27K(1-\nu)-8(7-8\nu)]l_*^6+2[9K(8\nu^2-11\nu+3)\\ &\quad-8(8\nu^2-24\nu+11)]l_*^5+[9K(4\nu^2-9\nu+5)\\ &\quad-32(10\nu^2-13\nu+5)]l_*^4-4[40\nu^2-48\nu+12\\ &\quad+9K(3\nu^2-5\nu+2)]l_*^3+[16(5\nu^2-6\nu+2)\\ &\quad-9K(4\nu^2-9\nu+5)]l_*^2+2(1-\nu)[9K(1-\nu)\\ &\quad+8(3-5\nu)]l_*+16(1-\nu)^2=0 \end{aligned} \tag{30}$$

Since no analytic expression of the critical and optimal modes can be derived from Eqs. (29) and (30), their variations versus applied stress T_0 have been numerically estimated in Fig. 3 in the case of a pore of initial radius $r_0=1$ μm. It is observed that for a given radius, the two integers l_c and l_M increase as the stress increases. It is then assumed that as the applied stress increases, the development of harmonics of higher order is favored leading to the formation of a rough cavity. It can finally be emphasized that the linear stability analysis presented in this paper only holds for the early beginning of the surface evolution and does not provide any information on the long time evolution of the amplitude of the roughness. It would be interesting to solve numerically the time evolution Eq. (12) of the sphere and to determine in which conditions of stress the development of high order perturbations Y_l^m (with $l>2$) can lead to the formation of cracks or to more complex evolutions such as the distribution of smaller cavities.

3 Conclusion

The present analysis provides the theoretical basis for understanding the effect of stress on the morphological evolution of a spherical pore embedded in a matrix. The major result of this work is that the spherical shape of the pore found to be stable for a stress-free matrix becomes unstable with respect to shape deviation from sphericity of the cavity as soon as a stress greater than a critical value is applied to the solid. The critical stress has been then determined as a function of the order of the fluctuation. The first fluctuation to appear by diffusion leads to the formation of an ellipsoidal shape. As the stress is increased, the development of more complex morphologies has been found to be favored on the surface of the cavity.

The results obtained in this paper in the linear regime of the surface evolution apply to any perturbation. The fluctuation can be expanded on a basis of spherical harmonics Y_l^m and the set of linear equations detailed in this work can be used to describe the time dependence of each coefficient of the shape development. It is also believed that stress driven surface instability can play a role during the sintering of powder compacts, the merging between two sufficiently closed spherical particles thus being achieved by developing a roughness on each of the two surfaces.

References

[1] Newcomb, S. A., and Tressler, R. E., 1993, "Slow Crack Growth in Sapphire Fibers at 800° to 1500°C," J. Am. Ceram. Soc., **76**, pp. 2505–2512.
[2] Kulinsky, L., Powers, J. D., and Glaeser, A. M., 1996, "Morphological Evolution of Pre-Perturbed Pore Channels in Sapphire," Acta Mater., **44**, pp. 4115–4130.
[3] Wakai, F., and Aldinger, F., 2003, "Sintering Through Surface Motion by the Difference in Mean Curvature," Acta Mater., **51**, pp. 4013–4024.
[4] Takahashi, Y., Takahashi, K., and Nishiguchi, K., 1991, "A Numerical Analysis of Void Shrinkage Processes Controlled by Coupled Surface and Interface Diffusion," Acta Metall. Mater., **39**, pp. 3199–3216.
[5] Mullins, W. W., and Sekerka, R. F., 1963, "Morphological Stability of a Particle Growing by Diffusion or Heat Flow," J. Appl. Phys., **34**, pp. 323–329.
[6] Asaro, R. J., and Tiller, W. A., 1972, "Interface Morphology Development During Stress Corrosion Cracking: Part I. Via Surface Diffusion," Metall. Trans., **3**, pp. 1789–1796.
[7] Grinfeld, M. A., 1986, "Instability of the Separation Boundary Between a Non-Hydrostatically Stressed Elastic Body and a Melt," Sov. Phys. Dokl., **31**, pp. 831–834.
[8] Srolovitz, D. J., 1989, "On the Stability of Surfaces of Stressed Solids," Acta Metall., **37**, pp. 621–625.
[9] Gao, H., 1994, "Some General Properties of Stress-Driven Surface Evolution in a Heteroepitaxial Thin Film Structure," J. Mech. Phys. Solids, **42**, pp. 741–772.
[10] Spencer, B. J., Voorhees, P. W., and Davis, S. H., 1991, "Morphological Instability in Epitaxially Strained Dislocation-Free Solid Films," Phys. Rev. Lett., **67**, pp. 3696–3699.
[11] Jonsdottir, F., 1995, "Computation of Equilibrium Surface Fluctuations in Strained Epitaxial Films Due to Interface Misfit Dislocations," J. Modelling Simul. Mater. Sci. Eng., **3**, pp. 503–520.
[12] Suo, Z., and Wang, W., 1994, "Diffusive Void Bifurcation in Stressed Solid," J. Appl. Phys., **76**, pp. 3410–3421.
[13] Wang, W., and Suo, Z., 1997, "Shape Change of a Pore in a Stressed Solid via Surface Diffusion Motivated by Surface and Elastic Energy Variation," J. Mech. Phys. Solids, **45**, pp. 709–729.
[14] Colin, J., Grilhé, J., and Junqua, N., 1997, "Morphological Instabilities of a

Stressed Cylindrical Pore Channel," Acta Mater., **45**, pp. 3835–3841.
[15] Kirill, D. J., Davis, S. H., Mikisis, M. J., and Voorhees, P. W., 2002, "Morphological Instability of Pores and Tubules," Interfaces Free Boundaries, **4**, pp. 371–394.
[16] Gao, H., 1991, "Stress Analysis of Smooth Polygonal Holes via a Boundary Perturbation Method," J. Appl. Mech., **58**, pp. 851–853.
[17] Gao, H., 1992, "Stress Analysis of Holes in Anisotropic Elastic Solids: Conformal Mapping and Boundary Perturbation," Q. J. Mech. Appl. Math., **45**, pp. 149–168.
[18] Leo, P. H., and Sekerka, R. F., 1989, "The Effect of Elastic Fields on the Morphological Stability of a Precipitate Grown from Solid Solution," Acta Metall., **37**, pp. 3139–3149.
[19] Leo, P. H., Iwan, J., Alaxender, D., and Sekerka, R. F., 1985, "Elastic Fields about a Perturbed Spherical Inclusion," Acta Metall. **33**, pp. 985–989.
[20] Caroli, B., Caroli, C., Roulet, B., and Voorhees, P. W., 1989, "Effect of Elastic Stresses on the Morphological Stability of a Solid Sphere Growing from a Supersatured Melt," Acta Metall., **37**, pp. 257–268.
[21] Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, 2nd ed., McGraw-Hill, New York, pp. 343–345.
[22] Larche, F. C., and Cahn, J. W., 1985, "The Interaction of Composition Stress in Crystalline Solids," Acta Metall., **33**, pp. 331–357.
[23] Mullins, W. W., 1957, "Theory of Thermal Grooving," J. Appl. Phys. **28**, pp. 333–339.
[24] Spencer, B. J., and Tersoff, J., 1997, "Equilibrium Shapes and Properties of Epitaxially Strained Islands," Phys. Rev. Lett. **79**, pp. 4858–4860.
[25] Johnson, W. C., and Cahn, W., 1984, "Elastically Induced Shape Bifurcation of Inclusions," Acta Metall., **32**, pp. 1925–1933.
[26] Sun, B., Suo, Z., and Evans, A. G., 1994, "Emergence of Crack by Mass Transport in Elastic Crystals Stressed at High Temperature," J. Mech. Phys. Solids, **42**, pp. 1653–1677.

Flow Analysis and Modeling of Field-Controllable, Electro- and Magneto-Rheological Fluid Dampers

Xiaojie Wang

Faramarz Gordaninejad[1]
Professor
e-mail: faramarz@unr.edu

Department of Mechanical Engineering,
University of Nevada,
Reno, NV 98557

This study combines a fluid mechanics-based approach and the Herschel-Bulkley constitutive equation to develop a theoretical model for predicting the behavior of field-controllable, magneto-rheological (MR), and electro-rheological (ER) fluid dampers. The goal is to provide an accurate theoretical model for analysis, design, and development of control algorithms of MR/ER dampers. Simplified explicit expressions for closed-form solution of the pressure drop across a MR fluid valve are developed. The Herschel-Bulkley quasi-steady flow analysis is extended to include the effect of fluid compressibility to account for the nonlinear dynamic behavior of MR/ER fluid dampers. The advantage of this model is that it only depends on geometric and material properties of the MR/ER material and the device. The theoretical results are validated by an experimental study. It is demonstrated that the proposed model can effectively predict the nonlinear behavior of field-controllable fluid dampers. [DOI: 10.1115/1.2166649]

1 Introduction

Semi-active vibration control devices have received significant attention because they can offer combined advantages of both passive and active control systems [1,2]. One class of semi-active control devices are field-controllable magneto-rheological (MR), and electro-rheological (ER), fluid dampers. Activation of the fluid in the damper causes a fast and dramatic change in the apparent viscosity of the MR/ER fluid contained in the device. These fluids can reversibly change from liquid to semisolid in milliseconds [3,4]. The result is a continuously variable controllable damper.

MR fluid dampers can be battery-operated, require minimal power, and have a broad range of capabilities, such as, the absence of mechanical valve (for flow control) in the damper, and insensitivity to impurities penetration [5]. Most importantly, a MR fluid dampers can be considered as a "*fail-safe*" device, that is, they can retain a minimum required damping capacity in the event of a power supply or electronic system failures. Another words, in the event of an electrical malfunction in the control hardware, the device behaves as a passive damper [6,7].

MR/ER fluid dampers demonstrate highly nonlinear behavior due to the inherent non-Newtonian behavior of these fluids. MR/ER fluids exhibit a strong field-dependent shear modulus and a shear yield stress that resists the material's flow until shear stress reaches a critical value. Bingham plastic model is often used to describe this phenomenon [5,8–10]. Based on the Bingham plastic constitutive equation, a quasi-steady, field-controllable damper model was developed to predict the damper's performance [11–15]. In this model, the fluid is assumed to be a Newtonian fluid in post-yield regime, with a constant plastic viscosity assumption. However, for cases where the fluid experiences post-yield shear thinning or shear thickening, the assumption of constant plastic viscosity is not valid.

Goodwin et al. [16] reported that the ER fluids prepared with particles of low conductivity showed pseudo-plastic behavior with power law index that was a function of conductivity. Mokeev et al. [17] employed the continuous inhomogeneous dielectric medium approximation to model Couette shear in high-concentration ER fluid under electric field. They obtained dependence of the apparent viscosity on the electric field and shear rate, which in the limiting case is reduced into the Bingham plastic model. Shulman and Korobko [18] considered refining the Bingham plastic model into a four-parametric Casson-type model with coefficients depending on electric field.

Stanway et al. [19] proposed a model to describe controllable fluid damper behavior. An experimental technique of nonlinear sequential filtering was used to estimate the parameters associated with an nth-power velocity model for the damping mechanism. Halsey et al. [20] and Felt et al. [21] studied a monodispers ER fluid and a monodispers MR fluid, respectively. Both studies found that the Bingham model does not accurately describe the shear thinning observed in theses systems. In order to incorporate the shear-thinning behavior of MR/ER fluids, the authors suggested using the Herschel-Bulkley model [22] which states that:

$$\tau = \tau_y + k|\dot{\gamma}|^{n-1}\dot{\gamma} \quad |\tau| > \tau_y$$

$$\dot{\gamma} = 0 \quad |\tau| \leqslant \tau_y \tag{1}$$

where τ is shear stress; $\dot{\gamma}$ is shear strain rate, and k and n are fluid index parameters. τ_y is the MR/ER fluid field-dependent yield stress. Li [23] studied the rheological properties of MR fluid of MRF-132LD under steady shear with a commercial MR rheometer. He compared experimental data and model-predicted curves utilizing both Herschel-Bulkley and Bingham plastic models, and concluded that the Herschel-Bulkley model predictions are more accurate than Bingham plastic model.

The quasi-steady models, which are based on Herschel-Bulkley steady flow analysis, was also successfully applied to a MR fluid damper to predict its force-displacement relationship [24,25]. However, similar to the Bingham model, it still cannot accurately

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 26, 2004; final manuscript received November 29, 2005. Review conducted by B. A. Younis. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

predict the force-velocity hysteretic behavior at low velocity. Accurate prediction of force-velocity is important for control analysis MR/ER fluid dampers.

In order to account for the force-velocity relationship of MR/ER fluid dampers, different models have been proposed [26–40]. Most models are based on this assumption that MR/ER fluids pre-yield viscoelastic properties, which are not accounted by Bingham (or Herschel Bulkley) model, could contribute to this highly nonlinear behavior, and therefore, more parameters associated with its pre-yield mechanism are added to make the model more complicated. Examples are the augmented six-parameter model proposed by Kamath and Wereley [26], and the extended Bingham model proposed by Gamota and Filisko [29]. These models can provide a good estimate of MR/ER fluid damper force-velocity response, however, the model parameters can only be obtained a posterior for every specific test and their effect on the model is difficult to assess.

Parametric or phenomenological models were also attempted to describe the behavior of electro- and magneto-rheological fluid dampers. These models are also based on determination of system parameters by curve fitting with experimental data [30–32].

Nonparametric approaches for modeling MR/ER fluid dampers have also been addressed [33–40]. For example, a Chebyshev polynomial curve fit method based on the damper velocity and acceleration was presented by Ehrgott and Masri [33], and Gavin et al. [34]. A neural network model with six input neurons, one output neuron and twelve neurons in the hidden layer were explored by Chang and Roschke [35].

Although, parametric and nonparametric models can satisfactorily emulate dynamic behavior of the ER/MR fluid damper, they cannot provide any valuable practical tools in damper design and analysis. In addition, the accuracy of these models relies primarily on the experimental data in the laboratory environment.

Peel et al. [38] and Sims et al. [39,40] extended their quasi-steady ER fluid damper model to include dynamic effects accounting for the hysterestic behavior of an ER long-stroke fluid damper. They developed a lumped parameter model consisting of a spring, mass, and damper connected in series. The spring stiffness and the mass are related to the fluid bulk modulus and the volume of working fluid and its density, respectively. Although the model was derived based on the material and geometric properties some model parameters, such as stiffness, yield stress, and viscosity still need to be identified using experimental data to improve the accuracy.

In this study, a theoretical model for predicting the behavior of field-controllable, electro-rheological (ER), and magneto-rheological (MR) fluid dampers is developed. This model is based on the physical parameters of a device, as well as, the properties of a MR fluid, making it a valuable tool for MR/ER devices' design and analysis and development of control systems. In Sec. 2, the Herschel-Bulkley model is utilized to establish relationship for MR/ER fluid flow in a pipe and between parallel plates in flow mode. A closed-form expression is presented for determining the pressure drop as a function of material properties, geometry, and volumetric flow rate for MR/ER fluids. In addition, a simplified model is developed for each case. In Sec. 3, pre-yield and post-yield behavior of MR/ER fluid and their effect on the dynamic performance of MR/ER fluids utilizing a piston-driven channel flow of MR device is developed. Comparing with the experimental results, the quasi-steady model based on Herschel-Bulkley constitutive equation was verified. In Sec. 4, the Herschel-Bulkley quasi-steady analysis is extended by incorporating the effect of fluid compressibility, and a fluid mechanics-based model is developed to account for a MR fluid damper dynamic behavior. In Sec. 5, the fluid mechanics-based model is compared with experiment results for a prototype MR fluid damper which is designed, constructed and tested at the University of Nevada, Reno (UNR). It is demonstrated that the proposed model can effectively capture the hysteresis characteristics of the MR fluid damper both in force-displacement and velocity-displacement over a wide working range.

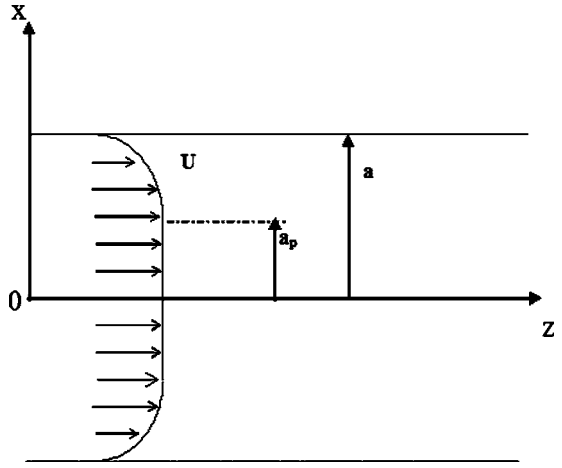


Fig. 1 Flow profile of a non-Newtonian fluid through a uniform circular or parallel plate cross section

2 Modeling of MR/ER Fluid Flow

The Herschel-Bulkley model is utilized to establish relationship for MR/ER fluid flow in flow mode in a pipe and between fixed parallel plates. In addition, a simplified model is developed for each case. The Herschel-Bulkley constitutive equation presented in Eq. (1) describes the flow of MR/ER fluids through circular or rectangular channels with a constant cross-sectional area.

Let us consider a steady, one-dimensional flow of an incompressible MR/ER fluid through a channel with rectangular cross section ($-a \leqslant x \leqslant a$), or a circular cross section ($0 \leqslant x \leqslant a$). Rectangular and cylindrical coordinate systems with notations shown in Fig. 1 are considered. The momentum equation of laminar flow for a continuous medium is:

$$\frac{dp}{dz} + \frac{1}{x^j}\frac{\partial(x^j\tau)}{\partial x} = 0 \tag{2}$$

with $\tau = \tau(x)$ denoting the distribution of tangential shear stress where $j=0$ and $j=1$ denote the rectangular and cylindrical coordinates, respectively, and the pressure gradient, dp/dz is considered constant along the flow direction. Integrating Eq. (2) yields shear stress as a function of the transverse coordinate x:

$$\tau = -\frac{1}{j+1}\frac{dp}{dz}x + \frac{c_0}{x^j} \tag{3}$$

where c_0 denoting a integral constant. The shear stress τ vanishes as $x \to 0$ and, therefore, $c_0 \equiv 0$. The flow has a nonyield region, which is signified by the "plug" radius shown in Fig. 1. The plug radius, a_p, can be determined by:

$$a_p = -\frac{(j+1)\tau}{dp/dz} \tag{3a}$$

The yield flow region is defined when $a_p \leqslant x \leqslant a$. Substitution of Eq. (1) into Eq. (3) yields:

$$\tau_y + k\left|\frac{du}{dx}\right|^n = -\frac{1}{j+1}\frac{dp}{dz}x \tag{4}$$

Since $du/dx < 0$, Eq. (4) can be written as:

$$\frac{du}{dx} = -\left(-\frac{1}{j+1}\frac{dp/dz}{k}x - \frac{\tau_y}{k}\right)^{1/n} \tag{5}$$

Solving Eq. (5) with boundary condition of $u=0$ at $x=a$, one obtains the velocity distribution in the yield flow region ($a_p \leqslant x \leqslant a$), as follows:

$$u = \frac{n}{n+1}\frac{(j+1)k}{dp/dz}\left[\left(-\frac{dp/dz}{(j+1)k}x - \frac{\tau_y}{k}\right)^{(n+1)/n} - \left(-\frac{dp/dz}{(j+1)k}a - \frac{\tau_y}{k}\right)^{(n+1)/n}\right] \tag{6}$$

In the plug area, $0 \leqslant x \leqslant a_p$, the velocity is constant. By letting $x = a_p$, the plug velocity can be expressed as:

$$u_p = \frac{n}{n+1}\frac{(j+1)k}{dp/dz}\left[\left(-\frac{dp/dz}{(j+1)k}a_p - \frac{\tau_y}{k}\right)^{(n+1)/n} - \left(-\frac{dp/dz}{2k}a - \frac{\tau_y}{k}\right)^{(n+1)/n}\right] \tag{7}$$

The volumetric flow rate can be derived from the following relation:

$$Q = 2\int_{a_p}^{a} u(\pi x)^j dx + \frac{2}{j+1}a_p^{j+1}\pi^j u_p \tag{8}$$

Substitution of Eqs. (6) and (7) into Eq. (8) yields:

$$Q_0 = \frac{n}{2n+1}\frac{1}{k^{1/n}(dp/dz)^2 a}\left(-\frac{dp}{dz}a - \tau_y\right)^{(n+1)/n}\left(\frac{n}{n+1}\tau_y - \frac{dp}{dz}a\right) \tag{9}$$

$$Q_1 = \frac{\left(-\frac{dp}{dz}\frac{a}{2} - \tau_y\right)^{(n+1)/n} a}{\left(\frac{pa}{2}\right)^3 k^{1/n}}\left[\frac{\left(-\frac{dp}{dz}\frac{a}{2} - \tau_y\right)^2}{\frac{3n+1}{n}} + \frac{2\tau_y\left(-\frac{dp}{dz}\frac{a}{2} - \tau_y\right)}{\frac{2n+1}{n}} + \frac{\tau_y^2}{\frac{n+1}{n}}\right] \tag{9a}$$

where $Q_0 = Q/2wa$ is the volume flow rate per unit area of a rectangular channel with a width of w, and $Q_1 = Q/\pi a^2$ is the volume flow rate per unit area of a cylindrical pipe with a constant radius of a. Equations (9) and (9*a*) can be approximated to provide explicit expressions for the pressure gradient in rectangular and circular channels presented in Eqs. (10) and (10*a*), respectively [22]:

$$\frac{dp}{dz} = A_j\frac{(j+1)\tau_y}{a} + \left(\frac{(jn+2n+1)}{n}\frac{Q_j}{a}\right)^n\frac{(j+1)k}{a} \quad \text{for } \frac{a_p}{a} \leqslant 0.5 \tag{10}$$

$$\frac{dp}{dz} = \frac{(j+1)\tau_y}{a\left[1 - B_j\left(\frac{(jn+2n+1)}{n}\frac{Q_j}{a}\right)^{n/\alpha_j}\left(\frac{k}{\tau_y}\right)^{1/\alpha_j}\right]} \quad \text{for } \frac{a_p}{a} > 0.5 \tag{10a}$$

where,

$$A_0 = \frac{2n+1}{n+1} - \frac{3n(1+2n)(1-n)}{16(n+1)^2} - \frac{3n(n^2+1)}{40(n+1)},$$

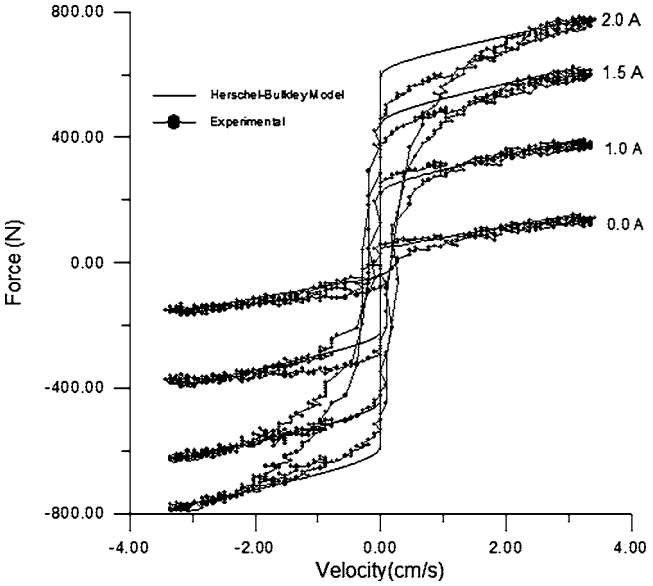


Fig. 2 Experimental and theoretical force-velocity results of UNR MRD-001 damper for a sinusoidal motion at frequency of 0.5 Hz and 1.016 cm amplitude

$$A_1 = \frac{3n+1}{2n+1} - \frac{3}{16}\frac{(3n+1)(1-n)}{(2n+1)^2(n+1)},$$

and B_j and α_j are also functions of fluids index n. The Herschel-Bulkley model developed in this study can be reduced to Bingham plastic model by setting $n=1$.

3 Pre-Yield Effect on MR Fluid Dampers

The presented approximated solutions using Herschel-Bulkley model can accurately predict the force capacity of ER/MR fluid dampers. In a previous study, using Herschel-Bulkley model and a quasi-steady assumption, the performance of a prototype MR fluid damper was validated by experimental result [24]. However, the quasi-steady assumption cannot capture the nonlinear force-velocity hysteresis, which is essential for control analysis of MR/ER fluid devices.

Figure 2 compares the experimental and theoretical results using Herschel-Bulkley model for the force-velocity response of UNR MRD-001 MR fluid damper. As can be seen, the quasi-static Herschel-Bulkley (or Bingham plastic) model is inadequate to fit the experimental data at low velocities. Most researchers were assumed that MR/ER fluids pre-yield viscoelastic properties, which are not accounted by Bingham (or Herschel-Bulkley) model, could contribute to this highly nonlinear behavior [26,28–30]. In the following Section the pre-yield and post-yield behavior of MR/ER fluid and their effect on the dynamic performance of these fluids with a piston-driven channel flow of MR device is studied. The pressure drop across the applied magnetic field is measured using two pressure transducers. The pressure drop and piston velocity under various magnetic field strengths is examined.

Figure 3 shows the piston-driven flow-mode type MR device used in this study with a flat channel of rectangular cross section. An electromagnet with 1200-turn coils and a low-carbon steel as its core material is built and located at the middle of the channel to permit the application of magnetic flux normal to the slit channel flow. The coil was activated by an operation-amplifier power supply in constant current mode. A Gauss meter measured the magnetic field strength inside the channel induced by the coil. The MR fluids was confined in the well-sealed channel cell and was pressurized to flow through the channel between two parallel-arranged magnet pole pieces. There is an external V-groove in the

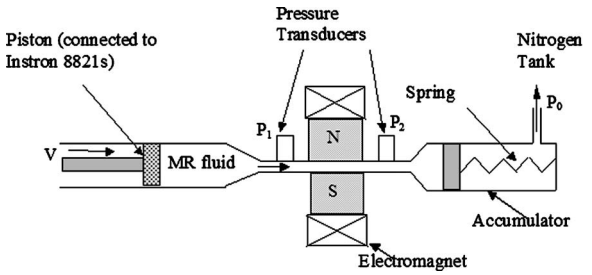


Fig. 3 Schematic of the MR channel flow experimental setup

middle of the test channel in order to place the electromagnetic coil and activate the MR fluid. The movement of the piston is controlled by a hydraulic Instron machine. Two fluid pressure transducers measure the pressure drop across the MR valve (this is the portion at which the magnetic field is applied). The MR valve is $h=1.0$ mm in height, $w=10$ mm in width, and $l=14$ mm in length. An accumulator, pressurized by a nitrogen tank, is used to seal the outlet of the flow. This is done to avoid the problem of collection of the extrude sample. Moreover, it can press the fluid repeatedly through the MR valve under the dynamic movement of the piston.

A series of sinusoidal inputs at frequency of 0.5 Hz with various input currents were applied using Instron servo-hydaulic machine Model 8821s at room temperature. Before each test, the accumulator was pressurized to 150 psi via nitrogen tank, so that the MR fluid can follow the movement of the piston in one cycle. This is done to avoid the formation of cavities inside the channel when the piston is pulled back. The input coil current for each velocity was 0, 0.25, 0.75, and 1.0 A, corresponding to 0.0, 120, 250, and 350 mT magnetic flux densities (B_0) inside the air gap measured by means of a Hall probe (F. W. Bell Gauss meter model 9500), respectively.

Typical results of the pressure drop across the MR valve versus the velocity of the piston for a sinusoidal input is shown in Fig. 4. The solid lines which are obtained from Eqs. (10) and (10*a*), presents the pressure drop across the MR valve. As can be seen, there is no pressure drop-velocity hysteresis loop exists. A Herschel-Bulkley quasi-static model can capture this behavior, accurately. This result is different from the test data of the MR fluid damper, although in both cases the flow mode is used. This may suggest that the nonlinear behavior of the MR fluid damper at low velocity is not due to the pre-yield property of the MR fluid.

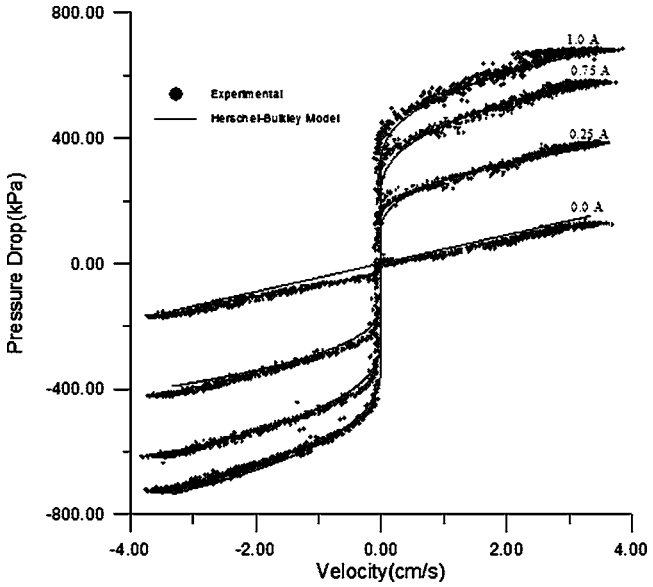


Fig. 4 Theoretical and experimental pressure drop-velocity results for the piston-driven MR channel flow for a sinusoidal motion at frequency of 0.5 Hz and 1.05 cm amplitude

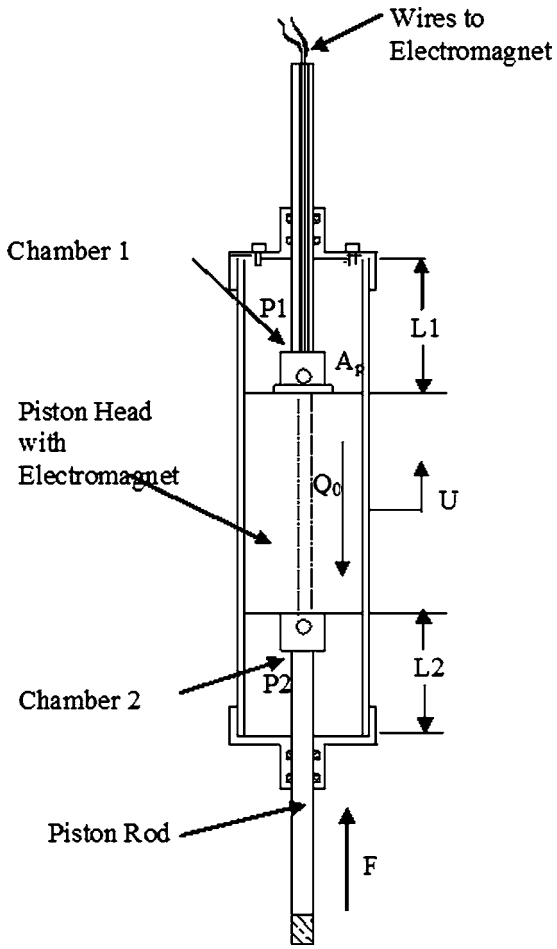


Fig. 5 Schematic of the prototype UNR MRD-001 MR fluid damper

4 Modeling of MR Dampers

In this section it is demonstrated that the compressibility of MR/ER fluids in the MR damper chambers plays an important role in the nonlinear behavior of MR fluid dampers. In this study, the quasi-steady analysis is extended into a dynamic model by taking into account the compressibility effects of the MR fluid. This model is based only on the physical parameters of the device and the properties of the MR fluid. The analysis provides further insight into the effects of various material properties and physical dimensions of the system on the performance of the damper.

The theoretical model is applied to a prototype MR fluid damper [6]. The schematic of this damper is shown in Fig. 5. The damper has a through-rod design, i.e., that the ratio of fluid volume to rod volume in the damper is constant over a stroke. The controllable MR valves are within the piston. An enlarged cross-sectional view of the MR fluid damper's piston/rod assembly is shown in Fig. 6. The fluid flow path through the piston and the geometric dimensions used in the theoretical formulation are also shown in Fig. 6. The mass flow rate continuity for the fluid volume is:

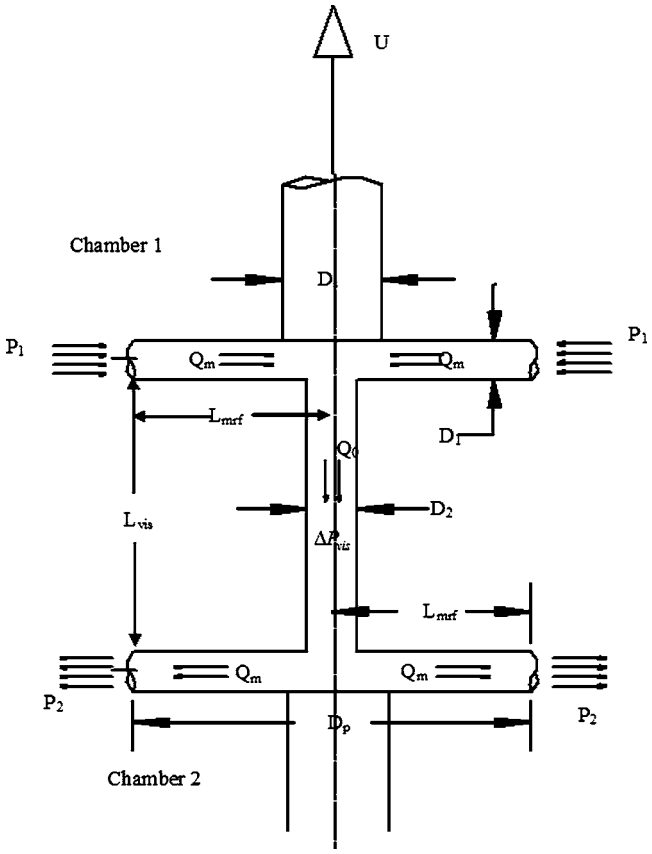


Fig. 6 Schematic of the flow path trough the UNR MRD-001 MR fluid damper

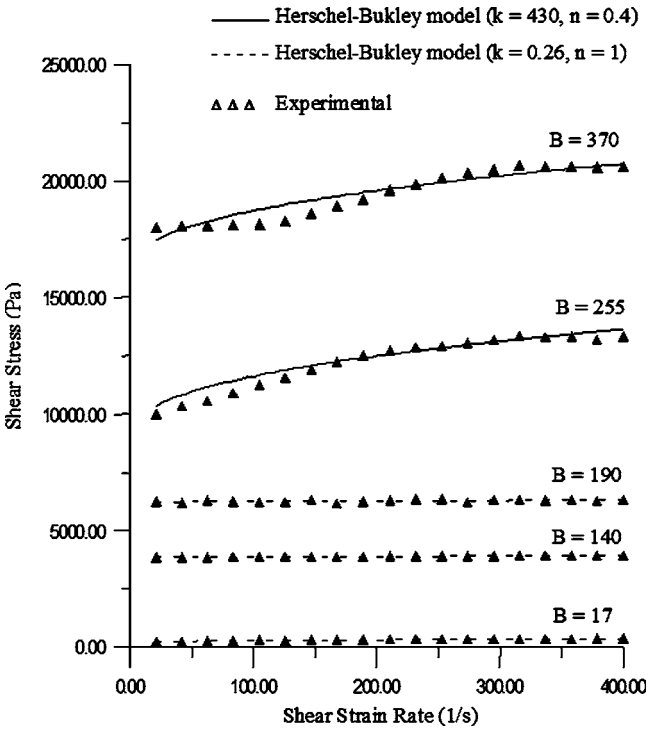


Fig. 7 Shear stress versus shear strain rate data at various magnetic field strengths

$$\frac{d}{dt}(m) = \frac{d}{dt}(\rho V) = \rho_{\text{in}} Q_{\text{in}} - \rho_{\text{out}} Q_{\text{out}} \tag{11}$$

where ρ is the density, m is the mass, V is the volume, Q_{in} is the input flow rate, and Q_{out} is output flow rate. Considering the compressibility of the fluid, one has:

$$\frac{d}{dt}(\rho V) = \rho\left(\frac{dV}{dt} + \frac{V}{\beta}\frac{dP}{dt}\right) \tag{12}$$

where P is the pressure, and β is the bulk modulus of the fluid. Combining Eqs. (11) and (12), and assuming a constant fluid density, one obtains:

$$\frac{dV}{dt} + \frac{V}{\beta}\frac{dP}{dt} = Q_{\text{in}} - Q_{\text{out}} \tag{13}$$

Equation (13) presents a mass flow rate continuity equation accounting for the fluid compressibility. For this specific MR fluid damper Q_{in} is zero, and the flow continuity Eq. (13) for chambers one and two are [41,42]:

$$\frac{dV_1}{dt} + \frac{V_1}{\beta_1}\frac{dP_1}{dt} = -Q_{\text{out}}$$

$$\frac{dV_2}{dt} + \frac{V_2}{\beta_2}\frac{dP_2}{dt} = Q_{\text{out}} \tag{14}$$

where

$$V_1 = (L_1 - U)A_p \text{ and}$$

$$V_2 = (L_2 + U)A_p \tag{15}$$

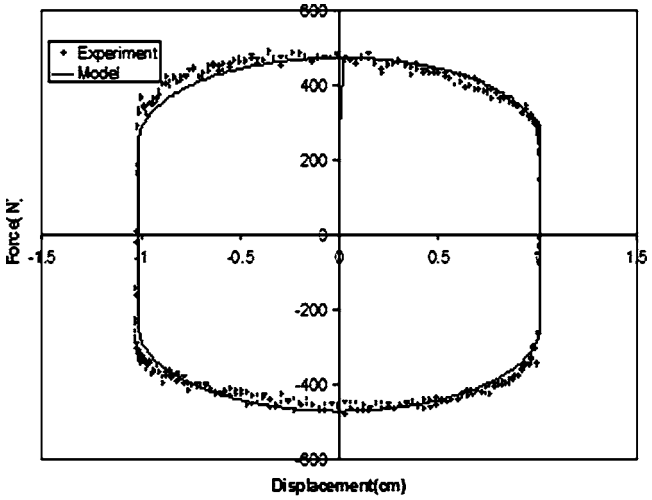
Here U is the piston displacement, A_p is the effective area of the piston, and β_1 and β_2 are the effective bulk moduli of the fluid in chambers one and two, respectively. The bulk modulus is sensitive to the variations in temperature, pressure, volumetric ratio of fluid, and air content. At room temperature, knowing the volume of air present per unit volume of oil and assuming a perfectly rigid container, the effective bulk modulus is estimated by Stringer [43], as follows:

$$\frac{1}{\beta} = \frac{1}{\beta_0} + \frac{V_{\text{air}}}{V_o}\frac{1}{P} \tag{16}$$

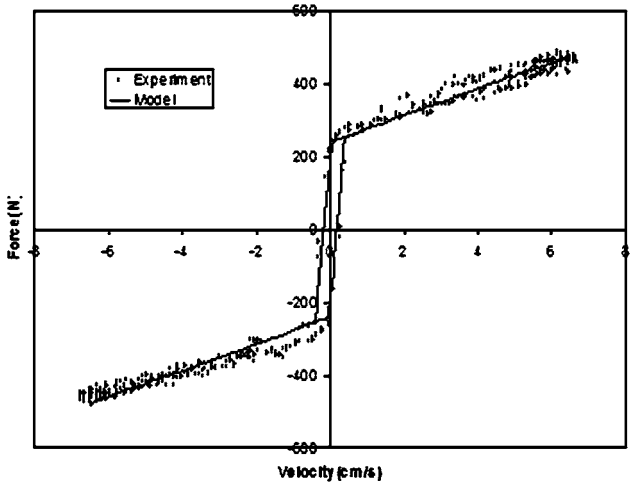
where β_0 is the bulk modulus of a pure oil and is about 17×10^8 N/m^2, V_{air}/V_o is ratio of a volume of air dispersed in a volume of oil (from 0.01% to 1%), and P is mean value pressure. In the MR fluid damper considered here, the chambers mean value pressure is 5.0×10^5 N/m^2. If 1% of air is entrapped in the fluid, the effective bulk modulus is 4.86×10^7 N/m^2, while the effective bulk modulus is 1.64×10^9 N/m^2 at 0.01% of air content. Therefore, it is evident that air content in the fluid can drastically reduces the effective bulk modulus. In practical applications, a small percentage of air is always present in the system, and the effectiveness of the design can significantly be affected, if the design

Table 1 Material and geometric properties of MR fluid and the MR fluid damper

k	430 Pa-s^n
n	0.4
N	4
μ	0.4 Pa-s
D_p	0.0508 m
D_s	0.0127 m
D_1	0.00635 m
D_2	0.00635 m
L_1	0.057 m
L_2	0.057 m
L_{mrf}	0.00292 m
A_p	1.9×10^{-3} m^2

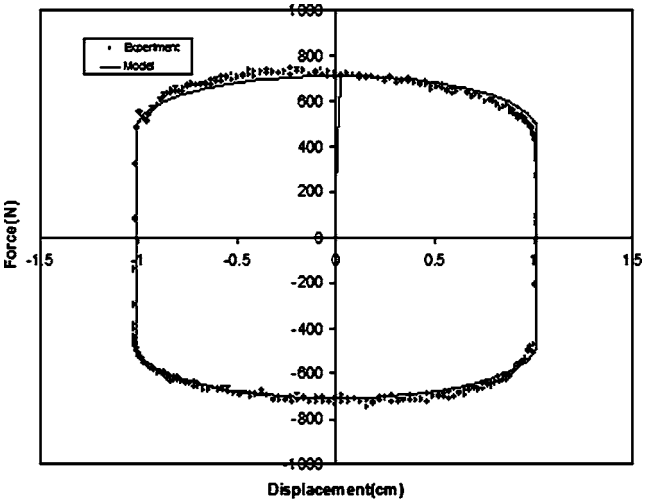


(a) Force - Displacement

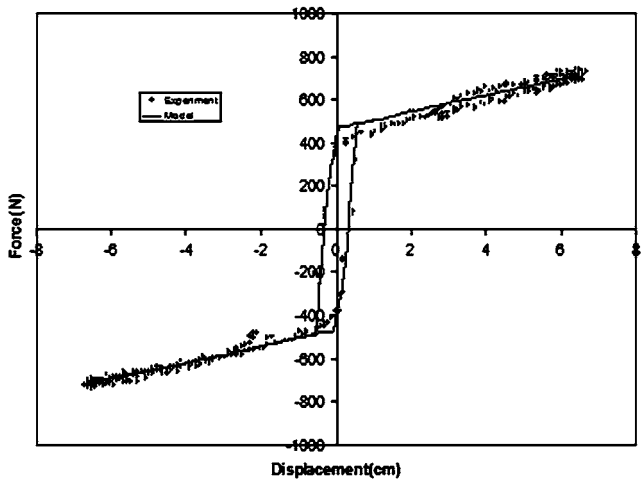


(b) Force - Velocity

Fig. 8 Comparisons between the proposed model and experimental results for a sinusoidal motion at 1.0 Hz and 1.016 cm (0.4 in) amplitude at 1.0 Amps electric current input. The effective bulk modulus is $\beta=4.0\times10^8$ N/m^2.



(a) Force - Displacement



(b) Force - Velocity

Fig. 9 Comparisons between the proposed model and experimental results for a sinusoidal motion at 1.0 Hz and 1.016 cm (0.4 in) amplitude at 1.5 Amps electric current input. The effective bulk modulus is $\beta=4.0\times10^8$ N/m^2.

process neglects the variation of compressibility. In this study, it is reasonable to assume that the effective bulk modulus β_1 and β_2 are the same constant. Assuming $\beta_1=\beta_2=\beta$, and letting $\Delta P=P_1-P_2$, Eqs. (14) and (15) can be rewritten as:

$$\frac{d\Delta P}{dt}=\beta(A_p\dot{U}-Q_{\text{out}})\left[\frac{1}{(L_1-U)A_p}+\frac{1}{(L_2+U)A_p}\right] \tag{17}$$

where $\dot{U}$ is the piston velocity. Equation (17) can be solved if Q_{out} is determined as a function of pressures P_1 and P_2. In this analysis, mass flow rate continuity with the incompressible fluid assumption is considered for the flow of a MR fluid through the channel. Based on the flow paths of the fluid through the piston shown in Fig. 4, the pressure drop P_1-P_2 between chamber one and chamber two is mainly related to the pressure drop ΔP_{vis} due to the fluid viscosity in the central channel, and the pressure drop ΔP_m across the MR valve normal to the axis of the shaft. When the MR fluid is activated, the viscous pressure drop ΔP_{vis} is much smaller than pressure drop ΔP_m in the MR valves. Therefore, the pressure drop in the central path is neglected. Thus, using Eq. (9*a*), the volume flow rate through MR valves can be expressed as the function of pressure drop, $\Delta P=P_1-P_2$:

$$Q_m=0 \quad \left(\tau_y\geqslant\frac{p'R}{2}\right) \tag{18a}$$

$$Q_m=\frac{\left(\frac{p'R}{2}-\tau_y\right)^{(n+1)/n}\pi R^3}{\left(\frac{p'R}{2}\right)^3 k^{1/n}}\left[\frac{\left(\frac{p'R}{2}-\tau_y\right)^2}{\frac{3n+1}{n}}+\frac{2\tau_y\left(\frac{p'R}{2}-\tau_y\right)}{\frac{2n+1}{n}}+\frac{\tau_y^2}{\frac{n+1}{n}}\right] \quad \left(\tau_y<\frac{p'R}{2}\right) \tag{18}$$

where $p'=(P_1-P_2)/2L_{\text{mrf}}$, and $R=D_1/2$.

Assuming incompressible fluid in the piston channel, the total volume flow rate is:

$$Q_{\text{out}}=NQ_m \tag{19}$$

where N is the number of MR valves. Using Eqs. (17)–(19), the pressure drop ΔP can be obtained. The pressure drop ΔP_{vis} gen-

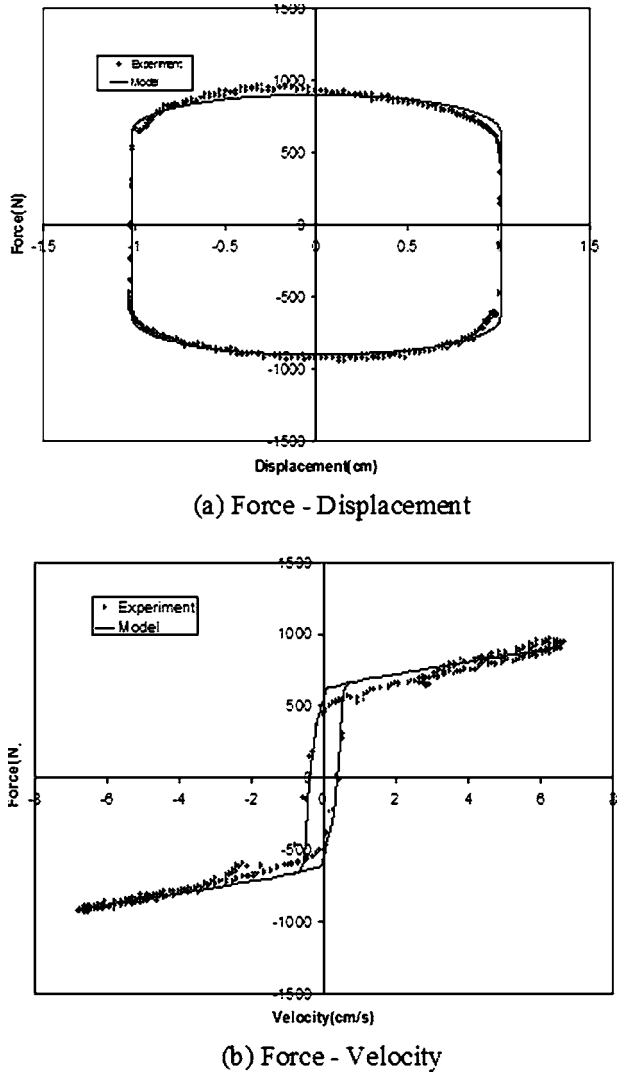


Fig. 10 Comparisons between the proposed model and experimental results for a sinusoidal motion at 1.0 Hz and 1.016 cm (0.4 in) amplitude at 2.0 Amps electric current input. The effective bulk modulus is $\beta=4.0\times10^8$ N/m^2.

erates the viscous damping force. Therefore, the total pressure drop across the piston can be expressed as:

$$F = A_p(\Delta P + \Delta P_{\text{vis}}) + F_f \,\text{sign}(\dot{U}) \tag{20}$$

where F_f is the damper's seal friction force (about 50 N). For a Newtonian Poiseuille flow in a circular channel, the pressure drop ΔP_{vis} is:

$$\Delta P_{\text{vis}} = \frac{128\mu Q L_{\text{vis}}}{\pi D_2^4} \tag{21}$$

where $Q=A_p\dot{U}$.

5 Validation of the Proposed Model

In this section, the theoretical results will be compared to the experimental data to examine the accuracy of the proposed model. The rheological property of a MR fluid (LORD 132LD) is measured by a universal dynamic spectrometer (MR-Rheometer), modified by Paar Physica [23]. Figure 7 presents the shear stress versus shear strain rate at various magnetic field strengths. It is found that at a range of low magnetic field strength, the MR fluid exhibit Bingham plastic behavior with a constant plastic viscosity

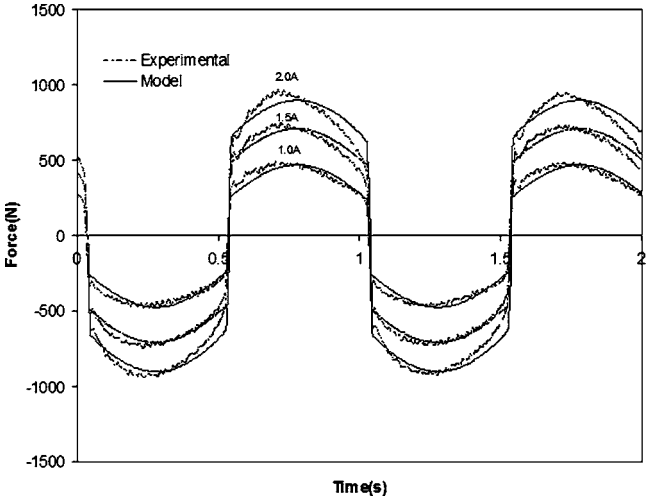


Fig. 11 Comparisons between the proposed model and experimental results for a sinusoidal motion at 1.0 Hz and 1.016 cm (0.4 in) amplitude at three different electric current input in time domain. The effective bulk modulus is $\beta=4.0\times10^8$ N/m^2.

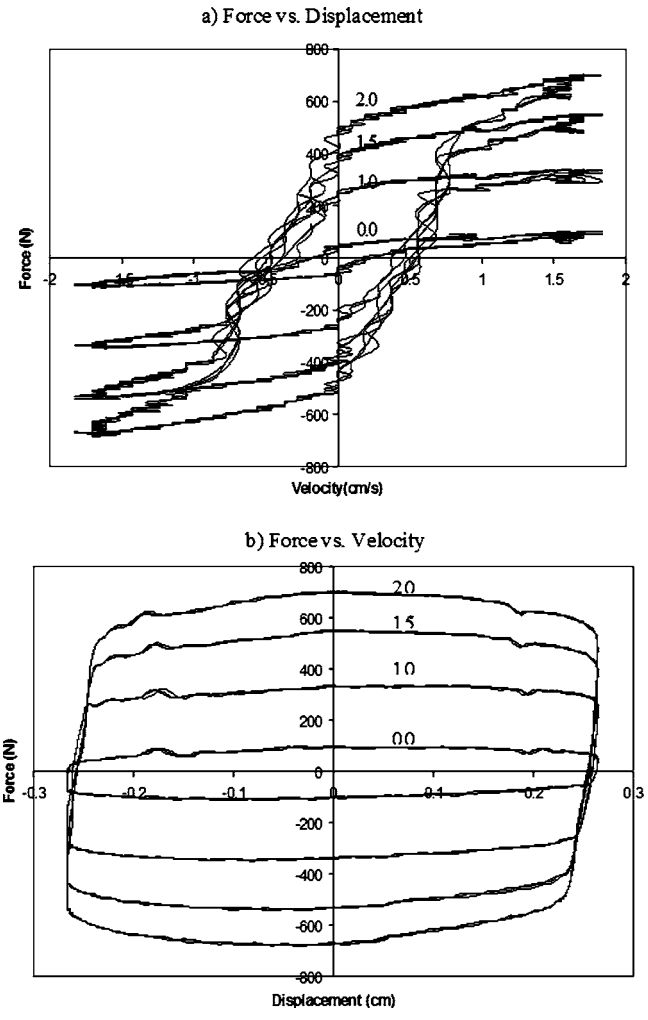
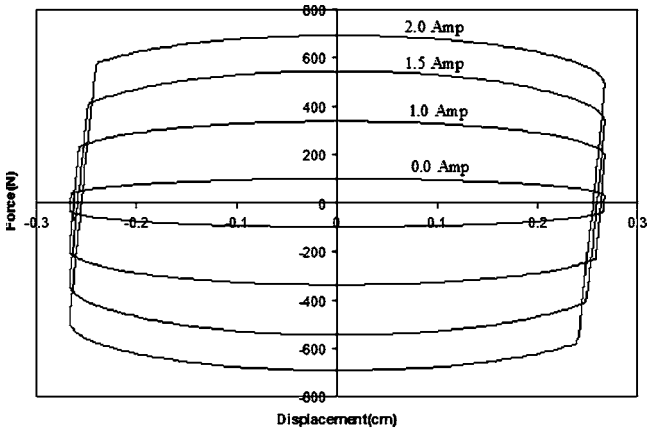
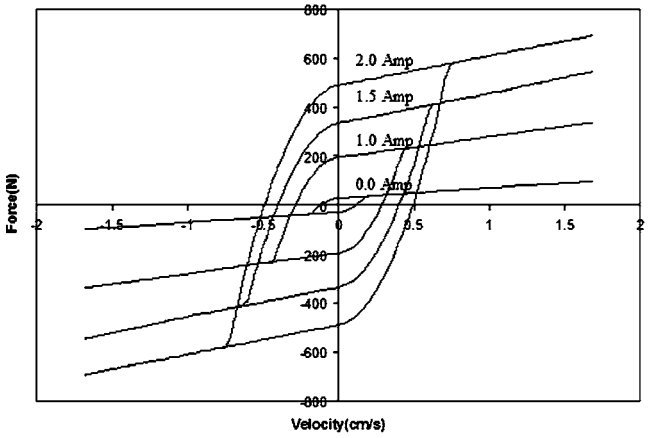


Fig. 12 Experimental results for hysteresis of UNR MRD-001 MR fluid damper subjected to 0.0, 1.0, 1.5, and 2.0 A input electric currents and harmonic motion at 1.0 Hz and 0.267 cm amplitude



a) Force vs. Displacement



b) Force vs. Velocity

Fig. 13 Theoretical results for hysteresis of UNR MRD-001 MR fluid damper subjected to 0.0, 1.0, 1.5, and 2.0 A input electric currents and harmonic motion at 1.0 Hz and 0.267 cm amplitude ($\beta=5.5\times10^7$ N/m^2)

being equal to the zero-field viscosity of Newtonian fluids. However, at higher magnetic field strengths, the MR fluid exhibits pseudo-plastic behavior with a field-dependent yield stress. Considering the shear thinning effect, the Herschel-Bulkley model is adapted to represent the measured shear stress versus shear strain rate data. The fluid index of the MR material is found to be: $k=430$ and $n=0.4$. Both the Bingham model and Herschel-Bulkley model are used to fit the experimental data. Herschel-Bulkley model presents more accurate representation of the MR fluid behavior.

A prototype, UNR MRD-001 MR fluid damper shown in Fig. 2 is tested on an Instron hydraulic dynamometer. The Instron has a maximum stroke of 15.24 cm (6.0 in), and it is equipped with a 22 KN (5000 lbs) load cell that measures the damping forces of the MR fluid damper. A displacement/velocity transducer measures the displacement and velocity of the hydraulic actuator. Experimental data is collected at a sampling rate of 200 Hz.

A series of tests is conducted to measure the response of the damper under various sinusoidal loading conditions. A performance test consists of frequency sweep held at constant displacement input. Each specific amplitude and frequency represents a run of the performance test. In each test, the input electric current applied to the prototype MR fluid damper is maintained at a constant level of 0, 1.0, 1.5, and 2.0 Amps. The geometric and material properties are given in Table 1. The physical parameters of the damper and the fluid properties used in this study are given in

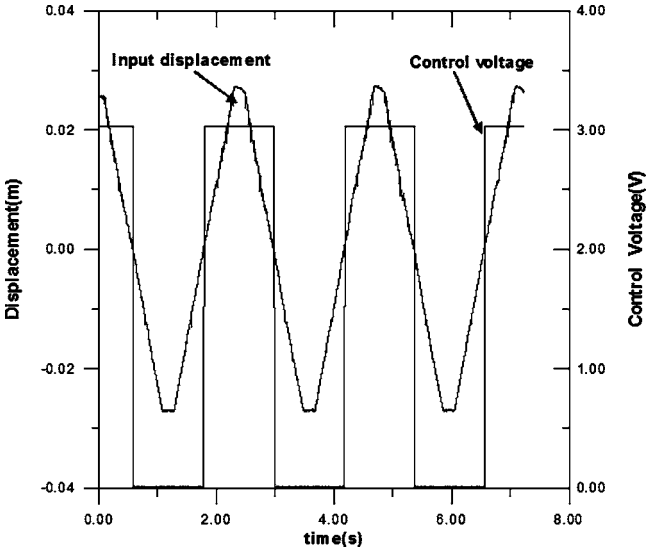


Fig. 14 Time history of input displacement and applied control voltage

Table 1. The yield stress τ_y is the function of the applied magnetic field strength. The shear yield stress, τ_y, and the input current I for the MR damper, as follows:

$$\tau_y = 2.7\times10^4 I^{1.5} \tag{22}$$

The units for τ_y and I, are Pa and Amp, respectively. This relation is obtained by analyzing the magnetic field distribution at the MR valve.

Figures 8–10 demonstrate the comparisons between experimental and analytical results for a sinusoidal input at a frequency of 1.0 Hz with three different current inputs of 1.0, 1.5, and 2.0 Amps. The analytical results are obtained by numerically integrating Eq. (17), and its substitution into Eq. (20). The value of effective bulk modulus β is 4.0×10^8 N/m^2 assuming that there is about 0.1% of air content in the MR fluid. As shown in these figures, the proposed model can accurately predict the dynamic response of the MR fluid damper for both force-displacement and force-velocity with variable input current. Figure 10 shows the same results in the force-time domain. As can be seen the damping force of the MR fluid damper increased by increasing the input currents. Also, the theoretical and experimental results agree well. Similar results are obtained for a range of displacements (0.254–1.016 cm) and frequencies (0.5–3.0 Hz).

The effect of fluid's bulk modulus on the dynamic response of MR damper is also examined by adding more air into the MR fluid. The MR fluid is exposed to air for seven days, and no vacuum procedure is applied to the MR fluids inside the damper before tested. Figures 11 and 12 present experimental and theoretical results for force-displacement and force-velocity hysteresis loops for the case where more air was entrapped in the MR fluid, with four different current inputs of 0.0, 1.0, 1.5, and 2.0 Amps. The value of bulk modus β in the theoretical analysis is 5.5×10^7 N/m^2 assuming that there is 1.0% of air content in the MR fluid. It can be seen the high nonlinearity in force-velocity hysteresis loops is mainly due to the compressibility of MR fluids. The proposed fluid mechanics-based model can provide a simple and direct method to capture this nonlinear behavior, if the value of fluid's bulk modulus is known. Both the experimental and theoretical results indicate that in addition to dissipating, the MR fluid dampers store energy. This characteristic may be important in the semi-active control design of these dampers.

Another series of experiments are conducted to verify the theoretical model under a step input motion. A ramp displacement input motion, resulting a constant velocity, is applied to the MR

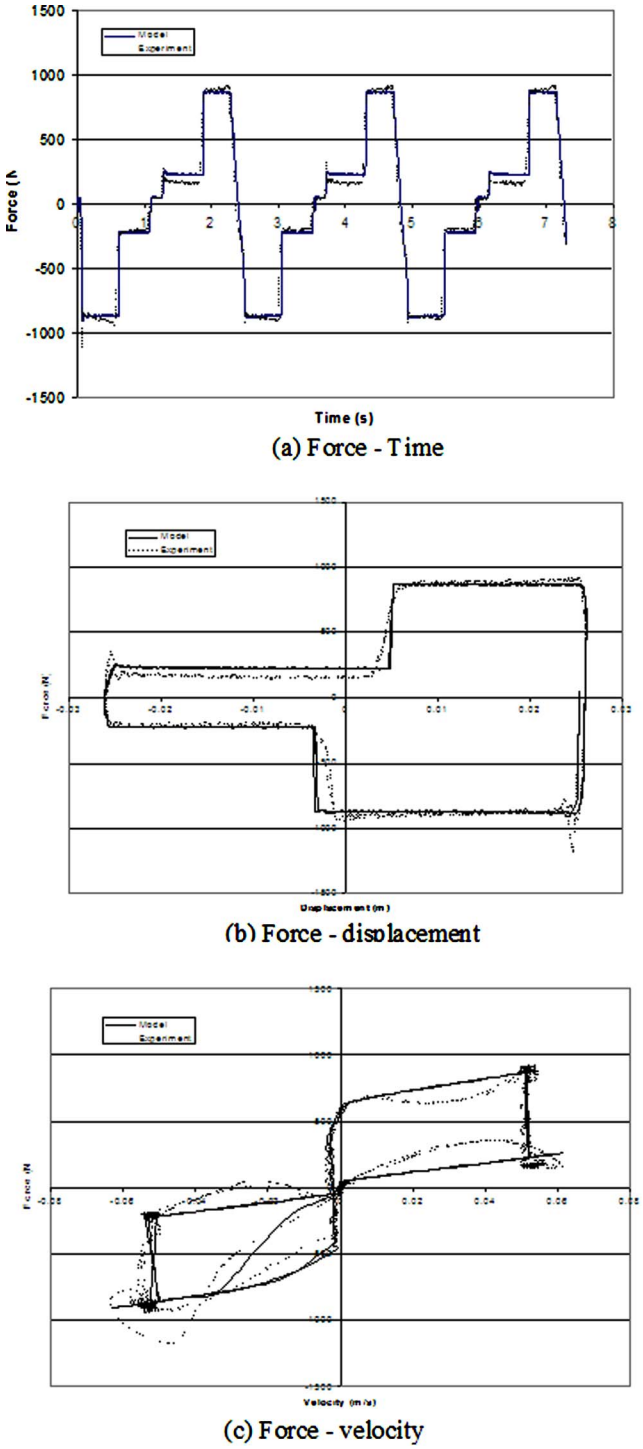


Fig. 15 Comparison between the theoretical model and experimental results

fluid damper. The input applied voltage from 0 to 3.0 V (corresponding from 0 to 2.0 Amps) is also applied to activate the damper based on a simpler feedback on-off control strategy. This control law is achieved by examining the damper movement. A step change in the applied voltage from 0 to 3.0 V is applied when the piston displacement passes through the center with a positive velocity. When the piston displacement passes through the center with a negative velocity, then the applied voltage is changed to zero. Additional linear motion transducer and a control board are used for these tests. The measured displacement and applied voltage are shown in Fig. 13. The experimental and theoretical predicted responses are compared in Figs. 14(*a*)–14(*c*) and 15. As can be seen, a good agreement is obtained between theoretical and experiment results except at the regions when the activation and de-activation occurs. This is due to the time delay. The control signal generation, the electromagnet activation and the MR fluid phase change are responsible for the total time delay. It should be noted that the response time of the MR fluid reaching its rheological equilibrium from current off to on states is longer than from that of on to off. In this study, the off-on response time is about 50 ms, while the on-off response time is about 25 ms.

6 Summary and Conclusions

In this study, an analysis of ER and MR controllable fluid flow through pipes or parallel plates was presented using the Herschel-Bulkley model. A steady, one-dimensional laminar flow was assumed. A simplified closed-form expression is presented for determining the pressure drop (gradient) as a function of material properties, geometry, and volumetric flow rate. In addition, a dynamic model is developed based on fluid mechanics and the Herschel-Bulkley flow analysis to predict the behavior of MR/ER fluid dampers.

The effect of fluid compressibility on the performance of the MR fluid damper is considered by the inclusion of the effective bulk modulus in the proposed model. The theoretical model is validated by comparing the analytical results with experimental data for a prototype MR fluid damper. It is demonstrated that the proposed fluid-mechanics based model can accurately predict the dynamic response of a MR fluid damper over a wide range of operating conditions.

Acknowledgment

This study is funded by the Army Research Office. The authors are thankful for the encouragement by Dr. Gary Anderson, the Program Director. Thanks are also extended to Dr. W. H. Li at Nanyang Technological University, Singapore, for providing the material properties of the MR fluid.

References

[1] Symans, M. D., and Constantinou, M. C., 1999, "Semi-Active Control Systems for Seismic Protection of Structures: A State-of-the-Art Review," Eng. Struct., **21**(6), pp. 469–487.
[2] Carlson, J. D., Catanzarite, D. M., and StClair, K. A., 1996, "Commercial Magneto-Rheological Fluid Devices," Int. J. Mod. Phys. B, **10**(23–24), pp. 2857–2865.
[3] Rabinow, J., 1951, "Magnetorheological Fluid," U.S. Patent No. 2,575,360.
[4] Winslow, W. M., 1949, "Induced Fibration of Suspensions," J. Appl. Phys., **20**, pp. 137–140.
[5] Ginder, J. M., 1996, "Rheology Controlled by Magnetic Fields," Encyclopedia of Applied Physics, **16**, pp. 487–503.
[6] Gordaninejad, F., and Breese, D. G., 2000, "Magneto-Rheological Fluid Dampers," U.S. Patent No. 6,019,201.
[7] Gordaninejad, F., and Kelso, S. P., 2001, "Fail-Safe Magneto-Rheological Fluid Dampers for Off-Highway, High-Payload Vehicles," J. Intell. Mater. Syst. Struct., **11**(5), pp. 395–406.
[8] Weiss, K. D., Duclos, T. G., Carlson, J. D., Chrazn, M. J., and Margalia, A. J., 1993, "High-Strength Magneto- and Electro-Rheological Fluids," *SAE Technical Paper Series No. 932451*, presented at the *1993 International Off-Highway and Powerplant Congress and Exposition*, Milwaukee, Wisconsin, USA.
[9] Ginder, J. M., 1998, "Behavior of Magnetorheological Fluids," MRS Bull., **23**, pp. 26–29.
[10] Rankin, P. J., Ginder, J. M., and Klingenberg, D. J., 1998, "Electro- and Magneto-Rheology," Curr. Opin. Colloid Interface Sci., **3**(4), pp. 373–381.
[11] Phillips, R. W., 1969, "Engineering Applications of Fluids With a Variable Yield Stress," Ph.D. Dissertation, University of California.
[12] Lou, Z., Ervin, R. D., and Filisco, F. E., 1994, "A Preliminary Parametric Study of Electrorheological Dampers," J. Fluid Mech. Eng., Trans. ASME, **116**, pp. 570–576.
[13] Gavin, H. P., Hanson, R. D., Filisko, F. E., 1996, "Electrorheological Dampers. 1. Analysis and Design," ASME J. Appl. Mech., **63**(3), pp. 669–675.
[14] Stanway, R., Sproston, J. L., and El-Wahed, A. K., 1996, "Application of Electrorheological Fluids in Vibration Control: A Survey," Smart Mater. Struct., **5**, pp. 464–482.
[15] Yang, G., Spencer, B. F., Carlson, J. D., and Sain, M. K., 2002, "Large-Scale MR Fluid Dampers: Modeling and Dynamic Performance Considerations,"

Eng. Struct., **24**(3), pp. 309–323.
[16] Goodwin, J. W., Markham, G. M., and Vincent, B., 1997, "Studies on Model Electrorheological Fluids," J. Phys. Chem. B, **101**, pp. 1961–1967.
[17] Mokeev, A. A., Korobko, E. V., and Vedernikova, L. G., 1992, "Structural Viscosity of Electrorheological Fluids," J. Non-Newtonian Fluid Mech., **42**, pp. 213–230.
[18] Shulman, Z. P., and Korobko, E. V., 1978, "Convective Heat Transfer of Dielectric Suspensions in Coaxial Cylindrical Channels," Int. J. Heat Mass Transfer, **21**(5), pp. 543–548.
[19] Stanway, R., Sproston, J. L., and Stevens, N. G., 1987, "Non-Linear Modeling of an Electrorheological Vibration Damper," J. Electrost., **20**, pp. 167–184.
[20] Halsey, T. C., Martin, J. E., and Adolf, D., 1992, "Rheology of Electrorheological Fluids," Phys. Rev. Lett., **68**(10), pp. 1519–1522.
[21] Felt, D. W., Hagenbuchle, M., Liu, J., and Richard, J., 1996, "Rheology of a Magnetorheological Fluid," J. Intell. Mater. Syst. Struct., **7**(5), pp. 589–593.
[22] Wang, X., and Gordaninejad, F., 1999, "Herschel-Bulkley Analysis of Electro- and Magneto-Rheological Controllable Fluids in Flow Mode," *Proceedings of the 7th International Conference on ER Fluids and MR Suspensions*, R. Tao, ed., World Scientific, Singapore, pp. 568–578.
[23] Li, W. H., 2000, "Rheology of MR Fluids and MR Damper Dynamic Response: Experimental and Modeling Approaches," Ph.D. Dissertation, School of Mechanical and Production Engineering, the Nanyang Technological University, Singapore.
[24] Wang, X., and Gordaninejad, F., 2000, "Study of Field-Controllable, Electro- and Magneto-Rheological Fluid Dampers in Flow Mode Using Herschel-Bulkley Theory," *Proceedings of the SPIE Smart Structure and Materials Conference*, Vol. 3989, Newport Beach, California, pp. 232–243.
[25] Lee, D. Y., and Wereley, N. M., 2000, "Analysis of Electro- and Magneto-Rheological Flow Mode Dampers Using Herschel-Bulkley Model," *Proceedings of the SPIE Smart Structures and Materials Conference*, Vol. 3989, Newport Beach, California, pp. 244–252.
[26] Kamath, G. M., and Wereley, N. M., 1997, "Nonlinear Viscoelastic-Plastic Mechanisms-Based Model of an Electrorheological Damper," J. Guid. Control Dyn., **20**(6), pp. 1125–1132.
[27] Kamath, G. M., Wereley, N. M., and Jolly, M. R., 1999, "Characterization of Magnetorheological Helicopter Lag Dampers," J. Am. Helicopter Soc., **44**(3), pp. 234–248.
[28] Li, W. H., Yao, G. Z., Chen, G., Yeo, S. H., and Yap, F. F., 2000, "Testing and Steady State Modeling of a Linear MR Damper Under Sinusoidal Loading," Smart Mater. Struct., **9**(1), pp. 95–102.
[29] Gamota, D. R., and Filisko, F. E., 1991, "Dynamic Mechanical Studies of Electrorheological Materials: Moderate Frequencies," J. Rheol., **35**, pp. 399–425.
[30] Burton, S. A., Makris, N., Konstantopoulos, I., and Antsaklis, P. J., 1996, "Modeling the Response of ER Damper: Phenomenology and Emulation," J. Eng. Mech., **122**(9), pp. 897–906.
[31] Spencer, B. F., Dyke, S. J., Sain, M. K., and Carlson, J. D., 1997, "Phenomenological Model for Magnetorheological Dampers," J. Eng. Mech., **123**(3), pp. 230–238.
[32] Choi, S. B., Lee, S. K., and Park, Y. P., 2001,"A Hysteresis Model for the Field-Dependent Damping Force of a Magnetorheological Damper," J. Sound Vib., **245**(2), pp. 375–383.
[33] Ehrgott, R. C., and Masri, S. F., 1992, "Modeling the Oscillatory Dynamic Behavior of ER Materials," Smart Mater. Struct., **8**(5), pp. 601–615.
[34] Gavin, H. P., Hanson, R. D., and Filisko, F. E., 1996, "Electrorheological Dampers. 2. Testing and Modeling," ASME J. Appl. Mech., **63**(3), pp. 676–682.
[35] Chang, C. C., and Roschke, P., 1998, "Neural Network Modeling of a Magnetorheological Damper," J. Intell. Mater. Syst. Struct., **9**(9), pp. 755–764.
[36] Leva, A., and Piroddi, L., 2002, "NARX-Based Technique for the Modeling of Magneto-Rheological Damping Devices," Smart Mater. Struct., **11**(1), pp. 79–88.
[37] Gang, J, Sain, M. K., Pham, K. D., Spencer, B. F., and Ramallo, J. C., 2001, "Modeling MR-Dampers: A Nonlinear Blackbox Approach," *Proceedings of the 2001 American Control Conference*, (Cat. No. 01CH37148), p. 429.
[38] Peel, D. J., Stanway, R., and Bullough, W. A., 1996, "Dynamic Modeling of an ER Vibration Damper for Vehicle Suspension Applications," Smart Mater. Struct., **5**(5), pp. 591–606.
[39] Sims, N. D., Peel, D. J., Stanway, R., Johnson, A. R., and Bullough, W. A., 2000, "The Electrorheological Long-Stroke Damper: A New Modeling Technique With Experimental Validation," J. Sound Vib., **229**(2), pp. 207–227.
[40] Sims, N. D., Stanway, R., Peel, D. J., Bullough, W. A., and Johnson, A. R., 1999, "Controllable Viscous Damping: An Experimental Study of an Electrorheological Long-Stroke Damper Under Proportional Feedback Control," Smart Mater. Struct., **8**(5), pp. 601–615.
[41] Patten, W. N., Mo, C., Kuehn, J., and Lee, J., 1998, "A Primer on Design of Semi-Active Vibration Absorbers (SAVA)," J. Eng. Mech., **124**(1), pp. 61–68.
[42] Symans, M. D., and Constantinou, M. C., 1997, "Experimental Testing and Analytical Modeling of Semi-Active Fluid Dampers for Seismic Protection," J. Intell. Mater. Syst. Struct., **8**(8), pp. 644–657.
[43] Stringer, J. D., 1976, *Hydraulic Systems Analysis: An Introduction*, John Wiley and Sons, New York.

Modeling the Tribochemical Aspects of Friction and Gradual Wear of Diamond-Like Carbon Films

Feodor M. Borodich
School of Engineering,
Cardiff University,
P.O. Box 925,
Queen's Buildings,
Cardiff CF24 3AT, UK

Chad S. Korach
Department of Mechanical Engineering,
State University of New York,
Stony Brook, NY 11794-2300

Leon M. Keer
Center for Surface Engineering and Tribology,
Technological Institute,
Northwestern University,
Evanston, IL 60208-3109
e-mail: l-keer@northwestern.edu

Problems in nanomechanics often need to combine mechanical approaches together with methods of physics and chemistry that are outside of the traditional mechanics scope. Recent experimental studies of dry sliding between two hydrogenated DLC (diamond-like carbon) coated counterparts in low oxygen environment showed that adsorbates have considerable influence on friction and the friction coefficient increases with the increasing of the time interval between contacts. The observed friction phenomena are assumed caused by a reaction between the adsorbate and carbon atoms of the coatings, and when the slider passes a point on the track, it removes mechanically some adsorbate from the surface. The mechanical action leads to reexposure of the surface to gases in the environment. This paper focuses on physical and tribochemical processes that occur in sliding contact between the DLC coated slider and the counterpart. We develop further our recently presented model of the process and assume that there is a transient short-life high temperature field at the vicinities of contacting protuberances that may cause various transformations of the surface. In particular, the sp^3 phase of DLC films may transform to graphite-like sp^2 carbon. Our model does not depend directly on the assumption that the adsorbate is oxygen. However, due to the prevalence of oxygen in atmospheric gas it is assumed that the adsorbate is oxygen in the model presented. We suppose that first an oxygen molecule becomes physically adsorbed to the surface and then due to rubbing the molecule dissociates into two chemically active oxygen atoms. This process leads to chemisorbtion between the carbon atoms of the coating and the "sticky" oxygen atoms. The latter atoms can interact with the counterpart. Our modeling established a direct connection between this kind of molecular friction and gradual wear. In particular, it is shown that the initial roughness of the DLC surface may have a considerable influence on the probability of breaking bonds during mechanical removal of adsorbate. Ab initio calculations of the bond dissociation energies between carbon atoms and carbon-oxygen atoms were performed using GAUSSIAN98 at the Møller-Plesset level of model chemistry. The bond dissociation energy found for the carbon-carbon bonds is 523 kJ/mol*, while for the carbon-oxygen bonds it is* 1447 kJ/mol*. It is assumed that carbon wear particles will not be formed during gradual degradation since the coating carbon molecules are dissolved within the environment gases. The model helps to explain how microscopic processes, such as the breaking and forming of interatomic bonds, may affect macroscopic phenomena, such as friction and wear.* [DOI: 10.1115/1.2172267]

1 Introduction

This paper deals with the modeling of dry sliding between two DLC (diamond-like carbon) coated counterparts in a low oxygen environment. It is known that understanding the physical and tribochemical processes that occur in a sliding contact is a question of particular interest for tribology [1]. The applicability of the Elovich (Roginsky-Zeldovich) equation for modeling the gas adsorption process when environmental molecules form bonds to the surface, and the modeling of friction and gradual wear when some part of the adsorbate is removed due to rubbing the surface by a slider are examined.

1.1 Amorphous Carbon Films. Amorphous carbon (*a*-C) films are widely used as protective overcoats in various industries. In particular, protective *a*-C coatings are used in the computer industry, when a coating of several nanometers thick is placed over the recording head and the magnetic alloy of a computer hard disk. Sometimes *a*-C films are confused with crystalline diamond coatings, implicitly assuming that *a*-C films must have either a nanocrystalline or larger microcrystalline structure. In fact, amorphous carbon is a disordered three-dimensional material where the sp^2 and sp^3 hybridizations are both randomly present. As an illustration, see Fig. 6 by Frauenheim et al. (p. 128 [2]) obtained by computer modeling of *a*-C films. It clearly shows that the structure is amorphous. Some sp^1 hybridization could also be present in the *a*-C films.

Because *a*-C has no exact geometric configuration, the following characteristics of *a*-C samples are commonly used: (i) the density of a sample; (ii) the sp^3/sp^2 ratio; and (iii) the hydrogen concentration. Amorphous hydrogenated carbon coatings are denoted as *a*-C:H. If the density and sp^3/sp^2 ratio are high, and the hydrogen concentration is low (up to 30%) then *a*-C:H coatings are called hydrogenated diamond-like carbon (DLC:H) coatings. Otherwise, they are soft *a*-C:H films. The ratio between the car-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 28, 2004; final manuscript received December 2, 2005. Review conducted by A. Maniatty. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

bon atoms in these hybridizations depends on deposition conditions and other factors. Quite often during deposition of DLC films, some metals, for example tungsten, are added. The properties of DLC:H films include high wear resistance, and high hardness. Friction coefficients for various DLC films span a range from 0.001 to more than 0.6; this large disparity is attributed mainly to the differences in sp^3/sp^2 ratio and in hydrogen content of the coatings [3].

1.2 Friction Forces. Friction occurs when two bodies rub each other. According to Kragelsky and Shchedrov [4], the main sources of friction were discovered already in the earlier studies of friction by G. Amontons, Ph. de la Hire, J. T. Desaguliers, by L. Euler and S. K. Kotelnikov [5,6]. These were the molecular adhesion of contacting solids and mechanical interlocking between protuberances of surfaces. The Amontons law of friction is

$$\mu = F_f/P \tag{1}$$

where μ is the coefficient of friction, F_f is the friction force, and P is the normal load (see also [4,7]). Later C. A. Coulomb [8] suggested that the total friction force could be represented as the sum of a constant force (A) depending on sticking of surfaces and a force that depends on the pressure

$$F_f = A + \mu P \tag{2}$$

Evidently, if A is small in comparison with μP then the Coulomb law (2) reduces to the Amontons law (1). G. A. Tomlinson [9] presented a model of the molecular friction that involves a lattice oscillation mechanism: Friction without wear occurs when atoms of one surface during sliding begin to oscillate due to interaction with atoms of the opposite surface, therefore producing the lattice vibrations. As a result the external mechanical energy (the driving force) dissipates as sound energy. Derjaguin [10] gave a molecular meaning to Coulomb's force A. He wrote

$$F_f = \mu(P + Sp_0) \tag{3}$$

where S is the true area of the interacting surface, and p is the specific attractive force. Hence, the term $A = \mu S p_0$ represents the tangential component of the force of molecular interactions. Derjaguin suggested distinguishing between the true friction coefficient μ_t and the apparent friction coefficient μ_a where

$$\mu_t = \frac{F_f}{P + Sp_0} \text{ and } \mu_a = \frac{F_f}{P}$$

Currently, it is common to present the total friction force (F_f) as the sum of the mechanical or deformational ($F_{f,\text{mech}}$) and molecular friction ($F_{f,\text{mol}}$) forces [11]

$$F_f = F_{f,\text{mech}} + F_{f,\text{mol}} \tag{4}$$

However, Akhmatov [12] noted that Derjaguin's model considered only crystal solids and, therefore, not all molecular mechanisms affecting friction were considered.

V. A. Zhuravlev [13] under supervision of N. N. Davidenkov gave an explanation of the Amontons law (1) using a statistical approach. The model assumes that the frictional force is caused by molecular interaction between atoms and is proportional to the true contact area S. Zhuravlev's theory of multiple contact assumed that all protuberances are spherical, the radii of the protuberances are equal to each other, and their heights are distributed by some statistical law. He showed that if the heights distribute linearly, then even for elastic contact, when the contact area under a single spherical protuberance is proportional to the power of the load with exponent 2/3, the linear relation occurs because as the load increases, the number of contacting protuberances also increases. A model similar to Zhuravlev's, is the Greenwood-Williamson model [14], which is commonly used by researchers.

Thus, when two bodies are in sliding contact, various mechanical, physical, chemical, and physiochemical processes take place at the interface [1,3,15–18]. In analogy to Eqs. (2) and (4), the measured friction coefficient may be presented as

$$\mu = \mu_{\text{mech}} + \mu_{\text{mol},1} + \mu_{\text{mol},2} \tag{5}$$

where $\mu_{\text{mol},1}$ is the part of the friction coefficient caused by the breaking of interatomic bonds between the adsorbate and a surface, $\mu_{\text{mol},2}$ is the part caused by all other molecular effects, in particular by the lattice oscillation mechanism, and μ_{mech} is the part caused by mechanical effects, in particular by interlocking protuberances of the roughness. Hence, it would be naïve to expect that one could model simultaneously all components of the friction force. In this paper we concentrate on modeling the part of the friction coefficient caused by breaking of interatomic bonds between the adsobate and a surface, $\mu_{\text{mol},1}$.

1.3 Environmental Effects on the Friction of DLC:H Films. As we have mentioned, dry sliding interaction between a slider and a coated surface leads to various chemical and physical transformations in DLC coatings. It is known that surfaces exposed to ambient air possess adsorbed layers of hydrocarbons and other small molecules or atoms [19]. The substance that adsorbs and desorbs is referred to as the adsorbate and the substance to which the adsorbate adsorbs is called the substrate. It was established that the friction and wear properties of carbon based coatings are dependent on the atmospheric conditions, the structure of the films, and surface chemistry [3,20–24]. In particular, the friction behavior of the studied coating materials strongly depend upon the presence of water vapor and oxygen molecules on coating layer surfaces and their associated ionic properties. However, various kinds of carbon based coatings, i.e., graphite, diamond, *a*-C, hydrogen-free DLC, and DLC:H, may demonstrate drastically different tribochemical and frictional properties under various environmental conditions. Adding oxygen to the experimental environment reduces friction of graphite [21], and hydrogen-free DLC, while friction of DLC:H films increases [3]. It is well-known that there are two types of adsorption to a surface, namely, chemisorption, when chemical bonds between the adsorbate and the substrate are formed, and physisorption, that involves van der Waals interactions such as dispersion forces or a dipolar interaction. Usually, the interface contains various species both physically adsorbed and chemically bonded [25].

Hydrogenated DLC surfaces exhibit inert behavior due to the satisfaction of surface sites by hydrogen atoms [3]. The existence of hydrogen atoms on the surface has two effects: (1) the surface energy is reduced due to the satisfaction of surface sites and (2) the adsorption of atmospheric gases is inhibited. The significance of the second point is important since the absorbed species are thought to control the frictional and wear properties of the surface. The mechanism behind the effect of adsorbates on the friction coefficient is by the breaking of atomic bonds. There is a certain probability associated with the mechanical bond breaking involving the strength of the molecular bonds and the probability and strength of adhesion between the counterface and the adsorbates. Erdemir [3] noted that the presence of dihydrated carbon atoms (two surface positive charged hydrogen nuclei bonded to one carbon atom) on the surface may provide better shielding or a higher degree of chemical passivation; a dipole configuration at the sliding interface may give rise to electrostatic repulsion between the hydrogen-terminated sliding surfaces of the films and reduce friction.

Recently, Erdemir et al. [26] have introduced a procedure for plasma-enhanced chemical vapor deposition based on the generation of hydrogen-rich methane plasmas. The procedure leads to the formation of super-low friction hydrogenated DLC films with friction coefficients as low as 0.003 [27]. Our studies are mainly concentrated on the friction of hydrogenated DLC films, and our models will be compared with some recent experimental results obtained by Heimberg et al. [23] that studied the friction of DLC:H films. A simplified modeling of the experiments was presented in an earlier paper [28]. The model is now refined with the

addition of an explicit analysis of the gas adsorption and desorption at points along the cycle, which then is averaged over an entire cycle.

2 Modeling Friction of DLC:H Films

2.1 Tribochemistry of Hydrogen Free and Hydrogenated DLC Films. Molecules or atoms adsorbed on surfaces of materials are viewed in modern models as a collection of point masses coupled to the surface by harmonic springs, so that they can oscillate about the adsorption bonds [29]. Dry sliding in the presence of oxygen and vapors can cause a reaction between an adsorbate and carbon atoms of the coatings. A possible explanation for this reaction is the dissociation of a physically adsorbed oxygen molecule into two oxygen atoms. During sliding, protuberances of surfaces interact with each other and generate oscillations of the adsorbates. As a result, high temperature at the sliding interfaces is generated. Due to transient short life high temperature fields and shear stresses, the sp^3 phase of DLC films can be transformed to graphite-like sp^2 carbon. This is the phenomena known as graphitization of DLC films that was observed under contact loading by various authors [30,31]. The local increase of temperature may also activate the dissociation of physically adsorbed molecular oxygen into atomic oxygen. The oxygen molecule O_2 first becomes physically adsorbed and then undergoes dissociation into two oxygen atoms due to activation by the sliding interface. This leads to chemisorbed O atoms to the active sites of the surface. This scenario agrees with a previous discussion of a similar process [25].

Ab initio calculations of the bond dissociation energies between carbon atoms and carbon-oxygen atoms were performed using GAUSSIAN98 [32]. A fourth order Møller-Plesset perturbation theory [33] model chemistry was used in an open-shell form. Calculations were performed using the 6−31+G(d) basis set in the GAUSSIAN98 software package, which allows for polarization and diffuse functions that may be necessary when analyzing bond dissociation energies. The bond dissociation energy computed for a single carbon-carbon bond is 523 kJ/mol, for the carbon-oxygen bonds it is 1447 kJ/mol, while for the carbon-hydrogen bonds it is 295 kJ/mol. The results compare well with published experimental bond energies for these pairs [34]. Bond distances, corresponding to position of lowest energy, computed for C-H and C-O are 1.12 Å and 1.16 Å, respectively. The bond distance for the carbon-oxygen bond is expected to be for a higher-order bond (triple bond), thus the bond distance for a carbon-oxygen double or single bond would be larger (1.2–1.43 Å). Hence, the bonds for O bonded to C are slightly longer than those of H bonded to C (mainly due to the atomic radius of O being larger than H). We may assume that the probability of an O adatom in contact with the counterpart on the same surface as H is higher. It is difficult to say that the slider will remove O adatoms before H since more energy is needed to remove the higher-order carbon-oxygen bond than the C-H bond.

The C-C bond energy is more complex than just breaking a single bond. The surface C atoms in hydrogen-free DLC films are bonded to their immediate neighbors with three σ bonds and the fourth bond may be free. The free bond may be passivated by an adatom (a water molecule, O, H, etc.) [3]. Even if any of the σ bonds are not as strong as a single C-C bond, these three bonds must be dissociated to remove a C molecule. Therefore, more energy is required to dissociate a C atom from the surface of hydrogen-free DLC than that of dissociating an O adatom from the surface.

The process is different for a hydrogenated DLC surface. Hydrogen entrapment in the DLC during deposition will change the film structure. Hence, a surface C atom in hydrogenated DLC films may be bonded not to its immediate C neighbors but to H atoms. Since the strength of the bond between H and C is less than that of the lead constituents of atmospheric gas (i.e., O, N, Ar) or water molecules, there is a greater probability that the hydrogen will be removed due to mechanical action of the slider. Once hydrogen is removed from the DLC surface, unoccupied surface sites will be present and gas components from the atmosphere may absorb to these sites. The atmospheric gas components having a larger bond energy with carbon will be more difficult to remove mechanically. This mechanism will increase the overall friction between the surfaces.

Scientific research in the literature has presented various and sometimes contradictory results concerning the friction behavior of DLC films. For example, Erdemir [3] reported that the friction coefficient of hydrogen-free DLC in dry nitrogen was approximately 0.65 and dropped to 0.25 after moist laboratory air was introduced. The same test for the highly hydrogenated DLC showed that the friction coefficient increased from 0.003 to 0.06. Marchon et al. [20] reported that the friction coefficient of a head/disc pair (DLC film in contact with either a Mn/Zn ferrite or $CaTiO_3$ ceramic slider) in pure nitrogen was 0.2 and increased to 1.2 when oxygen was introduced into the system. Marchon and coworkers explained the phenomenon by oxidation of the surface when exposed to oxygen, and the surface species can desorb as carbon monoxide and/or carbon dioxide. On the other hand, Donnet et al. [35] found that added oxygen does not change super-low friction behavior of hydrogenated DLC observed in ultrahigh vacuum. However, the addition of water vapor drastically increased the friction coefficient from about 0.01 to 0.1. Studies by Heimberg et al. [23] of DLC:H films in a nominally dry N_2 environment showed an increase in friction as sliding speed was decreased. This test was explained by gas adsorption to the surface. The results of Marchon et al. [20] studying hydrogen free amorphous carbon look as a contradiction to Erdemir's results. However, the results can be explained by noting that the chemistry of the DLC films and the sliders in these experiments were of a different chemical structure. This sensitivity to chemical structure of the DLC films leads to uncertainties in the application of the Donnet et al. [35] conclusion to hydrogenated DLC films used by Heimberg et al. [23]. Furthermore, Heimberg et al. [23] did not provide enough information to give preference to either oxygen or water vapor as sources of the increase in the friction coefficient. Perhaps both of them have influence and this is an area of future concern regarding DLC friction studies. Our model does not depend directly on the assumption that the adsorbate is O. However, due to the prevalence of O in atmospheric gas we will assume that the adsorbate is O in the model presented.

A competing process of hydrogen diffusion to the surface from the DLC film may be present, but is not taken into account in the model explicitly. Instead the mechanism will be included in the probability of hydrogen adsorbate removal. Mobility of surface species will be disregarded, although for high temperature applications this effect would need to be studied further. The presence of hydrogen in the DLC film and on the surface may have an effect on the chemical wear mechanism that has been discussed. The effects are largely focused on the removal of CO and CO_2 from the surface. Changes in the atomic structure of the film will cause localized effects such as bond energy and distance changes on the carbon bonds within the film that will introduce different probability values for the removal of CO and CO_2 molecules.

2.2 Kinetics of Adsorption. One of the main characteristics of the adsorption process is the fractional coverage of adsorbate θ, which can be defined as

$$\theta = \frac{N_c}{N}$$

where N_c is the number of surface sites occupied by adsorbate, N is the total number of active substrate adsorption sites, and $0 \leqslant \theta \leqslant 1$. Since the interactions between physisorbed species are very weak, the species with such bonds can be easily desorbed, for example, due to thermal desorption or by mechanically break-

ing the interacting forces between the adsorbate and substrate.

One of the most cited of the adsorption kinetics equations is an empirical equation (the Elovich equation) introduced by Roginsky and Zeldovich [36] in 1932 and actively used by Elovich since 1934. The equation is used to describe the adsorption process of chemical reactions involving chemical adsorption of gases on solid surfaces without the desorption of products

$$\frac{d\theta}{dt} = Be^{-\alpha\theta} \tag{6}$$

Here α and B are constants during any one experiment, and θ is the fractional part (the relative amount) of solute adsorbed at time t.

To describe the kinetics of gas adsorption in studies of friction between a ball and carbon based coatings during repeated passes of the ball over the same regions of a surface, Zaïdi et al. [21] considered a particular integral of Eq. (6) when the initial time $t_0=0$ and $\theta(0)=0$

$$\theta(t_p) = \frac{1}{\alpha}\ln(1 + B\alpha t_p)$$

where t_p is the period of the cycle. This formula gives the value of θ after one period if there was no adsorbate at the beginning of the cycle.

After integrating the Elovich Eq. (6), one has

$$\theta(t) = \frac{1}{\alpha}\ln[e^{\alpha\theta(t_0)} + B\alpha(t - t_0)] \tag{7}$$

where t_0 and $\theta(t_o)$ are the initial instant and the initial relative amount of adsorbate, respectively.

It should also be noted that the concepts of the total number of active sites N and the number of occupied surface sites N_c are not uniquely defined. If we consider the active sites of carbon, then initially they are all occupied by hydrogen. During sliding, some amount of hydrogen is removed. The created active sites can be occupied by water vapor or oxygen atoms. When the slider comes into contact with the considered region of the surface again, not only hydrogen atoms are removed from the surface but also the adsorbate. Hence, it is necessary to make additional assumptions in order to use the concept of the coverage θ.

Let us consider a section of surface of nominal size lxl, where $l=100$ nm. If we take into account the real roughness of the surface then the real area A is greater than 100×100 nm^2. Since the DLC bond distance is on the order of 1 A, then the surface A would have approximately 10^6 active carbon sites if all H adatoms were removed. Since the probability of adatom removal is quite low, if we assume the slider to remove n_HH atoms, where $n_H=1000$, during one pass (2 passes per a cycle), then it would remove $(100\text{ cycles})\times(1000\text{ atoms})\times(2\text{ passes/cycle})=2\times10^5$ atoms within A over 100 fast cycles. If θ is considered as related to the total coverage, then $N=10^6$ total surface sites. Thus, $N_c=(N-2\times10^5)$ and we obtain from the definition of coverage that $\theta=80\%$. This is the percentage of H covering the surface. Because we have neglected the adsorption of environmental gases in order to obtain this estimation, the coverage would be greater than this. Though, we are not interested in the hypothetical total coverage mainly caused by repulsive H atoms, but only in "sticky" O atoms or molecules of water vapor. Hence, take N as the number of hydrogen free active sites (i.e., $N=2\times10^5$) and N_c is the number of sites covered by "sticky" adatoms, then a surface with a mixture of H, O, and water vapor will be present.

We would like to repeat that the Elovich Eq. (6) is an empirical equation, where the constants B and α may be found from independent experiments. Due to the lack of experimental data, we may use indirect estimations for the constants. The constant B can be regarded as the initial rate since $(d\theta/dt)$ tends to B as θ tends to zero at $t=0$. Experiments by Valenzuela-Calahorro et al. [37] on adsorption of progesterone in an ethanol solution at surfaces of four carbonaceous materials showed that the adsorption specific rate constant k_1 varies from 5.29×10^{-4} to 44.85×10^{-4} s^{-1}. Hence, we can expect that B that has the same order as k_1, which is of order of magnitude 10^{-4} s^{-1}. If $\theta(0)=0$ then for a small interval Δt, we have $\theta=B\Delta t$. Let us take $B=2\cdot10^{-4}$ s^{-1} and $\Delta t=10$ s, then $\theta=2\cdot10^{-3}$. From the definition of coverage with $N=2\times10^5$ (the number of H free sites that has been estimated above) the number of surface sites occupied by adsorbates is $N_c=400$.

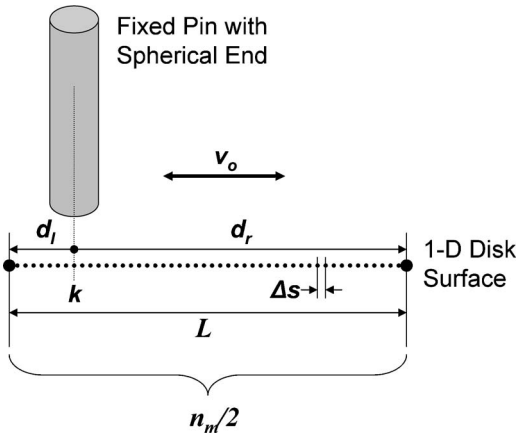


Fig. 1 Schematic of the pin-on-disk tribometer numerical simulation performed along a linear track with $n_m/2$ number of measurement points, Δs spacing between points, track length L, and distances d_l and d_r from the left and right, respectively, of the k-th point to the track endpoints. The relative speed between the fixed pin and oscillating disk is v_o. The pin has a spherical end cap.

We can estimate the upper bound for the number of removed H adatoms from the energy equality $\mu_{\text{mol},2}p_aAl\approx n_H N_A^{-1}J_{\text{C-H}}$, where p_a is the mean Hertzian pressure, $J_{\text{C-H}}$ is the bond dissociation energy for a carbon-hydrogen bond, and N_A is Avogadro's constant $(6.022\times10^{23}\text{ mol}^{-1})$. Taking $J_{\text{C-H}}$ as 295 kJ/mol, $p_a=1.1$ GPa, and $\mu_{\text{mol},2}=0.001$, the number of H atoms (n_H) to satisfy the equality is $n_H=2250$, which is an upper bound on the number of removed atoms. The value of N_c computed above is within this upper bound, so the value taken for B in the model is of the proper order.

Taking into account that the bond dissociation energy for a single carbon-oxygen bond $J_{\text{C-O}}\cong1447$ kJ/mol$\approx4J_{\text{C-H}}$, we obtain $\mu_{\text{mol},1}=\mu_{\text{mol},2}(J_{\text{C-O}}N_c/J_{\text{C-H}}n_H)\approx1.6\mu_{\text{mol},2}$ for the experiments with $\Delta t=10$ s. Even though an explicit relation for $\mu_{\text{mol},2}$ has been derived the model will use a constant value for $\mu_{\text{mech}}+\mu_{\text{mol},2}$, although the derived relation shows the uncertainty in taking $\mu_{\text{mech}}+\mu_{\text{mol},2}$ as a constant. Since the track length L is taken as 5 mm, then the speed of the slider should be less then 0.5 mm/s to allow a noticeable change in the friction coefficient. This value corresponds with the experimental results of Heimberg et al. [23], where a speed of 533 m/s showed little change in friction.

2.3 Kinetics of Adsorbate Removal. The model pertains to pin-on-disk tests when a DLC coated flat disk repeatedly slides over a fixed DLC coated ball. All values in the process will be averaged over each cycle.

Let n_m be the total number of friction measurements in a cycle, and thus $n_m/2$ measurements per track length, L (see Fig. 1). For example, if $n_m=1000$, then 1 to 500 measurements of μ are taken during sliding forward (from left to right in Fig. 1) and 501 to 1000 measurements of μ are taken during sliding in reverse (from right to left in Fig. 1). Hence, two measurements are taken at each

point of the track during every cycle. The span between measurement points Δs is defined as $\Delta s = 2L/n_m$. The distances d_l and d_r from the k-th point on the track to the left and the right ends of the track, respectively, are $d_l = \Delta s(k-1)$ and $d_r = L - \Delta s(k-1)$.

First, consider only the k-th point (a region) on the track during the i-th cycle. When the ball passes the point, it mechanically removes an amount of adsorbate from the surface and reexposes the surface to gases in the environment. Hence, the amount of adsorbate before the slider passes the point and after, i.e., after removing the adsorbate, is different. The total number of active sites N could also change because not only adatoms are removed during sliding but also H atoms from the surface. However, this effect is neglected because the overall change is relatively small. The coverage $\theta(k,i)$ increases in value continuously due to adsorption and decreases abruptly (the jump $(\theta(k,i))^{\pm}$) twice during a cycle when the disk passes the point during forward and backward sliding, i.e., $\theta(k,i)|_{t+0+} = \theta(k,i)|_{t+0-} - (\theta(k,i))^{\pm}$ where t is the moment in time of the passing.

3 Comparison With Experiments

An excellent data set on friction between two DLC:H coated counterparts was presented by Heimberg et al. [23]. An attempt to describe the experiments using a simplified approach was recently presented by Borodich and Keer [28]. The model [28] assumed that the ball had a circular track on a fixed disk. Hence, only one point on the track could be considered. Also the number of active sites N was fixed and restricted to only oxygen atoms. The current model is refined using the above estimations of bond energy. The refined model is employed to describe the experiments presented by Heimberg et al. [23].

In the experiments, the DLC coatings were deposited to a 1 mm thickness on H13 steel flats and either 6.35 mm sapphire balls or 12.7 mm diameter steels balls. Reciprocating pin-on-disk tests were performed in a nominally dry nitrogen environment with an oxygen level of 0.7%. The track length L was 5 mm. Each track was initially run-in for 1000 cycles at high sliding speeds (1–5 mm/s). Then, a series of "speed-dependent" and "time-delay" tests were performed:

(i) Each speed-dependent test began with 100 cycles at high speed (1–5 mm/s) followed by 20 cycles at a lower speed. The ball remained in contact with the disk, and sliding continued with no delay.

(ii) Each time-delay test was performed at a constant high sliding speed ($\geqslant$1 mm/s). The tests began with 100 cycles without stops, followed by 20 cycles with fixed delays at the endpoints of the track. The ball always remained in contact with the disk during the test.

The friction behavior of one of the DLC surface sets for a series containing seven different speed-dependent tests was presented in detail. At slow stages, the speed of the ball (v_0) was 513.0 μm/s, 303.0 μm/s, 100.0 μm/s, 75.0 μm/s, 50.0 μm/s, 30.0 μm/s, and 10 μm/s.

The increase of the adsorbate surface coverage depends on the time $\Delta t(k,i)$ between passes. For the first part of the cycle (sliding from left to right), $\Delta t(k,i) = d_l/\nu(i-1) + d_r/\nu(i)$, while for the second part of the cycle (sliding from right to left), $\Delta t(k,i) = 2d_r/\nu(i)$.

If the Elovich Eq. (6) is employed, then from, Eq. (7)

$$\theta(k,i)|_{t+\Delta t(k,i)+0+} = \ln[\exp(\alpha\{\theta(k,i)|_{t+0-} - [\theta(k,i)]^{\pm}\}) + \alpha B \Delta t(k,i)]/\alpha, \tag{8}$$

where $t+0-$ and $t+\Delta t(k,i)+0+$ are the moments just before and after the successive passing of the k-th point.

It is assumed that the jump in the coverage $(\theta(k,i))^{\pm}$ is proportional to $\theta(k,i)|_{t+0-}$ and to the probability $p(0 \leqslant p \leqslant 1)$ of the mechanical breaking of interacting bonds (removing the adsorbate from the surface). Hence, one can write

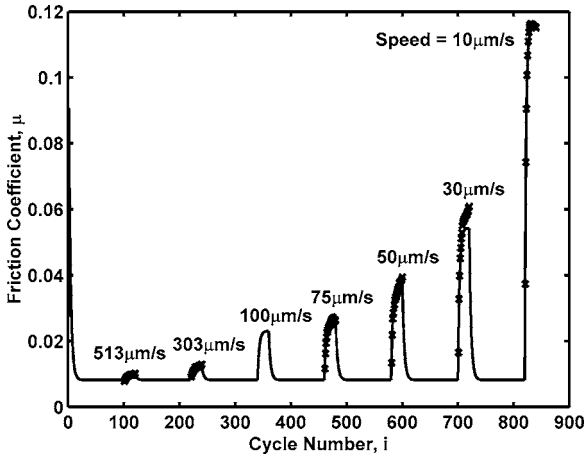


Fig. 2 Plot of the friction coefficient (μ) versus cycle number (i) for the entire experimental range. Data points (x) are the reciprocating pin-on-disk experimental data (from Heimberg et al. [23]) for DLC on DLC contact. Solid lines represent data obtained using numerical simulation of the friction coefficient based on the integral of the Elovich Eq. (11), with $B = 4.1 \cdot 10^{-4}$ s^{-1}, $\mu_c = 0.0065$, $\alpha = 0.6$, $p = 0.14$, $t_p = L/\nu_o$, $c = 0.81$, $t_o = 0$, and $\theta(0) = 0.001$, and relative speed between pin and disk (ν_o) of 10 μm/s, 30 μm/s, 50 μm/s, 75 μm/s, 100 μm/s, 303 μm/s, and 513 μm/s.

$$[\theta(k,i)]^{\pm} = p\,\theta(k,i) \tag{9}$$

Substituting Eq. (9) into, Eq. (8), we obtain

$$\theta(k,i)|_{t+\Delta t(k,i)+0+} = \ln\{\exp[\alpha\theta(k,i)(1-p)] + \alpha B \Delta t(k,i)\}/\alpha \tag{10}$$

The probability p may depend on the absolute temperature, the speed of the ball, roughness of the surfaces, and other parameters of the experiment. Indeed, the higher the absolute temperature, the higher the desorption rate as shown by the Boltzmann-Arrhenius equation. One can also expect that the higher the roughness of the contacting surfaces, the lower the value of p because the part of the molecules or atoms adsorbed on the surfaces would be located in valleys, where the atoms would be protected from mechanical removal by the protuberances of the surfaces. Thus, the probability p can vary during the observed process. As a first approximation, it is assumed that p is constant during an experiment with a particular sample, although p may vary from one sample to another one.

The model assumes that the molecular friction is proportional to the number of interacting bonds that were mechanically broken when the slider passes a point on the track. Hence, the friction coefficient $\mu_{\mathrm{mol},1}$ is proportional to $(\theta(tk,i))^{\pm}$ and the average friction coefficient that is measured on the ith cycle ($\mu(i)$) is

$$\mu(i) = \mu_c + \frac{1}{n_m}\sum_{k=1}^{n_m} \mu_{\mathrm{mol},1}(i,k) = \mu_c + \frac{c}{n_m}\sum_{k=1}^{n_m} [\theta(k,i)]^{\pm} \tag{11}$$

Where $\mu_c = \mu_{\mathrm{mech}} + \mu_{\mathrm{mol},2}$ is assumed to be constant for a specific sample, and c is a first-order constant.

In our numerical simulations of the speed-dependent test, the speed of the ball is taken as 1053 μm/s during the high speed cycles. If $B = 4.1 \cdot 10^{-4}$ s^{-1}, $\mu_c = 0.0065$, $\alpha = 0.6$, $p = 0.14$, $t_p = L/\nu_0$, $c = 0.81$, $t_0 = 0$, and $\theta(0) = 0.001$, and employing the model based on the integral of the Elovich equation, Eqs. (9)–(11), one can obtain Fig. 2 and observe an excellent agreement between the simulated and experimental friction coefficient data over the entire range of cycle numbers.

Using the same model and the same values of the model parameters in the previous paragraph, one can perform numerical

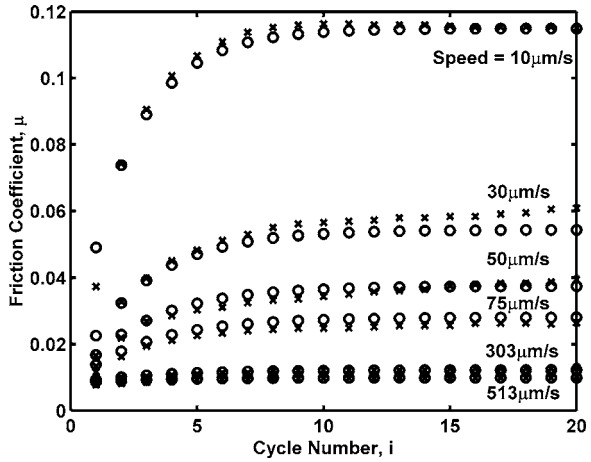


Fig. 3 Plot of the friction coefficient (μ) versus cycle number (i) for the slow speed ranges only. Data points (x) are the reciprocating pin-on-disk experimental data (from Heimberg et al. [23]) for DLC on DLC contact. Data points (o) represent data obtained using the numerical simulation of the friction coefficient based on the integral of the Elovich Eq. (11), with $B=4.1\cdot10^{-4}$ s^{-1}, $\mu_c=0.0065$, $\alpha=0.6$, $p=0.14$, $t_p=L/\nu_o$, $c=0.81$, $t_o=0$, and $\theta(0)=0.001$, and relative speed between pin and disk (ν_o) of 10 μm/s, 30 μm/s, 50 μm/s, 75 μm/s, 303 μm/s, and 513 μm/s.

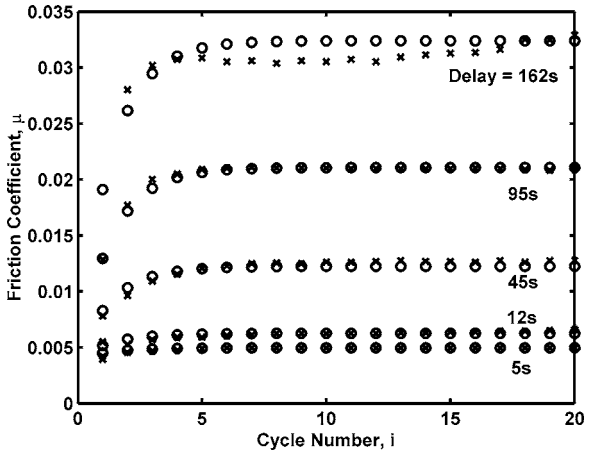


Fig. 4 Plot of the friction coefficient (μ) versus cycle number (i) for the time-delay tests. Data points (x) are the reciprocating pin-on-disk experimental data (from Heimberg et al. [23]) for DLC on DLC contact. Data points (o) represent data obtained using the numerical simulation of the friction coefficient based on the integral of the Elovich equation (11), with $B=2.3\cdot10^{-4}$ s^{-1}, $\mu_c=0.003$, $\alpha=0.6$, $p=0.3$, $t_p=L/\nu_o+t_d$, $c=0.81$, and relative speed between pin and disk $\nu_o=1053$ μm/s. Time delays at the track endpoints (t_d) were 5 μs, 12 μs, 45 μs, 95 μs, and 162 μs.

simulations of the friction coefficient at the slow speed stages. Figure 3 presents the graphs of the relation between friction coefficient and the cycle number at slow stages of the experiment, obtained using numerical simulations based on the integral of the Elovich equation.

In the time-delay test, the durations of the stops (t_d) at the endpoints of the track were 5.0 s, 12.0 s, 45.0 s, 95.0 s, and 162 s, respectively. It was assumed in the numerical simulations of the time-delay test that the fixed speed ν_0 was 1053 μm/s. The time $\Delta t(k,i)$ between passes by the slider of the k-th point is $\Delta t(k,i)=2d_l/\nu(i)+t_d$ for the first part of the cycle (sliding from left to right), and $\Delta t(k,i)=2d_r/\nu_0+t_d$ for the second part of the cycle (sliding from right to left).

Taking $\mu_c=0.003$, $B=2.3\cdot10^{-4}$ s^{-1}, $\alpha=0.6$, $t_p=L/\nu_0+t_d$, $p=0.3$, and $c=0.81$, and employing Eqs. (9)–(11), one can obtain Fig. 4, which shows a good agreement between the numerical simulations and the experimental data.

4 An Alternative Model

After submitting our paper, we learned that another model describing the same experimental results [23] was published by Dickrell et al. [39]. First, we note that the model by Dickrell et al. [39] is closer to our earlier published, simplified model [28] than to the model that we have presented in this paper. Both the Dickrell et al. model [39] and our simplified model [28] used the concept of the coverage θ without discussing the necessity of the additional assumptions as was discussed above. In addition, both models consider the process at a single point without averaging over all points of the track during an entire cycle. It is of interest to compare this alternative model [39] with our simplified model [28].

It is assumed in both models that the jump $[\theta_{k+1}]^{\pm}$ in the coverage after the slider passes a point is proportional to the coverage θ_k and a parameter. In [39] this parameter was considered as the ratio of the fraction of the surface covered at the exit of the pin contact to that at the entrance and denoted as $(1-\lambda)$, while in [28] this parameter was interpreted as the probability p of the mechanical breaking the interacting bonds (removing the adsorbate from the surface). Hence, Dickrell et al. [39] wrote $\theta_{k,\text{out}}=\lambda\theta_{k,\text{in}}$ while in our notations (see Eq. (9))

$$\theta_{k,\text{out}}=\theta_{k,\text{in}}-[\theta_{k+1}]^{\pm}=\theta_{k,\text{in}}-p\theta_{k,\text{in}}=(1-p)\theta_{k,\text{in}}$$

It is known (see, e.g., [28]) that if one considers the Taylor series expansion of the right-hand expression of the Elovich Eq. (6) and the series is truncated to the first order in θ, then one obtains the linear equation of adsorption kinetics

$$\frac{d\theta}{dt}=B(1-\alpha\theta). \tag{12}$$

Equation (12) can be easily integrated. Dickrell et al. [39] considered a particular case of the equation with $\alpha=1$ and attributed it to Langmuir giving reference to Langmuir's paper [40]. However, it should be noted that in fact Langmuir did not use Eq. (12) in [40]. The following physical interpretation to the parameter B was given [39], $B=\nu P$ where ν is the adsorption coefficient and P is the gas pressure. Integrating Eq. (12) with $\alpha=1$ from t_0 to t, one obtains

$$\frac{1-\theta(t)}{1-\theta(t_0)}=e^{-B(t-t_0)}$$

or $\theta(t)=1-[1-\theta(t_0)]e^{-B(t-t_0)}$.

Thus, in the framework of the above assumptions, the following equations are used:

$$\theta_{k,\text{out}}=\theta_{k,\text{in}}-[\theta_{k+1}]^{\pm}=\theta_{k,\text{in}}-p\theta_{k,\text{in}}=(1-p)\theta_{k,\text{in}}$$

$$\theta_{k+1,\text{in}}=1-[1-\theta_{k,\text{out}}]e^{-Bt_p} \tag{13}$$

Substituting the former equation into the latter, one obtains

$$\theta_{k+1,\text{in}}=1-[1-(1-p)\theta_{k,\text{in}}]e^{-Bt_p}. \tag{14}$$

Dickrell et al. [39] noted that Eq. (14) can written as $\theta_{k,\text{in}}=\theta_0+[(1-e^{-Bt_p})+\theta_0(q-1)]\Sigma_{j=1}^{k}q^{j-1}$, where $q=(1-p)e^{-Bt_p}$. One can check this using mathematical induction. Evidently, one can use the analytical formula for the sum of geometrical progression

$$\theta_{k,\text{in}}=\theta_0+[(1-e^{-Bt_p})+\theta_0(q-1)]\frac{1-q^k}{1-q}. \tag{15}$$

However, the analytical expression (15) is valid only in the framework of the above assumptions about the validity of the linear Eq. (12) with $\alpha=1$ and that the jump $(\theta_{k+1})^{\pm}$ of the coverage is proportional to the coverage θ_k. One cannot use Eq. (15) for the entire experiment when the speed of the slider is varied. Hence, one will need to use a numerical approach to describe the entire experiment as we did above.

It is interesting to note that Dickrell et al. [39] have also mentioned the Elovich Eq. (6) and gave expressions for the first three terms of θ_k. However, they did not present the general expression for θ_k as we did (see. Eq. (17) in [28]) and they did not apply the Elovich approach for description of experiments.

Probably the advantages of the Dickrell et al. model are (i) it gives a physical meaning to parameter B, namely, B is the product of the adsorption coefficient ν and the gas pressure P; (ii) it claims that the adsorption kinetics can be described by the linear Eq. (12) with $\alpha=1$ and, therefore, θ is always less than 1, as one can see from the integral of Eq. (12); and (iii) it reduces the adsorption-desorption kinetics of strictly periodic pin-on-disc experiments to a simple geometrical progression.

On the other hand, it seems to us that there are the following drawbacks in application of this model: (i) strictly speaking, the analytical expressions of the model are not applicable to experimental results [23] because in these experiments the slider was in contact with every point of the disc twice per cycle with nonequal time intervals between contacts; (ii) even for a circular motion of the slider, the analytical expression is not valid when the speed of the slider changes and therefore, one needs to employ a numerical approach to describe the entire experiment as we did in our simplified model [28]; and (iii) to describe the experiments with constant period (assuming the circular motion of the slider), it was necessary to change the parameter λ (this is the probability that the interacting bonds will not be broken) even for the same sample (see Figs. 2 and 4 in [39]). As one can see from the above consideration, in our model the probability p, $(p=1-\lambda)$ is constant for each experimental series.

5 Discussion and Conclusions

The results of the numerical model based on the Elovich equation have been compared with the experimental results for DLC friction from Heimberg et al. [23] and show that the application of the Elovich (Roginsky-Zeldovich) [36,38] equation to describe gas adsorption is in good agreement with the results. The current model includes a mechanism for gas desorption by modeling the adsorbate removal through mechanical sliding. Adsorbate removal is assumed proportional to number of broken bonds, the jump in coverage, and is calculated at discrete points along the pin track. Earlier versions of the desorption kinetic model [28] were treated in an average way over the cycle and were not able to capture the reciprocating nature of the experiments. The addition of a desorption feature in the model differs from previous models by Zaïdi et al. [21] and Heimberg et al. [23], which modeled the friction coefficient of carbon coatings by a gas adsorption mechanism only. The advantage of the model is that it can be applied to experiments with different speeds of the slider, while the model introduced by Dickrell et al. [39] is not applicable to these experiments.

The addition of an adsorbate removal mechanism allows for a direct connection between molecular interactions and friction through the $\mu_{\mathrm{mol},1}$ term. Ab initio calculations were performed to explain the effect of hydrogen in hydrogenated carbon films on the adsorbate removal. Due to the lower dissociation energy of C-H than C=O, the H is more readily removed from the surface by the slider creating more active sites for gaseous species (i.e., O and water vapor) to adsorb to the surface. Competing mechanisms such as hydrogen diffusion from the film to the surface, surface mobility, and temperature are not included explicitly in the model. The bond energies also are only first estimates. The structure of the DLC film surface will be more complex than the approach used here for estimating the bond energy of the adsorbates and a more detailed analysis is necessary for more precise results.

Using the new adsorption-desorption model, the numerical analyses of the friction coefficient were compared with experimental DLC friction results in Figs. 2–4. The numerical results showed that for a wide range of pin speeds (10–513 μm/s) the model compares well with the experiments. This is testament that the gas adsorption and desorption physical effects on friction are being captured accurately, even though the exact adsorbed species (atomic oxygen or water vapor) is not known. As the speed of the pin relative to the disk surface decreases the time for the pin to return to a previous point increases. The increase in time allows a larger number of active sites to be satisfied by gaseous "sticky" species in turn increasing the friction coefficient due to the larger number of bonds that must be broken as the pin moves across the disk surface. A steady-state friction value is reached after approximately ten cycles.

Gradual wear of hydrogenated DLC films is approached with the idea of removal of CO and CO_2 from the surface of the film and is an oxidational wear approach. While this model may not be applicable to carbon only films, because of the increase of energy necessary for molecular dissociation off the surface, the approach is applicable to studying DLC:H films because of the weakening of the carbon bonds in the film by hydrogen. From the ab intio calculations a carbon-carbon bond has larger dissociation energy than a carbon-hydrogen bond. Though the results are for only a single carbon-carbon bond, the dissociation of carbon from a carbon network (i.e., carbon film) would have increased dissociation energy due to the increased number of bonds to be broken. Likewise, for a carbon atom in a hydrogenated carbon film, the dissociation energy would be lower.

This paper has presented results for a friction model based on the integral of the Elovich equation discretely analyzed over a reciprocating pin-on-disk wear track of DLC on DLC contact. The model has the ability to capture the physics of adsorption and desorption kinetics related to molecular effects on friction. The use of ab intio software was successfully employed to explain the friction results from a molecular approach and to elucidate the effect of molecular friction modeling on gradual, tribochemical wear. Further studies, using techniques such as molecular dynamics [41], on the effects of molecular interactions on friction are required to fully understand the phenomena occurring, especially the effects of different gas species (e.g., oxygen, nitrogen, hydrogen, water vapor) on the analysis and the effects of film microstructure (e.g., at % hydrogen content, sp^3/sp^2) on the gas adsorption/desorption kinetics and tribochemical wear.

Acknowledgment

The authors are grateful for financial support of this research from the National Science Foundation. Support was also provided by the NSF Center for Surface Engineering and Tribology at Northwestern University and the NSF IGERT program in Virtual Tribology. Thanks are extended to Professor Y.-W. Chung for discussing the problem and his valuable comments.

References

[1] Krim, J., 2002, "Friction at Macroscopic and Microscopic Length Scales," Am. J. Phys., **70**, pp. 890–897.
[2] Frauenheim, T., Jungnickel, G., Porezag, D., Sternberg, M., and Köhler, T., 1997, "Atomistic Simulations of Carbon Systems Using a Density-Functional-Based Molecular-Dynamics Method," *The Physics of Diamond*, IOS Press, Amsterdam, pp. 105–132.
[3] Erdemir, A., 2001, "The Role of Hydrogen in Tribological Properties of Diamond-Like Carbon Films," Surf. Coat. Technol., **146**, pp. 292–297.
[4] Kragelsky, I. V., and Shchedrov, V. S., 1956, *Development of the Science about Friction*, USSR Academy of Sciences Press, Moscow.
[5] Amontons, G., 1699, "De la Résistance Causée Dans les Machines," Mém. de l'Académie Royale A, pp. 275–282.
[6] Kotelnikov, S. K., 1774, *A Book Containing the Doctrine on Equilibrium and Motion of Bodies*, Naval Noble College, Saint-Petersburg.
[7] Blau, P. J., 2001, "The Significance and Use of the Friction Coefficient,"

Tribol. Int., **34**, pp. 585–591.
[8] Coulomb, C. A., 1821, *Theorie des Machines Simples, en Ayant égard au Frottement de Leurs Parties et à la Roideur des Cordages*, Bachelier, Paris.
[9] Tomlinson, G. A., 1929, "A Molecular Theory of Friction," Philos. Mag. Ser. 7, **7**, pp. 905–939.
[10] Derjaguin, B., 1934, "Molekulartheorie der äußeren Reibung," Z. Phys., **88**, pp. 661–164.
[11] Mikhin, N. M., 1978, "Calculation of Coefficients of External Friction and Preliminary Displacement," *Friction, Wear and Lubrication*, I. V. Kragelsky and V. V. Alisin, eds., Mir Publishers, Moscow, 1, pp. 54–101.
[12] Akhmatov, A. S., 1963, *Molecular Physics of Boundary Friction*, GIFML, Moscow. (English Translation, 1966, Israel Program for Sc. Translations, Jerusalem).
[13] Zhuravlev, V. A., 1940, "On Question of Theoretical Justification of the Amontons-Coulomb Law for Friction of Unlubricated Surfaces," Zh. Tekh. Fiz., **10**, pp. 1447–1452.
[14] Greenwood, J. A., and Williamson, J. B. P., 1966, "Contact of Nominally Flat Surfaces," Proc. R. Soc. London, Ser. A, **295**, pp. 300–319.
[15] Heinicke, G., 1984, *Tribochemistry*, Academie-Verlag, Berlin.
[16] Quinn, T. F. J., 1991, *Physical Analysis for Tribology*, Cambridge University Press, Cambridge.
[17] Goryacheva, I. G., 1998, *Contact Mechanics in Tribology*, Kluwer Academic Publishers, Dordrecht.
[18] Chung, Y.-W., 2001, *Practical Guide to Surface Science and Spectroscopy*, Academic Press, San Diego.
[19] He, G., Müser, M. H., and Robbins, M. O., 1999, "Adsorbed Layers and the Origin of Static Friction," Science, **284**, pp. 1650–1652.
[20] Marchon, B., Heiman, N., and Khan, M. R., 1990, "Evidence for Tribochemical Wear of Amorphous Carbon Thin Films," J. Exp. Mar. Biol. Ecol., **26**, pp. 168–170.
[21] Zaïdi, H., Paulmier, D., and Lepage, J., 1990, "The Influence of the Environment on the Friction and Wear of Graphitic Carbons. II. Gas Coverage of Wear Debris," Appl. Surf. Sci., **44**, pp. 221–233.
[22] Dugger, M. T., Wahl, K. J., Chung, Y.-W., Bhushan, B., and Rothschild, W., 1991, "An Investigation of Environmental Effects on the Wear and Surface Composition of Thin Film Magnetic Discs,"Society of Tribologists and Lubrication Engineers, Park Ridge, IL, SP31, pp. 43–53.
[23] Heimberg, J. A., Wahl, K. J., Singer, I. L., and Erdemir, A., 2001, "Superlow Friction Behavior of Diamond-Like Carbon Coatings: Time and Speed Effects," Appl. Phys. Lett., **78**, pp. 2449–2451.
[24] McGuiggan, P. M., Hsu, S. M., Fong, W., Bogy, D., and Bhatia, C. S., 2002, "Friction Measurements of Ultra-Thin Carbon Overcoats in Air," J. Tribol., **124**, pp. 239–244.
[25] Adamson, A. W., 1990, *Physical Chemistry of Surfaces*, Wiley, New York.
[26] Erdemir, A., Eryilmaz, O. L., and Fenske, G. J., 2000, "Synthesis of Diamond-Like Carbon Films with Superlow Friction and Wear Properties," J. Vac. Sci. Technol. A, **18**, pp. 1987–1992.
[27] Sanchez-Lopez, J. C., Erdemir, A., Donnet, C., and Rojas, T. C., 2003, "Friction-Induced Structural Transformations of Diamond-Like Carbon Coatings Under Various Atmospheres," Surf. Coat. Technol., **163**, pp. 444–450.
[28] Borodich, F. M., and Keer, L. M., 2004, "Modeling Effects of Gas Adsorption and Removal on Friction During Sliding Along Diamond-Like Carbon Films," Thin Solid Films (in press).
[29] Pykhtin, M. V., Lewis, S. P., Mele, E. J., and Rappe, A. M., 1998, "Collective Motion and Structural Order in Adsorbate Vibrational Dynamics," Phys. Rev. Lett., **81**, pp. 5940–5943.
[30] Ramirez, A. G., and Sinclair, R., 1999, "Wear-Induced Modifications of Amorphous Carbon in the Presence of Magnetic Media," J. Appl. Phys., **85**, pp. 5597–5599.
[31] Jahanmir, S., Deckman, D. E., Ives, L. K., Feldman, A., and Farabaugh, E., 1989, "Tribological Characteristics of Synthesized Diamond Films on Silicon-Carbide," Wear, **133**, pp. 73–81.
[32] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, Jr., J. A., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Rega, N., Salvador, P., Dannenberg, J. J., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Andres, J. L., Gonzalez, C., Head-Gordon, M., Replogle, E. S., and Pople, J. A., 2001, *Gaussian 98* (Revision A.11.2), Gaussian, Inc., Pittsburgh, PA, USA.
[33] Møller, C., and Plesset, M. S., 1934, "Note on an Approximation Treatment for Many-Electron Systems," Phys. Rev., **46**, pp. 618–622.
[34] Weast, R. C., 1980, *CRC Handbook of Chemistry and Physics*, 60th ed., CRC Press, Boca Raton, FL.
[35] Donnet, C., Le Mogne, T., Ponsonnet, L., Belin, M., Grill, A., Patel, V., and Jahnes, C., 1998, "The Respective Role of Oxygen and Water Vapour on the Tribology of Hydrogenated Diamond-Like Carbon Coatings," Tribol. Lett., **4**, pp. 259–265.
[36] Roginsky, S. Z., 1949, *Adsorption and Catalysis on Heterogeneous Surfaces*, Akademia Nauk USSR, Moscow.
[37] Valenzuela-Calahorro, C., Navarrete-Guijosa, A., Stitou, M., and Cuerda-Correa, E. M., 2003, "Retention of Progesterone by Four Carbonaceous Materials: Study of the Adsorption Kinetics," Colloids Surf., A **224**, pp. 135–147.
[38] Elovich, S. Y., and Zhabrova, G. M., 1939, Zh. Fiz. Khim., **13**, pp. 1775–1786 (English Abstract, 1941, Chem. Abstr. **35**, pp. 373).
[39] Dickrell, P. L., Sawyer, W. G., and Erdemir, A., 2004, "Fractional Coverage Model for the Adsorption and Removal of Gas Species and Application to Superlow Friction Diamond-Like Carbon," J. Tribol., **126**, pp. 615–619.
[40] Langmuir, I., 1916, "The Constitution and Fundamental Properties of Solids and Liquids. Part I. Solids," J. Am. Chem. Soc., **38**, pp. 2221–2295.
[41] Zhang, S. L., Wagner, G., Medyanik, S. N., Liu, W. K., Yu, Y. H., and Chung, Y.-W., 2004, "Experimental and Molecular Dynamics Simulation Studies of Friction Behavior of Hydrogenated Carbon Films," Surf. Coat. Technol., **177**, pp. 818–823.

A Greenwood-Williamson Model of Small-Scale Friction

Reese E. Jones
Sandia National Laboratories,
Livermore, CA 94551-0969

A Greenwood and Williamson based model for interfacial friction is presented that incorporates the presliding transition phenomenon that can significantly affect small devices. This work builds on previous similar models by developing: an analytical estimate of the transition length in terms of material and surface parameters, a general recursion formula for the case of slip in one direction with multiple reversals and constant normal loading, and a numerical method for the general three-dimensional loading case. In addition, the proposed model is developed within a plasticity-like framework and is shown to have qualitative similarities with published experimental observations. A number of model problems illustrate the response of the proposed model to various loading conditions. [DOI: 10.1115/1.2172269]

1 Introduction

Two tribological phenomena make the behavior of small devices significantly different than that of macro-devices, namely relatively large (a) pre-sliding tangential deflection (PSTD) and (b) friction at zero external load due to adhesion. PSTD, the small hysteretic displacement displayed before large scale slip, was first quantified by Courtney-Pratt and Eisner [1].[1] It has been the topic of many subsequent experimental studies (e.g., [2–6]) and its phenomenology in micro-devices has recently been explored in [7], for instance. This work will concentrate on a micro-mechanically motivated model for friction incorporating PSTD. Extensions to include the effects of normal adhesion will be discussed in Sec. 3.

A simple and often employed paradigm for the contact of microscopically rough surfaces was introduced in 1966 by Greenwood and Williamson (GW) [8]. It generates the response of the contacting surfaces by integrating the mechanical response of a representative asperity over the contacting population, which is determined from the height distribution of the asperities and the mean relative distance of the two surfaces. A GW model with the particular assumption of constant shear strength adhesive junctions was used in [9], [10, §13.4], and [11], for example, to argue that the original GW conception of contacting rough surface was compatible with the classical (macroscopic) Amontons-Coulomb law of friction. The key to this comparison is the (nearly) linear increase in real contact area with increasing pressure predicted by the GW theory. In this formulation, all the asperities are expected to slip at once and there is no PSTD effect. In the work presented here, evolving subpopulations of the contacting asperities stick or slip based on their individual contact area which leads to a length scale over which the entire population slips. This is still a very simple conception of the tribological process involved in PSTD since mechanisms such as ploughing and wear particle formation are ignored. The idea of evolving subpopulations of sticking and slipping asperities was first presented by Fujimoto et al. [12,13] and Olofsson and Hagman [14]. The current work will build upon theirs by developing (1) an analytical estimate of the transition length in terms of material and surface parameters, (2) a general recursion formula for the case of slip in one direction with multiple reversals and constant normal loading, and (3) a numerical method for the general three-dimensional loading case. In addition, this work will demonstrate qualitative similarities with experimental observations and the corresponding empirical model presented in [15]. Furthermore, it will be shown that the proposed model has a plasticity-like framework.

[1]The term "pre-sliding tangential deflection" is used here to denote the frictional transition region of a pair of surfaces and to distinguish this phenomenon from the "micro-slip" of an individual microscopic protuberance on a counterface.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 9, 2005; final manuscript received December 7, 2005. Review conducted by A. Maniatty. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 A Greenwood-Williamson Model of Friction

The basic premise of the proposed GW model is that each asperity is in a state of complete stick until a constant shear strength is exceeded, as in [11,12,16] for example. The inherent simplifying assumption is that no micro-slip takes place and no Coulomb coefficient exists at the micro-scale (a conjecture proposed in [17]), which is in contrast to models that employ Mindlin's solution [18] at the asperity level (such as [12,19]). This work is an attempt to obtain the representative qualitative features and scaling of the PSTD phenomenon without modeling the details of the single asperity solution. Given the level of idealization in treating the geometry of each asperity as a hemisphere, as is commonly done, this additional simplification seems appropriate.

Hertzian theory (see, e.g., [10, §4.2]) gives the normal force-displacement relationship for an linear elastic hemispherical asperity as

$$p_{\text{Hertz}}(\delta) = \tfrac{4}{3}E^*\sqrt{R}\delta^{3/2}, \tag{1}$$

where $E^* = \{[1-(\nu_1)^2]/E_1 + [1-(\nu_2)^2]/E_2\}^{-1}$ is an effective elastic compression modulus for the contacting bodies (denoted by subscripts 1, 2), E and ν being Young's modulus and Poisson's ratio, respectively. In addition, the effective radius is denoted by $R = (1/R_1 + 1/R_2)^{-1}$ and δ is the (far field, relative) normal displacement. The contacting area for an elastic asperity is given by

$$a = a(\delta) = \pi R\delta, \tag{2}$$

from which the relationship between contact radius α and the normal displacement δ results

$$\alpha^2 = R\delta. \tag{3}$$

From [10, §7.2] the solution for the fully sticking tangential traction field leads to the following tangential force relationship

$$\mathbf{q} = -8G^*\alpha\boldsymbol{v}, \tag{4}$$

where $G^* = [(2-\nu_1)/G_1 + (2-\nu_2)/G_2]^{-1}$ is an effective elastic shear modulus and $\boldsymbol{v}$ is the (far-field, relative) tangential displacement. By substituting (3) into (4), the following expression for the force-displacement relationship

$$\mathbf{q}_{\text{stick}} = \mathbf{q}_{\text{stick}}(\delta, \boldsymbol{v}) = -8G^* \sqrt{R}\delta^{1/2}\boldsymbol{v} \tag{5}$$

results, which is linear in shear displacement $\boldsymbol{v}$ but grows with the one-half power of the (compressive) normal displacement δ as the area of contact increases.

It is assumed that slip will ensue when the (average) shear stress[2] reaches a critical value κ

$$\kappa a = \kappa \pi R \delta \geqslant \|\mathbf{q}\| = 8G^* \sqrt{R}\delta^{1/2}\|\boldsymbol{v}\|, \tag{6}$$

as in [20], from which a critical normal displacement as a function of tangential displacement can be derived

$$\delta_c(\boldsymbol{v}) = \frac{(8G^*)}{R(\pi\kappa)^2}\boldsymbol{v}\cdot\boldsymbol{v} \tag{7}$$

(and conversely the magnitude of a critical tangential displacement is given by $v_c(\delta) = (\pi\kappa/8G^*)\sqrt{R}\delta^{1/2}$). From this relation it is seen that asperities with a large contact area stick, and those with contact area smaller than the critical value based on δ_c slip, as in [12,14]. It is assumed that magnitude and direction of the slip is consistent with the imposed uniform tangential displacement of the surface comprising the population of asperities. For a slipping asperity, it is further assumed that the force will take on the limiting value

$$\mathbf{q}_{\text{slip}} = \mathbf{q}_{\text{slip}}(\delta, \boldsymbol{v}) = -\kappa a \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = -\kappa\pi R\delta\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}, \tag{8}$$

so that the tangential force is continuous in the stick-slip transition.[3] Note that Eq. (8) is presently restricted to tangential displacement monotonically increasing in a constant direction. Extension to reversals and two-dimensional tangential displacement will be treated later in this section.

A GW model with summit height distribution $\phi(h)$ (scaled by the number of asperities in a nominal area $\mathcal{A}$) and a uniform asperity radius R is employed. The reduction of the statistics of two contacting rough surfaces to a single equivalent height distribution $\phi(h)$ is based on the fact that contact only involves the relative distance between high points on the opposing surfaces. Specifically, it requires that the relative geometry of the surfaces has, say, a Gaussian character regardless of relative alignment (see [21,22] for further details). This assumption does not limit the development of the adhesion-based model of friction under consideration, but would be restrictive if interlocking and the relative alignment of asperities on the opposing surfaces were to be included. Herein, the relative alignment of asperities is assumed to be immaterial to the response in that each asperity with geometry and properties derived from the equivalent "sum surface" [22] (i.e., $R = (1/R_1 + 1/R_2)^{-1}$, etc) is effectively contacting a flat foundation. This can be seen as another consequence of assuming that statistics of the contacting surfaces are stationary and not dependent on their gross relative displacement.

For ease of representing the ensemble response, moments of the distribution are defined as

$$M_n(d,l) = \int_l^\infty (h-d)^n \varphi(h)\mathrm{d}h = \int_0^\infty (h-d+l)^n \varphi(h+l)\mathrm{d}h, \tag{9}$$

where it is necessary to distinguish the lower integration limit l from the variable d in the kernel. In the case where the first argument and the second are identical a simpler notation $M_n(d) = M_n(d,d)$ will be used, so that $M_0(d)$ represents the number of asperities with heights over d, for example. With this in hand, the normal force due to elastic compression of asperities over the nominal area $\mathcal{A}$ is given by

$$\mathcal{P}(d) = \int_d^\infty p_{\text{Hertz}}(h-d)\varphi(h)\mathrm{d}h = \frac{4}{3}E^*\sqrt{R}M_{3/2}(d), \tag{10}$$

where $h \in [d, \infty)$ represents the height range of the contacting population, $\delta = h - d$ is the compression of an individual asperity, and d is the mean separation or the "approach" of the rough surfaces. Likewise, the tangential force for the contacting ensemble is given by

$$\begin{aligned}
\mathcal{Q}(d,\boldsymbol{v}) &= \int_{d+\delta_c}^\infty \mathbf{q}_{\text{stick}}(h-d,\boldsymbol{v})\phi(h)\mathrm{d}h \\
&\quad + \int_d^{d+\delta_c} \mathbf{q}_{\text{slip}}(h-d)\phi(h)\mathrm{d}h\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \\
&= -\int_{d+\delta_c}^\infty 8G^*\sqrt{R}(h-d)^{1/2}\varphi(h)\mathrm{d}h\boldsymbol{v} \\
&\quad - \int_d^{d+\delta_c} \kappa\pi R(h-d)\varphi(h)\mathrm{d}h\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \\
&= -8G^*\sqrt{R}M_{1/2}(d, d+\delta_c(\boldsymbol{v}))\boldsymbol{v} - \kappa\pi R(M_1(d) \\
&\quad - M_1(d, d+\delta_c(\boldsymbol{v})))\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}.
\end{aligned} \tag{11}$$

In particular, a Gaussian distribution is typically used to represent the height variation, namely

$$\varphi(h) = \frac{N_t}{\sqrt{2\pi}\sigma}\exp\left(-\frac{h^2}{2\sigma^2}\right), \tag{12}$$

where N_t is the total number of asperities identified in the nominal area $\mathcal{A}$ and σ is the standard deviation of their summit heights. As shown in [19], the ensemble response of the surface is not very sensitive to the exact shape of the distribution as long as the distribution satisfies the physically based criterion of having a (asymptotically) zero probability of contacting an asperity far above the mean surface. As in [8], the radii of all the asperities are assumed to be equal and constant. This may not be the most physically representative statistical description, as Greenwood and Wu point out in [23]. An extension of the current model to a joint probability distribution of heights and radii is straightforward (but not explored here).

Note that the standard deviation σ is a measure of roughness and a length scale. In order to illustrate the qualitative aspects of the model (11), it is non-dimensionalized by $c = \pi\kappa/8G^*$, a ratio of shear strength to shear modulus, and the independent variables are transformed by $\bar{x} = d/\sigma$ and $\bar{\mathbf{y}} = \boldsymbol{v}/\sqrt{\sigma R}$. The nondimensional total tangential force is then given by

$$\tilde{\mathcal{Q}} = -M_{1/2}(\bar{x}, \bar{x} + \bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)\bar{\mathbf{y}} - c[M_1(\bar{x}) - M_1(\bar{x}, \bar{x} + \bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)]\frac{\bar{\mathbf{y}}}{\|\bar{\mathbf{y}}\|}, \tag{13}$$

where $8G^*R\sigma N_t\tilde{\mathcal{Q}}(d/\sigma, \boldsymbol{v}/\sqrt{\sigma R}) = \mathcal{Q}(d,\boldsymbol{v})$ and the moments are evaluated for $\sigma = 1$, $N_t = 1$, namely $\phi_{\sigma=1}(z) = (1/\sqrt{2\pi})\exp(-\frac{1}{2}z^2)$. The nondimensionalized counterpart to the normal force (10) is simply

[2]The solution to the fully sticking problem (4) includes singularities in the traction field. It is assumed that the total tangential force is only representative of the actual force on the asperity on average. A discussion of how normal adhesion is related to this tangential adherence is delayed to the last section.

[3]The continuity of $\mathbf{q}$ with respect to $\boldsymbol{v}$ is not specifically required by the theoretical framework up to this point, as is demonstrated in [16]. Furthermore, a Hertz-Mindlin solution (with a corresponding steady-state value of tangential traction) would merely soften the ramp-like transition of stick to large displacement slip, see Figure 7.8 of [10].

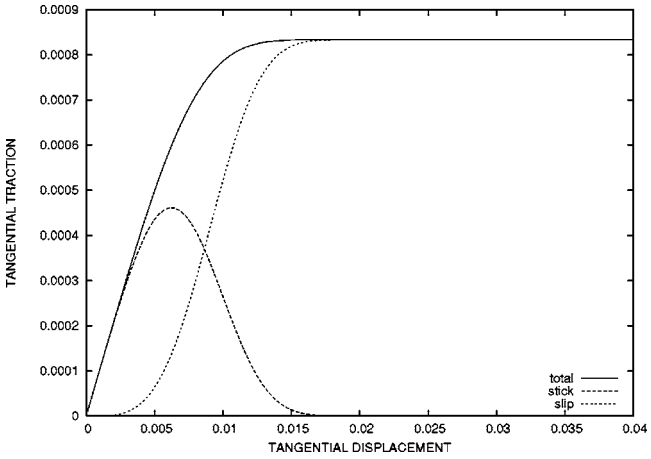


Fig. 1 Tangential force vs. tangential displacement for $c=0.01$ and $\bar{x}=1$

$$\tilde{\mathcal{P}}(\bar{x}) = M_{3/2}(\bar{x}), \tag{14}$$

where $\frac{4}{3}E^*\sqrt{R}\sigma^{3/2}N_t\tilde{\mathcal{P}}(d/\sigma)=\mathcal{P}(d)$. Note that the nondimensional version of (7) is

$$x_c(\bar{\mathbf{y}}) = \frac{1}{c^2}\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}. \tag{15}$$

Figure 1 shows the stick and slip contributions of the contacting population to the total force, using $c=10^{-2}$[4] and $\bar{x}=1$. Two obvious defining qualities of the model are (a) the steady-state value attained at large slip, which is related to the macroscopic Coulomb coefficient, and (b) the length of the transition region, which corresponds to the PSTD effect. With regard to the first property, the graph in Fig. 1 clearly shows a constant steady-state value after an initial transition region for the given approach (and therefore pressure). It should be noted that the area under this graph illustrates the amount of energy that is either stored elastically, in the case of the "stick" curve representing the first quantity on the right-hand side of (13), or dissipated, in the case of the "slip" curve representing the second quantity on the right-hand side of (13). Clearly, after the transition length has been exceeded a (asymptotically) constant amount of energy is stored elastically in the deformation of the asperities. The PSTD predicted by this model is not (entirely) reversible, due to accompanying dissipation from the slipping subpopulation. In the large slip limit an effective Coulomb coefficient can be defined as

$$\mu_\infty(\bar{x}) = \lim_{y\to\infty}\frac{\mathcal{Q}(\bar{x},\bar{y})}{\mathcal{P}(\bar{x})} = \frac{3\pi}{4}\frac{\kappa}{E^*}\sqrt{\frac{R}{\sigma}}\frac{M_1(\bar{x})}{M_{3/2}(\bar{x})}.$$

It is well known that, for certain distributions, there is a nearly linear relationship between real contact area (and the steady-state tangential force), which is proportional to $M_1(\bar{x})$, and resultant normal pressure, which is proportional to $M_{3/2}(\bar{x})$ (see Fig. 13.10 of [10]). For a (normalized) Gaussian distribution, $M_1(\bar{x})/M_{3/2}(\bar{x})$ is a smooth monotonicaly decreasing function and a range of $\bar{x}\in[-4,4]$ leads to $M_1(\bar{x})/M_{3/2}(\bar{x})\in[0.488404,1.6096]\approx\mathcal{O}(1)$. Consequently, $\kappa/G^*\in[10^{-3},10^{-2}]$ and $R/\sigma\in[10^2,10^4]$ will result in $\mu_\infty\in[10^{-2},10^0]$. So there are physical bounds on the dimensionless parameters κ/G^* and R/σ that will lead to the observed range of Coulomb coefficients. With regard to the second property, the transition length is dependent on c (and the scaling of v that resulted in $\bar{y}$). Clearly, as $c\to 0$, all asperities slip at once and the Coulomb behavior of sharp release is attained. Examining (13), both $M_{1/2}(\bar{x},\bar{x}+\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)$ and $M_1(\bar{x},\bar{x}+\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)$ go to zero within the transition length. The latter,

$$M_1(\bar{x},\bar{x}+\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(\bar{x}+\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)^2\right) - \frac{1}{2}\bar{x}\,\mathrm{erfc}\left(\frac{1}{\sqrt{2}}(\bar{x}+\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)^2\right),$$

is clearly bounded from above by $\phi(\bar{x}+\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)=(1/\sqrt{2\pi})\times\exp\left[-\frac{1}{2}(\bar{x}+\bar{\mathbf{y}}\cdot\bar{\mathbf{y}}/c^2)^2\right]$ since it is also bounded from below by zero. Hence

$$\exp(-k)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\bar{x}\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(\bar{x}+\bar{y}_T^2/c^2)^2\right)$$

leads to a (nondimensional) estimate of the transition length,

$$\bar{y}_T(\bar{x}) = c\sqrt{\sqrt{\bar{x}^2+2k}-\bar{x}}.$$

If the constant k is set equal to $k=4$, for example, the cutoff value is approximately 1.8% of the starting value $\phi(\bar{x})$ and $\bar{y}_T(\bar{x}=1)\approx 0.014$, which is in good agreement with the case shown in Fig. 1. Note that $\bar{y}_T$ is directly proportional to c and is monotonically decreasing with increasing $\bar{x}$. The corresponding dimensional transition length is

$$v_T(d) = \frac{\pi\kappa}{8G^*}\sqrt{R\sqrt{d^2+2k\sigma^2}-Rd}$$

For macro-devices this length may not be significant, but for micro-machines (with similar roughness characteristics) the transition length can be quite large in relation to their size.

2.1 Discrete Distribution. In micro-devices, the nominal area of contact and normal loads are typically small, which both lead to small populations of contacting asperities. To investigate the effects a small population size has on (13) and (14), it is now assumed that the summit height distribution takes the discrete form

$$\varphi_N(h) = \sum_{i=1}^{N_t}\hat{\delta}(h_i-h),$$

where $\{h_i\}$ are random samples of a continuous Gaussian distribution with standard deviation $\sigma=1$ and $\hat{\delta}(h)$ is the Dirac delta.[5] The discrete counterpart to the nondimensional normal force (14) calculated from a continuous distribution is

$$\tilde{\mathcal{P}}_N(\bar{x}) = M_{3/2}(\bar{x},\bar{x}) = \sum_{i|x_i\geq\bar{x}}(x_i-\bar{x})^{3/2} \tag{16}$$

and, likewise, the nondimensional tangential force is given by

$$\begin{aligned}\tilde{\mathcal{Q}}_N(\bar{x},\bar{\mathbf{y}}) &= -M_{1/2}(x,x+\bar{y}^2/c^2)\bar{\mathbf{y}} - c(M_1(x)-M_1(x,x+\bar{y}^2/c^2))\frac{\bar{\mathbf{y}}}{\|\bar{\mathbf{y}}\|}\\ &= -\bar{\mathbf{y}}\sum_{i|x_i\geq x+\bar{y}^2/c^2}(x_i-\bar{x})^{1/2} - c\frac{\bar{\mathbf{y}}}{\|\bar{\mathbf{y}}\|}\sum_{i|x_i<x+y^2/c^2}(x_i-\bar{x})\end{aligned} \tag{17}$$

Here the moments defined by (9) become

[4]This is taken as a representative value for c given that mechanical yield strengths are typically two to three orders of magnitude less than the corresponding elastic moduli.

[5]The symbol $\hat{\delta}(h)$ should not to be confused with the symbol δ used to denote the normal displacement of an asperity.

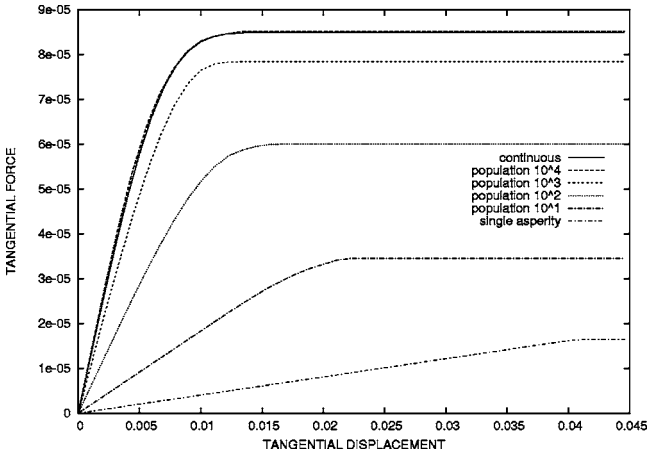


Fig. 2 Tangential force vs. tangential displacement for various population sizes and a fixed normal force

$$M_n(d,l) = \int_l^\infty (h-d)^n \varphi_d(h)\mathrm{d}h = \sum_{i|h_i \geqslant l} (h_i - d)^n \tag{18}$$

where the factor $1/N_t$ would be necessary to normalize the discrete distribution.

Figure 2 plots $\tilde{\mathcal{Q}}_N$ as a function of tangential displacement for a progression of population sizes $[N_t, 10^{-1}N_t, \ldots, 1]$, for $N_t = 10^4$. Each of these discrete populations is composed of the appropriate number of random samples of the continuous distribution $\phi(h)$, i.e., they have approximately the same mean and standard deviation as $\phi(h)$. The nominal pressure is the same for all populations, with the chosen value being given by $N_t\tilde{\mathcal{P}}(\bar{x}=2)$ for the continuous distribution and the magnitude of each $\tilde{\mathcal{Q}}_N$ plotted in Fig. 2 normalized by N_t. The trend shows that the transition length of $\tilde{\mathcal{Q}}_N$ increases as population size decreases, starting with that of the continuous response for a population of 10^4 and approaching that of the single asperity response. This is a direct result of the increasing average normal force on the contacting asperities for smaller populations, and the fact that the transition length is tied to largest contact area, i.e., the one with highest load. Note that only about 2% of the largest population is bearing the contact load but this number of contacting asperities is still greater than in the smaller populations where all the asperities are in contact. Interestingly, for this normal loading state, the total real contact area has a decreasing trend with decreasing population size as evidenced by the steady-state values in Fig. 2. Lastly, note that this numerical experiment is only meant to show trends, since it clearly may take the Hertzian asperity model beyond the limits of small, linear elastic deformations.

2.2 Cyclic Tangential Loading. Up until this point the tangential loading has been monotonic. In order to allow for reversal a history variable, the (tangential) slip, $\mathbf{s}$, is needed in the individual asperity response:

$$\mathbf{q} = \mathbf{q}_{\text{stick}}(\delta, \boldsymbol{v}-\mathbf{s}) = -8G^*\sqrt{R}\delta^{1/2}(\boldsymbol{v}-\mathbf{s}). \tag{19}$$

In the given form, this constitutive response is consistent with the classical theory of plasticity and is illustrated in Fig. 3. Indeed, this response can be cast in a (strain space-based) plasticity framework where the force-displacement relation is given by (19), and the yield criterion is given by the slip condition

$$\kappa a = \kappa\pi R\delta \geqslant \|\mathbf{q}\| = 8G^*\sqrt{R}\delta^{1/2}\|\boldsymbol{v}-\mathbf{s}\|. \tag{20}$$

To complete the model an evolution equation for $\mathbf{s}$ must be specified. In the case of tangential loading in a constant direction with

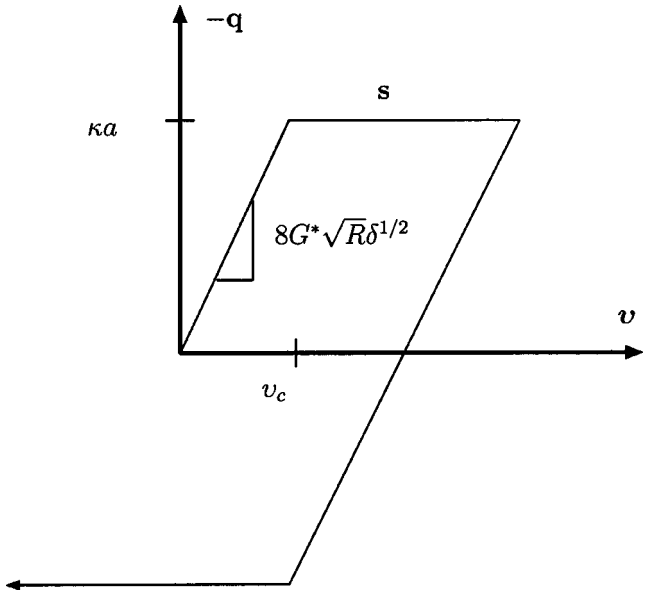


Fig. 3 Tangential force vs. tangential displacement for constant normal displacement

constant compression, $\dot{\mathbf{s}} = \dot{\boldsymbol{v}}$ leads to the behavior shown in Fig. 3.[6] In the general case where the tangential loading is two-dimensional, a flow rule analogous to that of elastic-perfect plasticity is plausible, specifically

$$\dot{\mathbf{s}} = \gamma(\boldsymbol{v}-\mathbf{s}). \tag{21}$$

This nonassociative flow rule uses the normal to the yield surface projected into the tangential slip space and is the strain-space equivalent of the constitutive rule used in [24] to model classical Amontons-Coulomb friction.

If the asperity is in a loading state, i.e., $\delta = \delta_c(\boldsymbol{v}-\mathbf{s})$ and

$$(\boldsymbol{v}-\mathbf{s})\cdot\dot{\boldsymbol{v}} - \left(\frac{\pi\kappa}{8G^*}\right)^2\frac{R}{2}\dot{\delta} \geqslant 0,$$

then consistency requires

$$(\boldsymbol{v}-\mathbf{s})\cdot(\dot{\boldsymbol{v}}-\dot{\mathbf{s}}) = \left(\frac{\pi\kappa}{8G^*}\right)^2\frac{R}{2}\dot{\delta}$$

which leads to a solution for the consistency parameter γ,

$$\gamma = \frac{(\boldsymbol{v}-\mathbf{s})\cdot\dot{\boldsymbol{v}}}{\|\boldsymbol{v}-\mathbf{s}\|^2} - \frac{\dot{\delta}}{2\delta}.$$

The asperity response given in (19) leads to an ensemble response analogous to (11),

$$\begin{aligned}\mathcal{Q}(d,\boldsymbol{v}-\mathbf{s}) &= \int_{d+\delta_S}^\infty \mathbf{q}_{\text{stick}}(h-d,\boldsymbol{v}-\mathbf{s})\varphi(h)\mathrm{d}h \\ &\quad + \int_d^{d+\delta_S} \mathbf{q}_{\text{slip}}(h-d)\varphi(h)\frac{\boldsymbol{v}-\mathbf{s}}{\|\boldsymbol{v}-\mathbf{s}\|}\varphi(h)\mathrm{d}h \\ &= -\int_{d+\delta_S}^\infty 8G^*\sqrt{R}(h-d)^{1/2}(\boldsymbol{v}-\mathbf{s})\varphi(h)\mathrm{d}h \\ &\quad - \int_d^{d+\delta_S} \kappa\pi R(h-d)\frac{\boldsymbol{v}-\mathbf{s}}{\|\boldsymbol{v}-\mathbf{s}\|}\varphi(h)\mathrm{d}h \end{aligned} \tag{22}$$

where $\mathbf{s} = \mathbf{s}(h)$ is a slip distribution and δ_S is the current upper

[6] Since $\boldsymbol{v}$, in fact, represents a *relative* displacement of the contacting surfaces, the material time derivative of $\boldsymbol{v}$ is objective.

limit of the slipping population. Since $\mathbf{q}(\delta, \boldsymbol{v}-\mathbf{s})$ is continuous and $\mathbf{s}$ is such that $\mathbf{q}$ satisfies the slip condition (20),

$$\mathcal{Q}(d, \boldsymbol{v}-\mathbf{s}) = \mathcal{Q}_{\text{stick}}(d, \boldsymbol{v}-\mathbf{s}) = -\int_d^\infty 8G^*\sqrt{R}(h-d)^{1/2}[\boldsymbol{v}-\mathbf{s}(h)]\varphi(h)\mathrm{d}h \quad (23)$$

can be used in place of (22). This form illustrates that only the elastic part of the tangential displacement, i.e., $\boldsymbol{v}-\mathbf{s}$, generates forces, and the product of two distributions, one describing the number of asperities in a certain range of heights and another tracking their slip history, is required to determine the ensemble response. It should also be noted that the ensemble response also fits into a plasticity-like framework[7] with (23) being the force-displacement relation,

$$\kappa\pi R(h-d) \geqslant 8G^*\sqrt{R}(h-d)^{1/2}\|\boldsymbol{v}-\mathbf{s}(h)\| \quad (24)$$

being the yield condition, and the flow rule (21), likewise, being applied to $\mathbf{s}(h)$. The nondimensional version of (23) is

$$\tilde{\mathcal{Q}}(\bar{x}, \bar{\mathbf{y}}-\boldsymbol{\varsigma}) = -\int_{\bar{x}}^\infty (x-\bar{x})^{1/2}[\bar{\mathbf{y}}-\boldsymbol{\varsigma}(x)]\varphi_{\sigma=1}(x)\mathrm{d}x, \quad (25)$$

where $\boldsymbol{\varsigma}=\mathbf{s}\sqrt{\sigma R}$ and, likewise, the condition (24) defining the yield envelope becomes

$$x-\bar{x} \geqslant \frac{1}{c^2}\|\bar{\mathbf{y}}-\boldsymbol{\varsigma}(x)\|^2. \quad (26)$$

For constant approach, and therefore pressure, and in two dimensions, analytical solutions exist for the "elastic" slip distribution, $\bar{\mathbf{y}}-\boldsymbol{\varsigma}(x)$. Here these functions are defined recursively as

$$\bar{y}-\varsigma_i(x) = \begin{cases} 0 & \text{if } x-\bar{x}<0 \\ y_c(x-\bar{x})\text{sign}(\bar{y}-\bar{y}_i^R) & \text{if } 0<x-\bar{x}<x_i^S \\ \bar{y}-\varsigma_{i-1} & \text{else} \end{cases} \quad (27)$$

where $\bar{y}$ is the signed magnitude of $\bar{\mathbf{y}}$. The form of this solution embodies the ideas that an individual asperity will continue to deform elastically with changes in the applied tangential displacement $\bar{y}$ up until the elastic slip state reaches the yield surface and that the shorter asperities with smaller contact areas slip first. The index i refers to the number of past reversals of the time derivative of $\bar{y}$, with the corresponding values of $\bar{y}$ being $\{\bar{y}_1^R, \bar{y}_2^R, \ldots, \bar{y}_i^R\}$. In addition, the upper limit of the slip region, x_i^S, is given relative to $\bar{x}$, the lower limit. The location of x_i^S is given by the intersection of the slip surface $y=y_c(x-\bar{x})$ and the elastic slip distribution at the previous reversal $\bar{y}-\varsigma_{i-1}$ shifted by $\bar{y}-\bar{y}_i^R$. The translation of the solution for $\bar{y}-\varsigma$ at the last reversal by the increment in $\bar{y}$ since the last reversal can be alternately interpreted as the simple translation of ς_{i-1} by $\bar{y}$ since

$$(\bar{y}_i^R-\varsigma_{i-1})+\bar{y}-\bar{y}_i^R = \bar{y}-\varsigma_{i-1}.$$

This geometric problem is illustrated in Fig. 4 where the intersection of each branch of slip distribution with the opposite sign of the current loading increment with the yield surface is shown. (Related concepts are found in [12] for rough surfaces and [25] for a punch contacting a foundation, albeit without the geometric interpretation.) Each of these branches is incorporated into the slip distribution $\bar{\mathbf{y}}-\boldsymbol{\varsigma}(x)$ by previous loading and translated by the tangential displacement history. The appropriate intersection point in this case is clearly at

$$\left(x^S = x_c\left(\tfrac{1}{2}(\bar{y}-\bar{y}_{N-1}^R)\right), \tfrac{1}{2}(\bar{y}-\bar{y}_{N-1}^R)\right)$$

and in general x_i^S is given by

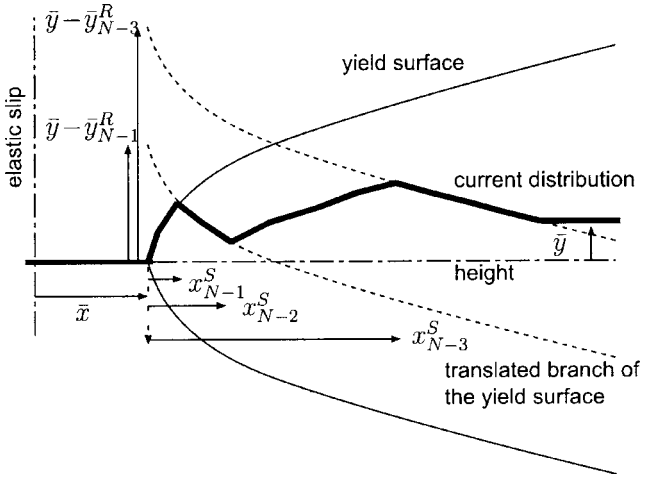


Fig. 4 Intersection of the yield surface and the negative slip branches

$$x_i^S = x_c\left(\tfrac{1}{2}(\bar{y}_i^R-\bar{y}_j^R)\right),$$

where $j=\max_j\{i-1, i-3, \ldots, \text{s.t. } x_c(\frac{1}{2}(\bar{y}_i^R-\bar{y}_j^R)) \leqslant x_j^S\}$ or $x_i^S = x_c(\bar{y}_i^R)$ if $x_c(\frac{1}{2}(\bar{y}_i^R-\bar{y}_j^R)) > x_1^S$ The first intersection is always given by $x_1^S=x_c(\bar{y}_1^R)$ and the first two solutions of the recursion are $\varsigma_0(x)=0$ and

$$\bar{y}-\varsigma_1(x) = \begin{cases} 0 & \text{if } x-\bar{x}<0 \\ y_c(x-\bar{x})\text{sign}\,\bar{y}_1^R & \text{if } 0<x-\bar{x}<x_1^S \\ \bar{y}-\varsigma_0 & \text{else.} \end{cases} \quad (28)$$

Figure 5 shows the actual distributions for a sequence of reversals. For the first loading shown ($t=40$) the solution for $\bar{\mathbf{y}}-\boldsymbol{\varsigma}(x)$ is given by (28) (i.e., the intersection of the horizontal line at $\bar{y}-\varsigma_0$ and the yield surface), for the second loading ($t=100$) the solution is the intersection of the previous one (translated by the current loading) and the yield surface, and so forth. Figure 6 shows the corresponding force response for this entire loading history. Note that [12] develops the solutions (27) for the first few reversals, which result in hysteresis loops similar to those shown in Fig. 6. The qualitative behavior displayed in Fig. 6 also corresponds with the linear displacement experiments of Courtney-Pratt and Eisner [1] and the angular displacement experiments of Hsieh and Pan [15]. Furthermore, the micro-mechanics of the proposed paradigm correlates well with qualitative observations in [15] and provides a rational explanation of the measured phenomenology. For instance, in Fig. 5, the abrupt disappearance of the elastic slip branch in the interval (x_{N-2}, x_{N-3}) while loading from $t=140$ to $t=170$ corresponds to the most unexpected of the five empirical rules stated in [15], namely that reversal points can be "wiped out from memory" by loading past the completion of a hysteresis loop. Essentially, what is happening in this event is the overloading of a portion of the asperity population that is currently sticking but has stored elastic slips from a previous (partial) slip event.

Lastly, the work dissipated for a tangential displacement cycle (in a fixed direction and at a fixed approach $\bar{x}$) is calculated for the nondimensional version (25) of the tangential reaction force $\tilde{\mathcal{Q}}$ as

$$W_{\text{cycle}}(\bar{x}, \bar{y}_{\max}) = \oint_{-\bar{y}_{\max}}^{\bar{y}_{\max}} \tilde{\mathcal{Q}}(\bar{x}, \bar{y}-\varsigma)\mathrm{d}\bar{y} = 2\int_{-\bar{y}_{\max}}^{\bar{y}_{\max}} \tilde{\mathcal{Q}}(\bar{x},\ \bar{y}-\varsigma))\mathrm{d}\bar{y},$$

where ς has the form (28) at the limits of the cycle (not $\varsigma=\varsigma_0=0$). The results for $\bar{x}=1$ and a range of $\bar{y}_{\max} \in (0, 0.04)$ are shown in Fig. 7. The work corresponding to the area enclosed by the tangential force cycle is relatively small for cycle magnitudes below the transition length $\bar{y}_T(\bar{x})$ since very little of the contacting population slips. On the other hand, there is essentially a linear

[7]This is similar to crystal plasticity, where a distribution of states contributes to the stress response at a homogenized material "point."

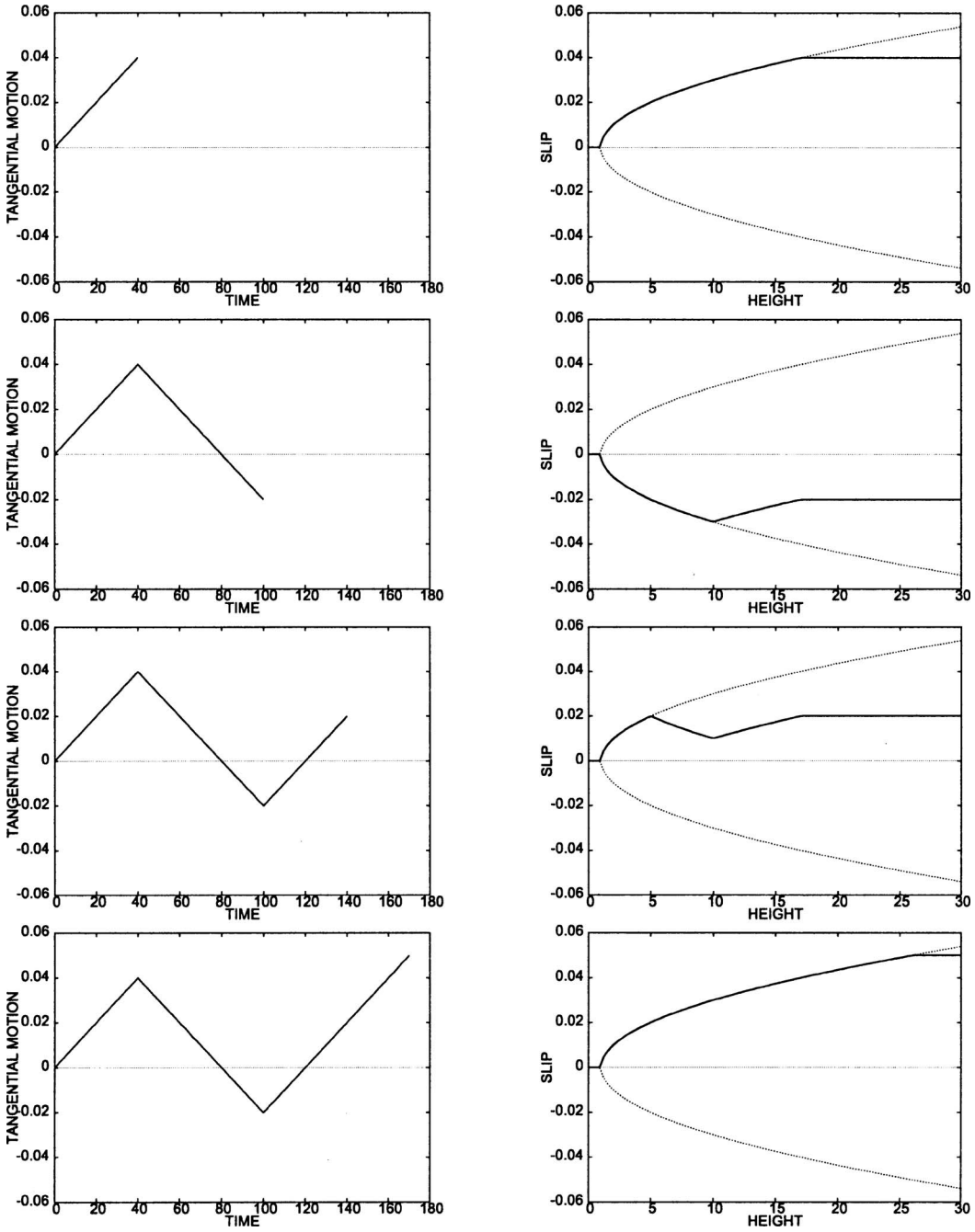


Fig. 5 Prescribed tangential displacement history and slip distribution at t=40,100,140,170

increase in dissipation per cycle after the maximum tangential displacement of the cycle exceeds the transition length, which is consistent with Amontons-Coulomb behavior.

2.3 General Loading. For general loading and the three-dimensional case, a numerical return map is necessary to integrate the constitutive equations. The proposed algorithm for maintaining consistency between the function $\mathbf{s}$ and the feasible region (defined by the slip surface (26)) is given in Table 1, where $\Delta\bar{\mathbf{y}}$ is the discrete increment in applied slip. This scheme is very similar to the predictor-corrector return-mapping algorithms from classical plasticity (see, for example, [26]), with a few important differences. First of all, a function

$$\mathbf{y} = \mathbf{y}(x) = \bar{\mathbf{y}} - \mathbf{s}(x)$$

is being manipulated, as opposed to a pointwise variable like plastic strain. Here, a piecewise linear function

$$\mathbf{y}(x) = \begin{cases} \mathbf{0} & \text{if } x < \bar{x} \\ \bar{\mathbf{y}} & \text{if } x > x_1^S \\ \Psi_i(x)\mathbf{y}_i & \text{else} \end{cases}$$

is used to represent $\mathbf{y}=\bar{\mathbf{y}}-\mathbf{s}$, where Ψ_i are the common linear interpolatory functions from finite elements (see for example [27]).[8] Secondly, since $\mathbf{y}(x)$ is assumed to represent a continuous function, its selective projection onto the convex slip surface (26) requires some care. For instance, the appropriate return path for a

[8]It is possible for the size of $\{\mathbf{y}_i\}$ to grow with the number of time steps; however, the number of points necessary to approximate $\mathbf{y}$ can be limited by noting that the region $x \in (\bar{x}, x_1^S)$ is of finite extent. Additionally, the highest portion of the asperity population may not contribute significantly to $y(x)\phi(x)$ and therefore to $\mathcal{Q}$ if the value of the population distribution $\phi(x)$ decreases rapidly as the independent variable increases (which is certainly the case for the Gaussian).

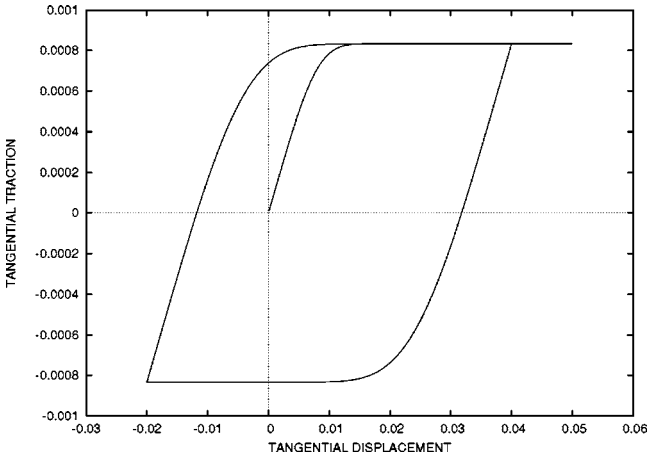


Fig. 6 Tangential force vs. tangential displacement for reverse slip for $\bar{x}=1$

Table 1 The return mapping algorithm for slip distribution

- Assume stick and integrate, $\mathbf{y}+=\Delta\bar{\mathbf{y}}$.
- Check yield condition of the first node:
  if $x_c(\mathbf{y}_1)>x_1-\bar{x}$, then prepend $(\bar{x}+x_c(\mathbf{y}_1),\mathbf{y}_1)$ to set.
- Return via:
  For $i=1,\ldots,N$, where N is the current number of nodes in the set.
  If $x_c(\mathbf{y}_i)>x_i-\bar{x}$, then
  1. If $x_c(y_{i-1})<x_{i-1}-\bar{x}$, then intersect surface using line segment to node $i-1$, i.e., solve:
     $$(\mathbf{y}_i-\mathbf{y}_{i-1})(x-x_{i-1})=(x_i-x_{i-1})(\mathbf{y}-\mathbf{y}_{i-1}),$$
     $$x-\bar{x}=x_c(\mathbf{y})$$
  2. Else project normally to surface by solving:
     $$(x-x_i)\partial_{(\mathbf{y})_\alpha}(x_c)=-(\mathbf{y}-\mathbf{y}_i)_\alpha\partial_x(x_c),\ \text{for}\ \alpha=1,\ldots,2$$
     $$x-\bar{x}=x_c(\mathbf{y})$$
     replace node i with new value $(x,\mathbf{y})$.
- Append apex of slip surface $(\bar{x},\mathbf{0})$ to the set.
- Drop misordered elements of the set:
  For $i=1,\ldots,N$, if $x_i+tol_x\geqslant x_{i+1}$, then drop element i from set (where tol_x is a small positive constant).

node may not be normal to the surface and may be multiply defined, as is the case for a node at a cusp or a kink in the function. In addition, the proposed algorithm includes a provision for the nodes representing $\mathbf{y}(x)$ to become misordered upon return to the feasible region. Note that in Table 1 the current values of the applied normal and tangential (far-field) displacement are assumed to be known (i.e., the ensemble is kinematically controlled) so the current location of the yield surface is also known. Figure 8 shows the return map for a general loading case were $\bar{x}$ and $\bar{\mathbf{y}}$ are both changing, with the predictor state being simply the previous state incremented by the new tangential load step $\Delta\bar{\mathbf{y}}$. Note the added and dropped nodes as the updated representation distribution $\mathbf{y}(x)$ (demarcated by gray nodes) is returned to the feasible part of slip space. Once the elastic slip distribution $\mathbf{y}=\bar{\mathbf{y}}-\mathbf{s}$ is obtained, Eq. (25) is used, together with the trapezoidal rule as a quadrature, to compute the resulting forces.

Three illustrative simulations were performed. The first compares the numerical solution generated by the algorithm in Table 1 with the two-dimensional analytical solution given by Eqs. (25) and (27). Employing the same loading as used in Fig. 5, Fig. 9 shows the numerical solution of the tangential force response superimposed on the analytical response shown in Fig. 6. Overall there is good correspondence, with the worst correlation being at the tangential load reversals. Better accuracy could be achieved with a smaller loading step size.

The second problem is used to demonstrate that the algorithm is capable of solving for the force response in the case of nonconstant normal separation. Here, the approach starts at $\bar{x}=1$, is held constant, then is linearly increased to $\bar{x}=2$, is held constant again, and then is linearly increased to its original value. During this

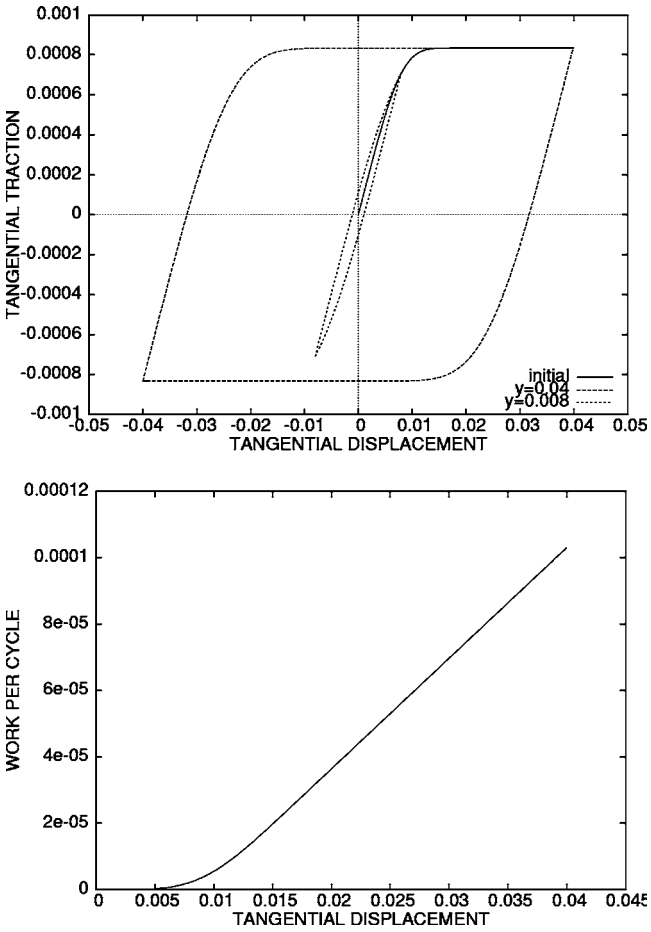


Fig. 7 Tangential force and work per cycle for cyclic loading in a fixed tangential direction

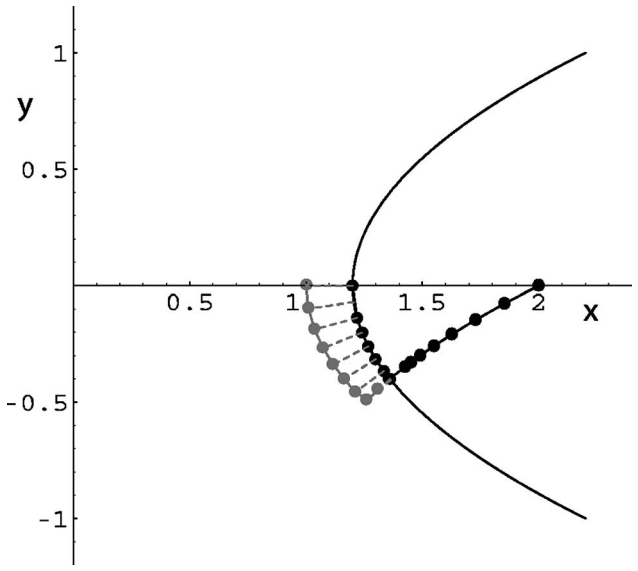


Fig. 8 Return map for a specific step

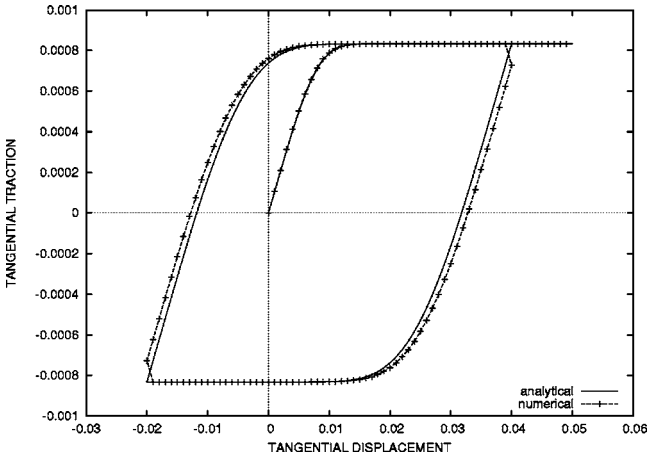


Fig. 9 Comparison of tangential force vs. normal force reverse slip for $\bar{x}=1$

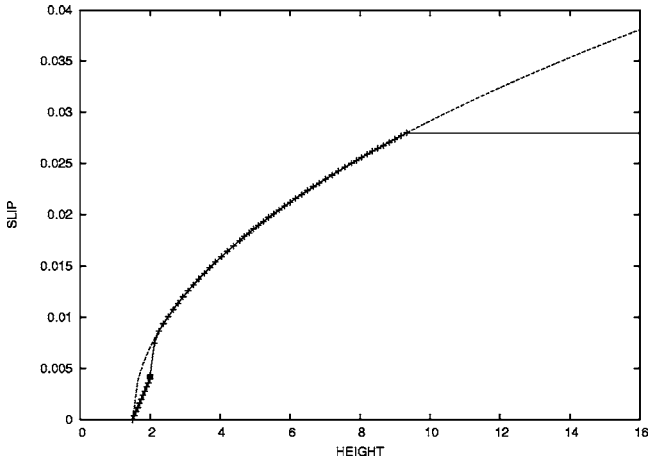


Fig. 11 Elastic slip distribution at $\bar{y}=1.5$ and $\bar{y}=0.028$

whole sequence, the tangential displacement is being steadily increased in a constant direction. The normal and tangential responses to this loading are shown in Fig. 10. The normal force follows the trend of the normal approach, as expected. The tangential force follows the same trend as seen in Fig. 1 until the approach starts to change, which happens before the tangential response reaches steady state. As the normal force decreases, the yield envelope is translating in the positive x direction (refer to Fig. 8) and the height threshold for contact is increasing commensurately, so that all previously slipping asperities are still slipping (or have left the contacting subpopulation) and increasingly higher asperities slip. So at the new value of $\bar{x}=2$ a steady state of slip has already been attained. Now when the approach is decreased to its original value, the yield envelope translates in the negative x direction, allowing newly contacting asperities to start to slip from a state of zero "elastic" slip, i.e., their previous slip history has been erased. A typical elastic slip distribution for this stage is shown in Fig. 11. This phenomenon accounts for the fact that the phase with decreasing approach appears to create a slip state that is slightly different from the reverse of the increasing approach phase. When the final, constant approach stage is reached, the yield envelope is where it started and the slip state of the population is similar to that shown in Fig. 11. Consequently, the remaining asperities near the apex of the yield surface finally slip after a small transition and the steady state appropriate for the normal loading is reached.

The third simulation involved simulating the tangential response to circular (tangential) displacement and constant normal load ($\bar{x}=1$). For the first instance, the radius of the circle is chosen to be 0.004, much less than the transition length of the model for a straight line tangential displacement (approximately equal to $c=0.01$). Figure 12 shows the evolution of the elastic slip distribution projected onto the $x-y_1$ plane. The stick-slip transition point of the elastic slip distribution follows a roughly helical path around the parabolic yield envelope till it reaches a steady state at about $x=1.63$. A parametric plot of the tangential force in Fig. 13 shows that nearly a half cycle of displacement is needed for the transition behavior to attain a limit cycle. Furthermore, the limiting circle is not centered about zero force, indeed the center is at $(-0.000021,-0.00029)$. This is due to the fact that at this small displacement a significant portion of the contacting population never slips and so the total force response has an elastic component that is centered at the origin of the displacement cycle. In actuality, if no significant portion of the contact populous slipped, the tangential force response would follow the prescribed tangential displacement which is centered on the negative vertical axis and passes through the origin. At the other extreme, if the radius of the displacement circle is chosen to be 0.04, which is greater than the straight line transition length, a limit cycle that is compatible with classical Coulomb behavior is displayed, as in Fig. 14. After an initial transition stage where the stick-slip transition

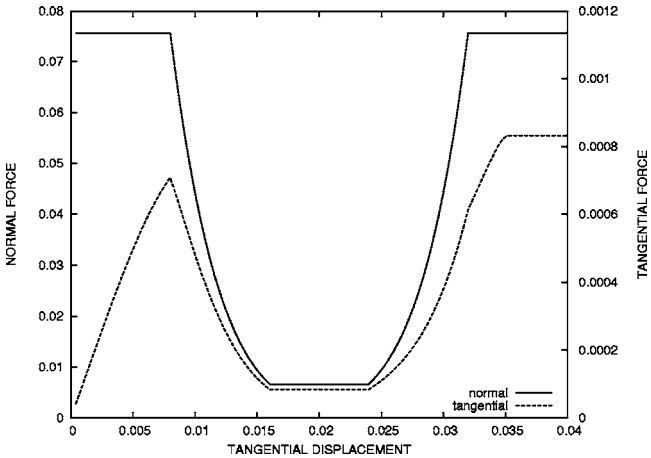


Fig. 10 Normal and tangential force for varying approach and increasing applied slip

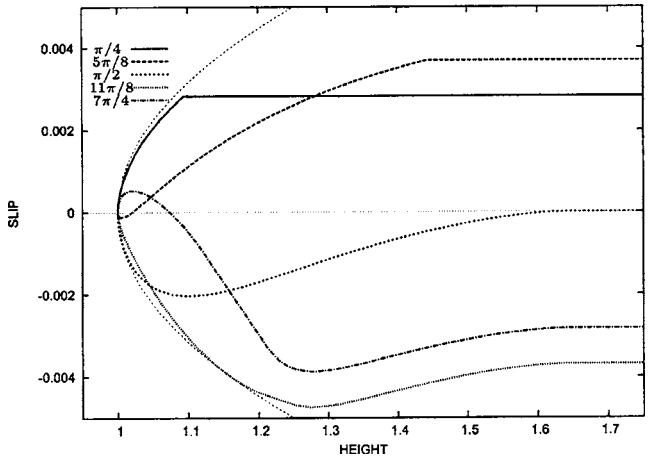


Fig. 12 Evolution of slip distribution for circular displacement of radius 0.004

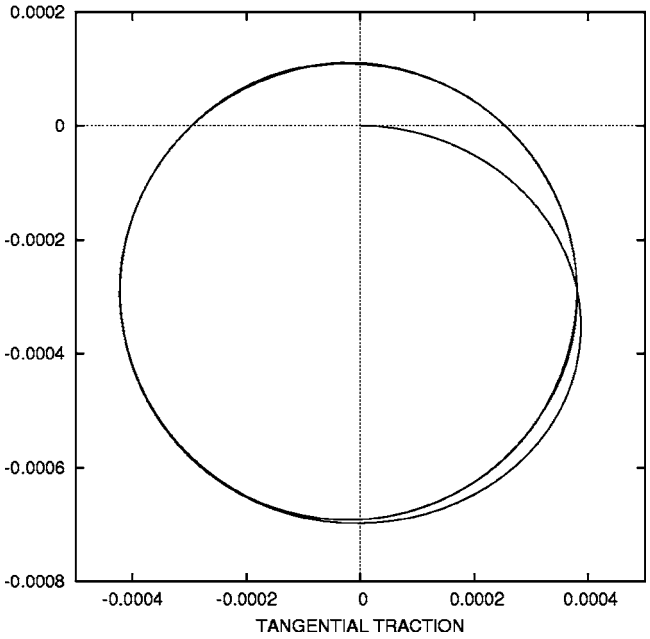


Fig. 13 Tangential force history for circular displacement of radius 0.004

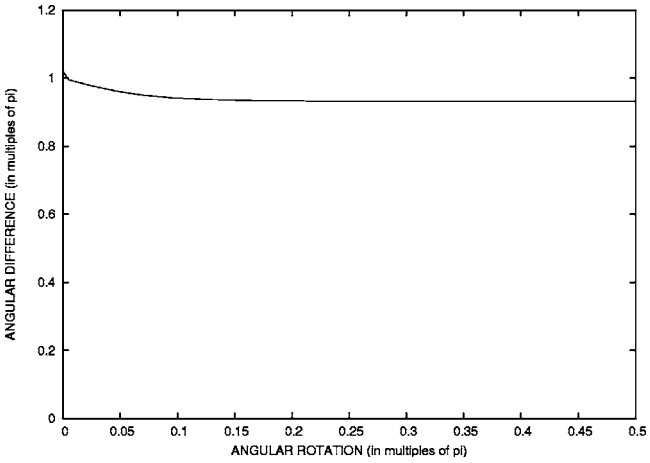


Fig. 15 Angular difference between displacement and force for circular displacement of radius 0.04

point of the population is changing, the force response is equally distributed about zero force. Figure 15 shows that even for this example, there is a steady-state deviation from direct opposition of the ensemble tangential force and the direction of tangential displacement. This deviation is inversely proportional to the radius of the displacement circle and proportional to the radius of a typical asperity's yield surface (which is dependent on c and the normal loading). A plausible explanation is illustrated in Fig. 16, which shows when the radius of the asperity's yield surface is comparable to the radius of curvature of the displacement path there will be a finite discrepancy between the elastic slip and the loading direction. (A material point with an elastic-perfect plastic constitution would display similar behavior under analogous loading.) This discrepancy disappears as the path's curvature becomes large relative to the radii of the population of contacting asperities that are significantly contributing to the total force and, in this case, Coulomb behavior is recovered.

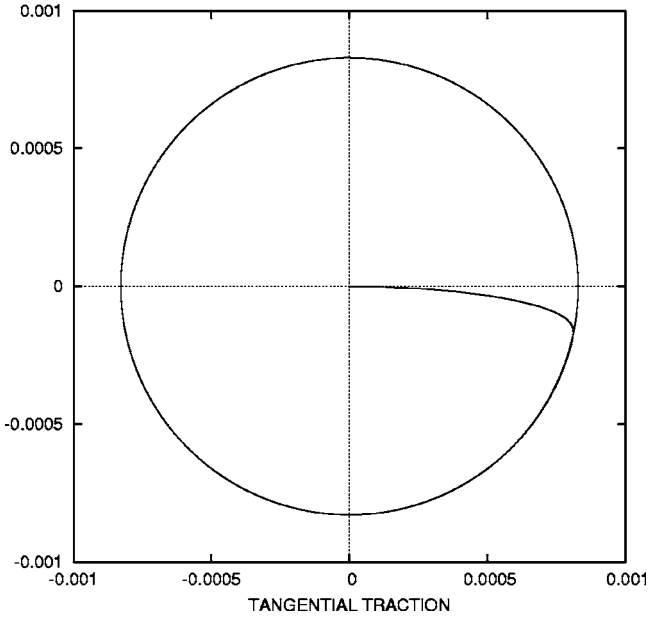


Fig. 14 Tangential force history for circular displacement of radius 0.04

3 Discussion

The treatment of the individual asperity response in this work is admittedly simple. It should be possible to apply the methodology developed herein to a more sophisticated and representative asperity response, such as the analytical solution described in [28,29] or an empirical response derived from an atomic force microscope, since the details of the traction fields need not be known.

Aside from the details of the asperity tractions, the first qualitative feature that needs to be included to represent the behavior of micro-devices is normal adhesion. This addition raises the basic question of whether the same mechanism is dominating both normal adhesion and the tangential adherence assumed in the model of friction described in this work. There are multiple sources of adhesion in small devices. For silicon-based MEMS, there seems to be three primary sources of adhesion: capillary, electrostatic, and van der Waals (as discussed in [30]). In polymer-coated

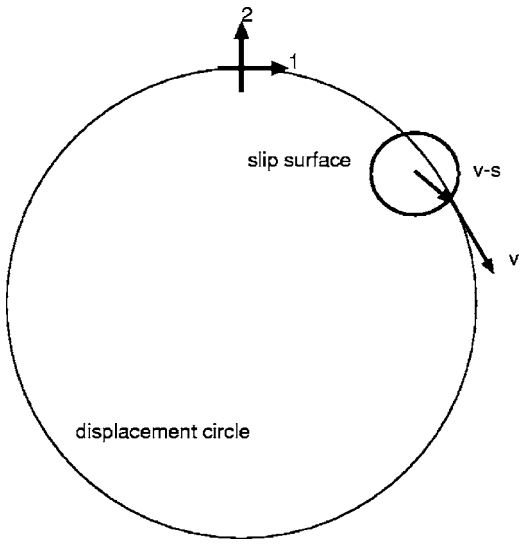


Fig. 16 The circular path and the yield surface for a typical asperity

(SAMS) surfaces, the van der Waals forces tend to dominate. In metallic LIGA-based devices, the metallic junction formation seen in larger metallic devices can come into play.

In the Derjaguin-Muller-Toporov (DMT) [31] limit, the adhesive force is weak relative to the elastic forces, and consequently does not affect the contact radius, and the extension of the proposed model is straightforward. The individual asperity normal response becomes

$$p(\delta) = p_{\text{Hertz}}(\delta) - 2\pi R w$$

where w is the work of adhesion per unit area on an asperity. For the surface, nondimensionalization similar to that used in Eq. (14) leads to

$$\tilde{\mathcal{P}}(\bar{x}) = M_{3/2}(\bar{x}) - \frac{2\pi}{\theta} M_0(\bar{x})$$

where the Tabor parameter $\theta = \frac{4}{3} E^* \sigma^{3/2} R^{1/2} / wR$ is a ratio of elastic to adhesive force (for a more detailed development see [11]). Since the contact area, Eq. (2), is unaffected by DMT-type adhesion, the tangential response (13) would be unchanged. Only the effective Coulomb coefficient would change, going to infinity at zero applied normal load.

Another limiting case is given by the Johnson-Kendall-Roberts (JKR) solution [32], where the adhesive force is weak relative to the elastic force. The effect of this type of adhesion on a GW-type contact model was explored by Fuller and Tabor in [33]. The JKR solution alters the relationship between compression and contact area, Eq. (2), and is complicated by an implicit definition. Lastly, if the range of adhesion is much greater than the surface roughness, the adhesion may be directly dependent on the separation and, unlike the asperity-based solutions, the nominal area (see the solutions for parallel plates in [34], for example).

4 Conclusion

In summary, despite the basic shortcomings of GW models (e.g., identifying "asperities" (see, e.g., [23]), the noninteraction of individual contacts, the lack of accounting for interlocking and alignment of asperities) and the ones inherent in this work, the qualitative aspects of the proposed model correspond well with experimental observations. The proposed model and methodology have also predicted phenomenology, i.e., the response to normal loading and unloading as surfaces are displaced tangentially and the response to small radius circular slip, that has yet to be investigated experimentally. With this being said, it is entirely possible that in certain cases PSTD behavior may be dominated by some other mechanism than asperity adherence. The extension of this work to physically interlocking asperities with adhesion is left for future work.

Acknowledgment

The author would like to acknowledge several valuable and insightful discussions of this work with Arthur Brown and Richard Regueiro. This development was funded by the Laboratory Directed Research and Development program at Sandia National Laboratories. This support is gratefully acknowledged. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract No. DE-ACO4-94AL85000.

References

[1] Courtney-Pratt, J. S., and Eisner, E., 1957, "The Effect of a Tangential Force on the Contact Metallic Bodies," Proc. R. Soc. London, Ser. A, **238**, pp. 529–550.

[2] Burdekin, M., Back, N., and Cowley, A., 1978, "Experimental Study of Normal and Shear Characteristics of Machined Surfaces in Contact," J. Mech. Eng. Sci., **20**, pp. 129–132.

[3] Menq, C.-H., Griffin, J. H., and Bielak, J., 1986, "The Influence of Microslip on Vibratory Response, Part II: A Comparison with Experimental Results," J. Sound Vib., **107**, pp. 295–307.

[4] Berthoud, P., and Baumberger, T., 1998, "Shear Stiffness of a Solid-Solid Multicontact Interface," Proc. R. Soc. London, Ser. A, **454**, pp. 1615–1634.

[5] Ni, J., and Zhu, Z. Q., 2001, "Experimental Study of Tangential Micro Deflection of Interface of Machined Surfaces," ASME J. Manuf. Sci. Eng., **123**, pp. 365–367.

[6] Filippi, S., Akay, A., and Gola, M. M., 2004, "Measurement of Tangential Contact Hysteresis During Microslip," ASME J. Tribol., **126**, pp. 482–489.

[7] Corwin, A. D., Street, M. D., Carpick, R. W., Ashurst, W. R., and de Boer, M. P., 2004, "Pre-Sliding Tangential Deflections Can Govern Friction in MEMS Devices," *Proceedings of the ASME/STLE Int'l Joint Tribology Conference*, Oct. 24-27, Long Beach, CA.

[8] Greenwood, J. A., and Williamson, J. B. P., 1966, "Contact of Nominally Flat Surfaces," Proc. R. Soc. London, Ser. A, **295**, pp. 300–319.

[9] Phan-Thien, N., 1981, "On an Elastic Theory of Friction," ASME J. Appl. Mech., **48**, pp. 438–439.

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

[11] Tas, N. R., Gui, C., and Elwenspoek, M., 2003, "Static Friction in Elastic Adhesion Contacts in MEMS," J. Adhes. Sci. Technol., **17**, pp. 547–561.

[12] Fujimoto, T., Kagami, J., Kawaguchi, T., and Hatazawa, T., 1995, "Tangential Microdisplacement and Stiffness in Contact," *Proceedings of the International Tribology Conference*, Japanese Society of Tribologists, Han Lim, ed., Won Pub. Co., Yokohama, Japan.

[13] Fujimoto, T., Kagami, J., Kawaguchi, T., and Hatazawa, T., 2000, "Micro-Displacement Characteristics Under Tangential Force," Wear, **241**, pp. 136–142.

[14] Olofsson, U., and Hagman, L., 1997, "A Model for Micro-Slip Between Flat Surfaces Based on Deformation of Ellipsoidal Elastic Bodies," Tribol. Int., **30**, pp. 599–603.

[15] Hsieh, C., and Pan, Y.-C., 2000, "Dynamic Behavior and Modelling of the Pre-Sliding Static Friction," Wear, **242**, pp. 1–17.

[16] Walsh, J. B., 2003, "A Theoretical Analysis of Sliding of Rough Surfaces," J. Geophys. Res., **108**, pp. ESE6-1–15.

[17] Bureau, L., Caroli, C., and Baumberger, T., 2003, "Elasticity and the Onset of Frictional Dissipation," Proc. R. Soc. London, Ser. A, **459**, pp. 2787–2805.

[18] Mindlin, R. D., and Deresiewicz, H., 1953, "Elastic Spheres in Contact Under Varying Oblique Forces," ASME J. Appl. Mech., **20**, pp. 327–344.

[19] Björklund, S., 1997, "A Random Model for Micro-Slip Between Nominally Flat Surfaces," ASME J. Tribol., **119**, pp. 726–732.

[20] Bowden, F. P., and Tabor, D., 1954, *The Friction and Lubrication of Solids*, rev. ed., Clarendon Press, Oxford.

[21] Greenwood, J. A., and Tripp, J. H., 1971, "The Contact of Two Nominally Flat Rough Surfaces," Proc. Inst. Mech. Eng., **185**, pp. 625–633.

[22] Francis, H. A., 1977, "Application of Spherical Indentation Mechanics to Reversible and Irreversible Contact Between Rough Surfaces," Wear, **45**, pp. 221–269.

[23] Greenwood, J. A., and Wu, J. J., 2001, "Surface Roughness and Contact: An Apology," Meccanica, **36**, pp. 617–630.

[24] Duvaut, G., and Lions, J.-L., 1976, *Inequalities in Mechanics and Physics* (Vol. 219, *Grundlehren der mathematischen Wissenschaften*), Springer-Verlag, Berlin.

[25] Jäger, J., 1998, "A New Principle in Contact Mechanics," ASME J. Tribol., **120**, pp. 677–184.

[26] Simo, J. C., and Hughes, T. J. R., 1998, *Computational Inelasticity*, Springer, New York.

[27] Hughes, T. J. R., 2000, *The Finite Element Method Linear Static and Dynamic Finite Element Analysis*, Dover, Mineola, NY.

[28] Hurtado, J. A., and Kim, K.-S., 1999, "Scale Effects in Friction of Single Asperity Contacts. I. From Concurrent Slip to Single-Dislocation-Assisted Slip," Proc. R. Soc. London, Ser. A, **455**, pp. 3363–3384.

[29] Hurtado, J. A., and Kim, K.-S., 1999, "Scale Effects in Friction of Single Asperity Contacts. II. Multiple-Dislocation-Cooperated Slip," Proc. R. Soc. London, Ser. A, **455**, pp. 3385–3400.

[30] Komvopoulos, K., and Yan, W., 1997, "A Fractal Analysis of Stiction in Microelectromechanical Systems," ASME J. Tribol., **119**, pp. 391–400.

[31] Derjaguin, B. V., Muller, V. M., and Toporov, Y. P., 1975, "Effect of Contact Deformations on the Adhesion of Particles," J. Colloid Interface Sci., **53**, pp. 314–326.

[32] Johnson, K. L., Kendall, K., and Roberts, A. D., 1971, "Surface Energy and the Contact of Elastic Solids," Proc. R. Soc. London, Ser. A, **324**, pp. 301–313.

[33] Fuller, K. N. G., and Tabor, D., 1975, "The Effect of Surface Roughness on the Adhesion of Elastic Solids," Proc. R. Soc. London, Ser. A, **345**, pp. 327–342.

[34] Israelachvili, J. N., 1992, *Intermolecular and Surface Forces*, 2nd ed., Academic Press, San Diego.

Three-Dimensional Sharp Corner Displacement Functions for Bodies of Revolution

C. S. Huang[1]
Department of Civil Engineering,
National Chiao Tung University,
1001 Ta-Hsueh Road,
Hsinchu, Taiwan, ROC
e-mail: cshuang@mail.nctu.edu.tw

A. W. Leissa
Department of Mechanical Engineering,
Colorado State University,
Fort Collins, CO 80523

Sharp corner displacement functions have been well used in the past to accelerate the numerical solutions of two-dimensional free vibration problems, such as plates, to obtain accurate frequencies and mode shapes. The present analysis derives such functions for three-dimensional (3D) bodies of revolution where a sharp boundary discontinuity is present (e.g., a stepped shaft, or a circumferential V notch), undergoing arbitrary modes of deformation. The 3D equations of equilibrium in terms of displacement components, expressed in cylindrical coordinates, are transformed to a new coordinate system having its origin at the vertex of the corner. An asymptotic analysis in the vicinity of the sharp corner reduces the equations to a set of coupled, ordinary differential equations with variable coefficients. By a suitable transformation of variables the equations are simplified to a set of equations with constant coefficients. These are solved, the boundary conditions along the intersecting corner faces are applied, and the resulting eigenvalue problems are solved for the characteristic equations and corner functions.
[DOI: 10.1115/1.2178358]

Introduction

Williams [1–3] showed a typical procedure to determine singular corner functions for the two-dimensional (2D) problems of plane elasticity and classical plate bending theory. These analyses have been used for a half century to determine stresses in the vicinity of sharp corners. The corner functions are either Airy stress functions (plane elasticity) or transverse displacement functions (plate bending) which are exact solutions of the partial differential equations of equilibrium. Satisfying the boundary conditions along the two radial edges which form the sharp corner results in an eigenvalue problem. Determining the roots (eigenvalues) of the characteristic equation and substituting them back into the boundary conditions yields the corner functions (eigenfunctions). Their second derivatives are the stresses. Extending Williams' works, numerous researchers used different solution schemes to determine the characteristic equations for a thin wedge consisted of two materials [4–7] or for three-dimensional elastic problems [8–10]. Nevertheless, no corner functions were explicitly provided in these works.

Subsequently, the corner functions themselves were used for plate vibration problems which are solved by the well-known Ritz [11,12] method. To a series of smooth algebraic polynomials for the transverse displacement (w) is added a series of corner functions. The latter accelerate the convergence of the solution for the desired free vibration frequencies and mode shapes because they represent the behavior well in the vicinity of a sharp corner. Because of the singularities there, the algebraic polynomials do not. This approach has been used to obtain accurate (i.e., almost exact) frequencies and mode shapes for sectorial plates [13], circular plates with V notches or sharp radial cracks [14], cantilevered skewed plates [15], and rhombic plates [16]. Three sets of corner functions (transverse displacement and two bending rotations) were also derived for 2D Midlin plate theory [17] and used with algebraic polynomials to analyze the vibrations of thick, cantilevered skewed plates [18]. It was found for many of the plate configurations, thin and thick, that the use of corner functions to supplement the algebraic polynomials greatly accelerated the convergence of solutions. In some cases, without them, accurate frequencies could not be reasonably achieved.

Corner functions were used not only in the Ritz method but also in other numerical approaches. In a finite element approach, Yosibash and Schiff [19] developed a singular superelement by using corner functions for plane elasticity and evaluated the stress intensity factors for a V-notched plate under different in-plane loading. The singular superelement overcomes the difficulties in accurately determining the stresses in the neighborhood of the tip of the V notch by using a traditional finite element approach. Corner functions or parts of corner functions were also used in the mesh-free Galerkin method [20,21] and a partition of unity method [22] to approximate crack tip displacement field and to determine intensity factors.

In recent years, because of the increase in computer speeds and storage capability, it has been possible to obtain accurate solutions for three-dimensional (3D) problems, especially for bodies of revolution, using algebraic polynomials for the three displacement components. For example, accurate frequencies have been achieved for cylinders [23], hollow cones [24], and spheres [25,26]. Frequencies for a fixed-free cylinder were found to converge much slower than for a free-free one [23], due to the stress singularities at the fixed end, and the lack of corner functions. For such shapes there were no abrupt changes in the boundary shapes. However, shapes with abrupt changes (e.g., a stepped circular shaft or a cone segment bonded to a circular cylinder (Fig. 1)) cause stress singularities, and adding suitable corner functions to the analysis can be essential. The purpose of this paper is to derive these functions.

One analysis of the stresses in the vicinity of boundary discontinuities was summarized by Zak [27] in a Brief Note here four decades ago, for the special case of axisymmetric loading. His analysis used a stress function presented by Love [28] for axisymmetric problems.

The present work approaches the 3D elasticity problem from the standpoint of displacements, which will ultimately be the corner functions sought. Beginning with the 3D equations of equilib-

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 26, 2004; final manuscript received April 26, 2005. Review conducted by H. D. Espinosa. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

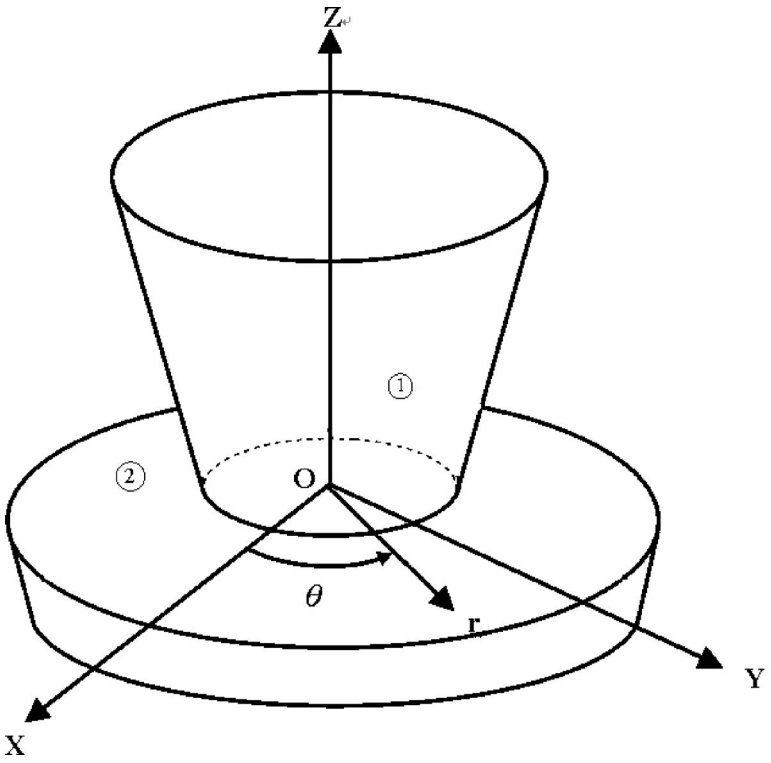


Fig. 1 Body of revolution with a sharp corner boundary discontinuity

rium for elastic bodies of revolution, capable of arbitrary displacements (including axisymmetric ones as a special case), expressed in cylindrical coordinates, they are transformed to a new axis system having its origin at the sharp corner. An asymptotic analysis in the vicinity of the sharp corner reduces the equations to a set of coupled, ordinary differential equations with variable coefficients. By a suitable transformation of variables the equations are simplified to a set of equations with constant coefficients. These are solved, the boundary conditions along the intersecting corner faces are applied, and the resulting eigenvalue problems are solved for the characteristic equations and corner functions.

Equilibrium Equations for a Sharp Corner

The 3D equations of equilibrium, expressed in terms of cylindrical coordinates (r,θ,z), are (cf. [29])

$$\sigma_{r,r}+\frac{\tau_{r\theta,\theta}}{r}+\tau_{rz,z}+\frac{\sigma_r-\sigma_\theta}{r}=0 \tag{1a}$$

$$\tau_{r\theta,r}+\frac{\sigma_{\theta,\theta}}{r}+\tau_{\theta z,z}+\frac{2\tau_{r\theta}}{r}=0 \tag{1b}$$

$$\tau_{rz,r}+\frac{\tau_{\theta z,\theta}}{r}+\sigma_{z,z}+\frac{\tau_{rz}}{r}=0 \tag{1c}$$

where the σ_i and τ_{ij} are normal and shear stresses, respectively, and the subscript "$,\beta$" denotes the differential with respect to the independent variable β. To express them, instead, in terms of displacement components, one uses the stress-strain equations for an isotropic material

$$\sigma_i=\bar{\lambda}e+2G\varepsilon_i,\quad \tau_{ij}=G\gamma_{ij} \tag{2}$$

where the ε_i and γ_{ij} are normal and shear strains, respectively, and

$$e=\varepsilon_r+\varepsilon_\theta+\varepsilon_z,\quad G=\frac{E}{2(1+\upsilon)},\quad \bar{\lambda}=\frac{2\upsilon G}{1-2\upsilon} \tag{3}$$

where $\bar{\lambda}$ is the Lamé parameter; G, E, and υ are the shear modulus, Young's modulus, and Poisson's ratio for the material, respectively, and the strain-displacement relations (cf. [29])

$$\varepsilon_r=u_{,r},\quad \varepsilon_\theta=\frac{1}{r}(v_{,\theta}+u),\quad \varepsilon_z=w_{,z},\quad \gamma_{r\theta}=\frac{u_{,\theta}}{r}+v_{,r}-\frac{v}{r},$$

$$\gamma_{rz}=w_{,r}+u_{,z},\quad \gamma_{\theta z}=v_{,z}+\frac{w_{,\theta}}{r} \tag{4}$$

where u, v, and w are displacement components in the r, θ, and z directions, respectively. Substituting Eqs. (2)–(4) into Eq. (1), one obtains

$$2(1-\upsilon)u_{,rr}+\frac{2(1-\upsilon)}{r}u_{,r}-2(1-\upsilon)\frac{u}{r^2}+\frac{1-2\upsilon}{r^2}u_{,\theta\theta}+(1-2\upsilon)u_{,zz}+\frac{1}{r}v_{,r\theta}-\frac{3-4\upsilon}{r^2}v_{,\theta}+w_{,rz}=0 \tag{5a}$$

$$\frac{1}{r}u_{,r\theta}+\frac{3-4\upsilon}{r^2}u_{,\theta}+(1-2\upsilon)v_{,rr}+(1-2\upsilon)\frac{v_{,r}}{r}-(1-2\upsilon)\frac{v}{r^2}+\frac{2(1-\upsilon)}{r^2}v_{,\theta\theta}+(1-2\upsilon)v_{,zz}+\frac{1}{r}w_{,\theta z}=0 \tag{5b}$$

$$u_{,rz}+\frac{u_{,z}}{r}+\frac{v_{,\theta z}}{r}+2(1-\upsilon)w_{,zz}+(1-2\upsilon)w_{,rr}+\frac{1-2\upsilon}{r}w_{,r}+\frac{1-2\upsilon}{r^2}w_{,\theta\theta}=0 \tag{5c}$$

These equations are also found, for example, in the paper by Chaudhuri and Xie [30].

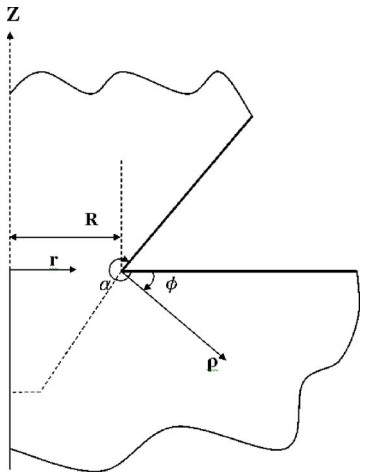


Fig. 2 Cylindrical (r,z) and sharp corner (ρ,ϕ) coordinates

Because solutions of Eqs. (5) will be applied to bodies of revolution subjected to arbitrary static or dynamic loads, it is propitious to assume them as Fourier components in θ

$$u=\sum_{n=0,1} U_n(r,z)\cos n\theta,\quad v=\sum_{n=0,1} V_n(r,z)\sin n\theta \text{ and} \tag{6}$$

$$w=\sum_{n=0,1} W_n(r,z)\cos n\theta$$

Substituting them into Eqs. (5) yields

$$2(1-\nu)U_{n,rr}+\frac{2(1-\nu)}{r}U_{n,r}-[2(1-\nu)+n^2(1-2\nu)]\frac{U_n}{r^2}+(1-2\nu)U_{n,zz}+\frac{n}{r}V_{n,r}-\frac{n(3-4\nu)}{r^2}V_n+W_{n,rz}=0 \tag{7a}$$

$$-\frac{n}{r}U_{n,r}-\frac{(3-4\nu)n}{r^2}U_n+(1-2\nu)V_{n,rr}+\frac{1-2\nu}{r}V_{n,r}-[(1-2\nu)+2n^2(1-\nu)]\frac{V_n}{r^2}+(1-2\nu)V_{n,zz}-\frac{n}{r}W_{n,z}=0 \tag{7b}$$

$$U_{n,rz}+\frac{1}{r}U_{n,z}+\frac{n}{r}V_{n,z}+2(1-\nu)W_{n,zz}+(1-2\nu)W_{n,rr}+\frac{1-2\nu}{r}W_{n,r}-\frac{n^2(1-2\nu)}{r^2}W_n=0 \tag{7c}$$

To investigate the stress singularities at a sharp corner along the circumference of the body, (r, z) coordinates are transformed to (ρ,ϕ) coordinates as shown in Fig. 2. The relations between the two coordinate systems are

$$\rho=\sqrt{(r-R)^2+z^2},\quad \phi=\tan^{-1}\left(\frac{-z}{r-R}\right) \tag{8a}$$

and

$$r-R=\rho\cos\phi,\quad z=-\rho\sin\phi. \tag{8b}$$

Utilizing Eqs. (8) with chain rule differentiation, Eqs. (7) become

$$\begin{aligned}&[2(1-\nu)\cos^2\phi+(1-2\nu)\sin^2\phi]U_{n,\rho\rho}\\&+\left[2(1-\nu)\left(\frac{\sin^2\phi}{\rho}+\frac{\cos\phi}{\rho\cos\phi+R}\right)+(1-2\nu)\frac{\cos^2\phi}{\rho}\right]U_{n,\rho}\\&-\frac{\sin 2\phi}{\rho}U_{n,\rho\phi}-\frac{2(1-\nu)+(1-2\nu)n^2}{(\rho\cos\phi+R)^2}U_n\\&+\left[2(1-\nu)\frac{\sin^2\phi}{\rho^2}+(1-2\nu)\frac{\cos^2\phi}{\rho^2}\right]U_{n,\phi\phi}\\&+\left[-\frac{2(1-\nu)}{\rho\cos\phi+R}\frac{\sin\phi}{\rho}+\frac{\sin 2\phi}{\rho^2}\right]U_{n,\phi}+\frac{n\cos\phi}{\rho\cos\phi+R}V_{n,\rho}\\&-\frac{n\sin\phi}{\rho(\rho\cos\phi+R)}V_{n,\phi}-\frac{n(3-4\nu)}{(\rho\cos\phi+R)^2}V_n-\sin\phi\cos\phi W_{n,\rho\rho}\\&+\frac{\sin\phi\cos\phi}{\rho}W_{n,\rho}-\frac{\cos 2\phi}{\rho}W_{n,\rho\phi}\\&+\frac{\sin\phi\cos\phi}{\rho^2}W_{n,\phi\phi}+\frac{\cos 2\phi}{\rho^2}W_{n,\phi}=0\end{aligned} \tag{9a}$$

$$\begin{aligned}&-\frac{n}{\rho\cos\phi+R}\left(\cos\phi U_{n,\rho}-\frac{\sin\phi}{\rho}U_{n,\phi}\right)-\frac{n(3-4\nu)}{(\rho\cos\phi+R)^2}U_n\\&+(1-2\nu)V_{n,\rho\rho}+(1-2\nu)\left(\frac{1}{\rho}+\frac{\cos\phi}{\rho\cos\phi+R}\right)V_{n,\phi}\\&+\frac{1-2\nu}{\rho^2}V_{n,\phi\phi}-\frac{1-2\nu}{\rho\cos\phi+R}\frac{\sin\phi}{\rho}V_{n,\phi}\\&-\frac{(1-2\nu)+2n^2(1-\nu)}{(\rho\cos\phi+R)^2}V_n+\frac{n}{\rho\cos\phi+R}\\&\times\left(\sin\phi W_{n,\rho}+\frac{\cos\phi}{\rho}W_{n,\phi}\right)=0\end{aligned} \tag{9b}$$

$$\begin{aligned}&-\sin\phi\cos\phi U_{n,\rho\rho}-\frac{\cos 2\phi}{\rho}U_{n,\rho\phi}\\&+\left(\frac{\sin\phi\cos\phi}{\rho}-\frac{\sin\phi}{\rho\cos\phi+R}\right)U_{n,\rho}\\&+\frac{\sin\phi\cos\phi}{\rho^2}U_{n,\phi\phi}+\left(\frac{\cos 2\phi}{\rho^2}-\frac{\cos\phi}{\rho(\rho\cos\phi+R)}\right)U_{n,\phi}\\&-\frac{n}{\rho\cos\phi+R}\left(\sin\phi V_{n,\rho}+\frac{\cos\phi}{\rho}V_{n,\phi}\right)\\&+[2(1-\nu)\sin^2\phi+(1-2\nu)\cos^2\phi]W_{n,\rho\rho}+\frac{\sin 2\phi}{\rho}W_{n,\rho\phi}\\&+\left[2(1-\nu)\frac{\cos^2\phi}{\rho}+(1-2\nu)\left(\frac{\sin^2\phi}{\rho}+\frac{\cos\phi}{\rho\cos\phi+R}\right)\right]W_{n,\rho}\\&+[2(1-\nu)\cos^2\phi+(1-2\nu)\sin^2\phi]\frac{1}{\rho^2}W_{n,\phi\phi}\\&+\left[-2(1-\nu)\frac{\sin 2\phi}{\rho^2}+(1-2\nu)\left(\frac{\sin 2\phi}{\rho^2}\right.\right.\\&\left.\left.-\frac{\sin\phi}{\rho(\rho\cos\phi+R)}\right)\right]W_{n,\phi}-\frac{n^2(1-2\nu)}{(\rho\cos\phi+R)^2}W_n=0\end{aligned} \tag{9c}$$

Now assume

$$U_n(\rho,\phi)=\sum_{m=0,1}^{\infty}\rho^{\lambda+m}\hat{U}_{nm}(\phi),$$

$$V_n(\rho,\phi)=\sum_{m=0,1}^{\infty}\rho^{\lambda+m}\hat{V}_{nm}(\phi),$$

$$W_n(\rho,\phi)=\sum_{m=0,1}^{\infty}\rho^{\lambda+m}\hat{W}_{nm}(\phi), \tag{10}$$

where λ is a yet-undetermined parameter, which can be a complex number. The real part of λ should be positive to make displacement components finite at $\rho=0$. Multiplying through Eqs. (9) by $(\rho\cos\phi+R)^2$, substituting Eqs. (10), and retaining only those terms with the lowest degree of ρ, one obtains the following equations to describe the behavior at the sharp corner:

$$[\lambda(\lambda-1)\cos^2\phi+\lambda^2(1-2\nu)+\lambda\sin^2\phi]\hat{U}_{n0}-(\lambda-1)\sin 2\phi\hat{U}_{n0,\phi}+(\sin^2\phi+1-2\nu)\hat{U}_{n0,\phi\phi}+\sin\phi\cos\phi\hat{W}_{n0,\phi\phi}+(1-\lambda)\cos 2\phi\hat{W}_{n0,\phi}+\lambda(2-\lambda)\sin\phi\cos\phi\hat{W}_{n0}=0 \tag{11a}$$

$$\hat{V}_{n0,\phi\phi}+\lambda^2\hat{V}_{n0}=0 \tag{11b}$$

$$\lambda(2-\lambda)\sin\phi\cos\phi\hat{U}_{n0}+(1-\lambda)\cos 2\phi\hat{U}_{n0,\phi}+\sin\phi\cos\phi\hat{U}_{n0,\phi\phi}+[\lambda(\lambda-1)\sin^2\phi+\lambda^2(1-2\nu)+\lambda\cos^2\phi]\hat{W}_{n0}+(\lambda-1)\sin 2\phi\hat{W}_{n0,\phi}+(\cos^2\phi+1-2\nu)\hat{W}_{n0,\phi\phi}=0 \tag{11c}$$

Equations (11) are independent of n, which means that the stress singularities at ρ approaching to zero are expected to be the same as those for axisymmetric problems ($n=0$).

The solution of Eq. (11*b*) is simply

$$\hat{V}_{n0}=A_1\cos\lambda\phi+A_2\sin\lambda\phi, \tag{12}$$

where A_1 and A_2 are coefficients to be determined from boundary conditions. Equations (11*a*) and (11*c*) are two coupled ordinary differential equations with variable coefficients.

Equations (11*a*) and (11*c*) are simplified to a set of equations with constant coefficients by the following transformation. Define new functions $\bar{U}_{n0}$ and $\bar{W}_{n0}$ such that

$$\hat{U}_{n0}(\phi)=\cos\phi\bar{U}_{n0}(\phi)-\sin\phi\bar{W}_{n0}(\phi) \tag{13a}$$

$$\hat{W}_{n0}(\phi)=-\sin\phi\bar{U}_{n0}(\phi)-\cos\phi\bar{W}_{n0}(\phi) \tag{13b}$$

Substituting Eqs. (13*a*) and (13*b*) into Eqs. (11*a*) and (11*c*) with careful rearrangement yields

$$(1-2\nu)\cos\phi\bar{U}_{n0,\phi\phi}-(3+\lambda-4\nu)\sin\phi\bar{U}_{n0,\phi}+2(\lambda^2-1)(1-\nu)\cos\phi\bar{U}_{n0}-2(1-\nu)\sin\phi\bar{W}_{n0,\phi\phi}+(-3+\lambda+4\nu)\cos\phi\bar{W}_{n0,\phi}+(1-\lambda^2)(1-2\nu)\sin\phi\bar{W}_{n0}=0 \tag{14a}$$

$$-(1-2\nu)\sin\phi\bar{U}_{n0,\phi\phi}-(3+\lambda-4\nu)\cos\phi\bar{U}_{n0,\phi}-2(\lambda^2-1)(1-\nu)\sin\phi\bar{U}_{n0}-2(1-\nu)\cos\phi\bar{W}_{n0,\phi\phi}-(-3+\lambda+4\nu)\sin\phi\bar{W}_{n0,\phi}+(1-\lambda^2)(1-2\nu)\cos\phi\bar{W}_{n0}=0 \tag{14b}$$

Multiplying Eq. (14*a*) by $\cos\phi$, and subtracting Eq. (14*b*) multiplied by $\sin\phi$ yields

$$(1-2\nu)\bar{U}_{n0,\phi\phi}+2(\lambda^2-1)(1-\nu)\bar{U}_{n0}+(-3+\lambda+4\nu)\bar{W}_{n0,\phi}=0 \tag{15a}$$

Similarly, summing Eq. (14*a*) multiplied by $\sin\phi$ and Eq. (14*b*) multiplied by $\cos\phi$ yields

$$-(3+\lambda-4\nu)\bar{U}_{n0,\phi}-2(1-\nu)\bar{W}_{n0,\phi\phi}+(1-\lambda^2)(1-2\nu)\bar{W}_{n0}=0 \tag{15b}$$

Thus, Eqs. (15*a*) and (15*b*) are two coupled ordinary differential equations with constant coefficients. The solutions can be easily obtained by standard procedures for solving linear differential equations, and they are

$$\bar{U}_{n0}(\phi)=B_1\sin(\lambda+1)\phi-B_2\cos(\lambda+1)\phi+\gamma B_3\sin(\lambda-1)\phi-\gamma B_4\cos(\lambda-1)\phi \tag{16a}$$

$$\bar{W}_{n0}(\phi)=B_1\cos(\lambda+1)\phi+B_2\sin(\lambda+1)\phi+B_3\cos(\lambda-1)\phi+B_4\sin(\lambda-1)\phi \tag{16b}$$

where $\gamma=(-3+\lambda+4\nu)/(3+\lambda-4\nu)$, and B_i $(i=1,2,3,4)$ are coefficients to be determined from boundary conditions.

In a brief summary, the solutions of Eqs. (5) are

$$u(\rho,\theta,\phi)=\sum_n\rho^{\lambda}[\cos\phi\bar{U}_{n0}(\phi)-\sin\phi\bar{W}_{n0}(\phi)]\cos n\theta+O(\rho^{\lambda+1})=\tilde{u}(\rho,\theta,\phi)+O(\rho^{\lambda+1}) \tag{17a}$$

$$v(\rho,\theta,\phi)=\sum_n\rho^{\lambda}(A_1\cos\lambda\phi+A_2\sin\lambda\phi)\sin n\theta+O(\rho^{\lambda+1})=\tilde{v}(\rho,\theta,\phi)+O(\rho^{\lambda+1}) \tag{17b}$$

$$w(\rho,\theta,\phi)=\sum_n\rho^{\lambda}[-\sin\phi\bar{U}_{n0}(\phi)-\cos\phi\bar{W}_{n0}(\phi)]\cos n\theta+O(\rho^{\lambda+1})=\tilde{w}(\rho,\theta,\phi)+O(\rho^{\lambda+1}) \tag{17c}$$

where $\bar{U}_{n0}(\phi)$ and $\bar{W}_{n0}(\phi)$ are given in Eqs. (16*a*) and (16*b*), and $O(\rho^{\lambda+1})$ are terms of higher order in ρ.

Boundary Conditions, Characteristic Equations, and Corner Functions

Having solutions to the equilibrium equations, attention is now turned to the boundary conditions along the edges of the sharp corner. These surfaces, $\phi=0$ and $\phi=\alpha$ (see Fig. 2) may each be either free or fixed. For example, for $\phi=\alpha$

(a) Free (traction forces are equal to zero)

$$T_r=\sigma_r\sin\alpha+\tau_{rz}\cos\alpha=0,\quad T_z=\tau_{zr}\sin\alpha+\sigma_z\cos\alpha=0,\quad T_\theta=\tau_{\theta r}\sin\alpha+\tau_{\theta z}\cos\alpha=0 \tag{18a}$$

(b) Fixed

$$u(\rho,\theta,\alpha)=v(\rho,\theta,\alpha)=w(\rho,\theta,\alpha)=0 \tag{18b}$$

Substituting the displacements (Eqs. (17)) into Eqs. (2)–(4), as ρ approaches to zero, the singular stress components can be asymptotically expressed for each Fourier component (n) as

$$\sigma_r=\frac{\nu E}{(1+\nu)(1-2\nu)}\Lambda+\frac{E}{1+\nu}\left(\tilde{u}_{,\rho}\cos\phi-\frac{\sin\phi\,\tilde{u}_{,\phi}}{\rho}\right)+O(\rho^{\lambda}) \tag{19a}$$

Table 1 Characteristic equations and corner functions for all combinations of fixed or free intersecting surface

Boundary conditions			
$\phi=0$	$\phi=\alpha$	Characteristic equations	Corner functions
Free	Free	$\sin\lambda\alpha=0$ $\sin\lambda\alpha=\pm\lambda\sin\alpha$	$v=\rho^{\lambda}\cos\lambda\phi\sin n\theta$ $u=\rho^{\lambda}(\cos\phi U_1-\sin\phi W_1)\cos n\theta$ $w=\rho^{\lambda}(-\sin\phi U_1-\cos\phi W_1)\cos n\theta$
Free	Fixed	$\cos\lambda\alpha=0$ $\sin^2\lambda\alpha=\dfrac{4(-1+\upsilon)^2-\lambda^2\sin^2\alpha}{3-4\upsilon}$	$v=\rho^{\lambda}\cos\lambda\phi\sin n\theta$ $u=\rho^{\lambda}(\cos\phi U_2-\sin\phi W_2)\cos n\theta$ $w=\rho^{\lambda}(-\sin\phi U_2-\cos\phi W_2)\cos n\theta$
Fixed	Fixed	$\sin\lambda\alpha=0$ $\sin\lambda\alpha=\pm\dfrac{\lambda\sin\alpha}{3-4\upsilon}$	$v=\rho^{\lambda}\sin\lambda\phi\sin n\theta$ $u=\rho^{\lambda}(\cos\phi U_3-\sin\phi W_3)\cos n\theta$ $w=\rho^{\lambda}(-\sin\phi U_3-\cos\phi W_3)\cos n\theta$

Note:
$U_1=-(-1+\lambda)/(3+\lambda-4\upsilon)\eta_1\sin(\lambda+1)\phi+(1+\lambda)/(3+\lambda-4\upsilon)\cos(\lambda+1)\phi-\gamma\eta_1\sin(\lambda-1)\phi-\gamma\cos(\lambda-1)\phi$,
$W_1=-(-1+\lambda)/(3+\lambda-4\upsilon)\eta_1\cos(\lambda+1)\phi-(1+\lambda)/(3+\lambda-4\upsilon)\sin(\lambda+1)\phi-\eta_1\cos(\lambda-1)\phi+\sin(\lambda-1)\phi$,
$U_2=-(-1+\lambda)/(3+\lambda-4\upsilon)\eta_2\sin(\lambda+1)\phi+(1+\lambda)/(3+\lambda-4\upsilon)\cos(\lambda+1)\phi+\gamma\eta_2\sin(\lambda-1)\phi-\gamma\cos(\lambda-1)\phi$,
$W_2=-(-1+\lambda)/(3+\Lambda-4Y)\eta_2\cos(\lambda+1)\phi-(1+\lambda)/(3+\lambda-4\upsilon)\sin(\lambda+1)\phi+\eta_2\cos(\lambda-1)\phi+\sin(\lambda-1)\phi$,
$U_3=-\eta_3\sin(\lambda+1)\phi+(-3+\lambda+4\upsilon)/(3+\lambda-4\upsilon)\cos(\lambda+1)\phi+\gamma\eta_3\sin(\lambda-1)\phi-\gamma\cos(\lambda-1)\phi$,
$W_3=-\eta_3\cos(\lambda+1)\phi-(-3+\lambda+4\upsilon)/(3+\lambda-4\upsilon)\sin(\lambda+1)\phi+\eta_3\cos(\lambda-1)\phi+\sin(\lambda-1)\phi$,
$\eta_1=[\lambda\sin(\lambda-2)\alpha-(2+\lambda)\sin\lambda\alpha]/[2\lambda\sin\alpha\sin(\lambda-1)\alpha]$, $\eta_2=[2(-1+\upsilon)\cos\alpha\cos\lambda\alpha+(-1+\lambda+2\upsilon)\sin\alpha\sin\lambda\alpha]/[(-2+\lambda+2\upsilon)\sin\alpha\cos\lambda\alpha+(1-2\upsilon)\cos\alpha\sin\lambda\alpha]$,
$\eta_3=[-\gamma[\cos(\lambda+1)\alpha-\cos(\lambda-1)\alpha]]/[-\sin(\lambda+1)\alpha+\gamma\sin(\lambda-1)\alpha]$.

$$\sigma_z=\frac{\upsilon E}{(1+\upsilon)(1-2\upsilon)}\Lambda+\frac{E}{1+\upsilon}\left(-\sin\phi\tilde{w}_{,\rho}-\frac{\cos\phi\tilde{w}_{,\phi}}{\rho}\right)+O(\rho^{\lambda}) \tag{19b}$$

$$\sigma_\theta=\frac{\upsilon E}{(1+\upsilon)(1-2\upsilon)}\Lambda+O(\rho^{\lambda}) \tag{19c}$$

$$\tau_{r\theta}=\tau_{\theta r}=\frac{E}{2(1+\upsilon)}\left(\cos\phi\tilde{v}_{,\rho}-\frac{\sin\phi\tilde{v}_{,\phi}}{\rho}\right)+O(\rho^{\lambda}) \tag{19d}$$

$$\tau_{z\theta}=\tau_{\theta z}=\frac{E}{2(1+\upsilon)}\left(-\sin\phi\tilde{v}_{,\rho}-\frac{\cos\phi\tilde{v}_{,\phi}}{\rho}\right)+O(\rho^{\lambda}) \tag{19e}$$

$$\tau_{rz}=\tau_{zr}=\frac{E}{2(1+\upsilon)}\left(\cos\phi\tilde{w}_{,\rho}-\frac{\sin\phi\tilde{w}_{,\phi}}{\rho}-\sin\phi\tilde{u}_{,\rho}-\frac{\cos\phi\tilde{u}_{,\phi}}{\rho}\right)+O(\rho^{\lambda}) \tag{19f}$$

where $\Lambda=\cos\phi\tilde{u}_{,\rho}-\dfrac{\sin\phi\tilde{u}_{,\phi}}{\rho}-\sin\phi\tilde{w}_{,\rho}-\dfrac{\cos\phi\tilde{w}_{,\phi}}{\rho}$.

Substituting either the displacements of Eqs. (17), or the stresses of Eqs. (19), as needed, into the boundary conditions of either Eq. (18*a*) or (18*b*) at the faces $\phi=0$ and $\phi=\alpha$ yields six homogeneous, linear algebraic equations in the six coefficients A_1, A_2, B_1, B_2, B_3, and B_4. For a nontrivial solution the determinant of the coefficient matrix is set to zero, from which the eigenvalues (λ) are obtained. Moreover, the coefficients A_1 and A_2 are uncoupled from the remaining four, so two sets of λ are determined, one set from the second order determinant of A_1 and A_2, and one set from the fourth order (B_1, B_2, B_3, B_4) determinant.

For example, if the surface $\phi=0$ and $\phi=\alpha$ are both fixed, then one obtains from the second of Eqs. (18*b*)

$$\begin{bmatrix}\cos\lambda\alpha & \sin\lambda\alpha\\ 1 & 0\end{bmatrix}\begin{Bmatrix}A_1\\A_2\end{Bmatrix}=0 \tag{20}$$

Whence $\sin\lambda\alpha=0$ is the characteristic equation for this set of λ. Applying the first and third of Eqs. (18*b*) at $\phi=0$ and $\phi=\alpha$ results in a set of four equations

$$\begin{bmatrix}0 & 1 & 0 & \gamma\\ 1 & 0 & 1 & 0\\ \sin(\lambda+1)\alpha & -\cos(\lambda+1)\alpha & \gamma\sin(\lambda-1)\alpha & -\gamma\cos(\lambda-1)\alpha\\ \cos(\lambda+1)\alpha & \sin(\lambda+1)\alpha & \cos(\lambda-1)\alpha & \sin(\lambda-1)\alpha\end{bmatrix}\times\begin{Bmatrix}B_1\\B_2\\B_3\\B_4\end{Bmatrix}=0 \tag{21}$$

Evaluating the determinant, using some trigonometric identities, and simplifying it further, one arrives at

$$\sin\lambda\alpha=\pm\frac{\lambda\sin\alpha}{3-4\upsilon} \tag{22}$$

Having the characteristic equations, the eigenvalues λ can be determined from them and substituted into Eqs. (20) and (21) to obtain the eigenvectors A_1/A_2 and B_1/B_4, B_2/B_4, B_3/B_4. The resulting eigenfunctions are the desired corner functions, having arbitrary amplitudes which may be taken as unity.

The characteristic equations for all three combinations of fixed or free corner surfaces have been obtained. They are given in Table 1. The equations corresponding to u and w in Table 1 were obtained by Zak [27] for the axisymmetric ($n=0$) case only, using a different approach. These equations are also identical to those derived by Williams [1] for the sharp corner plane elasticity problems.

The corner functions for the three combinations of fixed or free edges are also presented in Table 1. They have not been found in the previously published literature.

Equations (19*d*) and (19*e*) show that the shear stresses $\tau_{r\theta} = \tau_{\theta r}$and $\tau_{z\theta} = \tau_{\theta z}$ depend only on v, whereas the other stresses depend upon u and w. Because the characteristic equations for v are different from those for u and w, the singularities for the former stresses are different from those for the latter ones.

Concluding Remarks

The objective of this work was to derive corner displacement functions (u, v, w) which represent well the stresses and deformation of an elastic body of revolution in the vicinity of a sharp boundary corner. This was accomplished by means of an asymptotic analysis, which resulted in the characteristic equations and their corresponding corner functions summarized in Table 1.

The corner functions will be used in future 3D studies to determine accurate free vibration frequencies and mode shapes of bodies having such boundary discontinuities. These occur frequently in practice when rods or bars are machined (e.g., a circumferential V notch, or an abrupt diameter change). Although other corner functions have been used to advantage for vibration studies of 2D continuous systems [13–16,18], and the present ones are expected to be suitable for 3D vibration problems, they can also be used for static stress and deformation analysis, especially for determining the stress intensity factors for a V notch.

The present work deals with single homogeneous bodies. It would also be useful to have corner functions for bimaterial bodies, such as when a truncated cone is bonded to a cylinder of other material, as shown in Fig. 1. This will be the subject of a future study.

References

[1] Williams, M. L., 1952, "Stress Singularities Resulting from Various Boundary Conditions in Angular Corners of Plates in Extension," Am. J. Sci., **19**, pp. 526–528.

[2] Williams, M. L., 1953, "Discussion of 'Stress Singularities Resulting from Various Boundary Conditions in Angular Corners of Plates in Extension'," Am. J. Sci., **20**, pp. 590.

[3] Williams, M. L., 1952, "Surface Stress Singularities Resulting from Various Boundary Conditions in Angular Corners of Plates under Bending," *Proceedings of the First U.S. National Congress of Applied Mechanics*, ASME, New York, pp. 325–329.

[4] Sih, G. C., and Rice, J. R., 1964, "The Bending of Plates of Dissimilar Materials with Cracks," Am. J. Sci., **31**, pp. 477–482.

[5] Hein, V. L., and Erdogan, F., 1971, "Stress Singularities in a Two-Material Wedge," Int. J. Fract. Mech., **7**, pp. 317–330.

[6] Bogy, D. B., and Wang, K. C., 1971, "Stress Singularities at Interface Corners in Bonded Dissimilar Isotropic Elastic Materials," Int. J. Solids Struct., **7**, pp. 993–1005.

[7] Ting, T. C. T., 1990, "Interface Cracks in Anisotropic Bimaterials," J. Mech. Phys. Solids, **38**, pp. 505–513.

[8] Hartranft, R. J., and Sih, G. C., 1969, "The Use of Eigenfunction Expansions in the General Solution of the Three-Dimensional Crack Problems," J. Math. Mech., **19**, pp. 123–138.

[9] Su, X. M., and Sun, C. T., 1996, "On Singular Stress at the Crack Tip of a Thick Plate Under In-Plane Loading," Int. J. Fract., **82**, pp. 237–252.

[10] Glushkov, E., Glushkova, N., and Lapina, O., 1999, "3-D Elastic Stress Singularity at Polyhedral Corner Points," Int. J. Solids Struct., **36**, pp. 1105–1128.

[11] Ritz, W., 1908, "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik," J. Reine Angew. Math., **135**, pp. 1–61.

[12] Ritz, W., 1909, "Theorie der Transversalschwingungen einer quadratische Platte mit freien Rändern," Ann. Phys., **28**, pp. 737–786.

[13] Leissa, A. W., McGee, O. G., and Huang, C. S., 1993, "Vibrations of Sectorial Plates Having Corner Stress Singularities," Am. J. Sci., **60**, pp. 134–140.

[14] Leissa, A. W., McGee, O. G., and Huang, C. S., 1993, "Vibrations of Circular Plates Having V-Notches or Sharp Radial Cracks," J. Sound Vib., **161**, pp. 227–239.

[15] McGee, O. G., Leissa, A. W., and Huang, C. S., 1992, "Vibrations of Cantilevered Skew Plates with Corner Stress Singularities," Int. J. Numer. Methods Eng., **35**, pp. 409–423.

[16] Huang, C. S., McGee, O. G., Leissa, A. W., and Kim, J. W., 1995, "Accurate Vibration Analysis of Simply Supported Rhombic Plates by Considering Stress Singularities," ASME J. Vibr. Acoust., **117**, pp. 245–251.

[17] Huang, C. S., 2003, "Stress Singularities in Angular Corners in First-Order Shear Deformation Plate Theory," Int. J. Mech. Sci., **45**, pp. 1–20.

[18] Huang, C. S., Leissa, A. W., and Chang, M. J., 2005, "Vibrations of Skewed Cantilevered Triangular, Trapezoidal and Parallelogram Mindlin Plates with Considering Corner Stress Singularities," Int. J. Numer. Methods Eng., **62**, pp. 1789–1806.

[19] Yosibash, Z., and Schiff, B., 1993, "A Superelement for Two-Dimensional Singular Boundary Value Problems in Linear Elasticity," Int. J. Fract., **62**, pp. 325–340.

[20] Belytschko, T., Krongauz, Y., Fleming, M., Organ, D., and Liu, W. K., 1996, "Smoothing and Accelerated Computations in the Element Free Galerkin Method," J. Comput. Appl. Math., **74**, pp. 111–126.

[21] Fleming, M., Chu, Y. A., Moran, B., and Belyschko, T., 1997, "Enriched Element Free Galerkin Methods for Crack Tip Fields," Int. J. Numer. Methods Eng., **40**, pp. 1483–1504.

[22] Dolbow, J., Möse, N., and Belyschko, T., 2000, "Discontinuous Enrichment in Finite Element with a Partition of Unity Method," Finite Elem. Anal. Design, **36**, pp. 235–260.

[23] Leissa, A. W., and So, J., 1995, "Comparisons of Vibration Frequencies of Rods and Beams from One-Dimensional and Three-Dimensional Analyses," J. Acoust. Soc. Am., **98**, pp. 2122–2135.

[24] Leissa, A. W., and So, J., 1995, "Three-Dimensional Vibrations of Truncated Hollow Cones," J. Vib. Control, **1**, pp. 145–158.

[25] Leissa, A. W., and Kang, J.-H., 1999, "Three-Dimensional Vibration Analysis of Thick Shells of Revolution," J. Engrg. Mech. Div., **125**, pp. 1365–1372.

[26] Leissa, A. W., and Kang, J.-H., 2000, "Three-Dimensional Vibrations of Thick Spherical Shell Segments with Variable Thickness," Int. J. Solids Struct., **37**, pp. 4811–4823.

[27] Zak, A. R., 1964, "Stresses in the Vicinity of Boundary Discontinuities in Bodies of Revolution," Am. J. Sci., **31**, pp. 150–152.

[28] Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., The Macmillan Co, New York (reprinted by Daver Publications, 1944).

[29] Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, 2nd ed., McGraw-Hill Book, New York.

[30] Chaudhuri, R. A., and Xie, M., 2000, "A Novel Eigenfunction Expansion Solution for Three-Dimensional Crack Problems," Compos. Sci. Technol., **60**, pp. 2565–2580.

Rigid Body Dynamics, Constraints, and Inverses

Hooshang Hemami
Department of Electrical and Computer Engineering,
The Ohio State University,
Columbus, OH 43210

Bostwick F. Wyman
Department of Mathematics,
The Ohio State University,
Columbus, OH 43210

Rigid body dynamics are traditionally formulated by Lagrangian or Newton-Euler methods. A particular state space form using Euler angles and angular velocities expressed in the body coordinate system is employed here to address constrained rigid body dynamics. We study gliding and rolling, and we develop inverse systems for estimation of internal and contact forces of constraint. A primitive approximation of biped locomotion serves as a motivation for this work. A class of constraints is formulated in this state space. Rolling and gliding are common in contact sports, in interaction of humans and robots with their environment where one surface makes contact with another surface, and at skeletal joints in living systems. This formulation of constraints is important for control purposes. The estimation of applied and constraint forces and torques at the joints of natural and robotic systems is a challenge. Direct and indirect measurement methods involving a combination of kinematic data and computation are discussed. The basic methodology is developed for one single rigid body for simplicity, brevity, and precision. Computer simulations are presented to demonstrate the feasibility and effectiveness of the approaches presented. The methodology can be applied to a multilink model of bipedal systems where natural and/or artificial connectors and actuators are modeled. Estimation of the forces is accomplished by the inverse of the nonlinear plant designed by using a robust high gain feedback system. The inverse is shown to be stable, and bounds on the tracking error are developed. Lyapunov stability methods are used to establish global stability of the inverse system. [DOI: 10.1115/1.2178359]

1 Introduction

Three problems associated with rigid body dynamics are stability [1], control of constraints [2,3], and inverses for control and measurement. Stability is an issue for both the system and its inverse. The state space formulation here yields itself to systematic Lyapunov studies. The issue of noninvasive measurement of joint forces and torques [4–6] is important in the study of human location and biped models. Instruments and sensors are invasive and often undesirable. Instruments can be large, clumsy, and difficult to insert. They interfere with natural or intended function, and they may also require undesirable harnesses. A feasible alternative solution is to rely on computation and indirect measurement where quantities that are easy to measure are sensed, and measurements and computations are combined to estimate other quantities of interest [7,8].

Part of the directly and easily measurable quantities are ground reaction forces measured using instrumented force plates. Recently, better instrumentation of platforms supplies both force and moment of force measurements. Certain important parameters such as center of pressure (cop), zero moment point (zmp), and foot rotation indicator (fri) can be estimated by computation [5]. All these measurements are in the inertial (global) coordinate system (ics). Some of these quantities must be transformed to a body coordinate system (bcs), whose origin is at the center of mass, and whose axes are along the principal axes of the body [9].

The computational methods of interest here are used for calculating ground reaction forces and input torques (or moment of force).

This paper proposes the use of concepts from functional analysis [10,11] and high gain systems to construct the inverse. The application of high gain systems to linear control systems for the purposes of compensation, robustness, stability, and design has been known for some time [12,13]. Nonlinear feedback systems with high gains can be theoretically studied within the framework of singular perturbation [14,15]. However, the emphasis and development here is on the application of inverse systems [11,12]. In this paper we use the inverse system to estimate joint forces and torques without using invasive methods.

Here it is assumed that the whole state of the system is available for input to the inverse system. The measurement of the state of the system is based on a camera-computer vision system. With some computation, we are able to arrive at the three Euler angles and their first and second derivatives with respect to time. Planar cases of a multisegment system that demonstrates these issues are studied in Refs. [7,8].

This work is relevant to diagnostic procedures where a subject stands on a platform and performs specified maneuvers or his posture is disturbed by deliberate random motions of the platform. The objective is to estimate joint or muscular forces. The three issues are platform measurements, computation methods for joint force and torque estimation, and combinations of the two. The problem requires recursive procedures and inverse system applications to multi-segment systems. For simplicity and precision, the discussions and formulations of this paper are limited to a single rigid body with one idealized (resultant) vector of torque actuation and a massless stationary foot. More realistic multisegment models with natural attributes of muscles, ligaments, and skin tissue constitute future endeavors.

The inverse construction has other applications in understanding control mechanisms used by natural systems.

Maintaining upright stability and generating purposeful movements require knowledge of the movement of various body segments. The sensory modalities, namely, the proprioceptive, visual, and vestibular mechanisms, provide the angles and angular velocities [16,17]. Some of the angles and angular velocities are sensed directly (e.g., head velocity by the vestibular system) while others are sensed indirectly (e.g., ankle angle by the proprioceptive and somatosensory receptors). Based on this information, the central

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 3, 2004; final manuscript received December 19, 2005. Review conducted by I. Mezic. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

nervous system decides which muscles must be activated to generate the desired joint motions. Although the exact timing and magnitude of muscle activation are most likely determined locally, a more global control strategy is necessary to coordinate movements of different body segments. Determination of joint torques is the essential first step to understand the feedback mechanisms that map the sensory information onto the motor outputs.

A related application of inverse system methodology is to the concept of preprogramming in the central nervous system (CNS). Horak and Nashner [18] have proposed a hierarchical organization of postural control mechanisms in which a limited set of preprogrammed motor routines is used to generate multidimensional movements. In the most basic form, these movements take the form of ballistic movements where the limb trajectories are determined by the initial burst of the neural activity [19]. The ballistic movements are required for executing fast movements. Postural adjustments seem to fall under the category of "planned" movements. For these movements, the motor control system operates in a closed-loop manner to achieve a greater accuracy while minimizing the effects of neural delays. This view of the postural control system requires the CNS to maintain an approximate model of the intended movement and the internal system dynamics [20]. This model is developed through a "learning" process. During the execution of planned movements, the "stored" or "learned" movements are constantly compared with the information arriving from the sensory mechanisms. As long as the intended and actual movements are within a tolerance limit, the programmed routines can proceed without intervention. When conflicts arise, such as when an unanticipated postural disturbance is encountered, the CNS will first attempt to modify the parameters of the preprogrammed routine to achieve the intended movement. If the conflict persists, such as when a lesion develops in the sensorimotor mechanisms, the CNS must actively use the available sensory feedback mechanisms to maintain stability while at the same time modifying its internal model of the system dynamics. The cost of such conflicts is increased response time and reduced accuracy of resulting movements. We hope to apply our methods to the preprogramming situation in a later paper.

2 Rigid Body Dynamics

Rigid body dynamics and control can be formulated with the recently developed and elegant geometric tools [21,22]. Here we apply the Newton-Euler method in order to derive the equations of a single free rigid body. Let Θ and Ω be, respectively, the Euler angles and the angular velocity vector of the body expressed in the previously defined body coordinate system (bcs). Let X and V be the translational vectors of position and velocity of the center of gravity of the body, expressed in the inertial coordinate system (ics) system. With reference to vector R, we define the skew symmetric 3×3 matrix $\breve{\mathcal{R}}$ [23,24]. The vectors of force G and Λ are, respectively, the gravity vector and an equivalent or resultant vector of all forces [9] acting on the rigid body. Similarly N is an equivalent or resultant couple of all forces acting on the rigid body. Alternatively speaking, N is the sum of all the couples: stabilizing couples, trajectory control couples, and the moment of all forces acting on the rigid body. The asymptotic stability of the rotational system by nonlinear feedback has been discussed in Refs. [1,23], and will be briefly presented later in this paper. The coordinate systems are defined in the Appendix . Based on these simplifications, the equations of motion of the single rigid body are[25,26]

$$\begin{aligned}\dot{\Theta} &= B(\Theta)\Omega \\ J\dot{\Omega} &= f(\Omega) + N \\ \dot{X} &= V \\ m\dot{V} &= G + \Lambda\end{aligned} \tag{1}$$

where J is the diagonal moment of inertia matrix, expressed in bcs, B is a 3×3 matrix defined in the appendix , and where

$$f(\Omega) = \breve{\Omega}J\Omega$$

2.1 Rolling and Gliding. When a rigid body contacts another rigid body, the two surfaces can be approximated by rigid body spheres. As an example, suppose a player kicks a soccer ball such that the front of the shoe touches the ball, i.e., a toe kick takes place. The collision can be approximated by the viscoelastic collision of one sphere with another. When the player kicks the ball with the top of his foot, the foot surface can be approximated by a plane, a cylinder, or a sphere.

Let a uniformly dense rigid body sphere with radius Q be centered at the origin of the inertial system. This means the center of the sphere and its center of gravity coincide. Let another uniformly dense sphere of radius q roll or glide on the first sphere. The contact eliminates one of the three degrees of translational freedom of the moving sphere. At the point of contact, the two spheres' surfaces can be approximated by their respective tangent planes. This means two of the degrees of freedom of the moving rigid body are thus constrained. The remaining degree of freedom is the self- rotation of the rigid body along the center to center line of the two spheres. In order to control the motions of gliding and rolling, the equations of constraint are needed [2,3].

Let Eq. (1) describe the motion of the moving sphere. In both gliding and rolling, the two spheres' being in contact is described by the holonomic constraint

$$X'X - (Q+q)^2 = 0 \tag{2}$$

Alternatively, this equation is written as

$$X'\dot{X} = 0 \tag{3}$$

2.2 Pure Gliding. In gliding a fixed vector S in the body coordinate system (bcs) keeps contact with the stationary sphere. Let p and $\dot{p}$ be, respectively, the vector of the contact point, and its velocity in ics:

$$\begin{aligned}p &= X + A(\Theta)S \\ \dot{p} &= \dot{X} + A(\Theta)\breve{\Omega}S\end{aligned} \tag{4}$$

In gliding, the velocity of the moving body at the point of contact is in the tangent plane of the stationary sphere, therefore the inner product of $\dot{p}$ and X is zero. Also the velocities of the point of contact and the center of gravity of the moving sphere are in parallel, because the instantaneous center of rotation of the moving sphere is the origin of the ics, namely, the center of the stationary sphere. The latter requirement means

$$\dot{p} = [q/(Q+q)]\dot{X} \tag{5}$$

From Eqs. (4) and (5), the constraints for gliding can be obtained

$$q\dot{X} + (Q+q)A(\Theta)\breve{\Omega}S = 0 \tag{6}$$

2.3 Rolling Motion. Rolling is a special case of the more general class of nonholonomic systems [2], Chap. 3. Nonholonomic constraints either appear as part of the structure of a mechanical system, or are part of the specifications for control in order to simplify implementation of coordinated movement.

In rolling motion of the sphere, the instantaneous center of rotation of the moving sphere is the point of contact. From this information, and the above equations, one can obtain the three constraints for rolling

$$Q\dot{X} + qA(\Theta)\breve{\Omega}A'(\Theta)X = 0 \tag{7}$$

With the equations of the rigid body and the constraints given above, the questions of the derivation of the necessary forces of constraint in order to maintain the constraint, and deriving feedback in order to implement the constrained motion can be carried out as described in Refs. [2,3]. These issues are not further pursued here.

2.4 Illustrative Example. Suppose the motion is limited to the x_1x_3 plane. Suppose further that the spheres are initially stacked on the top of each other so that

$$p(0) = [0,0,Q]'$$

and

$$X(0) = [0,0,Q+q]'$$

The equations of motion can be derived by setting

$$x_2 = \omega_1 = \omega_3 = 0$$

The gliding constraints in this case become

$$\begin{aligned} \dot{x_1} &= (Q+q)\dot{\theta}_2\cos(\theta_2) \\ \dot{x_3} &= -(Q+q)\dot{\theta}_2\sin(\theta_2) \end{aligned} \tag{8}$$

The rolling constraints in this case become

$$\begin{aligned} Q\dot{x_1} &= q\dot{\theta}_2 x_3 \\ Q\dot{x_3} &= -q\dot{\theta}_2 x_1 \end{aligned} \tag{9}$$

2.5 Rotational Motion. Consider the rigid body formulation in Eq. (1). Suppose the center of gravity is fixed in the ics, and only the rotational motion is of interest. For an alternative state space formulation of rigid body dynamics, see [24].

It is assumed here that the range of Θ is limited such that the Lipschitz condition is satisfied. With the latter assumption, it has been shown that with state feedback

$$N = -B'(\Theta)K_1\Theta - L_1\Omega$$

where both K_1 and L_1 are positive definite [23] that the rotational rigid body motion is asymptotically stable. Isidori ([24], Lemma 2, Appendix B, Sec. 2.2), and Khalil ([15], Chaps. 5,6) have shown that the system is stable under nonvanishing persistent disturbance. Equivalently, one can prove that the rotational system of Eq. (1) produces uniformly bounded outputs when the inputs is uniformly bounded. This means the above system is bounded-input bounded-output (BIBO) stable.

3 Inverse System

3.1 Basics. The inverse system can be formally designed based on concepts from functional analysis [11] and nonlinear operator algebra [10] or from more recent geometrically and algebraically based methods [24,27]. The functional analysis method for computation of the inverse is based on feedback [11,12]. Since natural, robotic, and humanoid systems use feedback for stability, control, and tracking, use of methods based on feedback appear to be more natural, intuitive, and less sensitive to system parameter variations. Besides, the natural structure of the system to be inverted is used for arriving at the inverse. In addition, it is important that the inverse system be stable. It is not clear, from the development of the modern techniques [24,27], whether additional stabilizing mechanisms are needed for the inverse system, and how these additional mechanisms may disturb the inverse system, and interfere with its operation. The approach here also does not require development of zero dynamics [24] since it is assumed that the state and hence the initial state are known for the inverse system.

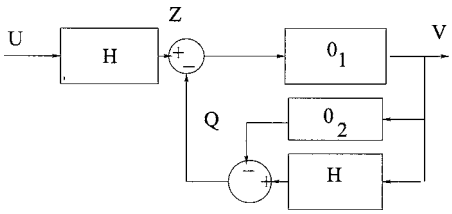


Fig. 1 A continuous nonlinear system *H*, and its inverse composed of a replica of *H* and two operators O_1 and O_2

The method developed below considers global stability and inversion issues simultaneously. It is shown how global stability may effect the performance of the inverse system, and how an exact inverse may affect global stability. First a brief introduction is made to the feedback structure of the general inverse systems. The structure is applied to the rotational dynamics of a single rigid body next. Global stability is guaranteed by perturbation of the inverse.

Let U be the input and Z the state output of a nonlinear continuous invertible system described operationally by a nonlinear operator H. Following the formulation of [10,11], the system can be described by the following equation:

$$Z = HU \tag{10}$$

We assume that conditions for the existence and uniqueness of the solution for Z, and the inverse system [11] are satisfied. Let O_1 and O_2 be two operators such that the product of the two operators is equal to identity

$$O_1O_2 = I$$

where I is the identity operator. The block diagram of the system above followed by the inverse system is given in Fig. 1. The output of the inverse system is V. The equations of the inverse system are

$$\begin{aligned} Q &= (H - O_2)V \\ V &= O_1(Z - Q) \end{aligned} \tag{11}$$

Elimination of Q from the above two equations results in

$$Z = HV \tag{12}$$

From Eqs. (10) and (12), it follows that

$$U = V$$

In other words the tandem connection of the system and its inverse results in the identity operator. In practical terms if Z is the measured state of a musculo-skeletal system, and if it is desired to estimate the input muscular vector of forces, i.e., vector U, the output V of the inverse system is an estimate of the unknown muscular forces. This inverse system can be constructed by the block diagram of Fig. 1, and requires an exact replica of the original system, i.e., the operator H.

3.2 Construction. We construct here the inverse of the rotational system discussed above. The inverse system is based on the above nonlinear operator principles and is defined by a nonlinear feedback system with one forward component and two feedback components as shown in Fig. 1.

Consider the rotational system alone

$$\begin{aligned} \dot{\Theta} &= B(\Theta)\Omega \\ J\dot{\Omega} &= f(\Omega) + U \end{aligned} \tag{13}$$

Consider the operator H to be represented by Eq. (13). The state

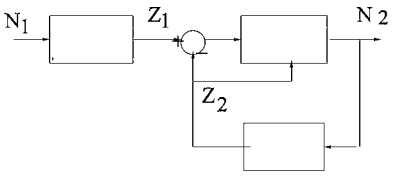


Fig. 2 The free rigid body in rotation with input N_1 and output Z_1, (Eq. (13)) in series with the nonlinear inverse with high gain feed forward and a model of the system in the feedback path

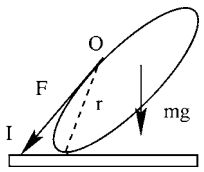


Fig. 3 The inverted pendulum in contact with a massless foot. An actuator with origin O on the pendulum and insertion point on the foot applies a force F to the pendulum along the line OI.

$$Z = [\Theta', \Omega']'$$

and the input is

$$U$$

In this formulation Z is of dimension six, and U of dimension three. For ease of formulation, we assume U is a sum of two couples so that it is also of dimension six

$$U = [V_1', V_2']'$$

The feed forward component of the inverse system is a nonlinear amplifier with high gain. Let K and L be two 3×3 positive definite matrices, and define the 6×6 positive definite matrix G_1 with K and L along its diagonal

$$G_1 = [K,0,0,G] \tag{14}$$

Similarly, with B as defined in Eq. (1), and I as the 3×3 identity matrix, define the 6×6 matrix α as follows:

$$\alpha = [B',0,0,I] \tag{15}$$

With these definitions, the operator O_1 is defined

$$O_1 = \alpha G_1 \tag{16}$$

The operator O_2 is simply the inverse of O_1

$$O_2 = (G_1)^{-1}(\alpha^{-1}) \tag{17}$$

The other feedback component is a replica of the original system slightly modified to account for adding together the vectors N_1 and N_2 as a single input vector.

$$\dot{\Theta}_2 = B(\Theta_2)\Omega_2$$
$$J\dot{\Omega}_2 = f(\Omega_2) + V_1 + V_2 \tag{18}$$

3.3 Stability. Stability of the inverse system is proven here. From the derivation of the inverse, it follows that the inverse system above is BIBO stable [15].

It is relatively easy to consider asymptotic stability of the inverse system as a high gain system by setting

$$O_2 = 0.$$

We show here that this approximate inverse system is asymptotically stable. The block diagram of the rotational system and its inverse is shown in Fig. 2. Let the Lyapunov function for the inverse system be the sum of the kinetic and elastic energy of the system

$$v = 0.5\Omega_2' J\Omega_2 + 0.5(\Theta_2)' K(\Theta_2). \tag{19}$$

It is easy to show that the derivative of the Lyapunov function is

$$\dot{v} = -\Omega_2' L\Omega_2 \tag{20}$$

This derivative is negative semidefinite, but a nonzero Θ cannot be a solution to the inverse system equation. Therefore by invoking La Salle's Theorem [28], the global stability of the inverse system is assured, subject to the limited range of Θ due to the Lipschitz condition.

3.4 Approximate Inverses. Due to the large choices for O_1 and O_2, one can develop a range of inverses. Of particular interest are those when a fraction of O_2 is used in the feedback path, namely, an amplifier with a small gain precedes or follows the operator O_2.

Assume a particular O_1 is chosen (see examples later). Let ϵ be a number

$$0 \leqslant \epsilon \leqslant 1$$

and let

$$E = \epsilon O_2$$

It can be shown that the system that is inverted is $H-(1-\epsilon)O_2$. This means, for a given H and given ϵ the operator O_1 can be selected sufficiently large enough and, consequently, the operator O_2 sufficiently small enough for the error in V to be acceptable. The computational detail of how to carry this step out are not presented here.

Another bound can be established for the steady state system state error. Assuming that the gains K and L are sufficiently high so that $U=V$, and letting E_1 and E_2 be, respectively, the errors in angle Θ and angular frequency Ω, it follows that

$$V_1 + V_2 = B'KE_1 + LE_2 \tag{21}$$

From this equation, approximate upper bounds can be established for the errors

$$|E_1| = < K^{-1}|V_1 + V_2|,$$

and

$$|E_2| = < L^{-1}|V_1 + V_2|$$

This means that the higher the gains K and L are, the smaller the error. Of course, higher values for the gains demand higher sampling rate, and therefore more on-line computations.

4 The Inverted Pendulum

Next, a rigid body is considered [23] as shown in Fig. 3. It is assumed that a point on the body with coordinates

$$R = [0,0,-r]$$

in the body coordinate system is attached to the origin of the ics, and that the vector of constraint force at the point of attachment, expressed in the ics is Λ. A system of muscle-like actuators connect the rigid body to a weightless foot. One actuator is shown in Fig. 3. Let the resultant moment of all the actuator forces, relative to the contact point, as shall be described later, be vector N in the bcs. Let the sum of the actuator forces, in the ics be H. The rotational motion of the rigid body about its center of gravity is governed by

$$\dot{\Theta} = B(\Theta)\Omega$$
$$J\dot{\Omega} = f(\Omega) + N + \breve{\mathcal{R}}A'\Lambda \tag{22}$$

The translational motion of the center of gravity is governed by

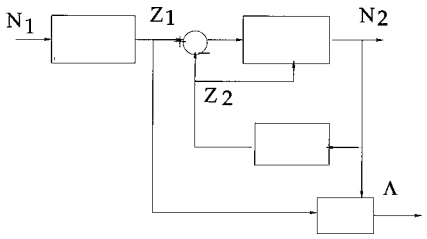


Fig. 4 The inverted pendulum with input N_1 and output Z_1 (Eq. (22)) in series with the nonlinear inverse with high gain feed forward path and a model of the system in the feedback path. The estimate of total input torque is the output N_2. The ground reaction force Λ (Eq. (23)) is constructed from Z_1 and N_t.

$$m\dot{V} = G + H + \Lambda$$

where G is the gravity vector in ics. It can also be shown [1] that the latter equation can be rewritten, involving variables Λ and $\dot{\Omega}$

$$\Lambda + G + H = -mA(\check{\mathcal{R}}\check{\Omega})\Omega + mA\check{\mathcal{R}}\dot{\Omega} \tag{23}$$

Equations (22) and (23) can be simultaneously solved for $\dot{\Omega}$ and Λ. The solution for $\dot{\Omega}$ is the state space equation for the pendulum relative to a body coordinate system whose origin is at the base, and whose axes are parallel with the principal axes.

$$\dot{\Theta} = B(\Theta)\Omega$$
$$J_b\dot{\Omega} = [f_b(\Omega) + N - \check{\mathcal{R}}A'G] \tag{24}$$

Here J_b is the moment of inertia matrix relative to a (principal axes) body coordinate system centered at the point of attachment or contact; and

$$f_b = \check{\Omega}J_b\Omega$$

The solution for Λ is the force of constraint as a function of inputs, i.e., N and gravity, and the state(i.e., Θ and Ω). For ease of reference the latter two equations are symbolically written as

$$\dot{\Omega} = O_3(\Theta,\Omega,N_t)$$
$$\Lambda = O_4(\Theta,\Omega,N_t) \tag{25}$$

where N_t is the total applied torque to the pendulum, including stabilizing torques, actuating torques, gravity, etc.

4.1 The Inverse System. The problem of the inverse system, defined before, is slightly generalized here: Given the quantities Θ, and Ω, estimate N_t and Λ. We assume again that $H=0$. This means that the resultant torque of these forces, namely, N is the same whether it is relative to the center of gravity of the inverted pendulum or the point of contact.

Suppose the same inverse system as in Fig. 2 is utilized, it follows that

$$N_t = N_2$$
$$\Lambda = O_4(\Theta,\Omega,N_t) \tag{26}$$

It is clear from Eq. (23) that Λ can be computed from the state Z_1, and the derivative of Ω with respect to time. This alternative way to arrive at an expression for Λ is to compute it from Eq. (1)

$$\Lambda = m\dot{V} = -G + m[A\check{\mathcal{R}}\dot{\Omega} - A\check{\Omega}^2R] \tag{27}$$

The structure of the inverse system is shown in Fig. 4.

4.2 Measurement of N and Λ. It is instructive to consider direct measurement of the torque N and force Λ. For this purpose, an instrumented platform is needed on which the inverted pendulum is installed. The instruments measure the total torque and force applied by the rigid body to the platform. Here, we consider the stationery case. For a discussion of the dynamic case, see [5]. For a detailed exposition of the planar case of the inverse system in a multilink biped in the sagittal plane see [7,8]. This means the platform instrumentation avails N_1 and Λ. Let us assume that the pendulum is attached to the ground by a massless and stationery foot. Further, if the foot is not attached rigidly to the ground, we assume that the friction forces are large enough, (i.e., larger than the constraint forces) so that the foot does not move or slide. Suppose there are six actuators connecting the inverted pendulum to the foot—each agonist-antagonist pair providing control for the roll, pitch and yaw movements. The arm of the actuator force, shown in Fig. 3 is vector r, expressed in the bcs that extends from the contact point to the origin of the actuator, i.e., point O. Suppose the force of this actuator, namely, vector F, is expressed in the ics. The torque of this actuator, relative to the point of contact, expressed in bcs, is

$$\check{R}A'F$$

The vector N is the sum of the six torques. One needs the points of origin and insertion of the six actuators in ics, in order to arrive at the direction of these forces, and one needs the insertion coordinates and the contact point coordinates, in the bcs, in order to arrive at the R vectors. The massless foot, in turn is acted upon by $-N$, expressed for convenience, in the ics. Suppose, the interaction between this foot and the ground is by a set of n discrete three-vector ground reaction forces: τ_i where index i runs from 0 to n. These forces act on the bottom of the foot, respectively, at vectors s_i in the ics. Let the ankle joint apply a single force of constraint Λ to the massless foot. The equilibrium of the foot results in the sum of the ground reaction forces being equal to Λ plus the sum of the actuator forces; and the moment of all the ground reaction forces relative to the point of contact, i.e., the origin of the ics, being equal to N, expressed in the ics. When the sum of the actuator forces is equal to zero: $H=0$, the sum of the ground reaction forces under the foot is exactly equal to Λ.

Now, from the instrumentation on the platform, the sum of the ground reaction forces under the foot and the moment of all these reaction forces relative to the point of contact can be computed. This means the quantities Λ and N can be measured from the instrumentation on the platform.

The main point of the discussion here is that the inverse of the inverted pendulum is derived with the dynamics of the inverted pendulum formulated about the point of contact. The inverse system estimates the total torque applied to the rigid body. To arrive at the applied torque, the contribution of the torque of the constraint force Λ must be subtracted.

5 Simulation Results

Several computer simulations are presented here in order to demonstrate the efficiency, viability, and feasibility of this formulation. The first simulation deals with a body rotating about its center of gravity. The numerical parameters and gains are listed in the Appendix .

5.1 Rotating Rigid Body. The rotational rigid body motion about its center of gravity is considered. First, state feedback of the position Euler angles and angular velocity vector, i.e., Ω is used from Sec. 2.5 to stabilize the motion. The stabilizing position and velocity gains are the $3by3$ diagonal matrices: $K_1 = \mathrm{Diag}(400,400,320)$, and $L_1 = \mathrm{Diag}(60,60,42)$.

Three impulsive torques are assumed to have been applied to this rigid body in order to set up the following initial conditions:

$$[0,0,0,2.5,-2.5,2.5]$$

With these initial states the behavior of the system is simulated for about 0.1 s. The vector of total torque acting on the body and the state are computed and also recorded. This transient behavior of

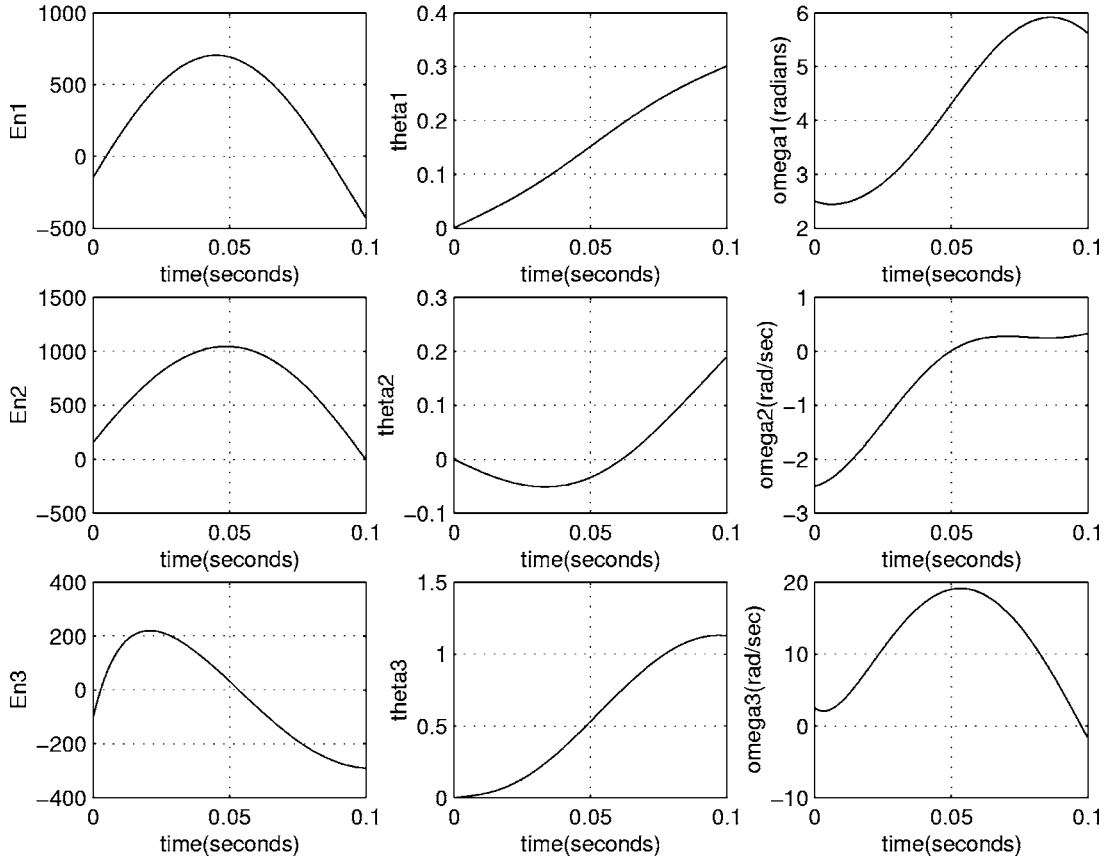


Fig. 5 The input torque *En*, the angles, and angular velocities as functions of time for one rigid body anchored at its center of gravity

the stabilized rigid body is documented in Fig. 4. The figure shows the input torques, the angles Θ, and angular velocities Ω, all as functions of time. The angles and angular velocities are now input to the inverse system. It is assumed that the physical parameters of the inverse system are the same as those of the rotating rigid body. The gains in the forward loop, namely, K and G in Eq. (14) are chosen to be the same diagonal 3×3 matrix, i.e.,

$$K = \mathrm{Diag}(10000, 5000, 500)$$

and

$$G = \mathrm{Diag}(10000, 5000, 500)$$

The output of the inverse system is given in Fig. 5, along with the error signals in Θ and Ω. The comparison of the actual and the estimated torques in the previous two figures show that the inverse system is very good in tracking the torques. There are initial errors in all three torques because, for the inverse system, the initial states and outputs are all zero. However, as can be seen in the figure, the tracking of the torques is very good after about 0.004–0.008 s.

5.2 The Inverted Pendulum. In this simulation, the rigid body is anchored to the ground at a point with coordinates R as an inverted pendulum. The initial conditions are all zero. It is subjected to stabilizing feedback torques as before. These feedback gains are given by the diagonal 3×3 matrices: $K_2 = \mathrm{Diag}(1200, 1200, 960)$, and $L_2 = \mathrm{Diag}(180, 180, 126)$.

In addition, there are periodic torques acting at the base of the pendulum for one second. The three components of the torque vector are equal

$$n = 1000\sin(10\pi t).$$

The input torques and the ground reaction force are to be the desired quantities to be estimated. They are recorded and plotted in Fig. 6. The angles and angular velocities of the inverted pendulum are the outputs of the system and are plotted in Fig. 7. It can be seen that the system assumes a steady state in less than 0.2 s. The structure of the inverse is taken as that of Fig. 8. The gains in the forward loop, namely, K and G in Eq. (14) are chosen to be the same diagonal 3×3 matrix, i.e.,

$$K = \mathrm{Diag}(17232, 15232, 200)$$

and

$$G = \mathrm{Diag}(17232, 15232, 200)$$

The inputs to this inverse system are the position angles and the angular velocities recorded as outputs of the previous simulation. With these inputs and zero initial conditions, the inverse system is simulated for one second. The outputs of the inverse system are the ground reaction forces and the total torques acting on the system. These estimated total torques and estimated ground reaction force are shown in Fig. 9. The state trajectories of the inverse system are shown in Fig. 10.

6 Conclusions

We have considered dynamics, inverse, and control of a single rigid body system with constraints. We have developed a methodology to construct Lyapunov stable inverses for these systems. The procedures here can be extended to multi-body systems that are better approximations to systems involved in medicine, robotics and human movement.

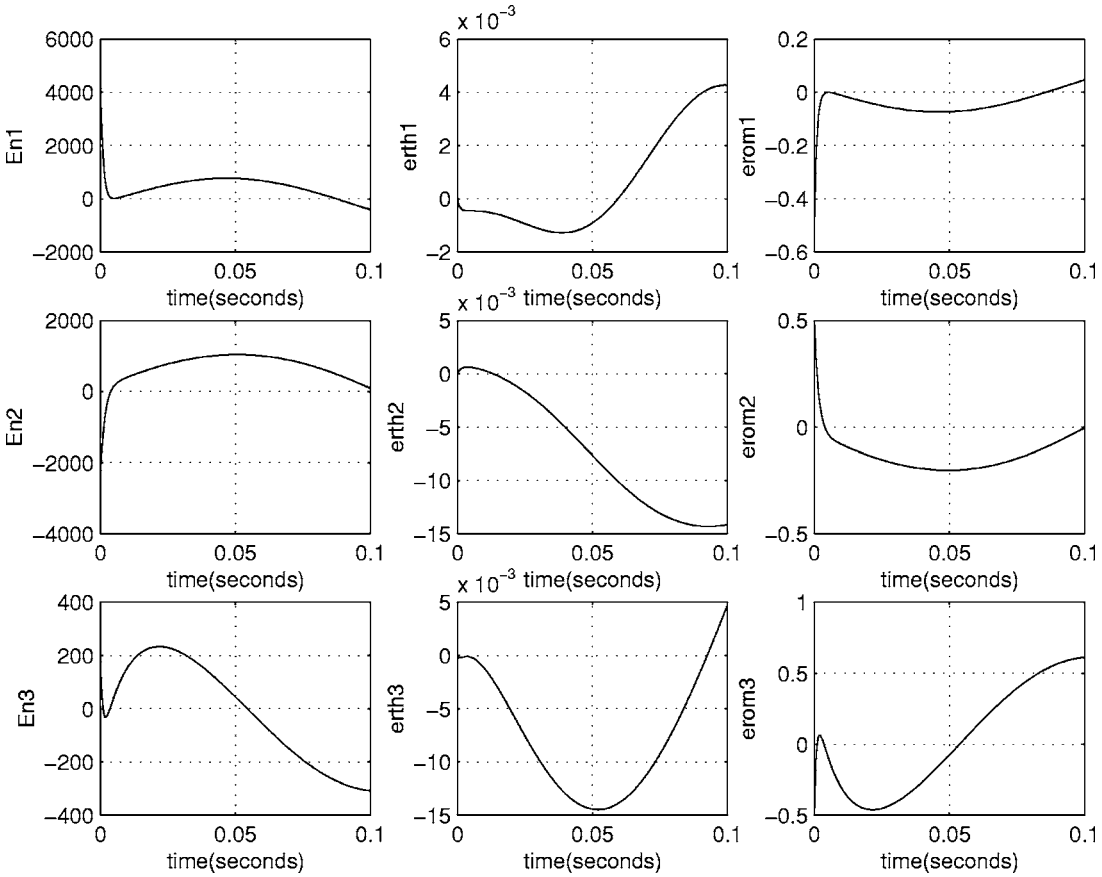


Fig. 6 The estimated torques as outputs of the inverse system as functions of time. The error signals in Θ and Ω, as inputs to the high gain forward component of the inverse are also plotted as functions of time.

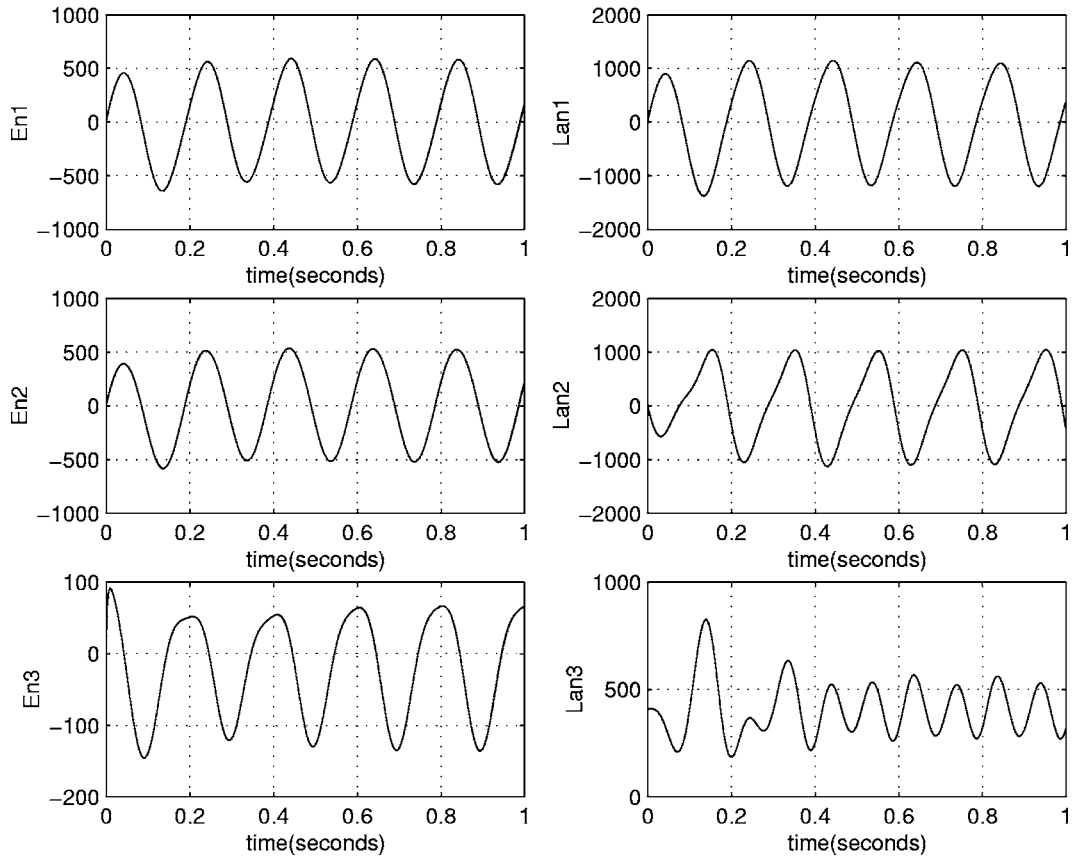


Fig. 7 The total applied torques to the inverted pendulum, and the resulting ground reaction force as functions of time

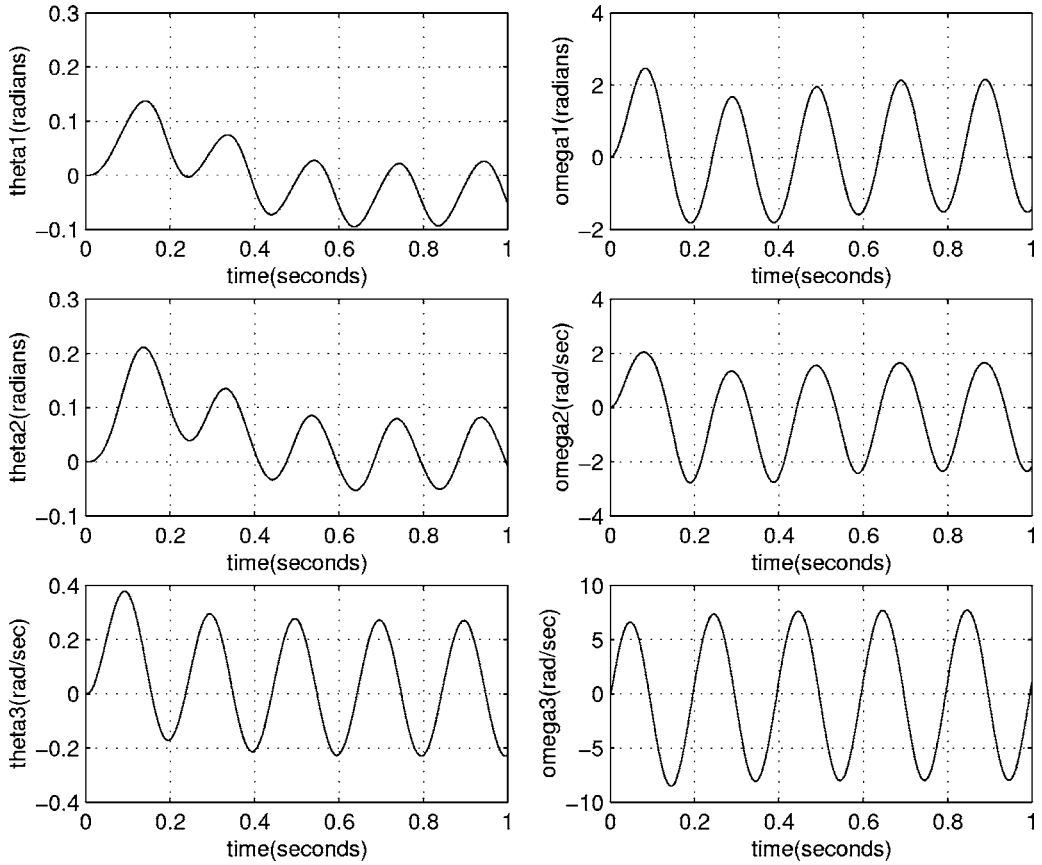


Fig. 8 The state trajectories of the inverted pendulum: Θ and Ω as functions of time

The computational methods pose interesting challenges for multilink three-dimensional systems:

1. If the trajectories of the state of the system (i.e., position and velocity variables) and either the input torques or the accelerations are known, the ground reaction and joint forces should be computable as shown in Ref. [2] for very simple cases.
2. The computation of input torques to the multilink system, in order for the system to follow a desired trajectory, or the estimation of such torques, when the trajectories are known, can be implemented with inverses.
3. In certain instances, both the trajectories and the constraint forces (ground reaction forces are an example) are specified. This class of problems has not been studied from a theoretical point of view. Some heuristic methods have been applied before [29]. The problem is tractable if there are a sufficient number of inputs to control the forces and the trajectories.

We have shown relevance of functional analysis, inverse system theory, and high gain systems to formulating and solving the above problems.

Acknowledgment

The authors would like to thank the reviewers for their careful reading of the manuscript, and for their valuable and constructive criticism.

Appendix

1 Single Rigid Body

The inertial coordinate system is defined as follows. The x_1 axis is to the front of a biped, or the front of an airplane —the longitudinal axis. The x_2 coordinate is in the direction of the extended left hand of the biped or the pitch axis of the airplane. The third axis x_3 is vertically upward, the yaw axis in an airplane. The Euler angle sequence corresponds to roll, pitch, and yaw.

Let A be the 3×3 orthonormal matrix that transforms vectors from the bcs to ics. The inverse of A is A'. The matrices $A(\Theta)$, $B(\Theta)$, and $\breve{\mathcal{R}}$ are given below.

Let $A_1(\theta_1)$, $A_2(\theta_2)$, and $A_3(\theta_3)$ be defined by

$$A_1(\theta_1)=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta_1 & -\sin\theta_1\\ 0 & \sin\theta_1 & \cos\theta_1\end{bmatrix}$$

$$A_2(\theta_2)=\begin{bmatrix}\cos\theta_2 & 0 & \sin\theta_2\\ 0 & 1 & 0\\ -\sin\theta_2 & 0 & \cos\theta_2\end{bmatrix}$$

$$A_3(\theta_3)=\begin{bmatrix}\cos\theta_3 & -\sin\theta_3 & 0\\ \sin\theta_3 & \cos\theta_3 & 0\\ 0 & 0 & 1\end{bmatrix}$$

Now $A(\Theta)$ can be defined

$$A(\Theta)=A_1(\theta_1)A_2(\theta_2)A_3(\theta_3)$$

The matrix $B(\Theta)$ is given by

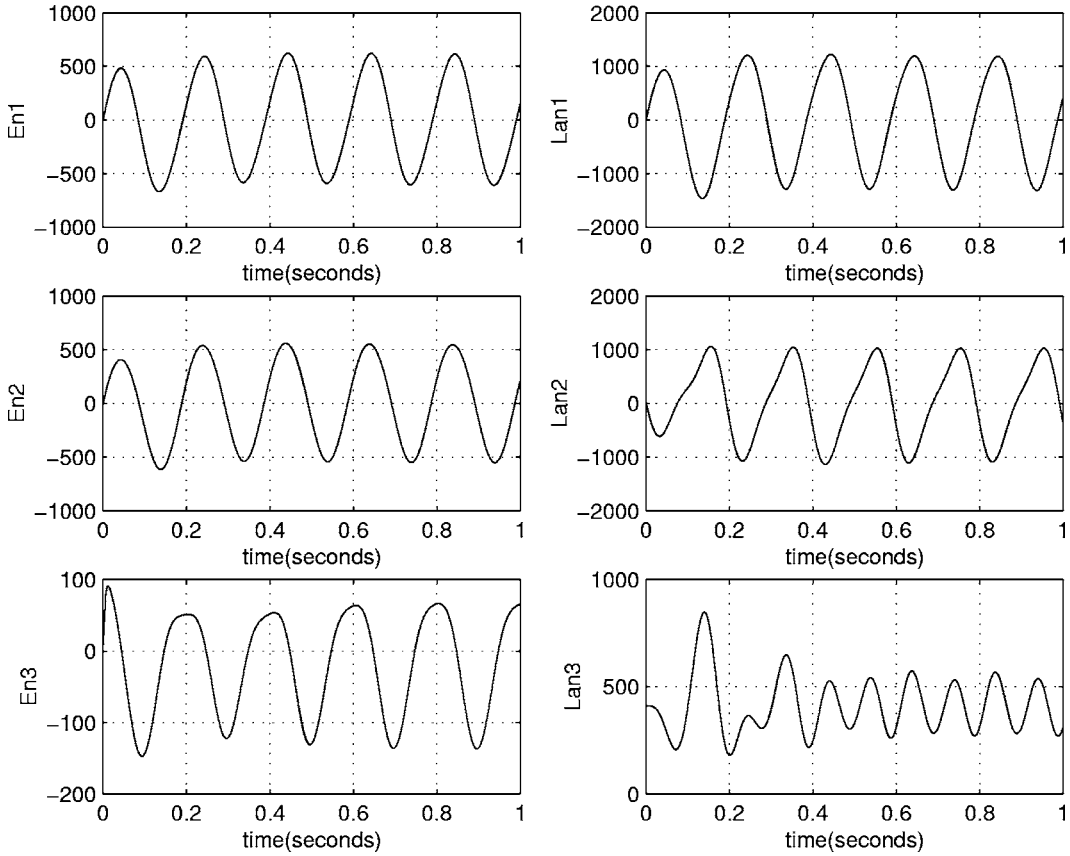


Fig. 9 The estimated total torque applied to the inverted pendulum, the estimated ground reaction forces as functions of time

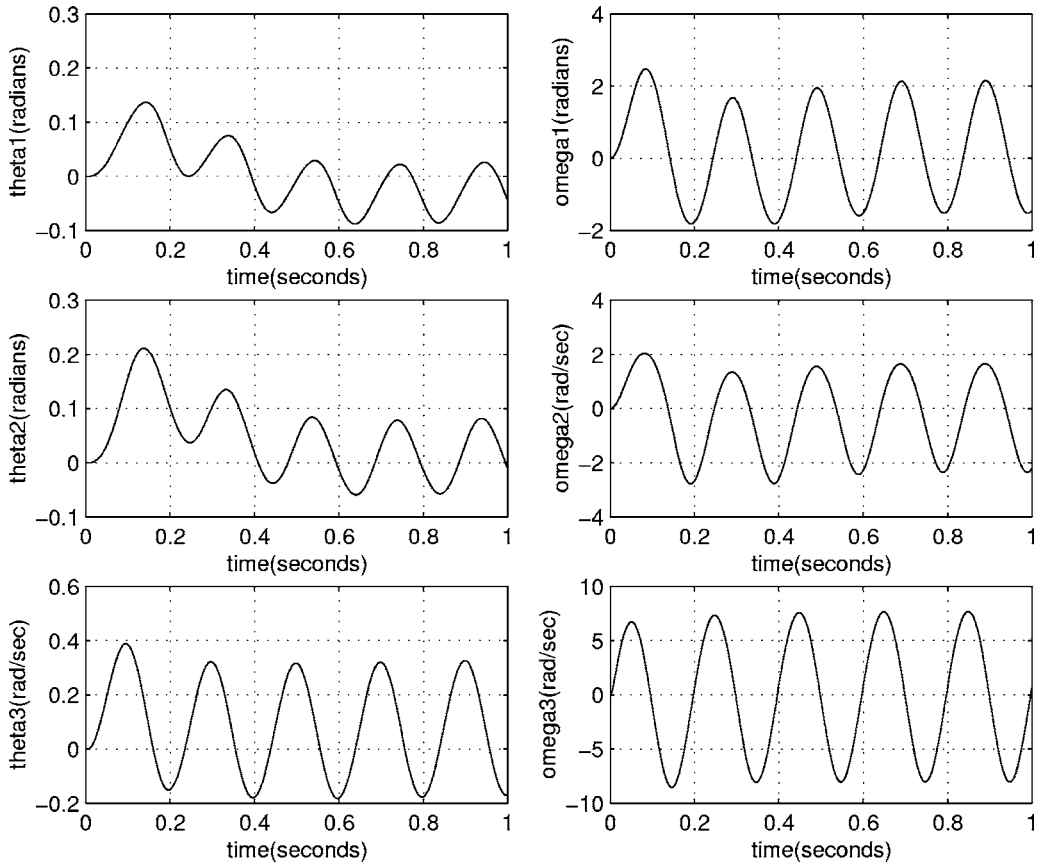


Fig. 10 The state of the estimator as a function of time

Table 1 Definition, symbols, and numerical values for one rigid body

	Symbol	Value	Unit
Mass.	m	41.00	Kg
Principal moment of inertia	j1	10.0	Kg m^2
Principal moment of inertia	j2	8.0	Kg m^2
Principal moment of inertia.	j3	0.4	Kg m^2
Center of gravity	l	0.42	m
Gravity constant	g	10.0	m/s^2

$$B(\Theta) = \begin{bmatrix} \dfrac{\cos\theta_3}{\cos\theta_2} & \dfrac{-\sin\theta_3}{\cos\theta_2} & 0 \\ \sin\theta_3 & \cos\theta_3 & 0 \\ \dfrac{-\sin\theta_2\cos\theta_3}{\cos\theta_2} & \dfrac{\sin\theta_2\sin\theta_3}{\cos\theta_2} & 1 \end{bmatrix}$$

It is to be observed that matrix B is of finite norm only for a limited range of θ_2, i.e., θ_2 must lie in the open interval: $-\pi/2$ and $+\pi/2$. This is the range for which the Lipschitz condition holds. For θ1 and θ_3, respectively, corresponding to roll and yaw, the range is from $-\pi$ to $+\pi$- corresponding to the range for the physical three-dimensional world. All the development and results in this paper are limited to these ranges.

Let vector R have components r_1, r_2, and r_3. The skew symmetric matrix $\breve{\mathcal{R}}$ is defined as

$$\breve{\mathcal{R}} = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix}.$$

2 Numerical Values

The definitions, symbols, and numerical values for a single rigid body system [23,30] are given in Table 1.

References

[1] Hemami, H., and Utkin, I., 2002, "On the Dynamics and Iyapunov Stability of Constrained and Embedded Rigid Bodies," (unpublished).

[2] Hemami, H., and Wyman, B., 1979, "Modeling and Control of Constrained Dynamic Systems With Application to Biped Locomotion in the Frontal Plane," IEEE Trans. Autom. Control, **24**, pp. 526–535.

[3] Hemami, H., and Wyman, B., 1980, "Indirect Control of the Forces of Constraint in Dynamic Systems," J. Dyn. Syst., Meas., Control, **24**(4), pp. 355–360.

[4] Winter, D., 1990, *Biomechanics and Motor Control of Human Movement*, Wiley, New York.

[5] Goswami, A., 1999, "Foot Rotation Indicator(fri) Point: A New Gait Planning Tool to Evaluate Postural Stability of Biped Robots," *Proc. ICRA 1999*, Detroit, Vol. 1, IEEE Robotic Society, New York, pp. 47–52.

[6] Dariush, B., Hemami, H., and Parnianpour, M., 2000, "Analysis and Synthesis of Human Motion From External Measurements," *Proc. ICRA 2000*, San Francisco, IEEE Robotic Society, New York, pp. 4015–4020.

[7] Dariush, B., Hemami, H., and Parnianpour, M., 2000 "A Well-Posed, Embedded Constraint Representation of Joint Movement From Kinesiological Measurements," J. Biomech. Eng., **122**, pp. 437–445.

[8] Dariush, B., Hemami, H., and Parnianpour, M., 2001, "Multi-Modal Analysis of Human Motion From External Measurements," J. Dyn. Syst., Meas., Control, **123**, pp. 272–278.

[9] McCuskey, S., 1959, *An Introduction to Advanced Dynamics*, Addison-Wesley.

[10] George, D., 1959, "Continuous Nonlinear Systems," Tech. Rep. 355, Research Laboratory of Electronics, Mass. Inst. of Technology, Cambridge, MA.

[11] Zames, G., 1960, "Nonlinear Operators for System Analysis," Tech. Rep. 370, Research Laboratory of Electronics, Mass. Inst. of Technology, Cambridge, MA.

[12] Smith, O., 1958, *Feedback Control Systems*, McGraw-Hill, New York.

[13] Astrom, K. J., and Wittenmark, B., 1989, *Adaptive Control*, Addison-Wesley.

[14] Kokotovic, P., Khalil, H. K., and O'ReillyJ., 1986, *Singular Perturbation Methods in Control*, Academic Press, New York.

[15] Khalil, H., 1996, *Nonlinear Systems*, Prentice-Hall, Englewood Cliffs, N. J.

[16] Barin, K., and Stockwell, C., 1986, "Parameter Estimation of a Model of Human Postural Control," *Proc. of the Eigth Annual Conference of the IEEE/Engineering in Medicine and Biology Society*, Forth Worth, Texas, Nov 7-10, 1986, IEEE, New York, pp. 1571–1574.

[17] Barin, K., 1989, "Evaluation of a Generalized Model of Human Postural Dynamics and Control in the Sagittal Plane," Biol. Cybern. **61**, pp. 37–50.

[18] Horak, F., and Nashner, L., 1986, "Central Programming of Postural Movements:Adaptation to Altered Support-Surface Configurations," J. Neurophysiol., **55**, pp. 1369–1381.

[19] Brooks, V., 1986, *The Neural Basis of Motor Control*, Oxford University Press, New York.

[20] Nashner, L., 1981, "Stance Posture in Humans," *Handbook of Behavioral Neurobiology*, A. L. Towe and E. Luschi, eds., Plenum Press, New York, pp. 527–566.

[21] Bloch, A., 1994, *Hamiltonian and Gradient Flows, Algorithms and Control*, American Mathematical Society, Providence, Rhode Island.

[22] Bullo, F., and Lewis, A., 2005, *Geometric Control of Mechanical Systems*. Springer, New York.

[23] Hemami, H., and Katbab, A., 1982, "Constrained Inverted Pendulum for Evaluating Upright Stability," J. Dyn. Syst., Meas., Control, **104**, pp. 343–349.

[24] Isidori, A., 1989, *Nonlinear Control Systems*. Springer-Verlag, Berlin.

[25] Hemami, H., 1982, "A State Space Model for Interconnected Rigid Bodies," IEEE Trans. Autom. Control, **27**, pp. 376–382.

[26] Hemami, H., 2002, "A General Framework for Rigid Body Dynamics, Stability and Control," J. Dyn. Syst., Meas., Control, **27** pp. 241–251.

[27] Krstic, M., Kanellakopoulos, I., and Kokotovic, P., 1995, *Nonlinear Adaptive Control Design*, Wiley, New York.

[28] Atherton, D. P., 1981, *Stability of Nonlinear Systems*, Research Studies Press, Chichester, UK.

[29] Hemami, H., Wongchaisuwat, C., and Brinker, J. L., 1987, "A Heuristic Study of Relegation of Control in Constrained Robotic Systems," ASME J. Dyn. Syst., Meas., Control, **109**, pp. 224–231.

[30] Kim, J., and Hemami, H., 1998, "Coordinated Three-Dimensional Motion of the Head and Torso by Dynamic Neural Networks," IEEE Trans. Syst., Man, Cybern., Part B: Cybern. **28**(5), pp. 653–666.

Craig A. Steeves
Research Associate
Department of Mechanical and Aerospace Engineering,
Princeton University,
Engineering Quad,
Olden Street,
Princeton, NJ 08544

Haydn N. G. Wadley
Professor
Department of Materials Science and Engineering,
University of Virginia,
116 Engineer's Way,
P.O. Box 499745,
Charlottesville, VA 22904

Richard B. Miles
Professor
Department of Mechanical and Aerospace Engineering,
Princeton University,
Engineering Quad,
Olden Street,
Princeton, NJ 08544

Anthony G. Evans
Professor
Department of Mechanical Engineering,
University of California,
Engineering II,
Room 2361A,
Santa Barbara, CA 93106

A Magnetohydrodynamic Power Panel for Space Reentry Vehicles

During reentry from space, a layer of high temperature air (>3000 K) *is formed extending tens of centimeters from the surface of the vehicle, well out into the high speed flow regime. Magnetohydrodynamics (MHD) can then be used to generate power by projecting magnetic fields outside the vehicle into the conducting air stream and collecting the resulting current. Here, we analyze a multifunctional MHD panel which generates the requisite magnetic fields, protects the vehicle from high temperatures, and is structurally stiff and strong. The analysis shows that a magnetic system weighing approximately* 110 kg *can generate* 0.6 MW *of power for* 1000 s. [DOI: 10.1115/1.2178360]

1 Introduction

This paper explores the potential for generating large amounts of electrical power during reentry from space by exploiting magnetohydrodynamics (MHD). While the analysis below will demonstrate that there are several unresolved design issues that must be addressed before MHD generation will be usable, the authors believe that these issues are soluble given current technology. The concept envisages a vehicle whose skin is a multifunctional structure that generates magnetic fields of sufficient strength to produce useful power, while simultaneously sustaining the high heat flux, as well as the aerodynamic and structural loads, all at acceptable mass. The analysis is performed at the subelement (or panel) level. The input for design of the panel emanates from aerothermochemistry calculations described elsewhere [1], which are based upon a wedge-shaped vehicle configuration [2,3].

Magnetohydrodynamic power generation is described by Faraday's law. Since the ionized gas flowing over the surface of the vehicle is conductive, a magnetic field extending from the vehicle into this plasma is able to accelerate charged particles (as shown in Fig. 1) creating a transverse current density $\mathbf{J}=\sigma\mathbf{u}\times\mathbf{B}$ (neglecting, for the moment, both Hall effects and ion slip), where σ is the conductivity of the plasma, $\mathbf{u}$ is the velocity of the plasma, and $\mathbf{B}$ is the local magnetic flux density. For a comprehensive discussion of power generation in a magnetohydrodynamic context, see Rosa [4].

The ensuing analysis of the usable power is based upon the following assumptions:

1. Air at $u=7$ km/s and at densities and temperatures associated with flow around a 12 deg half-angle wedge at 45 km altitude is seeded with NaK (properties as per [5]) imbuing the plasma enshrouding the vehicle with sufficient conductivity to generate ~1 MW of power per m^2 of surface area, provided that the magnetic field just outside the vehicle exceeds 0.2 T. This assumption is based on the aerothermodynamic model results.
2. The surface temperature of the MHD panel reaches a "steady state" of about 1500 K.
3. During the reentry period, which is assumed to last up to $t_f=1000$ s, the temperature at the magnets must remain below a maximum allowable operational limit.

The need to project a significant magnetic field beyond the surface of the vehicle encourages the use of thin walled vehicle structures that differ conceptually from the thick insulating (or ablative) tiles used in current (passive) thermal protection systems (TPS) (see, for a discussion, [6]). This paper analyzes a conceptual system comprised of a thin ceramic thermal barrier coating (TBC) strongly adhered to a metallic substrate and an active cooling system. Because of the thinness of the TBC, the interface between the insulating material and the structural alloy attains

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 10, 2004; final manuscript received January 6, 2006. Review conducted by G. B. A. Younis. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

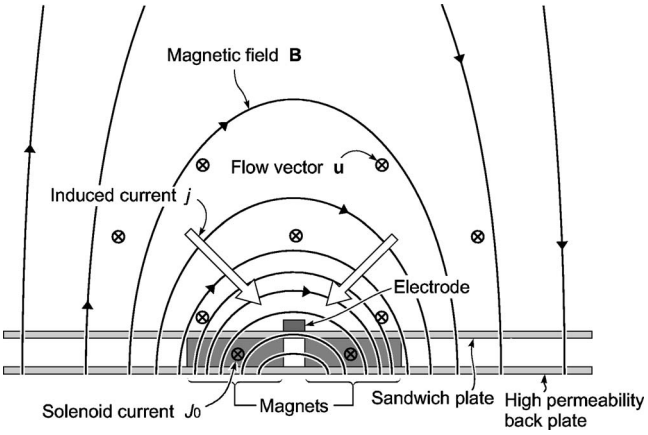


Fig. 1 Relative directions of flow, magnetic field, and generated current

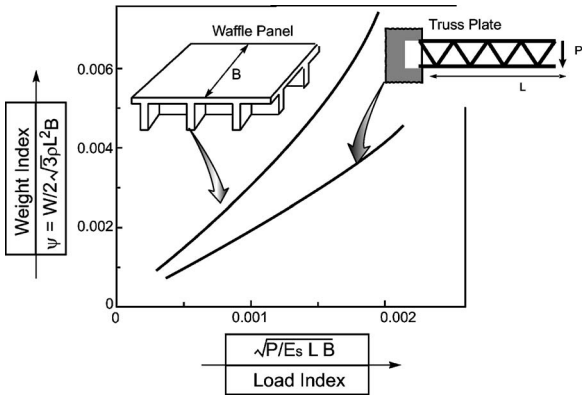


Fig. 2 Mass of truss-core panels as a function of load capacity, compared to solid beams

high temperatures. The thermostructural consequences are analogous to those encountered in gas turbines, which use a TBC concept in conjunction with internal cooling to control the rate of thermomechanical and oxidative damage. For gas turbine applications, a robust multilayer system has been devised with two primary attributes: It satisfies thermodynamic compatibility and is tolerant to thermal and mechanical strain misfits between layers. It consists of an insulating oxide, typically yttria stabilized zirconia (YSZ), deposited as a 100–200 μm thick coating onto an aluminum-rich 20 μm thick intermetallic layer (referred to as a bond coat) that slowly oxidizes to form $\alpha-Al_2O_3$. This concept has been chosen because YSZ and $\alpha-Al_2O_3$ are thermodynamically compatible [7,8]. The bond coat is in turn deposited on a structural nickel-based superalloy, with the constraint that the interdiffusion between the alloy and the bond coat be minimal during the expected life of the system. The alloy is actively cooled by air from the intake, directed through embedded serpentine channels. While the TBC surface temperature can be as high as 1425 K in this application, the underlying materials can be as much as 200 K cooler.

A comparable TBC system is proposed for the MHD panel, but now the structural system should be nonmagnetic, lightweight, and capable of supporting loads at high temperature. These requirements, coupled with a material selection algorithm [9,10], suggest titanium alloys as a primary candidate for the vehicle skin. A bond coat concept for Ti alloys and a suitable strain-tolerant TBC oxide have yet to be devised. For the present assessment we invoke a Ti-based intermetallic bond coat that forms $\alpha-Al_2O_3$ upon oxidation, combined with a ternary oxide with lower thermal conductivity than YSZ and more resistance to sintering at the surface temperatures experienced during reentry. The candidates are, respectively, $TiAl_3$ and $Gd_2Zr_2O_7$ [11]. When deposited as a columnar structure using electron beams, $Gd_2Zr_2O_7$ has a thermal conductivity, $k=0.5$ W/mK [11], but this conductivity can potentially be lowered [12] by manipulating the morphology of the porosity, possibly by as much as a factor of two [13,14].

The TBC would be deposited onto a multifunctional panel that performs four functions: (a) impedes heat transfer into the vehicle; (b) supports an integral cooling system; (c) contains embedded magnets; and (d) withstands large compressive and bending loads. A truss core sandwich design made using a Ti alloy is proposed. When optimized for load capacity, the weight of such panels as a function of the load is summarized in Fig. 2, relative to a solid panel of the same mass per unit area [15,16]. The open channels in the core allow cooling and permit permanent or solenoid magnets to be incorporated, as elaborated later.

This study explores the feasibility of using such a design to generate power during reentry, and will propose a systematic means for design and optimization. The analysis reveals that a MHD panel competes effectively with other power sources in terms of power density and energy density (see Fig. 3).

The paper is organized in the following manner. (a) The design concept is outlined. (b) The power that can be generated is related to magnetic field strength and the flow conditions. (c) Magnetic field strengths are calculated using finite elements. (d) The optimization approach is described. (e) The panels are optimized and trends in net power generated are established. The trends are used to provide a focus for future research that addresses feasibility and validation of the models.

2 Design Concept

The panel concept, illustrated in Fig. 4, embeds either resistive solenoid magnets (shown here) or permanent magnets (not displayed). (Superconducting magnets are discounted at this stage due to unresolved cooling requirements.) The basic panel design is a 50 mm thick titanium sandwich panel, comprising two 5 mm thick faces and a 40 mm thick truss core. The reentry maneuvers determine the maximum structural loads; these are mission-dependent and are not explicitly analyzed. Based on previous assessments of the strength and stiffness of truss core panels [15,17], it is believed that the proposed panel can be designed to sustain the required loads at acceptable weight. Because the core of the panel contains open channels, and because the truss struc-

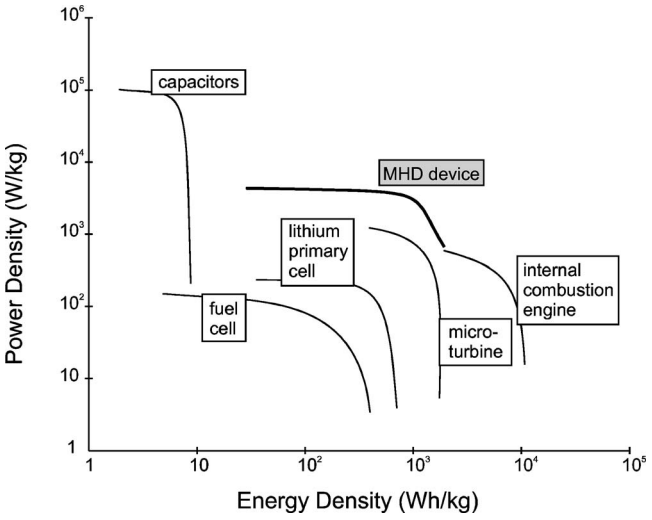


Fig. 3 Ragone plot for various power sources for a reentry vehicle including the proposed MHD device

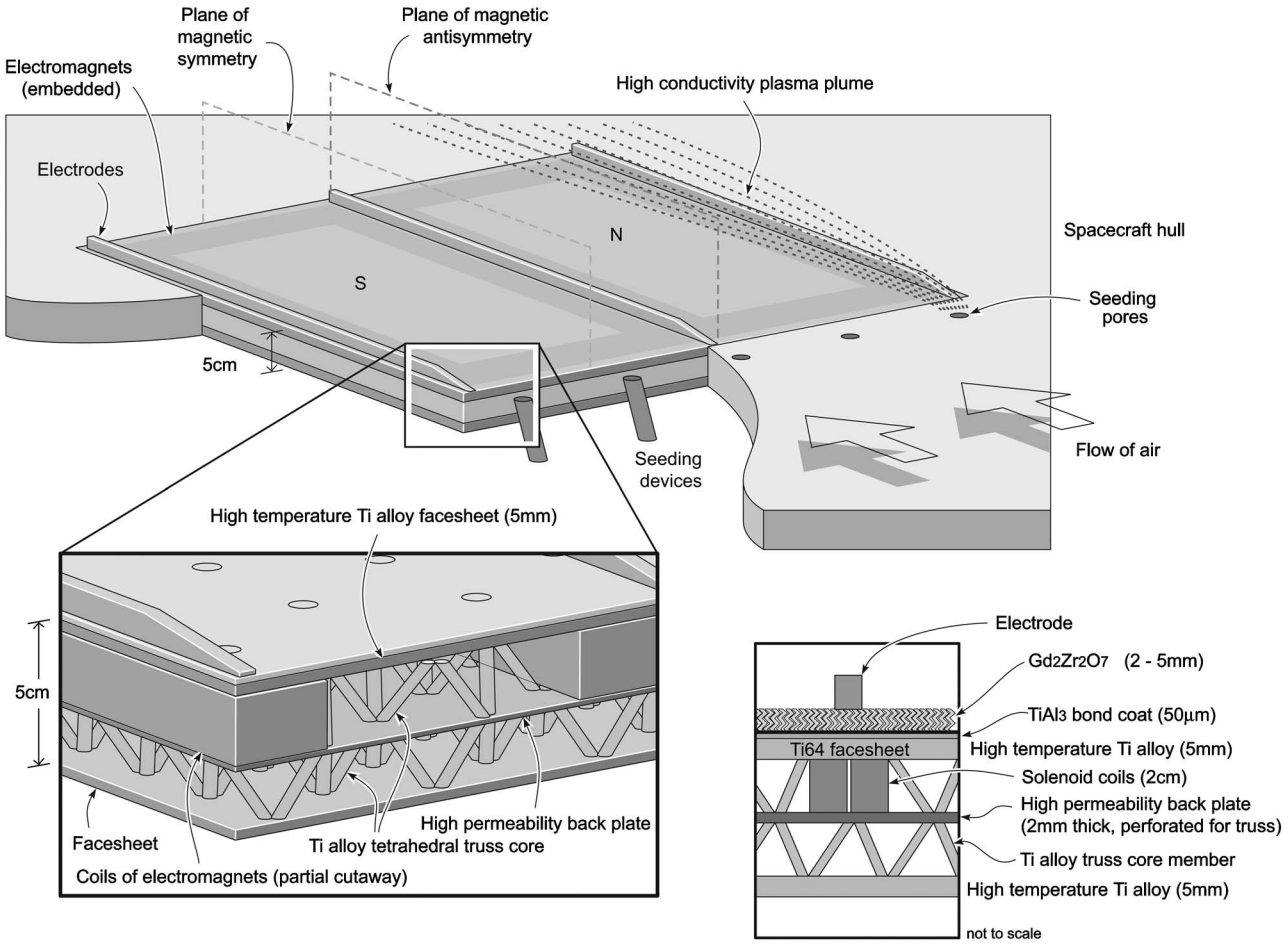


Fig. 4 Array of rectangular solenoids contained in a multifunctional truss-core sandwich structure

tures efficiently conduct heat from the faces, this panel enables simultaneous thermal management by active cooling. The magnetic array is relatively long in the direction of the flow stream, whereupon edge effects at the front and back can be neglected. The array is periodic in the width direction, which increases the magnetic field strengths [18,19]. To realize this potential, adjacent magnets must be oriented with alternating north and south poles, creating flux loops. By placing electrodes at the junctions between the magnets, the current is directed to these electrodes and collected, thereby generating power, as sketched in Fig. 1. Moreover, the symmetry simplifies the calculations required. The magnetic array is separated from the surface of the vehicle by a titanium alloy face and a TBC layer.

For the resistive solenoid design in Fig. 4, the north and south poles of the magnetic elements alternate. This particular array contains two periodic elements. Electrodes for power collection are mounted on the surface of the TBC, interlaced between the magnet poles. The design variables are the thickness h and total width $2w$ of the solenoid, the width of the arms of a single solenoid b, and the current density in the solenoid J_0. Again, it is assumed that the length L of the array is much larger than any other geometric variable, permitting a two-dimensional approximation. Additional variables, such as the geometry of the high-permeability backing plate, will not be addressed. The overall configuration of a permanent magnet array is similar. Comparable fields can be achieved with a variety of alternative configurations.

Thin layers of insulation are used in conjunction with active cooling of the sandwich face. The titanium alloy face layer beneath the TBC uses an embedded planar heat pipe to assure adequate creep strength by preventing the temperature from exceeding 900 K. Active cooling protects the upper surfaces of the magnets, preventing them from exceeding 500 K, their maximum operational temperature. This design uses water, which is heated from 273 K to 373 K, evaporated, and expelled from the vehicle. There are several precedents for a cooling system of this type, including Faghri [20]. In such a system, each kilogram of water would expel 2.8 MJ of heat from the vehicle. The mass of the cooling system, m_{cool} is considered to be that of the water m_w, plus a re-circulation system ϵm_w such that $m_{\text{cool}}=(1+\epsilon)m_w$. Active cooling is required to dissipate the heat associated with the operation of resistive solenoids.

3 Estimation of Power Generated

The power P_0 delivered to the load, per unit volume of flow over the vehicle, is

$$P_0 = K(1-K)\tilde{\sigma}u^2B^2 \tag{1}$$

where K is the load factor which is between zero and one, $\tilde{\sigma}$ is an effective conductivity of the gas which accounts for Hall effects, u is the magnitude of the velocity of the flow stream, and B is the norm of magnetic flux density; $B=\|\mathbf{B}\|$. The velocity $\mathbf{u}$ is perpendicular to the magnetic field $\mathbf{B}$. The maximum power that can be extracted from the flow occurs when the load impedance equals the resistance of the current passing through the air, and $K=0.5$. The use of Eq. (1), which is derived for a one-dimensional magnetic field, can be justified because, for the magnetic fields examined here, at all locations the vector product of the fluid velocity and the magnetic field is directed toward the electrodes. Ion slip effects have been shown to be negligible at 46 km altitude by

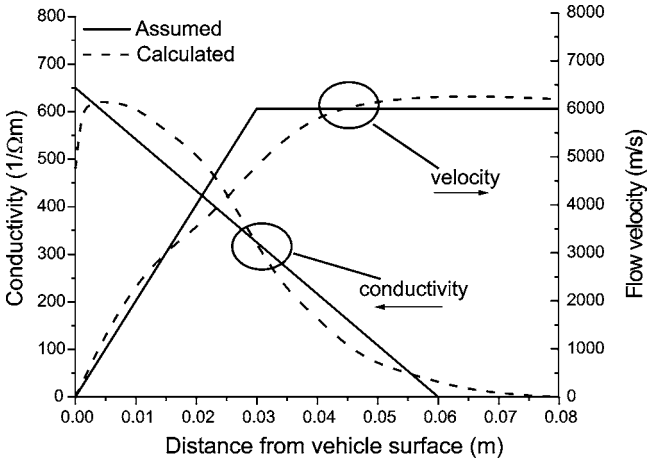


Fig. 5 Comparison of assumed and calculated conductivity and velocity profiles with distance from vehicle surface

Macheret et al. [1]. The remaining elements of Eq. (1) are addressed below.

3.1 Plasma Conductivity. The conductivity is a function of the density and composition of the flow stream, the configuration of the shock system, the gas temperature, and the presence or absence of artificial seeding. We assume that the conductivity arises principally because of artificial seeding, and that the seeding device projects the seed 0.06 m from the vehicle surface. Calculation of conductivities requires a multispecies nonequilibrium flow model, which is beyond the scope of this paper but has been performed by [1]. The seeding model used by [1] assumed 1% by mass NaK entrained in the flow ahead of the MHD region; during a 1000 s flight, this would require only 1–2 kg of seed plus some additional mass for an injection system which would be similar to fuel injection systems for ramjet engines. Because our purpose is to compare magnetic arrays, we simplify by regarding the conductivity as varying linearly within the region of interest

$$\sigma(z) = \sigma_0[(z_0 - z)/z_0], \tag{2}$$

where $\sigma_0 = 650/\Omega\text{m}$ and $z_0 = 0.06$ m. This is consistent with gas properties at an altitude of approximately 45 km, and reflects the distribution of temperature and seed material through the boundary layer in the calculations by Macheret et al. [1]. A comparison of the assumed distribution of conductivity with the distribution calculated by Macheret et al. [1] is shown in Fig. 5.

At high magnetic fields and low collision frequencies, the electrons circulate around the magnetic field lines rather than migrate between the electrodes leading to a Hall current and a significant reduction of the effective conductivity. If the Hall parameter Ω_e (the ratio of the cyclotron frequency of the electrons to the collision frequency) is greater than one, the Hall effect must be taken into account. In this case, the effective conductivity $\tilde{\sigma}$ is written

$$\tilde{\sigma} = \frac{\sigma}{1 + \Omega_e^2} = \frac{\sigma}{1 + (\mu_e B)^2} \tag{3}$$

where μ_e is the electron mobility (with $\mu_e B$ an alternate representation of the Hall parameter). Assuming an electron mobility $\mu_e = 10/\text{T}$, the Hall parameter is unity when flux density $B = 0.1$ T. Operation at magnetic field strengths much above 0.2 T does not lead to a significant increase in power extraction without the introduction of complex electrode segmentation and the potential of interelectrode voltage breakdown. In this analysis it is assumed that continuous Faraday electrodes are used. It should also be noted that 0.2 T is close to the limit of the magnetic field strength that can be generated from reasonable steady state electromagnet configurations.

3.2 Flow Velocity. Various assumptions need to be made about the flow over the vehicle.

1. The velocity increases linearly with distance z from the vehicle surface, starting at zero and reaching the velocity in the shock layer u_0 at the edge of the boundary layer. Outside the boundary layer, the velocity is constant and equal to the shock layer velocity.
2. The thickness of the boundary layer, $t_{\text{BL}} = 0.03$ m [1].
3. The magnetic Reynolds number is small, and hence the flow does not signifiacntly perturb the magnetic field.

The resulting velocity function is

$$u(z) = \begin{cases} u_0\left(\dfrac{z}{t_{BL}}\right) & \text{if } z \leqslant t_{BL} \\ u_0 & \text{if } z > t_{BL} \end{cases} \tag{4}$$

This velocity function is compared with detailed aerothermodynamic calculations [22] in Fig. 5.

The power generated by the MHD device, per unit volume, thus becomes

$$P_0 = \frac{\sigma(z) u(z)^2 B^2}{4[1 + (10B)^2]}, \tag{5}$$

where $u(z)$ is given by Eq. (4), $\sigma(z)$ is given by Eq. (2), and the magnetic flux density B is determined as detailed below. This function is integrated over the width of a magnetic element and the thickness of the conductive layer in order to determine the power generated, per unit length, by a single periodic magnetic element.

4 Magnetic Field Calculations

Finite element calculations performed with FEMLAB [23] have been used to develop a relationship for the dependence of power generation on the geometry and composition of the magnetic array. These computations provide the relationship between geometric variables (thickness h, element width $2w$, and solenoid arm width b), internal variables (current density J_0 or magnetization M_0), and the magnetic fields. The finite element results generate a database of magnetic fields as a substitute for an explicit functional relationship between the power generated and the geometric and array parameters.

An example is provided for the electromagnetic arrays. Because of the magnetic symmetries and antisymmetries and the assumption of periodicity, as shown in Fig. 4, only half of one periodic element needs to be modeled. Consider a magnetic field, intensity $\mathbf{H}$, and a boundary with normal $\hat{\mathbf{n}}$. Along a plane of magnetic symmetry, the condition $\hat{\mathbf{n}} \cdot \mathbf{H} = 0$ holds; (magnetic field lines do not cross the boundary). Similarly, on a plane of magnetic antisymmetry, the boundary condition is $\hat{\mathbf{n}} \times \mathbf{H} = 0$ (magnetic field lines crossing the boundary have no component parallel to the boundary). The magnetic flux density is given by $\mathbf{B} = \mu \mathbf{H}$, where μ is the permeability of the medium.

A typical result is presented in Fig. 6 for a solenoid of thickness $h = 0.02$ m, arm width $b = 0.12$ m, solenoid half-width $w = 0.3$ m, with current density $J_0 = 15\ \text{MA/m}^2$. The backing plate is assumed to be 0.01 m thick with relative permeability $\mu_r = 20{,}000$. It would be more realistic to use a thinner, more highly permeable back plate, but this choice would diminish the computational efficiency. The titanium face sheet and TBC layer have relative permeability near unity, and are not modeled explicitly. The plane $x = 0$ m exhibits magnetic symmetry, while the plane $x = 0.3$ m is a plane of magnetic antisymmetry. The region from which power can be extracted is marked "region of high conductivity".

For each such FEMLAB calculation, the expression (5) is integrated over the region of high conductivity, and the total power generated per periodic tile is twice the value of this integral (since only half of one periodic element is modeled). This power gen-

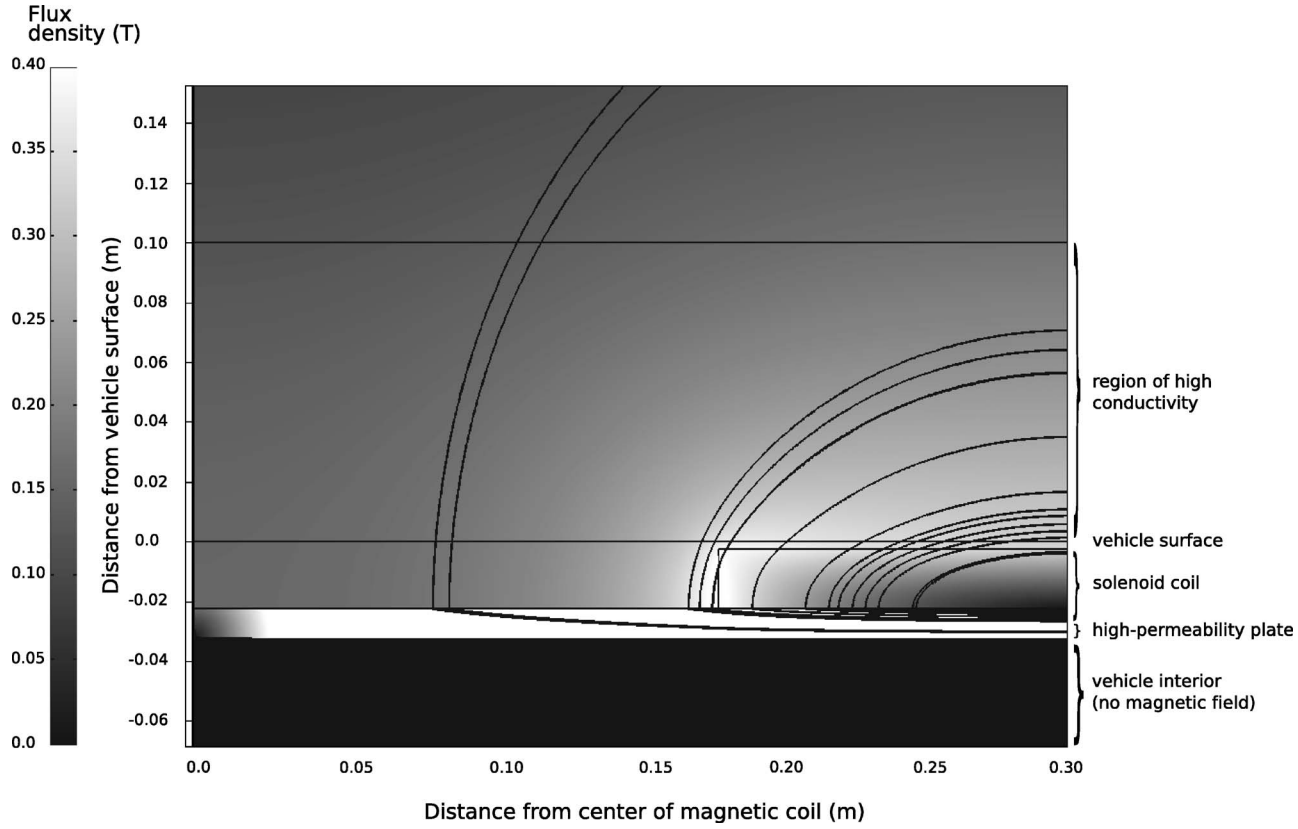


Fig. 6 Typical FEMLAB calculated result showing magnetic flux density (T) (with bar to the left) and a representative set of magnetic field lines. This configuration is for a solenoid with current density J_0=15 MA/m^2, of thickness h=0.02 m, arm width b=0.12 m, and solenoid half-width w=0.3 m. Note that the magnetic flux density in the back plate is very high, while the magnetic flux density in the vehicle in negligible.

eration is an input to the optimization procedure. Further finite element calculations suggest that the power generation capacity for a tile at the edge of a periodic array is reduced by approximately 15%. Thus, an array of N periodic tiles will produce power per meter length equal to

$$P_{\text{tot}} = (N - 0.3)P \tag{6}$$

where P refers to the integral of P_0 (Eq. (5)) over the width of the magnetic element and the thickness of the conductive layer.

5 Optimization Scheme

5.1 Power/Mass Exchange Constant. To optimize the configuration of the magnetic array, we define the effective power $\hat{P}$ which becomes the objective function

$$\hat{P} \equiv P_{\text{net}} - \gamma m_{\text{tot}} \tag{7}$$

In this formula, P_{net} is the net power generated by the panel (the total power generated less the power required to run the magnets) and m_{tot} is the total mass of the magnets plus the cooling system. The power/mass exchange constant γ is interpreted as the penalty associated with each unit of mass of the MHD device. Accordingly, the effective power $\hat{P}$ is the difference between the net power generated and the total penalty associated with the mass of the device. For specified γ, the optimal design is that maximizing the effective power $\hat{P}$.

If γ is progressively varied, a locus of optimal designs in $(m_{\text{tot}}, P_{\text{net}})$ space arises, as shown in Fig. 7. Note that a minimum mass of magnets m_{min} is required before any power can be generated. Graphically, the net power density (the ratio of net power to total mass) is maximized at the point where a ray from the origin is tangent to the locus, which implies that the slope of the locus must be equal to the net power density. At every point on the locus, since $\hat{P}$ is a maximum

$$\delta\hat{P} = \delta P_{\text{net}} - \gamma \delta m_{\text{tot}} = 0 \tag{8}$$

such that γ is the slope of the locus. Hence we must find the point where the net power density is equal to γ. When $\hat{P}$=0

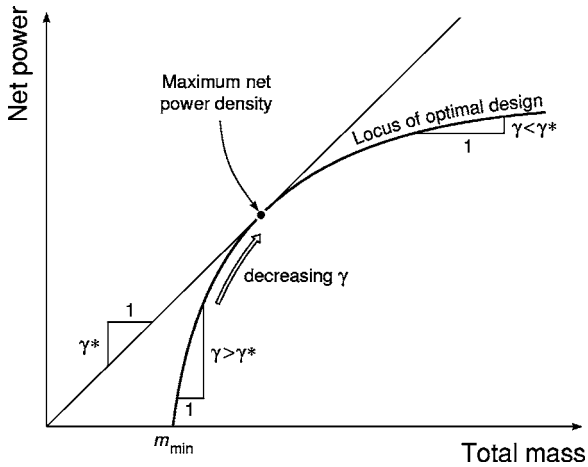


Fig. 7 The locus of optimal designs in $(m_{\text{tot}}, P_{\text{net}})$ space, showing the construct to obtain the maximum net power density. The maximum net power density is found when a ray from the origin is tangent to the locus.

$$\gamma|_{\hat{P}=0} = \frac{P_{\text{net}}}{m_{\text{tot}}} = \gamma^* \tag{9}$$

Thus γ can also be interpreted as the maximum net power density by finding the value of $\gamma=\gamma^*$ for which the maximum of $\hat{P}=0$.

5.2 The Optimization. Optimization involves an exchange between total mass (the MHD device plus the coolant) and power generation, dependent upon flight time t_f and heat flux q. Because the mass of the structural elements and insulation will be present whether or not a MHD device is incorporated, this mass need not be considered for optimization. To determine the number of periodic elements N (noting that fractional elements are not permissible), the maximum width of the array is taken to be 1.2 m. The power surplus P_{net} (per unit length) is the power generated P_{tot} less the power expended by the magnetic array P_{exp}. For permanent magnets $P_{\text{exp}}=0$, while for a resistive solenoid system

$$P_{\text{exp}} = 2N\frac{J_0^2 hb}{\sigma_s} \tag{10}$$

where σ_s is the conductivity of the solenoid material (at 373 K, the expected operating temperature of the solenoids, for copper $\sigma_s=4.6\times10^7/\Omega\text{m}$, while for aluminum, $\sigma_s=2.8\times10^7/\Omega\text{m}$). A resistive electromagnet generates heat equal to the power required to run the solenoid P_{exp}.

The mass of the solenoid array (per unit length) is

$$m_{\text{mag}} = 2Nbh\rho \tag{11}$$

where ρ is the density of the material used in the array (for copper, $\rho=8960\ \text{kg/m}^3$, and for aluminum, $\rho=2702\ \text{kg/m}^3$), and the mass of a neodymium-iron-boron permanent magnet array is

$$m_{\text{mag}} = 2Nwh\rho \tag{12}$$

with $\rho=7500\ \text{kg/m}^3$.

The steady-state flow of heat through the TBC is dictated by its thermal conductivity k, the difference in temperature between the two sides ΔT, and its thickness t_{TBC}

$$H_{\text{surf}} = 2Nw\frac{k\Delta T}{t_{\text{TBC}}} \tag{13}$$

where $2Nw$ is the total width of the panel. Transients are considered to be of short duration. The temperature difference is considered to be governed by the temperature attained on the external surface of the TBC (1500 K), ascertained by the aerothermodynamic calculations,[1] and the maximum temperature that a titanium alloy can sustain to assure adequate creep strength (900 K). The total heating input is

$$H_{\text{tot}} = H_{\text{surf}} + P_{\text{exp}} \tag{14}$$

The mass of the cooling system is

$$m_{\text{cool}} = (1+\epsilon)\frac{H_{\text{tot}} t_f}{C_{\text{water}}} \tag{15}$$

where C_{water} is the energy per unit mass to heat and evaporate the water. It is assumed that the water vapor is immediately vented from the vehicle through an insulated system, such that no further heating of the vehicle occurs. The total mass of the system is thus $m_{\text{tot}}=m_{\text{mag}}+m_{\text{cool}}$, and the effective power generation can be evaluated for each candidate design. Note that the quantities calculated here are per unit length of the MHD panel.

[1]Calculations by Candler [22] indicate that the total heating rate on the vehicle surface is approximately in equilibrium with radiative cooling and heat flux into the vehicle when the surface temperature is 1500 K.

Table 1 Standard values for optimization parameters

Parameter	Value
Exchange constant γ	5 kW/kg
Flight time t_f	1000 s
TBC thermal conductivity k	0.5 W/mK
Temperature gradient through TBC ΔT	600 K
Cooling system mass coefficient ϵ	0.15
Total panel width w_{tot}	1.2 m
Shock layer velocity u_0	6 km/s
Plasma conductivity σ_0	100/Ωm
Reference TBC thickness	2.5 mm
Potassium seeding (by mass)	1%

6 Optimization Results

The initial optimization will be performed using the parameters in Table 1. Thereafter the separate influences of several parameters will be explored.

A wide search through probable candidates for geometric and internal variables (including all combinations of those in Table 2) indicates that an optimal design is achieved by using an aluminum solenoid with element width $2w=0.6$ m, solenoid arm width $b=0.12$ m, solenoid thickness $h=0.02$ m, TBC thickness $t_{\text{TBC}}=0.0025$ m, and current density $J_0=12\ \text{MA/m}^2$ (equivalent to 28,800 amp turns). This array generates a net power $P_{\text{net}}=538$ kW/m. The total power generated is 588 kW/m, with 50 kW/m expended to power the solenoid. The mass of the solenoid is 26 kg/m, and the mass of coolant is 79 kg/m. At very long flight times, $t_f>6000$ s, copper becomes preferable to aluminum because of its higher conductivity. Permanent magnetic arrays are uniformly rejected because the magnetic fields are not sufficiently large to be competitive.

The procedure is illustrated by results for the net power/total mass at fixed TBC thickness (Fig. 8) with each curve representing a specific thermal conductivity. Each point on every curve represents a design with maximum effective power. Note that the shapes of the curves are similar to that shown in Fig. 7.

Variations of net power density with power are illustrated in Fig. 9. This plot shows that power density can be significantly increased by increasing the thickness of the TBC. It is also apparent that there is a maximum net power density for each value of TBC thickness.

Trends in the maximum are reexpressed in Fig. 10, using the loci through the maxima of each curve in Fig. 9. This result provides direct guidelines for the material to be used as the thermal barrier (based on the thermal conductivity range) and the thickness required to achieve the maximum net power density. It is

Table 2 Values of design parameters used in the finite element simulations. Note that the values used for the solenoid arm width *b* are dependent upon the solenoid half-width *w*; *b* is taken to be the integral part of the calculation.

Design parameter	Values simulated
Solenoid half-width w (m)	0.1, 0.12, 0.15, 0.20, 0.25, 0.30, 0.60
Solenoid thickness h (m)	0.01, 0.015, 0.02, 0.0225, 0.025, 0.03, 0.04, 0.045
Solenoid arm width b (m)	w/7 x (1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6)
Thickness of TBC layer t_{TBC} (mm)	1, 1.5, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 4.5, 5
Current density J_0(MA/m^2)	10, 11.25, 12, 13.25, 14.25, 15, 15.5, 16.25, 17.5, 20, 22.5, 25, 30, 35, 40

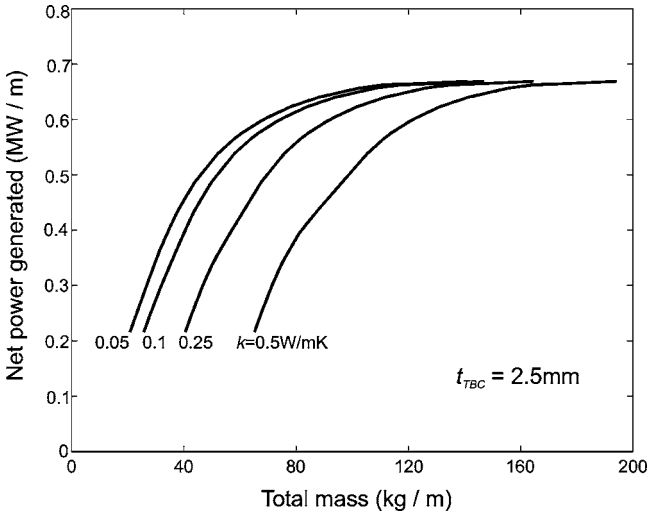


Fig. 8 Net power generation versus total mass at constant TBC thickness t_{TBC}=2.5 mm for varying TBC thermal conductivity

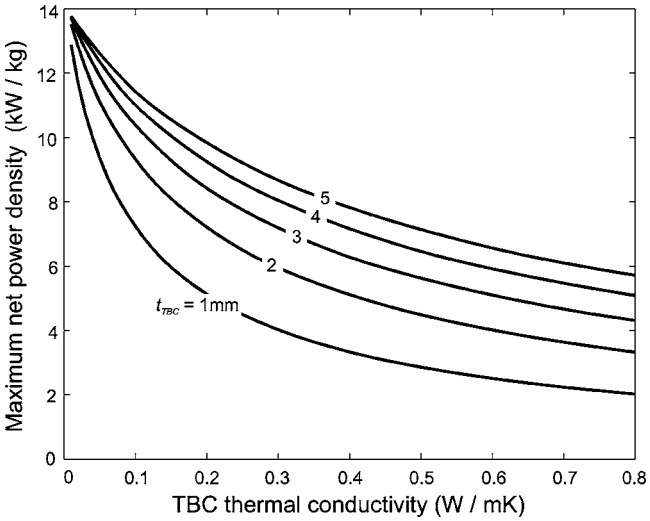


Fig. 10 Maximum net power density as a function of TBC thermal conductivity for varying TBC thickness

clear that maximum net power density increases rapidly as k is decreased below 0.2 W/mK; when k is very low, the TBC thickness becomes less important.

Varying the number of magnetic elements reveals that the maximum effective power occurs when $N=2$. This optimum arises because the increase in the flux density near the surface of the vehicle upon increasing N is counteracted by the rate at which the flux density declines with distance from the vehicle. The enhancement in net power enabled by reducing the extra weight of the cooling system is indicated on Fig. 11. At low γ, the design is not sensitive to the cooling system efficiency, because the additional mass of cooling water does not carry a large weight penalty. As γ increases, the design becomes more sensitive to cooling system efficiency. This figure expresses the benefit of designing a lightweight pumping system.

7 Implications

The analysis has demonstrated that MHD power panels located within a re-entry vehicle have the potential to generate usable power. The optimization indicates that it is possible to design a 1.2 m wide panel which will generate 0.6 MW/m length at a total mass of approximately 110 kg/m. However, stringent technological challenges need to be satisfied before such a panel can be brought to fruition.

1. Effective injection and mixing of the seed material are essential and remain to be demonstrated.
2. Water must be stored on board and a lightweight recirculation system developed that pumps the water through the panel and then ejects it from the vehicle.
3. A thermal barrier material with low thermal conductivity, such as $Gd_2Zr_2O_7$, must be deposited on the surface of the panel at unprecedented thickness. It must resist spalling during manufacturing, as well as when subject to a thermal gradient on reentry.
4. A bond coat must be developed for titanium alloys that survives manufacturing and remains intact when exposed to reentry.

The importance of thermal management is emphasized: For the example cases, approximately 75% of the total mass of the MHD

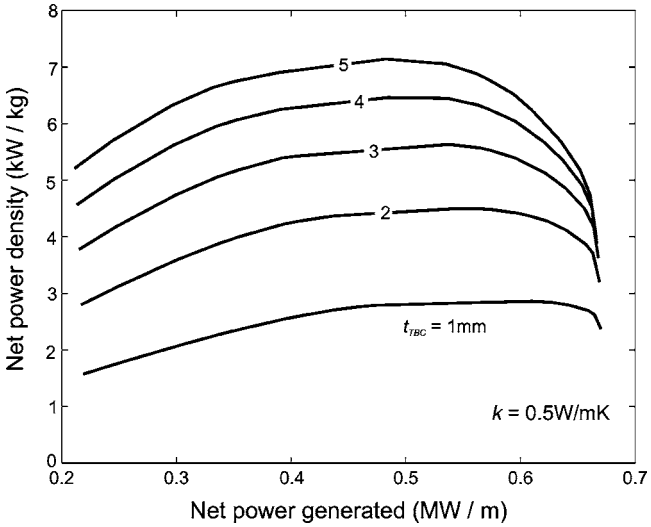


Fig. 9 Net power density as a function of net power for constant TBC thermal conductivity k=0.5 W/mK and varying TBC thickness

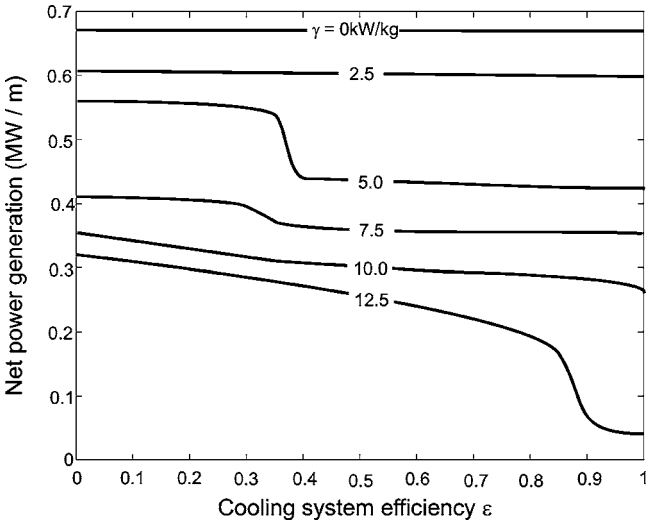


Fig. 11 Variation of net power generation with the cooling system mass coefficient ϵ

system is coolant. Preliminary materials and manufacturing assessments indicate that $Gd_2Zr_2O_7$ having the requisite thickness can be deposited onto superalloys by using plasma-assisted, directed vapor deposition (DVD) [24]. It remains to determine whether $Gd_2Zr_2O_7$ can be deposited on titanium alloys at the required thickness and pore structure and adhere, as well as resist delamination in a thermal gradient. Depositing a bond coat on the alloy that oxidizes to form $\alpha - Al_2O_3$, such as $TiAl_3$, would be essential. These materials and manufacturing issues would need to be clearly specified and resolved before embarking on MHD panel development.

Acknowledgment

The authors are grateful for financial support from DARPA—Defense Science Office, and for extensive discussions with and numerical modeling by Prof. Graham Candler at the University of Minnesota, and Dr. Mikhail Shneider and Dr. Sergey Macheret at Princeton University.

References

[1] Macheret, S. O., Shneider, M. N., and Candler, G. V., 2004, "Modelling of MHD Power Generation on Board Reentry Vehicles," Paper No. AIAA-2004-1024.
[2] Hankey, W. L., and Elliot, G. A., 1968, "Hypersonic Lifting Body Optimization," J. Spacecr. Rockets, **5**(12), pp. 1463–1467.
[3] Kolodziej, P., Bowles, J. V., and Roberts, C., 1998, "Optimizing Hypersonic Sharp Body Concepts from a Thermal Protection System Perspective," 8th, AIAA Paper No. 98-1610.
[4] Rosa, R. J., 1968, *Magnetohydrodynamic Energy Conversion*, McGraw-Hill, New York.
[5] Callery Chemical Company, 2002, "Potassium-Sodium Alloy (NaK)," Product data sheet.
[6] Woods, B. 2003, "Heated Debates: A History of the Development of the Space Shuttle's Thermal Protection System: 1970-1981," Quest, **10**(3), pp. 39–55.
[7] Miller, R. A., 1984, "Oxidation-Based Model for Thermal Barrier Coating Life, J. Am. Ceram. Soc., **67**, pp. 517–521.
[8] Evans, A. G., Mumm, D. R., Hutchinson, J. W. Meier, G. H., and Pettit, F. S., 2001, Mechanisms Controlling the Durability of Thermal Barrier Coatings," Prog. Mater. Sci., **46**, pp. 505–553.
[9] Ashby, M. F., 1989, "Materials Selection in Conceptual Design," Mater. Sci. Technol., **5**(6), pp. 517–525.
[10] Ashby, M. F., 1999, *Material Selection in Mechanical Design*, 2nd ed., Butterworth-Heinemann, Oxford, UK.
[11] Maloney, M. J., 2001, "Thermal Barrier Coating Systems and Materials," U.S. Patent 6,177,200, United Technologies Corporation, Hartford, CT.
[12] Nicholls, J. R., Lawson, K. J., Johnstone, A., and Rickerby, D. S., 2002, "Methods to Reduce the Thermal Conductivity of eb-pvd tbcs," Surf. Coat. Technol., **151–152**, pp. 383–391.
[13] Lu, T. J., Levi, C. G., Wadley, H. N. G., and Evans, A. G., 2001, "Distributed Porosity as a Control Parameter for Oxide Thermal Barriers Made by Physical Vapor Deposition," J. Am. Ceram. Soc., **84**(12), pp. 2937–2946.
[14] Hass, D. D. Slifka, A. J., and Wadley, H. N. G., 2001, "Low Thermal Conductivity Vapor Deposited Zirconia Microstructures," Acta Mater., **49**(6), pp. 973–983.
[15] Wicks, N., and Hutchinson, J. W., 2001, "Optimal Truss Plates," Int. J. Solids Struct., **38**(30–31), pp. 5165–5183.
[16] Evans, A. G., 2001, "Lightweight Materials and Structures," Mater. Res. Soc. Bull., **26**(10), pp. 790–797.
[17] Chiras, S., Mumm, D. R., Evans, A. G., Wicks, N., Hutchinson, J. W., Dharmasena, H. N., Wadley, H. N. G., and Fichter, S., 2002, "The Structural Performance of Near-optimized Truss Core Panels," Int. J. Solids Struct., **39**(15), pp. 4093–4115.
[18] Kovalev, K. I., and Markina, T. A., 1995, "Power Facility with a Built-in Multipole MHD Generator," High Temp., **33**, pp. 463–472.
[19] Bityurin, V. A. Bocharov, A. N., and Lineberry, J. T., 1999, "MHD Aerospace Applications," *13th International Conference on MHD Power Generation and High Temperature Technologies*, Vol. III, Beijing, Oct. 12–15, IEE CAS, Beijing, pp. 793–814.
[20] Faghri, A., 1995, *Heat Pipe Science and Technology*, Taylor & Francis, London.
[21] Sudmeijer, K. J., Buursink, J., and Lopes, C., 2003, "A Low-cost Active Cooling Mechanism to be Used on Future Launchers," Space Technol., **23**(2–3), pp. 87–103.
[22] Candler, G. 2003, Private communication.
[23] FEMLAB, 2003, Comsol Inc., Burlington MA.
[24] Hass, D. D., Groves, J. F., and Wadley, H. N. G., 2001, "Reactive Vapor Deposition of Metal Oxide Coatings," Surf. Coat. Technol., **146**, pp. 85–93.

Electromagnetoelastic Dynamic Response of Transversely Isotropic Piezoelectric Hollow Spheres in a Uniform Magnetic Field

H. L. Dai[1]
e-mail: hldai520@sina.com

Y. M. Fu

T. X. Liu

Department of Engineering Mechanics,
Hunan University,
Changsha 410082,
Hunan Province, P. R. C.

An analytical method is presented to solve the problem of electromagnetoelastic dynamic response of transversely isotropic piezoelectric hollow spheres in a uniform magnetic field, subjected to arbitrary mechanical load and electric excitation. Exact expressions for the dynamic responses of stresses, perturbation of magnetic field vector, electric displacement, and electric potential in piezoelectric hollow spheres are obtained by means of Hankel transform, Laplace transform and their inverse transforms. An interpolation method is applied to solve the Volterra integral equation of the second kind involved in the exact expression, which is caused by interaction between electric-elastic field and electric-magnetic field. Thus, an analytical solution for the problem of dynamic response of a transversely isotropic piezoelectric hollow sphere in a uniform magnetic field is obtained. Finally, some numerical examples are carried out, and may be used as a reference to solve other dynamic coupled problems of electromagneto-elasticity.
[DOI: 10.1115/1.2178361]

1 Introduction

The analyses for dynamic problems of elastic bodies are important and interesting research fields for engineers and scientists. Being the common structural form, the applications for transversely isotropic hollow spheres have continuously increased in some engineering areas, including aerospace, offshore and submarine structures, chemical vessel, and civil engineering structures. The interaction of electric potential, electric displacement, perturbation of magnetic field vector, and elastic deformation in spherical structures is studied due to its many engineering applications, such as geophysics for understanding the effect of the Earth's magnetic field on seismic waves, damping of acoustic waves in a magnetic field, emissions of electromagnetic radiations from nuclear devices, development of a highly sensitive superconducting magnetometer, electrical power engineering, optics, etc.

For nonpiezoelectric media, Hu [1] first initiated using a separation method and presented a general theory of elasticity for spherically isotropic medium. Many subsequently important analyses were inspired based on Hu's elegant investigations on some static problems such as a concentrated force in an infinite medium, stress concentration due to a spherical cavity, and a steadily rotating shell. Hata [2–4] obtained the transient thermal stress responses in a uniformly heated isotropic spherical shell and solid sphere, as well as transversely isotropic solid sphere by using the ray theory. Sherief and Ezzat [5] used the Laplace transform technique to find the distribution of thermal stresses and temperature in a generally thermoelastic and electrically conducting half space under sudden thermal shock and permeated by a primarily uniform magnetic field. By means of the Hankel transform and Laplace transform technique, Wang [6] obtained dynamic thermostress-concentration effect in a spherically isotropic sphere. While for piezoelectric media, Heyliger and Wu [7] studied the displacement and electrostatic potential fields in a static and free vibration response of layered piezoelectric spheres. By virtue of the separation of variables method, Ding et al. [8,9] investigated transient responses of a pyroelectric hollow sphere under radial deformation and a functionally graded pyroelectric hollow sphere for spherically symmetric problems. Utilizing the same method of Ref. [6], Dai and Wang [10] presented an analytical solution for the interaction of electric potentials, electric displacements, elastic deformations and mechanical loads, and described electromagnetoelastic responses and perturbation of magnetic field vector in a piezoelectric hollow cylinder subjected to sudden mechanical load and electric potential. Recently, Dai and Wang [11,12] investigated transient wave propagations in piezoelectric hollow spheres and laminated piezoelectric spherical shells subjected to thermal shock and electric excitation. However, the investigation of electromagnetoelastic dynamic response of transversely isotropic piezoelectric spheres in a uniform magnetic field is a more complex problem that has not been investigated as extensively as it deserves so far.

In this paper, an analytical method is developed for electromagnetoelastic interactions in transversely isotropic piezoelectric hollow spheres placed in a uniform magnetic field, subjected to mechanical load and electric excitation. The electromagnetoelastic dynamic equation of the piezoelectric hollow spheres may be decomposed into a quasistatic term and a dynamic term. First, using the method described by Lekhnitskii [13], the quasistatic question can be solved. Second, the solution to the inhomogeneous dynamic question is obtained by utilizing Hankel transform (Cinelli [14]), Laplace transform and their inverse transforms. Thus, the exact solution for electromagnetoelastic dynamic responses of the transversely isotropic piezoelectric hollow spheres is obtained and this solution is illustrated with numerical examples to demonstrate the methodology is simple and effective.

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 13, 2005; final manuscript received October 18, 2005. Review conducted by Z. Suo. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 Basic Formulations of the Problem

A transversely isotropic piezoelectric hollow sphere placed initially in a uniform magnetic field $\vec{H}(0,0,H_u)$. A spherical coordinate system (r,θ,φ) with the origin identical to the center of the sphere is used for a spherically symmetric problem where the transversely isotropic piezoelectric hollow sphere with internal radius a and external radius b, so the strain-displacement relations are expressed as

$$\epsilon_{rr}=\frac{\partial u(r,t)}{\partial r},\quad \epsilon_{\theta\theta}=\epsilon_{\varphi\varphi}=\frac{u(r,t)}{r},\quad \epsilon_{r\theta}=\epsilon_{\theta\varphi}=\epsilon_{r\varphi}=0 \tag{1}$$

where ϵ_{ij} $(i,j=r,\theta,\varphi)$ are strain components, and $u(r,t)$ expresses a radial displacement. The constitutive relations of the spherically transversely isotropic pyroelectric medium are expressed as (Sinha [15]; Chen and Shioya [16])

$$\sigma_r=c_{11}\epsilon_{rr}+2c_{12}\epsilon_{\theta\theta}+e_{11}\frac{\partial\phi}{\partial r} \tag{2a}$$

$$\sigma_\theta=c_{12}\epsilon_{rr}+(c_{22}+c_{23})\epsilon_{\theta\theta}+e_{12}\frac{\partial\phi}{\partial r} \tag{2b}$$

$$D_r=e_{11}\epsilon_{rr}+2e_{12}\epsilon_{\theta\theta}-\beta_{11}\frac{\partial\phi}{\partial r} \tag{2c}$$

where c_{ij}, e_{1i}, $(i=1,2,j=1,2,3)$, and β_{11} are elastic constants, piezoelectric constants, and dielectric constants, respectively. σ_i, $(i=r,\theta)$, and D_r are the component of stress and radial electric displacement, respectively.

Assuming that the magnetic permeability μ (Ezzat [17]) of the transversely isotropic piezoelectric hollow sphere equals the magnetic permeability of the medium around it, the governing electrodynamic Maxwell equations (John [18]) are given by

$$\vec{J}=\nabla\times\vec{h},\quad \nabla\times\vec{e}=-\mu\frac{\partial\vec{h}}{\partial t},\quad \operatorname{div}\vec{h}=0,$$
$$\vec{e}=-\mu\left(\frac{\partial\vec{U}}{\partial t}\times\vec{H}\right),\quad \vec{h}=\nabla\times(\vec{U}\times\vec{H}) \tag{3}$$

Applying a uniform magnetic field vector $\vec{H}(0,0,H_u)$ in the hollow spherical coordinate (r,θ,φ) system to Eq. (3), yields

$$\vec{U}=(u(r,t),0,0),\quad \vec{e}=-\mu\left(0,H_u\frac{\partial u}{\partial t},0\right) \tag{4a}$$

$$\vec{h}=(0,0,h_u),\quad \vec{J}=\left(0,-\frac{\partial h_u}{\partial r},0\right),\quad h_u=-H_u\left(\frac{\partial u}{\partial r}+\frac{2u}{r}\right) \tag{4b}$$

In the absence of free charge density, the charge equation of electrostatics is expressed as (Dunn and Taya [19])

$$\frac{\partial D_r(r,t)}{\partial r}+\frac{2D_r(r,t)}{r}=0 \tag{5}$$

From Eq. (5), we obtain

$$D_r(r,t)=\frac{1}{r^2}d(t) \tag{6}$$

where $d(t)$ is an undetermined function to time t.

Utilizing Eq. (6), we rewrite Eq. (2c), gives

$$\frac{\partial\phi}{\partial r}=\frac{e_{11}}{\beta_{11}}\frac{\partial u}{\partial r}+\frac{2e_{12}}{\beta_{11}}\frac{u}{r}-\frac{1}{\beta_{11}}\frac{d(t)}{r^2} \tag{7}$$

Then, substituting Eq. (7) into Eqs. (2a) and (2b), yields

$$\sigma_r=\left(c_{11}+\frac{e_{11}^2}{\beta_{11}}\right)\frac{\partial u}{\partial r}+2\left(c_{12}+\frac{e_{11}e_{12}}{\beta_{11}}\right)\frac{u}{r}-\frac{e_{11}}{\beta_{11}}\frac{d(t)}{r^2} \tag{8a}$$

$$\sigma_\theta=\left(c_{12}+\frac{e_{11}e_{12}}{\beta_{11}}\right)\frac{\partial u}{\partial r}+\left(c_{22}+c_{23}+\frac{2e_{12}^2}{\beta_{11}}\right)\frac{u}{r}-\frac{e_{12}}{\beta_{11}}\frac{d(t)}{r^2} \tag{8b}$$

The electromagnetoelastic dynamic equation of the transversely isotropic piezoelectric hollow sphere is (Lekhnitskii [13])

$$\frac{\partial\sigma_r}{\partial r}+\frac{2(\sigma_r-\sigma_\theta)}{r}+f_u=\rho\frac{\partial^2 u}{\partial t^2} \tag{9}$$

where ρ is the mass density, f_u is defined as Lorentz's force (John [18]), which can be written as

$$f_u=\mu(\vec{J}\times\vec{H})=\mu H_u^2\frac{\partial}{\partial r}\left(\frac{\partial u}{\partial r}+\frac{2u}{r}\right) \tag{10}$$

Substituting Eqs. (8) into the motion equation of a spherically symmetric problem, the basic displacement equation of electromagnetoelastic motion of the transversely isotropic piezoelectric hollow sphere is expressed as

$$\frac{\partial^2 u(r,t)}{\partial r^2}+\frac{2}{r}\frac{\partial u(r,t)}{\partial r}-\frac{M^2u(r,t)}{r^2}=\frac{1}{C_L^2}\frac{\partial^2 u(r,t)}{\partial t^2}+I\frac{d(t)}{r^3} \tag{11}$$

where

$$M^2=\frac{2[(c_{22}+c_{23}-c_{12}+\mu H_u^2)\beta_{11}+2e_{12}^2-e_{11}e_{12}]}{(m+\mu H_u^2)\beta_{11}},$$
$$I=-\frac{2e_{12}}{(m+\mu H_u^2)\beta_{11}},\quad m=c_{11}+\frac{e_{11}^2}{\beta_{11}},\quad C_L^2=\frac{m+\mu H_u^2}{\rho} \tag{12}$$

Boundary conditions of stress and electric are, respectively, expressed as

$$\sigma_r(a,t)=p_a(t),\quad \sigma_r(b,t)=p_b(t) \tag{13a}$$

$$\phi(a,t)=\phi_a(t),\quad \phi(b,t)=\phi_b(t) \tag{13b}$$

The transversely isotropic piezoelectric hollow sphere at rest prior to time $t=0$ and the initial conditions are

$$u(r,t)=\frac{\partial u(r,t)}{\partial t}=0\quad \text{at } t=0 \tag{14}$$

3 Solution of the Problem

Assume that the general solution of the basic displacement Eq. (11) may be expressed in the form (Eringen and Suhubi [20]; Dai and Wang [10–12])

$$u(r,t)=u_q(r,t)+u_d(r,t) \tag{15}$$

where $u_q(r,t)$ and $u_d(r,t)$ are the quasistatic solution and the dynamic solution of Eq. (11), respectively.

First, the quasistatic solution $u_q(r,t)$ must satisfy the following equation and the corresponding inhomogeneous boundary conditions

$$\frac{\partial^2 u_q(r,t)}{\partial r^2}+\frac{2}{r}\frac{\partial u_q(r,t)}{\partial r}-\frac{M^2}{r^2}u_q(r,t)=I\frac{d(t)}{r^3} \tag{16a}$$

$$\left[\frac{\partial u_q(r,t)}{\partial r}+h\frac{u_q(r,t)}{r}\right]_{r=i}=\theta_i(t)\quad (i=a,b) \tag{16b}$$

where

$$h=\frac{2(c_{12}\beta_{11}+e_{11}e_{12})}{m\beta_{11}},\quad \theta_i(\tau)=\frac{1}{m}\left[\frac{e_{11}}{\beta_{11}}\frac{d(t)}{i^2}+p_i(t)\right],\quad (i=a,b) \tag{17}$$

Using the method in Lekhnitskii [13], the general solution of Eq. (16*a*) is expressed as

$$u_q(r,t)=A_1r^{n-0.5}+A_2r^{-(n+0.5)}-\frac{I}{M^2}\frac{d(t)}{r} \tag{18}$$

where A_1 and A_2 are unknown constants which can be determined by making use of the inhomogeneous boundary conditions (16*b*), and $n=\sqrt{0.25+M^2}$. Thus, (18) is rewritten as

$$u_q(r,t)=B_1(r)p_a(t)+B_2(r)p_b(t)+B_3(r)d(t) \tag{19}$$

where

$$B_1(r)=\frac{b^{-(n+1.5)}}{(n-0.5+h)L_1m}r^{n-0.5}+\frac{b^{n-1.5}}{(n+0.5-h)L_1m}r^{-(n+0.5)}$$

$$B_2(r)=\frac{-a^{-(n+1.5)}}{(n-0.5+h)L_1m}r^{n-0.5}+\frac{-a^{n-1.5}}{(n+0.5-h)L_1m}r^{-(n+0.5)}$$

$$B_3(r)=\frac{a^{-2}b^{-(n+1.5)}-b^{-2}a^{-(n+1.5)}}{(n-0.5+h)L_1}L_2r^{n-0.5}+\frac{a^{-2}b^{n-1.5}-b^{-2}a^{n-1.5}}{(n+0.5-h)L_1}L_2r^{-(n+0.5)}-\frac{I}{M^2r}$$

$$L_1=a^{n-0.5}b^{-(n+1.5)}-b^{n-0.5}a^{-(n+1.5)},\quad L_2=\frac{(h-1)I}{M^2}+\frac{e_{11}}{m\beta_1} \tag{20}$$

Second, substituting Eq. (15) into Eq. (9) and utilizing Eq. (13) provides an inhomogeneous dynamic equation with homogeneous boundary conditions and the corresponding initial conditions for $u_d(r,t)$

$$\frac{\partial^2u_d(r,t)}{\partial r^2}+\frac{2}{r}\frac{\partial u_d(r,t)}{\partial r}-\frac{H^2}{r^2}u_d(r,t)=\frac{1}{C_L^2}\left[\frac{\partial^2u_d(r,t)}{\partial t^2}+\frac{\partial^2u_q(r,t)}{\partial t^2}\right] \tag{21a}$$

$$\left[\frac{\partial u_d(r,t)}{\partial r}+h\frac{u_d(r,t)}{r}\right]_{r=a,b}=0 \tag{21b}$$

$$u_d(r,0)+u_q(r,0)=0,\quad \frac{\partial u_d(r,0)}{\partial t}+\frac{\partial u_q(r,0)}{\partial t}=0 \tag{21c}$$

For solving the $u_d(r,t)$, Eq. (21*a*) may be transformed into a normal Bessel equation by assuming

$$u_d(r,t)=r^{-0.5}f(r,t) \tag{22}$$

Substituting Eq. (22) into Eqs. (21), yields

$$\frac{\partial^2f(r,t)}{\partial r^2}+\frac{1}{r}\frac{\partial f(r,t)}{\partial r}-\frac{N^2}{r^2}f(r,t)=\frac{1}{C_L^2}\left[\frac{\partial^2f(r,t)}{\partial t^2}+\frac{\partial^2u_{q1}(r,t)}{\partial t^2}\right] \tag{23a}$$

$$\frac{\partial f(a,t)}{\partial r}+h_af(a,t)=0,\quad \frac{\partial f(b,t)}{\partial r}+h_bf(b,t)=0 \tag{23b}$$

$$f(r,0)+u_{q1}(r,0)=u_1,\quad \frac{\partial f(r,0)}{\partial t}+\frac{\partial u_{q1}(r,0)}{\partial t}=v_1 \tag{23c}$$

where

$$u_{q1}(r,t)=C_1(r)p_a(t)+C_2(r)p_b(t)+C_3(r)d(t),\qquad C_i(r)=r^{0.5}B_i(r),\quad (i=1,2,3) \tag{24a}$$

$$N^2=0.25+M^2,\quad h_a=\frac{(h-0.5)}{a},\quad h_b=\frac{(h-0.5)}{b} \tag{24b}$$

Let $u_{q1}(r,t)=0$ in Eq. (23*a*), then this homogeneous equation with homogeneous boundary conditions (23*b*) may be solved by assuming

$$f(r,t)=g(r)\exp(i\omega t) \tag{25}$$

From Eqs. (23) and (25), we have the following eigenequation:

$$J_aY_b-J_bY_a=0 \tag{26}$$

where

$$J_a=k_iJ_N'(k_ia)+h_aJ_N(k_ia)\quad J_b=k_iJ_N'(k_ib)+h_bJ_N(k_ib)$$

$$Y_a=k_iY_N'(k_ia)+h_aY_N(k_ia)\quad Y_b=k_iY_N'(k_ib)+h_bY_N(k_ib) \tag{27}$$

and $J_N(k_ir)$ and $Y_N(k_ir)$ are, respectively, the first and the second kind of the Nth-order Bessel function, k_i $(i=1,2,\ldots,n)$ express a series of positive roots of the natural eigenequation (26), and the natural frequencies are $\omega_i=C_Lk_i$.

From Cinelli [14], define $\bar{f}(k_i,t)$ as the finite Hankel transform of $f(r,t)$, yields

$$\bar{f}(k_i,t)=H[f(r,t)]=\int_a^b rf(r,t)G_N(k_ir)dr \tag{28}$$

Then the inverse transform to Eq. (28) is given by

$$f(r,t)=\sum_{k_i}\frac{\bar{f}(k_i,t)}{F(k_i)}G_N(k_ir) \tag{29}$$

where

$$F(k_i)=\int_a^b r[G_N(k_ir)]^2\,dr=\frac{J_a^2}{J_b^2}\frac{2}{k_i^2\pi^2}\left\{h_b^2+k_i^2\left[1-\left(\frac{N}{k_ib}\right)^2\right]\right\}-\frac{2}{k_i^2\pi^2}\left\{h_a^2+k_i^2\left[1-\left(\frac{N}{k_ia}\right)^2\right]\right\} \tag{30a}$$

$$G_N(k_ir)=J_N(k_ir)Y_a-J_aY_N(k_ir) \tag{30b}$$

Applying the finite Hankel transform (28) to Eq. (23*a*) and utilizing the homogeneous boundary condition (23*b*), yields

$$-k_i^2\bar{f}(k_i,t)=\frac{1}{C_L^2}\left[\frac{\partial^2\bar{f}(k_i,t)}{\partial t^2}+\frac{\partial^2\bar{u}_{q1}(k_i,t)}{\partial t^2}\right] \tag{31}$$

where

$$\bar{u}_{q1}(k_i,t)=H[u_{q1}(r,t)] \tag{32}$$

Applying Laplace transform to the two sides of Eq. (31) and utilizing the initial conditions (23*c*), yields (the detailed solution processes is given in the Appendix)

$$\bar{f}^*(k_i,L)=-\bar{u}_{q1}^*(k_i,L)+\frac{\omega_i^2}{(\omega_i^2+L^2)}\bar{u}_{q1}^*(k_i,L) \tag{33}$$

where L is the parameter of Laplace transform.

The inverse Laplace transform of Eq. (33) is expressed as

$$\bar{f}(k_i,t)=-\bar{u}_{q1}(k_i,t)+\omega_i[\bar{u}_{q1}(k_i,t)\sin(\omega_it)] \tag{34}$$

where

$$\bar{u}_{q1}(k_i,t)\sin(\omega_it)=\int_0^t\bar{u}_{q1}(k_i,\tau)\sin[\omega_i(t-\tau)]d\tau \tag{35}$$

Substituting the first item of Eq. (24*a*) into Eq. (32), yields

$$\bar{u}_{q1}(k_i,t)=\bar{C}_1(k_i)p_a(t)+\bar{C}_2(k_i)p_b(t)+\bar{C}_3(k_i)d(t) \tag{36}$$

where $\bar{C}_j(k_i)=H[C_j(r)]$, $(j=1,2,3)$.

Substituting Eq. (36) into Eq. (35), and utilizing Eq. (34), gives

$$\bar{f}(k_i,t) = \bar{C}_1(k_i)I_{1i}(k_i,t) + \bar{C}_2(k_i)I_{2i}(k_i,t) + \bar{C}_3(k_i)I_{3i}(k_i,t) \quad (37)$$

where

$$I_{1i}(k_i,t) = -p_a(t) + \omega_i \int_0^t p_a(\tau)\sin[\omega_i(t-\tau)]d\tau \quad (38a)$$

$$I_{2i}(k_i,t) = -p_b(t) + \omega_i \int_0^t p_b(\tau)\sin[\omega_i(t-\tau)]d\tau \quad (38b)$$

$$I_{3i}(k_i,t) = -d(t) + \omega_i \int_0^t d(\tau)\sin[\omega_i(t-\tau)]d\tau \quad (38c)$$

Combining Eqs. (37) and (29), the dynamic solution for inhomogeneous Eq. (23) with homogeneous boundary conditions is given by

$$f(r,t) = \sum_{k_i} \frac{G_N(k_i r)}{F(k_i)}[\bar{C}_1(k_i)I_{1i}(k_i,t) + \bar{C}_2(k_i)I_{2i}(k_i,t) + \bar{C}_3(k_i)I_{3i}(k_i,t)] \quad (39)$$

Thus, utilizing Eqs. (15), (19), (22), and (39), the solution of the basic displacement equation of electromagnetoelastic motion in the transversely isotropic piezoelectric hollow sphere is expressed as

$$u(r,t) = B_1(r)p_a(t) + B_2(r)p_b(t) + B_3(r)d(t) + \sum_{k_i} \frac{r^{-0.5}G_N(k_i r)}{F(k_i)}[\bar{C}_1(k_i)I_{1i}(k_i,t) + \bar{C}_2(k_i)I_{2i}(k_i,t) + \bar{C}_3(k_i)I_{3i}(k_i,t)] \quad (40)$$

It is noted that in the above expression $d(t)$ still is an unknown function that is the relation to the electric displacement. It is necessary to determine $d(t)$ in the following. Integrating Eq. (7) and utilizing the corresponding electric boundary conditions (13b), yields

$$\phi(r,t) = \Phi_1(r)p_a(t) + \Phi_2(r)p_b(t) + \Phi_3(r)d(t) + \sum_i \Phi_{4i}(r)F_i(t) + \phi_a(t) \quad (41)$$

where

$$\Phi_1(r) = \frac{e_{11}}{\beta_{11}}\left[B_1(r) - B_1(a) - \sum_{k_i} \frac{[r^{-0.5}G_N(k_i r) - a^{-0.5}G_N(k_i a)]}{F(k_i)}\bar{B}_1(k_i)\right] + \frac{2e_{12}}{\beta_{11}}\int_a^r \frac{1}{r}\left[B_1(r) - \sum_{k_i} \frac{r^{-0.5}G_N(k_i r)}{F(k_i)}\bar{B}_1(k_i)\right]dr \quad (42a)$$

$$\Phi_2(r) = \frac{e_{11}}{\beta_{11}}\left[B_2(r) - B_2(a) - \sum_{k_i} \frac{[r^{-0.5}G_N(k_i r) - a^{-0.5}G_N(k_i a)]}{F(k_i)}\bar{B}_2(k_i)\right] + \frac{2e_{12}}{\beta_{11}}\int_a^r \frac{1}{r}\left[B_2(r) - \sum_{k_i} \frac{r^{-0.5}G_N(k_i r)}{F(k_i)}\bar{B}_2(k_i)\right]dr \quad (42b)$$

$$\Phi_3(r) = \frac{e_{11}}{\beta_{11}}\left[B_3(r) - B_3(a) - \sum_{k_i} \frac{[r^{-0.5}G_N(k_i r) - a^{-0.5}G_N(k_i a)]}{F(k_i)}\bar{B}_3(k_i)\right] + \frac{2e_{12}}{\beta_{11}}\int_a^r \frac{1}{r}\left[B_3(r) - \sum_{k_i} \frac{r^{-0.5}G_N(k_i r)}{F(k_i)}\bar{B}_3(k_i)\right]dr + \frac{1}{\beta_{11}}\frac{1}{r} \quad (42c)$$

$$\Phi_{4i}(r) = \frac{e_{11}}{\beta_{11}}\frac{[r^{-0.5}G_N(k_i r) - a^{-0.5}G_N(k_i a)]}{F(k_i)} + \frac{2e_{12}}{\beta_{11}}\int_a^r \frac{1}{r^{1.5}}\frac{G_N(k_i r)}{F(k_i)}dr \quad (42d)$$

$$F_i(t) = F_{1i}(t) + \bar{B}_3(k_i)\omega_i \int_0^t d(\tau)\sin[\omega_i(t-\tau)]d\tau \quad (42e)$$

$$F_{1i}(t) = \bar{B}_1(k_i)\omega_i \int_0^t p_a(\tau)\sin[\omega_i(t-\tau)]d\tau + \bar{B}_2(k_i)\omega_i \int_0^t p_b(\tau)\sin[\omega_i(t-\tau)]d\tau \quad (42f)$$

When $r=b$, Eq. (41) can be rewritten as

$$\phi_b(t) = \Phi_1(b)p_a(t) + \Phi_2(b)p_b(t) + \Phi_3(b)d(t) + \sum_i \Phi_{4i}(b)F_i(t) + \phi_a(t) \quad (43)$$

Substituting $t=0$ into Eq. (43), yields

$$d(0) = \frac{\phi_b(0) - \phi_a(0) - \Phi_1(b)p_a(0) - \Phi_2(b)p_b(0) - \sum_i \Phi_{4i}(b)F_i(0)}{\Phi_3(b)} \quad (44)$$

Substitution Eq. (42d) into Eq. (43), gives

$$\vartheta(t) = M_1 d(t) + \sum_i M_{2i} \int_0^t d(\tau)\sin[\omega_i(t-\tau)]d\tau \quad (45)$$

where

$$\vartheta(t) = \phi_b(t) - \phi_a(t) - \Phi_1(b)p_a(t) - \Phi_2(b)p_b(t) - \sum_i \Phi_{4i}(b)F_{1i}(t)$$

$$M_1 = \Phi_3(b), \quad M_{2i} = \Phi_{4i}(b)\bar{B}_3(k_i)\omega_i \quad (46)$$

It is seen that Eq. (45) is Volterra integral equation of the second kind (Kress [21]). In the following, Eq. (45) is solved by using the recursion formula based on linear interpolation function. Dividing the time interval $[0,t]$ into n subintervals, the discrete time points are $t_0=0,t_1,t_2\cdots t_n$. The interpolation function at the time interval $[t_{j-1},t_j]$ is expressed as

$$d(t) = \xi_j(t)d(t_{j-1}) + \eta_j(t)d(t_j), \quad (j=1,2\cdots n) \quad (47)$$

where

$$\xi_j(t) = \frac{t-t_j}{t_{j-1}-t_j}, \quad \eta_j(t) = \frac{t-t_{j-1}}{t_j-t_{j-1}}, \quad (j=1,2\cdots n) \quad (48)$$

Substituting Eq. (47) into Eq. (45), gives

$$\vartheta(t_j) = M_1 d(t_j) + \sum_i M_{2i} \sum_{k=1}^{j} [R_{ijk}d(t_{k-1}) + S_{ijk}d(t_k)] \quad (49)$$

where

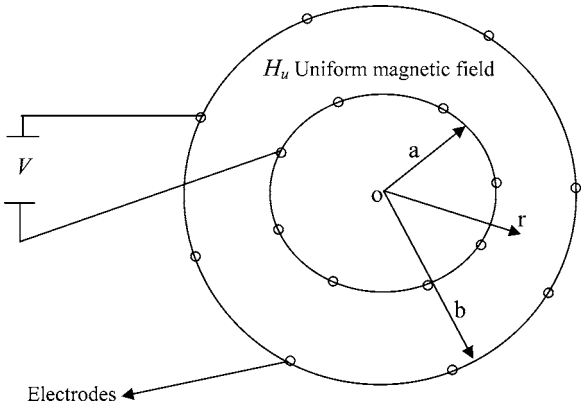


Fig. 1 A geometric graph of the piezoelectric hollow sphere placed in a uniform magnetic field, subjected to mechanical load and electric excitation

$$R_{ijk} = \int_{t_{k-1}}^{t_k} \xi_k(\tau)\sin[\omega_i(t-\tau)]d\tau \qquad (k = 1,2\cdots j, \quad j = 1,2\cdots n)$$
$$S_{ijk} = \int_{t_{k-1}}^{t_k} \eta_k(\tau)\sin[\omega_i(t-\tau)]d\tau, \tag{50}$$

Solving Eq. (49), gives

$$d(t_j) = \frac{\vartheta(t_j) - \sum_i M_{2i}\sum_{k=1}^{j-1}[R_{ijk}d(t_{k-1}) + S_{ijk}d(t_k)] - d(t_{j-1})\sum_i M_{2i}R_{ijj}}{M_1 + \sum_i M_{2i}S_{ijj}}, \qquad (j = 1,2\cdots n) \tag{51}$$

Substituting $d(0)$ in Eq. (44) into Eq. (51), $d(t_j)$, $(j=1,2\cdots n)$ can be determined step by step. Thus, the exact expression of the dynamic displacement $u(r,t)$ is obtained. The dynamic stresses $\sigma_r(r,t)$, $\sigma_\theta(r,t)$, the perturbation of magnetic field vector $h_u(r,t)$, the dynamic electric displacement $D_r(r,t)$, and the dynamic electric potential $\phi(r,t)$ are easily obtained from Eqs. (4) to (8).

4 Numerical Examples and Discussions

Electromagnetoelastic interaction in a transversely isotropic piezoelectric hollow sphere placed in a uniform magnetic field, subjected to mechanical load and electric excitation is considered. In numerical calculations, material constants for the transversely isotropic piezoelectric hollow sphere are taken as (Ding et al. [22]): $c_{11}=c_{33}=110.0$ GPa, $c_{22}=220$ GPa, $c_{12}=77.8$ GPa, $c_{13}=c_{23}=74.3$ GPa, $e_{12}=15.1(\mathrm{C/m^2})$, $e_{11}=e_{13}=-5.2(\mathrm{C/m^2})$, $\beta_{11}=5.62\times10^{-9}(\mathrm{C^2/Nm^2})$, and $\rho=4350(\mathrm{Kg/m^2})$. In all numerical examples, magnetic permeability is taken as $\mu=4\pi\times10^{-7}(\mathrm{H/m})$ and magnetic intensity is taken as $H_z=3\times10^8(\mathrm{A/m})$, the dimensionless radial coordinate $R=(r-a)/(b-a)$, the dimensionless time $\tau=C_Lt/(b-a)$, the response time is taken as $\tau=10$, and the ratio of the external radii to the internal radii is taken as $b/a=2$. (See Fig. 1.)

Example 1. Electromagnetoelastic interaction in the transversely isotropic piezoelectric hollow sphere, the boundary conditions are

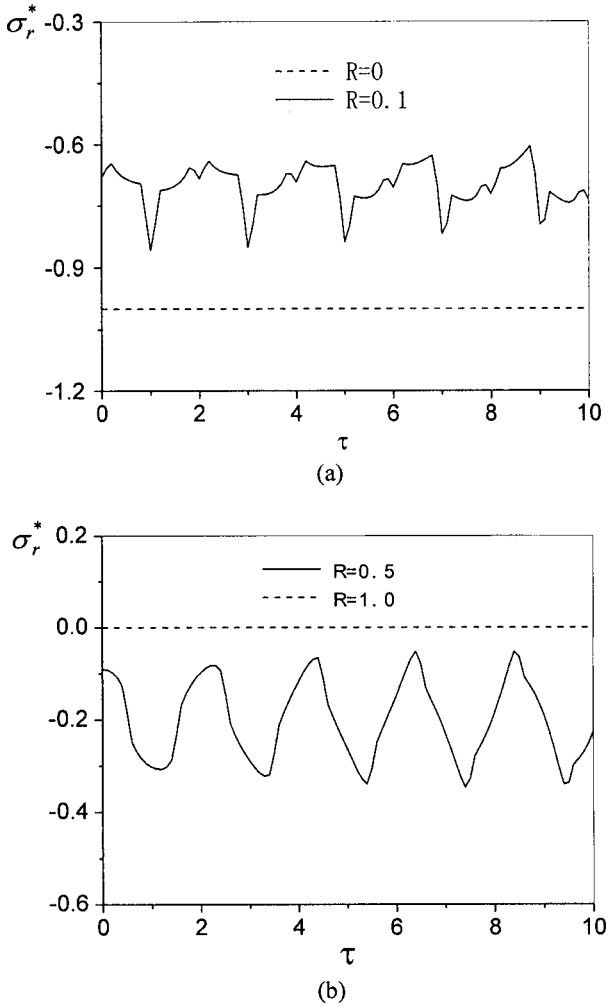


Fig. 2 Response histories of the dynamic radial stresses σ_r^*, where $R=(r-a)/(b-a)$, $\tau=C_Lt/(b-a)$, $p_a(t)=p_0$, and $\sigma_r^*=\sigma_r/p_0$. (*a*) $R=0$ and $R=0.1$; (*b*) $R=0.5$ and $R=1$.

$$\sigma_r(a,t) = p_0H(t), \quad \sigma_r(b,t) = 0, \quad \phi_a(a,t) = 0, \quad \phi_b(b,t) = 0 \tag{52}$$

where $H(t)$ expresses the Heaviside function, and $\sigma_i^*=\sigma_i/p_0$ $(i=r,\theta)$, $h_u^*=h_u/(p_0H_u)$, $D_r^*=D_r/p_0$, and $\phi^*=\phi/p_0$ are introduced in Figs. 2–6.

In this example, the response time is taken as $\tau=10$, when the responded time is taken as $t\geqslant1$, because of the small wall thickness $a/b=0.5$, and the effects of wave reflected between the inner wall and outer wall have been produced. Figures 2 show, respectively, the response histories of radial stresses at $R=0$, 0.1, 0.5, and 1. From Figs. 2, it is seen that the radial stresses at $R=0$ and $R=1$ are, respectively, equal to -1 and zero, which satisfies the internal and external boundary conditions (52), so the correction of the numerical results is clarified in this respect. The dynamic response histories of radial stresses at $R=0.1$ and $R=0.5$ oscillate dramatically because of the reflected effect of wave between the inner wall and outer wall. Figures 3–5 show, respectively, the dynamic response histories of hoop stresses, perturbation of magnetic field vector, and electric displacement at $R=0$, 0.5, and 1, it is seen easily from Figs. 3–5 that the peak values of hoop stresses, perturbation of magnetic field vector, and electric displacement decrease gradually from inner wall to outer wall at the identical time τ. From Fig. 6, it is also seen that the electric potential ϕ^* at the internal and external boundary equal zero, which satisfy the

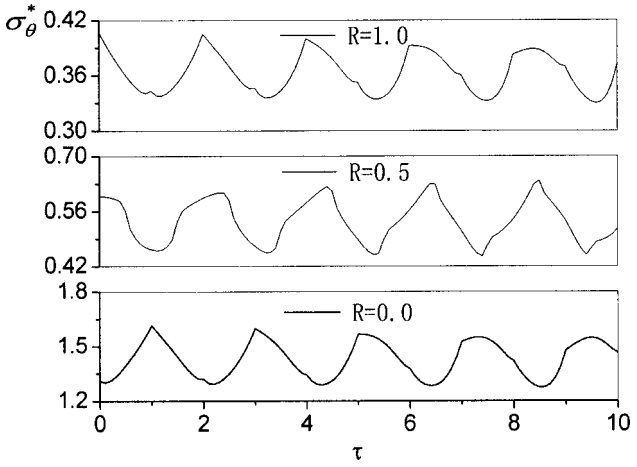


Fig. 3 Response histories of the dynamic hoop stresses σ_θ^* at $R=0$, $R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $p_a(t)=p_0$, and $\sigma_\theta^*=\sigma_\theta/p_0$

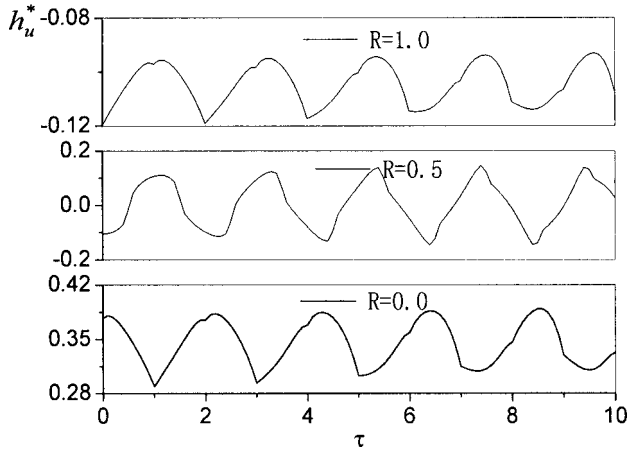


Fig. 4 Response histories of perturbation of magnetic field vector h_u^* at $R=0$, $R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $p_a(t)=p_0$, and $h_u^*=h_u/(H_u p_0)$

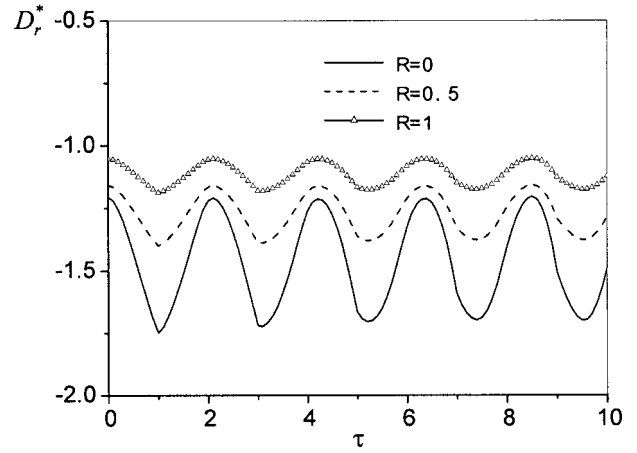


Fig. 5 Response histories of the dynamic electric displacements D_r at $R=0$, $R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $p_a(t)=p_0$, and $D_r^*=D_r/p_0$

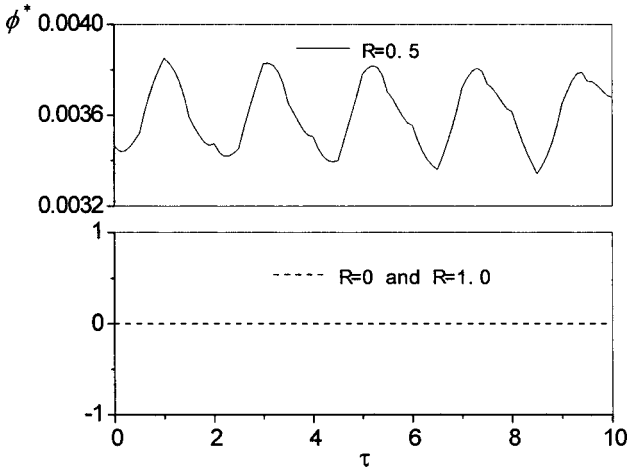


Fig. 6 Distributions of the dynamic electric potentials ϕ^* at $R=0$, $R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $p_a(t)=p_0$, and $\phi^*=\phi/p_0$

prescribed electric boundary conditions in Eq. (52), and the dynamic response histories of electric potential at $R=0.5$ are similar to that of the radial stress as shown in Fig. 2(*b*).

Example 2. Consider that the electromagnetoelastic interaction in the transversely isotropic piezoelectric hollow sphere is induced by electric excitation. The corresponding boundary conditions are expressed as

$$\sigma_r(a,t)=0, \quad \sigma_r(b,t)=0, \quad \phi_a(a,t)=0, \quad \phi_b(b,t)=\phi_0 H(t) \tag{53}$$

In the calculation, $\sigma_i^*=\sigma_i/\phi_0$ $(i=r,\theta)$, $h_u^*=h_u/(\phi_0 H_u)$, $D_r^*=D_r/\phi_0$, and $\phi^*=\phi/\phi_0$ are introduced.

From Figs. 7 and 11, it is seen that the radial stresses and the electric potential at the boundaries $R=0$, 1 satisfy the given boundary conditions (53). Figures 7–11 show that except the points at the given boundary condition, dynamic responses at other points oscillate dramatically around the corresponding quasistatic values because of the effects of wave reflected between the inner wall and outer wall. From Fig. 7, it is seen that the dynamic response histories of radial stresses at $R=0.1$ and $R=0.5$ are similar to that of the radial stress as shown in Figs. 2, Comparing Fig. 7 and Figs. 2, it is seen that the respondent amplitude of radial

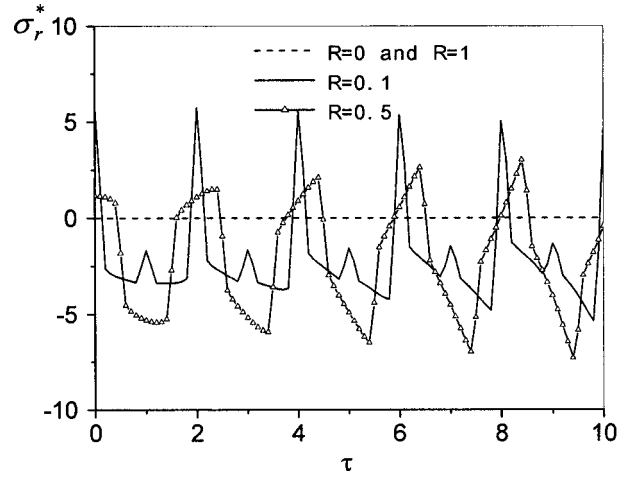


Fig. 7 Response histories of the dynamic radial stresses σ_r^* at $R=0$, $R=0.1$,$R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $\phi_b(t)=\phi_0$, and $\sigma_r^*=\sigma_r/\phi_0$

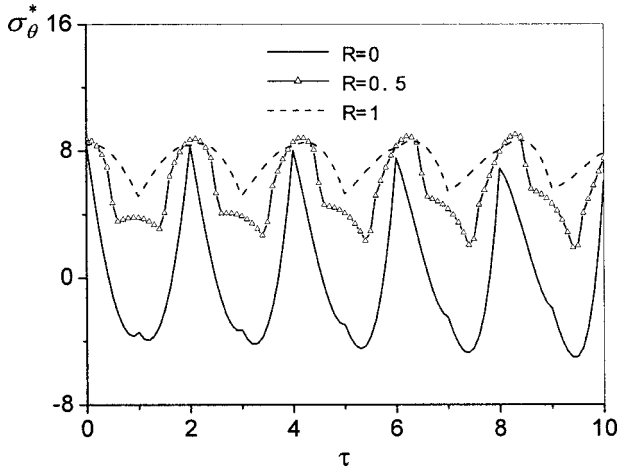


Fig. 8 Response histories of the dynamic hoop stresses σ_θ^* at $R=0$, $R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $\phi_b(t)=\phi_0$, and $\sigma_\theta^*=\sigma_\theta/\phi_0$

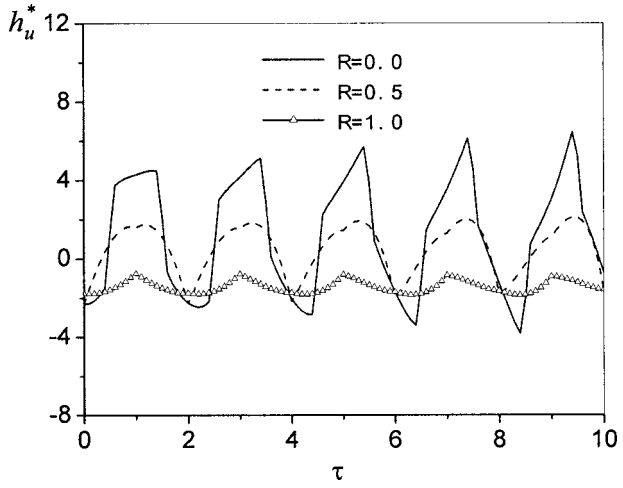


Fig. 9 Response histories of perturbation of magnetic field vector h_u^* at $R=0$, $R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $\phi_b(t)=\phi_0$, and $h_u^*=h_u/(H_u\phi_0)$

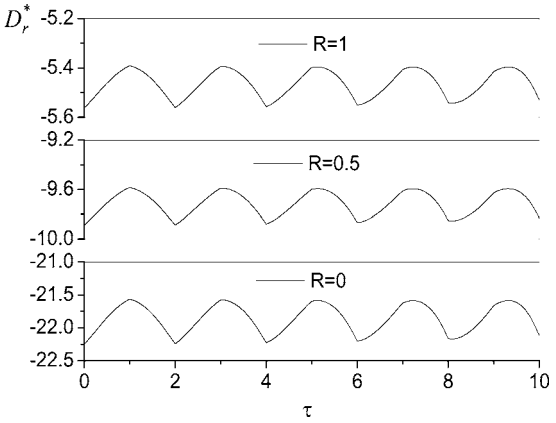


Fig. 10 Response histories of the dynamic electric displacements D_r at $R=0$, $R=0.5$, and $R=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $\phi_b(t)=\phi_0$, and $D_r^*=D_r/\phi_0$

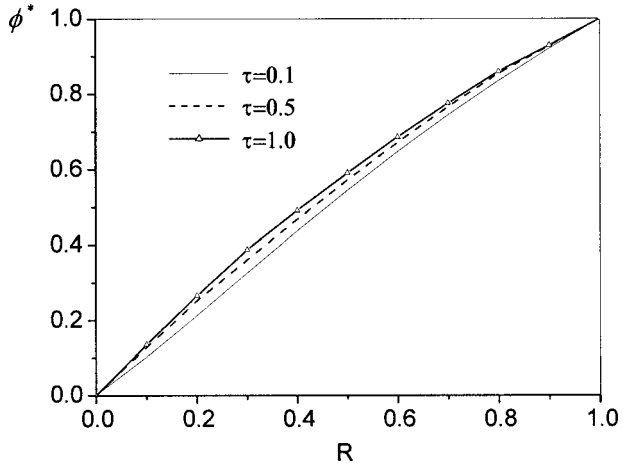


Fig. 11 Distributions of the dynamic electric potentials ϕ^* at $\tau=0.1$, $\tau=0.5$, and $\tau=1$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $\phi_b(t)=\phi_0$, and $\phi^*=\phi/\phi_0$

stress caused by the sudden unit electric potential is larger than that caused by the sudden unit pressure. Figures 8–10 show, respectively, the dynamic response histories of hoop stresses, perturbation of magnetic field vector and electric displacement at $R=0$, 0.5, and 1. From Fig. 8, it is seen that the amplitude of hoop compression stress caused by the sudden unit electric potential is smaller than that of hoop tension stress. It is seen easily in Fig. 9 that perturbation of magnetic field vector show oscillate around the vicinity of zero value at different radial point. The dynamic response histories of electric displacement at $R=0$, 0.5, and 1 are negative as shown in Fig. 10. Figure 11 depicts the response of electric potential ϕ^* along radius is weak non-linear at different nondimensional time τ.

Example 3. For the sake of comparison, we consider a purely elastic hollow sphere with the same material constants and boundary conditions as given Example 1. In the calculation, all dimensionless quantities are introduced in Figs. 12 and 13 as given above.

Figures 12 and 13 show the responses of σ_r^* and σ_θ^* at $R=0.5$ (the middle surface) in the piezoelectric and elastic hollow sphere due to a sudden internal mechanical load, respectively. From Fig. 12, we can see the peak values of radial dynamic stresses in the piezoelectric hollow sphere are larger than those in the elastic one. The response curves of the piezoelectric and elastic hollow sphere are different with each other. It is seen easily in Fig. 13 that the

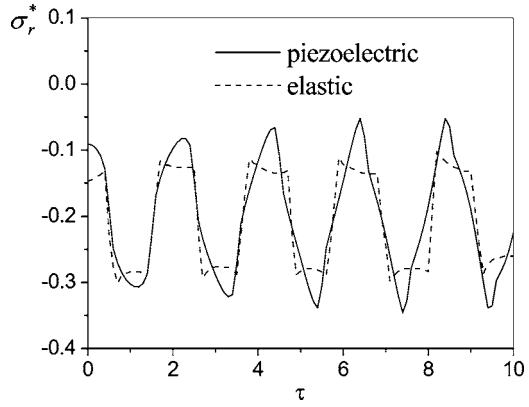


Fig. 12 Response histories of the dynamic radial stresses σ_r^* at $R=0.5$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $p_a(t)=p_0$, and $\sigma_r^*=\sigma_r/p_0$

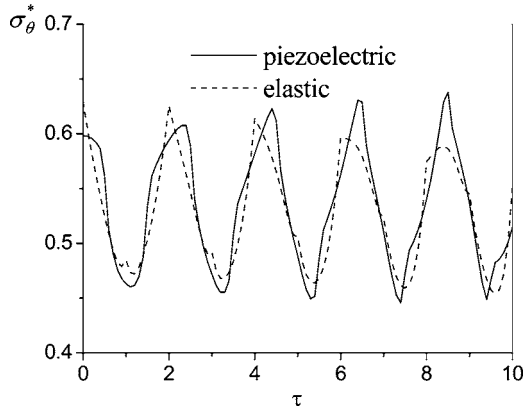


Fig. 13 Response histories of the dynamic hoop stresses σ_θ^* at $R=0.5$, where $R=(r-a)/(b-a)$, $\tau=C_L t/(b-a)$, $p_a(t)=p_0$, and $\sigma_\theta^*=\sigma_\theta/p_0$

difference is little between the response of the piezoelectric hollow sphere and that of the elastic one at the middle surface.

Example 4. Although the dynamic responses of coupled fields have been studied by a number of authors, no published experiment results can be used for a comparison with the present model. To our knowledge, no experimental results on the problem of dynamic response of the transversely isotropic piezoelectric hollow sphere placed in a uniform magnetic field, subjected to mechanical load and electric excitation are available in the literature. This is apparently due to the fact that the experimental research on the electromagnetoelastic dynamic response of a piezoelectric hollow sphere remains a formidable task.

In order to prove further the correctness of analytical results in the paper, the present method can be applied to solve the dynamic problem of spherically symmetric motion in piezoelectric hollow spheres. For ease of comparison with Ref. [8], the same dynamic problem of spherically symmetric motion in the piezoelectric hollow sphere is considered and the same material parameters are taken as Ref. [8]. The response of radial stress σ_r at $R=0.5$ and the response of hoop stress σ_θ at $R=0$ are depicted in Figs. 14, respectively, which appears in that the results from the two different methods are nearly the same.

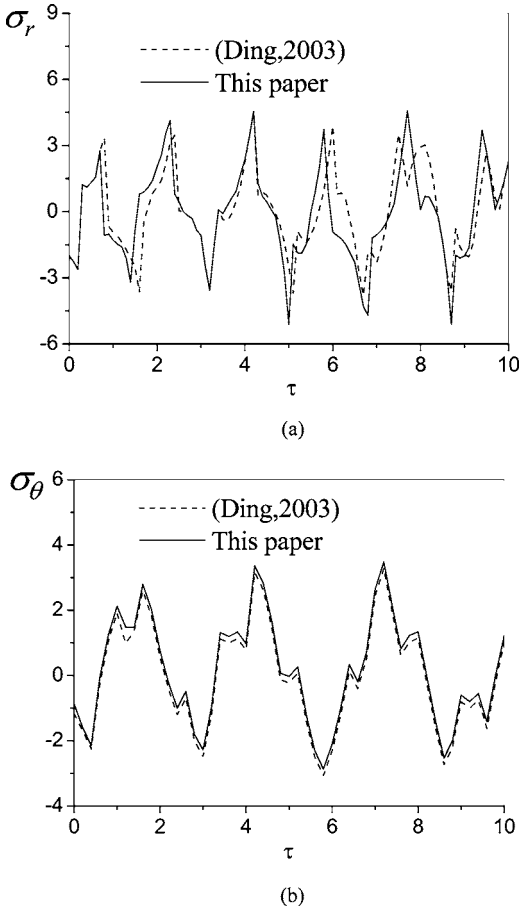


Fig. 14 (*a*) Response histories of dynamic radial stress σ_r at $R=0.5$; (*b*) response histories of dynamic hoop stress σ_θ at $R=0$

5 Conclusions

1. Because of the interaction between elastic deformation, electric field, and magnetic field, a sudden mechanical load induces the response of perturbation of magnetic field vector, electric displacement, and electric potential in a piezoelectric hollow sphere. Likewise, a sudden electric potential also causes the dynamic stresses responses and perturbation of magnetic field vector in the piezoelectric hollow sphere.
2. Comparing Example 1 with Example 2, it is seen that the response histories and distributions of stresses, perturbation of magnetic field vector, electric displacement, and electric potential in the transversely isotropic piezoelectric hollow sphere are obviously different for two kinds of boundary conditions which are, respectively, shown in Eqs. (52) and (53). Thus, it is possible to control the response histories and distribution of stresses in the transversely isotropic piezoelectric hollow sphere by applying a suitable mechanical load and electric excitation to the structure, or to assessment the response histories and distribution of stress in the transversely isotropic piezoelectric hollow sphere by measuring the response histories of electric potential in the structures.
3. It is concluded from the above analyses and discussions that the presented method is valid. So the solving method may be used as a reference to solve other dynamic coupled problems in transversely isotropic piezoelectric hollow spheres in a uniform magnetic field. From the knowledge of the response histories of dynamic stresses, perturbation of magnetic field vector, electric displacement and electric potential in transversely isotropic piezoelectric hollow spheres, one can design various electromagnetoelastic elements placed in a uniform magnetic field, subjected to mechanical load and electric excitation to meet specific engineering requirements.

Acknowledgment

The authors appreciate the valuable comments from the reviewers.

Nomenclature

ϵ_{ij} = components of strains
$\vec{U}, u$ = displacement vector and radial displacement $[m]$
c_{ij} = elastic constants (N/m^2)
e_{ij} = piezoelectric constants (C/m^2)
β_{11} = dielectric constants (C^2/Nm^2)
σ_i, D_r = components of stresses (N/m^2) and radial electric displacement (C/m^2)
$\phi(r,t)$ = electric potential (W/A)
ρ = mass density (kg/m^3)
t = time variable (s)
r = radial variable (m)

$\vec{H}$ = magnetic intensity vector
$\vec{h}$ = perturbation of magnetic field vector
$\vec{J}$ = electric current density vector
$\vec{e}$ = perturbation of electric field vector
μ = magnetic permeability (H/m)
f_u = Lorentz's force ($kg/m^2\,s^2$)
a,b = internal and external radii of piezoelectric hollow sphere (m)
C_L = electromagnetoelastic wave speed (m/s)
ω = the inherent frequency of the piezoelectric hollow sphere (1/s).

Appendix

Applying Laplace transform to the two sides of Eq. (31)

$$\text{Laplace}[-k_i^2\bar{f}(k_i,t)] = \frac{1}{C_L^2}\text{Laplace}\left[\frac{\partial^2\bar{f}(k_i,t)}{\partial t^2} + \frac{\partial^2\bar{u}_{q1}(k_i,t)}{\partial t^2}\right] \tag{A1}$$

The left-hand side of Eq. (A1) may become

$$\text{Laplace}[-k_i^2\bar{f}(k_i,t)] = -k_i^2[\bar{f}^*(k_i,L)] \tag{A2}$$

Where L is the parameter of Laplace transform.

The right-hand side of Eq. (A1) may become

$$\text{Laplace}\left[\frac{\partial^2\bar{f}(k_i,t)}{\partial t^2} + \frac{\partial^2\bar{u}_{q1}(k_i,t)}{\partial t^2}\right]$$
$$= L^2\bar{f}(k_i,t) - L\bar{f}(k_i,0) - \bar{f}'(k_i,0) + L^2\bar{u}_{q1}(k_i,t) - L\bar{u}_{q1}(k_i,0) - \bar{u}'_{q1}(k_i,0) \tag{A3}$$

and utilizing the initial condition (23c), yields

$$\bar{f}(k_i,0) + \bar{u}_{q1}(k_i,0) = 0, \quad \bar{f}'(k_i,0) + \bar{u}'_{q1}(k_i,0) = 0 \tag{A4}$$

Substituting Eq. (A4) into Eq. (A3), and utilizing Eq. (A2), gives

$$-k_i^2[\bar{f}^*(k_i,L)] = \frac{1}{C_L^2}[L^2\bar{f}(k_i,t) + L^2\bar{u}_{q1}(k_i,t)] \tag{A5}$$

By means of the natural frequencies $\omega_i = C_L k_i$, we have

$$-(\omega_i^2 + L^2)\bar{f}^*(k_i,L) = L^{2^*}u_{q1}(k_i,L) \tag{A6}$$

Thus, Eq. (33) can be obtained.

References

[1] Hu, H. C., 1954, "On the General Theory of Elasticity for a Spherically Isotropic Medium," Acta Sci. Sin., **3**, pp. 247–260.
[2] Hata, T., 1991, "Thermal Shock in a Hollow Sphere Caused by Rapid Uniform Heating," ASME J. Appl. Mech., **58**, pp. 64–69.
[3] Hata, T., 1993, "Stress-Focusing Effect in a Uniformly Heated Transversely Isotropic Sphere," Int. J. Solids Struct., **30**, pp. 1419–1428.
[4] Hata, T., 1997, "Stress-Focusing Effect Due to an Instantaneous Concentrated Heat Source in a Sphere," J. Therm. Stresses, **20**, pp. 269–279.
[5] Sherief, H. H., and Ezzat, M. A., 1996, "Thermal-Shock Problem in Magnetothermoelaticity with Thermal Relaxation," Int. J. Solids Struct., **33**, pp. 4449–4459.
[6] Wang, X., 2000, "Dynamic Thermostress-concentration Effect in a Spherically Isotropic Sphere," Acta Mech. Sin., **32**, pp. 245–250 (in Chinese).
[7] Heyliger, P., and Wu, Y. C., 1999, "Electroelastic Fields in Layered Piezoelectric Spheres," Ind. J. Eng. Mater. Sci., **37**, pp. 143–161.
[8] Ding, H. J., Wang, H. M., and Chen, W. Q., 2003, "Dynamic Response of a Pyroelectric Hollow Sphere Under Radial Deformation," Eur. J. Mech. A/Solids, **22**, pp. 617–631.
[9] Ding, H. J., Wang, H. M., and Chen, W. Q., 2003, "Dynamic Response of a Functionally Graded Pyroelectric Hollow Sphere for Spherically Symmetric Problems," Int. J. Mech. Sci., **45**, pp. 1029–1051.
[10] Dai, H. L., and Wang, X., 2004, "Dynamic Responses of Piezoelectric Hollow Cylinders in an Axial Magnetic Field," Int. J. Solids Struct., **41**, pp. 5231–5246.
[11] Dai, H. L., and Wang, X., 2005, "Dynamic Wave Propagation in Piezoelectric Hollow Spheres Subjected to Thermal Shock and Electric Excitation," Struct. Eng. Mech., pp. 441–457.
[12] Dai, H. L., and Wang, X., 2005, "Stress Wave Propagation in Laminated Piezoelectric Spherical Shells Under Thermal Shock and Electric Excitation," Eur. J. Mech. A/Solids, **24**, pp. 263–276.
[13] Lekhnitskii, S. G., 1981, *Theory of Elasticity of an Anisotropic Body*, Mir Publishers, Moscow.
[14] Cinelli, G., 1965, "An Extension of the Finite Hankel Transform and Application," Ind. J. Eng. Mater. Sci., **3**, pp. 539–559.
[15] Sinha, D. K., 1962, "Note on the Radial Deformation of a Piezoelectric, Polarized Spherical Shell with a Symmetrical Temperature Distribution," J. Acoust. Soc. Am., **34**, pp. 1073–1075.
[16] Chen, W. Q., and Shioya, T., 2001, "Piezothermoelastic Behavior of a Pyroelectric Spherical Shell," J. Therm. Stresses, **24**, pp. 105–120.
[17] Ezzat, M. A., 1997, "Generation of Generalized Magnetothermoelastic Waves by Thermal Shock in a Perfectly Conducting Half Space," J. Therm. Stresses, **20**(6), pp. 633–647.
[18] John, K. D., 1984, *Electromagnetics*, McGraw-Hill, New York.
[19] Dunn, M. L., and Taya, M., 1994, "Electroelastic Field Concentrations in and Around Inhomogeneties in Piezoelectric Solids," Appl. Mech. Rev., **61**, pp. 474–475.
[20] Eringen, A. C., and Suhubi, E. S., 1975, *Electrodynamics*, Vol. 2. Academic Press, New York, p. pp. 440.
[21] Kress, R., 1989, *Linear Integral Equation* (Applied Mathematical Sciences, Vol. 82), Springer-Verlag, Berlin.
[22] Ding, H. J., Wang, H. M., and Hou, P. F., 2003, "The Dynamic Response of Piezoelectric Hollow Cylinders for Axisymmetric Plane Strain Problems," Int. J. Solids Struct., **40**, pp. 105–123.

Buckling and Sensitivity to Imperfection of Conical Shells Under Dynamic Step-Loading

Mahmood Jabareen
e-mail: cvjmah@techunix.technion.ac.il

Izhak Sheinman
e-mail: cvrnrsh@techunix.technion.ac.il

Faculty of Civil and Environmental Engineering,
Technion—Israel Institute of Technology,
Haifa 32000,
Israel

A general nonlinear dynamic analysis, based on Donnell's shell-type theory, is developed for an arbitrary imperfect isotropic conical shell. It is used for studying dynamic stability and imperfection sensitivity under dynamic step loading. The nonlinear dynamic time history and the sensitivity behavior are examined in parametric terms over a wide range of aspect ratios. A general symbolic code (using the MAPLE compiler) was programmed to create the differential operators. By this means the Newmark discretization, Galerkin procedure, Newton-Raphson iteration, and finite difference scheme are applied for automatic development of an efficient FORTRAN code for the parametric study, and for examining the correlation of the sensitivity behavior between two different dynamic stability criteria. An extensive parametric study of the effect of the cone semi-vertex angle on the stability and sensitivity to imperfection under dynamic step loading was carried out. It was found that the dynamic buckling can indeed be derived from the nonlinear static solution. [DOI: 10.1115/1.2178836]

Introduction

The behavior of shell-like structures under buckling is characterized by a limit point with a snap through phenomenon. Thus these structures are sensitive to initial imperfection, and insight into the post-buckling stage is essential in lightweight structural design.

A conical shell is an appropriate representative of the entire range of this sensitivity: from the extra-sensitive cylinder to the completely insensitive annular plate. Hence, the transition from sensitivity to insensitivity can be followed by varying the cone semi-vertex angle. Two approaches were applied in investigating the sensitivity: (I) parametric study of the conical shell in terms of its initial post-buckling (see [1]) and (II) tracing of the entire nonlinear equilibrium path with emphasis on the level and direction of change of the stiffness during loading (see [2]). In both cases the study was confined to static loading, but most loads on structural systems include dynamic effects, whose contribution to the sensitivity behavior is of vital importance.

The object of the present paper is investigation of the buckling and sensitivity of imperfect conical shells under dynamic step loading. The term "dynamic stability" encompasses many classes of problems, such as parametric resonance, pulse buckling, and nonconservativeness. The study here concerns a sudden large increase in the response with displacement resulting from a small increase of the load parameter. The main object of the study is prediction of the critical conditions under which dynamic instability may set in. A system is considered stable when its displacements are bounded and the critical load is the lowest at which its motion becomes unstable.

Dynamic buckling of conical shells has not been extensively discussed so far in the literature. Rossettos and Parisse [3] studied the dynamic response of conical panels, Srinivasan and Krishnan [4,5]—that of layered and stiffened conical shells, and Ganapathi et al. [6]—the dynamic instability of truncated conical shells under periodic in-plane loading. None of them considered the effect of initial imperfection on the dynamic response in terms of the dynamic buckling load. Dynamic buckling of imperfect conical shells, using a six-degrees-of-freedom model, was investigated by Shiau et al. [7], but this model is inappropriate for the general dynamic stability and sensitivity of such shells. Dumir et al.'s [8] study was confined to axisymmetric dynamic buckling of shallow conical shells.

The present paper deals with imperfect conical shells under sudden axial compression in terms of their sensitivity to the amplitude and shape of the imperfection. Furthermore, two dynamic buckling criteria are compared: full nonlinear dynamicity throughout [9] and pure nonlinear staticity throughout [10,11]. The purpose of the comparison is to show the advantage of applying static criteria to dynamic stability.

The nonlinear differential operators of the equations of motion are derived with the aid of Hamilton's principle and using Donnell's kinematic relations. The solution procedure is based on discretization of the unknown functions in time by the Newmark method, with expansion in Fourier series in the circumferential direction and finite differences in the axial direction. The Galerkin procedure is then applied for minimizing the error due to truncation of the Fourier series. Finally, Newton-Raphson iteration is applied to obtain the solution. The overall procedure is written as a special symbolic program using the MAPLE compiler, whose output is the FORTRAN code NDAICS (Nonlinear Dynamic Analysis of Imperfect Conical Shells), used for parametric study over a wide range of aspect ratios. It was found that the static criterion for dynamic stability under axial step loading is representative.

Dynamic Stability Criteria

In the criteria suggested by Budiansky and Roth [9], the dynamic buckling load is defined as the level at which a large increase occurs in the deflection amplitude when the nonlinear equations of motion of the system are solved for different load levels.

The other criterion applied in this paper is that suggested by Hoff and Bruce [10] and Simitses [11], whereby the dynamic critical load is determined by the static post-limit load under which the modified total potential energy is zero. This criterion corresponds to the lower bound of the critical conditions for external applied step load (sudden, with infinite duration), was ap-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 25, 2005; final manuscript received December 7, 2005. Review conducted by N. Triantafyllidis. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

plied for cylindrical shells by Simitses and Sheinman [12], and is used here for a conical shell, over the whole range of vertex half-angle, for examining the imperfection sensitivity behavior. Due to its importance and for the sake of completeness, it will be recapitulated here in brief:

On the basis of the conservation of energy principle, the total energy is a constant, C:

$$\Pi(u,v,w;\lambda) + T(\dot{u},\dot{v},\dot{w}) = C \tag{1}$$

where Π and T are the total potential and kinetic energy functionals, respectively; $\dot{(\,)}$ is the time derivative; u, v, and w are the displacement functions in the meridional, circumferential, and normal directions, respectively, and λ is the load parameter. The constant, C, is readily reduced to a zero value by properly defining a modified total potential energy:

$$\Pi_{\text{mod}}(u,v,w;\lambda) + T(\dot{u},\dot{v},\dot{w}) = 0 \tag{2}$$

The modified total potential excludes the amount which is directly related to the axial (nonbuckling) mode. Motion is possible only when the modified total potential is nonpositive:

$$\Pi_{\text{mod}}(u,v,w;\lambda) \leqslant 0 \tag{3}$$

and the dynamic criterion is then defined, in terms of the applied load $\lambda=\bar{\lambda}$, through static analysis for which:

$$\Pi_{\text{mod}}(u^*,v^*,w^*;\bar{\lambda}) = 0 \tag{4}$$

$\bar{\lambda}$ denoting the upper bound of all loads for which the motion remains "unbuckled" and the system remains dynamically stable—see Ref. [13].

Governing Equations

The equations of motion governing the nonlinear dynamic behavior of imperfect conical shells are based on the Kirchhoff-Love hypotheses and obtained with the aid of Hamilton's principle:

$$\delta\int_{t_1}^{t_2}(T-\Pi)dt = 0 \tag{5}$$

where Π is the total potential of the strain and external loads functional, T is the kinetic energy functional, and $[t_1,t_2]$ is the time interval. Let (x,θ,z) be the coordinate in the meridional, circumferential, and normal directions, respectively (see Fig. 1).

Using Gauss' theorem, with the rotary inertia neglected, the following nonlinear equations of motion are obtained:

$$-\rho h u_{,tt} + N_{xx,x} + \frac{N_{x\theta,\theta}}{r(x)} + \frac{\sin(\alpha)}{r(x)}(N_{xx}-N_{\theta\theta}) + q_u = 0$$

$$-\rho h v_{,tt} + \frac{N_{\theta\theta,\theta}}{r(x)} + N_{x\theta,x} + 2\frac{\sin(\alpha)}{r(x)}N_{x\theta} + q_v = 0$$

$$\begin{aligned}-\rho h w_{,tt} &+ M_{xx,xx} + \frac{2M_{x\theta,x\theta}}{r(x)} + \frac{M_{\theta\theta,\theta\theta}}{r^2(x)} - \frac{\cos(\alpha)}{r(x)}N_{\theta\theta} \\ &+ \frac{\sin(\alpha)}{r(x)}\left(2M_{xx,x} - M_{\theta\theta,x} + 2\frac{M_{x\theta,\theta}}{r(x)}\right) + \frac{1}{r(x)}[r(x)N_{xx}(w_{,x}+\hat{w}_{,x}) \\ &+ N_{x\theta}(w_{,\theta}+\hat{w}_{,\theta})]_{,x} + \frac{1}{r(x)}\left(N_{x\theta}(w_{,x}+\hat{w}_{,x}) + \frac{N_{\theta\theta}}{r(x)}(w_{,\theta}+\hat{w}_{,\theta})\right)_{,\theta} \\ &+ q_w = 0\end{aligned} \tag{6}$$

with the following boundary conditions:

$$N_{xx} = \bar{N}_{xx} \quad \text{or } u = \bar{u}$$

$$N_{x\theta} = \bar{N}_{x\theta} \quad \text{or } v = \bar{v}$$

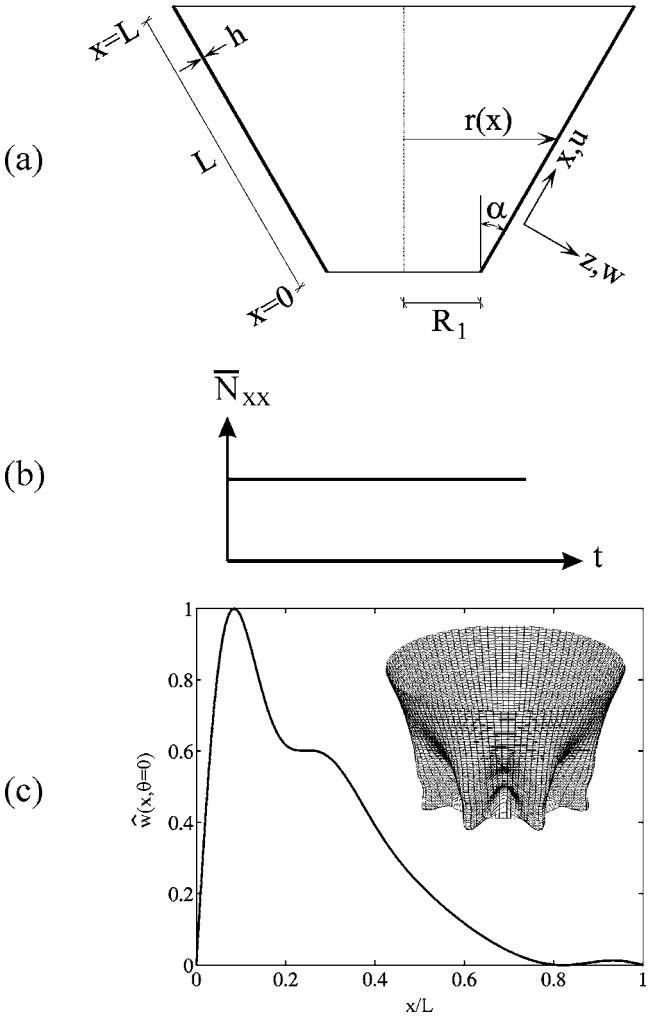


Fig. 1 (*a*) Geometry and sign convention, (*b*) step axial load, and (*c*) imperfection shape as the buckling mode

$$\begin{aligned}M_{xx,x} &+ \frac{2M_{x\theta,\theta}}{r(x)} + \frac{\sin(\alpha)}{r(x)}(M_{xx}-M_{\theta\theta}) + N_{xx}(w_{,x}+\hat{w}_{,x}) \\ &+ \frac{N_{x\theta}}{r(x)}(w_{,\theta}+\hat{w}_{,\theta}) = \bar{Q} \quad \text{or } w = \bar{w}\end{aligned}$$

$$M_{xx} = \bar{M}_{xx} \quad \text{or } w_{,x} = \bar{w}_{,x} \tag{7}$$

where u, v, and w are the components of the displacements functions of the shell in the x, θ, and z directions, respectively; $\hat{w}$ is the initial geometric imperfection shape; $(\,)_{,x}$, $(\,)_{,\theta}$, and $(\,)_{,t}$ denote the derivatives with respect to the axial and the circumferential coordinate and to time, respectively; $r(x)=R_1+x\sin(\alpha)$, α and ρ are the radius, cone semi-vertex angle, and mass density of the shell, respectively; q_u, q_v, and q_w are the external distributed loading in the three directions, respectively; and $\bar{(\,)}$ are the external applied forces or displacements at the boundaries.

The internal forces resultants $\{N\}=\{N_{xx},N_{\theta\theta},N_{x\theta}\}^T$ and $\{M\}=\{M_{xx},M_{\theta\theta},M_{x\theta}\}^T$ are defined by the following isotropic constitutive relation:

$$\begin{Bmatrix}\{N\}\\ \{M\}\end{Bmatrix} = \begin{bmatrix}[A] & [0]\\ [0] & [D]\end{bmatrix}\begin{Bmatrix}\{\varepsilon\}\\ \{\kappa\}\end{Bmatrix} \tag{8}$$

where $[A]$ and $[D]$ are the membrane and flexural stiffness coefficients, respectively. The membrane strains, ε, and the change of

curvature, κ, are defined via Donnell's [14] and Von Karman's approximations:

$$\{\varepsilon\} = \begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{\theta\theta} \\ \gamma_{x\theta} \end{Bmatrix} = \begin{Bmatrix} u_{,x} + \frac{1}{2} w_{,x}(w_{,x} + 2\hat{w}_{,x}) \\ \frac{v_{,\theta}}{r(x)} + \frac{\cos(\alpha)}{r(x)} w + \frac{\sin(\alpha)}{r(x)} u + \frac{w_{,\theta}}{2r(x)} \left(\frac{w_{,\theta}}{r(x)} + 2\frac{\hat{w}_{,\theta}}{r(x)} \right) \\ \frac{u_{,\theta}}{r(x)} + v_{,x} - \frac{\sin(\alpha)}{r(x)} v + \frac{w_{,x} w_{,\theta}}{r(x)} + \frac{\hat{w}_{,\theta} w_{,x}}{r(x)} + \frac{\hat{w}_{,x} w_{,\theta}}{r(x)} \end{Bmatrix}$$

$$\{\kappa\} = \begin{Bmatrix} \kappa_{xx} \\ \kappa_{\theta\theta} \\ \kappa_{x\theta} \end{Bmatrix} = \begin{Bmatrix} -w_{,xx} \\ -\frac{w_{,\theta\theta}}{r^2(x)} - \frac{\sin(\alpha)}{r(x)} w_{,x} \\ -\frac{w_{,x\theta}}{r(x)} + \frac{\sin(\alpha)}{r^2(x)} w_{,\theta} \end{Bmatrix} \tag{9}$$

Using the constitutive relations, Eqs. (8), the equations of motion and the appropriate boundary conditions (Eqs. (6) and (7)) can be rewritten in terms of the acceleration and the displacement components as:

$$\phi_p(\ddot{\mathbf{z}}, \mathbf{z}) = 0, \quad p = 1,2,3 \tag{10}$$

where $\mathbf{z}=\{u,v,w\}$ and ϕ_p consists of linear, quadratic, and cubic differential operators:

$$\begin{aligned} \phi_p = {} & L_p^1(u) + L_p^2(v) + L_p^3(w) + L_p^4(\hat{w},u) + L_p^5(\hat{w},v) + L_p^6(\hat{w},w) \\ & + L_p^7(u,w) + L_p^8(v,w) + L_p^9(w,w) + L_p^{10}(\hat{w},\hat{w},w) + L_p^{11}(\hat{w},w,w) \\ & + L_p^{12}(w,w,w) + q_p = 0, \quad p = 1,2,3 \end{aligned} \tag{11}$$

where $L_p^e(Q)$, $L_p^e(Q,S)$, and $L_p^e(Q,S,T)$ are the linear, quadratic, and cubic differential operators (see [15]):

$$L_p^e(Q) = \delta_p^e \rho Q_{,tt} + \sum_{i=0}^{4} \sum_{j=0}^{4-i} \Re_{ij}^{p,e} \frac{\partial^{(i+j)} Q}{\partial x^{(i)} \partial \theta^{(j)}}$$

$$L_p^e(Q,S) = \sum_{i=0}^{3} \sum_{j=0}^{3-i} \sum_{k=0}^{3} \sum_{\ell=0}^{\ell-k} \Re_{ijk\ell}^{p,e} \frac{\partial^{(i+j)} Q}{\partial x^{(i)} \partial \theta^{(j)}} \frac{\partial^{(k+\ell)} S}{\partial x^{(k)} \partial \theta^{(\ell)}}$$

$$\begin{aligned} L_p^e(Q,S,T) = {} & \sum_{i=0}^{2} \sum_{j=0}^{2-i} \sum_{k=0}^{2} \sum_{\ell=0}^{2-k} \sum_{m=0}^{2} \sum_{n=0}^{2-m} \Re_{ijk\ell mn}^{p,e} \\ & \times \frac{\partial^{(i+j)} Q}{\partial x^{(i)} \partial \theta^{(j)}} \frac{\partial^{(k+\ell)} S}{\partial x^{(k)} \partial \theta^{(\ell)}} \frac{\partial^{(m+n)} T}{\partial x^{(m)} \partial \theta^{(n)}} \end{aligned} \tag{12}$$

Here $\Re_{ij}^{p,e}$, $\Re_{ijk\ell}^{p,e}$, and $\Re_{ijk\ell mn}^{p,e}$ are coefficients of the elastic parameters (A_{ij} and D_{ij}) and the radius $r(x)$. δ_p^e is the Kronecker delta. This form of differential operators is especially suitable for symbolic programming.

Note that the equations of motion, Eq. (10), which are derived by the consistent variational principle, yield nonlinear terms of imperfection (see the differential operator $L_p^{10}(\hat{w},\hat{w},w)$ in Eq. (11)).

Solution Procedure

Given the solution at time step j (t_j), and using the Newmark method [16] for the time integration, we obtain the partial differential equations for the next time step $j+1$:

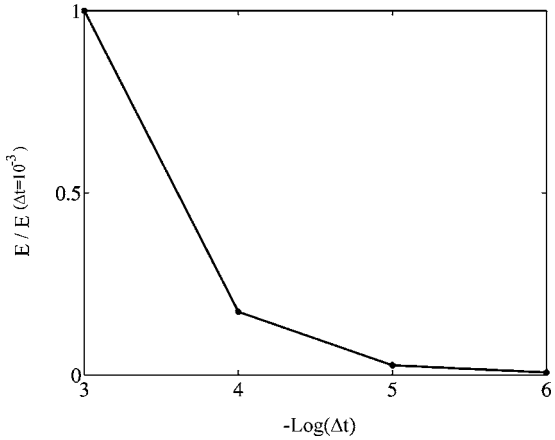


Fig. 2 Convergence curve; normalized total energy versus time step Δt

$$\bar{\phi}_p(\mathbf{z}_{j+1}, \ddot{\mathbf{z}}_j, \dot{\mathbf{z}}_j, \mathbf{z}_j) = 0, \quad p = 1,2,3 \tag{13}$$

This set is first reduced to one of ordinary differential equations by separating the variables and expansion in a truncated Fourier series as:

$$\{u_{j+1}(x,\theta), v_{j+1}(x,\theta), w_{j+1}(x,\theta)\} = \sum_{m=0}^{2N} \{u_{j+1}^m(x), v_{j+1}^m(x), w_{j+1}^m(x)\} g_m(\theta) \tag{14}$$

where $N=N_u$ or N_v or N_w (according to the variable) is the number of terms in the relevant series. The initial geometric imperfection and the external load $\{q_u, q_v, q_w\}$ can also be expanded in the same way:

$$\hat{w}(x,\theta) = \sum_{m=0}^{2N_w} \hat{w}^m(x) g_m(\theta) \tag{15}$$

$$\{q_u(x,\theta), q_v(x,\theta), q_w(x,\theta)\} = \sum_{m=0}^{2N_q} \{q_u^m(x), q_v^m(x), q_w^m(x)\} g_m(\theta) \tag{16}$$

The functions $g_m(\theta)$ are:

$$g_m(\theta) = \begin{cases} \cos(nm\theta) & m = 0,1,\ldots,N \\ \sin[n(m-N)\theta] & m = N+1,\ldots,2N \end{cases} \tag{17}$$

n denoting the characteristic circumferential wave number determined by that corresponding to the lowest total potential energy at a given load level.

The set of the equations of motion consists of three nonlinear differential equations, the first two being of second order and the third of fourth order. These were converted into four equations of second order in a new unknown function:

$$\xi_{j+1} = \frac{\partial^2 w_{j+1}}{\partial x^2} \tag{18}$$

Minimizing the errors due to the truncated Fourier series by applying the Galerkin procedure with cos() and sin() as weighting functions, we obtain the following nonlinear ordinary differential equations:

$$\Phi_r^q(\mathbf{Z}_{j+1}, \ddot{\mathbf{Z}}_j, \dot{\mathbf{Z}}_j, \mathbf{Z}_j; x) = \int_0^{2\pi} \bar{\phi}_r(\mathbf{z}_{j+1}, \ddot{\mathbf{z}}_j, \dot{\mathbf{z}}_j, \mathbf{z}_j) g_q(\theta) d\theta,$$

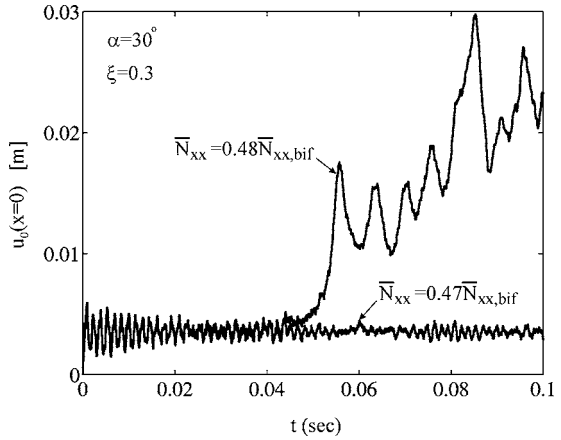


Fig. 3 Variation of average axial displacement at $x=0$ with time

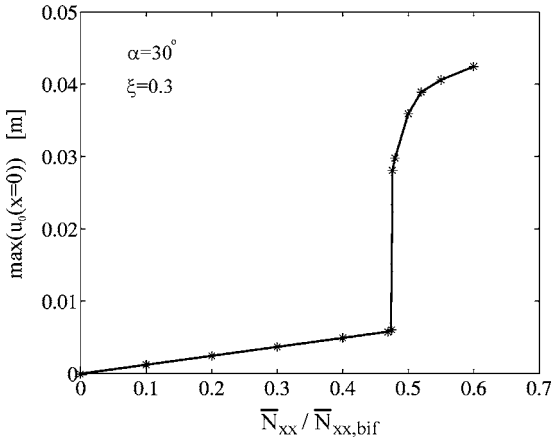


Fig. 5 Variation of maximum average axial displacement with applied load

$$q = 0,1,\ldots,2N, \quad r = 1,2,\ldots,4 \tag{19}$$

where Φ_r^q contains $2N_u+2N_v+4N_w+4$ nonlinear ordinary differential equations, and $\mathbf{Z}_{j+1}$ is the unknown vector function, defined by:

$$\mathbf{Z}_{j+1} = \{u_{j+1}^0,\ldots,u_{j+1}^{2N_u},v_{j+1}^0,\ldots,v_{j+1}^{2N_v},w_{j+1}^0,\ldots,w_{j+1}^{2N_w},\xi_{j+1}^0,\ldots,\xi_{j+1}^{2N_w}\}^T \tag{20}$$

Finally, using the finite-differences scheme, Eqs. (19) yields:

$$\mathbf{G}(\mathbf{u},\lambda) = \mathbf{0} \tag{21}$$

where $\mathbf{G}$ consists of nonlinear algebraic operators, $\mathbf{u}$ is the value of the unknown functions ($\mathbf{Z}_{j+1}$) at each point of the finite-difference scheme, and λ is the load-level parameter. The nonlinear solution of Eq. (21) is obtained by using the Newton-Raphson iteration, with the Riks [17] method resorted to for the post-limit-point solution.

Results and Discussion

The developed NDAICS code (Nonlinear Dynamic Analysis of Imperfect Conical Shells) covered the nonlinear static and dynamic analysis of any isotropic conical shell under arbitrary loading. An isotropic conical shell under step axial compression was considered with the following data: modulus of elasticity $E=7.24\times10^{10}$ N/m^2, Poisson's ratio $\nu=0.3$, mass density $\rho=2700$ kg-mass/m^3, radius at $x=0$ is $R_1=1.27$ m, length $L=2.54$ m, and thickness $h=0.0127$ m; boundary conditions: simply-supported, SS3 ($N_{xx}=-\bar{N}_{xx}, v=w=M_{xx}=0$) at $x=0$ and SS4 ($u=v=w=M_{xx}=0$) at $x=L$. The imperfection shape was taken as the buckling mode shape and plotted in Fig. 1. The bifurcation point, considering the effect of prebuckling nonlinearity (see Ref. [18]), was found as $\bar{N}_{xx,bif}=0.3603\times10^7$ N/m.

First, the criterion for the solution convergence with respect to the time step was examined and it was found that the one of the vanishing total energy ($E=T+\Pi=0$, the shell at rest at $t=0$) is the most representative. An example of this convergence is given in Fig. 2 for a conical shell with cone semi-vertex angle $\alpha=30$ deg and imperfection amplitude $\xi=\hat{w}/h=0.3$. It is obtained with a very small time interval due to the high frequency of the characteristic behavior as shown in Fig. 3.

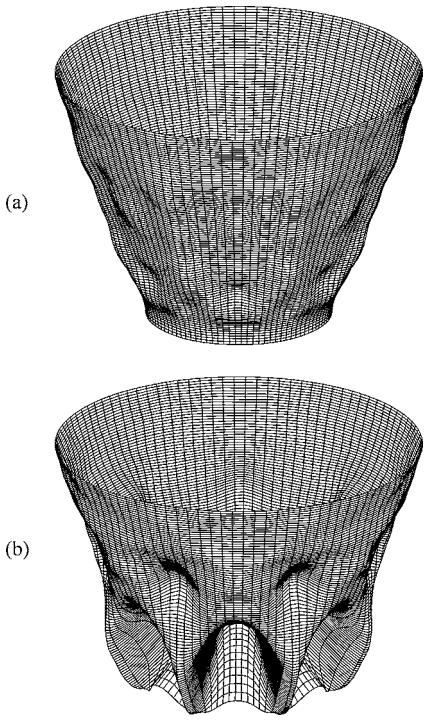


Fig. 4 Dynamic displacement pattern of conical shell with $\alpha=30$ deg under $\bar{N}_{xx}=0.48\bar{N}_{xx,bif}$; (*a*) at time 0.025 s; (*b*) at time 0.075 s

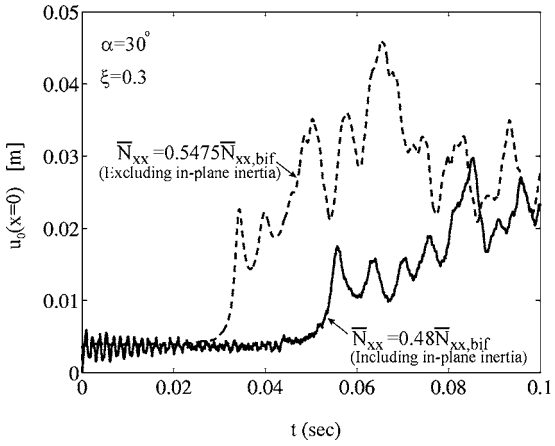


Fig. 6 Effect of in-plane inertia on time history

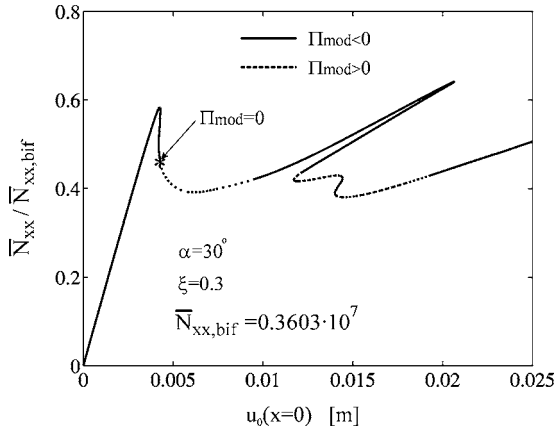


Fig. 7 Axial load versus average axial displacement

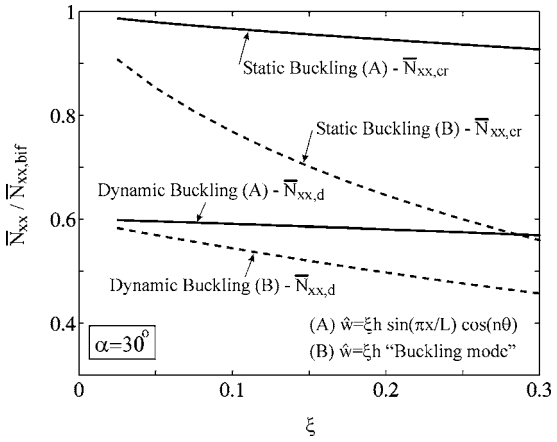


Fig. 10 Effect of the imperfection shape on static and dynamic buckling load

The above figure (Fig. 3) shows the time history of the average axial displacement at $x=0$ under an axial load of $0.47\bar{N}_{xx,\mathrm{bif}}$ and $0.48\bar{N}_{xx,\mathrm{bif}}$. It is seen that up to the load level $\bar{N}_{xx}=0.47\bar{N}_{xx,\mathrm{bif}}$ the oscillation is about the nonlinear static pre-buckling state, while for higher levels, $\bar{N}_{xx}>0.48\bar{N}_{xx,\mathrm{bif}}$, the shell snaps through to a new stable point in the post-buckling regime and oscillates with a large amplitude about the new state but is still bounded. It seems that the dynamic buckling load falls between $\bar{N}_{xx}=0.47\bar{N}_{xx,\mathrm{bif}}$ and $\bar{N}_{xx}=0.48\bar{N}_{xx,\mathrm{bif}}$. The deformation shape for $\bar{N}_{xx}=0.48\bar{N}_{xx,\mathrm{bif}}$ is presented in Fig. 4 at time $t=0.025$ s (oscillation about the nonlinear pre-buckling state) and $t=0.075$ s (oscillation about the nonlinear post-buckling state). The near-axisymmetric mode in the solution dominates during the pre-buckling (unbuckled motion) (Fig. 4(*a*)), while at the buckling level (Fig. 4(*b*)), the sharp asymmetric mode becomes dominant. Figure 5 illustrates the Budiansky-Roth criterion, with the maximum average axial displacement (at $x=0$) plotted versus the applied load, and the sharp jump at $\bar{N}_{xx}=0.476\bar{N}_{xx,\mathrm{bif}}$ clearly seen. The effect of the in-plane inertia on the dynamic behavior of the shell is shown in Fig. 6 (the dashed line

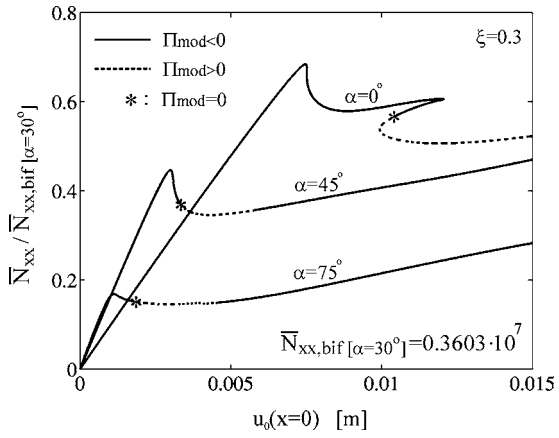


Fig. 8 Axial load versus average axial displacement for different semi-vertex angles

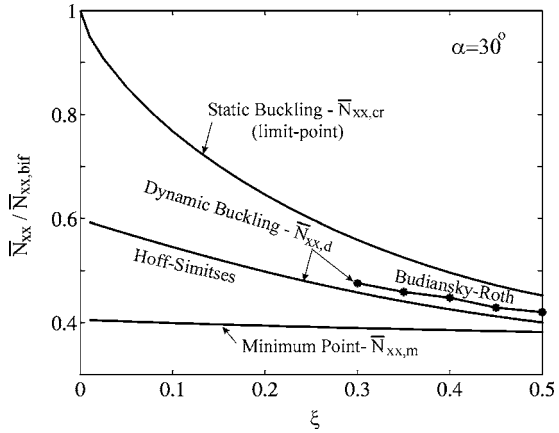


Fig. 9 Imperfection sensitivity of conical shell under static and dynamic step load

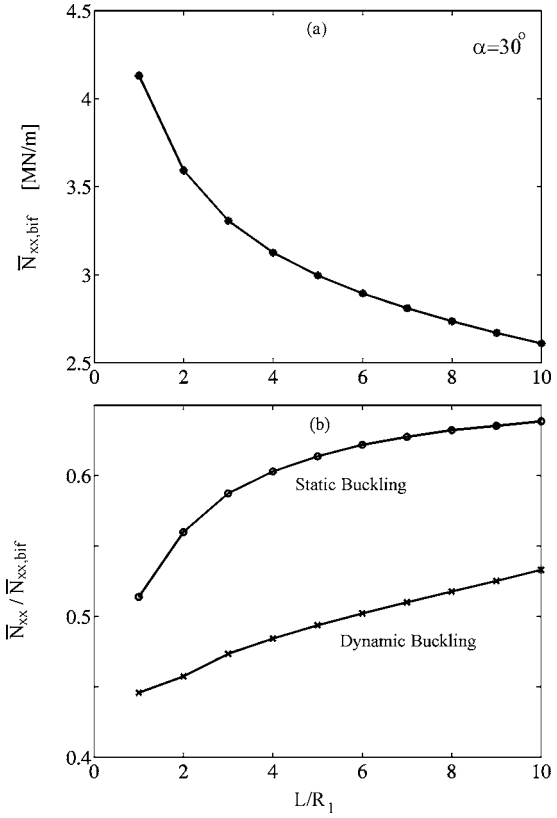


Fig. 11 Effect of aspect ratio L/R_1 on the (*a*) bifurcation point and (*b*) static and dynamic buckling load

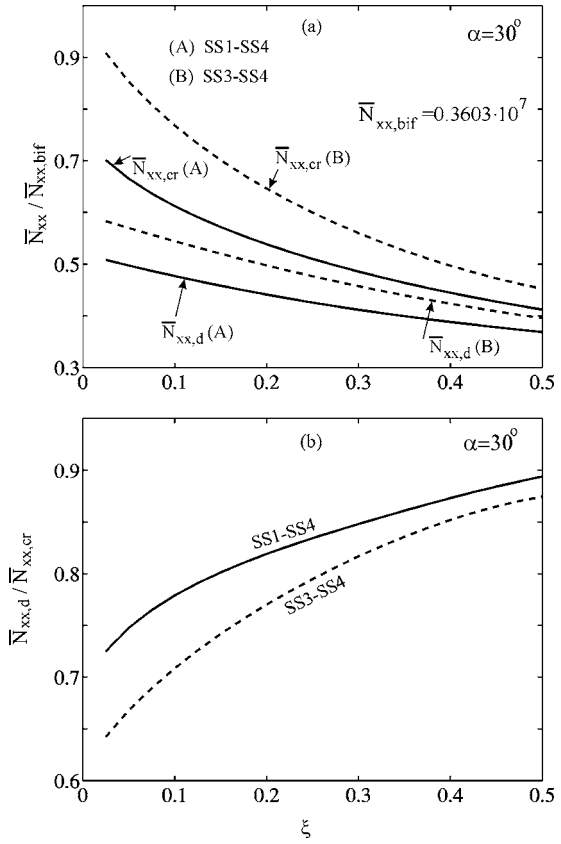


Fig. 12 Effect of in-plane boundary conditions on static and dynamic buckling load

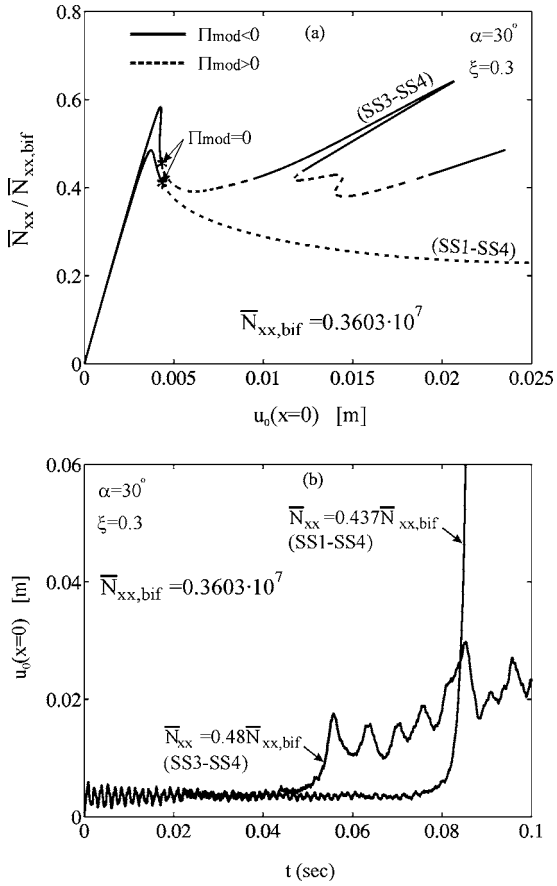


Fig. 13 Effect of in-plane boundary conditions on the (*a*) nonlinear static behavior and (*b*) dynamic behavior

refers to neglected in-plane inertia and the solid line to its retention). The dynamic buckling load in the case of neglected in-plane inertia is the higher of the two, as expected. Hence, formulations that neglect the in-plane inertia, like the WF (where F is the Airy stress function and W is the radial displacement, see Ref. [19]) may lead to inaccurate results.

Figure 7 illustrates the Hoff-Simitses criterion whereby the modified total potential energy is zero at $\bar{N}_{xx}=0.459\bar{N}_{xx,\text{bif}}$, which is slightly lower than the Budiansky-Roth result ($\bar{N}_{xx}=0.476\bar{N}_{xx,\text{bif}}$). The entire nonlinear static pattern in this figure was obtained with the aid of the code of Jabareen and Sheinman [2], with the solid lines representing the equilibrium path with $\Pi_{\text{mod}}<0$ (bounded motion) and the dashed lines that with $\Pi_{\text{mod}}>0$ (unbounded) see Eqs. (3) and (4). The same criterion is shown for other configurations $\alpha=0$ deg, $\alpha=45$ deg, and $\alpha=75$ deg in Fig. 8.

The dynamic sensitivity to imperfection according to the different criteria is summarized in Fig. 9. It is seen that the results are quite close (up to 5% difference). The curves of the static buckling load (limit-point) and the minimum point (representing the possible lowest load level at the equilibrium path) are also plotted in this figure and serve as the upper and lower bounds for the dynamic buckling load.

The effect of the imperfection shape is given in Fig. 10. The solid line refers to the results for the symmetric imperfection shape taken from Simitses [20], $\hat{w}=\xi h\sin(\pi x/L)\cos(n\theta)$, and the dashed line to those for the shape from the buckling mode. The imperfection of the mode shape caused a larger reduction in the static buckling (limit-point) and dynamic buckling.

Figure 11 shows the effect of the length-to-radius ratio (L/R_1). The bifurcation point is plotted in Fig. 11(*a*), and the normalized static and dynamic buckling load with respect to the bifurcation level are given in Fig. 11(*b*). It is shown that the longer the shell, the lower the sensitivity to imperfection for both the static and dynamic cases.

In view of the pronounced effect of the in-plane boundary conditions, the static and dynamic buckling loads were studied and summarized in Fig. 12 for the simply-supported boundary conditions: SS1 ($N_{xx}=-\bar{N}_{xx}, N_{x\theta}=w=M_{xx}=0$) or SS3 ($N_{xx}=-\bar{N}_{xx}, v=w=M_{xx}=0$) at $x=0$, and SS4 ($u=v=w=M_{xx}=0$) at $x=L=2.54$ m, and the imperfection taken as the buckling mode. The solid lines represent the case of SS1-SS4 boundary conditions, and the dashed lines that of SS3-SS4. The normalized static (limit-point) and dynamic buckling loads are given in Fig. 12(*a*), and the sensitivity—defined by the ratio of the dynamic and static buckling loads—is plotted in Fig. 12(*b*). It is shown that SS3-SS4 yields a higher static and dynamic buckling level, with rather more sensitivity to imperfection.

The nonlinear behavior of the SS1-SS4 case is completely different from that of SS3-SS4 in terms of its post-buckling hardening (see Fig. 13(*a*)). In the SS1-SS4 case, from the dynamic buckling level on, the total modified potential energy is positive ($\Pi_{\text{mod}}>0$), while for SS3-SS4 it has some regions of $\Pi_{\text{mod}}<0$ even after the dynamic buckling level (switching from the dashed to the solid line and back, as shown in Fig. 13(*a*)). In other words, SS3-SS4 is characterized by some hardening regions, while SS1-SS4 has no hardening at all. That of course affects the real time history behavior in which beyond the dynamic buckling level SS3-SS4 oscillates around a new equilibrium state (see the curve of $\bar{N}_{xx}=0.48\bar{N}_{xx,\text{bif}}$ in Fig. 13(*b*)), while SS1-SS4 shows no oscillations at all after the buckling, and the motion grows unbounded.

The effect of the semi-vertex angle on the static and dynamic buckling loads is plotted in Fig. 14. The pattern curves reflect the

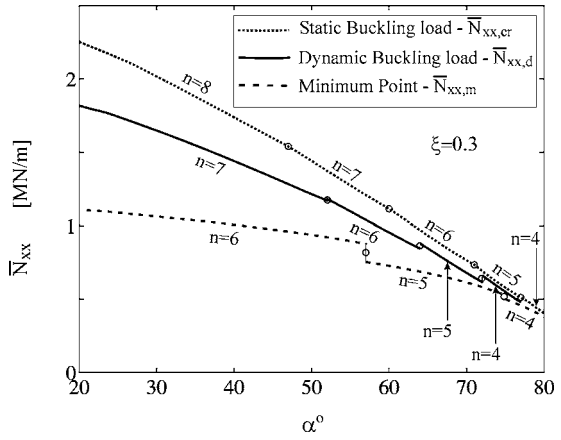


Fig. 14 Effect of vertex half-angle on static (limit-point) and dynamic buckling load

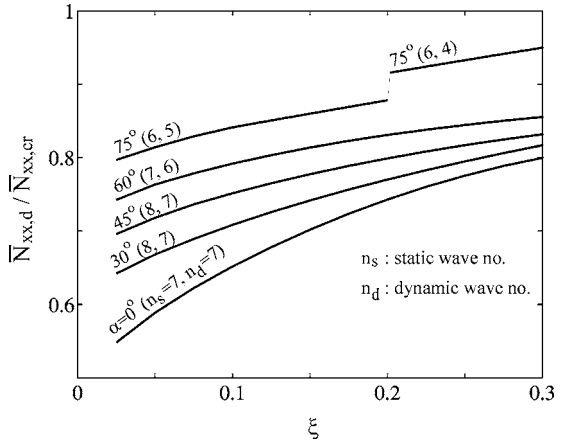


Fig. 15 Dynamic sensitivity to imperfection for different values of semi-vertex angles

change of the circumferential characteristic wave number, which decreases as the angle increases. It is also seen that as the angle increases the two loads (the static and dynamic buckling curves) draw closer together and for an annular plate they coincide—in other words, as the angle increases the sensitivity to imperfection decreases. Finally, the dynamic sensitivity to imperfection is plotted in Fig. 15 for different values of the semi-vertex angle. It is seen that as the angle increases, the ratio of dynamic to static buckling load increases—meaning less dynamic sensitivity to imperfection.

Concluding Remarks

A symbolic procedure for creating a numerical code is presented for the nonlinear dynamic behavior of conical shells. It is used here to compare different dynamic stability criteria and to examine the effect of the imperfection on the nonlinear dynamic stability and sensitivity behavior for different aspect ratios. The following conclusions can be drawn from the results:

- The dynamic buckling load obtained by the Hoff-Simitses criterion (via only static analysis) is acceptable and slightly lower than that obtained by the Budiansky-Roth criterion (via complete dynamic analysis).
- The Hoff-Simitses criterion is quite efficient, thus obviating any need for dynamic analysis.
- The static and dynamic sensitivity behavior is imperfection-amplitude dependent: the larger the amplitude, the higher the sensitivity.
- The imperfection shape as the buckling mode leads to rigorous sensitivity to imperfection compared to other shapes.
- The larger the length-to-radius ratio, the smaller the sensitivity to imperfection.
- The cone semi-vertex angle plays an important role in the sensitivity to imperfection, in both the static and the dynamic cases: the larger the angle, the lower the sensitivity.

Acknowledgment

This study was partially supported by the Fund for Promotion of Research at the Technion. The authors are indebted to Ing. E. Goldberg for editorial assistance.

References

[1] Goldfeld, Y., Sheinman, I., and Baruch, M., 2003, "Imperfection Sensitivity of Conical Shells," AIAA J., **41**(3), pp. 517–524.
[2] Jabareen, M., and Sheinman, I., 2006, "Post-buckling Analysis of Geometrically Imperfect Conical Shells," submitted for publication.
[3] Rossettos, J. N., and Parisse, R. F., 1969, "The Dynamic Response of Cylindrical and Conical Panels," Trans. ASME, J. Appl. Mech., **36**, pp. 271–276.
[4] Srinivasan, R. S., and Krishnan, P. A., 1989, "Dynamic Response of Layered Conical Shell Panel Using Integral Equations Technique," Comput. Struct., **31**(6), pp. 897–905.
[5] Srinivasan, R. S., and Krishnan, P. A., 1989, "Dynamic Analysis of Stiffened Conical Shell Panels," Comput. Struct., **33**(3), pp. 831–837.
[6] Ganapathi, M., Patel, B. P., Sambandam, C. T., and Touratier, M., 1999, "Dynamic Instability Analysis of Circular Conical Shells," Compos. Struct., **46**, pp. 59–64.
[7] Shiau, A. C., Roth, R. S., and Soong, T. T., 1974, "Dynamic Buckling of Conical Shells With Imperfections," AIAA J., **12**(6), pp. 755–760.
[8] Dumir, P. C., Dube, G. P., and Mullick, A., (2003), "Axisymmetric Static and Dynamic Buckling of Laminated Thick Truncated Conical Cap," Int. J. Non-Linear Mech., **37**, pp. 903–910.
[9] Budiansky, B., and Roth, R. S., 1963, "Axisymmetric Dynamic Buckling of Clamped Shallow Spherical Shells," in "Collected Papers on Instability of Shell Structures," NASA TN, D-1510, pp. 597–606.
[10] Hoff, N. J., and Bruce, V. C., 1954, "Dynamic Analysis of the Buckling of Laterally Loaded Flat Arches," J. Math. Phys., **32**, pp. 276–288.
[11] Simitses, G. J., 1967, "Axisymmetric Dynamic Snap-Through Buckling of Shallow Spherical Caps," AIAA J., **5**, pp. 1019–1021.
[12] Simitses, G. J., and Sheinman, I., 1982, "Dynamic Buckling of Shell Structures: Concepts and Applications," Acta Astronaut., **9**(3), pp. 178–182.
[13] Simitses, G. J., 1990, *Dynamic Stability of Suddenly Loaded Structures*, Springer Veralg, New York.
[14] Donnell, L. H., 1933, "Stability of Thin-Walled Tubes Under Torsion," NACA TR-479, National Advisory Committee for Aeronautics, Washington, DC.
[15] Sheinman, I., and Goldfeld, Y., 2001, "Buckling of Laminated Cylindrical Shell in Terms of Different Theories and Formulations," AIAA J., **39**(9), pp. 1773–1781.
[16] Newmark, N. M., 1959, "A Method of Computation for Structural Dynamics," J. Engrg. Mech. Div., **85**(EM3), pp. 67–94.
[17] Riks, E. 1979, "An Incremental Approach to the Solution of Snapping and Buckling Problems," Int. J. Solids Struct., **15**(7), pp. 529–551.
[18] Jabareen, M., and Sheinman, I. 2005, "Effect of the Nonlinear Pre-buckling State on the Bifurcation Point of Conical Shells," Int. J. Solids Struct., **43**(7–8), pp. 2146–2159.
[19] Shaw, D., Shen, Y. L., and Tsai, P., 1993, "Dynamic Buckling of an Imperfect Composite Circular Cylindrical Shell," Comput. Struct., **48**(3), pp. 467–472.
[20] Simitses, G. J. 1986, "Buckling and Post-Buckling of Imperfect Cylindrical Shells: A Review." Appl. Mech. Rev., **39**(10), pp. 1517–1524.

Yueming Liang
Alexander V. Spuskanyuk
Shane E. Flores

Mechanical Engineering Department,
Materials Department,
University of California,
Santa Barbara, CA 93106

David R. Hayhurst
School of Mechanical,
Aerospace and Civil Engineering,
The University of Manchester,
Manchester, M60 1QD, UK

John W. Hutchinson
Division of Engineering
and Applied Sciences,
Harvard University,
Pierce Hall, Cambridge, MA 02138

Robert M. McMeeking
Anthony G. Evans

Mechanical Engineering Department,
Materials Department,
University of California,
Santa Barbara, CA 93106

The Response of Metallic Sandwich Panels to Water Blast

Metallic sandwich panels subject to underwater blast respond in a manner dependent on the relative time scales for core crushing and water cavitation. This article examines the response at impulses representative of an (especially relevant) domain: wherein the water cavitates before the core crushes. Three core topologies (square honeycomb, I-core, and corrugated) have been used to address fundamental issues affecting panel design. Their ranking is based on three performance metrics: the back-face deflection, the tearing susceptibility of the faces, and the loads transmitted to the supports. The results are interpreted by comparing with analytic solutions based on a three-stage response model. In stage I, the wet face acquires its maximum velocity with some water attached. In stage II, the core crushes and all of the constituents (wet and dry face and core) converge onto a common velocity. In stage III, the panel deflects and deforms, dissipating its kinetic energy by plastic bending, stretching, shearing, and indentation. The results provide insight about three aspects of the response. (i) Two inherently different regimes have been elucidated, designated strong (STC) and soft (SOC), differentiated by a stage II/III time scale parameter. The best overall performance has been found for soft-core designs. (ii) The foregoing analytic models are found to underestimate the kinetic energy and, consequently, exaggerate the performance benefits. The discrepancy has been resolved by a more complete model for the fluid/structure interaction. (iii) The kinetic energy acquired at the end of the second stage accounts fully for the plastic dissipation occurring in the third stage, indicating that the additional momentum acquired after the end of the second stage does not affect panel performance. [DOI: 10.1115/1.2178837]

1 Introduction

The response of metallic sandwich panels to the impulse caused by underwater blast has been the subject of several recent assessments [1–7]. The scenario of interest is depicted in Fig. 1. An impulse from the water impinges on a panel rigidly supported around its perimeter. The blast wave causes the core to compress and the panel to deflect. It is desirable to design panels that (for equivalent weight per area) beneficially affect the following three performance metrics, in order of importance:

A. The incidence and extent of dynamic tearing of the front (or wet) face.
B. The center deflection of the back (or dry) face.
C. The load imparted to the supports.

The benchmark is a solid plate of the same material with identical weight per area. The intent is to probe the influence of topology on all three performance metrics and thereby provide insights into the design of optimal panels. In the present paper, numerical simulation is used to ascertain responses that can be expressed in a form amenable to comparison with available analytic solutions. Results for a plane impulse provide the fundamental perspective [2–7].

The analysis originates with a temporally distinct three-stage model for air blast (Fig. 1) [2–5]. In stage I, the impulse imparts a momentum to the front face which dictates its kinetic energy. In stage II, the front face decelerates as the core and back face accelerate, until a common velocity is attained. At the end of this stage, the momentum and kinetic energy are the same as those experienced by a solid plate having the same mass per area. Finally, in stage III, the kinetic energy is dissipated by plastic bending, stretching, and indentation of the panel.

While the response to water blast may not be temporally separable in such explicit manner (and for certain panels may be inseparable even for air blast), the three stages remain a useful concept. The extension to water blast has been pursued through the incorporation of the fluid/structure interaction (FSI) [4]. This assessment has identified four different domains differentiated by their cavitation and core crushing characteristics. To set the objectives of the present study, the key features of this FSI analysis are described in the following section.

2 Background Analysis

2.1 Initial Response. Consider a panel with total mass per unit area,

$$m_{\text{total}} = m_f + m_b + m_c \tag{1}$$

with m_f the mass per unit area of the front face, m_b that for the back face, and m_c the mass per unit area of the core. The impinging wave has the form [3–5]:

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 9, 2005; final manuscript received December 2, 2005. Review conducted by H. D. Espinosa. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

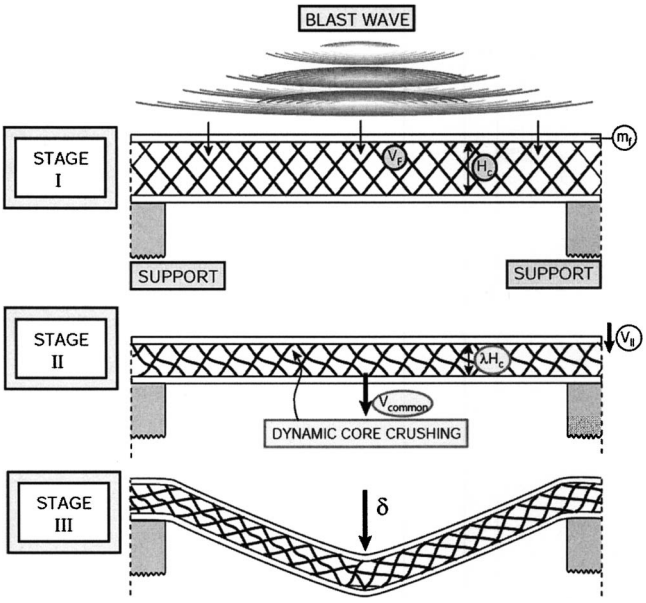


Fig. 1 A schematic showing the three temporally distinct stages that accompany a panel subject to air blast [3]

$$p_I = p_0 e^{\xi},$$

$$\xi = \frac{x}{c_w t_0} - \frac{t}{t_0},$$

with t_0 the characteristic decay time for the pulse and t the time after the blast first arrives at the fluid/structure interface (at $x=0$). Thereafter, the pressure in the water is (for $\zeta = x/c_w t_0 + t/t_0 > 0$) [4,5]:

$$p(x,t) = p_0 \left[e^{\xi} - \left(\frac{2\beta}{1-\beta} + \frac{\sigma_{YD}^c}{p_0} \right) e^{-\beta\zeta} + \frac{1+\beta}{1-\beta} e^{-\zeta} + \frac{\sigma_{YD}^c}{p_0} \right], \tag{2a}$$

and (for $\zeta < 0$)

$$p(x,t) = p_0 e^{\xi} \tag{2b}$$

with the fluid/structure interaction parameter, $\beta = \rho_w c_w t_0 / m_f$. The corresponding velocity distribution is (for $\zeta > 0$):

$$v(x,t) = \frac{p_0}{\rho_w c_w} \left[e^{\xi} + \left(\frac{2\beta}{1-\beta} + \frac{\sigma_{YD}^c}{p_0} \right) e^{-\beta\zeta} - \frac{1+\beta}{1-\beta} e^{-\zeta} - \frac{\sigma_{YD}^c}{p_0} \right] \tag{3a}$$

and (for $\zeta < 0$)

$$v(x,t) = \frac{p_0}{\rho_w c_w} e^{\xi}. \tag{3b}$$

The water begins to cavitate at location $x=x_c$ after time $t=t_c$, when the pressure in the water first satisfies:

$$p(x_c,t_c) = 0 \tag{4a}$$

and

$$\frac{\partial p(x_c,t_c)}{\partial x} = 0. \tag{4b}$$

This event coincides with the end of stage I. The fluid pressure and velocity profiles at the end of this stage are thus obtained from Eqs. (2*a*) and (3*a*) as:

$$\begin{aligned} p'(x) &= p(x,t_c) \\ v'(x) &= v(x,t_c) \end{aligned} \tag{5}$$

This occurs at time:

$$t_c \approx t_0 (\ln \beta)/(\beta - 1).$$

The pressure and velocity characteristics at this time for $\beta = 3.125$ and $\sigma_{YD}^c/p_0 = 0.18$ are shown in Fig. 2(*a*), together with finite element calculations elaborated later. Subsequently, a cavitation front moves through the water, away from the panel, exemplified by the pressure and velocity distributions at $t=2t_c$ plotted in Fig. 2(*b*).

At this stage, the momentum of the front face and attached water layer (added mass) ($x_c < x < 0$) is:

$$M_F = m_f v_{\text{face}} + \int_{x_c}^{0} \rho_w v'(x) dx, \tag{6}$$

while that acquired by the core plus back face is [5]:

$$M_B = \sigma_{YD}^c t_c. \tag{7}$$

In the preceding formulas, σ_{YD}^c is the "dynamic strength" of the core and is assumed constant in time. It is the stress induced in the core, immediately adjacent to the front face, at the onset of cavitation. For a material with linear hardening, tangent modulus, E_T (typical for stainless steels), the following approximate form has been proposed [8]:

$$\frac{\sigma_{YD}^c}{\sigma_Y \bar{\rho}} \equiv \lambda_D \approx 1 + \sqrt{\frac{E_T}{E}} \left(\frac{v_{\text{face}}}{c_{el} \varepsilon_Y} - 1 \right). \tag{8}$$

Here $c_{el} = \sqrt{E/\rho}$ is the relevant elastic wave speed in the constituent material, σ_Y is its yield strength, and ε_Y is its yield strain (at the relevant strain-rate). To determine σ_{YD}^c from (8), the front face velocity, v_{face}, must be determined from (3*a*) at $x=0$, $t=t_c$, leading to an implicit expression. As an alternative, the velocity can be estimated from (3*a*) with the core yield strength set to zero; the result is almost the same. It will be shown below that (8) provides a reasonably accurate measure of σ_{YD}^c for each of the cores examined in this investigation.

The pressure pulses caused by underwater explosions are more complex. The growth and collapse of bubbles caused by the expansion and compression of the explosion gases, as well as the interaction of the buoyant bubble and its acoustic signals with the sea bottom, the water surface, and the structure [9–12], lead to pressure waves having complicated forms. Nevertheless, the initial pressure rise and its decay at locations below the surface are well characterized by the preceding forms. Most of the differences between the present model and the reality of an underwater explosion occur after the initial pressure rise has decayed to moderate levels [9–12]. The secondary pressure spikes that arrive later do not cause significant damage to a well-designed structure that has survived the initial impulse, as described later. So we focus on the initial pressure wave and its effect on the structures of interest. It is also true that the simple exponential decay after the sharp spike of pressure is an idealization, with some oscillation of the pressure occurring superposed on the exponential form [9]. However, these pressure oscillations are minor and do not significantly alter the effect of the impulse on the deforming structure.

2.2 Impulsive Domains. The response of the panel in stage II depends on the magnitudes of the impulse, the FSI parameter, β, and the mass [4]. Key aspects have been elucidated by envisaging homogenized cores at appropriate relative densities [4]. Four domains emerge, illustrated using two additional parameters: a nondimensional impulse [4],

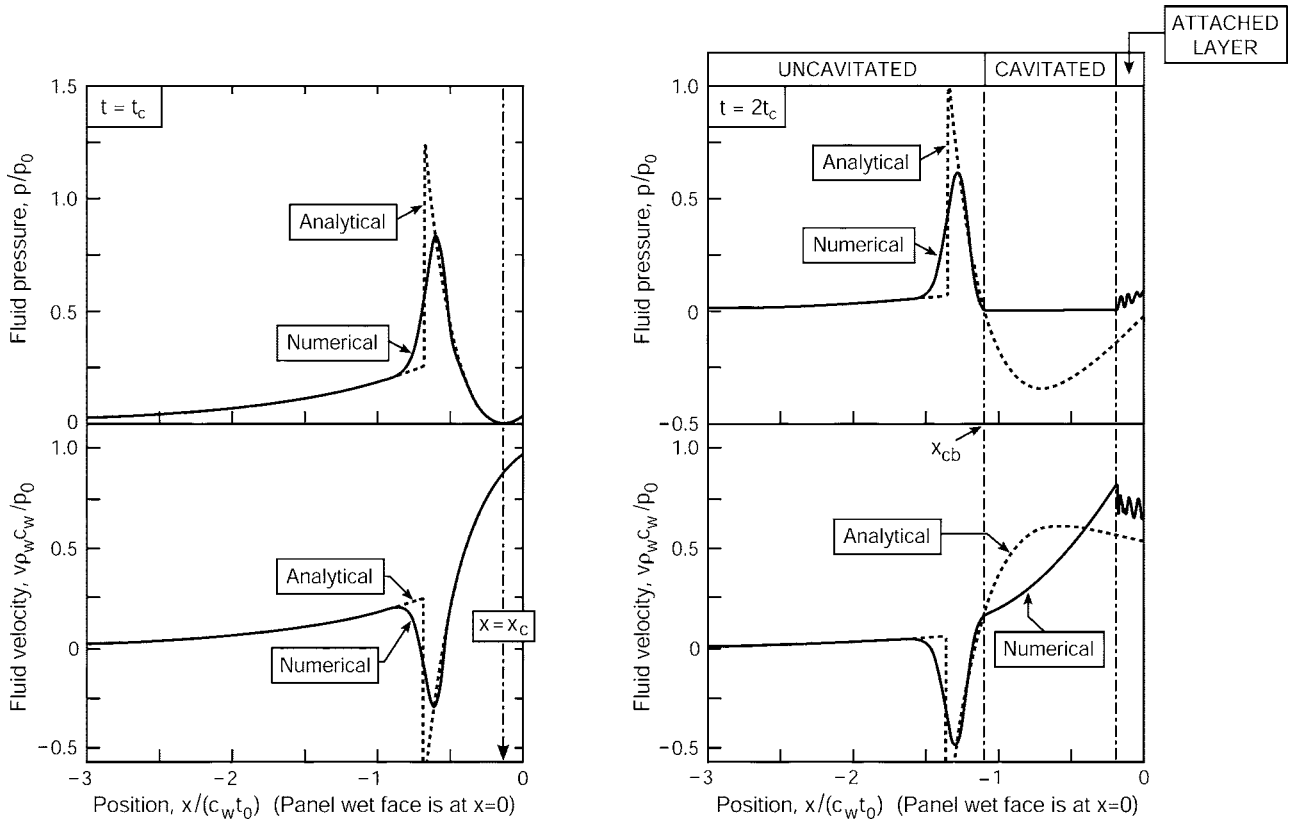


Fig. 2 The trends in pressure and velocity in the water at times t_c and $2t_c$ (with t_c designating the instant when cavitation commences). The plots compare the analytic solution with a calculation conducted using ABAQUS/Explicit. After t_c, cavitation fronts propagate through the water (towards and away from the panel) leaving a zone of cavitated water in their wakes. Note that, at the cavitation front, the water has positive velocity in the direction of motion of the panel. The equivalent thickness of the sandwich panel is 20 mm. The core has a relative density of 0.03 and yield strength of $\sigma^c_{YD}/p_0=0.18$. The thickness of the front face is 6 mm so that $\beta=3.125$ and that of the back face is 8 mm.

$$\bar{I}=\frac{2I_o t_o}{m_f H_c} \tag{9a}$$

and a nondimensional mass,

$$\bar{m}=\frac{m_c}{m_f}, \tag{9b}$$

where H_c is the core height. The domains are illustrated on Fig. 3 [4] for levels of impulse and panel dimensions representative of those to be explored in the present study ($\bar{I}\approx 0.5, \bar{m}\approx 1$). They comprise a high strength domain (IV) in which the core is completely resistant to crushing and a low strength, small β, domain (I) in which the core completely crushes before the water cavitates. Two intermediate domains (II and III), in which the cores partially crush, are especially relevant. In domain II, *cavitation occurs in the water* (at location $x_c<0$) before core crushing is complete. Domain III also involves partial crushing, but differs in the sense that cavitation occurs at the front face ($x_c=0$), not in the water. It arises when the cores are stronger. Contours of transmitted impulse (Fig. 3) indicate how the domains affect the response. Because of its importance to ongoing investigations of blast resistant panels [1–7], *the present study has been designed to probe domain II for representative core topologies*. That is, all of the impulse levels and core topologies result in domain II responses. Subsequent investigations will examine other domains as well as transitions between domains.

2.3 Analytic Formulas for Domain II. Analysis conducted for impulses and cores representative of domain II [5] have provided formulas that benchmark the ensuing numerical results. Formulas are available for a design with front and back faces having the same mass [5] expressing the transmitted impulse, I_T, as well as the momentum acquired by the front face, M_F, both inclusive of the mass of attached water between the initial cavitation plane and the structure, and the momentum induced in the core plus back face, M_B. This attached water mass is [5]:

$$m_w=0.71 m_f \beta \frac{\sigma^c_{YD}}{p_0}. \tag{10}$$

such that

$$\frac{I_T}{I_0}\equiv F_T=2f+1.27(1-f)\frac{\sigma^c_{YD}}{p_0} \tag{11a}$$

$$\frac{M_B}{I_0}\equiv F_B=3.64 f\frac{\sigma^c_{YD}}{p_0} \tag{11b}$$

$$M_F\equiv I_0 F_F=I_T-M_B \tag{11c}$$

with $f=\beta^{\beta/(1-\beta)}$.

The kinetic energies acquired by the panel can be determined from these momenta. Inclusive of the mass of the attached water, the KE at the end of stage I, before the core begins to crush, is [5];

$$\mathrm{KE}_I=\frac{M_F^2}{2(m_f+m_w)}+\frac{M_B^2}{2(m_b+m_c)} \tag{12a}$$

and, after stage II, when core crushing is complete [5]:

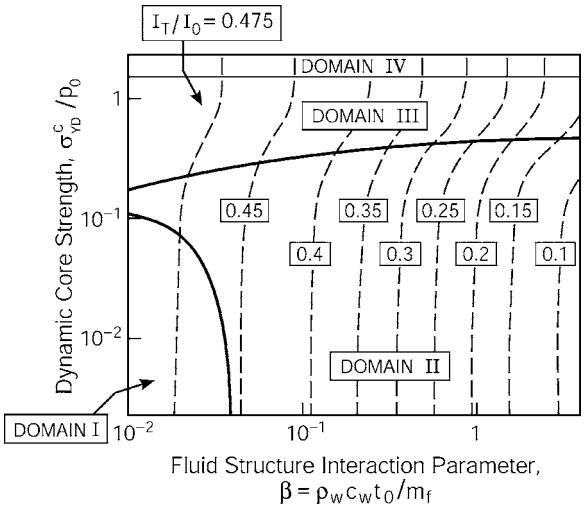


Fig. 3 Fluid structure interaction map ($\bar{I}$=0.5, $\bar{m}$=1.0) with axes of core dynamic strength and the Taylor fluid structure interaction parameter β. The four impulse domains are marked on the map. Contours of the impulse transmitted into the sandwich plate at first cavitation are also included [4].

$$\mathrm{KE}_{II} = \frac{I_T^2}{2(m_f + m_b + m_c + m_w)}. \tag{12b}$$

In some of the following numerical assessments, the kinetic energies in the structure only (that is, excluding the attached water) are also determined: these are designated KE_I^* and KE_{II}^*.

The times associated with these stages are [5]:

$$t_I = t_c \approx t_0(\ln\beta)/(\beta - 1),$$

$$t_{II} \approx \frac{I_T}{2\sigma_{YD}^c}, \tag{13}$$

$$t_{III} \approx L\sqrt{\rho/\sigma_Y}.$$

At the end of stage II, when temporally distinct from stage III, all constituents attain a common velocity,

$$v_{\mathrm{common}} = \sqrt{\frac{2\mathrm{KE}_{II}}{m_{\mathrm{total}} + m_w}}. \tag{14}$$

During stage II, the crushing strain ε_c of the core is dictated by its ability to absorb the kinetic energy differential, $\mathrm{KE}_I - \mathrm{KE}_{II}$, through plastic dissipation in the core:

$$W_c \approx \sigma_{YD}^c \varepsilon_c H_c. \tag{15}$$

Equating W_c to $\mathrm{KE}_I - \mathrm{KE}_{II}$ leads to an expression for the crushing strain [5].

The back face deflection is dictated by the ability of the sandwich to absorb KE_{II}, during stage III. The dissipation involves plastic bending, stretching, shearing, and indentation, subject to the prior core crushing in stage II. When bending and stretching predominate (no shear resistance in the core and no indentation), the dissipation for a panel with both faces having equal thickness rigidly supported at the ends is given by [5]:

$$W_{pl}^{\mathrm{total}} = \frac{2}{3}\sigma_Y h_f(2 + \lambda_s m_c/m_f)\left(\frac{\delta_b}{L}\right)^2 + 4\sigma_Y h_f \frac{H_c(1-\varepsilon_c)}{L}\frac{\delta_b}{L} \tag{16}$$

where $\lambda_s = \sigma_s/\sigma_Y\bar{\rho}$ is a measure of the stretch resistance of the core, with σ_s the stretch strength. The results for other support conditions will be discussed elsewhere [13].

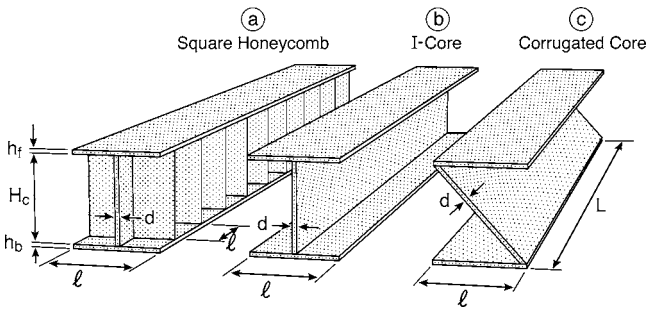


Fig. 4 The geometries of the three core topologies used in the analysis

2.4 Strong and Soft Cores. The possibility that two regimes exist has already been postulated [4]. These regimes emerge vividly in the present study. Their relative incidence is anticipated by the relative time scales for stages II and III expressed by the ratio [14]:

$$\Pi \equiv \frac{t_{II}}{t_{III}} \approx \frac{2I_T}{L\sigma_{YD}^c}\sqrt{\frac{\sigma_Y}{\rho}}. \tag{17}$$

The differing responses are differentiated by a critical value, Π_{th}. When $\Pi < \Pi_{th}$, stages II and III are temporally distinct, enabling the panels to attain a common velocity at the end of stage II, whereupon the response can be analyzed with modifications to the preceding analytical formulas. Panels that respond in this manner are designated *strong core designs* (STC), because high core strength contributes to the sign of the Π inequality. Note that the distance between the supports also affects the transition, so that the STC designation should not be construed to be solely governed by the dynamic strength of the core.

When $\Pi > \Pi_{th}$, stages II and III merge, causing a change in regime. The alternative response is conceptually closer to the buffer plate/crushable core concept [15], with no common velocity. Panels having this response are designated *soft core designs* (SOC), with the caviat that, again, the support length is also involved. It will be demonstrated that, for all three performance-governing metrics (A, B and C, above), SOC designs are preferable. Estimates of Π_{th} will be provided.

3 Scope of the Calculations

Three different core topologies have been selected, based on their geometric versatility and the range in their dynamic strength (Fig. 4). When the core height H_c is significant relative to the span of the panel, it can be considered to represent a double hull system in naval architecture. However, we do not consider our concepts to be restricted to either single or double hull systems, but applicable as appropriate to both possibilities.

i. A high strength orthotropic core based on the square honeycomb (Fig. 4(*a*)). In the double hull setting, this architecture is known as orthogonally stiffened.
ii. An I-core with axial characteristics comparable to the honeycomb, but entirely different transverse properties, rendering it amenable to a comprehensive parameter study (Fig. 4(*b*)). In the double hull case, this design is known as unidirectionally-stiffened.
iii. A corrugated core amenable to a wide sensitivity assessment, upon varying the included angle and the relative density (Fig. 4(*c*)). This design is reminiscent of the Navtruss. However, our architecture differs both in detail and in parametric range.

All calculations are performed without initial imperfections. The role of imperfections, which can be substantial, will be examined in a separate study [13].

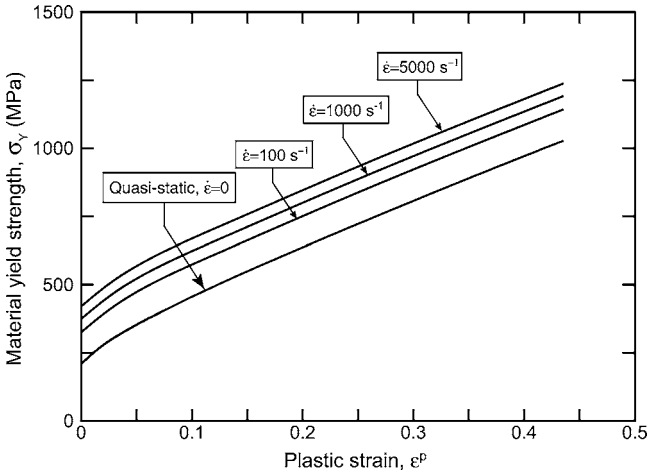


Fig. 5 Dynamic stress/strain curves for 304 stainless steel used in the simulations [8]

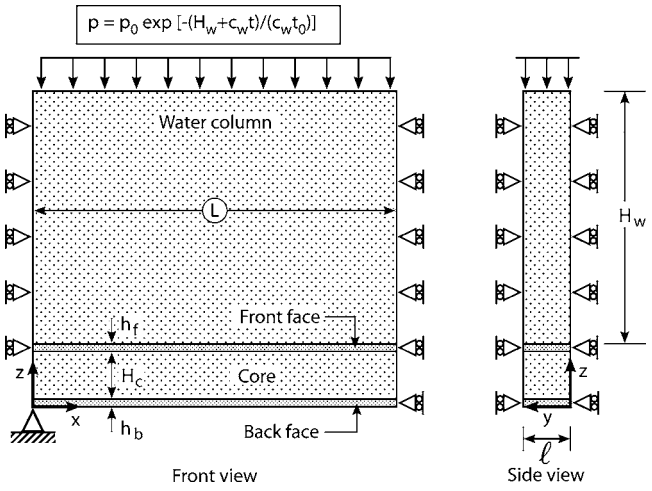


Fig. 6 A schematic of the numerical model used in ABAQUS/Explicit

Two different boundary conditions have been pursued.

Condition I. The two faces and the core are rigidly held at the supports. This condition has been used in most prior investigations [1–3,5–7].

Condition II. Only the back face is fixed at two outer supports. Periodic boundary conditions are imposed on the front face, but, otherwise, this face is free to displace into the core and toward the supports.

Note that the applicability of the analytic results (above) is unaffected by the choice of boundary condition through stage II. Differences arise in stage III.

A preliminary numerical assessment, used to refine the scope of the investigation, has revealed the following two characteristics.

a. The back face displacements attained for condition I are less than those for condition II. The difference is associated with the diminished front face stretching occurring for the latter.
b. The plastic strains in the front face are lower for condition II.

Given the apparently greater importance of tearing than deflection, *condition II has been chosen for most of the ensuing analysis*.

All calculations are performed using a representative free field impulse with pressure/time characteristics (p_0=100 MPa, t_0 =0.1 ms) indicative of domain II. The consequences of larger impulses and the occurrence of different domains (Fig. 3) will be elaborated in future assessments. The calculations are carried out for panels consisting of 304 stainless steel having the approximately bilinear stress-strain curves depicted on Fig. 5, with appreciable strain-rate sensitivity. All properties and constitutive law details are as summarized in [8]. In addition to calculations for edge-supported beams, some simulations have been carried out for freestanding panels in the form depicted in the side view of Fig. 6 (i.e., without kinematic constraint parallel to the direction of the impulse). These simulations provide results of the type shown as "numerical" on Fig. 2.

The ensuing sections of this article are organized in the following manner. In Sec. 4, the core topologies to be used are geometrically specified. In Sec. 5, the numerical scheme is described and some calibration results presented. Thereafter, in Sec. 6, mechanism identification is used to provide criteria that distinguish STC and SOC responses, emphasizing the different performance metrics in the two domains. In Sec. 7, an alternative analytical model is used to provide a consistent reinterpretation. In Sec. 8, the deflections and the plastic strains in the faces are examined and related to analytic results.

4 Geometries

All panels have fixed mass per unit area, m_{total}=160 kg/m^2, corresponding to a solid plate thickness, H_{eq}=2 cm. The half span is taken to be representative, L=1 m, width l=0.1 m, with the constraint that the core thickness, $H_c/L<0.4$. The three core topologies (Fig. 4) allow a wide range of geometric options for fixed m_{total}.

For the *square honeycomb*, the benchmark is a core with relative density, $\bar{\rho}$=0.03, and faces with equal mass, subject to a fixed spacing between core members, l=0.1 m. Around this benchmark design, the core relative density is allowed to vary between $0.01 \leqslant \bar{\rho} \leqslant 0.04$ while the spacing, l, is fixed. This is achieved by adding/subtracting mass to the faces in order to retain m_{total}. The ratio of the back to front face thickness, $\Delta=h_b/h_f$, is also varied between 1 and 6 for each $\bar{\rho}$. The latter is used to highlight the influence of the faces on the fluid/structure interaction.

The *I-core* density was varied between $0.01 \leqslant \bar{\rho} \leqslant 0.06$, with particular emphasis on relative core height in the range, $0.1 \leqslant H_c/L \leqslant 0.3$. The face thickness ratio Δ was varied between 1 and 7, again achieved by redistributing the mass between the faces and the core to maintain m_{total}. Two specific geometries provide the clearest distinction between strong and soft responses. The STC response is illustrated by a system having dimensions h_f

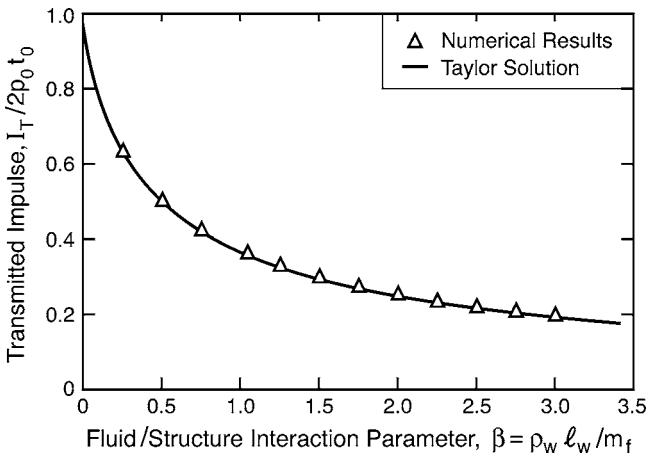


Fig. 7 The predictions of the transmitted impulse conducted for a solid plate using ABAQUS/Explicit and the comparison with the analytic solution given by Taylor [17]

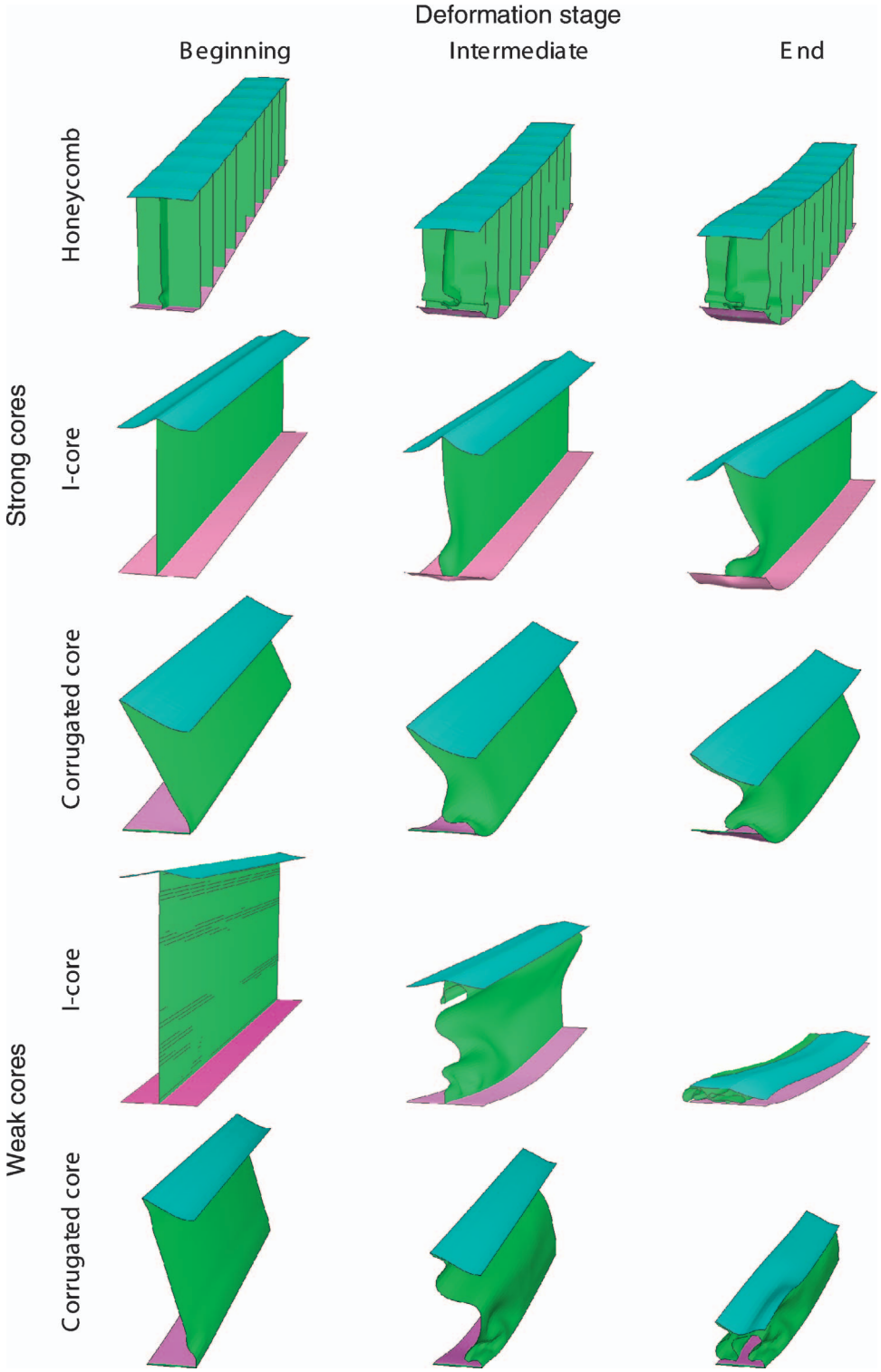


Fig. 8 The deformations predicted using the numerical model for the three cores shown in Fig. 4. Results for the I-cores and the corrugated cores are shown for both strong and soft responses.

=2 mm, h_b=6 mm, H_c=20 cm, and $\bar{\rho}$=0.06. The SOC response is demonstrated using dimensions: h_f=4 mm, h_b=12 mm, H_c =30 cm, and $\bar{\rho}$=0.013.

For the *corrugated core*, the relative densities were in the range $0.01 \leqslant \bar{\rho} \leqslant 0.05$. The height of the core was varied within $0.05 \leqslant H_c/L \leqslant 0.4$ by changing the angle, θ. The face thickness ratio Δ was allowed to vary subject to the constraint, $h_f \geqslant 2$ mm. Again, two geometries are used to distinguish the responses. The SOC response is demonstrated using the dimensions, $\bar{\rho}$=0.02, Δ=5.5, and H_c/L=0.3. The STC response is found using dimensions $\bar{\rho}$ =0.05, Δ=4, and H_c/L=0.2.

Given that the I-cores and the corrugations are anisotropic, choices must be made regarding the orientation. The present assessment is conducted in the "stronger" orientation. That is, for

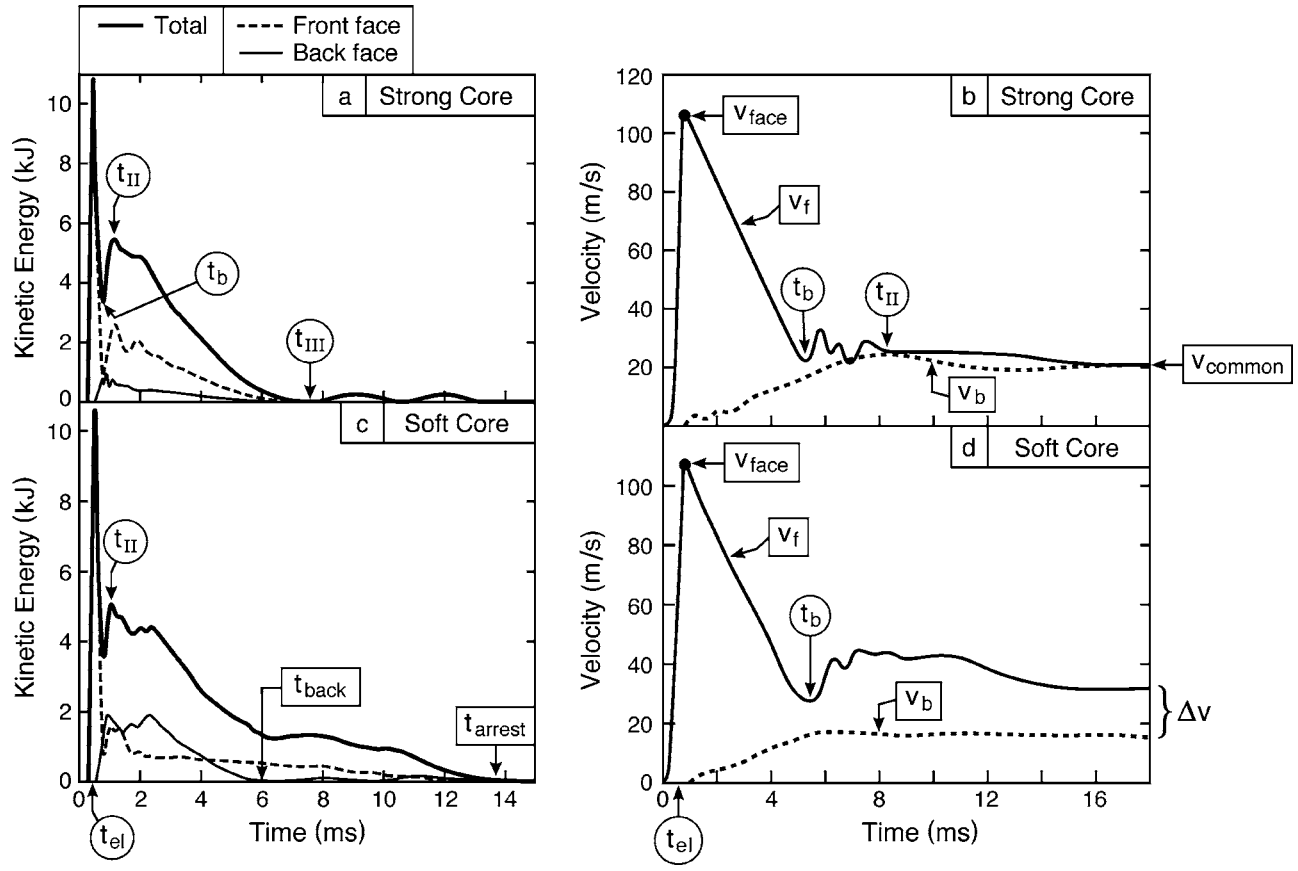


Fig. 9 The constituent velocities and kinetic energies obtained for corrugated cores: (*a*) and (*b*) refer to a strong core with relative density, $\bar{\rho}=0.05$, $\Delta=4$, and $\beta=9.375$ while (*c*) and (*d*) refer to a soft core with $\bar{\rho}=0.02$, $\Delta=5.5$, and $\beta=8.70$.

the I-and corrugated cores, the members are axial and continuous between the supports (Fig. 4). This orientation is orthogonal to that analyzed by Rabczuk et al. [7], who studied panels having the core members aligned transversely. With all other parameters held fixed, the differing orientations generate very different trends, as discussed later. The rationale for the present choice has been motivated by 3-D simulations (Appendix A), which suggest that the dynamic responses of square panels are largely governed by the behavior along the "strong" direction. Future assessments will elaborate and clarify the role of anisotropy.

5 The Numerical Scheme and Calibration Tests

The numerical model (Fig. 6) consists of a water column above the sandwich panel. Contact is enforced at the interface between the water and the panel. Symmetry boundary conditions are applied at all surfaces, except at the back face support (at $x=0$), which is clamped. A uniform pressure boundary condition is imposed on the top surface of the water column. Numerical tests have ascertained that, to correctly capture the fluid/structure interactions, the height of the column must satisfy: $H_w \geq 4H_c$.

The commercial code, ABAQUS/Explicit [16], is used. Eight-node 3D brick elements with reduced integration (C3D8R) are employed to model the water, while four-node shell elements (S4R), with five integration points through the thickness, are used to model the faceplates and the core members. The water is assumed to be linear elastic under compression, with zero tensile strength and zero shear modulus. (Some simulations have been performed with a small finite shear modulus, $G/c_w^2\rho_w=10^{-6}$, to affirm that the results are negligibly different from those with zero shear modulus). Thus, the pressure (p) in the water is given by:

$$p = -c_w^2\rho_w\varepsilon_V,\ (\varepsilon_V < 0)$$
$$p = 0,\ (\varepsilon_V < 0) \tag{18}$$

where $\varepsilon_V=\int_0^t(\partial v_i/\partial x_i)dt$ is the fluid volume strain. Thus when $\varepsilon_V \geq 0$, all stresses in the water become zero, causing cavitation.

Careful consideration has been given to the artificial bulk viscosity coefficients. The default coefficients designated in ABAQUS/Explicit give excessive dissipation in the water that diminishes the pressure before the impulse reaches the structure. Consequently, the coefficients have been systematically reduced until the blast wave pressure closely matches p_0 when it confronts the panel. The values utilized are $b_1=0.02$ and $b_2=0.2$ [16]. To affirm the fidelity of the approach, numerical results for a free-standing metal plate have been obtained and compared with the analytic solution given by Taylor [17] (Fig. 7).

For each simulation, the following parameters are obtained:

i. The final mid-point deflections of the faces: front, δ_f, and back, δ_b.
ii. The core crushing strain, ε_c.
iii. The velocities of the faces (front, v_f and back, v_b) as a function of time, t, after the impulse contacts the front face.
iv. The kinetic energies of the two faces (KE_{front} and KE_{back}) and the core (KE_{core}) and the total (KE_{total}), as a function of time.
v. The plastic dissipation, $W_{pl}^{constituent}$, in each constituent as a function of time.
vi. The reaction forces, P_{react}, at the supports as a function of time.

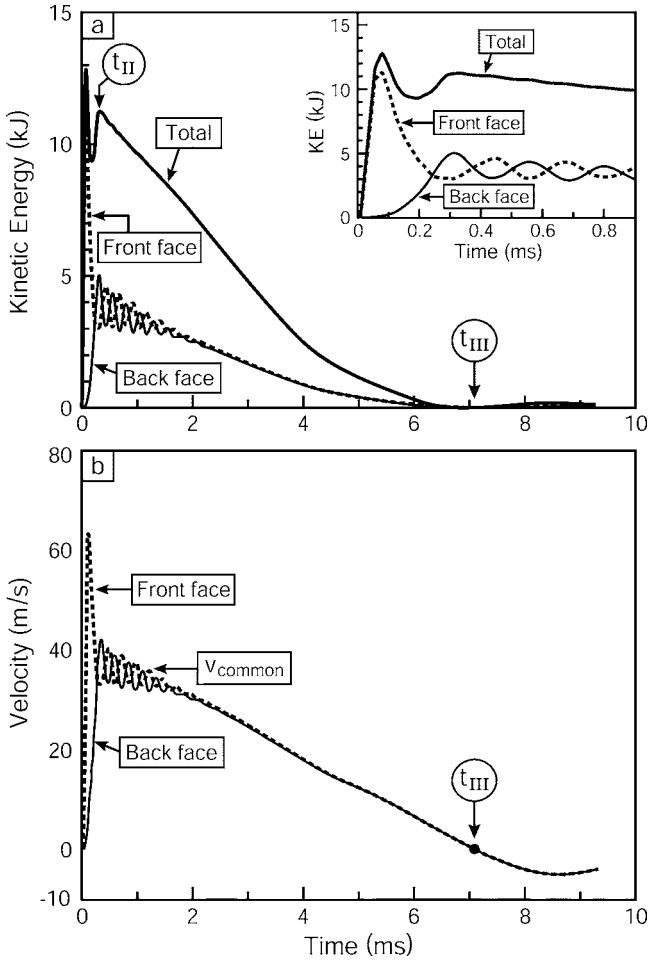


Fig. 10 The kinetic energies and velocities for a square honeycomb core with $\bar{\rho}=0.03$, $\Delta=1$, and $\beta=2.68$

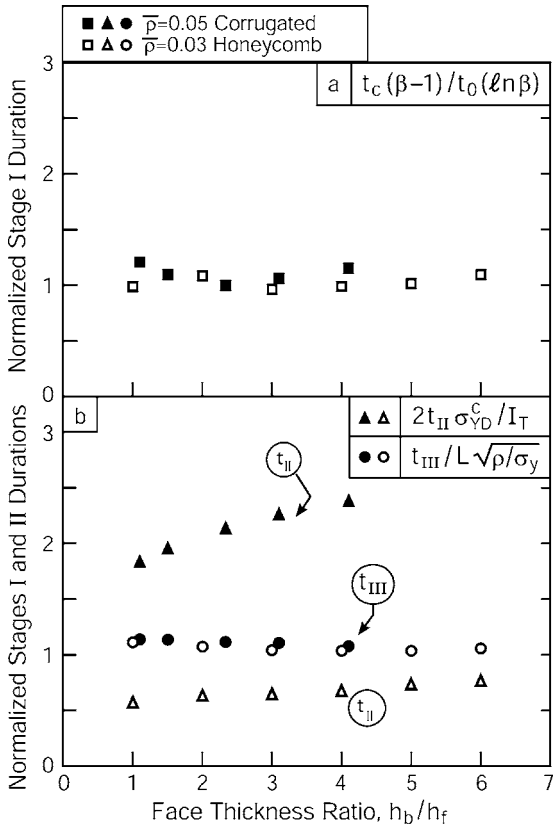


Fig. 11 The durations of stages I, II, and III normalized by the expressions derived using the analytic model (Sec. 1): $t_c/t_0 \ln\beta/(1-\beta)$, $t_{II}/(I_T/2\sigma^c_{YD})$, and $t_{III}/L\sqrt{\rho/\sigma_Y}$

vii. The largest plastic strain, $\varepsilon^{\max}_{pl}$, in the two faces.

The reaction force, P_{react}, is taken as the summation of all nodal forces on the clamped edge of the back face.

At any instant, the average speed, v_i, of element i is obtained through finite element interpolation, once the velocities of its nodes have been ascertained. Then its kinetic energy is obtained ($\text{KE}_i=(1/2)m_i v_i^2$, with m_i the mass of the element) and the total kinetic energy of the constituent determined from the summation:

$$\text{KE}_{\text{constituent}} = \sum_{i=1,n_E} \text{KE}_i, \tag{19a}$$

where n_E is the number of elements in the constituent. The total kinetic energy in the structure is ascertained as:

$$\text{KE}^* = \text{KE}_{\text{front}} + \text{KE}_{\text{back}} + \text{KE}_{\text{core}} \tag{19b}$$

The kinetic energy in the attached water at the end of stage II is also evaluated for STC designs. For this purpose, the average velocity of the panel, v_{common}, at the end of stage II is first obtained. Then, by monitoring the fluid velocity at various locations near the wet surface, the thickness, $|x_a|$, of the attached layer adjacent to the panel (namely the thickness of the fluid having the same velocity) is determined. The kinetic energy of the attached layer is then $\text{KE}_{\text{water}}=lL\rho_w|x_a|v^2_{\text{common}}/2$.

The plastic dissipation in each element is calculated using:

$$W^i_{pl} = \Delta V \int_0^{\varepsilon_p} \sigma_Y d\varepsilon_p, \tag{20}$$

with ε_p the effective plastic strain and ΔV the volume of the element. Thereafter, the total plastic dissipation in each constituent is given by the summation,

$$W^{\text{constituent}}_{pl} = \sum_{i=1,n_E} W^i_{pl}. \tag{21}$$

In ABAQUS, such summations are performed automatically.

Since the bending deflections of the panel are negligible in stages I and II, the velocities can be ascertained from the associated kinetic energies: $v_f=\sqrt{2\text{KE}_{\text{front}}/m_f}$ and $v_b=\sqrt{2\text{KE}_{\text{back}}/m_b}$.

6 Mechanism Identification

General Features. Many different calculations have been used to probe the response space. Only the distinctive results are presented. The displacement sequences (Fig. 8) affirm that two inherently different (STC and SOC) responses exist. The most obvious distinction is that the soft cores collapse during stage III. For the range of topologies examined, honeycomb cores are always STC, while cores with I- and corrugated topologies exhibit both responses, dependent on geometry. The major distinction between the two domains can be ascertained from typical plots of the velocities and kinetic energies (Figs. 9 and 10) of the constituents. To be complementary, and to reveal aspects of the response over different time domains, the KE results in Fig. 9 are presented over the entire structural response time, while the velocity results are confined to shorter times (of order t_{II}). For clarity of presentation, the KE plots for the core have been excluded (since they can be readily inferred from the total KE and the KE for the two faces).

Strong Cores. In *all STC designs*, exemplified by the corrugated

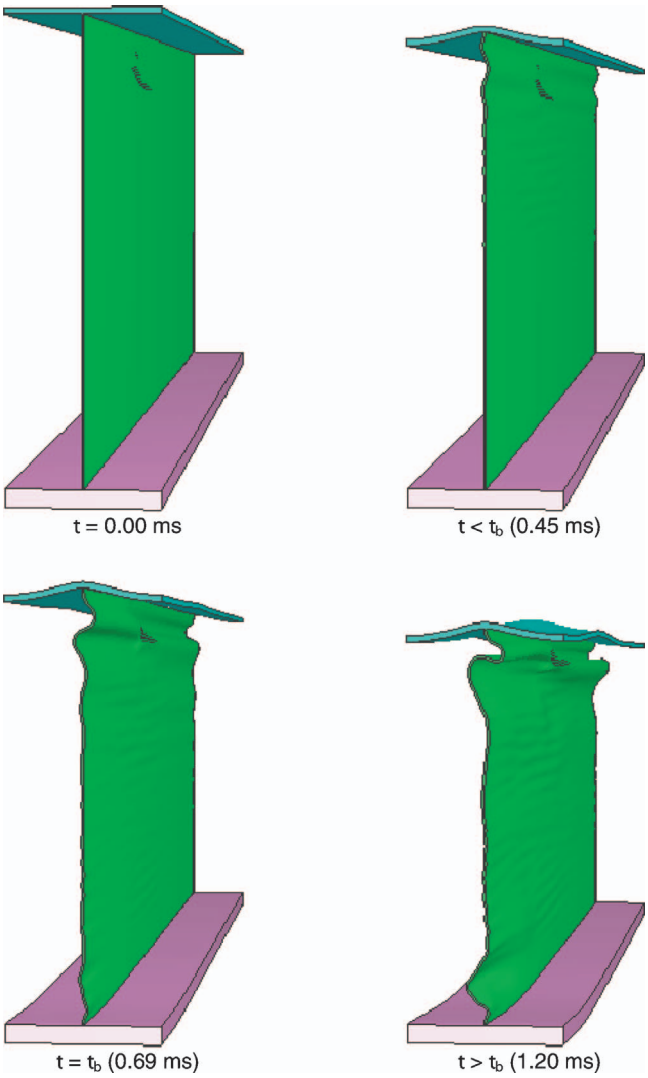


Fig. 12 A deformation sequence for a soft I-core showing the dynamic elastic buckling of the core near the back face at instant t_b (H_c=0.3L, $\bar{\rho}$=0.013, Δ=3, and β=4.69).

core with $\bar{\rho}$=0.05, Δ=4, and β=9.375 (Figs 9(*a*) and 9(*b*)) and the honeycomb with $\bar{\rho}$=0.03, Δ=1, and β=2.68 (Fig. 10), the total kinetic energy in the structure, KE^*_{II}, exhibits two peaks. The first is coincident with time t_c. The second coincides with the end of stage II, t_{II}, when all constituents have attained a common velocity, v_{common}, apparent from plots of the front and back face velocities (Figs. 9 and 10). Note that, for some strong cores, the front face acceleration stops *before the end of stage II*, at time $t_b < t_{II}$ (Fig. 9). Conversely, the back face continues to accelerate up to t_{II}, before attaining v_{common}. The significance of t_b will become apparent later. In stage III, the KE's and velocities of all constituents decrease as energy is dissipated by plastic deformation of the faces and the core. *They reach zero simultaneously in all constituents*. This occurs at time t_{III}, coincident with the back face reaching its final, permanent deflection (Figs. 9 and 10). Some differences among core topologies are apparent at durations close to t_{II} (cf. Figs. 9 and 10). Namely, the square honeycomb and I-cores exhibit distinctive elastic oscillations, duration t_{el}, as the two faces settle into a common velocity (Fig. 10). Similar oscillations have been reported by Deshpande and Fleck [4]. Conversely, the oscillations found for the strong corrugated cores are less distinctive (Fig. 9) and the two faces approach the common velocity differently. The importance of this distinction will become apparent later. The durations of the stages (Fig. 11) merit comment. Those for stage I, $t_c/t_0 \ln \beta/(1-\beta)$, and stage III, $t_{III}/L\sqrt{\rho/\sigma_Y}$, are the same for all strong cores and completely consistent with the analytic models (Eqs. (13)). However, $t_{II}/(I_T/2\sigma^c_{YD})$ not only differs from the analytic model, but also has different values for honeycomb and corrugated cores. The implications will become apparent later.

Soft Cores. In *SOC designs*, exemplified by a corrugated core with $\bar{\rho}$=0.02, Δ=5.5, and β=8.70 (Figs. 9(*c*) and 9(*d*)), there are marked differences in the kinetic energies and velocities from the STC examples. While KE_{back} still approaches zero at time $t \approx t_{III}$, *the front face and the core continue to move and their kinetic energies only become zero at longer times*, $t \approx t_{\text{arrest}} \approx 2t_{III}$. However, beyond t_{III} the back face is elastic and dissipation occurs only due to core and front face plasticity. The temporal pattern of velocities also differs. Most notable, the accelerations of the faces stop at time, $t \approx t_b < t_{II}$ (Fig. 9(*d*)). One consequence is that the velocity acquired by the back face is smaller than that found in comparable strong cores. Additionally, between t_b and t_{II}, the front face velocity *increases*. The implication is that the soft response is coincident with events occurring at time t_b. An investigation of the phenomena occurring in soft I-cores with $\bar{\rho}$=0.013, Δ=3, and β=4.69 (Fig. 12) reveals that, *exactly at* t_b (0.55 *ms in this case*), the core buckles at the intersection with the back face.

Displacements. The back face displacements are distinctive (Fig. 13), visualized using displacement surfaces with the following coordinates: core density, $\bar{\rho}$, and relative thickness, Δ. *For all strong cores*, δ_b *is lowest at large* $\bar{\rho}$. *Conversely*, *for all soft cores*, δ_b *is lowest for small* $\bar{\rho}$. All cores result in smaller displacements at larger Δ.

Plastic Strains. Preliminary assessment of the face tearing susceptibility uses the plastic strain as a metric. A more complete ranking awaits incorporation of a dynamic failure criterion into the FE code. The premise is that *the largest equivalent plastic strain in the faces*, at any lengthwise location, averaged over the width, l, provides the relevant scaling. (Tearing of the core is regarded as relatively benign.) A series of plastic strain results is presented in Fig. 14. Several features emerge.

a. The strains in the front face are considerably lower than those found in the equivalent impulsive loading of the same panels subject to support condition I (results not shown).
b. For the front face, in all cases (that is, for both STC and SOC), the strains are lowest for panels having similar front and back face thickness, $\Delta \approx 1$, and cores with lowest relative density, $\bar{\rho} \approx 0.01$. Consequently, there is a conflict with designs based on deflection (Fig. 13). *The conflict is least for soft cores, which demonstrate benefit in both tearing and deflection at low relative density*. The remaining conflict is in front face thickness, because for thin faces, relatively large plastic strains arise due to bulging between core members (Fig. 8). A compromise will be required.
c. The strains in the back face are invariably larger than at the front, especially in the vicinity of the supports. But again, the SOC designs result in smaller strains.

Force and Impulses. The reaction forces at the supports reveal corresponding features (Fig. 15(*a*)). The strong cores generate relatively large forces over duration t_{III}. *During the same interlude*, *the soft cores impart lower forces*. However, for SOC designs the forces persist to longer times, $t > t_{III} \approx 6$ ms, and may become larger after t_{III}. Trends in the peak force with core design, plotted in Fig. 16, demonstrate the specific benefits of soft cores. A further assessment of reaction forces will be described elsewhere [13]. The integral of the forces over time provides the total transmitted impulse, I_{total} (Fig. 15(*b*)). This measure of the impulse is found to be approximately the same as the free field

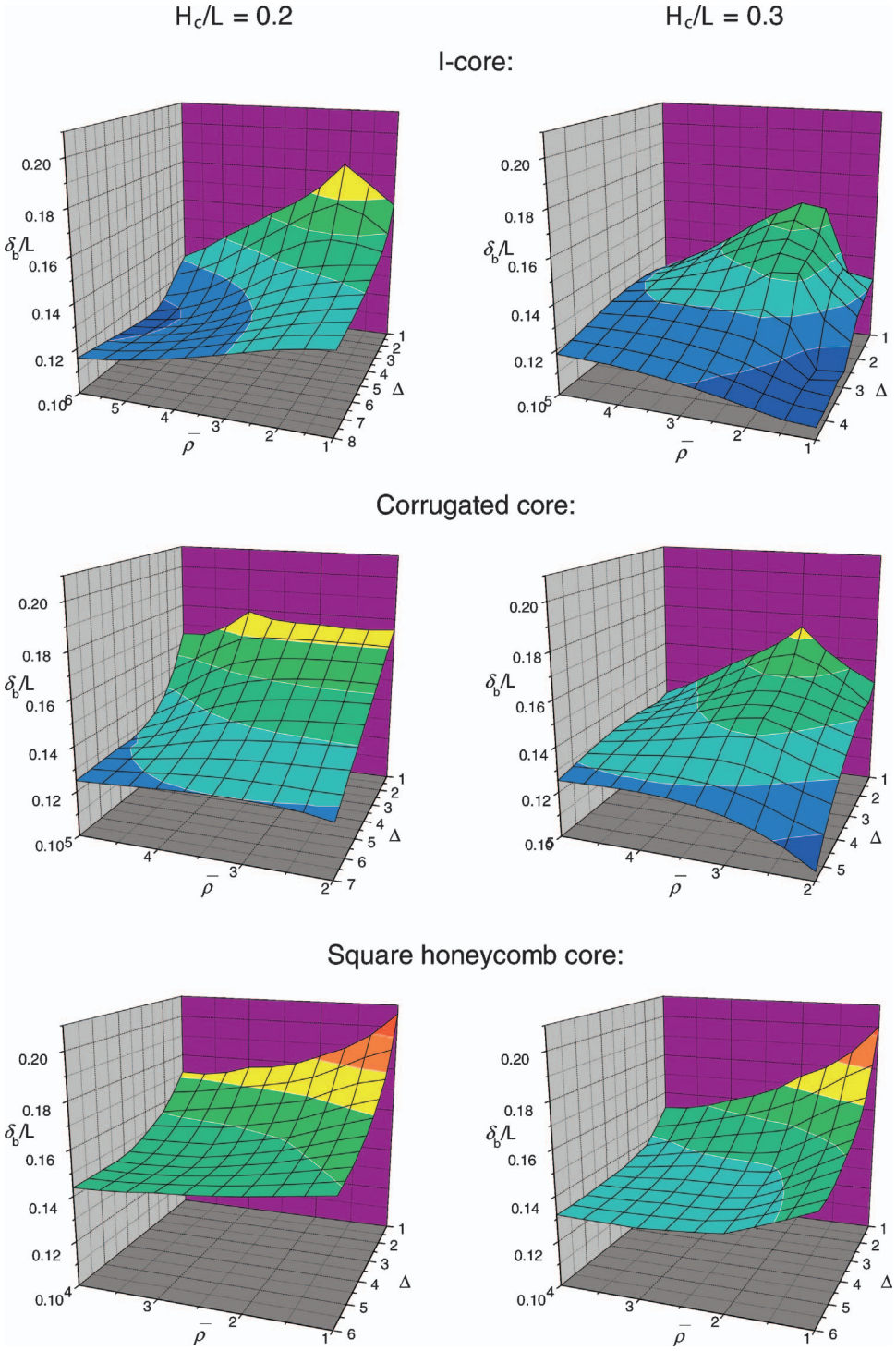


Fig. 13 A synopsis of back face displacement ascertained for a wide range of strong and soft cores. The coordinates are the ratio of back to front face thickness, Δ, and the relative density of the cores, $\bar{\rho}$.

momentum $I_0=p_0t_0$ in all cases. As noted earlier, the back face is elastic for $t>t_{III}$, indicating that the impulse beyond t_{III} does not make a significant contribution to the permanent deformation of the panel. The momentum acquired by the structure, plus the attached water, at the end of stage II is more pertinent to the goals of the present assessment. It is the analog of I_T, specified by the analytic model (11*a*), which dictates KE_{II}. The momentum determined numerically will be given a separate designation, M_T. Trends in M_T and I_{total} are summarized in Fig. 17 for a range of core designs.

To compare these results with those predicted by the analytic model, the dynamic strength of the core, σ^c_{YD}, must be ascertained. A detailed assessment has found that, *for the present designs*, (8) *provides acceptable fidelity* upon incorporating the front face velocity ascertained at t_c from the modified Taylor formula (3*a*).

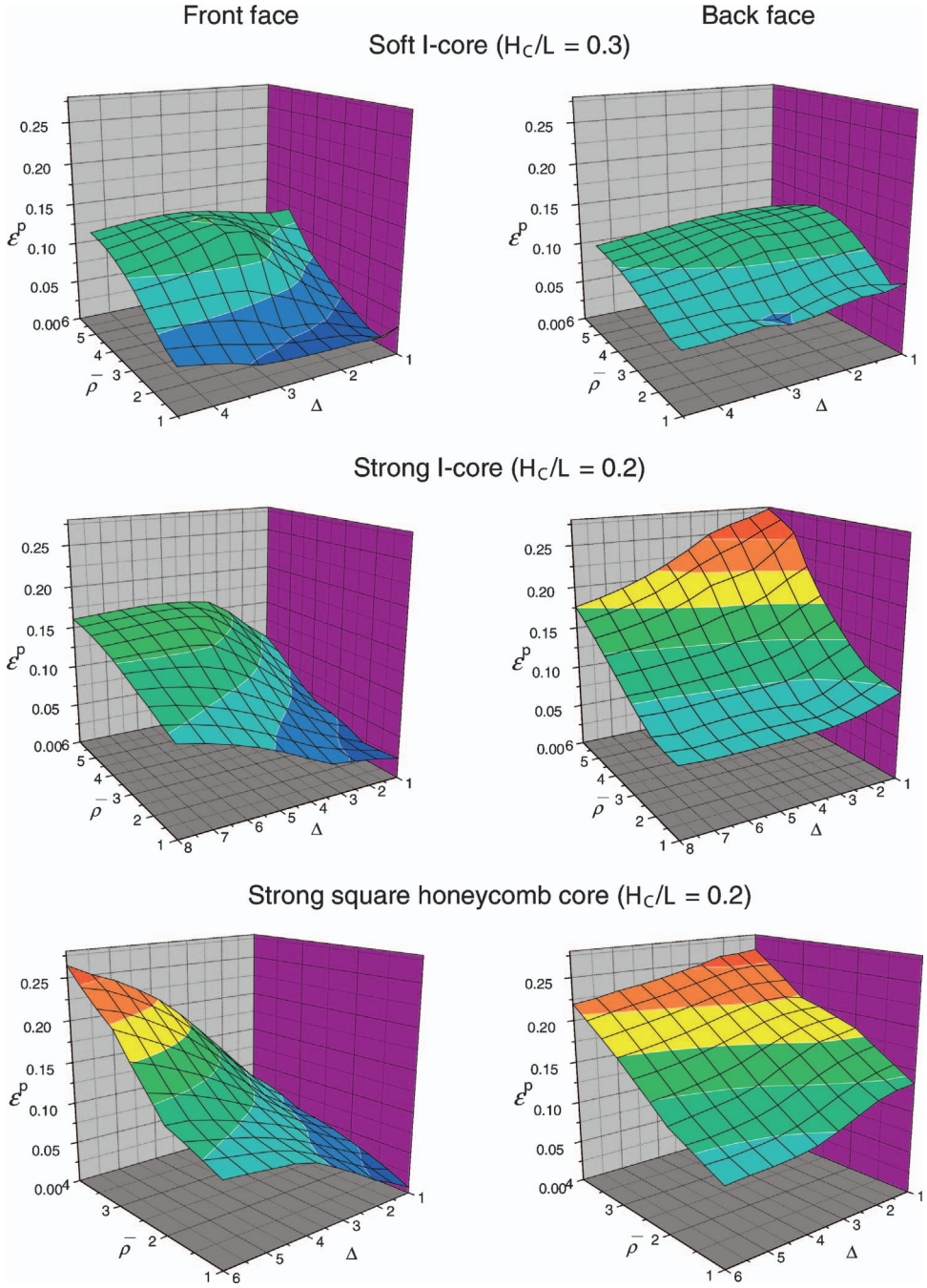


Fig. 14 A synopsis of trends in the maximum plastic strain in the front and back faces, ascertained for a range of strong and soft cores. The coordinates are the ratio of back to front face thickness, Δ, and the relative density of the cores, $\bar{\rho}$.

Initially, the momentum M_T is compared with I_T determined from (11*a*). It is apparent (Fig. 17) that the analytic formula substantially underestimates the momentum. The magnitude of the underestimate becomes most pronounced for designs with thin front faces and for cores with lower dynamic strength. *The discrepancy suggests that the water imposes a larger momentum than assumed by the existing model*.

Kinetic Energies. Trends in the kinetic energy KE_{II} acquired at the end of stage II with attached water included are plotted in Fig. 18 for a range of STC designs. Comparison with the analytic predictions again reveals that the model substantially underestimates the numerical results.

Transition. The time scales associated with all of the foregoing results have been used to present a plot of the proposed transition parameter, Π (see Eq. (17)), as a function of core relative density $\bar{\rho}$ and ratio of back to front face thickness Δ (Fig. 19). Overlaying the soft and strong core responses indicates that most of the results can be distinguished by a critical value, $\Pi_{th} \approx 0.2$. That is, smaller values of Π result in strong responses and *vice versa*. However, there are discrepancies, and it remains to establish a rigorous criterion for distinguishing soft and strong responses.

In summary, STC designs can be distinguished by the following characteristics.

a. The KE of the two faces and the core all approach zero at time after impact, $t \approx t_{III}$.

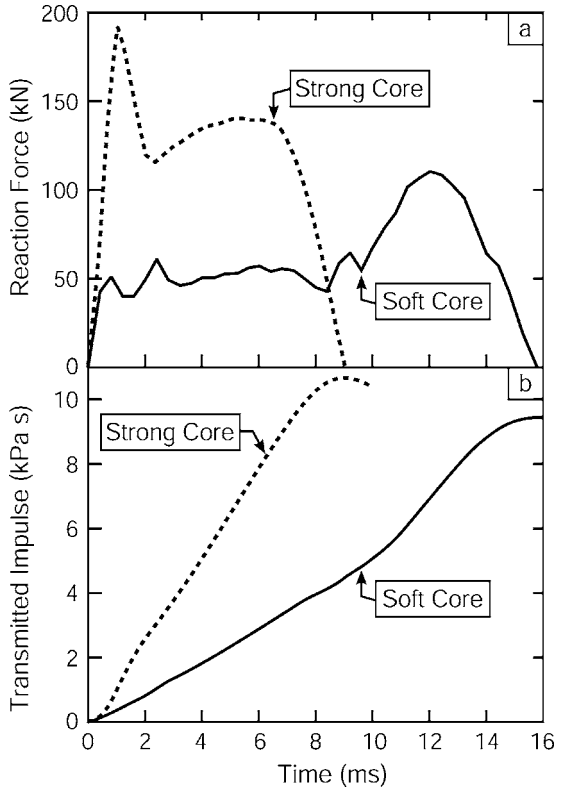


Fig. 15 (*a*) The reaction forces at the supports typifying the difference between strong and soft cores (results for I-cores are shown). (*b*) The corresponding values of the impulse transmitted to the structure determined from the reaction forces.

b. The front and back faces attain a common velocity at time, $t \approx t_{II}$.

The corresponding SOC characteristics are as follows.

a. The front face and the core retain appreciable KE at times after impact, $t > t_{III}$.
b. The acceleration of the faces stops at time after impact, $t \approx t_b < t_{II}$.
c. Between $t_b < t < t_{II}$, the front face accelerates.

The analytic model consistently underpredicts the acquired momentum and kinetic energy, attributed to an underestimate of the water attached to the wet face at the end of stage II. An alternative FSI model that rectifies this discrepancy for STC designs is presented in the next section. *The remainder of the article is confined to issues affecting the responses of strong core designs.*

7 Alternative Fluid/Structure Interaction Model

7.1 The New Hypothesis. An examination of the velocities and pressures in the water (Fig. 2) indicates that, beyond stage I, a zone of cavitated water exists that widens rapidly with time. Consequently, during stage II, the structure is in contact with cavitated water moving in the same direction. The modified Taylor solution (Eqs. (2) and (3)) does not characterize the response of this zone. An alternative model is needed. The salient features are as follows. Once formed, the left boundary of the cavitation zone (at $x = x_{cb}$) acquires an initial rate of translation, $v_{cb} \approx 10 c_w$, away from the panel and then slows asymptotically to c_w (Fig. 20). This speed is supersonic [4] but, more importantly, is initially two orders of magnitude larger than the peak velocity acquired by the front face. Consequently, the cavitation zone expands on a time frame much shorter than the panel structural response, t_{III} [3]. In addition, the rightmost boundary of the cavitation zone moves quickly towards the panel. Inside the cavitation zone, the pressure is zero, yet the cavitated fluid has *positive velocity* everywhere, causing it to move toward the panel at velocity $v_r(x)$ (Fig. 21). That is, during stage II, a point in the cavitated fluid, initially at x, retains a fixed positive velocity v_r. At locations in the fluid close to the face, v_r exceeds the velocity of the panel, which is decelerating. Such regions within the cavitated fluid (visualized as a porous medium [4]) may thus reattach to the panel, adding momentum. As this happens, these fluid regions decelerate to a velocity approximately equal to the rate of translation of the face at the instant of reattachment. The subsequent fluid pressure in this reattached layer is positive: however, the level oscillates (Fig. 2) due to acoustic interactions with the face and the ongoing reattachment of water. A model incorporating these effects, but eliding some of the complications, will be examined in the following section. The hypothesis is that *any regions within the cavitated fluid having velocity exceeding the ultimate common speed of the structure (attained at the end of stage II) will reattach by the end of that stage and impart additional momentum.*

This hypothesis enables the following three-step analysis.

Step I. Find the location of the cavitation front, x_{cb}, from (2*a*) by imposing $p = 0$.

Step II. Determine the velocity of the water at the cavitation front $v_{cb}(x_{cb})$ by inserting x_{cb} into (3*a*). This velocity is found to be insensitive to the time during stage II.

Step III. Assert that this is the velocity of the cavitated water, $v_r(x)$, relative to its location x prior to the arrival of the incoming blast wave.

7.2 Velocity of Cavitated Water. The development is pursued by evaluating the instant, $t_{cb}(x) \geq t_c$, at which cavitation occurs at location x in the water, by solving (2*a*) at zero pressure:

$$p(x,t_{cb}) = p_0 \left[\exp\left(\frac{x - c_w t_{cb}}{c_w t_o} \right) - \left(\frac{2\beta}{1-\beta} + \frac{\sigma_{YD}^c}{p_0} \right) \times \exp\left(- \frac{\beta (x + c_w t_{cb})}{c_w t_o} \right) + \frac{1+\beta}{1-\beta} \exp\left(- \frac{x + c_w t_{cb}}{c_w t_o} \right) + \frac{\sigma_{YD}^c}{p_0} \right] = 0 \qquad (22)$$

The ensuing time histories of the location of the cavitation front at time t_{cb} and its rate of translation ($v_{cb} = |dx_{cb}/dt|$) are plotted in Fig. 20. Note that v_{cb} is large, consistent with the preceding discussion. The fluid velocity $v_r(x)$ at the cavitation boundary at the instant of cavitation is now obtained by inserting $t = t_{cb}$ into Eq. (3*a*):

$$v_r(x) = v(x,t_{cb}) = \frac{2p_o}{\rho_w c_w} \exp\left(\frac{x - c_w t_{cb}(x)}{c_w t_o} \right). \qquad (23)$$

The velocity, $v_r(x)$, obtained in this manner refers to the position, x, of the fluid in its undistorted configuration (in which the density is ρ_w). Since the pressure and pressure gradient in the cavitated zone are zero for $t \geq t_{cb}$ (at least until possible reattachment to the panel), the fluid velocity, $v_r(x)$, remains constant at times $t_c \leq t \leq t_{II}$. It, therefore, represents the actual residual velocity of the water in the cavitation zone. This time invariant velocity profile, $v_r(x)$, is plotted in Fig. 21. Observe the relatively high velocity of the fluid near the sandwich panel.

7.3 The Momentum Transfer. The momentum transferred to the system at the end of stage II becomes:

$$M_T = (m_f + m_c + m_b - \rho_w x_a) v_c. \qquad (24)$$

The equivalent momentum for the panel plus the same mass of water is:

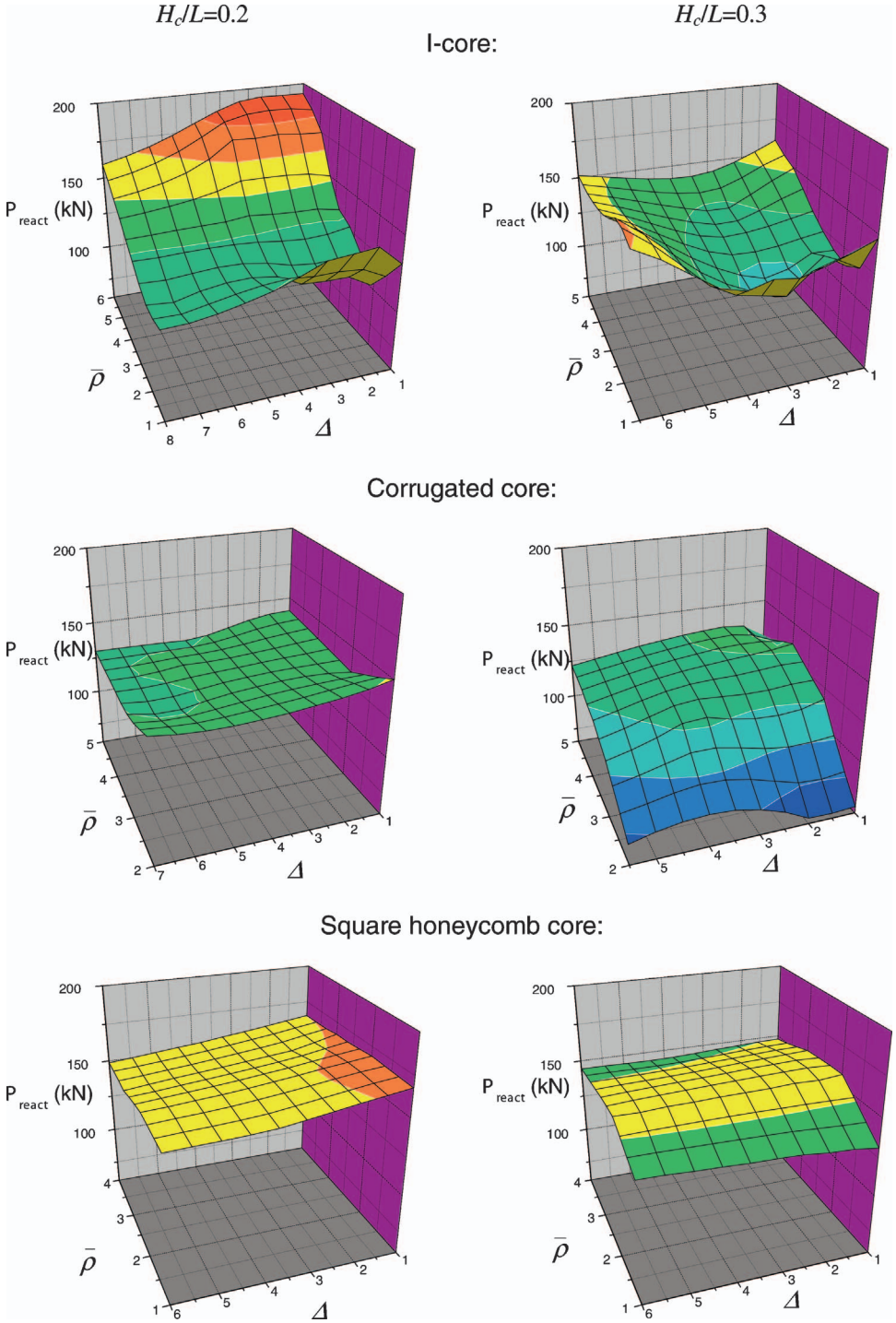


Fig. 16 A synopsis of trends in the peak reaction force with relative density and ratio of back to front face thickness

$$M_T = \rho_w \int_{x_a}^{x_c} v_r(x)dx + M_F + M_B. \tag{25}$$

Equating (24) and (25) gives

$$v_c = \frac{\rho_w \int_{x_a}^{x_c} v_r(x)dx + M_F + M_B}{m_f + m_c + m_b - \rho_w x_a}. \tag{26}$$

Combining (26) with the velocity requirement, $v_r(x_a) = v_c$, gives an equation

$$(m_f + m_c + m_b - \rho_w x_a)v_r(x_a) - \rho_w \int_{x_a}^{x_c} v_r(x)dx = M_F + M_B \tag{27}$$

that can be solved numerically to obtain x_a and thus v_c. Approximate analytic formulas are given in Appendix B.

The fidelity of this model is tested by comparing the velocity distribution predicted by the improved model with that given by a finite element calculation for a freestanding foam core sandwich panel. The results are presented in Fig. 22. The excellent consis-

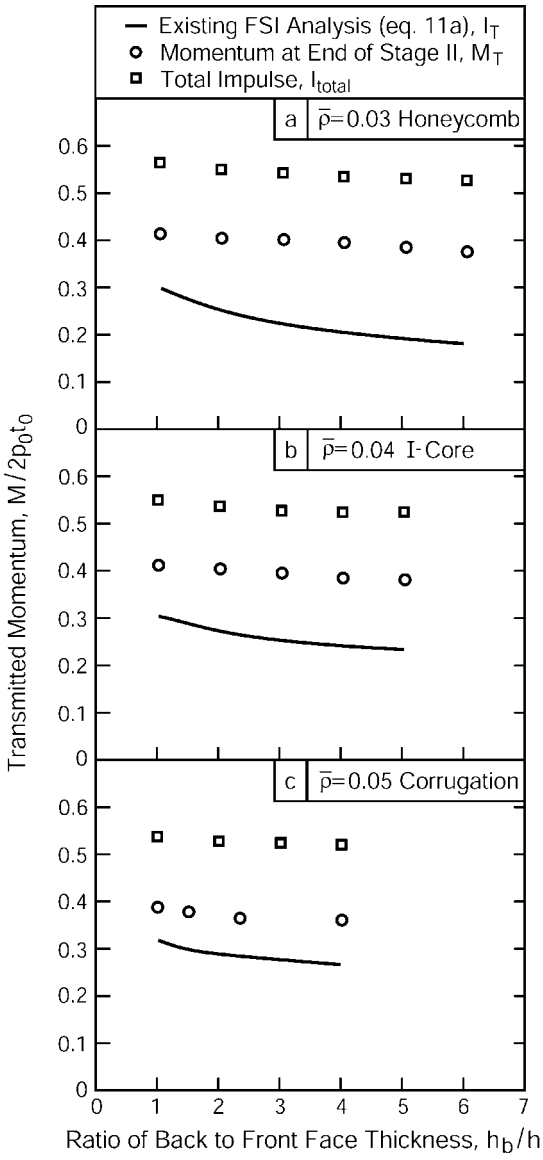


Fig. 17 Trends in the momentum acquired at the end of stage II (structure plus attached water), designated M_T, with ratio of back to front face thickness for a range of strong cores. The total impulse I_{total} transmitted to the system is also plotted. Comparisons with the predicted I_T from the existing analytic model, Eq. (11*a*), are included.

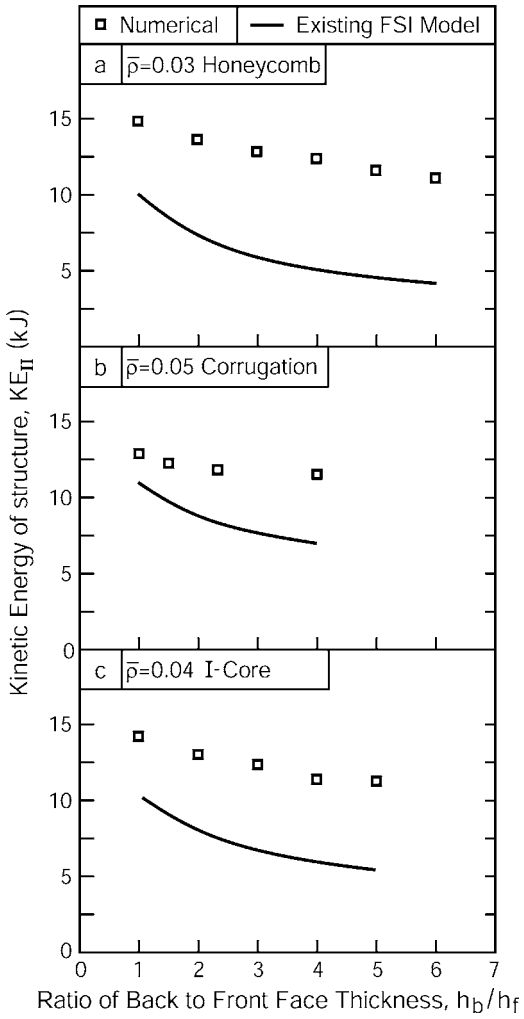


Fig. 18 Trends in the kinetic energies acquired at the end of stage II with ratio of back to front face thickness for a range of strong cores. Results for KE_{II} (structure plus attached water) are plotted. A comparison with the existing analytic model (Eq. (12*b*)) is included.

tency between the velocity profiles from the model and the simulations justifies the postulate about the common velocity and the extent of reattached water. Another validation is provided by a comparison with numerical results for panels with *foam cores* and steel front and back faces (Fig. 23) [4]. The momentum has been calculated for a freestanding sandwich panel in which the core yield strength parameter, σ^c_{YD}/p_0, was varied from 10^{-3} to 0.5 [4], embracing domains II and III. Inspection of Fig. 23 reveals excellent agreement between the improved model and the finite element simulations.

7.4 Reinterpretation of the Acquired Momentum. Based on this new interpretation of the momentum transfer, the transmitted impulse and the kinetic energy have been recomputed and compared with the present numerical results for strong cores in Figs. 24 and 25. The evident consistency between the numerical and analytic results *affirms the applicability of hypothesis to structured cores*. Note that the total momentum, inclusive of that in the water (Fig. 24), is appreciably larger than that in the structure and has a different dependence on geometry. Moreover, additional momentum is transferred after the end of stage II, accounting for the difference between the total transmitted impulse (ascertained from the reaction forces) and M_T, evident in the Fig. 17. It will be apparent in the following section that *the deformation of the structure in stage III is controlled by* M_{Total} *and not the total impulse*. The results also reveal that the benefit of the thin front face in terms of the deflection metric (Fig. 13) is not attributable to its influence on momentum transfer [2,3], but, rather, to the reduced deformations that occur in stage III because of the thick back face [13].

The momentum acquired by the corrugated core compares least favorably with the model, in the sense that the model overpredicts the momentum by ~15%. This discrepancy appears to be linked to the difference in t_{II} between this core and the others (Fig. 11) for reasons yet to be understood. Discrepancies remain for the SOC designs, which do not attain a common velocity. A model capable of predicting the momentum transfer for such cores remains to be developed.

Now that a viable model has been devised for the impulse transmitted to a strong core structure through stage II, the model can be used as input for evaluation of the stage III response, described in the next section.

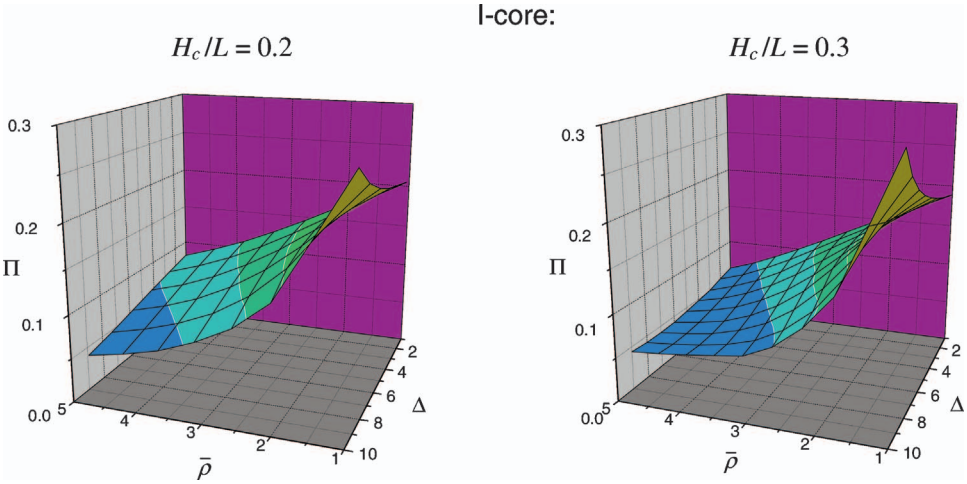


Fig. 19 Trends of the transition parameter Π for I-core design

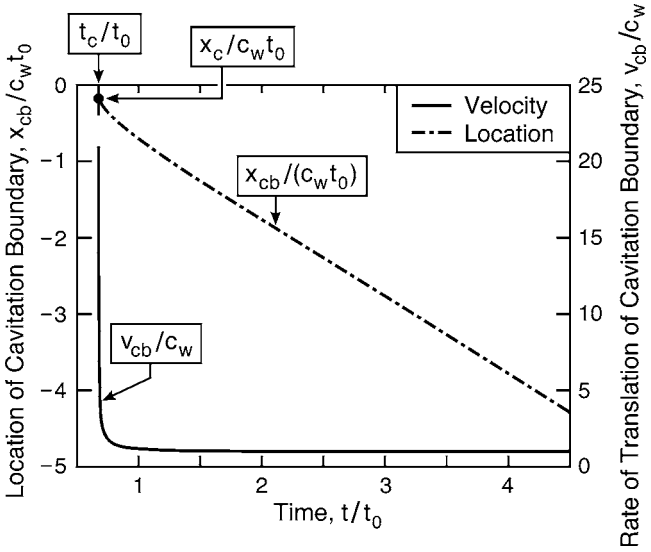


Fig. 20 Location and rate of translation of the cavitation boundary as a function of time for the case shown in Fig. 2

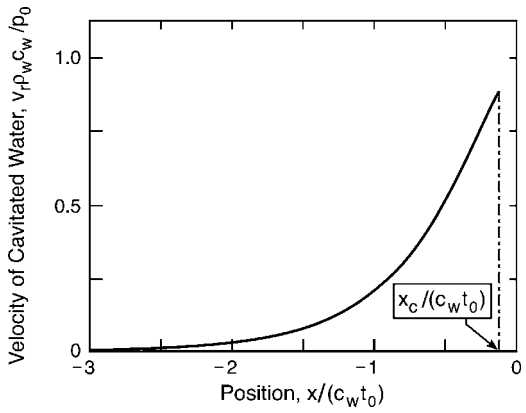


Fig. 21 Characteristic velocity of cavitated fluid for the case shown in Fig. 2

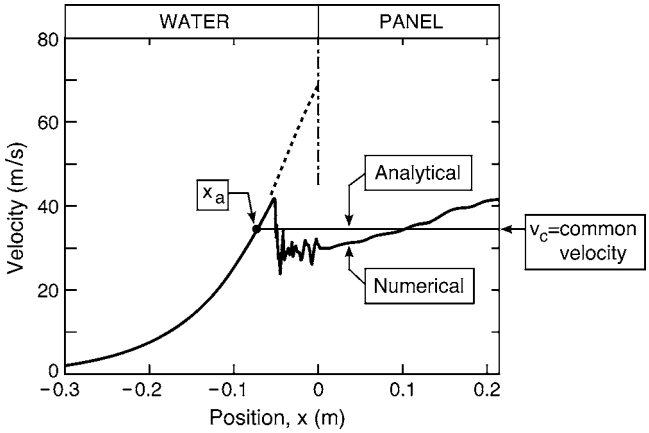


Fig. 22 The velocity distributions at the end of stage II for a foam-core panel ascertained from the new analytic model compared with the result obtained using ABAQUS/Explicit. The core has relative density of $\bar{\rho}_c$=0.03, strength of σ^c_{YD}/p_0=0.18, and height of H_c=0.2 m. The thickness of the front face is h_f =6 mm so that β=3.125 and that of the back face is h_b=8 mm.

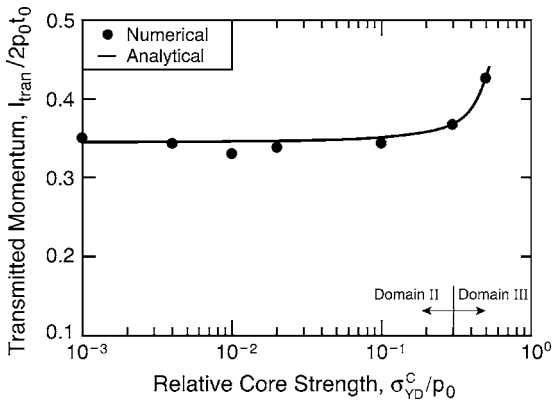


Fig. 23 Comparisons of transmitted momentum calculated using the present model and that obtained numerically by Deshpande and Fleck [4]. The freestanding panel unit has face thickness h_f=h_b=10 mm so that β=1.875, core height H_c=0.1 m, the density of the parent metal is 8000 kg/m^3, and the relative density of the core is $\bar{\rho}_c$=0.1.

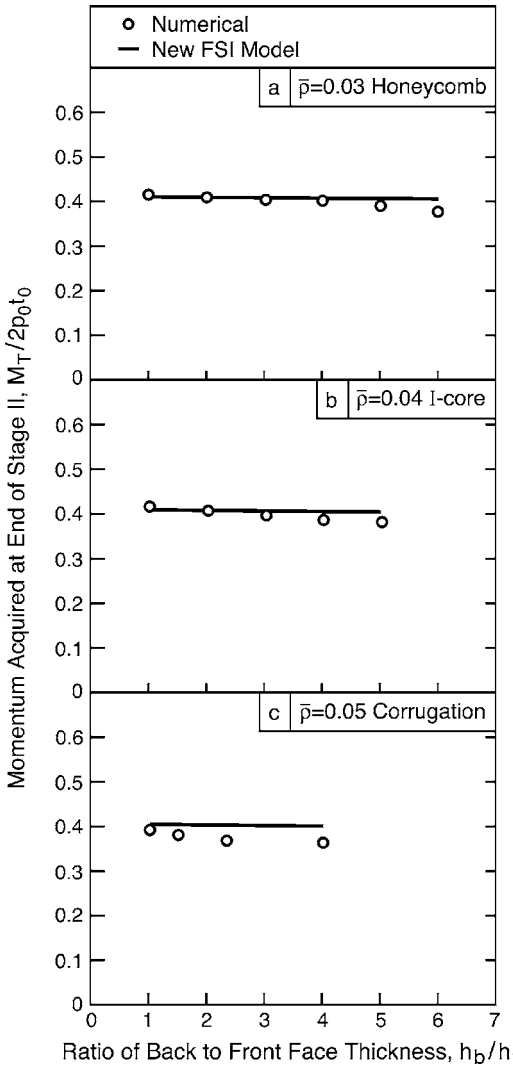


Fig. 24 The momentum at the end of stage II (structure plus attached water) determined numerically and the comparison with the new analytic model (using Eq. (24))

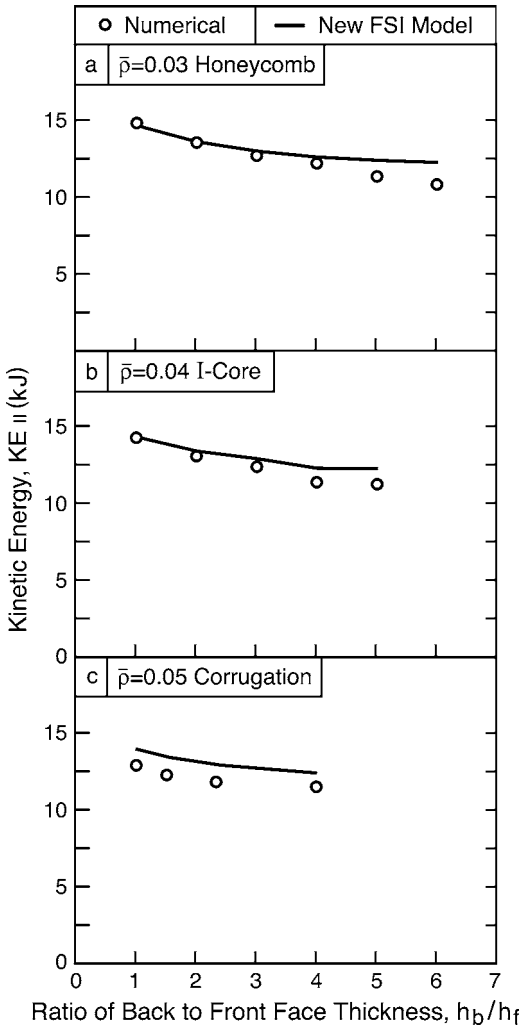


Fig. 25 The kinetic energy at the end of stage II (structure plus attached water) determined with the new analytic model and comparison with numerical calculations for STC design

8 Deflections

The first step in the analysis of the deflections is to ascertain whether KE_{II} (determined using the new model) governs the plastic dissipation, W_{pl}, occurring in stage III, that is, whether the extra momentum transferred during stage III can be discounted. With this objective, plots of the trends in KE_{II} and in the total plastic dissipation, W_{pl}^{total} with panel design (Fig. 26) demonstrate that, when KE_{II} includes the correct contribution from the water, it slightly exceeds W_{pl}^{total}. The slight excess is consistent with a small contribution to stage III dissipation from the elastic reverberations shown on Figs. 9 and 10. *The major implication is that the momentum transferred after the end of stage II does not contribute to the deformation of the panel*. This momentum is transmitted directly to the supports and induces only elastic reverberations in the panel. This finding, in conjunction with the new FSI model, provides a firm basis for predicting the deflections and deformations that occur in stage III, discussed next.

A comparison of the deflections determined numerically with those predicted by (16) has revealed major discrepancies [13]. A deviation is not surprising, given that the deformation modes for the present supports (condition II) differ from the condition I supports used to derive (16). For condition II, most of the deforma-

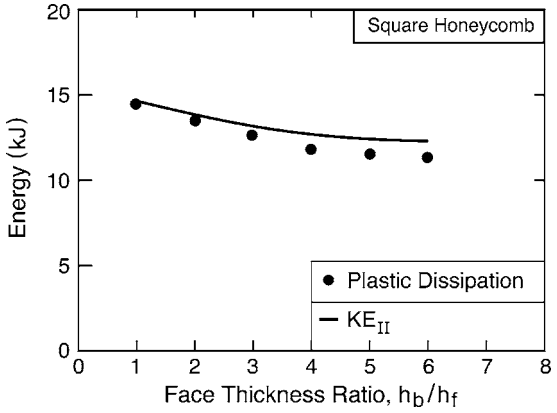


Fig. 26 Comparison of KE_{II} determined using the new analytic model with the total plastic dissipation in stage III calculated numerically. The height of the core is H_c=0.2L and the relative density of the core is $\bar{\rho}$=0.03.

tion occurs through indentation of the back face and the core by the supports (Fig. 8). The sources of the discrepancy and alternative stage III models are discussed elsewhere [13].

9 Concluding Comments

The present assessment has addressed aspects of the response of metallic sandwich structures to underwater blast, emphasizing a domain wherein the water cavitates before the core crushes. It has unearthed several issues affecting the design of high performance panels.

i. Two regimes have been illustrated exhibiting different trends in displacement, plastic strain and reaction force. One regime, designated STC, has characteristics consistent with an analytic three-stage response model, albeit with quantitative discrepancies. The other, designated SOC, appears to exhibit the best performance. The present results predict that I- and corrugated cores having low relative density, incorporating back faces somewhat thicker than the front, are preferred.

ii. The behavior of STC designs has been compared with the predictions of a three-stage analytic model [5], which correctly describes the participating events, but has deficiencies. Foremost among these is an underestimate of the momentum imparted to the system by the end of the second stage, attributed to the previously overlooked characteristics of the cavitated water formed by the reflected pressure wave. A new model that accounts fully for the momentum imparted by the end of stage II rectifies the deficiency and accurately predicts the kinetic energy acquired by the structure and attached water. The caviat is that the model only applies to STC designs: characterized by the occurrence of a common velocity for all constituents at the end of stage II. Soft cores do not acquire a common velocity and it remains to extend the present postulate in such a manner that it encompasses these designs.

iii. For STC designs, the kinetic energy acquired by the end of the second stage accounts fully for the plastic dissipation that occurs in stage III (by bending, stretching, shearing, and indentation of the panel). That is, the extra momentum imparted by the water during this stage does not contribute to the panel deformation. This finding enables the analytically derived kinetic energy at the end of stage II to be used to predict the stage III response. In practice, this opportunity has yet to be realized because the deformation modes differ from those assumed in prior analytical models. Alternative models that rectify this deficiency are explored elsewhere [13].

iv. Panels supported only at the back face develop smaller front face strains than those supported at both front and back faces, rendering this support system less susceptible to front face tearing. The strains in the back face are larger than those at the front, especially at the supports. The strains in both front and back faces are lowest when SOC designs are used, with implications for designing cores giving the best tearing resistance.

One nuance concerning the momentum transfer merits additional comment. As noted here and elsewhere [2–5], the classical Taylor [17] model is entirely satisfactory for solid plates because cavitation initiates at the plate wet surface, the (constant) velocity acquired by the plate at first cavitation exceeds that for the cavitated water, and there is no layer of attached water on the plate. Thus the momentum and kinetic energy solely in the plate upon first cavitation define the subsequent plastic deformation. The situation differs for the wet face of a sandwich panel *even if the core has negligible strength*. In this case, after acquiring its velocity upon initial cavitation, the front face decelerates, because it compresses the core (which still has mass) and accelerates the back face. Now the cavitated water can catch the decelerating front face. Consequently, the classical Taylor model [17] underestimates the mass of water by neglecting the reattachment process that occurs during stage II.

Finally we note that the present results for strong cores differ from those presented by Rabczuk et al. [7], who found the lowest center displacement for their corrugated core panels at the lowest relative density. Since their corrugated panels and ours should behave similarly through stage II (despite the orientation orthogonality), the difference is tentatively attributed to the plastic dissipation in stage III, governed especially by the differences in core shear and indentation caused by orientation.

Acknowledgment

This work was supported by the ONR MURI program on Blast Resistant Structures through a subcontract from Harvard to the University of California, Santa Barbara (Contract No. 123163-03). Also, David R. Hayhurst acknowledges financial support for his Sabbatical leave at the Materials and Mechanical Engineering Departments, University of California, Santa Barbara, provided through a Global Research Award by the Royal Academy of Engineering, UK.

Nomenclature

c_{el} = elastic wave speed in base material
c_w = sound speed in water
E, E_T = Young's modulus and plastic tangent modulus of base material
h_f, h_b = thickness of front face and back face, respectively
H_c = height of the core
H_{eq} = equivalent thickness of the sandwich panel
H_w = height of the water column used in the calculations
I_0 = free field momentum, $I_0 = p_0 t_0$
I_T = transmitted momentum (structure plus water) at the end of stage I
I_{total} = total momentum transmitted to supports
$\bar{I}$ = nondimensional impulse, $\bar{I} = 2I_0 t_0/(m_f H_c)$
$\text{KE}_{\text{constituent}}$ = kinetic energy of a constituent (e.g., front face, core, or back face)
$\text{KE}_I, \text{KE}_{II}$ = total kinetic energy (structure and attached water) at the end of stages I and II, respectively
KE_{II}^* = kinetic energy of the structure (excluding the water) at the end of stage II
l = spacing between core members
l_w = characteristic length of incident pressure pulse in water, $l_w = c_w t_0$
L = half-width of the sandwich beam
M_B, M_F = momentum of the core plus back face, and that of the front face plus attached water at $t = t_c$
M_T = total momentum of the structure and attached water calculated at t_{II}
m_f, m_c, m_b = mass/area of front face, core, and back face
m_w = mass/area of attached water (added mass) at $t = t_c$
$\bar{m}$ = nondimensional mass, $\bar{m} = m_c/m_f$
p = fluid pressure
p_0 = peak pressure of free field impulse
p' = fluid pressure at $t = t_c$
P_{react} = total reaction force at support
t_0 = characteristic time of incident pressure pulse
$t_I = t_c, t_{II}, t_{III}$ = durations of stages I, II, and III, respectively
t_{arrest} = time at which all constituents are arrested in cases with soft cores
t_b = time at which the back face acceleration stops
t_{cb} = time at which the cavitation boundary arrives at location $x = x_{cb}$

v = fluid velocity used in the extended Taylor model
v' = fluid velocity at $t=t_c$
v_b, v_f = average velocities of back face and front face, respectively
v_c = common velocity of structure and attached water in the new FSI model
v_{cb} = rate of translation of the cavitation boundary
v_{common} = common average velocity of all constituents at the end of stage II
v_{face} = velocity of front face at $t=t_c$
v_r = residual velocity of cavitated water used in the new FSI model
$W_{pl}^{\text{constituent}}$ = plastic dissipation in a constituent
W_{pl}^{total} = total plastic dissipation in the beam during stage III
x = distance from front face of the panel, with water residing in $x<0$
x_a = width of the attached water from the new FSI model at the end of stage II
x_c = location of first cavitation in water
x_{cb} = location of the cavitation boundary
β = fluid-structure interaction parameter, $\beta=\rho_w c_w t_0/m_f$
δ_f, δ_b = mid-span deflection of front and back face, respectively
Δ = ratio of face plate thickness, $\Delta=h_b/h_f$
ε_c = core crush strain
$\varepsilon_{pl}^{\max}$ = maximum plastic strain in face plates
ε_Y = yield strain of base material
ζ, ξ = dimensionless parameters used for wave propagation in Sec. 2
$\bar{\rho}$ = relative density of the core
ρ, ρ_w = density of base metal and that of water, respectively
σ_Y = yield strength of base material
σ_{YD}^c = dynamic yield strength of the core.

Appendix A: A Synopsis of Numerical Simulations for Square Panels

The numerical model described in Sec. 5 has been used to simulate square panels for each of the core topologies. For the simulations, the panels are supported along the entire perimeter around the back face. An example of a simulation for a strong I-core, relative to that for a beam is presented in Fig. 27: the deformed shapes at the bottom (beam) and top (panel) have been obtained using separate simulations. Note that the central areas of the plates experience about the same deflections as the mid-span areas of the beams. Moreover, the buckling modes of the core members are identical. This result is typical of many such comparative simulations.

Appendix B: Analytic Approximations

Many numerical simulations have been carried out for a wide parameter range ($0<\sigma_{YD}^c/p_0<0.35$ and $0<\beta<8$) encompassing domain II. The results have been fitted into analytical forms that can be used with Eqs. (24) and (25) to determine the final momentum for any sandwich design within this domain. The first cavitation plane is located at

$$\bar{x}_c \equiv \frac{x_c}{c_w t_0} \approx 4.8\left(\frac{\sigma_{YD}^c}{p_0}\right)^{1.1}(0.24-0.015\beta+0.0012\beta^2). \quad \text{(B1)}$$

An exponential function is most applicable for the reference field:

$$\bar{v}_r \equiv \frac{v_r}{p_0/(\rho_w c_w)} = A\exp(\bar{x}/T), \quad \text{(B2)}$$

with

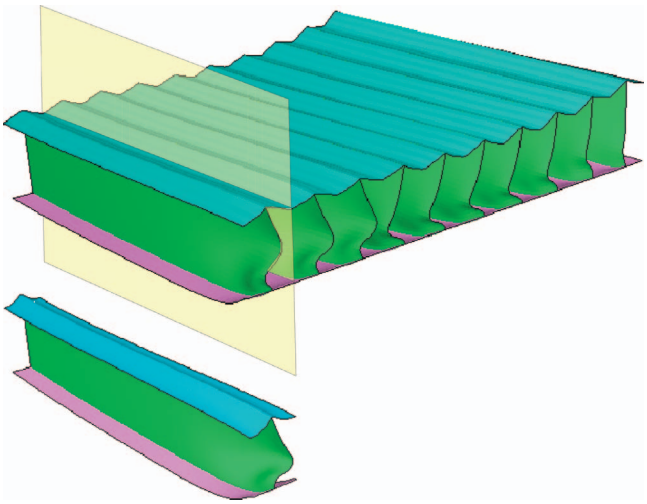


Fig. 27 Deformed shape of the square I-core plate (top picture, $\frac{1}{4}$ of the plate shown) and the I-core beam (bottom picture, $\frac{1}{2}$ of the beam shown)

$$A \approx (0.08+2.2\beta)f(\beta)\left[1-0.28\frac{\sigma_{YD}^c}{p_0}-0.18\left(\frac{\sigma_{YD}^c}{p_0}\right)^2\right],$$

$$T \approx (0.63-0.028\beta+0.003\beta^2)\left(1-0.047\frac{\sigma_{YD}^c}{p_0}\right),$$

where $f(\beta)=\beta^{\beta/(1-\beta)}$. Since, at location $\bar{x}=\bar{x}_a$, $\bar{v}_c=\bar{v}_r(\bar{x}_a)$, therefore

$$\bar{x}_a = T(\ln\bar{v}_c - \ln A) \quad \text{(B3)}$$

With Eqs. (B1) to (B3), Eq. (26) can be rewritten as

$$\bar{v}_c = \frac{I_T + A\cdot T\cdot\exp(\bar{x}_c/T)}{\dfrac{m_f+m_c+m_b}{\rho_w c_w t_0} - T(\ln\bar{v}_c - \ln A - 1)}, \quad \text{(B4)}$$

which can be solved for $\bar{v}_c$ and thus $\bar{x}_a$.

References

[1] Xue, Z., and Hutchinson, J. W., 2003, "Preliminary Assessment of Sandwich Plates Subject to Blast Loads," Int. J. Mech. Sci., **45**, pp. 687–705.
[2] Xue, Z., and Hutchinson, J. W., 2004, "A Comparative Study of Impulse-Resistant Metallic Sandwich Plates," Int. J. Impact Eng., **30**, pp. 1283–1305.
[3] Fleck, N. A., and Deshpande, V. S., 2004, "The Resistance of Clamped Sandwich Beams to Shock Loading," ASME J. Appl. Mech., **71**, pp. 386–401.
[4] Deshpande, V. S., and Fleck, N. A., 2005, "One-Dimensional Shock Response of Sandwich Plates," J. Mech. Phys. Solids, **53**, pp. 2347–2383.
[5] Hutchinson, J. W., and Xue, Z., 2005, "Metal Sandwich Plates Optimized for Pressure Impulses," Int. J. Mech. Sci., **47**, pp. 545–569.
[6] Qiu, X., Deshpande, V. S., and Fleck, N. A., 2003, "Finite Element Analysis of the Dynamic Response of Clamped Sandwich Beams Subject to Shock Loading," Eur. J. Mech. A/Solids, **22**, pp. 801–814.
[7] Rabczuk, T., Samaniego, E., and Belytschko, T., 2005, "Simplified Model for Predicting Impulse Loads on Submerged Structures to Account for Fluid-Structure Interaction," Int. J. Impact Engng., in press.
[8] Xue, Z., and Hutchinson, J. W., 2005, "Crush Dynamics of Square Honeycomb Sandwich Cores," Int. J. Numer. Methods Eng., **65**, pp. 2221–2245.

[9] Cole, R. H., 1948, *Underwater Explosions*, Princeton University Press, Princeton, NJ.
[10] Hunter, K. S., and Geers, T. L., 2004, "Pressure and Velocity Fields Produced by an Underwater Explosion," J. Acoust. Soc. Am., **115**, pp. 1483–1496.
[11] Shin, Y. S., 2004, "Ship Shock Modeling and Simulation for Far-Field Underwater Explosion," Comput. Struct., **82**, pp. 2211–2219.
[12] Klaseboer, E., Hung, K. C., Wang, C., Wang, C. W., Khoo, B. C., Boyce, P., Debono, S., and Charlier, H., 2005, "Experimental and Numerical Investigation of the Dynamics of an Underwater Explosion Bubble Near a Resilient/Rigid Structure," J. Fluid Mech., **537**, pp. 387–413.
[13] Hayhurst, D. R., Flores, S., McMeeking, R. M., and Evans, A. G., 2005, "The Response of Metallic Panels to Water Blast: Deflections," submitted for publication.
[14] Deshpande, V. S., and Fleck, N. A., 2005, private communication.
[15] Ashby, M. F., Evans, A. G., Fleck, N. A., Gibson, L. J., Hutchinson, J. W., and Wadley, H. N. G., 2000, *Metal Foams: A Design Guide*, Butterworth-Heinemann, Boston.
[16] ABAQUS/Explicit User's Manual, Version 6.4. Hibbitt, Karlsson and Sorensen Inc., 2003.
[17] Taylor, G. I., 1963, "The Pressure and Impulse of Submarine Explosion Waves on Plates," in *The Scientific Papers of G.I. Taylor*, Vol. III, Cambridge University Press, Cambridge, UK, pp. 287–303.

A Mathematical Model for Frictional Elastic-Plastic Sphere-on-Flat Contacts at Sliding Incipient

L. Chang

H. Zhang

Department of Mechanical and Nuclear Engineering,
The Pennsylvania State University,
University Park, PA 16802

This paper presents a mathematical model for frictional elastic-plastic sphere-on-flat contacts at sliding incipient. The model is developed based on theoretical work on contact mechanics in conjunction with finite-element results. It incorporates the effects of friction loading on the contact pressure, the mode of deformation, and the area of contact. The shear strength of the contact interface is, in this paper, assumed to be proportional to the contact pressure with a limiting value that is below the bulk shear strength of the sphere. Other plausible interfacial-shear-strength characteristics may also be implemented into the contact model in a similar manner. The model is used to analyze the frictional behavior of a sphere-on-flat contact where the experimental data suggest that the interfacial shear strength is similar in nature to the one implemented in the model. The theoretical results are consistent with the experimental data in all key aspects. This sphere-on-flat contact model may be used as a building block to develop an asperity-based contact model of rough surfaces with friction loading. It may also serve in the modeling of boundary-lubricated sliding contacts where the interfacial shear strength in each micro-contact is coupled with its flash temperature and related to the lubricant/surface physical-chemical behavior. [DOI: 10.1115/1.2178838]

Introduction

Levinson et al. [1] have carried out an experimental study to evaluate the contact and friction theories presented in Refs. [2–4]. In the experiments, a copper-rod specimen with a spherical end is brought into contact with either a steel or sapphire flat surface. The copper is significantly softer than the steel or sapphire with negligible strain hardening so that the contact system in the experiment is consistent with the theoretical models of an elastic-perfectly-plastic sphere against a rigid flat. Furthermore, the surfaces are well polished to comply with another assumption of perfectly smooth contact.

The specimens are thoroughly cleaned and the experiments are carried out under ambient conditions. The friction force in the contact is measured and recorded at the sliding inception, along with the applied normal load. The experiments cover a wide range of loading conditions from elastic contact to contact with significant plastic deformation. The friction-normal-load results are nondimensionalized according to the theory in [4], which then exhibit a fairly similar behavior independent of the spherical diameters of the copper and the copper-steel or copper-sapphire contact systems. Figure 7 of Ref. [1] shows that the static coefficient of friction is about 0.3 for an elastic contact (i.e., $P/Pc<1.0$), exhibits a steep reduction to about 0.1 as the normal load increases towards $P/Pc=50$, and levels off to around 0.06 and 0.07 as P/Pc goes beyond 150 at which the contact becomes significantly plastic. The results are shown in Fig. 1, where P is the applied normal load and Pc is the normal load causing initial plastic yielding in the contact with no friction.

While the nondimensionalization makes the results independent of the sizes and materials of the specimens, the frictional behavior derived from the experiments appears to be characteristically different from that predicted by the theory of Kogut and Etsion [4]. A comparison of the theory and experiment is presented in Fig. 8 of Ref. [1], showing a characteristic departure of the friction behavior when the applied normal load is beyond a relatively small value of $P/Pc=9$. The results and the theory-experiment comparison are shown in Fig. 2. A key assumption of the theories finished in [2,4] is that the contact junction cannot support a tangential force when a high normal load brings the contact into significant plastic deformation. The disagreement between the theory and the experiment prompted the authors of [1] to question this key theoretical assumption. It also suggests the need for further theoretical studies of elastic-plastic point contacts with friction loading.

This paper presents a mathematical model of frictional spherical contact at sliding incipient. It is based on the theories of frictional contact advanced by Tabor [5] and Johnson [6] in conjunction with some recent analytical and numerical results of Zhao et al. [7] and Zhang et al. [8]. The effects of friction loading on the contact pressure, the mode of deformation, and the area of contact are modeled. With a plausible description of the interfacial shear-strength characteristics suggested by the experimental data, the model yields results of a frictional sphere-on-flat contact that appear to be consistent with those of the experiments of [1] in every key aspect. The point-contact model developed in this paper may be used as a building block to develop an asperity-based contact model of rough surfaces with friction loading. It may also serve in the modeling of boundary-lubricated sliding contacts where the interfacial shear strength in each micro-contact is coupled with its flash temperature and depends on the lubricant/surface physical-chemical properties. The modeling concepts and techniques are presented next followed by results and discussions.

Modeling

The frictional contact between an elastic-perfectly-plastic sphere and a rigid flat at sliding inception is modeled. The contact surfaces are assumed to be perfectly smooth. The sphere deforms elastically at a small normal approach between the sphere and the

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 27, 2005; final manuscript received December 9, 2005. Review conducted by A. Maniatty. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

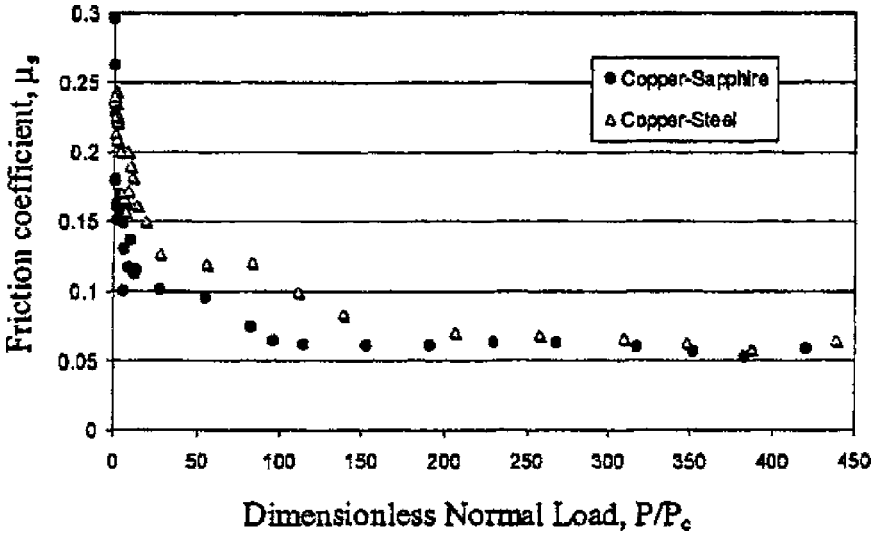


Fig. 1 Experimental results shown in Fig. 7 of Ref. [1]

rigid flat. As the normal approach is increased beyond a first critical value, the contact becomes elastic-plastic. As the normal approach reaches a second critical value, it becomes fully plastic. The values of the two critical normal approaches depend on the friction force in the contact. Figure 3 shows the finite-element results [8] of the critical normal approaches as functions of the friction coefficient in the contact between an elastic-perfectly-plastic cylinder and a rigid flat. The modeling presented below includes four components: the interfacial shear strength, the critical normal approaches, the contact pressure, and the area of contact.

Interfacial Shear Strength. The friction force that can be developed at the sliding inception depends on the shear strength of the contact interface. While the shear strength characteristics can be complex, experiments [9,10] and fundamental analyses [11] have shown that the interfacial shear strength is approximately proportional to the contact pressure with various surface-film materials. Furthermore, Johnson ([6], p. 236) states that, due to contamination or lubrication, the maximum shear stress that the interface can sustain is less than that of the bulk solid. In this paper, the interfacial shear strength is assumed to be proportional to the contact pressure until it reaches a limiting value that is below the shear strength of the sphere bulk. It is given by the following equation:

$$\tau = \begin{cases} cp & cp < \tau_m \\ \tau_m & cp \geqslant \tau_m \end{cases} \tag{1}$$

where τ is the interfacial shear strength of the contact, p is the contact pressure, and τ_m is the limiting value of the interfacial shear strength. Two dimensionless input parameters are used to calculate the friction force with a given applied normal load. One is the shear-strength-pressure proportionality constant, c, and the other is the ratio of the limiting interfacial shear strength to the shear strength of the solid bulk, $\bar{\tau}_m$. The static coefficient of friction is defined as the ratio of the friction force and the normal force at sliding incipient of the contact. In this paper, the average pressure and shear stress in the contact junction are used as the modeling variables. The friction coefficient may then be expressed as:

$$\mu = \tau/p \tag{2}$$

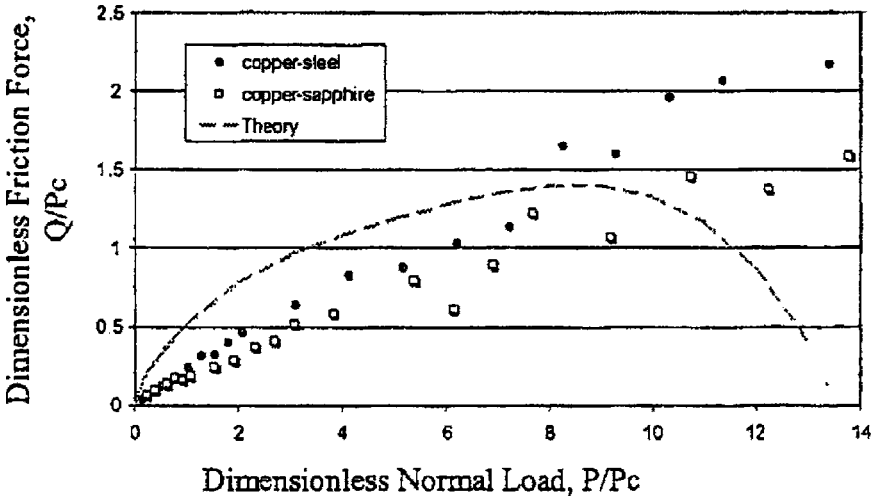


Fig. 2 Experimental and theoretical results shown in Fig. 8 of Ref. [1] (the theoretical curve is generated by the model of Ref. [4])

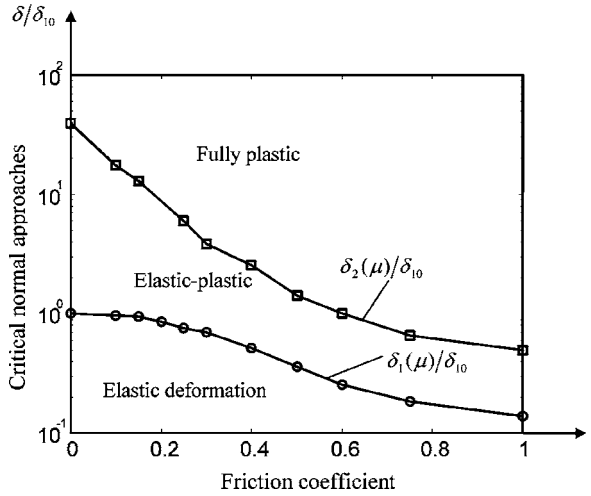


Fig. 3 Effects of friction on the critical normal approaches and modes of deformation in an elastic-perfectly-plastic cylinder-on-flat contact

It should be pointed out that other plausible interfacial-shear-strength characteristics can be implemented as well under the modeling principle presented in this paper. Of particular interest is the boundary-lubricated sliding contact where the interfacial shear strength depends on the contact temperature and the lubricant/surface physical-chemical properties.

Critical Normal Approaches. The first and second critical normal approaches divide the contact deformation into three modes: elastic, elastic-plastic, and fully plastic. Referring to Fig. 3, both of them decrease as the friction coefficient increases. The first critical normal approach, $\delta_1(\mu)$, corresponds to the initial yielding of the spherical contact. For a frictionless contact with the von Mises' yielding criterion, it is given by [6]

$$\delta_1(\mu)|_{\mu=0} = \delta_{10} = \left(\frac{3k_Y\pi}{4}\right)^2\left(\frac{Y}{E}\right)^2 R \tag{3}$$

where R is the radius of the sphere, Y and E are the uni-axial tensile strength and Young's modulus, and k_y is a constant with a typical value of 1.067. For a frictional contact and following the work of Hamilton [12], it can be shown that the second stress invariant, J_2, at any point in the contacting sphere is proportional to the normal approach and is a certain function of the friction coefficient. For the state of initial yielding in the contact, it may be expressed as:

$$J_2 = \frac{Y^2}{3} = J_{2L\,\max}(\mu)\left(\frac{4E}{3\pi}\right)^2\frac{\delta_1(\mu)}{R} \tag{4}$$

where $J_{2L_{\max}}(\mu)$ is a dimensionless function of the friction coefficient at the location in the sphere where J_2 reaches its maximum. Dividing Eq. (4) by its frictionless counterpart and rearranging yields

$$\delta_1'(\mu) = \frac{\delta_1(\mu)}{\delta_1(0)} = \frac{J_{2L\,\max}(0)}{J_{2L\,\max}(\mu)} \tag{5}$$

Equation (5) may be evaluated numerically. With a Poisson's ratio of $\nu=0.3$, the results are best fitted by the following equation:

$$\log[\delta_1'(\mu)] = 3.80\mu^3 - 6.51\mu^2 + 0.67\mu - 0.01 \tag{6}$$

The results are also very similar to the finite-element results obtained for a cylinder-on-flat contact shown in Fig. 3.

The second critical normal approach $\delta_2(\mu)$ defines the onset of fully plastic deformation. For a frictionless contact, Johnson ([6], p. 179) shows that the load causing the spherical contact to become fully plastic is about 400 times the load causing the initial yielding. This relation leads to the following expression for the frictionless second critical normal approach:

$$\delta_2(\mu)|_{\mu=0} = \delta_{20} = 800\left(\frac{Y}{E}\right)^2 R \tag{7}$$

For a frictional contact, the second critical approach is determined based on the results that the theoretically determined $\delta_1'(\mu)$ is closely matched by the finite-element results for a cylindrical contact. It is assumed that the normalized second critical approach, $\delta_2'(\mu)=\delta_2(\mu)/\delta_{20}$, is also similar to that obtained from the finite-element calculation. An approximate expression can then be derived for $\delta_2'(\mu)$ by curve-fitting the finite-element results of [8] and is given by

$$\log[\delta_2'(\mu)] = 4.40\mu^2 - 8.83\mu + 0.03 \tag{8}$$

Contact Pressure. The contact pressure may be significantly reduced by the friction loading, especially in a plastic contact [5,6]. Expressions for the average pressure in a spherical contact are derived in this section as functions of the contact normal approach and friction coefficient, $p(\delta,\mu)$. Consider an elastic contact, which corresponds to a normal approach less than $\delta_1(\mu)$. The friction force generally has small effects on the contact pressure [6]. Therefore, it is assumed to be only dependent on the normal approach and is given by [6]

$$p(\delta,\mu) = \frac{4E}{3\pi}\left(\frac{\delta}{R}\right)^{1/2} \quad \delta \leqslant \delta_1(\mu) \tag{9}$$

When δ is increased beyond the second critical normal approach, $\delta_2(\mu)$, plastic flow occurs. For a frictionless contact, the contact pressure at $\delta=\delta_2(\mu)|_{\mu=0}$ or δ_{20} reaches its maximum possible value or the indentation hardness of the sphere, H. Thus, the frictionless contact pressure for $\delta \geqslant \delta_{20}$ can be written as:

$$p(\delta,\mu)|_{\mu=0} = H \quad \delta \geqslant \delta_{20} \tag{10}$$

With friction loading, the pressure that can be developed in a fully plastic contact is dependent on the interfacial shear strength. The pressure and the shear strength may be related by the Tabor equation [5]:

$$p^2 + \alpha\tau^2 = H^2 \quad \delta \geqslant \delta_2(\mu) \tag{11}$$

where α is a constant. Combining this equation with Eq. (2) yields a general expression for the pressure in a fully plastic contact

$$p(\delta,\mu) = \frac{H}{(1+\alpha\mu^2)^{1/2}} \quad \delta \geqslant \delta_2(\mu) \tag{12}$$

With the contact pressure determined for both $\delta \leqslant \delta_1(\mu)$ and $\delta \geqslant \delta_2(\mu)$, a pressure expression can be obtained for a contact in elastic-plastic deformation. For a frictionless elastic-plastic contact, Francis [13] characterizes the pressure as a logarithmic function of the normal approach. Based on that, Zhao et al. [7] derive the following expression for the pressure:

$$p(\delta,\mu)|_{\mu=0} = p_{Y0} + (H - p_{Y0})\frac{\ln\delta - \ln\delta_{10}}{\ln\delta_{20} - \ln\delta_{10}} \quad \delta_{10} < \delta < \delta_{20} \tag{13}$$

where p_{Y0} is the pressure at the initial yielding. It is assumed that the logarithmic relation also holds when friction loading is applied. Equation (13) may then be generalized to calculate the pressure of a frictional contact in the elastic-plastic regime. For a given normal approach and friction coefficient, the pressure expression is given by

$$p(\delta,\mu) = p_Y(\mu) + [p_F(\mu) - p_Y(\mu)]\frac{\ln\delta - \ln\delta_1(\mu)}{\ln\delta_2(\mu) - \ln\delta_1(\mu)}$$

$$\delta_1(\mu) < \delta < \delta_2(\mu) \tag{14}$$

where $p_Y(\mu)$ is the pressure at $\delta=\delta_1(\mu)$ calculated with Eq. (9) and $p_F(\mu)$ is the pressure at $\delta=\delta_2(\mu)$ determined by Eq. (12).

Area of Contact. Expressions to calculate the area of a spherical contact are derived in this section as functions of the normal approach and friction coefficient. For an elastic, frictionless contact, the area is given by [6]

$$A(\delta,\mu)|_{\mu=0} = \pi R\delta \quad \delta \leqslant \delta_{10} \tag{15}$$

where R is the radius of the sphere. For a fully plastic, frictionless contact, it may be calculated by [14]

$$A(\delta,\mu)|_{\mu=0} = 2\pi R\delta \quad \delta \geqslant \delta_{20} \tag{16}$$

For an elastic-plastic contact, Zhao et al. [7] and Jeng and Wang [15] model the area of contact using a polynomial function that smoothly joins Eqs. (15) and (16). The resulting expression is given by

$$A(\delta,\mu)|_{\mu=0} = (1+3\bar{\delta}^2-2\bar{\delta}^3)\pi R\delta \quad \delta_{10} < \delta < \delta_{20} \tag{17}$$

where $\bar{\delta}=(\delta-\delta_{10})/(\delta_{20}-\delta_{10})$.

The area of contact with friction loading is now modeled. Based on experimental and theoretical studies [5,6], friction may cause the growth of the contact junction. The amount of junction growth depends on the interfacial shear stress and the mode of deformation. Thus, the area of contact may be expressed as the frictionless area multiplied by a junction-growth factor that is a function of both the normal approach and the friction coefficient:

$$A(\delta,\mu) = k_A(\delta,\mu)A(\delta,0) \tag{18}$$

A model for $k_A(\delta,\mu)$ is developed below. For an elastic contact, the area, similar to the pressure, is assumed to be unaffected by the tangential force. Furthermore, the junction growth is absent with $\mu=0$. Therefore,

$$k_A(\delta,\mu) = 1.0 \quad \delta \leqslant \delta_1(\mu) \text{ or } \mu = 0 \tag{19}$$

Next, for a fully plastic contact with a given friction coefficient, the pressure and interfacial shear stress no longer change as δ further increases beyond $\delta_2(\mu)$. Therefore, it is assumed that $k_A(\delta,\mu)$ also reaches an upper bound at $\delta=\delta_2(\mu)$:

$$k_A(\delta,\mu) = k_{Al}(\mu) \quad \delta \geqslant \delta_2(\mu) \tag{20}$$

In the range $\delta_1(\mu)<\delta<\delta_2(\mu)$ of the normal approach with a given value of $\mu=\tau/p$, the shear stress in the contact would be given by an expression similar to Eq. (14) for the pressure. Since the junction growth is induced by the shear stress, $k_A(\delta,\mu)$ may be approximated by an equation comparable to Eq. (14). It is given by

$$k_A(\delta,\mu) = 1 + (k_{Al}(\mu)-1)\frac{\ln\delta - \ln\delta_1(\mu)}{\ln\delta_2(\mu) - \ln\delta_1(\mu)} \quad \delta_1(\mu) < \delta < \delta_2(\mu) \tag{21}$$

The upper-bound junction growth function, $k_{Al}(\mu)$, defined in Eq. (20) needs to be modeled. This function may be determined by first transforming it into a function of the interfacial shear stress, $k'_{Al}(\tau)$. For a fully plastic contact, Eq. (12) in conjunction with Eq. (2) yields a relation between the shear stress and the friction coefficient:

$$\tau = \frac{\mu H}{(1+\alpha\mu^2)^{1/2}} \quad \delta \geqslant \delta_2(\mu) \tag{22}$$

Now consider a contact that is subjected to both normal and tangential loading and is fully plastic. Under such a condition, the characteristics of the junction growth may be captured by the slip-line field solution of a rigid-perfectly-plastic wedge. As shown by Johnson [16], the tangential force causes the plastic zone to be shifted in the direction of the force and a volume of material to be agglomerated at the leading shoulder of the wedge. A similar shifting and agglomerating process is also revealed in the finite element results [8]. This process is intensified as the shear stress increases and is likely to be the cause of the friction-induced junction growth. As the interfacial shear stress approaches the shear strength of the bulk solid, k, the upper-bound junction-growth function reaches its maximum value, which is estimated next.

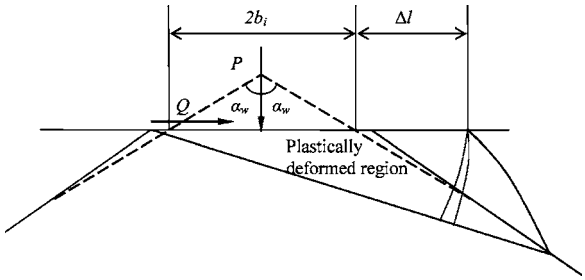


Fig. 4 Schematic of the slip-line field solution of a rigid-perfectly-plastic wedge under combined action of normal and tangential loading

Figure 4 shows a schematic of the slip-line field solution of a rigid-perfectly-plastic wedge with $\tau \approx k$. With such a high interfacial shear stress, the plastic deformation is largely confined to the thin surface layer [8,16]. Consequently, volume conservation requires that the material agglomerated at the leading edge occupies a volume equal to that of the apex segment of the wedge that would have penetrated into the flat surface. The slip-line solution further suggests that the shape of the agglomerated material is similar to that of the penetrated segment of the wedge [16], as illustrated in Fig. 4. Thus, the amount of the junction growth Δl may be approximated by

$$\Delta l = \frac{b_i}{\sin\alpha_w} \tag{23}$$

where b_i is the semi-width of the frictionless contact at the given normal approach of the wedge. The size of contact with friction is then given by

$$l = \left(1 + \frac{1}{2\sin\alpha_w}\right)2b_i \tag{24}$$

The maximum junction-growth factor k_{Al0} is the ratio of l to $2b_i$ and so

$$k_{Al0} = 1 + \frac{1}{2\sin\alpha_w} \tag{25}$$

For a relatively large semi-wedge angle, α_w, corresponding to a cylindrical contact, Eq. (25) yields $k_{Al0} \approx 1.5$. A value of $k_{Al0}=1.4$ is chosen in this study to model the junction growth of a spherical contact. The choice is based on the above order-of-magnitude analysis in conjunction with the consideration that the friction loading should not increase the load capacity of the point contact.

For a fully plastic contact, the upper-bound junction growth function, $k'_{Al}(\tau)$ or $k_{Al}(\mu)$, increases from unity to k_{Al0} as the interfacial shear stress, τ, increases from zero to k. This increase may be divided into two stages based on the analysis of the junction growth by Kayaba and Kato [17] and the finite-element results in [8]. In the first stage, the junction growth is mild before the shear stress reaches a high value of $\tau=0.8-0.9k$. In the second stage of $\tau \to k$, it increases rapidly to reach the maximum value of k_{Al0}. Therefore, the following piecewise linear function is used to model $k'_{Al}(\tau)$:

$$k'_{Al}(\tau)=\begin{cases}1+(k_{Alc}-1)\dfrac{\tau}{\tau_c} & 0\leqslant\tau<\tau_c\\ k_{Alc}+(k_{Al0}-k_{Alc})\dfrac{\tau-\tau_c}{k-\tau_c} & \tau\geqslant\tau_c\end{cases}\tag{26}$$

In this paper, $k_{Alc}=1.1$ and $\tau_c=0.85k$ are used to describe the mild junction growth in the first stage. Finally, transforming $k'_{Al}(\tau)$ into $k_{Al}(\mu)$ using Eq. (22) yields

$$k_{Al}(\mu)=\begin{cases}1+(k_{Alc}-1)\dfrac{\mu H}{\tau_c(1+\alpha\mu^2)^{1/2}} & 0\leqslant\mu<\mu_c\\ k_{Alc}+(k_{Al0}-k_{Alc})\dfrac{\mu H-\tau_c(1+\alpha\mu^2)^{1/2}}{(k-\tau_c)(1+\alpha\mu^2)^{1/2}} & \mu\geqslant\mu_c\end{cases}\tag{27}$$

where μ_c is related to τ_c by Eq. (22) to give

$$\mu_c=\left[\left(\frac{H}{\tau_c}\right)^2-\alpha\right]^{-1/2}\tag{28}$$

Equations (18), (21), and (27) form a complete set to model the contact area and the friction-induced junction growth of a spherical contact.

Summary and Nondimensionalization. The key equations of the four components of the model are summarized below:

Interfacial shear strength

$$\tau=\begin{cases}cp & cp<\tau_m\\ \tau_m & cp\geqslant\tau_m\end{cases}\tag{1'}$$

$$\mu=\pi/p\tag{2'}$$

Critical normal approaches

$$\delta_1(\mu)|_{\mu=0}=\delta_{10}=\left(\frac{3k_Y\pi}{4}\right)^2\left(\frac{Y}{E}\right)^2R\tag{3'}$$

$$\log[\delta'_1(\mu)]=3.80\mu^3-6.51\mu^2+0.67\mu-0.01\tag{6'}$$

$$\delta_2(\mu)|_{\mu=0}=\delta_{20}=800\left(\frac{Y}{E}\right)^2R\tag{7'}$$

$$\log[\delta'_2(\mu)]=4.40\mu^2-8.83\mu+0.03\tag{8'}$$

Contact pressure

$$p(\delta,\mu)=\begin{cases}\dfrac{4E}{3\pi}\left(\dfrac{\delta}{R}\right)^{1/2} & \delta\leqslant\delta_1(\mu)\\ p_Y(\mu)+[p_F(\mu)-p_Y(\mu)]\dfrac{\ln\delta-\ln\delta_1(\mu)}{\ln\delta_2(\mu)-\ln\delta_1(\mu)} & \delta_1(\mu)<\delta<\delta_2(\mu)\\ \dfrac{H}{(1+\alpha\mu^2)^{1/2}} & \delta\geqslant\delta_2(\mu)\end{cases}\tag{29}$$

Area of contact

$$A(\delta,\mu)=k_A(\delta,\mu)A(\delta,0)\tag{18'}$$

$$A(\delta,0)=A(\delta,\mu)|_{\mu=0}=\begin{cases}\pi R\delta & \delta\leqslant\delta_{10}\\ (1+3\bar{\delta}^2-2\bar{\delta}^3)\pi R\delta & \delta_{10}<\delta<\delta_{20}\\ 2\pi R\delta & \delta\geqslant\delta_{20}\end{cases}\tag{30}$$

$$k_A(\delta,\mu)=\begin{cases}1.0 & \delta\leqslant\delta_1(\mu)\\ 1+[k_{Al}(\mu)-1]\dfrac{\ln\delta-\ln\delta_1(\mu)}{\ln\delta_2(\mu)-\ln\delta_1(\mu)} & \delta_1(\mu)<\delta<\delta_2(\mu)\\ k_{Al}(\mu) & \delta\geqslant\delta_2(\mu)\end{cases}\tag{31}$$

For most metals obeying the von Mises' yield criterion, the tensile strength, Y, the shear strength, k, and the indentation hardness, H, of the material may be related by $Y=H/3$ and $k=Y/\sqrt{3}$ [6]. Furthermore, the Poisson's ratio may be assumed to be $\nu=0.3$. Then the model equations summarized above along with other key equations can all be made dimensionless and independent of the radius and basic material constants of the sphere such as the Young's modulus and tensile strength. The following normalizations are carried out. The normal approach is normalized by $R(Y/E)^2$. The contact pressure and the interfacial shear stress are normalized by Y. The area of contact is normalized by $R^2(Y/E)^2$. And the normal and friction forces are normalized by $YR^2(Y/E)^2$.

Tabor's constant α in Eq. (11) may be estimated by considering an extreme situation. Under high vacuum conditions, a high friction coefficient of the order of 10 or higher is obtained for clean metal surfaces [5,18]. In this case, the shear stress approaches the bulk shear strength and the shear flow is observed. As a result, the real area of contact increases substantially and the pressure is greatly reduced. In the extreme, Eq. (11) yields $\alpha\approx(H/k)^2=27$.

The model equations are solved iteratively with a numerical method. The input is the normal approach of the sphere to the rigid flat. The solution yields a normal force, a friction force, and thus a friction coefficient corresponding to sliding incipient of the contact.

Results and Discussions

The model developed in this paper is used to further study the frictional behavior of the sphere-on-flat contact problems of [1]. Dimensionless results are obtained and presented in the same format as the experimental results for easy comparison and analysis.

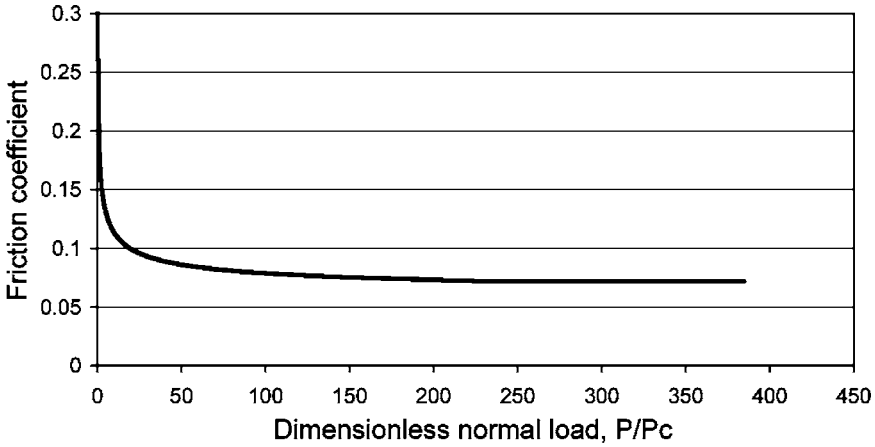


Fig. 5 Theoretical results corresponding to the experimental data of Fig. 1 (c=0.3 and $\bar{\tau}_m$=0.35)

Values for the two dimensionless parameters, c and $\bar{\tau}_m$, which characterize the interfacial shear strength with Eq. (1), need to be chosen for the calculation. That the friction coefficient shown in Fig. 1 starts around 0.3 at a low normal load suggests that a sensible value for the shear-strength-pressure proportionality constant is $c \approx 0.3$. The value for the ratio of the limiting interfacial shear strength to the bulk shear strength is determined by matching the calculated friction coefficient to that in the experiment at a very high normal load. This determination yields a value of $\bar{\tau}_m \approx 0.35$.

Figure 5 shows the friction coefficient as a function of the dimensionless normal load, P/Pc. The results are shown to be consistent with the experimental data of Fig. 1 for both the trend and the level. The results are rearranged to show a dimensionless friction force, Q/Pc, against the dimensionless normal load, P/Pc, for small values of P/Pc up to 14, as is done in Fig. 8 of [1]. Figure 6 shows this result, which also exhibits a consistent trend and magnitude with the experimental data of Fig. 2. It should be pointed out that the interfacial parameters of c=0.3 and $\bar{\tau}_m$ =0.35 are chosen based on the initial and final values of the measured friction coefficient of the contact. They do not totally dictate the trend of the friction coefficient as the applied normal load increases. This trend can be significantly affected by the contact pressure and the area of contact in addition to the characteristics of the interfacial shear strength. The reasonable match in the trend between the theory and the experiment offers support to the mathematical model, in addition to the underlying contact-mechanics fundamentals.

Levinson et al. [1] have also measured the areas of contact under various conditions to evaluate the theoretical results of Kogut and Etsion [3] and the theory of friction-induced junction growth of Tabor [5]. In particular, they measure the residual contact area after unloading from a high normal load of P/Pc=220, which generates a large amount of contact plastic deformation in the copper specimen. They carry out another area measurement with a new copper specimen, this time unloading from P/Pc =220 *with sliding friction*. The measured areas of contact are essentially the same with or without friction loading, which prompts the authors of [1] to question Tabor's theory of junction growth.

With the values of c=0.3 and $\bar{\tau}_m$=0.35, which are used to obtain the results of Figs. 5 and 6, the areas of contact are also calculated under various conditions and compared with the experimental results of [1]. At the normal load of P/Pc=220, the calculation yields an area of contact normalized by the area of contact at P/Pc=1.0 to be A/Ac=84.4. For the frictionless contact, the calculated area ratio is 85.9. Physically, the former should not be smaller than the latter. The small percentage difference between the two results is evidently due to the modeling imperfection as the contact-area junction growth is enforced in the model for a given normal approach instead of a given normal load. The two area results calculated above essentially suggest that the current model predicts a negligible junction growth under the given normal-loading and friction conditions, consistent with the experimental results of [1]. That the junction growth is negligible with a low friction coefficient is also suggested in Johnson's theory [6]. Using the model of a rigid-perfectly-plastic wedge against a rigid

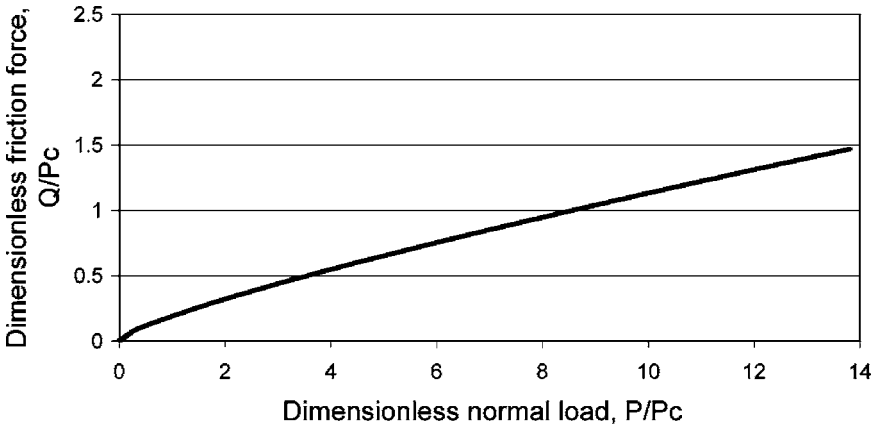


Fig. 6 Theoretical results corresponding to the experimental data of Fig. 2 (c=0.3 and $\bar{\tau}_m$=0.35)

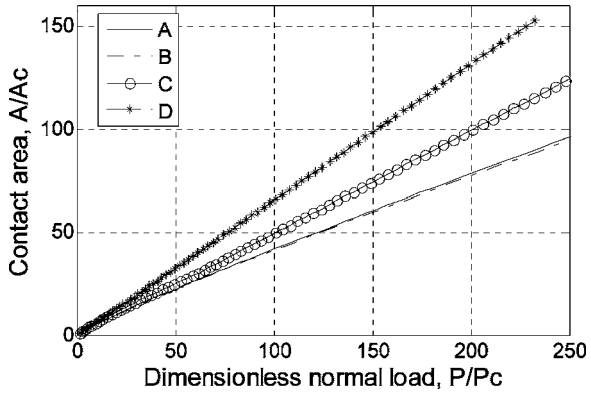


Fig. 7 Contact area and friction-induced junction growth with $c=0.3$ (A—frictionless, B—$\bar{\tau}_m=0.35$, C—$\bar{\tau}_m=0.7$, and D—$\bar{\tau}_m=1.0$)

flat, Johnson ([6], Fig. 7.16) shows that the junction growth is only a few percent if the friction coefficient is in the 0.06–0.07 range.

It may be premature to draw a conclusion that friction generally does not cause a junction growth based on limited experimental data or theoretical results. Figure 7 presents the contact-area results calculated under a wide range of normal-loading and friction conditions. The areas obtained with $c=0.3$ and $\bar{\tau}_m=0.35$ are basically identical to those obtained for the frictionless contact. Furthermore, the results match very well with the experimental data shown in Fig. 4 of [1]. However, the theoretical analysis predicts a significant junction growth with a high level of contact friction. At $P/Pc=200$, for example, the friction-induced junction growth is predicted to be about 25% with $\bar{\tau}_m=0.7$ and 65% with $\bar{\tau}_m=1.0$. Large friction-induced junction growths were experimentally observed with chemically clean metallic surfaces [19,20], with which both the shear-strength-pressure slope (i.e., c) and the limiting interfacial shear strength in the contact can be of high values.

The results presented in this paper are relatively brief, focusing on comparison with some experimental results and analysis of the observed frictional behavior. Nevertheless, the results and their agreement with single point-contact experimental data lend further credence to the model in addition to the theoretical foundation of [5,6] on which it is based. This sphere-on-flat contact model may serve as a key building block to develop a contact model of engineering surfaces with friction loading. One approach to model nominally flat rough-surface contacts is to envision the roughness into spherical-tipped asperities of some statistical height distribution, which is pioneered by Greenwood and Williamson [21] and further advanced by many others. The majority of asperity-based models, to the authors' knowledge, have been for frictionless contacts. Thus, it would be an original contribution to the modeling of rough-surface contacts to incorporate friction loading into the model. The frictional sphere-on-flat model developed in this paper can be readily transported into the surface model at the asperity contacts. In addition, the interfacial-shear-strength relation described in this paper may be easily replaced by other plausible shear-strength characterizations so that a larger class of the contact problems may be modeled. An example is the contact of metallic surfaces with boundary lubrication where the shear strength in an asperity micro-contact is significantly influenced by its flash temperature and the lubricant/surface physical-chemical behavior. Work in this direction is in progress.

Conclusion

This paper presents a mathematical model for frictional elastic-plastic sphere-on-flat contacts at sliding incipient. The model is developed based on theoretical work on contact mechanics in conjunction with finite-element results. It incorporates the effects of friction loading on the contact pressure, the mode of deformation, and the area of contact. The shear strength of the contact interface is assumed to be proportional to the contact pressure with a limiting value that is below the bulk shear strength of the sphere. Other plausible shear-strength characteristics can be equally implemented into the contact model in a similar manner. The model is used to analyze the frictional behavior of a sphere-on-flat contact studied in Refs. [1,22]. The theoretical results are consistent with the experimental data in all key aspects. The analysis reveals that relatively low limiting shear strength of the contact interface is likely the key factor leading to the steep reduction in the measured friction coefficient as the applied normal load is increased. This low limiting shear strength is also likely to be the main reason for the negligible junction growth observed in the experiments. This sphere-on-flat contact model may be used as a building block to develop an asperity-based contact model of rough surfaces with friction loading. It may also serve in the modeling of boundary-lubricated sliding contacts where the interfacial shear strength in each micro-contact is coupled with its flash temperature and related to the lubricant/surface physical-chemical behavior.

References

[1] Levinson, O., Etsion, I., and Halperin, G., 2003, "An Experimental Investigation of Elastic Plastic Contact and Friction of a Sphere on Flat," *Proc. of Contact Mechanics—Friction: Modeling and Experiment*, STLE/ASME Joint International Tribology Conference, pp. 19–23.
[2] Chang, W. R., Etsion, I., and Bogy, D. B., 1988, "Static Friction Coefficient Model for Metallic Rough Surfaces," ASME J. Tribol., **110**, pp. 57–63.
[3] Kogut, L., and Etsion, I., 2002, "Elastic-Plastic Contact Analysis of a Sphere and a Rigid Flat," ASME J. Appl. Mech., **69**, pp. 657–662.
[4] Kogut, L., and Etsion, I., 2003, "A Semi-Analytical Solution for the Sliding Inception of a Spherical Contact," ASME J. Tribol., **125**, pp. 499–506.
[5] Tabor, D., 1959, "Junction Growth in Metallic Friction: The Role of Combined Stresses and Surface Contamination," Proc. R. Soc. London, Ser. A, **251**, pp. 378–393.
[6] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.
[7] Zhao, Y., Maietta, D., and Chang, L., 2000, "An Asperity Micro-Contact Model Incorporating the Transition From Elastic Deformation to Fully Plastic Flow," ASME J. Tribol., **122**, pp. 86–93.
[8] Zhang, H., Chang, L., Webster, M. N., and Jackson, A., 2003, "Effects of Friction on the Contact and Deformation Behavior in Sliding Asperity Contacts," Tribol. Trans., **46**, pp. 514–521.
[9] Briscoe, B. J., Scrutton, B., and Willis, R. F., 1973, "The Shear Strength of Thin Lubricant Films," Proc. R. Soc. London, Ser. A, **333**, pp. 99–114.
[10] Timsit, R. S., and Pelow, C. V., 1992, "Shear Strength and Tribological Properties of Stearic Acid Films," ASME J. Tribol., **114**, pp. 150–166.
[11] Tabor, D., 1982, "The Role of Surface and Intermolecular Forces in Thin Film Lubrication," in *Microscopic Aspects of Adhesion and Lubrication*, J.-M. Georges, ed., Elsevier, Amsterdam, pp. 651–679.
[12] Hamilton, G. M., 1983, "Explicit Equations for the Stresses Beneath a Sliding Spherical Contact," Proc. Inst. Mech. Eng., Part C: Mech. Eng. Sci., **197**, pp. 53–59.
[13] Francis, H. A., 1976, "Phenomenological Analysis of Plastic Spherical Indentation," ASME J. Eng. Mater. Technol., **76**, pp. 272–281.
[14] Abbott, E. J., and Firestone, F. A., 1933, "Specifying Surface Quality—A Method Based on Accurate Measurement and Comparison," Mech. Eng. (Am. Soc. Mech. Eng.), **55**(9), pp. 569–572.
[15] Jeng, Y. R., and Wang, P. Y., 2003, "An Elliptical Microcontact Model Considering Elastic, Elastoplastic, and Plastic Deformation," ASME J. Tribol., **125**, pp. 232–240.
[16] Johnson, K. L., 1968, "Deformation of a Plastic Wedge by a Rigid Flat Die Under the Action of a Tangential Force," J. Mech. Phys. Solids, **16**, pp. 395–340.
[17] Kayaba, T., and Kato, K., 1978, "Theoretical Analysis of Junction Growth," Technol. Rep., Tohoku Univ., **43**, pp. 1–10.
[18] McFadden, C. F., and Gellman, A. J., 1998, "Metallic Friction: The Effect of Molecular Adsorbates," Surf. Sci., **409**, pp. 171–182.
[19] Bowden, F. P., and Young, J. E., 1951, "Friction of Clean Metals and the Influence of Adsorbed Films," Proc. R. Soc. London, Ser. A, **208**, pp. 311–325.
[20] Parker, R. C., and Hatch, D., 1950, "The Static Coefficient of Friction and the Area of Contact," Proc. Phys. Soc. London, Sect. B, **63**, pp. 185–197.
[21] Greenwood, J. A., and Williamson, J. B. P., 1966, "Contact of Nominally Flat Surfaces," Proc. R. Soc. London, Ser. A, **295**, pp. 300–319.
[22] Etsion, I., Levinson, O., Halperin, G., and Varenberg, M., 2005, "Experimental Investigation of the Elastic-Plastic Contact Area and Static Friction of a Sphere on Flat," ASME J. Tribol., **127**, pp. 47–50.

On the Admissibility of Given Acceleration-Dependent Forces in Mechanics

Michael M. Zhechev
Leading Researcher
Department of Vehicle Statistical Dynamics,
Institute of Technical Mechanics of the National
Academy of Sciences of Ukraine
and the National Space Agency of Ukraine,
15 Leshko-Popel Street,
Dniepropetrovsk 49005,
Ukraine
e-mail: zhechev@optima.com.ua

In his book "A Treatise on Analytical Dynamics," Pars asserted that acceleration-dependent forces are inconsistent with one of the fundamental principles of mechanics, namely, with the superposition principle, thus spreading among mechanical scientists the idea that such forces are not admissible in mechanics. This article demonstrates that given forces that depend on acceleration or higher derivatives are admissible in mechanics and shows that this assertion in Pars's book is fallacious and the only condition for the applicability of such forces is the equation of motion possessing a unique solution. [DOI: 10.1115/1.2187528]

1 Problem Statement

Traditionally, given forces in mechanics are considered as functions of position, velocity, and time, i.e., $F=F(x,\dot{x},t)$, where x are generalized coordinates. However, from time to time the question arises as to whether given forces can depend on acceleration as well: $F=F(x,\dot{x},t;\ddot{x})$, i.e., as to whether traditional methods of dynamics hold for acceleration-dependent forces.

In 1964, Pars asserted in his book [1] that acceleration-dependent forces are inconsistent with one of the fundamental principles of mechanics, namely, with the superposition principle. Well-known and widely used as a textbook, that book has spread among mechanical scientists the idea that such forces are not admissible in mechanics. Nevertheless, both theoretical and applied works in which acceleration-dependent forces are used are published too.

Before Pars's book was published, Birkhoff in his well-known book [2] wrote generalized forces as functions that depend not only on position, velocity or time, but on acceleration as well. After Pars's book had been published, works that used such forces were published too. The use of acceleration-dependent forces in Ref. [3] made it possible to explain the secular retardation of the Earth's rotation. There are examples of the use of acceleration-dependent forces in an attitude controller [4] and in space mechanical arm control systems [5].

Thus, on the one hand, acceleration-dependent forces do exist and are used, but on the other hand, it is argued (and proved by Pars) that such forces are not admissible in mechanics. The objective of this article is to analyze this contradiction.

Section 2 cites Pars's proof [1] on the basis of which he inferred the inadmissibility of acceleration-dependent forces in mechanics. Section 3 shows that the admissibility of these forces may be demonstrated using traditional methods of mechanics. Section 4 demonstrates that Pars's assertion is fallacious and shows the causes of this fallacy. Section 5 demonstrates the admissibility of forces that depend not only on acceleration, but on higher derivatives as well provided that the equation of motion possesses a unique solution. In Sec. 6, it is noted that this condition may not be satisfied in the general case. However, this does not suggest the inadmissibility of a particular force; this only points to the necessity of refining the mathematical model by including "finer" factors that affect the motion of the system under consideration or, where control actions are involved, to the fact that this force is physically unrealizable.

2 Pars's Proof

In support of his conclusion that given acceleration-dependent forces are not admissible in mechanics, Pars gives the following proof using the rectilinear motion of a particle of mass m (Fig. 1(a)) as an example [1].

"Consider two forces: $m\varphi(f)$ and $m\psi(f)$ where f is the acceleration, $f=\ddot{x}$. The functions φ and ψ may also involve x, $u=\dot{x}$ and t, but it is the dependence on f that primarily concerns us at the moment and that is emphasized in the notation. We now consider three experiments. In the first experiment the particle is acted on by the force $m\varphi$, in the second by $m\psi$, and in the third by $m(\varphi+\psi)$. The values of x,u,t are the same in all three experiments. If we denote the accelerations in the three experiments by f_1,f_2 and f_3, we have

$$f_1=\varphi(f_1) \tag{1}$$

$$f_2=\psi(f_2) \tag{2}$$

$$f_3=\varphi(f_3)+\psi(f_3) \tag{3}$$

The first point that catches our attention is that an equation such as (1) does not necessarily determine f_1 uniquely, a situation in itself foreign to the Newtonian outlook. However we can afford to ignore this point, because a much more serious difficulty confronts us in a moment. We will assume therefore that f_1,f_2, and f_3 are uniquely determined.

Now it is a fundamental postulate of Newtonian mechanics that when two forces act simultaneously on a particle the effect is the same as that of a single force equal to their (vector) sum. An equivalent form of the same postulate is that each force gives rise to the acceleration that it would produce if the other force were absent. Thus we must have

$$f_3=f_1+f_2 \tag{4}$$

and combining this with Eq. (3) we find

$$f_1+f_2=\varphi(f_1+f_2)+\psi(f_1+f_2) \tag{5}$$

Now it is easy to see that in general Eqs. (1), (2), and (5) are inconsistent. Equation (1) involves only the value of φ at f_1, and Eq. (2) involves only the value of ψ at f_2, and neither involves any reference to or any restriction on the value of φ or the value of ψ at f_1+f_2.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 10, 2004; final manuscript received December 15, 2005. Review conducted by I. Mezic. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

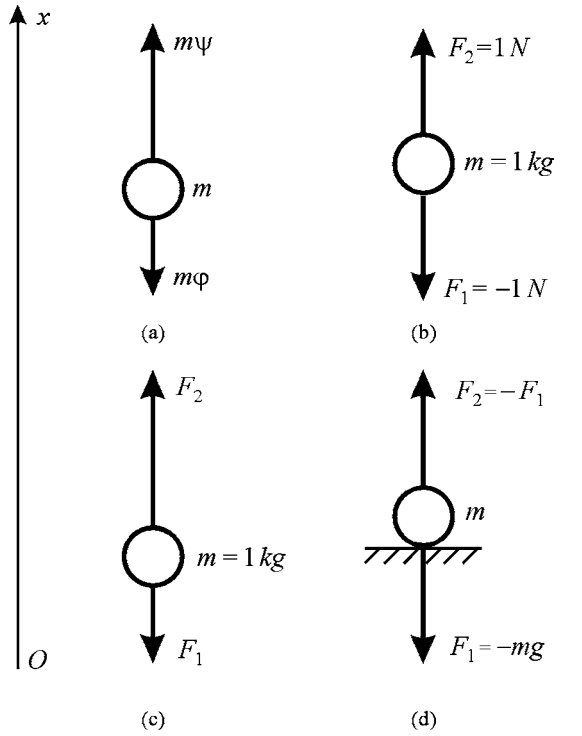


Fig. 1 Particle acted on by two forces; (a) acceleration-dependent forces (illustration to Pars's proof), (b) opposing forces equal in magnitude, (c) interrelated forces ($F_1+F_2=1$N), (d) particle resting on an immovable support

'…Thus forces depending on the acceleration are not admissible in Newtonian dynamics'"

3 Another Approach

If it were not for Pars's proof, the possibility of using acceleration-dependent forces in mechanics could be demonstrated by bringing forward the following rather simple arguments.

As in Pars's proof, consider the rectilinear motion of a particle of mass m on the line Ox, but now the particle is acted on by forces $m\varphi[x,\dot{x},t;f(x,\dot{x},t)]$ and $m\psi[x,\dot{x},t;f(x,\dot{x},t)]$ where φ and ψ are smooth functions of their arguments, and the function $f(x,\dot{x},t)$, in addition to the natural requirement of smoothness, must meet one more requirement: it must be uniquely determined.

Because the forces $m\varphi[x,\dot{x},t;f(x,\dot{x},t)$ and $m\psi[x,\dot{x},t;f(x,\dot{x},t)]$ depend only on the coordinate x, velocity $\dot{x}$ and time t and do not depend on the acceleration, Pars's proof does not hold for this case, and thus we are entitled, in accordance with traditional methods of mechanics, to write the equation of motion of the particle in the traditional form

$$\ddot{x} = \varphi[x,\dot{x},t;f(x,\dot{x},t)] + \psi[x,\dot{x},t;f(x,\dot{x},t)] \tag{6}$$

The equations of motion of the particle under consideration have this form regardless of the form of the function $f(x,\dot{x},t)$. Because of this, if $f(x,\dot{x},t)$ is the solution of the algebraic (in f) equation

$$f = \varphi(x,\dot{x},t;f) + \psi(x,\dot{x},t;f) \tag{7}$$

the motion of the particle is described by Eq. (6) too.

Let Eq. (7) determine f uniquely. (Note that by the implicit function theorem, Eq. (7) determines $f=f(x,\dot{x},t)$ uniquely if $\partial\varphi/\partial f+\partial\psi/\partial f\neq 1$).

It follows from Eqs. (6) and (7) that $f(x,\dot{x},t)\equiv\ddot{x}(x,\dot{x},t)$. Because of this, keeping in mind that the motion of a particle depends on the values of the forces acting thereon rather than on the notation used for these forces or their arguments, one should accept that

$$\ddot{x} = \varphi(x,\dot{x},t;\ddot{x}) + \psi(x,\dot{x},t;\ddot{x}) \tag{8}$$

describes the motion of the particle acted on by the acceleration-dependent forces $m\varphi(x,\dot{x},t;\ddot{x})$ and $m\psi(x,\dot{x},t;\ddot{x})$.

Thus, restricting ourselves to forces and methods traditional for mechanics, we have unambiguously derived the equation of motion of a particle acted on by two acceleration-dependent forces. This means that either the contradiction revealed by Pars is typical of traditional forces and methods as well or Pars's proof is fallacious.

4 Comments on Pars's Proof

The first point that catches our attention in Pars's proof is that he uses two interpretations of one and the same assumption that is sometimes called the *superposition principle*: "each force gives rise to the acceleration that it would produce if the other force were absent." The key phrase in this assumption is "if the other force were absent." The results of the application of this principle depend on how the absence of the other force is understood. Does it mean that

(i) when "the other force" is dropped, the retained force remains unchanged, i.e., its value (direction and magnitude) remains the same as in the presence of the dropped force (as in deriving Eq. (3)), or
(ii) when "the other force" is dropped, the retained force takes the value it would take if the dropped force were actually equal to zero (as in deriving Eq. (4))?

Here, we will show that the contradiction at which Pars arrived is due to the fact that in some cases the two interpretations of the superposition principle contradict each other, and this is true not only for acceleration-dependent forces, but for traditional forces as well.

Consider in more detail the difference between the two interpretations of the superposition principle.

In case (i), no account is taken of how the absence of one force can affect the other; that is, this interpretation only implies the following. *If several forces act simultaneously on a particle, the effect is the same as that of a single force equal to their vector sum.* In other words, it is only the current values of the forces that are of importance, and it makes no difference whether these forces are interdependent or not.

For example, if a particle of unit mass is acted on by two forces, $F_1=-1$N and $F_2=1$N (Fig. 1(*b*)), then, whatever the nature of these forces and no matter whether they may be interdependent or not, one should proceed as follows: to find the acceleration of the particle acted on by the force F_1 alone (-1 m/s^2), then find the acceleration of the particle acted on by the force F_2 alone (1 m/s^2) and add together these accelerations to obtain the acceleration produced by both forces acting simultaneously (0 m/s^2). Clearly with this approach it makes no difference at all whether the force depends on acceleration or not.

In case (ii), one should account for the interrelation between these forces, i.e., account for how the presence or absence of one force affects the other. In this case, in the example considered above we should find out whether there is any dependence between the forces F_1 and F_2, e.g., whether they are related to each other through acceleration, before writing the equation of motion of the particle. If the forces are independent, their current values are added together to give zero acceleration as in case (i). If the forces are interdependent, but this interdependence is unknown to us, the equation of motion of the particle acted on by these forces simultaneously cannot be written in principle. If this interdepen-

dence is known, we should add together the values that each of the forces would take if the other force were equal to zero rather than their current values.

Let us see how the two interpretations of the superposition principle are consistent with each other and with our empirical knowledge of the behavior of mechanical systems.

Let the forces be interrelated in some way (Fig. 1(*c*)). If, for example, $F_1+F_2=1\text{N}$, then, according to universally accepted ideas, the acceleration of the particle acted on by the two forces should be equal to $1\ \text{m/s}^2$. We will obtain the same if we use interpretation (i) of the superposition principle. However, if interpretation (ii) is used, the acceleration will be $2\ \text{m/s}^2$. To demonstrate, if $F_1=0\text{N}$, then $F_2=1\text{N}$, and vice versa.

We may also consider the situation where only F_1 affects F_2 while F_2 does not affect F_1 (Fig. 1(*d*)), for example, the situation where the particle rests on an immovable support, F_1 is the gravity force, and F_2 is the supporting force. In case (i), the acceleration will obviously be zero as it must. In case (ii), however, the situation is quite different. Let the supporting force be zero. Then the acceleration imparted to the particle by the gravity force will be $-g$ (here, g is the acceleration of gravity). Now let the gravity force be zero. In the absence of the gravity force, the supporting force is zero too. Hence, the acceleration imparted to the particle by the supporting force (in the absence of the gravity force) will be zero. Therefore, in this case the acceleration imparted to the particle by the two forces will be $-g$. In other words, interpretation (ii) of the superposition principle implies that a particle resting on an immovable support must execute a uniformly accelerated downward motion.

These examples show that in the case of interdependent forces the two interpretations of the superposition principle result in different motions. In other words, for interdependent forces these interpretations are non-equivalent and, what's more, they conflict with each other, this being true both for acceleration-dependent forces and for forces that depend on time alone and even for constant forces.

The basic fallacy in Pars's proof is that he equates the particle accelerations obtained from the two different interpretations of the superposition principle (Eq. (4)), which in the case of interdependent forces yield radically different results. (The interdependence of the forces considered by Pars [1] is obvious because the value of each of them depends on whether it acts singly as in Eqs. (1) and (2) or together with the other force as in Eq. (3)). On this basis, he concludes that acceleration-dependent forces are not admissible. But in just the same way, using a particle on an immovable support as an example, it could be concluded that the gravity force is not admissible or Newton's third law is not true.

In effect, the only question that could arise at this point is whether interpretation (ii) of the superposition principle in itself is admissible. The examples considered above show that this interpretation is inconsistent with what we know from experience and hence fallacious. Because of this, Pars's conclusions based on this interpretation are fallacious too, and thus they cannot be considered as the demonstration of the inadmissibility of acceleration-dependent forces in mechanics.

At the same time, as shown in Sec. 3, the equations of motion of a particle acted on by acceleration-dependent forces can be obtained from the traditional approach without any additional assumptions. This fact alone makes it possible to conclude that acceleration-dependent forces are every bit as admissible in mechanics as forces that depend only on position, velocity, and time. Let us consider this problem from another point of view.

5 Force as a Function of Time

According to interpretation (i) of the superposition principle, the acceleration of a particle at every instant of time is determined solely by the values of the forces acting thereon. In this case, it makes no difference whether the forces depend on time explicitly or through the mediation of other functions, for example, through the mediation of $x(t)$, $\dot{x}(t)$ and some function $f(t)$. In other words, of importance are the values of the forces at every instant of time rather than the algorithm by which they are calculated. Thus, for example, if a particle of mass m is acted on by forces $m\varphi[x(t),\dot{x}(t),t;f(t)]$ and $m\psi[x(t),\dot{x}(t),t;f(t)]$, then, regardless of the algorithm of calculation of the function $f(t)$, the equation of motion of the particle has the form

$$\ddot{x}(t)=\varphi[x(t),\dot{x}(t),t;f(t)]+\psi[x(t),\dot{x}(t),t;f(t)] \tag{9}$$

Note that the validity of this assertion does not depend by any means on how $f(t)$ is related to $\ddot{x}(t)$: does not coincide at any t, coincides at some instances of time or is identically equal thereto. For example, if $\varphi=1/2$ and $\psi=1/2$, the fact that in this case each of the functions is formally identically equal to $\ddot{x}/2$ can hardly cast any doubt upon the validity of the equation of motion $\ddot{x}=1$.

If, in one sense or another, our interest is in finding the function $f(t)$ that satisfies the condition $f(t)\equiv\ddot{x}(t)$, then, as follows from Eq. (9), it can be found from

$$f=\varphi[x(t),\dot{x}(t),t;f]+\psi[x(t),\dot{x}(t),t;f]$$

or, what is the same, we can put $f(t)=\ddot{x}(t)$ where $\ddot{x}(t)$ is determined by

$$\ddot{x}=\varphi[x(t),\dot{x}(t),t;\ddot{x}]+\psi[x(t),\dot{x}(t),t;\ddot{x}] \tag{10}$$

If Eq. (10) possesses a unique solution (as, for example, in the case of Eq. (6)), then, as discussed earlier, it uniquely determines the motion of the particle acted on by the forces $m\varphi(x,\dot{x},t;\ddot{x})$ and $m\psi(x,\dot{x},t;\ddot{x})$.

We may also define the function $f(t)$ otherwise, for example, $f(t)\equiv\dddot{y}(t)$ where $y(t)$ is the solution of

$$\ddot{y}=\varphi(y,\dot{y},t;\ddot{y})+\psi(y,\dot{y},t;\ddot{y}) \tag{11}$$

at the initial conditions $y(t_0)=x(t_0)$, $\dot{y}(t_0)=\dot{x}(t_0)$ and $\ddot{y}(t_0)=\ddot{x}(t_0)$. Substitution of $\dddot{y}(t)$ for $f(t)$ in Eq. (9) yields

$$\ddot{x}=\varphi[x,\dot{x},t;\dddot{y}(t)]+\psi[x,\dot{x},t;\dddot{y}(t)] \tag{12}$$

When Eq. (11) is compared with Eq. (12), it is apparent that $x=y(t)$ is the unique solution of Eq. (12), and thus $\dddot{x}(t)\equiv\dddot{y}(t)\equiv f(t)$. Substituting $\dddot{x}(t)$ for $f(t)$ in Eq. (9) and in the expressions for the forces, we can see that

$$\ddot{x}=\varphi(x,\dot{x},t;\ddot{x})+\psi(x,\dot{x},t;\ddot{x})$$

uniquely determines the motion of the particle acted on by the forces $m\varphi(x,\dot{x},t;\ddot{x})$ and $m\psi(x,\dot{x},t;\ddot{y})$ provided that this equation possesses a unique solution. Naturally, in this case, in addition to the traditional initial conditions $x(t_0)$ and $\dot{x}(t_0)$, we also have to specify $\ddot{x}(t_0)$.

Further generalizations are obvious. We may consider several functions $f_1(t),f_2(t),\ldots,f_k(t)$ requiring that $f_i(t)$ be identically equal to the $(i+1)$th derivative $x^{(i+1)}(t)$. Reasoning by analogy, we can show that in this case,

$$\ddot{x}=\varphi(x,\dot{x},t;\ddot{x},\dddot{x},\ldots,x^{(k+1)})+\psi(x,\dot{x},t;\ddot{x},\dddot{x},\ldots,x^{(k+1)}) \tag{13}$$

determines the motion of the particle acted on by the forces $m\varphi(x,\dot{x},t;\ddot{x},\dddot{x},\ldots,x^{(k+1)})$ and $m\psi(x,\dot{x},t;\ddot{x},\dddot{x},\ldots,x^{(k+1)})$ provided that Eq. (13) possesses a unique solution at the given initial conditions $x(t_0),\dot{x}(t_0),\ddot{x}(t_0),\ldots,x^{(k)}(t_0)$.

6 Conclusions

The above discussion shows that the question of whether forces that depend on acceleration or other higher derivatives of generalized coordinates are admissible in mechanics (in the sense in which it is considered in Pars's book, i.e., as a question that touches on the foundations of mechanics) is far-fetched. Ultimately, it is only the value of a force at every instant of time that is of importance, no matter whether the force depends on time

explicitly or through the mediation of other time functions, for example, through the mediation of $x(t)$, $\dot{x}(t)$, $\ddot{x}(t)$, $\dddot{x}(t)$, etc. If for some reason or other it is necessary that the control action depend on a higher derivative of a generalized coordinate or such dependence is brought about by our description of an external force, then there are no fundamental obstacles to the application of traditional methods of dynamics to such forces. The only thing that matters is that the force in question must be uniquely determined at every instant of time, i.e., the equation of motion must possess a unique solution. This condition may not be satisfied in the general case. However, this does not suggest the inadmissibility of a particular force; this only points to the necessity of refining the mathematical model by including "finer" factors that affect the motion of the system under consideration or, where control actions are involved, to the fact that this force is physically unrealizable.

References

[1] Pars, L. A., 1964, *A Treatise on Analytical Dynamics*, Heinemann, London.
[2] Birkhoff, G. D., 1927, Dynamical Systems, AMS.
[3] Ghosh, A., 1986, "Velocity-Dependent Inertial Induction and Secular Retardation of the Earth's Rotation," Pramana, J. Phys., **26**(1), pp. 1–8.
[4] Abgaryan, K. A., Kalyazin, E. L., Mishin, V. P., and Rapoport, I. M., 1990, *Rocket Dynamics* (in Russian), Mashinostroyeniye, Moscow, Russia.
[5] Zhechev, M. M., 2003, *Time-Shared Control* (in Russian), Naukova Dumka, Kiev, Ukraine.

Computer Simulation of Rapid Granular Flow Through an Orifice

Hojin Ahn
Department of Mechanical Engineering,
Yeditepe University,
34755 Kayışdağı/Istanbul, Turkey

Rapid granular flow through an orifice (nozzle-shaped flow restrictor) located at the bottom of a vertical tube has been studied using three-dimensional direct computer simulation with the purpose of investigating (1) characteristics of rapid granular flows through the flow restrictor, (2) the choking condition of rapid flow at the orifice and thus conditions at which the maximum discharge rate takes place for the given orifice, and (3) a functional relationship between the discharge rate and flow quantities such as granular temperature and solid fraction. In the present simulation, where the frictional hard-sphere collision operator was employed, it was possible to obtain both rapid and slow (choked) flows through the orifice by controlling the number of particles in the system. The results show that the profile of granular temperature in the vicinity of the orifice plays an important role in determining the choking condition at the orifice. Flow appears to be choked when an adverse granular conduction occurs locally at the orifice in the direction opposite to the mean flow. On the other hand, flow is not choked when the fluctuation energy is conducted in the mean flow direction near the orifice. When flow is not choked, the discharge rate through the orifice increases with increasing solid fraction or normal stress. Once the flow becomes choked, however, the discharge rate decreases as the solid fraction or normal stress increases. Also for inelastic, rough particles, the discharge rate is found to be proportional to the granular temperature to the power of 1.5 and inversely proportional to the gravitational acceleration and the tube length. [DOI: 10.1115/1.2187529]

Introduction

In recent years research on granular flows has received considerable attention due to both academic interests and their application. In fundamental research aspects, the constitutive equations for granular flow and its boundary conditions are of great interest. In application aspects, transports of dry chemicals and pharmaceutical granules (see Khanam and Nanda [1], for example), handling of grain and coal, flows in pebble bed nuclear reactors, and even snow avalanches (see Manna and Herrmann [2]) demand a better understanding of the characteristics of granular flows.

In particular, flows from hoppers and silos have drawn considerable interest since hoppers have been the most common device for the storage and discharge of granular materials. One of the early experimental works on the prediction of the discharge rate of mono-sized granular materials from orifices was conducted by Beverloo et al. [3]. They proposed a correlation for the discharge rate of solid materials through a hopper orifice as follows:

$$\dot{m} = C\rho g^{1/2}(D_o - kd)^{2.5}$$

where ρ is the bulk density of granular material, g is the gravitational acceleration, D_o is the diameter of the orifice at the bottom of the hopper, and d is the particle diameter. In this correlation, two empirical parameters, C and k, need to be determined experimentally. Beverloo et al. assumed ρ as the initial fill density of the hopper. Consequently, the discharge coefficient C was determined to be in a range between 0.55 and 0.65. Nedderman and Laohakul [4] found the value of k to be approximately 1.5 ± 0.1 for mono-sized spherical particles. Humby et al. [5] attemped to predict hopper discharge rates of binary granular mixtures and recognized the importance of predicting density near the orifice in determining the discharge rate of granules through a hopper orifice. Recently Grantham and Forsberg [6] investigated the flow behavior of a powder through a silo from the displacement measurements using Digital Speckle Radiography (DSR). Sielamowicz et al. [7] employed the digital particle image velocimetry (DPIV) technique to measure velocity fields of plane hopper flows.

Potapov and Campbell [8] conducted two-dimensional computer simulations of granular flows from plane hoppers. Hirshfeld and Rapaport [9] also employed discrete-particle simulations to examine three-dimensional granular flows from a silo. Both employed a soft-particle simulation technique which models particle interactions as linear springs and linear dashpots. In particular, the results of the computer simulations by Potapov and Campbell showed that hopper flows appear to be in an intermediate regime between quasistatic yield and rapid flow; consequently, hopper flows may neither be modeled with plasticity theory under the assumption that the material is always yielding, nor modeled as rapid flows.

Most of the research on hopper flows focuses on particles slowly flowing out of a hopper where particles are stored at high density. These flows are far from rapid flows. However, it may be interesting to study the transition from rapid flows to slow flows within a system. For example, when granular materials transported in high speed encounter a reduced flow area such as an orifice, the choking phenomenon may take place at the orifice for sufficiently high flow rates.

The present work simulates gravitational flows of granular materials through a nozzle-shaped flow restrictor (orifice) located at the bottom of a cylindrical tube. This paper attempts to address the choking condition of granular flows and thus to determine conditions for the maximum transport rate of granular materials to take place for a given orifice size. The effects of flow restrictor size (or orifice size), particle diameter, pipe length, and the magnitude of gravity are investigated using computer simulation. Furthermore, while most studies on prediction of discharge rates from hoppers make efforts to relate the discharge rate with the hopper orifice diameter and density, this paper presents an alternative way

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received October 13, 2005; final manuscript received January 18, 2006. Review conducted by S. Govindjee. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

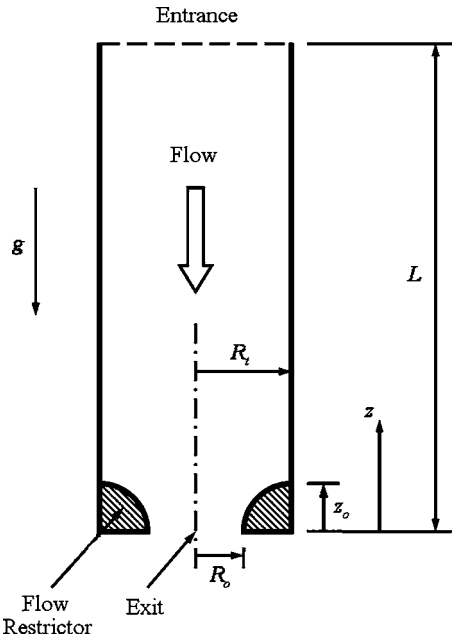


Fig. 1 Schematic of gravitational flow through the flow restrictor

of expressing the discharge rate as a function, not of orifice diameter, but of granular temperature and solid fraction.

Simulation Method

The present work simulates gravitational granular flow through a cylindrical tube and a nozzle-shaped flow restrictor (or orifice) located at the bottom of a tube as shown in Fig. 1. Granular particles start freefall under gravity at the entrance of the tube, accelerate through the vertical tube, and exit the system through the flow restrictor. The cross-sectional view of the flow restrictor is also shown in Fig. 1. The shape of the flow restrictor makes it relatively easy to calculate the collision times between particles and the wall of the flow restrictor. Also the shape has no horizontal surface. If a horizontal surface were present in the system, some inelastic particles would sit on the surface after many collisions with the solid surface (unless other particles hit the particle away from the surface). The current simulation method cannot handle the situation of particles staying in constant contact with solid walls or other particles.

As shown in the schematic, the tube length is denoted by L, and the radii of the tube and the orifice by R_t and R_o, respectively. In the present simulation, L was selected to be either 0.3 or 0.45 m, and R_o varied from 6 to 12 mm. A fixed value of 20 mm is selected for R_t. The dimensions of the system, such as the tube length and diameter, were selected to be small enough to provide a sufficiently high solid fraction for a given number of particles in the system. Most simulations were conducted with spherical particles 3 mm in diameter, which seemed to be large compared to the orifice size. However, the large size of the particle was necessary to obtain a high solid fraction for a given number of particles in the system. On the other hand, several simulations were conducted with 2-mm particles to investigate the effect of the particle diameter on flows through the orifice. The density of the particles was assumed to be 2500 kg/m^3 to calculate normal stresses and mass flow rates.

The current simulation employed the frictional hard-sphere collision operator described in Walton [10]. Consider two identical spheres of d in diameter with centers located at $\mathbf{r}_a$ and $\mathbf{r}_b$, traveling with velocities $\boldsymbol{\nu}_a$ and $\boldsymbol{\nu}_b$ and rotational velocities $\boldsymbol{\omega}_a$ and $\boldsymbol{\omega}_b$. Then the unit vector from sphere a to b at contact, the relative velocity, the normal and tangential direction relative velocities, and the relative surface velocity before collision are given by

$$\mathbf{r}_{ab} = (\mathbf{r}_b - \mathbf{r}_a)/d$$

$$\boldsymbol{\nu}_{ab} = \boldsymbol{\nu}_b - \boldsymbol{\nu}_a$$

$$\boldsymbol{\nu}_n = (\boldsymbol{\nu}_{ab} \cdot \mathbf{r}_{ab})\mathbf{r}_{ab}$$

$$\boldsymbol{\nu}_t = \boldsymbol{\nu}_{ab} - \boldsymbol{\nu}_n$$

$$\boldsymbol{\nu}_s = \boldsymbol{\nu}_t + \frac{d}{2}\mathbf{r}_{ab} \times (\boldsymbol{\omega}_a + \boldsymbol{\omega}_b)$$

The resulting changes in normal, tangential and rotational velocities for each sphere after collision are given by

$$\Delta\boldsymbol{\nu}_{na} = -\Delta\boldsymbol{\nu}_{nb} = \tfrac{1}{2}(1+e)\boldsymbol{\nu}_n$$

$$\Delta\boldsymbol{\nu}_{ta} = -\Delta\boldsymbol{\nu}_{tb} = \frac{K(1+\beta)}{2(K+1)}\boldsymbol{\nu}_s$$

$$\Delta\boldsymbol{\omega}_a = \Delta\boldsymbol{\omega}_b = \frac{(1+\beta)}{d(K+1)}\mathbf{r}_{ab} \times \boldsymbol{\nu}_s$$

where K is the square of the ratio of the radius of gyration to the radius of the sphere (K=0.4 for the sphere).

The coefficient of rotational restitution, β, is defined by the ratio of the relative surface velocity of two objects at the post-collision to that at the pre-collision. For a perfectly smooth surface, its value is −1, while a perfectly rough surface has the value of 1. In the present work, as described in Walton [10], the coefficient of rotational restitution is not constant but varies, depending on the conditions of the impact, that is, whether a particle slides or rolls at contact. The predetermined (constant) maximum coefficient of rotational restitution was chosen to be 0.0 and the friction coefficient 0.4 was used in the present study.

The coefficient of restitution both between two particles and between particles and the wall of the flow restrictor, e, was set as a constant value of 0.95. The value of 0.95 was chosen for two reasons: One reason was that a high value for the coefficient of restitution is desired to avoid situations in which particles slide down the wall surface of the restrictor while in constant contact with the wall. Another reason was to allow comparison with future experimental work with glass beads. Lun and Savage [11] show that the coefficient of restitution for glass beads is about 0.95 for impact velocity of 1 to 2 m/s. In the present simulation with a vertical tube 0.3 m in length, the maximum impact velocity is approximately 2 m/s; on the other hand, very low impact velocity is also present in the simulation, in which case the coefficient of restitution should be close to 1.0. Though most cases were run with the value of 0.95, several cases were also run with different values of the coefficient of restitution.

On the other hand, the coefficient of restitution between the tube wall and particles was selected to be 1.0. For low coefficients of restitution, some cases were observed where particles were falling down swirling along the cylindrical tube wall while in constant contact with the wall. The value of 1.0 for the coefficient of restitution eliminated those cases where particles are in constant contact with the side wall.

Computation procedures are as follows: Initially particles are randomly positioned in the tube. A statistically steady flow is achieved by recycling particles to the top of the tube after they are discharged from the orifice at the bottom. The particles enter the tube with an initial velocity (which was selected as 0.2 m/s in the present work), being uniformly distributed over the cross section of the tube. Convergence to a steady state was usually achieved within 0.5 s, which is about twice the time required for freefall through the tube 0.3 m in height. Also all the particles initially placed in the system are usually discharged from the system

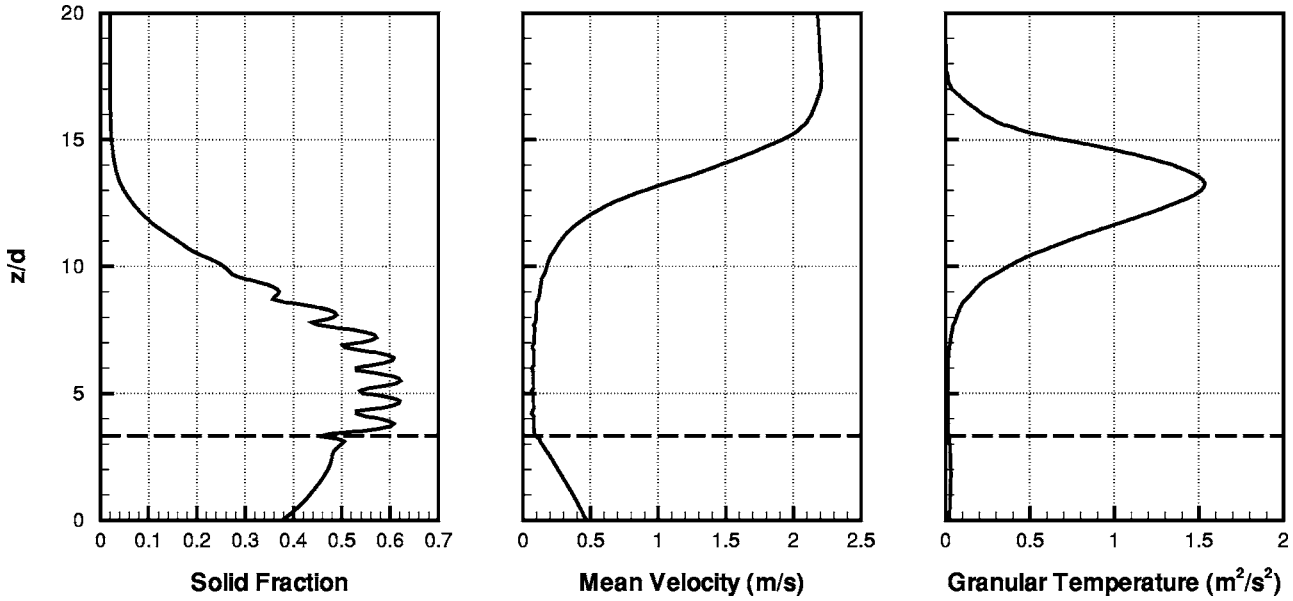


Fig. 2 Case of 2000 particles in the system with an orifice 10 mm in radius. The location of $z_o/d=3.3$ is also shown with a dashed line.

within the first 0.5 s. The system was simulated up to 2–30 s, depending on the number of particles in the system. The simulation time was selected such that the total number of collisions and the number of particles having been discharged from the orifice are high enough to ensure reasonable average processes. All collision times among particles and between particles and walls were initially calculated and stored. After the initial calculation, collision times were calculated only for particles that have just collided with other particles or walls. Careful programming was required to calculate collision times between particles and the wall of the flow restrictor. (Since the flow restrictor has the cross-sectional shape of a quadrant of a circle, there are a maximum of four possible roots for collision time between a particle and the wall. The least positive root is the desired collision time.)

The total number of particles in the system was selected from 100 to 2500 in the present study. The CPU time for each simulation varied from 1 min to 1 week using a personal computer equipped with 3.0 GHz Pentium IV, depending mainly on the number of particles in the system and the orifice size for the given tube dimensions and the particle diameter.

Simulation calculated solid fraction, three mean velocities (two of which should be zero), three components of fluctuational velocities, three components of mean rotational velocities (all of which should be zero) and their fluctuations. All these flow quantities were averaged over the cross-sectional area of the tube, and thus were measured as functions of the distance from the exit along the tube with the resolution of a tenth of the particle diameter. Mass flow rate or discharge rate from the orifice was also measured. Only a normal stress component exerted on the orifice was calculated in the present work. One of the highest collision frequencies averaged over the total simulation time (usually 2 s) was observed to be 0.4 μs.

Characteristics of Flows Through the Flow Restrictor

Characteristics of rapid granular flows through the flow restrictor shown in Fig. 1 were studied with 3-mm particles. Figures 2–4 show the simulation results with a vertical tube 0.3 m in length and 20 mm in radius and with an orifice 10 mm in radius. The vertical coordinate from the exit, z, was nondimensionalized by the particle diameter, d. The location of the beginning of the flow restrictor (the orifice) is denoted by z_o as shown in Fig. 1. With the sizes of the tube, the orifice and the particle in mind, therefore, it is noted that the flow restrictor starts at $z_o/d=3.3$ and ends at $z/d=0$ (at which particles completely exit the system).

The case of the simulation with 2000 particles is presented in Fig. 2. Solid fraction, mean velocity, and granular temperature which are averaged over the cross-sectional area of the tube are plotted as a function of the nondimensionalized vertical coordinate. In this study, the granular temperature is defined as the sum of mean-squares of three fluctuational velocities and three fluctuational rotational velocities times the moment of inertia divided by the particle mass. Particles start freefall in acceleration from $z/d=100$ (at which the entrance of the tube is located).

The increase of granular temperature near $z/d=17$ indicates that particles start to collide with one another near that location, resulting in the increase of solid fraction and the decrease of mean velocity. The maximum of the granular temperature is observed near $z/d=13$. Note that the granular temperature is very low in the regions above $z/d=18$ and below $z/d=6$.

The profile of solid fraction shows that the solid fraction is low before collisions become frequent and that the solid fraction increases rapidly as particles crowd the region upstream of the flow restrictor. In that region, particles are piled up to a thickness of about five particle diameters, the mean solid fraction being about 0.56. It is apparent, therefore, that the flow is choked at the flow restrictor. Several layers of the microstructure of particles upstream of the flow restrictor are also observed as particles form layers near a solid boundary (in this case, the wall of the flow restrictor). The layered microstructure has been observed in other literature as well (see Ahn et al. [12], for example). Due to this microstructure, it was difficult to define the value of the solid fraction just upstream of the flow restrictor. The solid fraction immediately upstream of the flow restrictor was, therefore, defined by the solid fraction averaged over the distance of one diameter immediately upstream from the flow restrictor. This average value for the solid fraction was used for various plots in the following sections.

The profile of mean velocity is also plotted in Fig. 2. The mean velocity increases as particles accelerate under gravity, and then decreases as particles start to collide with one another near $z/d=17$. Note that the mean velocity decreases rapidly between $z/d=17$ and 8 where the granular temperature is shown high. This indicates that the kinetic energy of particles in the mean velocity is converted into the fluctuational energy represented by the

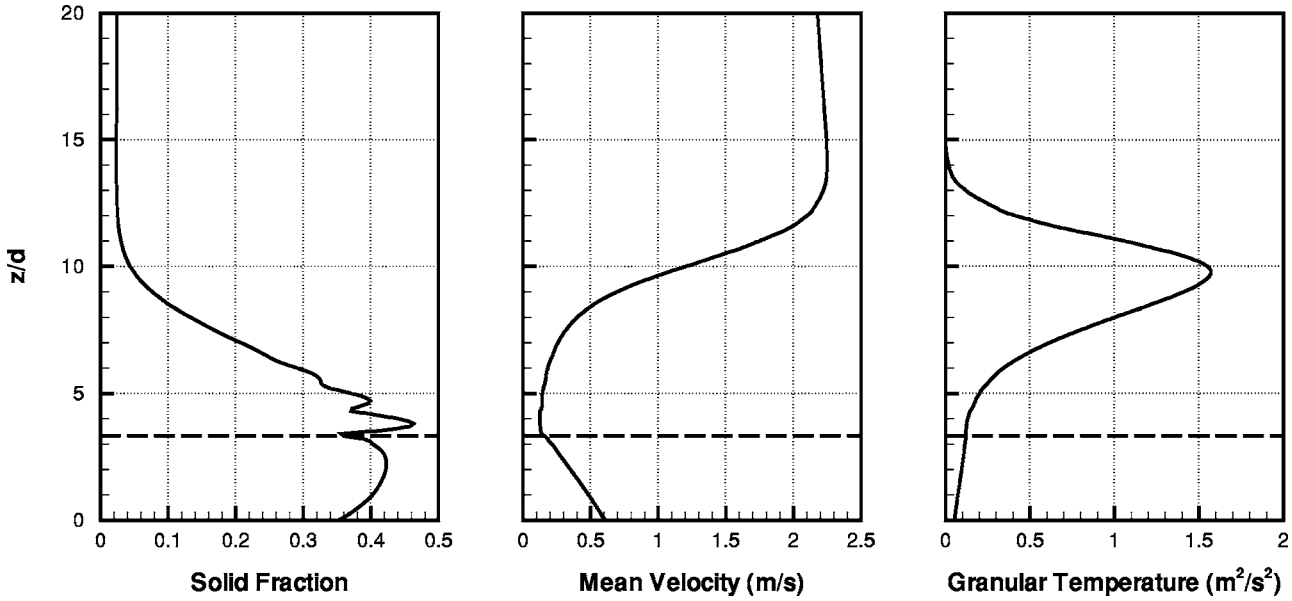


Fig. 3 Case of 1600 particles in the system with an orifice 10 mm in radius. The location of $z_o/d=3.3$ is also shown with a dashed line.

granular temperature (and some of the kinetic energy is dissipated due to inelastic collision and friction). The minimum of the mean velocity is observed immediately upstream of the flow restrictor (at $z/d=3.3$), after which the mean velocity increases as particles exit the flow restrictor in acceleration. It is interesting to note that the mean velocity upstream of the flow restrictor is nearly constant over the distance of several particle diameters for the case of the choked flow.

The case of 1600 particles is shown in Fig. 3. This is the case where the discharge rate for the orifice size of 10 mm in radius is the maximum. (This will be discussed more in the next section.) The mean solid fraction just upstream of the flow restrictor is about 0.42. Just as in Fig. 2, the microstructure of a few layers of particles is also observed in the plot of solid fraction against the normalized vertical coordinate. The profile of the mean velocity shares characteristics similar to those in Fig. 2. The mean velocity decreases as collisions become frequent, reaches a minimum just upstream of the flow restrictor, and increases again as particles exit in acceleration through the flow restrictor. The location of the maximum granular temperature is at $z/d=10$ which is lower than that of the case in Fig. 2. Also the granular temperature just upstream of the flow restrictor does not appear to be as low as shown in Fig. 2, but is nearly zero above $z/d=15$.

The simulation result with 500 particles is shown in Fig. 4. One distinct difference between this case in Fig. 4 and cases in Figs. 2 and 3 is that the collision rate is low in this case due to a small number of particles in the system compared to the previous cases. Therefore, particles can bounce up high from the flow restrictor wall without collision with other particles. As a result, the magni-

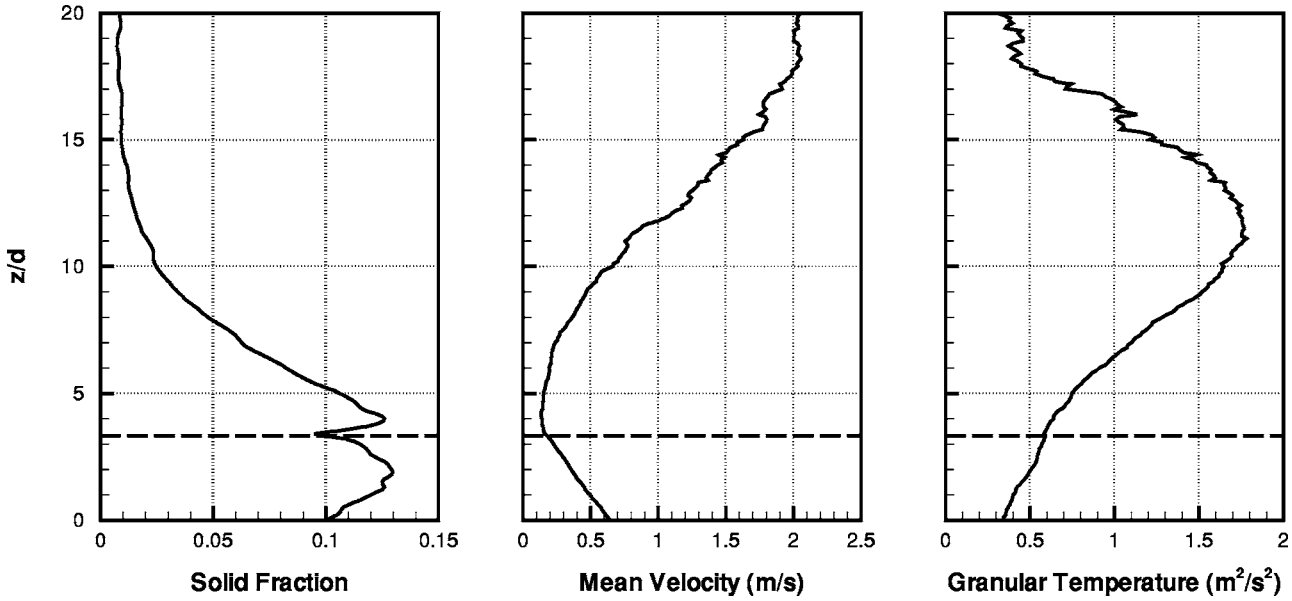


Fig. 4 Case of 500 particles in the system with an orifice 10 mm in radius. The location of $z_o/d=3.3$ is also shown with a dashed line.

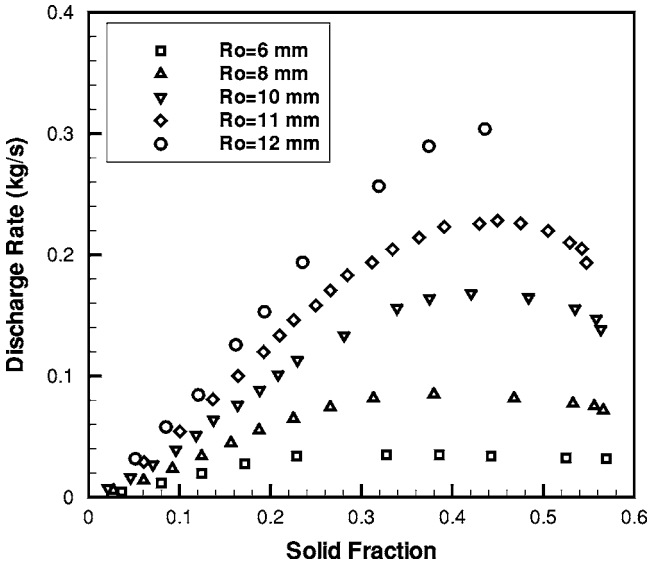


Fig. 5 Discharge rate from the orifice against the mean solid fraction upstream of the orifice for various orifice sizes

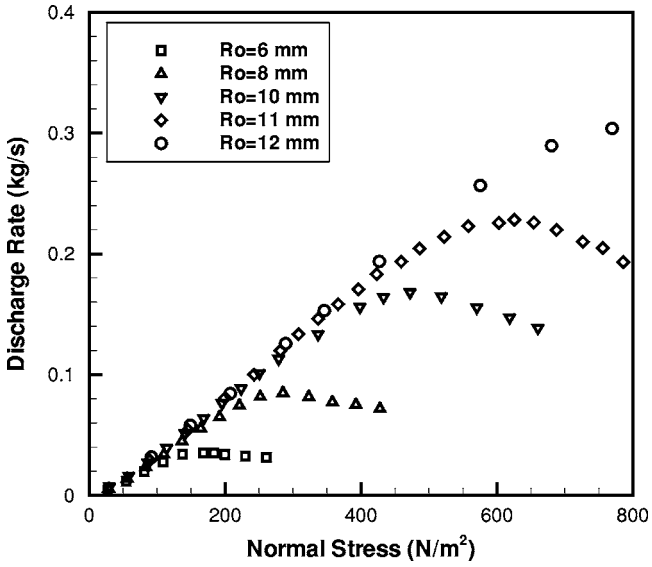


Fig. 6 Discharge rate from the orifice against the normal stress exerted on the orifice for various orifice sizes

tude of granular temperature even at $z/d=20$ is not negligible. Even near the wall of the flow restrictor, the collision rate among particles is low and granular temperature does not decrease quickly, compared to the other cases. The mean solid fraction just upstream of the flow restrictor is about 0.12. The mean velocity reaches a minimum just upstream of the flow restrictor just as in the other cases.

Preliminary Investigation on Choking Condition

The discharge rate from the orifice is presented as a function of the mean solid fraction just upstream of the orifice in Fig. 5. As mentioned in the previous section, the mean solid fraction is obtained by averaging solid fraction over the distance of one particle diameter immediately upstream from the flow restrictor. Several sizes of orifices were examined with 3-mm particles. For a given orifice, the discharge rate is almost linear with the solid fraction for the region of low solid fraction as shown in Fig. 5. As the solid fraction increases further, the discharge rate reaches a maximum. After the maximum discharge rate is achieved, however, the discharge rate decreases as the solid fraction increases. It is probably because particles at high solid fraction become locked with each other. Note the rapid decrease of the discharge rate in the cases with orifice sizes 10 mm and 11 mm in radius as the solid fraction approaches its maximum shear limit (about 0.6). Therefore, it appears that the choking situation starts to take place at the maximum discharge rate. Though it is difficult to determine the exact solid fraction at which choking takes place, the maximum discharge rate is observed at a solid fraction of 0.3 to 0.5, depending on the orifice size. When the orifice size is smaller, choking takes place at a lower solid fraction.

The discharge rate from the orifice is examined as a function of the normal stress exerted on the orifice in Fig. 6. For low normal stress, the discharge rate is a function only of normal stress, independent of the orifice size as shown in Fig. 6. All the data from different orifice sizes fall into a single line for low normal stress. But as normal stress increases, the discharge rate is not a function only of the normal stress but also of the orifice size. For a smaller orifice size, the data deviate from the single line at a lower normal stress. For a given orifice, as the normal stress increases, the discharge rate increases, reaches a maximum, and then decreases. Examine, for example, the case of the orifice 10 mm in radius. The discharge rate is almost linear with the normal stress until the normal stress is about 300 N/m^2. After 300 N/m^2, the discharge rate is still increasing but in a fashion deviating from the single line. Then it reaches a maximum at 470 N/m^2 after which the discharge rate decreases with further increase of the normal stress. This may be explained as follows: At approximately 300 N/m^2, particle collisions start locking particles to one another and thus preventing particles from being freely discharged from the orifice. Nevertheless, a higher normal stress pushes particles out of the exit in a stronger manner, resulting in a higher discharge rate. However, further increase in normal stress starts locking particles more severely, resulting in the decrease of the discharge rate.

The discharge rates calculated by the present simulation as shown in Fig. 6 were compared with the correlation by Beverloo et al. [3]. The discharge rates predicted by Beverloo et al. were lower than the simulation results by approximately 50%. The cause of the discrepancy remains to be investigated.

The present work attempts to investigate the conditions at which the maximum discharge rate occurs. The profiles of granular temperature are plotted in Fig. 7 as a function of the nondimensionalized distance from the exit for cases with different numbers of particles in the system with an orifice 10 mm in radius. Though not clear from Figs. 5 and 6, the maximum discharge rate for the orifice 10 mm in radius occurred when the number of particles in the system was 1600. As the number of particles in the system increased up to 1600, the discharge rate increased. However, further increase of the number of particles in the system resulted in the decrease of the discharge rate.

A detailed observation of the profiles of granular temperature is presented in Fig. 7. It is noted that the converging section of the flow restrictor starts at $z/d=3.3$ and the exit is located at $z/d=0$. The cases of 1400 and 1500 particles in the system show that the granular temperature monotonically decreases as particles move toward the exit. That is, granular temperature upstream of the flow restrictor is higher than that at the converging section of the flow restrictor. This shows that granular conduction occurs in the mean flow direction for these cases. Granular flows for these cases were not observed to be choked.

On the other hand, the cases of 1700, 1800, 1900, and 2000 particles show that granular temperature decreases in a flow direction, reaches a minimum at a location somewhere upstream of the flow restrictor, and then increases at the converging section of the flow restrictor before decreasing again at the exit. For these cases, therefore, initially granular conduction occurs in the bulk in the same direction with the mean flow. But as particles move near the flow restrictor, fluctuational energy is conducted backward from

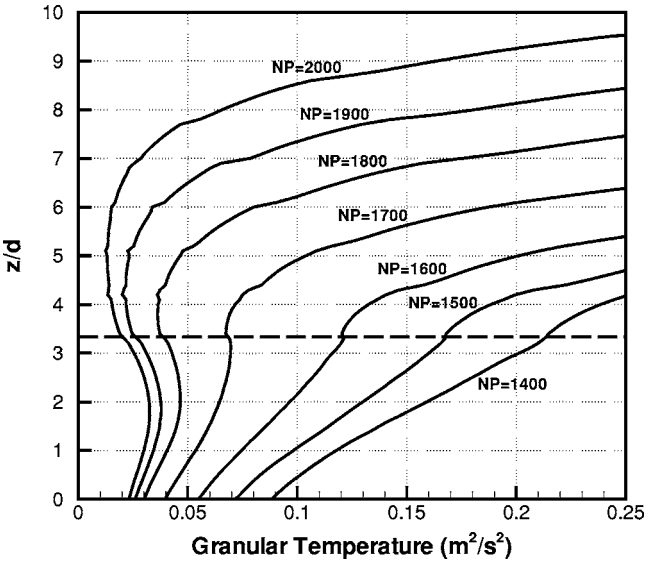


Fig. 7 Profiles of granular temperature as a function of the distance from the exit for cases with different numbers of particles in the system with an orifice 10 mm in radius. The location of $z_o/d=3.3$ is also shown with a dashed line.

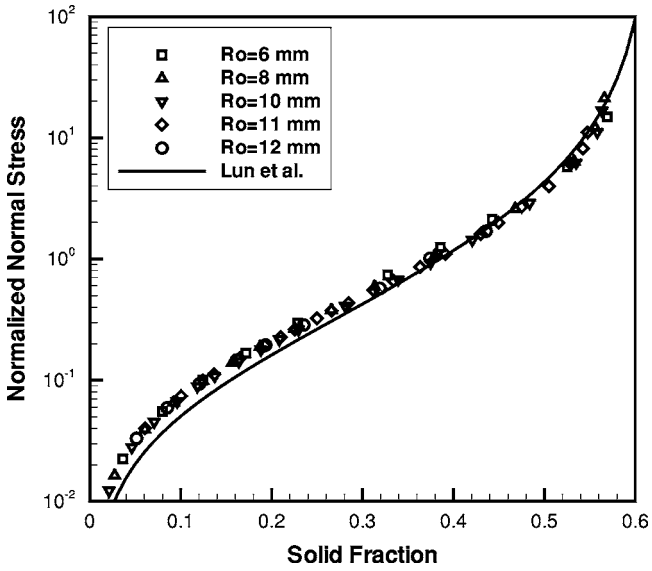


Fig. 8 Non-dimensionalized normal stress versus solid fraction for the case of $e_p=0.95$. Data from the current simulation results are compared with the result of Lun et al. [14].

the converging section of the flow restrictors to the bulk against the mean flow direction. That is, there is a local adverse granular conduction (though small in magnitude) at the vicinity of the orifice in the direction opposite to the mean flow. In these cases, flows were observed to be choked.

The case of 1600 particles, which yielded the maximum discharge rate, shows that there is no local granular conduction immediately upstream of the flow restrictor (at $z/d=3.3$ to 3.5). Therefore, it may be concluded from the above observation that the maximum discharge rate occurs when no local granular conduction occurs upstream of an orifice. When an adverse granular conduction locally exists against the mean flow direction, discharge rate decreases and flow becomes choked. On the other hand, when granular temperature is monotonically conducted from the bulk to the exit, flow is not choked and discharge rate increases as the solid fraction and normal stress increase as shown in Figs. 5 and 6.

It is interesting to examine some details in the profiles of the solid fraction from Figs. 2–4 in conjunction with the granular temperature profiles. The case of Fig. 2 represents a flow that is choked at the flow restrictor. It is noted that the mean solid fraction just upstream of the flow restrictor is about 0.56, which is higher than the maximum solid fraction (about 0.51) in the converging section of the flow restrictor (from $z/d=3.3$ to 0). On the other hand, Fig. 3 shows the case where the discharge rate is the maximum for the given orifice size. In this case, the mean solid fraction just upstream of the flow restrictor is about 0.42, which is essentially equal to the maximum solid fraction in the converging section of the flow restrictor. Finally, the case in which flow is not choked is shown in Fig. 4. The mean solid fraction upstream of the flow restrictor is 0.12, which is less than the maximum solid fraction (about 0.13) in the converging section of the flow restrictor. In general, the solid fraction is low in a region where granular temperature is high. Therefore, the fact that the solid fraction in the converging section is higher than that immediately upstream of the flow restrictor in Fig. 4 may indicate that the granular temperature at the converging section is lower than it is upstream of the flow restrictor. Similarly, it may be said for the case of Fig. 2 that the lower solid fraction in the converging section implies the granular temperature at the converging section is higher than immediately upstream of the flow restrictor, thus yielding an adverse granular conduction.

The choking of a granular flow, similar to conventional fluid flows, should occur when the wave speed of the granular material in motion exceeds the mean velocity of the flow. The wave speed through a granular material in motion is expected to be a function of the solid fraction and granular temperature. The wave speed must increase as the solid fraction and granular temperature increase. However, the propagation mechanism of pressure waves in granular materials appears very complicated (see Hostler and Brennen [13]). We still do not fully understand wave propagation in granular flows to describe the choking condition in terms of the wave propagation speed. As shown in Figs. 5 and 6, therefore, this paper attempted to examine the choking condition in terms of the mass flow rate instead of the wave speed. It has been shown that the choking phenomenon takes place when the mass flow rate reaches a maximum. In addition, the result shown in Fig. 7 implies that the maximum flow rate takes place when no local granular conduction exists just upstream of the orifice, and thus that choking takes place when an adverse granular conduction just appears upstream of the orifice. Further investigation on the choking condition of granular flows remains to be done.

Mass Flow Rate as a Function of Flow Quantities

Almost all research on the discharge rate through a hopper orifice relates the discharge rate as a function of the orifice size, as in Beverloo et al. [3]. Figures 5 and 6 also show the discharge rate as a function of the orifice size and solid fraction or normal stress. However, it may be desired to establish a universal representation of the discharge rate, which does not include the orifice size. The present study attempts to relate the discharge rate with flow quantities such as the granular temperature and solid fraction. There has been some work to obtain the granular temperature for granular flow out of a hopper. For example, Hirshfeld and Rapaport [9] calculated the distribution of kinetic energy in the flow from a silo (through an orifice). But they did not relate the kinetic energy to the flow rate. The current work may be the first attempt to relate the discharge rate to flow quantities instead of orifice size.

The rheological behavior of granular flows is briefly examined in the current study. Figure 8 shows the normal stress normalized by $\rho_p T$ against the solid fraction which is the average value over the distance of one particle diameter upstream from the flow restrictor. The density of the particle is denoted as ρ_p. Recall that the coefficient of restitution of 0.95 and the Coulomb friction coeffi-

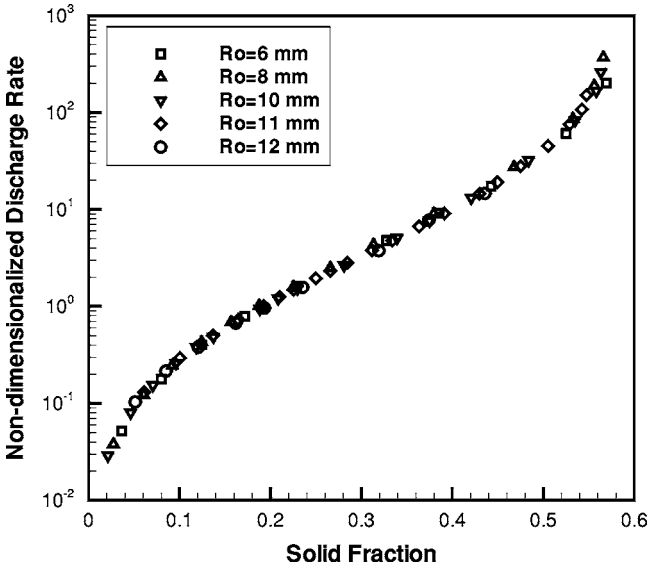


Fig. 9 The non-dimensionalized discharge rate against solid fraction for various orifice sizes

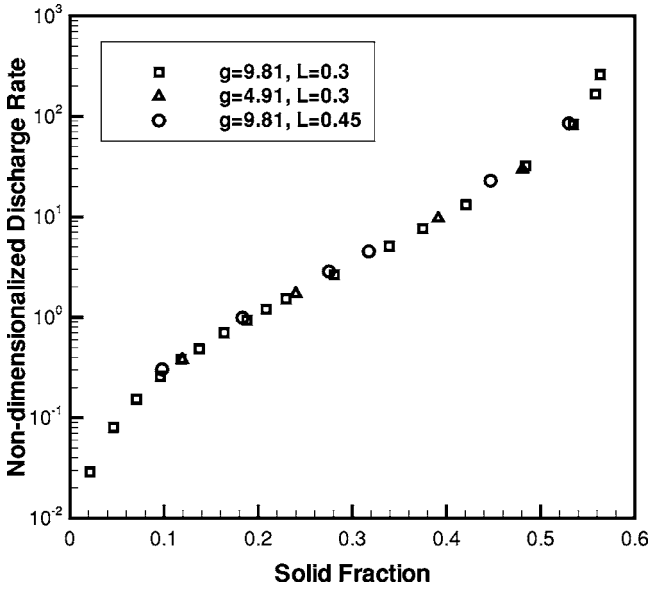


Fig. 10 Non-dimensionalized discharge rate as a function of solid fraction with various values for the gravitational acceleration and tube lengths

cient of 0.4 and the predetermined maximum coefficient of rotational restitution of 0.0 were used for each data point. All data are well correlated with a single curve regardless of various orifice sizes employed in the plot. Also the theoretical prediction by Lun et al. [14] for the case of $e_p=0.95$ is plotted in Fig. 8. There is good agreement between the current simulation results and those of Lun et al.

In order to propose an expression for the discharge rate which is independent of the orifice size, we limit ourselves to flows with inelastic, rough particles. In the current simulation setup, particles falling from the top of the tube will have enough collisions with other particles and also with walls that most of the initial potential energy of the particles at the entrance of the tube will be dissipated before the particles discharge through the orifice. However, the dissipation rate is known to be proportional to $T^{3/2}$ (see Lun et al. [14], for example). Therefore, we may relate the initial potential energy flux to the dissipation rate as follows:

$$\dot{m}gL \sim T^{3/2}$$

Thus the non-dimensionalized discharge rate is proposed to be

$$\frac{\dot{m}gL}{\rho_p T^{3/2} A} = f(\nu, e, \mu, \beta)$$

where A is the cross-sectional area of the tube, T is the granular temperature just upstream of the orifice, and f is a function of the solid fraction (ν), coefficient of restitution (e), friction coefficient (μ), and coefficient of rotational restitution (β). Noting that $\dot{m} = \rho_p \nu A \bar{V}$ where $\bar{V}$ is the mean velocity over the cross-sectional area of the tube, the mean velocity may be related to the granular temperature as follows:

$$\bar{V} = \frac{T^{3/2}}{gL} h(\nu, e, \mu, \beta)$$

where h is a function equal to f/ν.

The non-dimensionalized discharge rate as proposed above is plotted against the solid fraction in Fig. 9 for various orifice sizes. All data from the present simulation are remarkably correlated to a single curve, independent of the orifice size. This confirms that the dissipation rate is indeed proportional to the granular temperature to the power of 1.5. Also our basic assumption is valid that most of the initial potential energy flux at the entrance of the tube is dissipated before particles discharge through the orifice.

The effects of the gravitational acceleration and the tube length were also investigated as shown in Fig. 10. In these runs, the gravitational acceleration was selected as either $g=9.81$ or 4.91 m/s^2 and the tube length as either $L=0.3$ or 0.45 m. As shown in Fig. 10, all data with different values of the gravitational acceleration and the tube length are well correlated, thus confirming the proposed non-dimensionalization for the discharge rate.

It is tempting to non-dimensionalize the discharge rate by the particle diameter instead of the tube length L. However, several simulation results (though not included here) showed that the discharge rate is independent of the particle diameter, but rather depends on the vertical tube length as proposed above.

It should be recalled that the proposed non-dimensionalized discharge rate should be a function of several parameters such as the coefficient of restitution, the coefficient of rotational restitution and the friction coefficient as well as solid fraction. Though not included here, several runs with different values of the coefficient of restitution or of the friction coefficient do not correlate with the single curve in Fig. 9, as expected. For a lower coefficient of restitution, the discharge rate was higher than the case for a higher coefficient of restitution. Therefore, it seems that a system of particles with more dissipation has a higher discharge rate. More extensive study is necessary to investigate the effects of those parameters. In particular, the proposed non-dimensionalization for the discharge rate assumes the particles to be inelastic and rough (or frictional). Consequently the initial potential energy flux of particles at the entrance of the tube is assumed to be mostly dissipated before they are discharged through an orifice. Therefore, the proposed non-dimensionalization may not work for highly elastic and smooth particles.

It should also be noted that the current simulation used the coefficient of restitution of 0.95, which is highly elastic. Nevertheless, it was observed that the Coulomb friction coefficient of 0.4 and the pre-determined maximum coefficient of rotational restitution of 0.0 in the current study rendered significant dissipation in the control volume. Therefore, it appears that the dissipation due to collisions in tangential direction is as significant as the dissipation due to collisions in normal direction.

Conclusion

The present work simulates gravitational flows of granular materials through a nozzle-shaped flow restrictor (orifice) located at

the bottom of a cylindrical tube. In the present simulation setup, particles discharged from the orifice again enter the system at the top of the tube. By controlling the number of particles in the system, therefore, it was possible to obtain both rapid and slow (choked) flows through the orifice. The slow, choked flow had a region of dense particles to a thickness of about five particle diameters upstream of the orifice and might not be considered as a typical hopper flow (whose normal stress near the orifice is much larger than that of the current simulation due to particles being piled up to the height of hundreds or thousands of particle diameters in the hopper). Nevertheless, it was very encouraging that the rapid flow model employed in the current simulation could create slow flows with the local solid fraction of up to 0.6 near the orifice.

When flow is not choked, the discharge rate through the orifice increases with increasing solid fraction and normal stress. But once the flow becomes choked, the discharge rate decreases as solid fraction and normal stress increase. Thus the maximum discharge rate takes place when the flow starts choking. When the orifice size is small, the choking takes place at lower solid fraction and normal stress.

The profiles of granular temperature in the vicinity of the orifice were carefully examined to determine the choking condition at the orifice. Flow appears to be choked when an adverse granular conduction occurs locally at the orifice in the direction opposite to the mean flow. On the other hand, flow is not choked when the fluctuation energy is conducted in the mean flow direction near the orifice. Therefore, the maximum discharge rate takes place when no local granular conduction exists immediately upstream of the orifice.

This paper also presents an alternative way of expressing the discharge rate as a function, not of orifice diameter, but of granular temperature and solid fraction. For inelastic, rough particles, the discharge rate is found to be proportional to the granular temperature to the power of 1.5. This is because the initial potential energy flux at the entrance of the tube should dissipate with the granular temperature to the power of 1.5. The discharge rate is also a function of the solid fraction, inversely proportional to the gravitational acceleration and the height of the tube (which represent the initial potential energy of particles at the entrance of the tube). Since the dissipation rate depends on several parameters such as the coefficient of restitution, the coefficient of rotational restitution, and the friction coefficient, a different set of the correlation of the non-dimensionalized discharge rate versus the solid fraction is expected for a different set of parameters (as indicated in several simulation runs that are not included here).

References

[1] Khanam, J., and Nanda, A., 2005, "Flow of Granules Through Cylindrical Hopper," Powder Technol., **150**, pp. 30–35.

[2] Manna, S. S., and Herrmann, H. J., 2000, "Intermittent Granular Flow and Clogging With Internal Avalanches," Eur. Phys. J. E, **1**, pp. 341–344.

[3] Beverloo, W. A., Leniger, H. A., and van de Velde, J., 1961, "The Flow of Granular Solids Through Orifices," Chem. Eng. Sci., **15**, pp. 260–269.

[4] Nedderman, R. M., and Laohakul, C., 1980, "The Thickness of the Shear Zone of Flowing Granular Materials," Powder Technol., **25**, pp. 91–100.

[5] Humby, S., Tüzün, U., and Yu, A. B., 1998, "Prediction of Hopper Discharge Rates of Binary Granular Mixtures," Chem. Eng. Sci., **53**, pp. 483–494.

[6] Grantham, S. G., and Forsberg, F., 2004, "Measurement of Granular Flow in a Silo Using Digital Speckle Radiography," Powder Technol., **146**, pp. 56–65.

[7] Sielamowicz, I., Blonski, S., and Kowalewski, T. A., 2005, "Optical Technique DPIV in Measurements of Granular Material Flows, Part 1 of 3—Plane Hoppers," Chem. Eng. Sci., **60**, pp. 589–598.

[8] Potapov, A. V., and Campbell, C. S., 1996, "Computer Simulation of Hopper Flows," Phys. Fluids, **8**, pp. 2884–2894.

[9] Hirshfeld, D., and Rapaport, D. C., 2001, "Granular Flow From a Silo: Discrete-Partical Simulations in Three Dimensions," Eur. Phys. J. E, **4**, pp. 193–199.

[10] Walton, O. R., 1993, "Numerical Simulation of Inelastic, Frictional Particle-Particle Interactions," *Particulate Two-phase Flow*, M. C. Roco, ed., Butterworth-Heinemann, Boston, pp. 884–911.

[11] Lun, C. K. K., and Savage, S. B., 1986, "The Effect of an Impact Velocity Dependent Coefficient of Restitution on Stresses Developed by Sheared Granular Materials," Acta Mech., **63**, pp. 15–44.

[12] Ahn, H., Brennen, C. E., and Sabersky, R. H., 1991, "Measurements of Velocities, Velocity Fluctuations, Density, and Stresses for Chute Flows of Granular Materials," J. Appl. Mech., **58**(3), pp. 792–803.

[13] Hoster, S. R., and Brennen, C. E., 2005, "Pressure Wave Propagation in a Granular Bed," Phys. Rev. E, under review.

[14] Lun, C. K. K., Savage, S. B., Jeffrey, D. J., and Chepurniy, N., 1984, "Kinetic Theories for Granular Flow: Inelastic Particles in Couette Flow and Slightly Inelastic Particles in a General Flowfield," J. Fluid Mech., **140**, pp. 223–256.

Analysis of Wave Propagation in Beams With Transverse and Lateral Cracks Using a Weakly Formulated Spectral Method

N. Hu
Associate Professor
e-mail: hu@ssl.mech.tohoku.ac.jp

H. Fukunaga

M. Kameyama

Department of Aeronautics and Space Engineering,
Tohoku University,
6-6-01 Aramaki-Aza-Aoba,
Aoba-ku, Sendai 980-8579,
Japan

D. Roy Mahapatra

S. Gopalakrishnan

Department of Aerospace Engineering,
Indian Institute of Science,
Bangalore 560012,
India

In this paper, a novel numerical technique based on the global-local hybrid spectral element (HSE) method is proposed to study wave propagation in beams containing damages in the form of transverse and lateral cracks. The ordinary spectral element method is employed to model the exterior or far field regions, while a new type of element (HSE) is constructed to model the interior region containing damages. To develop this efficient new element for the damaged area, first, the flexural and the shear wave numbers are explicitly determined using the first-order shear deformation theory. These wave modes, in one of the two mutually orthogonal directions for two-dimensional transient elastodynamics, are then used to enrich the Lagrangian interpolation functions in context of displacement-based finite element. The equilibrium equation is then derived through the weak form in the frequency domain. Frequency-dependent stiffness and mass matrices can be accurately formed in this manner with a coarse discretization. The proposed method takes the advantage of using (i) a strong form for one-dimensional wave propagation and also (ii) a weak form by which a complex geometry can be discretized. Numerical verification is carried out to illustrate the effectiveness of the method. Finally, this method is employed to investigate the behaviors of wave propagation in beams containing various types of damages, such as multiple transverse cracks and lateral cracks. [DOI: 10.1115/1.2188015]

1 Introduction

In order to improve the safety, reliability, and operational life, it is an urgent task to monitor the integrity of structures. Therefore, the availability of efficient techniques for nondestructive damage detection is essential. To reduce the human interaction, while monitoring the integrity of structures, many works have been published based on the traditional modal analysis techniques [1–3]. These low-frequency techniques employ the information related to the changes in the modal signature due to the presence of a comparatively large size of damages. On the other hand, it is quite difficult to detect the small and local defects using the traditional modal analyses, since the effects of small damages on the low-frequency global responses are often impossible to amplify. Diagnosis of small cracks is necessary because under certain loading conditions they can grow rapidly and may lead to catastrophic failure of the structure before any precautionary measure can be taken. In the recent years, various new and efficient techniques have appeared using the high-frequency characteristics provided by smart materials, such as piezoelectrics. These approaches employ the high-frequency dynamic responses, such as Lamb wave propagation, to monitor the presence of small defects. A lot of works have mainly been concentrated on the experimental investigations [4–6]. Some novel theoretical studies have also been performed. For instance, Karim et al. [7] studied the Lamb wave scattering from cracks and inclusions in a plate due to a vertical Gaussian beam load using a hybrid finite element method (FEM) and normal function expansion method. Mal and Chang [8] investigated the scattering of Lamb waves from rivet holes and cracks in plates by using hybrid frequency domain FEM and normal mode expansion (called global-local FEM technique) followed by inverse fast Fourier transformation (FFT) to obtain the scattered field in time domain. Moreover, some recent researches [9–12] are focused on the improvement of the spectral element method for cracks and delaminations in beam structures. The spectral element method utilizes the exact solution to the strong form of the elastodynamics for finite element interpolation at discrete frequencies. Computationally, it is a very efficient and powerful method for analyzing the high-frequency responses.

In this paper, the authors propose a new hybrid spectral method based on the weak formulation of the finite two-dimensional (2D) transient elastodynamic problem. Two different types of interpolation bases are used in two orthogonal directions: (i) the superposed harmonic wave-type solution using wave vector (**k**) from an assumed kinematic theory (e.g., first-order shear deformation theory for the present beam problem), and (ii) Lagrange family of interpolation. While dealing with the wave propagation problem in complex geometry, the above description of displacement field has definite advantage over the exact spectral interpolation. As a consequence of the hybrid interpolation using (i) and (ii), a weak formulation of the boundary value problem becomes necessary. Therefore, we employ a frequency-domain variational approach and derive the frequency-dependent dynamic stiffness matrix, the mass matrix, and the consistent load vector. Transverse and horizontal cracks are modeled as interelement discontinuity. Ordinary spectral elements are used to model the "exterior regions" or far field regions. An appropriate number of hybrid spectral elements (HSEs) is used for modeling the region having cracks, i.e., the "interior region." Displacement continuity and equilibrium of forces between the "interior region" and "exterior region" are modeled using the global-local approach [7,8]. Numerical ex-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 16, 2004; final manuscript received January 18, 2006. Review conducted by O. M. O'Reilly. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

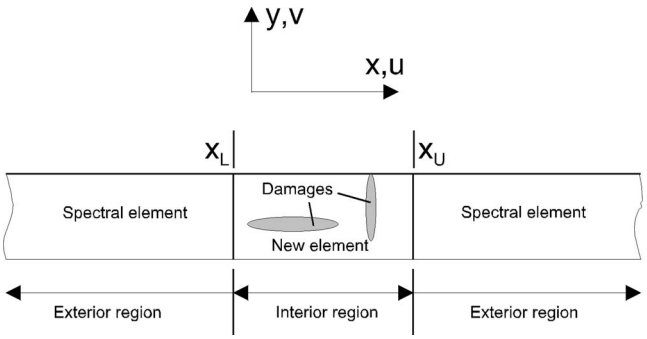


Fig. 1 Schematic diagram of the problem geometry

amples are shown to demonstrate the effectiveness of the HSEs and to study the behavior of wave propagation in the cracked beams.

2 Field Interpolation

2.1 Solution to One-Dimensional Wave Dispersion. In this section, the authors give a brief outline of the steps involved in obtaining the spectral family of interpolation functions by solving the characteristic system in terms of the wave number (k). These functions are then used to interpolate the displacements along the longitudinal direction (x) of the beam (see Fig. 1). In this paper, while arriving at the wave dispersion model, the authors apply first-order shear deformation theory (FSDT) for the beam considered as an equivalent single layer. The formulations given in Mahapatra and Gopalakrishnan [9] are employed.

First, the beam cross-sectional stiffness and inertial coefficients are defined as follows [9]:

$$[A_{kl},B_{kl},D_{kl}] = \sum_i \int_{z_i}^{z_{i+1}} \bar{Q}^i_{kl}[1,z,z^2]b_w dz \tag{1a}$$

$$[I_0,I_1,I_2] = \sum_i \int_{z_i}^{z_{i+1}} \rho[1,z,z^2]b_w dz \tag{1b}$$

where z is the thickness coordinate perpendicular to the beam reference plane, b_w the width of the beam, ρ the mass density, and $\bar{Q}^i_{kl}$ the elastic constitutive tensor of the ith layer. In addition, the following parameters are introduced

$$k_a = \frac{\omega_m}{c_a}, \quad k_b = \frac{\omega_m}{c_b}, \quad k_r = \frac{\omega_m}{c_s} \tag{2a}$$

$$c_a = \sqrt{\frac{A_{11}}{I_0}}, \quad c_b = \sqrt[4]{\frac{D_{11}\omega_m^2}{I_0}}, \quad c_s = \sqrt{\frac{A_{55}}{I_0}} \tag{2b}$$

$$r = \sqrt{\frac{B_{11}^2}{(A_{11}D_{11})}}, \quad s_1 = \omega_m\sqrt{\frac{I_2}{A_{55}}}, \quad s_2 = \sqrt{\frac{I_1^2}{(I_0I_2)}} \tag{2c}$$

For the coupled axial-flexural-shear deformation, without the thickness contractional modes, the wave numbers (k_j) associated with the individual wave modes at frequency ω_m are determined by solving the following equation:

$$\begin{vmatrix} (k_j^2-k_a^2) & 0 & \left(\frac{s_1s_2k_a^2}{k_r}-\frac{rk_ak_j^2}{k_r^2}\right) \\ 0 & (k_j^2-k_r^2) & -ik_j \\ \left(\frac{rk_r^2k_j^2}{k_ak_b^2}-s_1s_2k_r\right) & -ik_j & \left(s_1^2-1-\frac{k_r^2k_j^2}{k_b^4}\right) \end{vmatrix} = 0 \tag{3}$$

The characteristic system given by Eq. (3) can be expressed as follows:

$$ak_j^6 + bk_j^4 + ck_j^2 + d = 0 \tag{4}$$

where

$$a = 1 - r^2 \tag{5a}$$

$$b = \frac{2rs_1s_2k_ak_b^2}{k_r} - (1-r^2)k_r^2 - \frac{s_1^2k_b^4}{k_r^2} - k_a^2 \tag{5b}$$

$$c = k_a^2k_r^2 - 2rs_1s_2k_ak_rk_b^2 - (1-s_1^2)k_b^4 + \frac{s_1^2(1-s_2^2)k_a^2k_b^4}{k_r^2} \tag{5c}$$

$$d = [1-s_1^2(1-s_2^2)]k_a^2k_b^4 \tag{5d}$$

The six roots of Eq. (4) are given by

$$k_{1,2} = \pm\sqrt{\sqrt[3]{-\frac{q}{2}+\sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{2}\right)^3}}+\sqrt[3]{-\frac{q}{2}-\sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{2}\right)^3}}-\frac{b}{3a}} \tag{6a}$$

$$k_{3,4} = \pm\sqrt{\mu_2\sqrt[3]{-\frac{q}{2}+\sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{2}\right)^3}}+\mu_1\sqrt[3]{-\frac{q}{2}-\sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{2}\right)^3}}-\frac{b}{3a}} \tag{6b}$$

$$k_{5,6} = \pm\sqrt{\mu_1\sqrt[3]{-\frac{q}{2}+\sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{2}\right)^3}}+\mu_2\sqrt[3]{-\frac{q}{2}-\sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{2}\right)^3}}-\frac{b}{3a}} \tag{6c}$$

where

$$p = \frac{(27ca^2-9ab^2)}{27a^3}, \quad q = \frac{(27da^2-9abc+2b^2)}{27a^3} \tag{7a}$$

$$\mu_1 = \frac{-1+i\sqrt{3}}{2}, \quad \mu_2 = \frac{-1-i\sqrt{3}}{2} \tag{7b}$$

The three pairs of wave numbers in Eqs. (6a)–(6c) are distinct (except at few frequencies where cross-over may occur), since they represent the axial, the flexural and the shear modes, respec-

tively. The axial mode must strictly have real wave numbers. The reason can be explained as follows: first, assume that the wave numbers for axial modes are complex. Then they should exist in conjugate form, i.e., $k_1=a+ib$, $k_2=a-ib$. Since the bulk wave equations are of second order, even when they are coupled with flexural waves, they propagate with forward and backward wave components. Hence, the associated wave vector, whose amplitude is the wave number, must have an opposite sign. The real part of the wave number must have an opposite sign (the imaginary part is attenuation). Therefore, one must have $\mathrm{Re}[k_1]=\mathrm{Re}[a+ib]$, $\mathrm{Re}[k_2]=\mathrm{Re}[a-ib]=-\mathrm{Re}[k_1]=-\mathrm{Re}[a+ib]$. The last equation is a contradiction. Consequently, the assumption that the bulk wavenumbers are complex is not true.

Depending on the material configuration and geometry, the other modes may have complex wave numbers. Note that the shear mode starts propagating only after the following cutoff frequency [9]:

$$\omega_{\text{cutoff}} = \sqrt{\frac{A_{55}}{I_2(1-s_2^2)}} \tag{8}$$

When $\omega_m < \omega_{\text{cutoff}}$, the shear waves are evanescent in nature. Usually, if there are unsymmetric damages in beams, the components in the reflected waves contain the axial wave mode. However, in this paper, the authors deal with the damages, which are only symmetric with respect to the middle plane of the beam, i.e., where only the flexural and the shear waves need to be considered.

2.2 Enriched Interpolation of Displacement Field. After transforming the longitudinal displacement $u(x,y,t)$ and the transverse displacement $v(x,y,t)$ (see Fig. 1) from the time domain to the frequency-domain, their frequency domain counterpart are expanded as linear combination of orthogonal basis functions. The bases for interpolation parallel to the x-axis are the four wave modes $\psi_i, i=1,2,3,4$ (the forward and the backward propagating flexural wave modes, and the forward or the backward propagating/evanescent shear wave modes). The bases for interpolation parallel to the y-axis are the Lagrangian family of interpolation functions. Thus, the displacement components of u and v in the frequency domain are expressed as

$$\bar{u}_m(x,y,\omega_m) = \sum_{j=1}^{n}\sum_{l=1}^{NW} N_j(x,y)\psi_l(x)A_j^l \tag{9a}$$

$$\bar{v}_m(x,y,\omega_m) = \sum_{j=1}^{n}\sum_{l=1}^{NW} N_j(x,y)\psi_l(x)B_j^l \tag{9b}$$

where A_j^l and B_j^l are the boundary dependent coefficients for each of the wave modes $l=1,\ldots,4$. n is the number of elemental nodes, NW is the number of used wave numbers (k_1, k_2, k_3, and k_4), and N_j is the standard FE Lagrangian interpolation functions, i.e., the "old shape functions." In other words, ψ_l are the spectral enrichment functions over N_j. For the present beam problem, they are described as

$$\psi_1 = e^{-i[k_1(x-x_L)]}, \quad \psi_2 = e^{i[k_2(x_U-x)]} \tag{10a}$$

$$\psi_3 = e^{i[k_3(x-x_L)]}, \quad \psi_4 = e^{-i[k_4(x_U-x)]} \quad \forall\ \omega < \omega_{\text{cutoff}} \tag{10b}$$

$$\psi_3 = e^{-i[k_3(x-x_L)]}, \quad \psi_4 = e^{i[k_4(x_U-x)]} \quad \forall\ \omega \geqslant \omega_{\text{cutoff}} \tag{10c}$$

where k_1 and k_2 are the wave numbers of flexural waves, which are positive real, and negative real numbers, respectively. k_3 and k_4 are wave numbers of shear wave, which are positive imaginary and negative imaginary numbers when the frequencies are smaller than ω_{cutoff}. x_L and x_U denote the x coordinates of the left and right cross-sections of the interior region, respectively, as shown in Fig. 1. Generally, it is convenient to define them in the local coordinate system.

The enrichment functions ψ_l in Eqs. (10) occur in pairs, where one member of the pair is the "mirror image" of the other. In fact, the origin of this idea for the displacement field in Eqs. (9) can be traced to recent work for solving Helmholtz equations [13,14], where at each of the FE nodes, the potential is expanded into one or multiple discrete series of plane waves. However, no work has been reported for elastic waves in bounded media.

To prevent transverse shear locking while applying FSDT, the order of approximation has to be at least quadratic along the y-axis. Therefore, 8-noded or 12-noded isoparametric shape functions are used as the "old shape functions." Rewriting the enriched interpolation functions as $P_{j,l}=N_j\psi_l$, the displacement field can be expressed as

$$\begin{Bmatrix} u(x,y,t) \\ v(x,y,t) \end{Bmatrix} = \sum_{m=1}^{M} \aleph \bar{\mathbf{u}} e^{-i\omega_m t} \tag{11}$$

where M is the number of sampling points in the frequency domain, and

$$\aleph = \begin{bmatrix} \Re_1 & \mathbf{0} & \Re_2 & \mathbf{0} & \cdots & \Re_n & \mathbf{0} \\ \mathbf{0} & \Re_1 & \mathbf{0} & \Re_2 & \cdots & \mathbf{0} & \Re_n \end{bmatrix} \tag{12}$$

where

$$\Re_j = [P_{j,1}\ P_{j,2}\ \cdots\ P_{j,NW}] \tag{13}$$

$\bar{\mathbf{u}}$ is a vector consisting of nodal unknown coefficients,

$$\bar{\mathbf{u}} = \{\alpha_1\ \beta_1\ \cdots\ \alpha_n\ \beta_n\}^T \tag{14}$$

where

$$\alpha_j = \{A_j^1\ A_j^2\ \cdots\ A_j^{NW}\} \tag{15a}$$

$$\beta_j = \{B_j^1\ B_j^2\ \cdots\ B_j^{NW}\} \tag{15b}$$

3 Weak Formulation in Frequency Domain

With the above description of the enriched displacement field, the Hamiltonian for the two-dimensional elastodynamics is given by

$$\Pi_P = \int_{t_1}^{t_2} \left[\frac{1}{2}\int\int_{\Omega} \boldsymbol{\varepsilon}^T \mathbf{D}\boldsymbol{\varepsilon}\, dxdy - \int_S \mathbf{q}^T\delta ds - \sum_{k=1}^{K} \mathbf{P}_k^T\delta_k - \frac{1}{2}\int\int_{\Omega} \frac{d\delta^T}{dt}\vartheta\frac{d\delta}{dt}dxdy \right] dt \tag{16}$$

where t_1 and t_2 are the starting time and ending time, respectively, $\mathbf{D}$ is the elasticity matrix for the plane stress model adopted here, and δ is a vector containing the two displacement components and is defined as

$$\delta = \begin{Bmatrix} u \\ v \end{Bmatrix} \tag{17}$$

The diagonal matrix ϑ for mass density ρ can be expressed as

$$\vartheta = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} \tag{18}$$

The terms $\mathbf{P}_k$ and $\mathbf{q}$ in Eq. (16) are the applied concentrated loads and applied distributed loads, respectively. The strain vector $\boldsymbol{\varepsilon}$ in Eq. (16) is defined in the following form

$$\boldsymbol{\varepsilon} = \begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{xy} \end{Bmatrix} = \begin{Bmatrix} \dfrac{\partial u}{\partial x} \\ \dfrac{\partial v}{\partial y} \\ \dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x} \end{Bmatrix} = \sum_{m=1}^{M} \Theta \bar{\mathbf{u}} e^{-i\omega_m t} \tag{19}$$

where

$$\Theta = \begin{bmatrix} \dfrac{\partial \Re_1}{\partial x} & \mathbf{0} & \dfrac{\partial \Re_2}{\partial x} & \mathbf{0} & \cdots & \dfrac{\partial \Re_n}{\partial x} & \mathbf{0} \\ \mathbf{0} & \dfrac{\partial \Re_1}{\partial y} & \mathbf{0} & \dfrac{\partial \Re_2}{\partial y} & \cdots & \mathbf{0} & \dfrac{\partial \Re_n}{\partial y} \\ \dfrac{\partial \Re_1}{\partial y} & \dfrac{\partial \Re_1}{\partial x} & \dfrac{\partial \Re_2}{\partial y} & \dfrac{\partial \Re_2}{\partial x} & \cdots & \dfrac{\partial \Re_n}{\partial y} & \dfrac{\partial \Re_n}{\partial x} \end{bmatrix} \tag{20}$$

To obtain the strain-displacement matrix Θ, the derivatives in Eq. (20) are calculated as follows:

$$\frac{\partial P_{j,l}}{\partial x} = \left[\frac{\partial N_j}{\partial x} + ik_l N_j \right] \psi_l \tag{21a}$$

$$\frac{\partial P_{j,l}}{\partial y} = \left[\frac{\partial N_j}{\partial y} \right] \psi_l \tag{21b}$$

The kinetic energy term in Eq. (16) can further be expanded as

$$\Gamma = \int_{t_1}^{t_2} \frac{1}{2} \int \int_{\Omega} \frac{d\delta^T}{dt} \vartheta \frac{d\delta}{dt} dxdydt = \frac{1}{2} \int \int_{\Omega} \frac{d\delta^T}{dt} \vartheta \delta dxdy \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{1}{2} \int \int_{\Omega} \delta^T \vartheta \frac{d^2\delta}{dt^2} dxdydt \tag{22}$$

where

$$\frac{d^2\delta}{dt^2} = \begin{Bmatrix} \dfrac{d^2u}{dt^2} \\ \dfrac{d^2v}{dt^2} \end{Bmatrix} = -\sum_{m=1}^{M} \omega_m^2 \begin{Bmatrix} \bar{u}_m \\ \bar{v}_m \end{Bmatrix} e^{-i\omega_m t} = -\sum_{m=1}^{M} \omega_m^2 \aleph \bar{\mathbf{u}} e^{-i\omega_m t} \tag{23}$$

Subsequently, similar to the displacement field in Eq. (11), by employing FFT, the distributed load $\mathbf{q}$ and concentrated load $\mathbf{P}_k$ in the time domain are described as

$$\mathbf{q}(x,y,t) = \sum_{m=1}^{M} \bar{\mathbf{q}}(x,y,\omega_m) e^{-i\omega_m t},$$

$$\mathbf{P}_k(x_k,y_k,t) = \sum_{m=1}^{M} \bar{\mathbf{P}}_k(x_k,y_k,\omega_m) e^{-i\omega_m t} \tag{24}$$

For the convenience of description, the local elemental variables are transformed into the structural global ones as $\bar{\mathbf{u}} = T\bar{\mathbf{U}}$. Finally, Eq. (16) for an arbitrary element can be rewritten as

$$\Pi_P = \int_{t_1}^{t_2} \sum_{m=1}^{M} \left[\left(\frac{1}{2} \int \int_{\Omega} \bar{\mathbf{U}}^T T^T \Theta^T \mathbf{D} \Theta T \bar{\mathbf{U}} dxdy - \int_S \bar{q}^T \aleph T \bar{\mathbf{U}} ds - \sum_{k=1}^{K} \bar{\mathbf{P}}_k^T \aleph^k T \bar{\mathbf{U}} - \frac{1}{2} \omega_m^2 \int \int_{\Omega} \bar{\mathbf{U}}^T T^T \aleph^T \vartheta \aleph T \bar{\mathbf{U}} dxdy \right) \times (e^{-i\omega_m t})^2 \right] dt - \frac{1}{2} \int \int_{\Omega} \frac{d\delta^T}{dt} \vartheta \delta dxdy \Big|_{t_1}^{t_2} \tag{25}$$

Using a variational principle, the finite element equilibrium equation is obtained as

$$(\bar{\mathbf{K}} - \omega_m^2 \bar{\mathbf{M}}) \bar{\mathbf{U}} = \bar{\mathbf{P}} \tag{26}$$

in which, the complex matrices are expressed as follows:

$$\bar{\mathbf{K}} = \sum_{1}^{\mathrm{NE}} \int \int_{\Omega_e} T^T \Theta^T \mathbf{D} \Theta T dxdy \tag{27a}$$

$$\bar{\mathbf{M}} = \sum_{1}^{\mathrm{NE}} \int \int_{\Omega_e} T^T \aleph^T \vartheta \aleph T dxdy \tag{27b}$$

$$\bar{\mathbf{P}} = \sum_{1}^{\mathrm{NE}} \int_S \bar{\mathbf{q}}^T \aleph T ds - \sum_{k=1}^{K} \bar{\mathbf{P}}_k^T \aleph^k T \tag{27c}$$

where NE is the number of elements. Thus, the proposed HSE is obtained. The main advantage of the above formulation is that a complex distribution in the y-axis can be modeled easily due to Lagangian FE interpolation, and a long but finite span in the x-axis can be modeled accurately using a much smaller number of these new elements compared to the standard Lagrangian FE model. Moreover, the near-exact feature of wave propagation characteristics along the x-axis is ensured, which leads to spectral convergence for interpolation parallel to the x-axis. Subsequently, the h convergence can be expected for interpolation parallel to the y-axis, since the wave modes are only functions of the x-axis.

4 Numerical Implementation

In the elemental matrices in Eqs. (27*a*)–(27*c*), the integrals encountered are of the form

$$I_{lm} = \int_{-1}^{1} \int_{-1}^{1} f(\xi,\eta) e^{i(k_l x)} e^{i(k_m x)} |\mathbf{J}| d\xi d\eta \tag{28}$$

where $\mathbf{J}$ is the Jacobian matrix for isoparametric mapping. The expression $f(\xi,\eta)$ involves the product of the "old shape function," their derivatives, etc. In this section, the evaluation of these integrals is carried out using higher-order Gauss-Legendre integration scheme. The number of integration points depends on the element nodal spacing with respect to the smallest wavelength ($\lambda_l = 2\pi/k_l$) at a given frequency ω_m. As will be shown later, only a few of the proposed HSEs in the spanwise direction of the beam, and hence, a few degrees of freedom can produce sufficient accuracy even for high-frequency excitation if the proper numerical integration scheme is used. Therefore, unlike the Lagrangian FE method, the computational time is mainly consumed at the step of numerical integration while calculating the dynamic stiffness matrix and the mass matrix, but not at the step of FE system solution. Numerical implementations of the proposed HSEs require efficient integration algorithms, such as frequency-dependent and elemental size-dependent numerical integration scheme. Consider a section with unit length, i.e., $x \in [0,1]$, the number of cycles of sin and cos functions within the domain is around $0.1k_1$. Therefore, in order to integrate one single cycle accurately, at least 30 Gauss integration points are needed. For example, for more accuracy, if 60 Gauss integration points are taken for one cycle of sin and cos functions, the number of Gauss integration points (NG) for the element i in the x-axis is roughly determined as $\mathrm{NG}^i = 6L_e^i k_1^i(\omega_m)$, where $k_1^i(\omega_m)$ is the flexural wave number of the element i at frequency ω_m, and L_e^i is the length along the x-axis of the element i. However, the minimum number of Gauss integration points is set to be 6. Such a consistent choice of frequency-dependent and elemental size-dependent integration scheme can efficiently reduce the computational cost. In the direction parallel to the y-axis, the number of Gauss integration points is set to be 2 for eight-noded element and 3 for 12-noded element, respectively.

Having obtained $\bar{\mathbf{U}}$ containing A_j^l and B_j^l, the displacements in the frequency domain can be calculated using Eq. (9), therefore, the solution in the time domain is evaluated simply by applying the inverse FFT.

4.1 Assemblage of Interior and Exterior Regions Using Global-Local Approach. As will be verified later from the numerical examples, the proposed HSEs are very efficient for the quasi-two-dimensional waves propagation at low, as well as high frequencies. However, at high frequencies, the accuracy is not so high when the forward and the backward propagation exist simultaneously and the propagation distance is long. The main reason is the dissimilar orders in the diagonal elements of the stiffness and the mass matrices, which are contributed by the exponential terms, i.e., the terms determined by wave numbers k_3 and k_4. For high-frequency waves, the components of stiffness and mass matrices, which involve $\psi_3 = e^{i[k_3(x-x_L)]}$ in Eq. (9) for large x, and $\psi_4 = e^{-i[k_4(x_U-x)]}$ in Eq. (9) for large x_U and small x, are too small compared to the other elements in the matrices. These small elements, especially small diagonal elements of the stiffness and mass matrices, may cause the numerical instability. For instance, to deal with the boundary conditions for propagation in both directions, the forward and backward terms are coupled at the boundary. When x is large at the boundary, the forward terms involving $\psi_3 = e^{i[k_3(x-x_L)]}$ are very small. On the other hand, the backward terms involving $\psi_4 = e^{-i[k_4(x_U-x)]}$ are well conditioned. The numerical instability is thus inherent.

For the cases of unidirectional propagation of waves, e.g., only forward propagation alone, the elements of the stiffness matrix and the mass matrix involve only the term $\psi_3 = e^{i[k_3(x-x_L)]}$. As stated above, although this term has the similar exponential characteristics as x increases, by collocating the diagonal elements corresponding to the exponential decay in the matrix $(\bar{\mathbf{K}} - \omega_m^2 \bar{\mathbf{M}})$ in a decreasing sequence, the numerical instability can be effectively removed. However, this technique is not suited for the cases of the simultaneous propagation of forward and backward waves. For such double directional propagation of high-frequency waves, the present element is only applicable for short traveling distance. In fact, from the authors' numerical experience, the existence of the exponential decay causes many numerical troubles. To overcome this problem, the new HSE needs to be combined with the ordinary spectral element model or any other efficient discretized model of the exterior region. As shown in Fig. 1, this new HSE is employed for the smaller interior region, and the ordinary spectral elements are used to model the exterior region. Another purpose of assembling the ordinary spectral element for the uniform exterior region is to reduce the computational cost to a greater extent.

In the global-local approach, the displacement continuity at the boundary of the two regions must be ensured. With the help of Eq. (9) and by setting $N_j = 1$, the condition of displacement continuity between the interior and the exterior regions (see Fig. 1) at the node j in an arbitrary HSE can be expressed as

$$\bar{u}_m(x, y, \omega_m) = \sum_{l=1}^{NW} \psi_l A_j^l = -y\bar{\phi}(x, \omega_m) \tag{29a}$$

$$\bar{v}_m(x, y, \omega_m) = \sum_{l=1}^{NW} \psi_l B_j^l = \bar{w}(x, \omega_m) \tag{29b}$$

where $\bar{\phi}(x, \omega_m)$ and $\bar{w}(x, \omega_m)$ are the rotation and the transverse displacement degrees of freedom in the spectral element (see [9] for details).

There exist three different approaches for enforcing the interface displacement continuity in context of global-local finite element analysis. The first approach is the direct enforcement of the constraints at the interfaces as reported by Gopalakrishnan and Doyle [15]. The second approach is the weak enforcement using Lagrangian multiplier as reported by Halliday and Grosh [16].

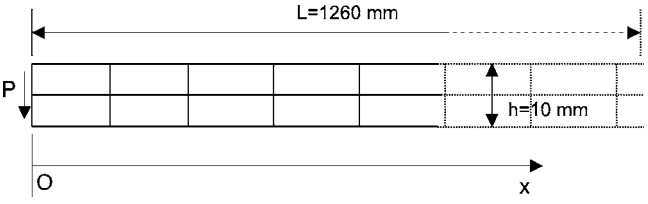


Fig. 2 Schematic diagram of a 1D problem

Finally, the third approach is the weak enforcement based on multipoint constraints as reported by Mahapatra and Gopalakrishnan [11]. While using two dissimilar models, e.g., standard Lagrangeian FE model for the interior region and the ordinary spectral element model for the exterior region, direct enforcement of interface constraints [15] appears computationally intensive and requires cross-check against numerical convergence. Again, from the authors' numerical experiences, it is found that the penalty function method [11] is not efficient while using dissimilar models for the exterior and the interior regions, since the choice of penalty parameters is very difficult and the introduction of large penalty parameters may result into the numerical instability. In the present global-local approach, the Lagrangian multiplier method is employed to enforce the displacement continuity conditions given in Eqs. (29*a*) and (29*b*).

5 Numerical Examples

5.1 Comparison to Traditional FEM. Consider a one-dimensional (1D) beam shown in Fig. 2, which is subjected to a transverse load $P(t)$ at the free end, i.e., $x = 0$, which is expressed as

$$P(t) = \begin{cases} 0.5[1 - \cos(2\pi ft/N)\sin(2\pi ft)], & t \leq N/f \\ 0, & t > N/f \end{cases} \tag{30}$$

where f is the central frequency in hertz and N is the number of sinusoidal cycles within a pulse.

To compare to the traditional FEM, the authors consider the low-frequency excitation case, where $f = 50$ Hz, and $N = 15$ cycles. The sample duration time T is 4.8 s. The Nyquist frequency is $N_q/2T$, where the number of sampling points is $N_q = 2^k, k = 12$, and is used consistently in the following several examples. The traditional 2D eight-noded isoparametric element is employed for comparison. The traditional FEM mesh possesses 84 elements with two elements in the thickness direction and 42 elements in the spanwise direction. Only six eight-noded HSEs are used with two elements in the thickness direction and three ele-

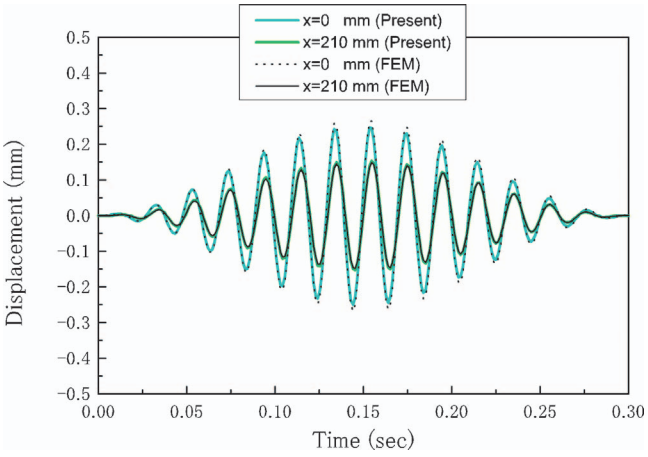


Fig. 3 Deflections of the present element and the traditional FEM at two measurement points

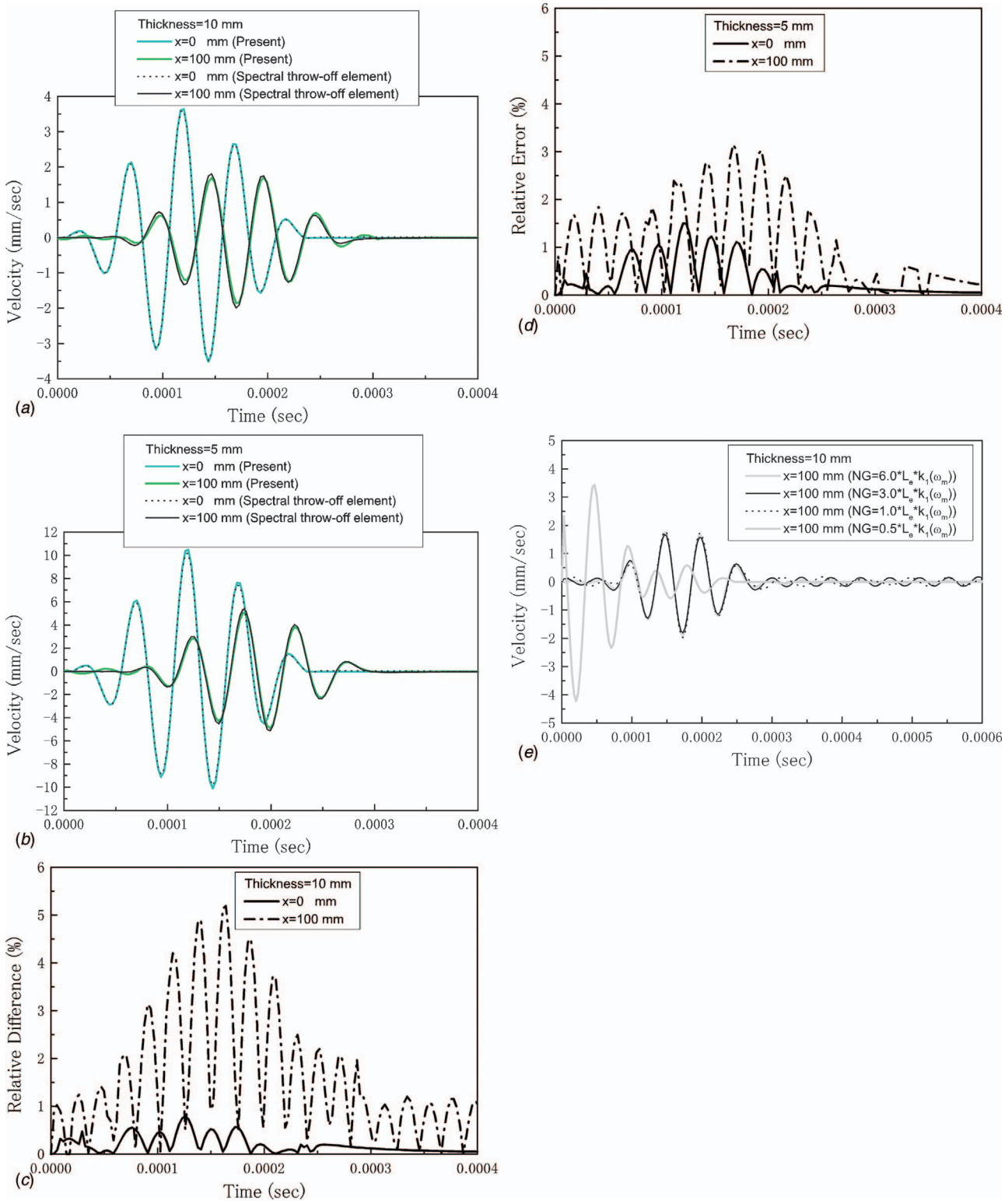


Fig. 4 (*a*) **Transverse velocities of the present element and the throw-off spectral element at two measurement points for a beam of thickness of 10 mm,** (*b*) **transverse velocities of the present element and the throw-off spectral element at two measurement points for a beam of thickness of 5 mm,** (*c*) **relative difference between velocities of the present element and the throw-off spectral element at two measurement points for a beam of thickness of 10 mm,** (*d*) **relative difference between velocities of the present element and the throw-off spectral element at two measurement points for a beam of thickness of 5 mm, and** (*e*) **convergence check of number of Gauss integration points along the *x*-axis for a beam of thickness of 10 mm**

ments in the spanwise direction. The same eight-noded HSEs will be used in all the following examples. The beam is made of aluminum with the following properties: $E=73.0$ GPa, $G=28.08$ GPa, $\nu=0.3$, and $\rho=2770$ kg/m^3. The thickness of the beam (h) is equal to 10.0 mm. The span of the beam for the present example is taken as 1260.0 mm. The width of the beam is taken as 10.0 mm. The above constants are consistently used in all the following examples. The responses at the far ends are assumed

to be zero, and the reflected wave modes are eliminated. To impose such absorbing boundary condition, the nodal degrees of freedom corresponding to the backward-propagating modes are set to be zero. It means that A_j^2, A_j^4, B_j^2, and B_j^4 in Eq. (9) are set to be zero at all the nodes. Furthermore, A_j^1, A_j^3, B_j^1, and B_j^3 in Eq. (9), which are the degrees of freedom corresponding to the forward propagation, are set to be zero at the nodes on the far end. Two measurement points are considered at $x=0.0$ mm, and $x=210.0$ mm away from the source of excitation. Comparison of the time histories of deflections at the measurement points is shown in Fig. 3. This result reveals that the present element can attain very high accuracy, although the mesh is very coarse. In fact, the present method is almost insensitive to the number of elements when a sufficiently higher-order Gauss integration scheme is employed in the x-axis.

5.2 Comparison With Throw-off Spectral Element. Consider a similar problem as shown in Fig. 2. The high-frequency case, where $f=20$ kHz and $N=5$ cycles is investigated. The sample duration time T is 0.012 s. The authors compare the performance of the present element to that of the throw-off spectral element [9], which only tackles the unidirectional wave propagation. In this example, the beam is considered to have enough length, i.e., 1200.0 mm. The response at the far end is imposed to be zero, and the reflected wave modes are excluded. Therefore, the unknown parameters, i.e., A_j^l and B_j^l are dealt in the same way as the previous example. The present approach employs 16 elements with two elements in the thickness direction and eight elements in the spanwise direction. For high-frequency cases, the amplitude of deflection becomes very small, and here the transverse velocity is selected for plotting. The transverse velocities at two measurement points, i.e., at $x=0.0$ mm, and $x=100.0$ mm are obtained. Comparison of the time histories of transverse velocity at these measurement points is shown in Fig. 4(*a*). It may be noted that the result obtained using the present element is in good agreement with those obtained using the spectral throw-off element. When the thickness of the beam is equal to 5.0 mm, the results of the two approaches are illustrated in Fig. 4(*b*). In this case, the amplitude of the transverse velocity increases and the velocity of traveling wave decreases due to the thinner beam. Furthermore, both results are in very good agreement. The absolute difference between the two kinds of results, which is normalized by the maximum absolute value in the results of spectral element, indicates relative difference between the results of the two methods. This relative difference is shown in Figs. 4(*c*) and 4(*d*). It is observed in these figures that the maximum relative difference is ~5% for 10 mm thickness in Fig. 4(*c*) and ~3% for 5 mm thickness in Fig. 4(*d*), respectively. This difference may be due to the fact that the method used here is for 2D plane problems, whereas the spectral element method is constructed from beam theory. The influence of the number of Gauss integration points on the results for $x=100.0$ mm, $h=10.0$ mm, is shown in Fig. 4(*e*), where, as stated before, for one single element, the number of Gauss integration points in the x-axis is calculated as: $NG^i=opL_e^i k_1^i(\omega_m)$ with a parameter op, but the minimum number of integration points is set to be 6. From this Fig. 4(*e*), it can be seen that $op=0.5$ cannot produce the correct result. With the increase in op, the result converges monotonically. In fact, the result for $op=3$ is almost identical to that of $op=6$.

5.3 Comparison to Ordinary Spectral Element. In this section, a cantilever beam as shown in Fig. 5(*a*) is considered and two ordinary spectral elements [9] are employed, i.e., one finite and the other throw-off. The material properties are the same as in the previous examples. $f=20$ kHz, and $N=5$ cycles are chosen in this example. The length of the beam from the fixed end to the point of application of transverse load is 2000.0 mm. As shown in Fig. 5(*b*), the entire domain in Fig. 5(*a*) is divided into three portions, i.e., two finite ordinary spectral elements with length of

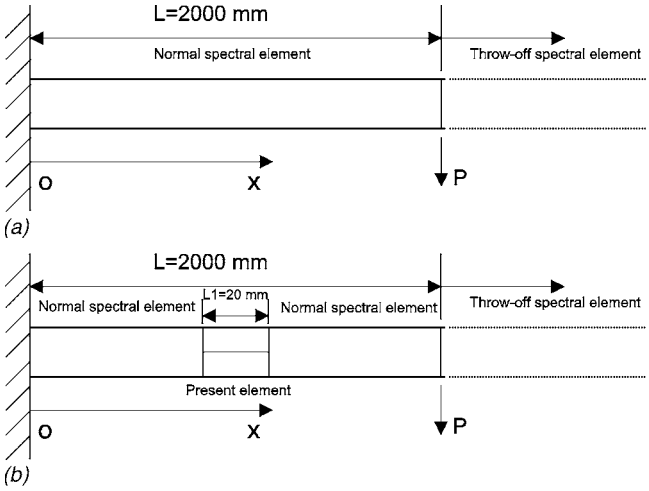


Fig. 5 (*a*) Schematic diagram of a cantilever beam using the spectral element only and (*b*) schematic diagram of a cantilever beam using the hybrid approach

990.0 mm, and one interior region with length of 20 mm discretized by the present HSEs. The two HSEs are only employed for the interior region between region 1 and region 2, where in the spanwise direction, only one HSE is used. The comparison of time histories of transverse velocity at the loading point is depicted in Fig. 6. This figure clearly shows both the incident and reflected waves. It can also be noted that both results, which are based on the meshes in Figs. 5(*a*) and 5(*b*), are in good agreement. However, there are some small oscillations while using the global-local approach. This phenomenon may be caused by the mismatch between the HSE dynamic stiffness and that of the ordinary spectral element due to the different assumptions for the displacement field interpolation. The fact is that there is no contraction in the thickness direction considered in the ordinary spectral element model. In fact, by observing Figs. 4(*c*) and 4(*d*), it can also be found that the results of the present HSE do not completely match with those of the throw-off spectral element. Generally, this mismatch does not cause a serious problem in the low-frequency cases. However, with the increase in the excitation frequency, the effects of mismatch become more obvious. From the authors' numerical experiences, increasing the number of the present HSEs along the spanwise and thickness directions to alleviate the effects of this mismatch is not effective.

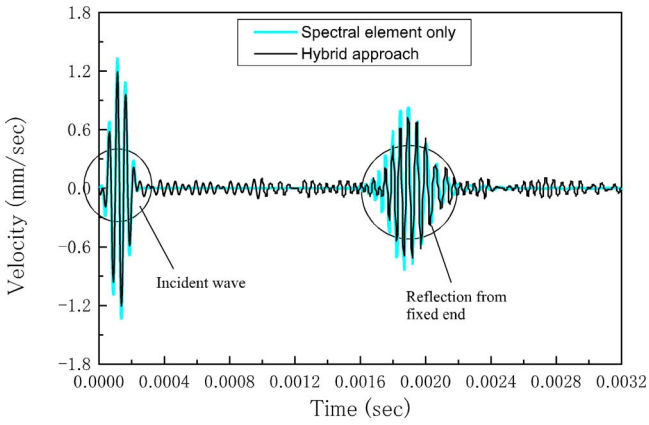


Fig. 6 Transverse velocities of the spectral element only and the hybrid approach at the load point

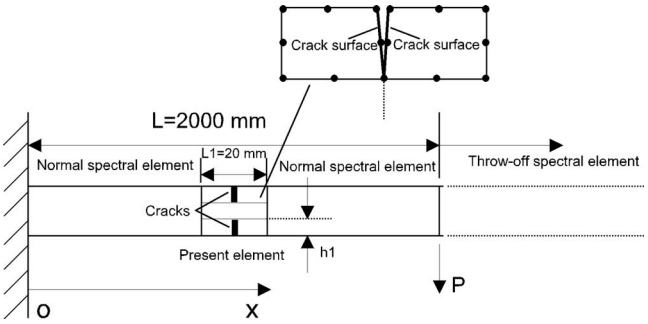


Fig. 7 Schematic diagram of a cantilever beam with two symmetric transverse cracks

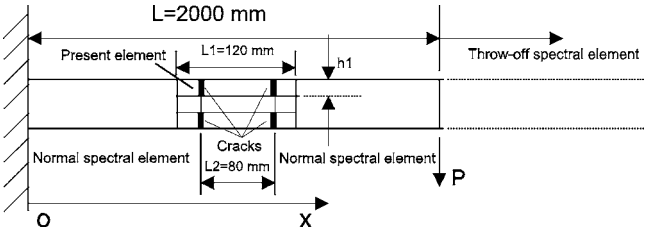


Fig. 9 Schematic diagram of a cantilever beam with four transverse cracks

5.4 Cantilever Beam With Two Symmetric Transverse Cracks. In this section, another example identical to the previous one is considered, except that there are two symmetric transverse cracks having depth $h1$ as shown in Fig. 7. Six present HSEs are employed to model the interior region. With the present elements, the crack geometry can be easily modeled. The placement of the element nodes at the interior region and the crack surfaces are shown in Fig. 7. The crack surfaces essentially form the interelement discontinuity. The effect of contact between the two crack surfaces is neglected. The time histories of transverse velocity of the intact and cracked beams ($h1=3.0$ mm) at the loading point are shown in Fig. 8(a). The first reflection from the cracks can be accurately identified in Fig. 8(a), which arrives at $t=0.001$ s. The first reflection from the fixed end can be identified for both the intact and cracked cases. However, the amplitude of the first reflection from the fixed end in the cracked beam is much lower than that of the intact beam. Furthermore, the second reflection from the fixed end can also be identified for the cracked beam. Here, the second reflection is the combination of (i) part of the first reflection from the fixed end, which is reflected by the cracks back to the fixed end and (ii) the second reflection stated in (i) by the fixed end, which finally arrives at the measurement point. This second reflection does not appear for the case of intact beam. When $h1=1.5$ mm, the results are shown in Fig. 8(b). It is observed that the amplitude of reflection from the cracks decreases and the amplitude of reflection from the fixed end increases as the crack depth decreases.

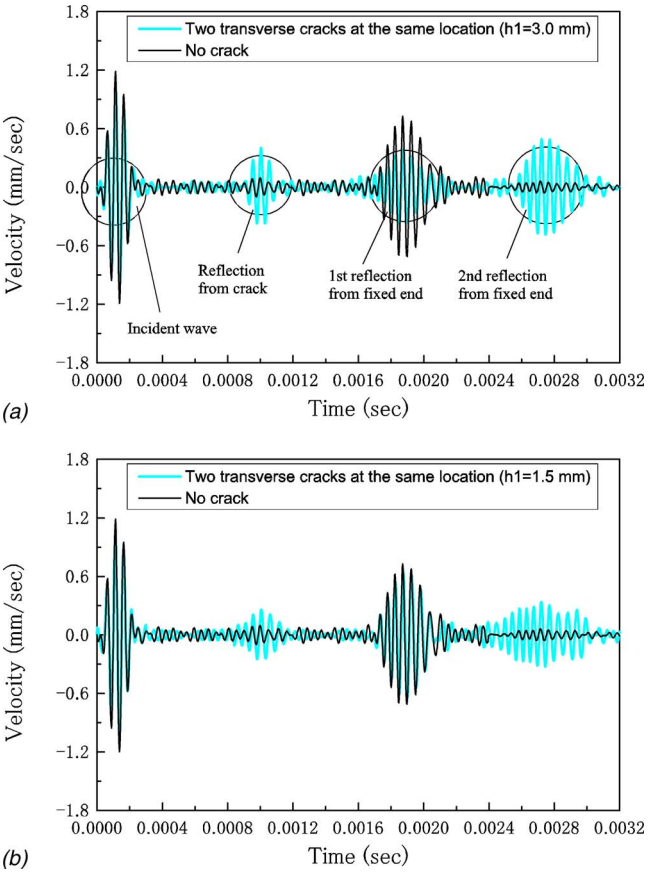


Fig. 8 (a) Transverse velocities of intact and cracked beams for $h1=3.0$ mm at the load point and (b) transverse velocities of intact and cracked beams for $h1=1.5$ mm at the load point

5.5 Cantilever Beam With Four Transverse Cracks. As shown in Fig. 9, four transverse cracks are considered. The interior region is 120 mm long and the distance between the two sets of transverse cracks is 80 mm. Nine present HSEs are employed for the interior region, and the crack depth $h1=3.0$ mm. Other conditions are identical to the above example. The transverse velocity at the loading point is shown in Fig. 10. It can be seen that the reflections from two sets of cracks overlap due to the short distance between the cracks. Compared to the results in Fig. 8(a) for two transverse cracks, here the amplitude of reflection from the four cracks is higher. However, the reflection from the fixed end has decreased significantly.

5.6 Cantilever Beam With a Lateral Crack at Midplane. Here, the case shown in Fig. 11 is considered, where a lateral crack is located at the midplane of the beam. Six present HSEs, with three elements in the spanwise direction and two elements in the thickness direction, are employed to model the interior region. Other conditions are identical to the above examples. The effect of contact between the two crack surfaces is neglected. Time histories of transverse velocity at the loading point for the two different crack lengths are shown in Fig. 12. Inspection of this figure reveals that the change in the shape of the scattered pulse is not so

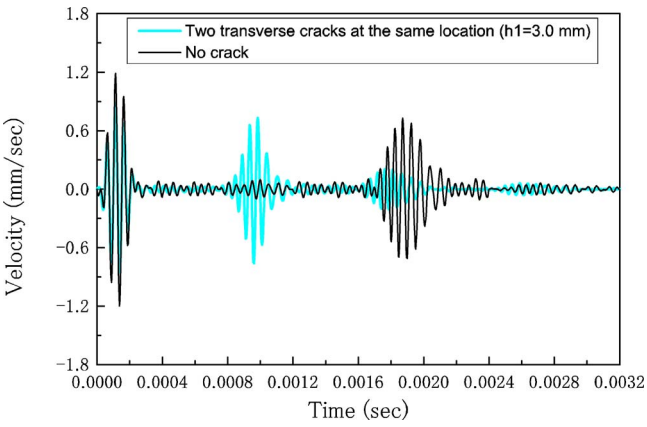


Fig. 10 Transverse velocities of intact and cracked beams at the load point

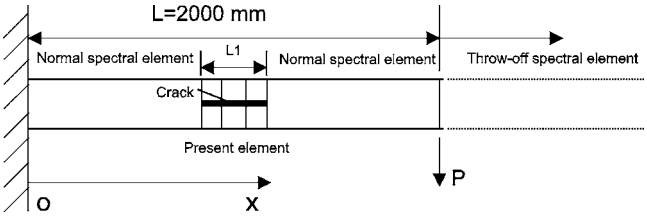


Fig. 11 Schematic diagram of a cantilever beam with a lateral crack at midplane

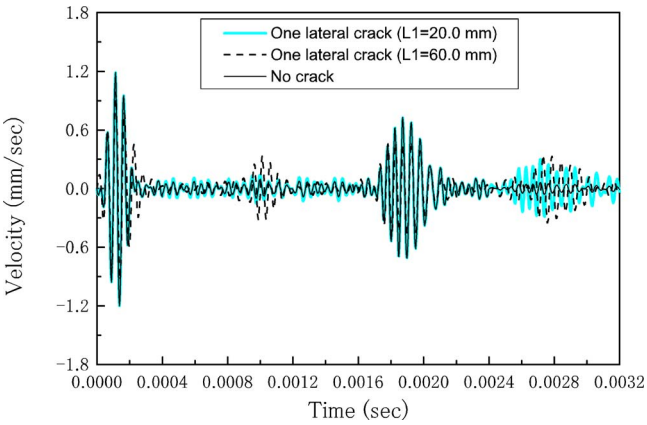


Fig. 12 Transverse velocities of intact and cracked beams at the load point

obvious compared to that of the intact beam, when the lateral crack is short, i.e., $L1=20$ mm. However, when $L1=60$ mm, the reflection from the lateral crack can be clearly identified. By comparing this result to those of the previous transverse cracks cases, it is observed that, under the present loading condition and frequency band of excitation, the scattered waves are not sensitive to the presence of lateral crack.

6 Conclusions

In this paper, a new global-local hybrid spectral element (HSE) method is presented to model the wave propagation in beams with various damages. In this approach, the ordinary spectral element method is employed to simulate the behavior of wave propagation in the exterior regions, meanwhile the proposed HSEs are used to model the interior region containing the damages in the form of cracks. Formulation of the proposed HSE is based on the hybrid interpolation scheme, where the Lagrangian family of interpolation bases is enriched by introducing four wave modes (two flexural wave modes and two shear wave modes) associated to the interpolation along one of the two mutually orthogonal coordinates for the 2D plane elastodynamic problem. Frequency-domain finite element model is then obtained by minimizing the Hamiltonian. The proposed approach essentially balances the advantages of (i) the ordinary spectral element model, which is the direct solution to the strong form, in which the solution is enforced in a point-wise sense, as well as (ii) the weakly formulated finite element, where the solution is enforced in a piecewise discrete sense. The advantage of (i) is in the high numerical efficiency while solving transient elastodynamics. The advantage of (ii) is in the efficient handling of complex geometry.

With this balanced approach, accurate analyses can be realized with a very coarse FE mesh. The proposed HSEs are highly efficient and hence can be used to model more complex problems. Some numerical examples are shown to illustrate the effectiveness of the present eight-noded element. Wave propagation in beams with various damages in the form of transverse and lateral cracks has been studied using the global-local technique. It is shown from the numerical results that the reflection from the damages can be identified, and the pattern of wave propagation in beams with complex crack configurations can be efficiently studied. Moreover, the wave response seems to be less sensitive to the presence of lateral crack compared to the presence of transverse crack.

References

[1] Cawley, P., Adams, R. D., Pye, C. J., and Stone, B. J., 1978, "A Vibration Technique for Non-Destructively Assessing the Integrity of Structures," J. Mech. Eng. Sci., **20**(2), pp. 93–100.
[2] Hu, N., Wang, X., Fukunaga, H., Yao, Z. H., Zhang, H. X., and Wu, Z. S., 2001, "Damage Assessment of Structures Using Modal Test Data," Int. J. Solids Struct., **38**, pp. 3111–3126.
[3] Fukunaga, H., Hu, N., and Chang, F. K., 2002, "Structural Damage Identification Using Piezoelectric Sensors," Int. J. Solids Struct., **39**, pp. 393–418.
[4] Maslov, K., and Kundu, T., 1997, "Selection of Lamb Modes for Detecting Internal Defects in Composite Laminates," Ultrasonics, **35**(2), pp. 141–150.
[5] Ghosh, T., Kundu, T., and Karpur, P., 1998, "Efficient Use of Lamb Modes for Detecting Defects in Large Plates," Ultrasonics, **36**(7), pp. 791–801.
[6] Toyama, N., and Takatsubo, J., 2004, "Lamb Wave Method for Quick Inspection of Impact-Induced Delamination in Composite Laminates," Compos. Sci. Technol., **64**(9), pp. 1293–1300.
[7] Karim, M. R., Awal, M. A., and Kundu, T., 1992, "Elastic Wave Scattering by Cracks and Inclusions in Plates: In Plane Case," Int. J. Solids Struct., **29**, pp. 2355–2367.
[8] Mal, A. K., and Chang, Z., 2000, "A Semi-Numerical Method for Elastic Wave Scattering Calculations," Geophys. J. Int., **143**, pp. 328–334.
[9] Mahapatra, D. R., and Gopalakrishnan, S., 2003, "A Spectral Finite Element Model for Analysis of Axial-Flexural-Shear Coupled Wave Propagation in Laminated Composite Beams," Compos. Struct., **59**, pp. 67–88.
[10] Nag, A., Mahapatra, D. R., and Gopalakrishnan, S., and Sankar, T. S., "A Spectral Finite Element With Embedded Delamination for Modeling of Wave Scattering in Composite Beams," Compos. Sci. Technol., (in press).
[11] Mahapatra, D. R., and Gopalakrishnan, S., 2004, "Spectral Finite Element Analysis of Coupled Wave Propagation in Composite Beams With Multiple Delaminations and Strip Inclusions," Int. J. Solids Struct., **41**, pp. 1173–1208.
[12] Krawczuk, M., Palacz, M., and Ostachowicz, W., 2003, "The Dynamic Analysis of a Cracked Timoshenko Beam by the Spectral Element Method," J. Sound Vib., **264**, pp. 1139–1153.
[13] Laghrouche, O., Bettess, P., and Astley, R. J., 2001, "Modelling of Short Wave Diffraction Problems Using Approximating Systems of Plane Waves," Int. J. Numer. Methods Eng., **54**, pp. 1501–1533.
[14] Melenk, J. M., and Babuška, I., 1996, "The Partition of Unity Finite Element Method: Basic Theory and Application," Comput. Methods Appl. Mech. Eng., **139**, pp. 289–314.
[15] Gopalakrishnan, S., and Doyle, J. F., 1995, "Spectral Super-Elements for Wave Propagation in Structures With Local Non-Uniformities," Comput. Methods Appl. Mech. Eng., **121**, pp. 77–90.
[16] Halliday, P. J., and Grosh, K., 1999, "Dynamic Response of Complex Structural Interconnections Using Hybrid Methods," ASME J. Appl. Mech., **66**, pp. 653–659.

Plane Analysis of Finite Multilayered Media With Multiple Aligned Cracks—Part I: Theory

Linfeng Chen

Marek-Jerzy Pindera

Civil Engineering Department,
University of Virginia,
Charlottesville, VA 22904-4742

Elasticity solutions are developed for finite multilayered domains weakened by aligned cracks that are in a state of generalized plane deformation under two types of end constraints. Multilayered domains consist of an arbitrary number of finite-length and finite-height isotropic, orthotropic or monoclinic layers typical of differently oriented, unidirectionally reinforced laminas arranged in any sequence in the plane in which the analysis is conducted. The solution methodology admits any number of arbitrarily distributed interacting or noninteracting cracks parallel to the horizontal bounding surfaces at specified elevations or interfaces. Based on half-range Fourier series and the local/global stiffness matrix approach, the mixed boundary-value problem is reduced to a system of coupled singular integral equations of the Cauchy type with kernels formulated in terms of the unknown displacement discontinuities. Solutions to these integral equations are obtained by representing the unknown interfacial displacement discontinuities in terms of Jacobi or Chebyshev polynomials with unknown coefficients. The application of orthogonality properties of these polynomials produces a system of algebraic equations that determines the unknown coefficients. Stress intensity factors and energy release rates are derived from dominant parts of the singular integral equations. In Part I of this paper we outline the analytical development of this technique. In Part II we verify this solution and present new fundamental results relevant to the existing and emerging technologies. [DOI: 10.1115/1.2201883]

1 Introduction

Multilayered media appear in many modern composites and advanced material systems, including protective coatings, laminated ceramics, microelectronic packages, layered nanofilms, functionally graded, and multifunctional materials and adhesive joints. A common failure mode present in such high-performance multilayered structures is the delamination of adjacent layers or spallation caused by interlaminar cracks or cracks parallel to the bounding horizontal surface. The interfacial crack problem, therefore, has attracted considerable attention aimed at improving the structural integrity and reliability of layered materials and structural components.

Investigations into analytical solutions of plane crack problems have been conducted by many researchers since the 1920s. In general, the various analytical approaches are based either on Muskhelishvili's [1,2] complex potential method or Sneddon's [3,4] Fourier transform and series techniques. A periodic array of cracks along a single horizontal row or a single vertical column in an infinite homogeneous and isotropic medium is amenable to the complex potential treatment by expressing the complex potential in terms of a singular integral that satisfies the required periodicity conditions, effectively reducing the problem to a single crack problem, Westergaard [5], Koiter [6], Sneddon and Srivastava [7], Lowengrub [8], as discussed by Sneddon and Lowengrub [9]. Erdogan [10] extended this approach to two collinear interacting cracks, while Rice and Sih [11] considered cracks along an interface separating two dissimilar half planes. Doubly periodic crack arrays were analyzed by Delameter et al. [12] using the singular integral equation and Green's function approach within a unit cell framework based on periodic boundary conditions.

For arbitrarily oriented interacting cracks, rigorous analytical treatment is difficult and thus different approximate approaches have been proposed. Isida [13] used a Laurent series representation of complex potentials and a perturbation technique for the problem of an infinite isotropic plate with randomly distributed cracks of arbitrary size and orientation. Chen [14] and Horii and Nemat-Nasser [15] proposed a pseudotraction method based on approximating tractions on interacting crack surfaces by suitable polynomials whose coefficients were obtained from the solution of a system of algebraic equations generated by the imposition of consistency conditions. Kachanov [16,17] used a similar but simplified approach based on approximating crack tractions by their averages. This approach was subsequently extended to anisotropic materials by Mauge and Kachanov [18]. Du and Aydin [19] used asymptotic crack-tip and far-field stress approximations and superposition to analyze interaction of crack arrays of various arrangements and geometrical configurations in infinite isotropic media under plane strain. Binienda [20] used the Fourier transform technique together with superposition to analyze an infinite isotropic plate with fully interacting multiple cracks subjected to farfield loading under plane stress or plane strain conditions. Shbeeb et al. [21] extended this approach by considering an inhomogeneous isotropic plate with exponentially varying Young's modulus along a fixed direction.

Erdogan and Gupta [22] and Erdogan [23] outlined a general solution procedure for a multilayered medium containing a single crack based on the Fourier transform technique, which allowed us to reduce the problem to a system of singular integral equations governing the crack opening displacement components. Numerical results were obtained for three-layer systems with a crack embedded in a homogeneous layer and between dissimilar layers. This approach was extended by Chatterjee et al. [24] and Chatterjee [25] to arbitrarily layered media with isotropic, orthotropic, and monoclinic plies containing multiple cracks along different interfaces using the local/global stiffness matrix approach. Nu-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 18, 2005; final manuscript received April 1, 2006. Review conducted by Y. Huang. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

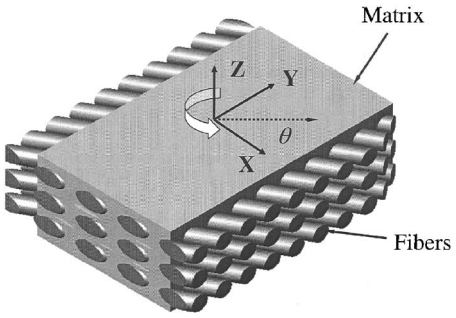


Fig. 1 A unidirectional composite layer rotated by an angle about axis perpendicular to its plane

merical results were limited to single delaminations due to considerable computational difficulties of dealing with interacting cracks. Pindera [26] illustrated the utility of this approach in analyzing interlaminar stress distributions in a bimaterial beam with two symmetrically positioned, noninteracting disbonds under three-point bending upon a comparison with experimental and finite-element results. Choi and Thangjitham [27] provided additional results following the approach of Chatterjee et al. [24] and Chatterjee [25].

In contrast, relatively little analytical work has been reported on multiple crack interaction in homogeneous or layered media with finite dimensions in the plane of analysis. Isida [28] used Laurent series representations of complex potentials together with a boundary collocation procedure to analyze width and length effects of rectangular plates on the stress intensity factors of centrally positioned cracks. Periodic arrays of cracks in plates with one finite dimension could be modeled using this approach by a suitable adjustment of the boundary conditions. Chatterjee [29] investigated two symmetric cracks in finite multilayered isotropic or orthotropic media pinned vertically at right and left ends under three point bending using a quarter-range Fourier series representation of the displacement field in each layer and a local/global stiffness matrix approach. Chen [30] proposed a semianalytical and numerical method to analyze multiple crack problems for finite homogeneous plates with an arbitrary contour configuration from elementary infinite solutions. This method is similar to the boundary force method based on the superposition of opposite tractions along an imaginary boundary in an infinite medium used to match boundary conditions of finite specimens. Zhan et al. [31] solved the problem of a finite homogeneous isotropic plate with multiple microcracks using a series expansion of complex potentials and a boundary collocation procedure based on a superposition scheme. Seelig et al. [32] presented a hybrid numerical-analytical method for finite homogeneous bodies with multiple cracks using Kachanov's pseudotraction technique together with a boundary element method for bounded domains.

In this paper, we present a unified solution methodology for finite-dimensioned multilayered configurations with two types of end constraints that contain any number of arbitrarily distributed interlaminar or horizontal cracks in the plane of analysis. The crack opening displacements result from external, internal, or combined normal loading. The solution admits (transversely) isotropic, orthotropic, or monoclinic layers. The latter are obtained by rotation of a transversely isotropic, or orthotropic layer about the layer's normal axis; Fig. 1. The employed displacement formulation, and solution of the Navier's equations, is based on suitable half-range Fourier series representations of the displacement components that admit end-face constraints that mimic horizontal and vertical pins. Following the treatment of layered media infinitely long in the horizontal direction by Erdogan [23], Chatterjee et al. [24], Chatterjee [25], and Pindera [26], the present finite-domain problem is reduced to the determination of the crack-opening displacement components for the pth crack situated along the αth interface in the interval $c_\alpha^{(p)} < x < d_\alpha^{(p)}$, which is governed by the standard system of coupled singular integral equations

$$\mathbf{T}_\alpha^+(x) = \bar{\mathbf{A}}_\alpha^* \boldsymbol{\Theta}_\alpha^{(p)}(x) + \frac{1}{\pi}\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \bar{\mathbf{B}}_\alpha^* \frac{\boldsymbol{\Theta}_\alpha^{(p)}(x')}{x'-x}dx' + \frac{1}{\pi}\sum_{\beta=2}^{n}\sum_{q=1}^{Q^{(\beta)}}\int_{c_\beta^{(q)}}^{d_\beta^{(q)}} \bar{\mathbf{K}}_{\alpha\beta}(x,x')\boldsymbol{\Theta}_\beta^{(q)}(x')dx' + \mathbf{F}_\alpha(x) \quad (1)$$

In the above, n in the summation limit is the total number of layers (assuming that in the most general case all interfaces are cracked), $Q^{(\beta)}$ is number of cracks on the βth interface, $\mathbf{T}_\alpha^+(x)$ is the traction vector specified on the bottom face of the pth crack on the cracked αth interface, $\boldsymbol{\Theta}_\alpha^{(p)}(x)$ is the unknown displacement discontinuity density vector (noting that $\boldsymbol{\Theta}_\beta^{(q)}(x')$ may be zero if the βth interface is uncracked), $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ are regular Fredholm kernels, $\mathbf{F}_\alpha(x)$ is the specified external load vector, and $\bar{\mathbf{A}}_\alpha^*$, $\bar{\mathbf{B}}_\alpha^*$ are constant square matrices whose elements depend only on the material properties of adjacent layers. Muskhelishvili [1,2] and Erdogan et al. [33] outlined a solution technique for such singular integral equations based on the properties of appropriate orthogonal polynomials. An explicit evaluation of the Fredholm kernels $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ for multiple interacting cracks presents considerable difficulties, which in the past have limited the implementation of this exact elasticity formulation to problems involving one or two cracks.

The reduction of the finite-domain multilayered interface crack problem to the standard form given by Eq. (1) is also based on the local/global stiffness matrix formulation (albeit in the harmonic parameter domain), and the extraction of the Cauchy kernel in the crack-face traction condition above is facilitated by the asymptotic behavior of the local stiffness matrix as the harmonic number $m \to \infty$. The regular Fredholm kernels $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ derived in explicit form include contributions resulting from interactions among all cracks, as well as from crack-vertical boundary interactions absent in the Fourier-transform treatment. A representation of the unknown displacement discontinuities in terms of Jacobi polynomials (or Chebyshev polynomials for cracks between adjacent layers of the same material) multiplied by unknown coefficients reduces the singular integral equations to a system of algebraic equations. A solution for the unknown coefficients produces, in general, mixed-mode stress intensity factors and strain energy release rates for each crack, as well as tractions along each cracked interface and stress and displacement fields over the entire multilayered domain. These are presented in Part II of this paper for a number of technologically important cases. The developed elasticity-based capability of extracting important fracture mechanics parameters for multiply cracked layered media in a wide range of materials applications can also be used as a verification tool for extracting these parameters from various conservation integral-based numerical approaches; cf. Kim and Paulino [34] for recent applications involving graded media.

2 Problem Definition and Formulation

The problem under consideration is illustrated in Fig. 2. It involves a finite-height, finite-length layered plate of dimensions H and L that extends to infinity in the out-of-plane direction. The total number of layers is n and the kth layer thickness is h_k. Each layer is fully bonded to its adjacent neighbors along each uncracked interface. The number of cracks along each cracked interface is $Q^{(\beta)}$, where β is the index associated with the cracked βth interface; see Eq. (1). The individual crack intervals along the βth interface are designated by $c_\beta^{(q)} < x < d_\beta^{(q)}$, where q is the index associated with the qth crack. Cracks within a layer are treated as interlaminar cracks between two layers of the same material prop-

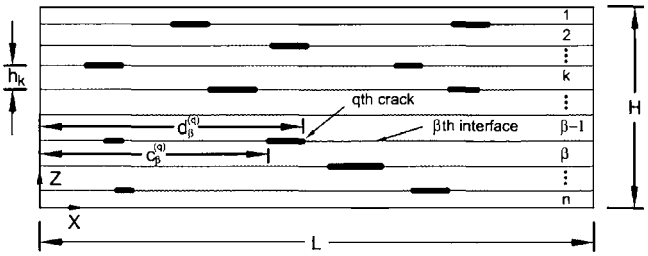


Fig. 2 Geometry of the finite multilayered media with multiple aligned cracks

erties. There is no limit on the number of cracks along a given interface, and no limit on the number of interfaces containing cracks.

The individual layers are either (transversely) isotropic, orthotropic, or monoclinic. Monoclinic layers are obtained by rotating a transversely isotropic or orthotropic layer about the vertical axis z by an angle θ such as would be done in a multilayered composite plate comprised of differently oriented unidirectional plies. Each layer is linearly elastic whose response is represented by the generalized Hooke's law expressed in matrix form

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\epsilon} \tag{2}$$

where $\boldsymbol{\sigma}=[\sigma_{xx},\sigma_{yy},\sigma_{zz},\sigma_{yz},\sigma_{xz},\sigma_{xy}]^T$ is the stress vector, $\boldsymbol{\epsilon}=[\epsilon_{xx},\epsilon_{yy},\epsilon_{zz},2\epsilon_{yz},2\epsilon_{xz},2\epsilon_{xy}]^T$ is the engineering strain vector, and the stiffness matrix $\mathbf{C}$ has the form given below for an orthotropic ply whose principal material coordinate system coincides with the global coordinate system $x-y-z$ shown in Fig. 2,

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix} \tag{3}$$

The stiffness matrix for a transversely isotropic ply with the $x-z$ plane of isotropy is obtained by setting $C_{12}=C_{23}$, $C_{11}=C_{33}$, $C_{44}=C_{66}$, and $C_{55}=1/2(C_{11}-C_{13})$, with similar relations for fully isotropic plies. A transversely isotropic or orthotropic ply rotated by an angle about the z axis behaves like a monoclinic ply in the global (fixed) coordinate system, and its stiffness matrix $\bar{\mathbf{C}}(\theta)$ is related to the stiffness matrix $\mathbf{C}$ in the principal coordinate system by the transformation equations

$$\bar{\mathbf{C}}(\theta) = \mathbf{T}_1\mathbf{C}\mathbf{T}_2^{-1} \tag{4}$$

The transformation matrices $\mathbf{T}_1$ and $\mathbf{T}_2$ relate stress and engineering strain quantities in the principal coordinate system, $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$, to the corresponding quantities in the rotated (primed) coordinate system, $\boldsymbol{\sigma}'$ and $\boldsymbol{\epsilon}'$ (i.e., $\boldsymbol{\sigma}'=\mathbf{T}_1\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}'=\mathbf{T}_2\boldsymbol{\epsilon}$) and are used to derive Eq. (4) from Hooke's law in the principal material coordinate system. Under the above transformation, the stiffness matrix $\bar{\mathbf{C}}(\theta)$ acquires the following form in the global coordinate system

$$\bar{\mathbf{C}}(\theta) = \begin{bmatrix} \bar{C}_{11} & \bar{C}_{12} & \bar{C}_{13} & 0 & 0 & \bar{C}_{16} \\ \bar{C}_{12} & \bar{C}_{22} & \bar{C}_{23} & 0 & 0 & \bar{C}_{26} \\ \bar{C}_{13} & \bar{C}_{23} & \bar{C}_{33} & 0 & 0 & \bar{C}_{36} \\ 0 & 0 & 0 & \bar{C}_{44} & \bar{C}_{45} & 0 \\ 0 & 0 & 0 & \bar{C}_{45} & \bar{C}_{55} & 0 \\ \bar{C}_{16} & \bar{C}_{26} & \bar{C}_{36} & 0 & 0 & \bar{C}_{66} \end{bmatrix} \tag{5}$$

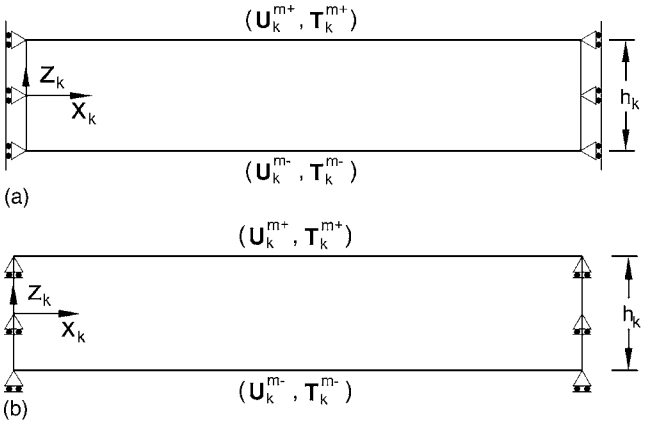


Fig. 3 A finite-dimension layer in the plane of analysis showing two types of end constraints: (*a*) Left and right ends horizontally pinned; (*b*) left and right ends vertically pinned

The loading applied to the layered plate in the $x-z$ plane on the top and bottom faces can be either in the form of normal traction specified by the stress component σ_{zz}, with the shear traction components σ_{yz},σ_{xz} set to zero, or the normal displacement w with the shear traction components set to zero. Each crack face can also be subjected to specified tractions or kept traction-free. The specified boundary conditions are independent of the out-of-plane coordinate y, and thus in the presence of monoclinic plies all three displacement components are present and are functions of the in-plane coordinates x and z, namely $u(x,z),v(x,z),w(x,z)$. Such a displacement field is called generalized plane deformation, Lekhnitskii [35], and reduces to plane strain in the absence of monoclinic plies.

In addition, two ways of constraining the ends are considered; Fig. 3. In the case of horizontally pinned ends,

$$u(0,z) = u(L,z) = 0, \quad v(0,z) = v(L,z) = 0 \tag{6}$$

This type of constraint produces no resultant vertical traction on the end faces, requiring vertical equilibrium to be maintained by appropriate boundary conditions applied on the top and bottom surface of the layered plate. In the case of vertically pinned ends,

$$w(0,z) = w(L,z) = 0 \tag{7}$$

vertical equilibrium is ensured by the end-face shear tractions that result from the applied normal tractions on the top and bottom surfaces.

The necessity of satisfying both the displacement and traction continuity across the interfaces separating adjacent layers along the fully bonded sections makes it desirable to formulate the problem in terms of displacements. The displacement formulation also facilitates the introduction of displacement discontinuity functions along the crack faces of embedded cracks that must be included in the solution procedure to account for the crack opening displacements due to external or internal loading. Using the generalized Hooke's law in conjunction with the strain-displacement equations

$$\epsilon_{xx} = u_{,x}, \quad \epsilon_{yy} = v_{,y}, \quad \epsilon_{zz} = w_{,z}$$

$$\epsilon_{xy} = \frac{1}{2}(u_{,y} + v_{,x}), \quad \epsilon_{xz} = \frac{1}{2}(w_{,x} + u_{,z}), \quad \epsilon_{yz} = \frac{1}{2}(v_{,z} + w_{,y}) \tag{8}$$

in the stress-equilibrium equations, the resulting Navier's equations for a generic monoclinic ply under generalized plane deformation are

$$\bar{C}_{11}u_{,xx} + \bar{C}_{55}u_{,zz} + \bar{C}_{16}v_{,xx} + \bar{C}_{45}v_{,zz} + (\bar{C}_{13} + \bar{C}_{55})w_{,xz} = 0$$

$$\bar{C}_{16}u_{,xx} + \bar{C}_{45}u_{,zz} + \bar{C}_{66}v_{,xx} + \bar{C}_{44}v_{,zz} + (\bar{C}_{36} + \bar{C}_{45})w_{,xz} = 0$$

$$(\bar{C}_{13}+\bar{C}_{55})u_{,xz}+(\bar{C}_{36}+\bar{C}_{45})v_{,xz}+\bar{C}_{55}w_{,xx}+\bar{C}_{33}w_{,zz}=0 \quad (9)$$

where the subscripted commas denote partial derivatives with respect to the specified coordinates, and $\bar{C}_{ij}$ are stiffness matrix elements obtained from the transformation equations given by Eq. (4). Coupling exists among the three displacement components for a monoclinic layer. If the principal material coordinate system coincides with the global coordinate system for an orthotropic or (transversely) isotropic layer, the stiffness matrix elements $\bar{C}_{16}$, $\bar{C}_{26}$, $\bar{C}_{36}$, and $\bar{C}_{45}$ vanish, thereby decoupling the out-of-plane displacement $v(x,z)$ from the inplane displacements $u(x,z)$ and $w(x,z)$ in the Navier's equations.

3 Solution Procedure

The solution to the defined problem is obtained by first solving the Navier's equations for the unknown displacement field in each layer without regard to the presence of interfacial cracks. The presence of interfacial cracks is taken into account during the application of interfacial continuity conditions between adjacent layers through the introduction of displacement discontinuity density functions for the individual cracks. These unknown functions are determined by constructing additional equations that specify prescribed traction conditions on the individual crack faces.

The local/global stiffness matrix approach provides a convenient framework for the systematic solution of multilayered media problems. This approach reduces the size of the global system of equations and facilitates an extraction of the correct form of the governing integral equations for the interfacial displacement discontinuities, as described in the sequel. The local/global stiffness matrix approach is similar to the flexibility matrix formulation proposed by Bufler [36] for isotropic layered media. This approach was reformulated by Rowe and Booker [37] in terms of the local stiffness matrix and applied to the plane analysis of nonhomogeneous isotropic layered soils with the horizontal dimension extending to infinity. It was subsequently extended to laminated composites with interlaminar cracks by Chatterjee and co-workers [24,25,29]. Pindera [26] applied it to the problems of a radially orthotropic layered disk, the delamination of a bi-material beam, and frictionless contact on layered half planes with further extensions given by Pindera and Lane [38], and Urquhart and Pindera [39]. Most recently it has been applied to the analysis of functionally graded and periodic materials, Bansal and Pindera [40,41].

3.1 Solution of the Navier's Equations for a Single Finite Layer. The solution for the displacement field within each layer is obtained in the local coordinate system with the origin centered vertically halfway at each layer's left end; Fig. 3. We obtain solutions to the Navier's equations under the aforementioned two types of end constraints, namely, horizontally and vertically pinned ends. These end constraints dictate the manner of Fourier series expansions of the displacement components $u(x,z), v(x,z), w(x,z)$ and the resulting stress components $\sigma_{zz}(x,z), \sigma_{xz}(x,z), \sigma_{yz}(x,z)$ used to construct the local stiffness matrix for each layer. In both cases, we use a combination of appropriate half-range cosine and sine expansions for the individual displacement components that generate the desired end constraint at $x=0,L$. For example, the half-range cosine and sine Fourier expansions for $u(x,z)$ are, respectively,

$$u(x,z)=\bar{u}_0(z)+\sum_{m=1}^{\infty}\bar{u}_m(z)\cos\left(\frac{m\pi x}{L}\right)$$

$$u(x,z)=\sum_{m=1}^{\infty}\bar{u}'_m(z)\sin\left(\frac{m\pi x}{L}\right) \quad (10)$$

with the respective Fourier coefficients $\bar{u}_0$, $\bar{u}_m$, $\bar{u}'_m$ given by the Euler formulas

$$\bar{u}_0(z)=\frac{1}{L}\int_0^L u(x,z)dx, \quad \bar{u}_m(z)=\frac{2}{L}\int_0^L u(x,z)\cos\left(\frac{m\pi x}{L}\right)dx$$

$$\bar{u}'_m(z)=\frac{2}{L}\int_0^L u(x,z)\sin\left(\frac{m\pi x}{L}\right)dx \quad (11)$$

For both types of end constraints, the solution for the displacement field can be expressed in the following general form

$$\mathbf{U}(x,z)=\bar{\mathbf{U}}^0(z)+\sum_{m=1}^{\infty}\boldsymbol{\zeta}^m(x)\bar{\mathbf{U}}^m(z), \quad (12)$$

where

$$\mathbf{U}(x,z)=\begin{bmatrix}w(x,z)\\u(x,z)\\v(x,z)\end{bmatrix}, \quad \bar{\mathbf{U}}^m(z)=\begin{bmatrix}\bar{w}_m(z)\\\bar{u}_m(z)\\\bar{v}_m(z)\end{bmatrix} \quad (13)$$

and where $\bar{\mathbf{U}}^0(z)$ and $\boldsymbol{\zeta}^m(x)$ for the horizontally pinned end constraint have the forms

$$\bar{\mathbf{U}}^0(z)=\begin{bmatrix}\bar{w}_0(z)\\0\\0\end{bmatrix},$$

$$\boldsymbol{\zeta}^m(x)=\begin{bmatrix}\cos\left(\frac{m\pi x}{L}\right)&0&0\\0&\sin\left(\frac{m\pi x}{L}\right)&0\\0&0&\sin\left(\frac{m\pi x}{L}\right)\end{bmatrix} \quad (14)$$

whereas for the vertically pinned end constraint they are

$$\bar{\mathbf{U}}^0(z)=\begin{bmatrix}0\\\bar{u}_0(z)\\\bar{v}_0(z)\end{bmatrix},$$

$$\boldsymbol{\zeta}^m(x)=\begin{bmatrix}-\sin\left(\frac{m\pi x}{L}\right)&0&0\\0&\cos\left(\frac{m\pi x}{L}\right)&0\\0&0&\cos\left(\frac{m\pi x}{L}\right)\end{bmatrix} \quad (15)$$

The negative sign in the above Fourier series representation of the displacement field for the vertically pinned case ensures that the Navier's equations reduce to the same system of ordinary differential equations in the unknown harmonic vectors $\bar{\mathbf{U}}^m(z)$ for both end constraints, thereby unifying the development of local/global stiffness matrices for these two cases. The solutions for each harmonic of the displacement components of these vectors are then sought in the form,

$$\bar{\mathbf{U}}^m(z)=\mathbf{U}_0^m e^{m\lambda z} \quad (16)$$

where the eigenvalues λ and the associated eigenvectors $\mathbf{U}_0^m$ are given in Appendix A for monoclinic, orthotropic, and (transversely) isotropic plies. The eigenvalues appear in $\pm\lambda_i$ pairs, allowing to express the solutions for the displacement components

in terms of hyperbolic eigenfunctions of z multiplied by unknown eigenvectors. This is the form given in Appendix A, which is employed in the construction of the mth harmonic of the local stiffness matrix described in the next section.

Knowledge of the displacement components produces stress components upon the use of the generalized Hooke's law and strain-displacement relations. In particular, the traction components associated with the top and bottom surfaces of a monoclinic layer can be evaluated from the stress components

$$\sigma_{zz}(x,z) = \bar{C}_{13}u_{,x} + \bar{C}_{33}w_{,z} + \bar{C}_{36}v_{,x}$$

$$\sigma_{xz}(x,z) = \bar{C}_{45}v_{,z} + \bar{C}_{55}(u_{,z} + w_{,x})$$

$$\sigma_{yz}(x,z) = \bar{C}_{44}v_{,z} + \bar{C}_{45}(u_{,z} + w_{,x}) \tag{17}$$

Thus the above stress components have the same functional form with respect to the variable x as the displacement components for each end-constraint case, namely

$$\mathbf{T}(x,z) = \bar{\mathbf{T}}^0(z) + \sum_{m=1}^{\infty} \frac{m\pi}{L}\boldsymbol{\zeta}^m(x)\bar{\mathbf{T}}^m(z) \tag{18}$$

where

$$\mathbf{T}(x,z) = \begin{bmatrix} \sigma_{zz}(x,z) \\ \sigma_{xz}(x,z) \\ \sigma_{yz}(x,z) \end{bmatrix}, \quad \bar{\mathbf{T}}^m(z) = \begin{bmatrix} \bar{\sigma}^m_{zz}(z)/\dfrac{m\pi}{L} \\ \bar{\sigma}^m_{xz}(z)/\dfrac{m\pi}{L} \\ \bar{\sigma}^m_{yz}(z)/\dfrac{m\pi}{L} \end{bmatrix} \tag{19}$$

and where $\bar{\mathbf{T}}^0(z)$ for the horizontally and vertically pinned ends has the respective forms

$$\bar{\mathbf{T}}^0(z) = \begin{bmatrix} \bar{\sigma}^0_{zz} \\ 0 \\ 0 \end{bmatrix}, \quad \bar{\mathbf{T}}^0(z) = \begin{bmatrix} 0 \\ \bar{\sigma}^0_{xz} \\ \bar{\sigma}^0_{yz} \end{bmatrix} \tag{20}$$

while $\boldsymbol{\zeta}^m(x)$ retains the same form as in the displacement expansions. The components of $\bar{\mathbf{T}}^m(z)$ are given as linear combinations of the components of $\bar{\mathbf{U}}^m(z)$ and their derivatives using Eqs. (17). The zeroth-harmonic components $\bar{\sigma}^0_{zz}$ and $\bar{\sigma}^0_{xz}, \bar{\sigma}^0_{yz}$ are constant and represent average normal and shear tractions along a given interface, i.e., $\bar{\sigma}^0_{zz} = (1/L)\int_0^L \sigma_{zz}(x,h/2)dx$, $\bar{\sigma}^0_{xz} = (1/L)\int_0^L \sigma_{xz}(x,h/2)dx$, $\bar{\sigma}^0_{yz} = (1/L)\int_0^L \sigma_{yz}(x,h/2)dx$ are the average tractions along any interface. When only normal tractions are applied on the top and bottom surfaces of the layered medium, $\bar{\sigma}^0_{xz} = \bar{\sigma}^0_{yz} = 0$ for each ply throughout the entire medium.

3.2 Local Stiffness Matrix for a Finite Layer. The local stiffness matrix for the kth ply is constructed by relating the mth harmonics of the three displacements on the top and bottom of the ply to the corresponding traction harmonics,

$$\bar{\mathbf{T}}^{m\pm}_k = \begin{bmatrix} \pm\bar{\sigma}^{m\pm}_{zz}/\dfrac{m\pi}{L} \\ \pm\bar{\sigma}^{m\pm}_{xz}/\dfrac{m\pi}{L} \\ \pm\bar{\sigma}^{m\pm}_{yz}/\dfrac{m\pi}{L} \end{bmatrix}_k, \quad \bar{\mathbf{U}}^{m\pm}_k = \begin{bmatrix} \bar{w}^{\pm}_m \\ \bar{u}^{\pm}_m \\ \bar{v}^{\pm}_m \end{bmatrix}_k \tag{21}$$

with $\pm$ in the superscript denoting respective harmonic terms at the top and bottom surfaces of a layer, i.e., $\bar{\sigma}^{m\pm}_{zz} = \bar{\sigma}^m_{zz}(z = \pm h_k/2)$, etc. By evaluating the harmonics of both displacement and stress components at the top and bottom layer surfaces and eliminating the common unknown eigenvectors, we obtain the local stiffness matrix for the kth layer expressed in the following symbolic form

$$\begin{bmatrix} \bar{\mathbf{T}}^{m+}_k \\ \bar{\mathbf{T}}^{m-}_k \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{m,k}_{11} & \mathbf{K}^{m,k}_{12} \\ \mathbf{K}^{m,k}_{21} & \mathbf{K}^{m,k}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{U}}^{m+}_k \\ \bar{\mathbf{U}}^{m-}_k \end{bmatrix} \tag{22}$$

where the submatrices $\mathbf{K}^{m,k}_{ij}$ have the following structure for a monoclinic layer

$$\begin{bmatrix} \mathbf{K}^{m,k}_{11} & \mathbf{K}^{m,k}_{12} \\ \mathbf{K}^{m,k}_{21} & \mathbf{K}^{m,k}_{22} \end{bmatrix} = \begin{bmatrix} k^m_{11} & k^m_{12} & k^m_{13} & k^m_{14} & k^m_{15} & k^m_{16} \\ k^m_{12} & k^m_{22} & k^m_{23} & -k^m_{15} & k^m_{25} & k^m_{26} \\ k^m_{13} & k^m_{23} & k^m_{33} & -k^m_{16} & k^m_{26} & k^m_{36} \\ k^m_{14} & -k^m_{15} & -k^m_{16} & k^m_{11} & -k^m_{12} & -k^m_{13} \\ k^m_{15} & k^m_{25} & k^m_{26} & -k^m_{12} & k^m_{22} & k^m_{23} \\ k^m_{16} & k^m_{26} & k^m_{36} & -k^m_{13} & k^m_{23} & k^m_{33} \end{bmatrix}_k \tag{23}$$

with the individual elements derived explicitly in terms of the harmonic number m, elastic constants and geometry of the layer, and given in Appendix B.

The asymptotic form of the local stiffness matrix for $m \to \infty$ plays a key role in reducing the governing equations for the crack-opening displacement to Cauchy-type singular integral equations. As $m \to \infty$, the off-diagonal submatrices $\mathbf{K}^{m,k}_{12}$ and $\mathbf{K}^{m,k}_{21}$ that comprise the local stiffness matrix vanish, and the elements of the diagonal submatrices $\mathbf{K}^{m,k}_{11}$ and $\mathbf{K}^{m,k}_{22}$ approach constant values, also given in Appendix B. In this case, the structure of the local stiffness matrix has the form

$$\begin{bmatrix} \mathbf{K}^{*,k}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}^{*,k}_{22} \end{bmatrix} = \begin{bmatrix} k^*_{11} & k^*_{12} & k^*_{13} & 0 & 0 & 0 \\ k^*_{12} & k^*_{22} & k^*_{23} & 0 & 0 & 0 \\ k^*_{13} & k^*_{23} & k^*_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & k^*_{11} & -k^*_{12} & -k^*_{13} \\ 0 & 0 & 0 & -k^*_{12} & k^*_{22} & k^*_{23} \\ 0 & 0 & 0 & -k^*_{13} & k^*_{23} & k^*_{33} \end{bmatrix}_k \tag{24}$$

which indicates decoupling of the top and bottom surface traction harmonics in the kth layer.

In the case of (transversely) isotropic and orthotropic plies, coupling between in-plane and out-of-plane displacement harmonics vanishes and the elements $k^m_{13}, k^m_{23}, k^m_{16}, k^m_{26}, k^m_{36}$ and their limiting counterparts $k^*_{13}, k^*_{23}, k^*_{16}, k^*_{26}, k^*_{36}$ become zero. The nonvanishing elements of the local stiffness matrix for these two cases are also given in Appendix B.

3.3 Global Stiffness Matrix for the Layered Medium. The assembly of the global stiffness matrix is carried out by applying interfacial traction and displacement continuity conditions at each interface. We start at the top surface (called the first interface) and proceed to the bottom surface, using the relations between interfacial tractions and displacements given in terms of the local stiffness matrix elements in Eq. (22).

Ply interface	Individual Traction Equations	Resultant Traction Equations
1^+	$\bar{\mathbf{T}}_1^{m+} = \mathbf{K}_{11}^{m,1}\bar{\mathbf{U}}_1^{m+} + \mathbf{K}_{12}^{m,1}\bar{\mathbf{U}}_1^{m-}$	$\mathbf{K}_{11}^{m,1}\bar{\mathbf{U}}_1^{m+} + \mathbf{K}_{12}^{m,1}\bar{\mathbf{U}}_1^{m-} = \bar{\mathbf{T}}_1^{m+}$
.	.	.
$(k-1)^-$	$\bar{\mathbf{T}}_{k-1}^{m-} = \mathbf{K}_{21}^{m,k-1}\bar{\mathbf{U}}_{k-1}^{m+} + \mathbf{K}_{22}^{m,k-1}\bar{\mathbf{U}}_{k-1}^{m-}$	$\mathbf{K}_{21}^{m,k-1}\bar{\mathbf{U}}_{k-1}^{m+} + \mathbf{K}_{22}^{m,k-1}\bar{\mathbf{U}}_{k-1}^{m-} +$
k^+	$\bar{\mathbf{T}}_k^{m+} = \mathbf{K}_{11}^{m,k}\bar{\mathbf{U}}_k^{m+} + \mathbf{K}_{12}^{m,k}\bar{\mathbf{U}}_k^{m-}$	$\mathbf{K}_{11}^{m,k}\bar{\mathbf{U}}_k^{m+} + \mathbf{K}_{12}^{m,k}\bar{\mathbf{U}}_k^{m-} = \mathbf{0}$
.	.	.
n^-	$\bar{\mathbf{T}}_n^{m-} = \mathbf{K}_{21}^{m,n}\bar{\mathbf{U}}_n^{m+} + \mathbf{K}_{22}^{m,n}\bar{\mathbf{U}}_n^{m-}$	$\mathbf{K}_{21}^{m,n}\bar{\mathbf{U}}_n^{m+} + \mathbf{K}_{22}^{m,n}\bar{\mathbf{U}}_n^{m-} = \bar{\mathbf{T}}_n^{m-}$

(25)

The resultant traction equilibrium equations are generated for the interfaces $k=2,\ldots,n$ by imposing the equilibrium requirement (or traction continuity condition)

$$\bar{\mathbf{T}}_{k-1}^{m-} + \bar{\mathbf{T}}_k^{m+} = \mathbf{0} \tag{26}$$

For uncracked plies, the displacement continuity condition requires that

$$\bar{\mathbf{U}}_{k-1}^{m-} - \bar{\mathbf{U}}_k^{m+} = \mathbf{0} \tag{27}$$

Thus the displacement continuity conditions for uncracked plies can be incorporated directly into the global system of equations by defining the common interfacial displacement vector at the kth uncracked interface as

$$\bar{\mathbf{U}}_{k-1}^{m-} = \bar{\mathbf{U}}_k^{m+} = \bar{\mathbf{U}}_k^{m} \tag{28}$$

For cracked plies, the traction continuity conditions at the cracked αth interface separating $\alpha-1$ and α plies still hold, but the displacement continuity conditions must be modified to account for the interfacial separations or displacement jumps associated with the cracks present along such an interface. We do so by defining the mth harmonic of the displacement discontinuity vector

$$\bar{\mathbf{U}}_{\alpha-1}^{m-} - \bar{\mathbf{U}}_\alpha^{m+} = \bar{\mathbf{U}}_\alpha^{m*} \tag{29}$$

Using the above definition, and examining the limiting behavior of the resultant traction equilibrium equations

$$\mathbf{K}_{21}^{m,\alpha-1}\bar{\mathbf{U}}_{\alpha-1}^{m+} + \mathbf{K}_{22}^{m,\alpha-1}\bar{\mathbf{U}}_{\alpha-1}^{m-} + \mathbf{K}_{11}^{m,\alpha}\bar{\mathbf{U}}_\alpha^{m+} + \mathbf{K}_{12}^{m,\alpha}\bar{\mathbf{U}}_\alpha^{m-} = \mathbf{0} \tag{30}$$

as $m\to\infty$ for the αth cracked interface given the limiting properties of the local stiffness submatrices $\mathbf{K}_{21}^{*,\alpha-1}, \mathbf{K}_{22}^{*,\alpha-1}, \mathbf{K}_{11}^{*,\alpha}, \mathbf{K}_{12}^{*,\alpha}$, the mth harmonics of the displacement vectors $\bar{\mathbf{U}}_{\alpha-1}^{m-}$ and $\bar{\mathbf{U}}_\alpha^{m+}$ can be expressed in terms of the mth harmonics of the common interfacial displacement vector $\bar{\mathbf{U}}_\alpha^{m}$ and the displacement discontinuity vector $\bar{\mathbf{U}}_\alpha^{m*}$ as follows

$$\begin{aligned}\bar{\mathbf{U}}_{\alpha-1}^{m-} &= \bar{\mathbf{U}}_\alpha^{m} + [\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{11}^{*,\alpha}\bar{\mathbf{U}}_\alpha^{m*}\\ \bar{\mathbf{U}}_\alpha^{m+} &= \bar{\mathbf{U}}_\alpha^{m} - [\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\bar{\mathbf{U}}_\alpha^{m*}\end{aligned} \tag{31}$$

where

$$[\mathbf{K}^{*,\alpha}]^{-1} = [\mathbf{K}_{22}^{*,\alpha-1} + \mathbf{K}_{11}^{*,\alpha}]^{-1} \tag{32}$$

The above representation of the mth harmonic of the interfacial displacement vector along interfaces separating cracked plies allows us to express the global system of equations for an n-layered media with multiple interlaminar cracks as follows

$$\begin{bmatrix} \mathbf{K}_{11}^1 & \mathbf{K}_{12}^1 & & & & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_{21}^1 & \mathbf{K}_{22}^1+\mathbf{K}_{11}^2 & & & & & \mathbf{0} \\ & & \ddots & & & & \\ & & \mathbf{K}_{21}^{k-1} & \mathbf{K}_{22}^{k-1}+\mathbf{K}_{11}^k & \mathbf{K}_{12}^k & & \\ & & & & \ddots & & \\ \mathbf{0} & & & & & \mathbf{K}_{22}^{n-1}+\mathbf{K}_{11}^n & \mathbf{K}_{12}^n \\ \mathbf{0} & \mathbf{0} & & & & \mathbf{K}_{21}^n & \mathbf{K}_{22}^n \end{bmatrix}^m \begin{bmatrix} \bar{\mathbf{U}}_1 \\ \bar{\mathbf{U}}_2 \\ \vdots \\ \bar{\mathbf{U}}_k \\ \vdots \\ \bar{\mathbf{U}}_n \\ \bar{\mathbf{U}}_{n+1} \end{bmatrix}^m = \begin{bmatrix} \bar{\mathbf{T}}_1^+ \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \bar{\mathbf{T}}_n^- \end{bmatrix}^m - \sum_{\beta=2}^{n} \begin{bmatrix} \vdots \\ 0 \\ \mathbf{L}_{(\alpha-1)\beta} \\ \mathbf{L}_{\alpha\beta} \\ \mathbf{L}_{(\alpha+1)\beta} \\ 0 \\ \vdots \end{bmatrix}^m \bar{\mathbf{U}}_\beta^{m*}, \quad (\alpha,\beta=2,\ldots,n) \tag{33}$$

where the submatrices of $\mathbf{L}_{(\alpha-1)\beta}^m, \mathbf{L}_{\alpha\beta}^m, \mathbf{L}_{(\alpha+1)\beta}^m$ assume the following nonzero expressions for $\alpha=\beta$ only

$$\begin{aligned}\mathbf{L}_{(\alpha-1)\alpha}^m &= \mathbf{K}_{12}^{m,\alpha-1}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{11}^{*,\alpha}\\ \mathbf{L}_{\alpha\alpha}^m &= \mathbf{K}_{22}^{m,\alpha-1}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{11}^{*,\alpha} - \mathbf{K}_{11}^{m,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\\ \mathbf{L}_{(\alpha+1)\alpha}^m &= -\mathbf{K}_{21}^{m,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\end{aligned} \tag{34}$$

Symbolically, the previous system of equations can be represented by

$$\mathbf{K}^m\bar{\mathbf{U}}^m = \bar{\mathbf{T}}^m - \mathbf{L}^m\bar{\mathbf{U}}^{m*} \tag{35}$$

where $\mathbf{K}^m$ denotes $3(n+1)\times 3(n+1)$ banded and symmetric global stiffness matrix, $\bar{\mathbf{U}}^m$ represents the unknown common interfacial displacement vector of length $3(n+1)$, $\bar{\mathbf{T}}^m$ is the prescribed interfacial traction vector of length $3(n+1)$, $\mathbf{L}^m$ is the $3(n+1)\times 3(n-1)$ matrix containing dimensions and material properties of layers adjacent to cracks, and $\bar{\mathbf{U}}^{m*}$ represents the displacement discontinuity vector of length $3(n-1)$. Equation (35) assumes the most general case that all interfaces are cracked. If an interface is uncracked, say the βth interface, we simply set $\bar{\mathbf{U}}_\beta^{m*}=\mathbf{0}$.

4 Determination of the Displacement Discontinuity Functions

Given the external loading represented by the mth harmonic $\bar{\mathbf{T}}^m$ of the force vector, the solution for the mth harmonic of the common interfacial displacement vector $\bar{\mathbf{U}}^m$ in the absence of interfacial cracks, $\bar{\mathbf{U}}^{m*}=\mathbf{0}$, is obtained in a straightforward manner by inverting the governing system of equations given by Eq. (35). The local displacement, strain, and stress fields can then be built up one harmonic at a time from the knowledge of the interfacial displacement harmonics by solving this system of equations sufficient number of times required for convergence. In the presence of interfacial cracks, the additional displacement discontinuity functions represented by the harmonics $\bar{\mathbf{U}}^{m*}$ must be determined from additional conditions involving the specification of tractions on the crack faces. For the αth interface containing a specified number of cracks, the mth harmonic of the traction vector on the top surface of the αth ply is obtained from Eq. (25), upon setting $k=\alpha$, in the form

$$\bar{\mathbf{T}}_\alpha^{m+}=\mathbf{K}_{11}^{m,\alpha}\bar{\mathbf{U}}_\alpha^{m+}+\mathbf{K}_{12}^{m,\alpha}\bar{\mathbf{U}}_\alpha^{m-} \tag{36}$$

Substituting for $\bar{\mathbf{U}}_\alpha^{m+}$ and $\bar{\mathbf{U}}_\alpha^{m-}$ using Eq. (31), and separating the singular contributions, i.e., contributions which do not vanish as $m\to\infty$, we obtain the following expression for the mth harmonic of the traction vector along the cracked αth interface

$$\begin{aligned}\bar{\mathbf{T}}_\alpha^{m+}=&\mathbf{K}_{11}^{m,\alpha}\bar{\mathbf{U}}_\alpha^{m}+\mathbf{K}_{12}^{m,\alpha}\bar{\mathbf{U}}_{\alpha+1}^{m}-\mathbf{K}_{11}^{*,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\bar{\mathbf{U}}_\alpha^{m*}\\&-\bar{\mathbf{K}}_{11}^{m,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\bar{\mathbf{U}}_\alpha^{m*}+\mathbf{K}_{12}^{m,\alpha}[\mathbf{K}^{*,\alpha+1}]^{-1}\mathbf{K}_{11}^{*,\alpha+1}\bar{\mathbf{U}}_{\alpha+1}^{m*}\end{aligned} \tag{37}$$

where $\bar{\mathbf{K}}_{11}^{m,\alpha}=\mathbf{K}_{11}^{m,\alpha}-\mathbf{K}_{11}^{*,\alpha}$ so that $\bar{\mathbf{K}}_{11}^{m,\alpha}\to\mathbf{0}$ as $m\to\infty$. These equations can be expressed solely in terms of the mth harmonics of the displacement discontinuity functions along all cracked interfaces and external loading by solving for the common interfacial displacements $\bar{\mathbf{U}}_\alpha^m$ and $\bar{\mathbf{U}}_{\alpha+1}^m$. Inverting symbolically the governing system of equations, Eq. (35), we express the unknown interfacial displacement harmonics $\bar{\mathbf{U}}^m$ in terms of external traction boundary conditions and unknown displacement discontinuities as follows

$$\bar{\mathbf{U}}^m=[\mathbf{K}^m]^{-1}\bar{\mathbf{T}}^m-\mathbf{G}^m\bar{\mathbf{U}}^{m*} \tag{38}$$

where $\mathbf{G}^m=[\mathbf{K}^m]^{-1}\mathbf{L}^m$. The solutions for $\bar{\mathbf{U}}_\alpha^m$ and $\bar{\mathbf{U}}_{\alpha+1}^m$ can be extracted from the above general solution in the form

$$\begin{aligned}\bar{\mathbf{U}}_\alpha^m&=[\mathbf{K}^m]_{\alpha1}^{-1}\bar{\mathbf{T}}_1^{m+}+[\mathbf{K}^m]_{\alpha(n+1)}^{-1}\bar{\mathbf{T}}_n^{m-}-\sum_{\beta=2}^{n}\mathbf{G}_{\alpha\beta}^m\bar{\mathbf{U}}_\beta^{m*}\\\bar{\mathbf{U}}_{\alpha+1}^m&=[\mathbf{K}^m]_{(\alpha+1)1}^{-1}\bar{\mathbf{T}}_1^{m+}+[\mathbf{K}^m]_{(\alpha+1)(n+1)}^{-1}\bar{\mathbf{T}}_n^{m-}-\sum_{\beta=2}^{n}\mathbf{G}_{(\alpha+1)\beta}^m\bar{\mathbf{U}}_\beta^{m*}\end{aligned} \tag{39}$$

Therefore, the mth harmonic of the traction on the top surface of the αth layer is expressed solely in terms of externally applied tractions and unknown displacement discontinuities

$$\bar{\mathbf{T}}_\alpha^{m+}=-\mathbf{B}_\alpha^*\bar{\mathbf{U}}_\alpha^{m*}-\sum_{\beta=2}^{n}\mathbf{A}_{\alpha\beta}^m\bar{\mathbf{U}}_\beta^{m*}+\mathbf{F}_{1\alpha}^m\bar{\mathbf{T}}_1^{m+}+\mathbf{F}_{n\alpha}^m\bar{\mathbf{T}}_n^{m-} \tag{40}$$

where

$$\begin{aligned}\mathbf{A}_{\alpha\beta}^m=&\mathbf{K}_{11}^{m,\alpha}\mathbf{G}_{\alpha\beta}^m+\mathbf{K}_{12}^{m,\alpha}\mathbf{G}_{(\alpha+1)\beta}^m+\bar{\mathbf{K}}_{11}^{m,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\delta_{\alpha\beta}\\&-\mathbf{K}_{12}^{m,\alpha}[\mathbf{K}^{*,\alpha+1}]^{-1}\mathbf{K}_{11}^{*,\alpha+1}\delta_{(\alpha+1)\beta}\\\mathbf{B}_\alpha^*=&\mathbf{K}_{11}^{*,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\end{aligned}$$

$$\begin{aligned}\mathbf{F}_{1\alpha}^m&=\mathbf{K}_{11}^{m,\alpha}[\mathbf{K}^m]_{\alpha1}^{-1}+\mathbf{K}_{12}^{m,\alpha}[\mathbf{K}^m]_{(\alpha+1)1}^{-1}\\\mathbf{F}_{n\alpha}^m&=\mathbf{K}_{11}^{m,\alpha}[\mathbf{K}^m]_{\alpha(n+1)}^{-1}+\mathbf{K}_{12}^{m,\alpha}[\mathbf{K}^m]_{(\alpha+1)(n+1)}^{-1}\end{aligned} \tag{41}$$

and where $\delta_{\alpha\beta}$ is the Kronecker delta.

The above equations apply to both the horizontally pinned and vertically pinned ends. Summing up all the harmonics of the traction vector on the cracked αth interface, Eq. (18) with $z=h_\alpha/2$, the traction vector itself becomes

$$\mathbf{T}_\alpha\left(x,\frac{h_\alpha}{2}\right)=\bar{\mathbf{T}}_\alpha^0+\sum_{m=1}^{\infty}\frac{m\pi}{L}\boldsymbol{\zeta}^m(x)\bar{\mathbf{T}}_\alpha^{m+} \tag{42}$$

where the elements of $\bar{\mathbf{T}}_\alpha^0$ are given by Eq. (20) for horizontally and vertically pinned ends,

$$\begin{aligned}\sum_{m=1}^{\infty}\frac{m\pi}{L}\boldsymbol{\zeta}^m\bar{\mathbf{T}}_\alpha^{m+}=&-\sum_{m=1}^{\infty}\frac{m\pi}{L}\boldsymbol{\zeta}^m\mathbf{B}_\alpha^*\bar{\mathbf{U}}_\alpha^{m*}-\sum_{m=1}^{\infty}\frac{m\pi}{L}\boldsymbol{\zeta}^m\sum_{\beta=2}^{n}\mathbf{A}_{\alpha\beta}^m\bar{\mathbf{U}}_\beta^{m*}\\&+\sum_{m=1}^{\infty}\frac{m\pi}{L}\boldsymbol{\zeta}^m(\mathbf{F}_{1\alpha}^m\bar{\mathbf{T}}_1^{m+}+\mathbf{F}_{n\alpha}^m\bar{\mathbf{T}}_n^{m-})\end{aligned} \tag{43}$$

and the matrix $\boldsymbol{\zeta}^m$ remains the same as before for both pinned cases.

4.1 Integral Form of the Displacement Discontinuities. In order to reduce the expression for the traction vector on the top surface of the αth ply acting on the face of the pth crack in the interval $c_\alpha^{(p)}<x<d_\alpha^{(p)}$, Eq. (42), to a system of integral equations, we introduce displacement discontinuity density vector $\boldsymbol{\Theta}_\alpha^{(p)}$ in the following manner

$$\mathbf{U}_{\alpha-1}^{-(p)}(x)-\mathbf{U}_\alpha^{+(p)}(x)=\mathbf{U}_\alpha^{*(p)}(x)=\begin{cases}\displaystyle\int_{c_\alpha^{(p)}}^{x}\boldsymbol{\Theta}_\alpha^{(p)}(x')dx' & c_\alpha^{(p)}\leqslant x\leqslant d_\alpha^{(p)}\\ \mathbf{0} & \text{otherwise}\end{cases} \tag{44}$$

with the constraint

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\boldsymbol{\Theta}_\alpha^{(p)}(x')dx'=\mathbf{0} \tag{45}$$

where $\boldsymbol{\Theta}_\alpha^{(p)}$ consists of displacement discontinuity density components associated with the z, x, and y directions,

$$\boldsymbol{\Theta}_\alpha^{(p)}(x')=\begin{bmatrix}\theta_z^{(p)}(x')\\ \theta_x^{(p)}(x')\\ \theta_y^{(p)}(x')\end{bmatrix}_\alpha \tag{46}$$

The corresponding Fourier series representation of the crack opening displacement $\mathbf{U}_\alpha^{*(p)}(x)$ for the pth crack along the αth cracked interface is then given by

$$\mathbf{U}_\alpha^{*(p)}(x)=\sum_{m=1}^{\infty}\boldsymbol{\zeta}^m\bar{\mathbf{U}}_\alpha^{m*(p)}\quad\text{where}\quad\bar{\mathbf{U}}_\alpha^{m*(p)}=\begin{bmatrix}\bar{w}_m^{*(p)}\\ \bar{u}_m^{*(p)}\\ \bar{v}_m^{*(p)}\end{bmatrix}_\alpha \tag{47}$$

Therefore, the mth harmonic $\bar{\mathbf{U}}_\alpha^{m*(p)}$ of the crack opening displacement $\mathbf{U}_\alpha^{*(p)}(x)$ is obtained in terms of the components of the displacement discontinuity density vector $\boldsymbol{\Theta}_\alpha^{(p)}(x')$ by multiplying the above equation by the appropriate orthogonal harmonic and integrating the result along the crack interval. For both horizontally and vertically pinned ends, $\bar{\mathbf{U}}_\alpha^{m*(p)}$ can be written symbolically as follows

$$\bar{\mathbf{U}}_{\alpha}^{m*(p)} = \frac{2}{m\pi}\int_{c_{\alpha}^{(p)}}^{d_{\alpha}^{(p)}} \hat{\boldsymbol{\zeta}}^{m}\boldsymbol{\Theta}_{\alpha}^{(p)}(x')dx' \tag{48}$$

where for the horizontally pinned ends,

$$\hat{\boldsymbol{\zeta}}^{m} = \begin{bmatrix} -\sin\left(\frac{m\pi x'}{L}\right) & 0 & 0 \\ 0 & \cos\left(\frac{m\pi x'}{L}\right) & 0 \\ 0 & 0 & \cos\left(\frac{m\pi x'}{L}\right) \end{bmatrix} \tag{49}$$

which is actually the same as $\boldsymbol{\zeta}^{m}$ for the vertically pinned case. Similarly, for the vertically pinned ends,

$$\hat{\boldsymbol{\zeta}}^{m} = -\begin{bmatrix} \cos\left(\frac{m\pi x'}{L}\right) & 0 & 0 \\ 0 & \sin\left(\frac{m\pi x'}{L}\right) & 0 \\ 0 & 0 & \sin\left(\frac{m\pi x'}{L}\right) \end{bmatrix} \tag{50}$$

which is actually the same as $-\boldsymbol{\zeta}^{m}$ for the horizontally pinned case.

The mth harmonic $\bar{\mathbf{U}}_{\alpha}^{m*}$ of the crack opening displacement for all cracks along the αth cracked interface appearing in Eq. (43) is then obtained by summing up all contributions from each crack as follows

$$\bar{\mathbf{U}}_{\alpha}^{m*} = \sum_{p=1}^{P^{(\alpha)}} \bar{\mathbf{U}}_{\alpha}^{m*(p)} \tag{51}$$

4.2 Reduction of Crack-Face Traction Condition to Singular Integral Equations. The integral representation of the crack opening displacement for the pth crack along the αth interface using the displacement discontinuity density vector $\boldsymbol{\Theta}_{\alpha}^{(p)}(x')$ allows us to express the crack face traction condition in the interval $c_{\alpha}^{(p)} < x < d_{\alpha}^{(p)}$ in terms of this unknown vector as follows

$$\begin{aligned}\mathbf{T}_{\alpha}\left(x, \frac{h_{\alpha}}{2}\right) = &-\frac{2}{L}\sum_{m=1}^{\infty}\boldsymbol{\zeta}^{m}\mathbf{B}_{\alpha}^{*}\sum_{p=1}^{P^{(\alpha)}}\int_{c_{\alpha}^{(p)}}^{d_{\alpha}^{(p)}}\hat{\boldsymbol{\zeta}}^{m}\boldsymbol{\Theta}_{\alpha}^{(p)}(x')dx' \\ &-\frac{2}{L}\sum_{m=1}^{\infty}\boldsymbol{\zeta}^{m}\sum_{\beta=2}^{n}\mathbf{A}_{\alpha\beta}^{m}\sum_{q=1}^{Q^{(\beta)}}\int_{c_{\beta}^{(q)}}^{d_{\beta}^{(q)}}\hat{\boldsymbol{\zeta}}^{m}\boldsymbol{\Theta}_{\beta}^{(q)}(x')dx' \\ &+\sum_{m=1}^{\infty}\frac{m\pi}{L}\boldsymbol{\zeta}^{m}(\mathbf{F}_{1\alpha}^{m}\bar{\mathbf{T}}_{1}^{m+} + \mathbf{F}_{n\alpha}^{m}\bar{\mathbf{T}}_{n}^{m-}) + \bar{\mathbf{T}}_{\alpha}^{0}\end{aligned} \tag{52}$$

using Eqs. (42) and (43), where the components of $\mathbf{T}_{\alpha}(x, h_{\alpha}/2)$ are specified (typically either zero or constant). The first term on the right hand side of this system of three-coupled integral equations,

$$-\frac{2}{L}\sum_{m=1}^{\infty}\boldsymbol{\zeta}^{m}\mathbf{B}_{\alpha}^{*}\sum_{p=1}^{P^{(\alpha)}}\int_{c_{\alpha}^{(p)}}^{d_{\alpha}^{(p)}}\hat{\boldsymbol{\zeta}}^{m}\boldsymbol{\Theta}_{\alpha}^{(p)}(x')dx' \tag{53}$$

reduces to a singular integral with a Cauchy-type kernel, as shown in the sequel, while the remaining two terms

$$-\frac{2}{L}\sum_{m=1}^{\infty}\boldsymbol{\zeta}^{m}\sum_{\beta=2}^{n}\mathbf{A}_{\alpha\beta}^{m}\sum_{q=1}^{Q^{(\beta)}}\int_{c_{\beta}^{(q)}}^{d_{\beta}^{(q)}}\hat{\boldsymbol{\zeta}}^{m}\boldsymbol{\Theta}_{\beta}^{(q)}(x')dx',$$

$$\sum_{m=1}^{\infty}\frac{m\pi}{L}\boldsymbol{\zeta}^{m}(\mathbf{F}_{1\alpha}^{m}\bar{\mathbf{T}}_{1}^{m+} + \mathbf{F}_{n\alpha}^{m}\bar{\mathbf{T}}_{n}^{m-}) + \bar{\mathbf{T}}_{\alpha}^{0} \tag{54}$$

are regular, with the elements of $\mathbf{A}_{\alpha\beta}^{m}$, $\mathbf{F}_{1\alpha}^{m}$, and $\mathbf{F}_{n\alpha}^{m}$ vanishing in the limit as $m \to \infty$.

4.2.1 Extraction of the Singular Kernel. The extraction of the Cauchy-type kernel from the first term forms the key step in the reduction of the crack-face traction condition to a system of singular integral equations. It is carried out by first expressing the four summations below, which appear in the integrals associated with the components of $\boldsymbol{\Theta}_{\alpha}^{(p)}(x')$ in Eq. (53), as follows

$$\begin{aligned}\sum_{m=1}^{\infty}\sin\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right) &= \frac{1}{2}\sum_{m=1}^{\infty}\sin\frac{m\pi(x+x')}{L} \\ &+ \frac{1}{2}\sum_{m=1}^{\infty}\sin\frac{m\pi(x-x')}{L}\end{aligned}$$

$$\begin{aligned}\sum_{m=1}^{\infty}\cos\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right) &= \frac{1}{2}\sum_{m=1}^{\infty}\sin\frac{m\pi(x+x')}{L} \\ &- \frac{1}{2}\sum_{m=1}^{\infty}\sin\frac{m\pi(x-x')}{L}\end{aligned}$$

$$\begin{aligned}\sum_{m=1}^{\infty}\sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right) &= -\frac{1}{2}\sum_{m=1}^{\infty}\cos\frac{m\pi(x+x')}{L} \\ &+ \frac{1}{2}\sum_{m=1}^{\infty}\cos\frac{m\pi(x-x')}{L}\end{aligned}$$

$$\begin{aligned}\sum_{m=1}^{\infty}\cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right) &= \frac{1}{2}\sum_{m=1}^{\infty}\cos\frac{m\pi(x+x')}{L} \\ &+ \frac{1}{2}\sum_{m=1}^{\infty}\cos\frac{m\pi(x-x')}{L}\end{aligned} \tag{55}$$

The products on the left hand side of the above equations are the four different combinations obtained by multiplying the elements of the matrices $\boldsymbol{\zeta}^{m}$ and $\hat{\boldsymbol{\zeta}}^{m}$ that appear in the first term (and others). The four summations on the right hand side of the above equations are then transformed such that known identities can be applied in order to extract the singular contributions, as shown in detail in Appendix C. Summarizing these manipulations, the integrals in Eq. (53) become

$$\begin{aligned}&\int_{c_{\alpha}^{(p)}}^{d_{\alpha}^{(p)}}\sum_{m=1}^{\infty}\sin\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx' \\ &= \int_{c_{\alpha}^{(p)}}^{d_{\alpha}^{(p)}}\left(-\frac{L}{2\pi}\frac{1}{x'-x} + \Omega(x',x)\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx'\end{aligned}$$

$$\begin{aligned}&\int_{c_{\alpha}^{(p)}}^{d_{\alpha}^{(p)}}\sum_{m=1}^{\infty}\cos\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx' \\ &= \int_{c_{\alpha}^{(p)}}^{d_{\alpha}^{(p)}}\left(\frac{L}{2\pi}\frac{1}{x'-x} + \Omega(x,x')\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx'\end{aligned}$$

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \sum_{m=1}^{\infty} \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$
$$= \int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \frac{L}{2}\cdot\delta(x-x')\,\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \sum_{m=1}^{\infty} \cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$
$$= \int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \frac{L}{2}\cdot\delta(x-x')\,\theta_{(\cdot)\alpha}^{(p)}(x')dx' \tag{56}$$

where the function $\Omega(x,x')$ is given below

$$\Omega(x,x') = \frac{1}{4}\cot\frac{\pi(x+x')}{2L} - \frac{x-x'}{4\pi L}\sum_{n=1}^{\infty}\frac{1}{\left(\frac{x-x'}{2L}\right)^2 - n^2} \tag{57}$$

and $0<x/L,x'/L<1$. We note that the sequence of x,x' in our definition of $\Omega(x,x')$ matters, i.e., $\Omega(x,x')\neq\Omega(x',x)$.

Explicit expressions for the three terms in the system of singular integral equations that represent the crack-face traction condition for the horizontally pinned and vertically pinned ends, Eq. (52), are given in the next subsection.

4.2.2 Singular Integral Equations. Using the results obtained in the preceding section to separate the singular and nonsingular kernel contributions in the integrals appearing in Eq. (52), the crack-face traction condition for the pth crack in the interval $c_\alpha^{(p)}<x<d_\alpha^{(p)}$ along the αth interface can be expressed as a system of singular integral equations with Cauchy-type kernels using the same symbolic notation as that of Eq. (1), repeated here for convenience

$$\mathbf{T}_\alpha\left(x,\frac{h_\alpha}{2}\right) = \bar{\mathbf{A}}_\alpha^*\boldsymbol{\Theta}_\alpha^{(p)}(x) + \frac{1}{\pi}\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\bar{\mathbf{B}}_\alpha^*\frac{\boldsymbol{\Theta}_\alpha^{(p)}(x')}{x'-x}dx'$$
$$+\frac{1}{\pi}\sum_{\beta=2}^{n}\sum_{q=1}^{Q^{(\beta)}}\int_{c_\beta^{(q)}}^{d_\beta^{(q)}}\bar{\mathbf{K}}_{\alpha\beta}(x,x')\boldsymbol{\Theta}_\beta^{(q)}(x')dx' + \mathbf{F}_\alpha(x) \tag{58}$$

The summation on $\beta=2,\ldots,n$ includes crack interactions on different interfaces. For both horizontally and vertically pinned ends, the structure of the matrices $\bar{\mathbf{A}}_\alpha^*$ and $\bar{\mathbf{B}}_\alpha^*$ is the same, namely

$$\bar{\mathbf{A}}_\alpha^* = \begin{bmatrix} 0 & -B_{12}^{*\alpha} & -B_{13}^{*\alpha} \\ B_{12}^{*\alpha} & 0 & 0 \\ B_{13}^{*\alpha} & 0 & 0 \end{bmatrix},\quad \bar{\mathbf{B}}_\alpha^* = \begin{bmatrix} B_{11}^{*\alpha} & 0 & 0 \\ 0 & B_{22}^{*\alpha} & B_{23}^{*\alpha} \\ 0 & B_{23}^{*\alpha} & B_{33}^{*\alpha} \end{bmatrix} \tag{59}$$

where $\bar{\mathbf{A}}_\alpha^*$ and $\bar{\mathbf{B}}_\alpha^*$ involve only material properties of the adjacent layers. The individual elements $B_{ij}^{*\alpha}$ that populate these two matrices are obtained from $\mathbf{B}_\alpha^*$ in Eqs. (41) in terms of the products of the asymptotic values of the stiffness matrix elements of the adjacent plies separated by the αth interface. The elements of the matrix $\bar{\mathbf{K}}_{\alpha\beta}$ and the force vector $\mathbf{F}_\alpha(x)$ depend on the manner of end support.

For the horizontally pinned ends, the elements of $\bar{\mathbf{K}}_{\alpha\beta}$ are

$$\bar{K}_{11}^{\alpha\beta} = \frac{2\pi}{L}\left(\kappa_{11}^{\alpha\beta} + \left(\Omega(x,x') + \frac{L}{2\pi}\frac{1-\delta_{pq}}{x'-x}\right)\delta_{\alpha\beta}B_{11}^{*\alpha}\right)$$

$$\bar{K}_{1j}^{\alpha\beta} = \frac{2\pi}{L}\kappa_{1j}^{\alpha\beta}$$

$$\bar{K}_{i1}^{\alpha\beta} = \frac{2\pi}{L}\kappa_{i1}^{\alpha\beta}$$

$$\bar{K}_{ij}^{\alpha\beta} = \frac{2\pi}{L}\left(\kappa_{ij}^{\alpha\beta} + \left(-\Omega(x',x) + \frac{L}{2\pi}\frac{1-\delta_{pq}}{x'-x}\right)\delta_{\alpha\beta}B_{ij}^{*\alpha}\right) \tag{60}$$

where

$$\kappa_{11}^{\alpha\beta} = \sum_{m=1}^{\infty} A_{11}^{m(\alpha\beta)}\cos\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right)$$

$$\kappa_{1j}^{\alpha\beta} = \sum_{m=1}^{\infty} (-A_{1j}^{m(\alpha\beta)})\cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right)$$

$$\kappa_{i1}^{\alpha\beta} = \sum_{m=1}^{\infty} A_{i1}^{m(\alpha\beta)}\sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right)$$

$$\kappa_{ij}^{\alpha\beta} = \sum_{m=1}^{\infty} (-A_{ij}^{m(\alpha\beta)})\sin\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right) \tag{61}$$

for $i,j=2,3$, and the elements $A_{ij}^{m(\alpha\beta)}$ are obtained by inverting the global stiffness matrix, as shown in Eq. (41). The terms in $A_{ij}^{m(\alpha\beta)}$ account for crack interactions due to cracks situated on different interfaces. The functions $\Omega(x,x')$ and $\Omega(x',x)$ describe crack-vertical boundary interactions caused by the constraints of the right and left ends (this type of interaction is the same as one arising from two collinear cracks if one imagines another crack located symmetrically across the layered medium's vertical boundary due to the chosen Fourier series displacement representation). The term $(1-\delta_{pq})/(x'-x)$ represents the effect of collinear crack interactions on the current pth crack by the qth crack situated on the same interface.

The load vector $\mathbf{F}_\alpha(x)$ is given by

$$\mathbf{F}_\alpha(x) = \begin{bmatrix} \bar{\sigma}_{zz\alpha}^0 \\ 0 \\ 0 \end{bmatrix} + \sum_{m=1}^{\infty}\begin{bmatrix} \cos\left(\frac{m\pi x}{L}\right)(F_{11,1}^{m(\alpha)}\bar{\sigma}_{zz1}^{m+} - F_{11,n}^{m(\alpha)}\bar{\sigma}_{zzn}^{m-}) \\ \sin\left(\frac{m\pi x}{L}\right)(F_{21,1}^{m(\alpha)}\bar{\sigma}_{zz1}^{m+} - F_{21,n}^{m(\alpha)}\bar{\sigma}_{zzn}^{m-}) \\ \sin\left(\frac{m\pi x}{L}\right)(F_{31,1}^{m(\alpha)}\bar{\sigma}_{zz1}^{m+} - F_{31,n}^{m(\alpha)}\bar{\sigma}_{zzn}^{m-}) \end{bmatrix} \tag{62}$$

when only normal traction loading is applied on the top and bottom surfaces of the layered media. This loading is represented by half-range Fourier cosine expansions with the constant and fluctuating contributions denoted by $\bar{\sigma}_{zz\alpha}^0$ and $\bar{\sigma}_{zz1}^{m+}$, $\bar{\sigma}_{zzn}^{m-}$, respectively. If only internal crack pressure is applied, $\mathbf{F}_\alpha(x)=0$.

For the vertically pinned ends, the elements of the matrix $\bar{\mathbf{K}}_{\alpha\beta}$ are

$$\bar{K}_{11}^{\alpha\beta} = \frac{2\pi}{L}\left(\kappa_{11}^{\alpha\beta} + \left(-\Omega(x',x) + \frac{L}{2\pi}\frac{1-\delta_{pq}}{x'-x}\right)\delta_{\alpha\beta}B_{11}^{*\alpha}\right)$$

$$\bar{K}_{1j}^{\alpha\beta} = \frac{2\pi}{L}\kappa_{1j}^{\alpha\beta}$$

$$\bar{K}_{i1}^{\alpha\beta} = \frac{2\pi}{L}\kappa_{i1}^{\alpha\beta}$$

$$\bar{K}_{ij}^{\alpha\beta} = \frac{2\pi}{L}\left(\kappa_{ij}^{\alpha\beta} + \left(\Omega(x,x') + \frac{L}{2\pi}\frac{1-\delta_{pq}}{x'-x}\right)\delta_{\alpha\beta}B_{ij}^{*\alpha}\right) \tag{63}$$

where

$$\kappa_{11}^{\alpha\beta}=\sum_{m=1}^{\infty}(-A_{11}^{m(\alpha\beta)})\sin\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right)$$

$$\kappa_{1j}^{\alpha\beta}=\sum_{m=1}^{\infty}\left(-A_{1j}^{m(\alpha\beta)}\right)\sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right)$$

$$\kappa_{i1}^{\alpha\beta}=\sum_{m=1}^{\infty}A_{i1}^{m(\alpha\beta)}\cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right)$$

$$\kappa_{ij}^{\alpha\beta}=\sum_{m=1}^{\infty}A_{ij}^{m(\alpha\beta)}\cos\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right) \tag{64}$$

for $i,j=2,3$.

The load vector $\mathbf{F}_\alpha(x)$ is given by

$$\mathbf{F}_\alpha(x)=\sum_{m=1}^{\infty}\begin{bmatrix}-\sin\left(\frac{m\pi x}{L}\right)(F_{11,1}^{m(\alpha)}\bar{\sigma}_{zz1}^{m+}-F_{11,n}^{m(\alpha)}\bar{\sigma}_{zzn}^{m-})\\ \cos\left(\frac{m\pi x}{L}\right)(F_{21,1}^{m(\alpha)}\bar{\sigma}_{zz1}^{m+}-F_{21,n}^{m(\alpha)}\bar{\sigma}_{zzn}^{m-})\\ \cos\left(\frac{m\pi x}{L}\right)(F_{31,1}^{m(\alpha)}\bar{\sigma}_{zz1}^{m+}-F_{31,n}^{m(\alpha)}\bar{\sigma}_{zzn}^{m-})\end{bmatrix} \tag{65}$$

when only normal traction loading is applied on the top and bottom surfaces of the layered media. The half-range Fourier sine expansion ensures that the zeroth order harmonic of constant normal traction is 0. As before, $\mathbf{F}_\alpha(x)=0$ is employed when only internal pressure is applied.

5 Solution of Singular Integral Equations

The system of equations defined by Eq. (58) is solved by the method proposed by Erdogan et al. [33]. The dominant system consisting of the Cauchy-type integrals and the free terms $\bar{\mathbf{A}}_\alpha^*\boldsymbol{\Theta}_\alpha^{(p)}(x)$ is first diagonalized and then expressed in normalized form with respect to each crack interval. Let

$$\boldsymbol{\Theta}_\alpha^{(p)}(x)=\mathbf{R}_\alpha^*\boldsymbol{\psi}_\alpha^{(p)}(x) \tag{66}$$

to diagonalize the singular integral equations. This procedure yields

$$\mathbf{g}_\alpha(x)+\mathbf{R}_\alpha^{*-1}\bar{\mathbf{B}}_\alpha^{*-1}\mathbf{T}_\alpha^+(x)=\boldsymbol{\Lambda}_\alpha\boldsymbol{\psi}_\alpha^{(p)}(x)+\frac{1}{\pi}\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\frac{\boldsymbol{\psi}_\alpha^{(p)}(x')}{x'-x}dx'+\frac{1}{\pi}\sum_{\beta=2}^{n}\sum_{q=1}^{Q^{(\beta)}}\int_{c_\beta^{(q)}}^{d_\beta^{(q)}}\mathbf{h}_{\alpha\beta}(x,x')\boldsymbol{\psi}_\beta^{(q)}(x')dx' \tag{67}$$

where

$$\boldsymbol{\Lambda}_\alpha=\mathbf{R}_\alpha^{*-1}\bar{\mathbf{B}}_\alpha^{*-1}\bar{\mathbf{A}}_\alpha^*\mathbf{R}_\alpha^*$$

$$\mathbf{g}_\alpha(x)=-\mathbf{R}_\alpha^{*-1}\bar{\mathbf{B}}_\alpha^{*-1}\mathbf{F}_\alpha(x)$$

$$\mathbf{h}_{\alpha\beta}(x,x')=\mathbf{R}_\alpha^{*-1}\bar{\mathbf{B}}_\alpha^{*-1}\bar{\mathbf{K}}_{\alpha\beta}\mathbf{R}_\beta^* \tag{68}$$

and $\boldsymbol{\Lambda}_\alpha$ and $\mathbf{R}_\alpha^*$ are diagonal and modal matrices, respectively, determined from the associated eigenvalue-eigenvector problem; see Appendix D. Then, let $x'=\frac{1}{2}(d_\beta^{(q)}-c_\beta^{(q)})\tau_\beta^{(q)}+\frac{1}{2}(d_\beta^{(q)}+c_\beta^{(q)})$ and $x=\frac{1}{2}(d_\alpha^{(p)}-c_\alpha^{(p)})t_\alpha^{(p)}+\frac{1}{2}(d_\alpha^{(p)}+c_\alpha^{(p)})$ to normalize the singular integral equations as follows

$$\mathbf{g}_\alpha(t_\alpha^{(p)})+\mathbf{R}_\alpha^{*-1}\bar{\mathbf{B}}_\alpha^{*-1}\mathbf{T}_\alpha^+(t_\alpha^{(p)})=\boldsymbol{\Lambda}_\alpha\boldsymbol{\psi}_\alpha^{(p)}(t_\alpha^{(p)})+\frac{1}{\pi}\int_{-1}^{+1}\frac{\boldsymbol{\psi}_\alpha^{(p)}(\tau_\alpha^{(p)})}{\tau_\alpha^{(p)}-t_\alpha^{(p)}}d\tau_\alpha^{(p)}+\frac{1}{\pi}\sum_{\beta=2}^{n}\sum_{q=1}^{Q^{(\beta)}}\frac{d_\beta^{(q)}-c_\beta^{(q)}}{2}\times\int_{-1}^{+1}\mathbf{h}_{\alpha\beta}(t_\alpha^{(p)},\tau_\beta^{(q)})\boldsymbol{\psi}_\beta^{(q)}(\tau_\beta^{(q)})d\tau_\beta^{(q)} \tag{69}$$

The solution to the above diagonalized and normalized system of equations can be expressed as the product of so-called fundamental function $\boldsymbol{\omega}_\alpha(t)$ reflecting the nature of the problem's singularity and some function that is bounded at $t=\pm1$, where

$$\boldsymbol{\omega}_\alpha(t)=(1-t)^{\mathbf{a}_\alpha}(1+t)^{\mathbf{b}_\alpha} \tag{70}$$

$\mathbf{a}_\alpha=\left[-\frac{1}{2},-\frac{1}{2}+i\boldsymbol{\epsilon}_\alpha,-\frac{1}{2}-i\boldsymbol{\epsilon}_\alpha\right]^T$, $\mathbf{b}_\alpha$ is the complex conjugate of $\mathbf{a}_\alpha$, and $\boldsymbol{\epsilon}_\alpha=(1/2\pi)\log[(\Lambda_\alpha+1)/(\Lambda_\alpha-1)]$. Since the fundamental function $\boldsymbol{\omega}_\alpha(t)$ is the weight of Jacobi polynomials, it is advantageous to expand the bounded function in Jacobi polynomials with unknown coefficients and the resulting representation of $\boldsymbol{\psi}_\alpha^{(p)}(t_\alpha^{(p)})$ is

$$\boldsymbol{\psi}_\alpha^{(p)}(t_\alpha^{(p)})=\boldsymbol{\omega}_\alpha(t_\alpha^{(p)})\sum_{j=0}^{\infty}\mathbf{C}_{j(\alpha)}^{(p)}P_j^{(\mathbf{a}_\alpha,\mathbf{b}_\alpha)}(t_\alpha^{(p)}) \tag{71}$$

where $\mathbf{C}_{j(\alpha)}^{(p)}=[C_{1,j(\alpha)}^{(p)}\ C_{2,j(\alpha)}^{(p)}\ C_{3,j(\alpha)}^{(p)}]^{\mathrm{T}}$ is the unknown influence coefficient vector.

In order to reduce the three coupled integral equations governing the crack opening displacements to a system of algebraic equations, we first substitute the above series representation of $\boldsymbol{\psi}_\alpha^{(p)}(t_\alpha^{(p)})$ into Eq. (69) and employ the Jacobi polynomial identity

$$\boldsymbol{\Lambda}_\alpha P_j^{(\mathbf{a}_\alpha,\mathbf{b}_\alpha)}(t_\alpha^{(p)})\boldsymbol{\omega}_\alpha(t_\alpha^{(p)})+\frac{1}{i\pi}\int_{-1}^{+1}\frac{P_j^{(\mathbf{a}_\alpha,\mathbf{b}_\alpha)}(t_\alpha^{(p)})\boldsymbol{\omega}_\alpha(\tau_\alpha^{(p)})}{\tau_\alpha^{(p)}-t_\alpha^{(p)}}d\tau_\alpha^{(p)}=\frac{\sqrt{1-\boldsymbol{\Lambda}_\alpha^2}}{2i}P_{j-1}^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}(t_\alpha^{(p)}) \tag{72}$$

to reduce the first two terms to a single summation in $P_j^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}(t_\alpha^{(p)})$. Multiplying both sides of the resulting reduced equation by $\hat{\boldsymbol{\omega}}_\alpha(t_\alpha^{(p)})P_j^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}(t_\alpha^{(p)})$ and taking advantage of the orthogonality of Jacobi polynomials $P_j^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}(t_\alpha^{(p)})$

$$\int_{-1}^{+1}\hat{\boldsymbol{\omega}}_\alpha(t_\alpha^{(p)})P_i^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}(t_\alpha^{(p)})P_j^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}(t_\alpha^{(p)})dt_\alpha^{(p)}=\begin{cases}\theta_j^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)} & \text{for } i=j\\ 0 & \text{for } i\neq j\end{cases} \tag{73}$$

where the weight functions $\hat{\boldsymbol{\omega}}_\alpha(t)$ are

$$\hat{\boldsymbol{\omega}}_\alpha(t)=(1-t)^{-\mathbf{a}_\alpha}(1+t)^{-\mathbf{b}_\alpha} \tag{74}$$

and

$$\theta_j^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}=\frac{2\Gamma(j+1-\mathbf{a}_\alpha)\Gamma(j+1-\mathbf{b}_\alpha)}{j!\,(j+1)\Gamma(j+2)} \tag{75}$$

the singular integral equations are reduced to the algebraic system of equations in the unknown coefficients $\mathbf{C}_{j(\alpha)}^{(p)}$

$$\frac{\pi}{2}\sqrt{1-\boldsymbol{\Lambda}_\alpha^2}\,\theta_j^{(-\mathbf{a}_\alpha,-\mathbf{b}_\alpha)}\mathbf{C}_{(j+1)(\alpha)}^{(p)}+\sum_{\beta=2}^{n}\sum_{q=1}^{Q^{(\beta)}}\sum_{k=0}^{N_j}\frac{(d_\beta^{(q)}-c_\beta^{(q)})}{2}\mathbf{D}_{jk(\alpha\beta)}^{(pq)}\mathbf{C}_{k(\beta)}^{(q)}=\mathbf{G}_{j(\alpha)}^{I(p)}+\mathbf{G}_{j(\alpha)}^{II(p)} \tag{76}$$

where $\Gamma(\cdot)$ is a gamma function and ! denotes factorial in Eq. (75), and $j=0,1,2,\ldots,N_j$. The constant matrices and vectors $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$, $\mathbf{G}_{j(\alpha)}^{I(p)}$, $\mathbf{G}_{j(\alpha)}^{II(p)}$ are given below

$$\mathbf{D}_{jk(\alpha\beta)}^{(pq)} = \int_{-1}^{+1} \left[\int_{-1}^{+1} \mathbf{h}_{\alpha\beta}(t_\alpha^{(p)}, \tau_\beta^{(q)}) P_j^{(-\mathbf{a}_\alpha, -\mathbf{b}_\alpha)}(t_\alpha^{(p)}) \hat{\boldsymbol{\omega}}_\alpha(t_\alpha^{(p)}) dt_\alpha^{(p)} \right] \times P_k^{(\mathbf{a}_\beta, \mathbf{b}_\beta)}(\tau_\beta^{(q)}) \boldsymbol{\omega}_\beta(\tau_\beta^{(q)}) d\tau_\beta^{(q)}$$

$$\mathbf{G}_{j(\alpha)}^{I(p)} = -\int_{-1}^{+1} \mathbf{R}_\alpha^{*-1} \bar{\mathbf{B}}_\alpha^{*-1} \mathbf{F}_\alpha(t_\alpha^{(p)}) P_j^{(-\mathbf{a}_\alpha, -\mathbf{b}_\alpha)}(t_\alpha^{(p)}) \hat{\boldsymbol{\omega}}_\alpha(t_\alpha^{(p)}) dt_\alpha^{(p)}$$

$$\mathbf{G}_{j(\alpha)}^{II(p)} = \int_{-1}^{+1} \mathbf{R}_\alpha^{*-1} \bar{\mathbf{B}}_\alpha^{*-1} \mathbf{T}_\alpha^+(t_\alpha^{(p)}) P_j^{(-\mathbf{a}_\alpha, -\mathbf{b}_\alpha)}(t_\alpha^{(p)}) \hat{\boldsymbol{\omega}}_\alpha(t_\alpha^{(p)}) dt_\alpha^{(p)} \quad (77)$$

The solution of the above system of equations is accomplished numerically. In practice, the series in Eq. (76) is truncated after the first N_j terms that must be large enough to yield accurate results. The size of the system of algebraic equations is $3 \times (N_j + 1) \times \Sigma_{\alpha=2}^n P^{(\alpha)}$. Numerical aspects of determining the coefficients given in Eq. (77) are discussed in Part II of this paper.

6 Strain Energy Release Rates and Stress Intensity Factors

Once the coefficients $\mathbf{C}_{j(\alpha)}^{(p)}$ of the Jacobi polynomials associated with the various cracks are known, the displacement discontinuities and thus tractions (and stresses) in the vicinity of crack tips can be evaluated. The energy release rates and stress intensity factors follow directly from the knowledge of the above quantities. Herein, we derive these quantities for the case $\bar{\mathbf{A}}_\alpha^* = \mathbf{0}$, that is $B_{12}^{*\alpha} = B_{13}^{*\alpha} = 0$, which results in $\boldsymbol{\Lambda}_\alpha = \mathbf{0}$. In this case, the Jacob polynomials $P_j^{(-\mathbf{a}_\alpha, -\mathbf{b}_\alpha)}$ reduce to Chebyshev polynomials of the first kind T_j since $\mathbf{a}_\alpha = \mathbf{b}_\alpha = \left[-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right]^T$. This is the case when the crack is situated between two layers of the same elastic properties, thereby sidestepping the problem of oscillatory crack-tip stress fields that has commanded much attention and effort by a number of researchers, cf. Wang and Choi [42,43], as discussed by Ting [44].

The strain energy release rate caused by an infinitesimal crack extension of the pth crack along the αth interface under self-similar crack growth is calculated by evaluating the integral given later

$$\frac{\partial U_\alpha^{(p)}}{\partial a} = \lim_{\delta \to 0} \frac{1}{\delta} \frac{1}{2} \int_{\pm 1}^{\pm 1 \pm \delta'} \frac{(d_\alpha^{(p)} - c_\alpha^{(p)})}{2} \mathbf{T}_\alpha^+(t_\alpha^{(p)}) \circ \mathbf{U}_\alpha^*(t_\alpha'^{(p)}) dt_\alpha^{(p)} \quad (78)$$

where the symbol $\circ$ denotes dot product operation, $\pm$ denotes the right or left tip of the crack, $t_\alpha'^{(p)} = t_\alpha^{(p)} - 2\delta/(d_\alpha^{(p)} - c_\alpha^{(p)})$, and $\delta' = 2\delta/(d_\alpha^{(p)} - c_\alpha^{(p)})$. In the above expression it is sufficient to consider only the limiting values of the traction $\mathbf{T}_\alpha^+(t_\alpha^{(p)})$ and crack opening displacement $\mathbf{U}_\alpha^*(t_\alpha^{(p)})$ at each crack location in the neighborhood of $t_\alpha^{(p)} = \pm 1$.

The expression for the crack-opening displacement inside the pth crack along the αth interface, $|t_\alpha^{(p)}| \leq 1$, is given by

$$\mathbf{U}_\alpha^*(t_\alpha^{(p)}) = \frac{(d_\alpha^{(p)} - c_\alpha^{(p)})}{2} \int_{-1}^{t_\alpha^{(p)}} \boldsymbol{\Theta}_\alpha^{(p)}(\tau_\alpha^{(p)}) d\tau_\alpha^{(p)} = \frac{(d_\alpha^{(p)} - c_\alpha^{(p)})}{2} \int_{-1}^{t_\alpha^{(p)}} (1 - \tau_\alpha^{(p)})^{-1/2} (1 + \tau_\alpha^{(p)})^{-1/2} \sum_{j=0}^{\infty} \frac{(2j)!}{2^{2j}(j!)^2} \mathbf{C}_{j(\alpha)}^{(p)} T_j(\tau_\alpha^{(p)}) d\tau_\alpha^{(p)} \quad (79)$$

In the vicinity of the right crack tip, the above integral is governed by the dominant part of the kernel, yielding

$$\mathbf{U}_\alpha^*(t_\alpha^{(p)}) = \frac{(d_\alpha^{(p)} - c_\alpha^{(p)})}{2} (1 - t_\alpha^{(p)})^{1/2} \frac{2}{\sqrt{2}} \sum_{j=0}^{\infty} \frac{(2j)!}{2^{2j}(j!)^2} \mathbf{C}_{j(\alpha)}^{(p)} \quad (80)$$

using a result from the complex variable theory and $T_j(1) = 1$. Similarly, for the same pth crack along the αth interface, the dominant or singular part of the stress field ahead of the crack tip obtained from Eq. (58) is given by

$$\mathbf{T}_\alpha^+(t_\alpha^{(p)}) \approx \frac{\bar{\mathbf{B}}_\alpha^*}{\pi} \int_{-1}^{+1} \frac{\boldsymbol{\Theta}_\alpha^{(p)}(\tau_\alpha^{(p)})}{\tau_\alpha^{(p)} - t_\alpha^{(p)}} d\tau_\alpha^{(p)} = \frac{\bar{\mathbf{B}}_\alpha^*}{\pi} \sum_{j=1}^{\infty} \frac{(2j)!}{2^{2j}(j!)^2} \mathbf{C}_{j(\alpha)}^{(p)} \int_{-1}^{+1} \times \frac{T_j(\tau_\alpha^{(p)})(1 - \tau_\alpha^{(p)})^{-1/2}(1 + \tau_\alpha^{(p)})^{-1/2}}{\tau_\alpha^{(p)} - t_\alpha^{(p)}} d\tau_\alpha^{(p)} \quad (81)$$

Using a result from the complex variable theory, the asymptotic behavior of this integral in the vicinity of the right crack tip for $|t_\alpha^{(p)}| > 1$ becomes

$$\mathbf{T}_\alpha^+(t_\alpha^{(p)}) = -(t_\alpha^{(p)} - 1)^{-1/2} \bar{\mathbf{B}}_\alpha^* \frac{1}{\sqrt{2}} \sum_{j=0}^{\infty} \frac{(2j)!}{2^{2j}(j!)^2} \mathbf{C}_{j(\alpha)}^{(p)} \quad (82)$$

The asymptotic expressions for the displacement and traction fields in the vicinity of the right crack tip contain common summations that are defined in the manner given below

$$-\frac{1}{\sqrt{2}} \sum_{j=0}^{\infty} \frac{(2j)!}{2^{2j}(j!)^2} \mathbf{C}_{j(\alpha)}^{(p)} = [A_\alpha^{(p)}, B_\alpha^{(p)}, C_\alpha^{(p)}]^T \quad (83)$$

Using this definition in the asymptotic expressions for the crack-tip tractions and displacements in the integral for the strain energy release rate, and separating the three contributions, we obtain the following integrals for the energy release rate due to the opening mode

$$\frac{\partial U_{I\alpha}^{(p)}}{\partial a} = \lim_{\delta \to 0} \frac{1}{\delta} \frac{(d_\alpha^{(p)} - c_\alpha^{(p)})^2}{4} B_{11}^{*\alpha} (A_\alpha^{(p)})^2 \times \int_1^{1+\delta'} (t_\alpha^{(p)} - 1)^{-1/2} (1 - t_\alpha'^{(p)})^{1/2} dt_\alpha^{(p)}$$

the sliding mode

$$\frac{\partial U_{II\alpha}^{(p)}}{\partial a} = \lim_{\delta \to 0} \frac{1}{\delta} \frac{(d_\alpha^{(p)} - c_\alpha^{(p)})^2}{4} (B_{22}^{*\alpha} B_\alpha^{(p)} + B_{23}^{*\alpha} C_\alpha^{(p)}) B_\alpha^{(p)} \times \int_1^{1+\delta'} (t_\alpha^{(p)} - 1)^{-1/2} (1 - t_\alpha'^{(p)})^{1/2} dt_\alpha^{(p)}$$

and the tearing mode

$$\frac{\partial U_{III\alpha}^{(p)}}{\partial a} = \lim_{\delta \to 0} \frac{1}{\delta} \frac{(d_\alpha^{(p)} - c_\alpha^{(p)})^2}{4} (B_{23}^{*\alpha} B_\alpha^{(p)} + B_{33}^{*\alpha} C_\alpha^{(p)}) C_\alpha^{(p)} \times \int_1^{1+\delta'} (t_\alpha^{(p)} - 1)^{-1/2} (1 - t_\alpha'^{(p)})^{1/2} dt_\alpha^{(p)}$$

Evaluating the above integrals using the following identity derived from the definition of beta function, $B(x,y) = \int_0^1 t^x (1-t)^{y-1} dt$, Gradsheteyn and Ryzhik ([45], p. 898)

$$\int_1^{1+2\delta/(d_\alpha^{(p)}-c_\alpha^{(p)})} (t_\alpha^{(p)}-1)^{-1/2}(1-t_\alpha^{(p)}-2\delta/(d_\alpha^{(p)}-c_\alpha^{(p)}))^{1/2}\,dt_\alpha^{(p)}$$

$$= \frac{\pi\delta}{d_\alpha^{(p)}-c_\alpha^{(p)}} \tag{84}$$

the following expressions for the strain energy release rates corresponding to each mode at the right crack tip are obtained

$$\frac{\partial U_{I\alpha}^{(p)}}{\partial a} = \frac{\pi}{4}(d_\alpha^{(p)}-c_\alpha^{(p)})B_{11}^{*\alpha}(A_\alpha^{(p)})^2$$

$$\frac{\partial U_{II\alpha}^{(p)}}{\partial a} = \frac{\pi}{4}(d_\alpha^{(p)}-c_\alpha^{(p)})(B_{22}^{*\alpha}B_\alpha^{(p)}+B_{23}^{*\alpha}C_\alpha^{(p)})B_\alpha^{(p)}$$

$$\frac{\partial U_{III\alpha}^{(p)}}{\partial a} = \frac{\pi}{4}(d_\alpha^{(p)}-c_\alpha^{(p)})(B_{23}^{*\alpha}B_\alpha^{(p)}+B_{33}^{*\alpha}C_\alpha^{(p)})C_\alpha^{(p)} \tag{85}$$

The corresponding energy release rates at the left tip are obtained by evaluating the Chebyshev coefficient sums in Eqs. (83) at $t_\alpha^{(p)}=-1$ instead of $t_\alpha^{(p)}=+1$.

The stress intensity factors are obtained by multiplying the asymptotic expressions of crack-tip tractions by either $(t_\alpha^{(p)}-1)^{1/2}$ for the right tip or by $(-1-t_\alpha^{(p)})^{1/2}$ for the left tip, and then taking the limit of these equations as $t\to\pm1$. Taking this limit at the right tip, $t_\alpha^{(p)}=1$, in Eq. (82), the stress intensity factors become

$$K_{I\alpha}^{(p)} = B_{11}^{*\alpha}A_\alpha^{(p)}$$

$$K_{II\alpha}^{(p)} = B_{22}^{*\alpha}B_\alpha^{(p)}+B_{23}^{*\alpha}C_\alpha^{(p)}$$

$$K_{III\alpha}^{(p)} = B_{23}^{*\alpha}B_\alpha^{(p)}+B_{33}^{*\alpha}C_\alpha^{(p)} \tag{86}$$

with similar expressions at the left tip.

7 Summary and Conclusions

A unified solution methodology has been developed for finite-dimensioned multilayered media laminated with (transversely) isotropic, orthotropic, and monoclinic plies containing arbitrarily distributed interacting cracks in a state of generalized plane deformation under two types of end constraints that mimic vertical and horizontal pins. Using the local/global stiffness matrix approach in the Fourier harmonic parameter domain, the mixed boundary-value problem has been reduced to the standard system of coupled singular integral equations for the crack opening displacements previously obtained by several investigators in the Fourier transform domain when the horizontal dimension of the layered medium is infinite. Explicit expressions for the elements of the local stiffness matrix and the singular and regular kernels associated with the crack opening displacements have been developed in the Fourier harmonic parameter domain that can easily be programmed. These kernels account for all crack interactions from cracks resident on every cracked interface, as well crack-vertical boundary interactions absent in the corresponding Fourier transform formulation.

Expressions for strain energy release rates and stress intensity factors have been developed based on the obtained solution that will be first verified in Part II of this paper, and then employed in several technologically significant applications for the first time.

Acknowledgement

The first author thanks the Civil Engineering Department at the University of Virginia and the Engineering Material Concepts, LLC for providing partial support in the course of this investigation.

Appendix A: Solutions to Navier's equations

The solutions for the displacement field in monoclinic, orthotropic, and (transversely) isotropic layers are obtained by substituting the assumed functions given in Eqs. (12)–(15) in the Navier's equations and applying orthogonality of a Fourier series. The zeroth-order, $m=0$, solutions to the Navier's equations for the horizontally and vertically pinned cases, Eqs. (14) and (15), respectively, are linear functions of the coordinate z as shown below

$$\bar{w}_0(z) = A_{30}+B_{30}z, \quad \bar{u}_0(z) = A_{10}+B_{10}z, \quad \bar{v}_0(z) = A_{20}+B_{20}z$$

For harmonics greater than zero, $m>0$, the solutions for the displacement field in monoclinic, orthotropic, and (transversely) isotropic layers are given below.

A.1 Monoclinic Layers

$$\bar{w}_m(z) = \sum_{j=1}^{3} R_j\left(G_j\cosh\left(\frac{m\pi}{L}\lambda_j z\right)+F_j\sinh\left(\frac{m\pi}{L}\lambda_j z\right)\right)$$

$$\bar{u}_m(z) = \sum_{j=1}^{3}\left(F_j\cosh\left(\frac{m\pi}{L}\lambda_j z\right)+G_j\sinh\left(\frac{m\pi}{L}\lambda_j z\right)\right)$$

$$\bar{v}_m(z) = \sum_{j=1}^{3} L_j\left(F_j\cosh\left(\frac{m\pi}{L}\lambda_j z\right)+G_j\sinh\left(\frac{m\pi}{L}\lambda_j z\right)\right)$$

where F_j, G_j, are the unknown eigenvectors, the coefficients L_i, R_i are

$$L_i = \frac{(\bar{C}_{11}-\bar{C}_{55}\lambda_i^2)(\bar{C}_{36}+\bar{C}_{45})-(\bar{C}_{16}-\bar{C}_{45}\lambda_i^2)(\bar{C}_{13}+\bar{C}_{55})}{(\bar{C}_{45}\lambda_i^2-\bar{C}_{16})(\bar{C}_{36}+\bar{C}_{45})-(\bar{C}_{44}\lambda_i^2-\bar{C}_{66})(\bar{C}_{13}+\bar{C}_{55})}$$

$$R_i = -\frac{(\bar{C}_{45}\lambda_i^2-\bar{C}_{16})(\bar{C}_{45}\lambda_i^2-\bar{C}_{16})-(\bar{C}_{44}\lambda_i^2-\bar{C}_{66})(\bar{C}_{55}\lambda_i^2-\bar{C}_{11})}{(\bar{C}_{45}\lambda_i^2-\bar{C}_{16})(\bar{C}_{36}+\bar{C}_{45})-(\bar{C}_{44}\lambda_i^2-\bar{C}_{66})(\bar{C}_{13}+\bar{C}_{55})}\frac{1}{\lambda_i}$$

and the eigenvalues λ_j are obtained from the characteristic equation $A\lambda^6+B\lambda^4+C\lambda^2+D=0$, where

$$A = \bar{C}_{33}(\bar{C}_{44}\bar{C}_{55}-\bar{C}_{45}^2)$$

$$B = \bar{C}_{44}(\bar{C}_{11}\bar{C}_{33}-\bar{C}_{13}^2)+\bar{C}_{55}(\bar{C}_{33}\bar{C}_{66}-\bar{C}_{36}^2) - 2\bar{C}_{45}(\bar{C}_{16}\bar{C}_{33}-\bar{C}_{13}\bar{C}_{36})-2\bar{C}_{13}(\bar{C}_{44}\bar{C}_{55}-\bar{C}_{45}^2)$$

$$C = (\bar{C}_{36}+2\bar{C}_{45})(\bar{C}_{11}\bar{C}_{36}-\bar{C}_{13}\bar{C}_{16})+(\bar{C}_{13}+2\bar{C}_{55})(\bar{C}_{13}\bar{C}_{66} - \bar{C}_{16}\bar{C}_{36})-\bar{C}_{11}(\bar{C}_{44}\bar{C}_{55}-\bar{C}_{45}^2)-\bar{C}_{33}(\bar{C}_{11}\bar{C}_{66}-\bar{C}_{16}^2)$$

$$D = \bar{C}_{55}(\bar{C}_{11}\bar{C}_{66}-\bar{C}_{16}^2)$$

Defining $\gamma=\lambda^2-B/3A$, the characteristic equation is transformed into the cubic form $\gamma^3+e\gamma+f=0$ with $e=-B^2/3A^2-C/A$ and $f=-2B^3/27A^3-BC/3A^2-D/A$. Assuming that $e^3/27+f^2/4<0$, which is the case for most unidirectional composites, the equation has three real and unequal roots $\gamma_j=2\sqrt{-e/3}\cos[1/3(\varphi+2(j-1)\pi)]$, $j=1,2,3$, with $\varphi=\cos^{-1}[-\sqrt{27}f/2(-e)^{3/2}]$. Therefore, the three positive roots of the original equation are $\lambda_j=\sqrt{\gamma_j+B/3A}$.

A.2 Orthotropic Layers

$$\bar{w}_m(z) = \sum_{j=1}^{2} R_j\left(G_j\cosh\left(\frac{m\pi}{L}\lambda_j z\right)+F_j\sinh\left(\frac{m\pi}{L}\lambda_j z\right)\right)$$

$$\bar{u}_m(z) = \sum_{j=1}^{2}\left(F_j\cosh\left(\frac{m\pi}{L}\lambda_j z\right)+G_j\sinh\left(\frac{m\pi}{L}\lambda_j z\right)\right)$$

$$\bar{v}_m(z) = H_1 \cosh\left(\frac{m\pi}{L}\lambda_3 z\right) + I_1 \sinh\left(\frac{m\pi}{L}\lambda_3 z\right)$$

where the coefficients R_i are

$$R_i = -\frac{C_{55}\lambda_i^2 - C_{11}}{C_{13} + C_{55}}\frac{1}{\lambda_i}$$

and the eigenvalues λ_j are obtained from the characteristic equations $B\lambda^4 + C\lambda^2 + D = 0$ and $\lambda^2 = C_{66}/C_{44}$ where $B = C_{33}C_{55}$, $C = C_{13}^2 - C_{11}C_{33} + 2C_{13}C_{55}$ and $D = C_{11}C_{55}$. If $C^2 - 4BD > 0$, which is the case for most unidirectional composites, the first characteristic equation has two positive real and unequal roots $\lambda_{1,2} = \sqrt{(-C \pm \sqrt{C^2 - 4BD})/2B}$ and the third positive root $\lambda_3 = \sqrt{C_{66}/C_{44}}$ comes from the second characteristic equation.

A.3 Isotropic Layers

$$\bar{w}_m(z) = -\left(\left(G_1 + \frac{R_1}{\frac{m\pi}{L}}F_2 + zG_2\right)\cosh\left(\frac{m\pi}{L}z\right) + \left(F_1 + \frac{R_1}{\frac{m\pi}{L}}G_2 + zF_2\right)\sinh\left(\frac{m\pi}{L}z\right)\right)$$

$$\bar{u}_m(z) = (F_1 + zF_2)\cosh\left(\frac{m\pi}{L}z\right) + (G_1 + zG_2)\sinh\left(\frac{m\pi}{L}z\right)$$

$$\bar{v}_m(z) = H_1 \cosh\left(\frac{m\pi}{L}z\right) + I_1 \sinh\left(\frac{m\pi}{L}z\right)$$

where

$$R_1 = -\frac{3C_{33} - C_{12}}{C_{11} + C_{12}}$$

and the eigenvalues λ_j are obtained from the characteristic equations $(\lambda^2 - 1)^2 = 0$ and $\lambda^2 = 1$ whose three positive roots are $\lambda_{1,2,3} = 1$.

Appendix B: Elements of Local Stiffness Matrix

The elements of local stiffness matrices for finite monoclinic, orthotropic, and isotropic layers in Fourier expansion for generalized plane deformation are given below.

B.1 Monoclinic Layers

$$k_{11}^m(k_{14}^m) = A_-^+[P_1(L_3 - L_2) + P_2(L_1 - L_3) + P_3(L_2 - L_1)]$$

$$k_{12}^m(k_{15}^m) = H_-^+(P_1L_2 - P_2L_1)R_3 + E_-^+(P_3L_1 - P_1L_3)R_2 + F_-^+(P_2L_3 - P_3L_2)R_1$$

$$k_{13}^m(k_{16}^m) = H_-^+(P_2 - P_1)R_3 + E_-^+(P_1 - P_3)R_2 + F_-^+(P_3 - P_2)R_1$$

$$k_{22}^m(k_{25}^m) = B_-^+(T_2R_1 - T_1R_2)L_3 + C_-^+(T_1R_3 - T_3R_1)L_2 + D_-^+(T_3R_2 - T_2R_3)L_1$$

$$k_{23}^m(k_{26}^m) = B_-^+(T_1R_2 - T_2R_1) + C_-^+(T_3R_1 - T_1R_3) + D_-^+(T_2R_3 - T_3R_2)$$

$$k_{33}^m(k_{36}^m) = B_-^+(R_2Q_1 - R_1Q_2) + C_-^+(R_1Q_3 - R_3Q_1) + D_-^+(R_3Q_2 - R_2Q_3)$$

where the minus sign in the notation A_-^+, etc., refers to the elements k_{ij} in the parentheses,

$$P_i = \bar{C}_{13} + \bar{C}_{33}\lambda_i R_i + \bar{C}_{36}L_i, \quad Q_i = \bar{C}_{44}\lambda_i L_i + \bar{C}_{45}(\lambda_i - R_i),$$

$$T_i = \bar{C}_{45}\lambda_i L_i + \bar{C}_{55}(\lambda_i - R_i)$$

and

$$A_-^+ = \frac{1}{2}\left(\pm\frac{1}{\Delta_1} + \frac{t_1t_2t_3}{\Delta_2}\right), \quad B_-^+ = \frac{1}{2}\left(\frac{t_1t_2}{\Delta_1} \pm \frac{t_3}{\Delta_2}\right), \quad C_-^+ = \left(\frac{1}{2}\frac{t_1t_3}{\Delta_1} \pm \frac{t_2}{\Delta_2}\right)$$

$$D_-^+ = \frac{1}{2}\left(\frac{t_2t_3}{\Delta_1} \pm \frac{t_1}{\Delta_2}\right), \quad E_-^+ = \frac{1}{2}\left(\frac{t_2}{\Delta_1} \pm \frac{t_1t_3}{\Delta_2}\right), \quad F_-^+ = \frac{1}{2}\left(\frac{t_1}{\Delta_1} \pm \frac{t_2t_3}{\Delta_2}\right)$$

$$H_-^+ = \frac{1}{2}\left(\frac{t_3}{\Delta_1} \pm \frac{t_1t_2}{\Delta_2}\right)$$

where

$$\Delta_1 = t_3(L_2 - L_1)R_3 + t_2(L_1 - L_3)R_2 + t_1(L_3 - L_2)R_1$$

$$\Delta_2 = t_1t_2(L_2 - L_1)R_3 + t_1t_3(L_1 - L_3)R_2 + t_2t_3(L_3 - L_2)R_1$$

and $t_i = \tanh((m\pi/L)\lambda_i h/2)$, h is the kth ply thickness, L is the ply half length, and m is the harmonic.

The asymptotic elements of the local stiffness matrix are:

$$k_{11}^* = \frac{1}{\Delta}[P_1(L_3 - L_2) + P_2(L_1 - L_3) + P_3(L_2 - L_1)]$$

$$k_{12}^* = \frac{1}{\Delta}[(P_1L_2 - P_2L_1)R_3 + (P_3L_1 - P_1L_3)R_2 + (P_2L_3 - P_3L_2)R_1]$$

$$k_{13}^* = \frac{1}{\Delta}[(P_2 - P_1)R_3 + (P_1 - P_3)R_2 + (P_3 - P_2)R_1]$$

$$k_{22}^* = \frac{1}{\Delta}[(T_2R_1 - T_1R_2)L_3 + (T_1R_3 - T_3R_1)L_2 + (T_3R_2 - T_2R_3)L_1]$$

$$k_{23}^* = \frac{1}{\Delta}[(T_1R_2 - T_2R_1) + (T_3R_1 - T_1R_3) + (T_2R_3 - T_3R_2)]$$

$$k_{33}^* = \frac{1}{\Delta}[(R_2Q_1 - R_1Q_2) + (R_1Q_3 - R_3Q_1) + (R_3Q_2 - R_2Q_3)]$$

where $\Delta = (L_2 - L_1)R_3 + (L_1 - L_3)R_2 + (L_3 - L_2)R_1$.

B.2 Orthotropic Layers

$$k_{11}^m(k_{14}^m) = A_-^+(P_2 - P_1)$$

$$k_{12}^m(k_{15}^m) = B_-^+P_1R_2 - E_-^+P_2R_1$$

$$k_{22}^m(k_{25}^m) = D_-^+(T_1R_2 - T_2R_1)$$

$$k_{33}^m(k_{36}^m) = C_-^+Q_3$$

where $P_i = -C_{13} + C_{33}\lambda_i R_i$, $T_i = C_{55}(\lambda_i + R_i)$, $Q_3 = C_{44}\lambda_3$, and

$$A_-^+ = \frac{1}{2}\left(\pm\frac{1}{\Delta_1} + \frac{t_1t_2}{\Delta_2}\right), \quad B_-^+ = \frac{1}{2}\left(\frac{t_1t_2}{\Delta_1} \pm \frac{1}{\Delta_2}\right), \quad C_-^+ = \frac{1}{2}\left(t_3 \pm \frac{1}{t_3}\right)$$

$$D_-^+ = \frac{1}{2}\left(\frac{t_2}{\Delta_1} \pm \frac{t_1}{\Delta_2}\right), \quad E_-^+ = \frac{1}{2}\left(\frac{t_1}{\Delta_1} \pm \frac{t_2}{\Delta_2}\right)$$

where $\Delta_1 = t_2R_2 - t_1R_1$, $\Delta_2 = t_1R_2 - t_2R_1$, and $t_i = \tanh[(m\pi/L)\lambda_i h/2]$.

The asymptotic elements of the local stiffness matrix are:

$$k_{11}^* = \frac{P_2 - P_1}{R_2 - R_1}$$

$$k_{12}^* = \frac{P_1R_2 - P_2R_1}{R_2 - R_1}$$

$$k_{22}^{*}=\frac{T_1R_2-T_2R_1}{R_2-R_1}$$

$$k_{33}^{*}=Q_3$$

B.3 Isotropic Layers

$$k_{11}^{m}(k_{14}^{m})=A_{-}^{+}P_2$$

$$k_{12}^{m}(k_{15}^{m})=D_{-}^{+}P_1+E_{-}^{+}P_2$$

$$k_{22}^{m}(k_{25}^{m})=B_{-}^{+}(T_1R_1-T_2)$$

$$k_{33}^{m}(k_{36}^{m})=C_{-}^{+}Q_3$$

where $P_1=C_{11}-C_{12}$, $P_2=C_{11}(1+R_1)$, $Q_3=C_{44}$, $T_1=2C_{55}$, $T_2=C_{55}(1+R_1)$, and

$$A_{-}^{+}=\frac{1}{2}\left(\pm\frac{1}{\Delta_1}+\frac{t_1^2}{\Delta_2}\right)\frac{L}{m\pi},\quad B_{-}^{+}=\frac{1}{2}\left(\frac{t_1^2}{\Delta_1}\pm\frac{1}{\Delta_2}\right)\frac{L}{m\pi},$$

$$C_{-}^{+}=\frac{1}{2}\left(t_1\pm\frac{1}{t_1}\right)$$

$$D_{-}^{+}=\frac{1}{2}(1\pm1),\quad E_{-}^{+}=\frac{1}{2}\left(\frac{t_1}{\Delta_1}\pm\frac{t_1}{\Delta_2}\right)\frac{L}{m\pi}$$

where $\Delta_1=h(1-t_1^2)/2+R_1t_1L/m\pi$, $\Delta_2=-h(1-t_1^2)/2+R_1t_1L/m\pi$, and $t_1=\tanh[(m\pi/L)h/2]$.

The asymptotic elements of the local stiffness matrix are:

$$k_{11}^{*}=C_{11}\frac{1+R_1}{R_1}$$

$$k_{12}^{*}=C_{12}+\frac{C_{11}}{R_1}$$

$$k_{22}^{*}=k_{11}^{*}$$

$$k_{33}^{*}=\frac{1}{2}(C_{11}-C_{12})$$

Appendix C: Extraction of the Singular Kernel

The four summations on the right hand side of Eq. (55) are transformed such that known identities given below, Gradsheteyn and Ryzhik ([45], p. 44), valid for $0<x<2\pi$, can be applied in order to extract the singular contributions.

$$\sum_{m=1}^{\infty}\frac{\sin(mx)}{m}=\frac{\pi-x}{2},\quad \sum_{m=1}^{\infty}\frac{\cos(mx)}{m}=\frac{1}{2}\log\frac{1}{2(1-\cos x)}$$

We start with the summations involving the sum $x+x'$ in Eq. (55). First, we rewrite the summation on $\sin(m\pi/L)(x+x')$ as follows

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\sum_{m=1}^{\infty}\sin\frac{m\pi(x+x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$=\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}-\frac{d}{dx}\sum_{m=1}^{\infty}\frac{L}{m\pi}\cos\frac{m\pi(x+x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

Using the identity

$$\sum_{m=1}^{\infty}\frac{L}{m\pi}\cos\frac{m\pi(x+x')}{L}=\frac{L}{2\pi}\log\frac{1}{2\left(1-\cos\dfrac{\pi(x+x')}{L}\right)}$$

which holds for $0<(\pi/L)(x+x')<2\pi$ and performing the required differentiation, we obtain

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\sum_{m=1}^{\infty}\sin\frac{m\pi(x+x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$=\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\frac{1}{2}\cot\frac{\pi(x+x')}{2L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

where the kernel is nonsingular since $0<(\pi/L)(x+x')<2\pi$. Applying the same procedure to the summation on $\cos(m\pi/L)(x+x')$, we obtain

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\sum_{m=1}^{\infty}\cos\frac{m\pi(x+x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$=\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\frac{d}{dx}\sum_{m=1}^{\infty}\frac{L}{m\pi}\sin\frac{m\pi(x+x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$=\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\frac{d}{dx}\left(\frac{L-(x+x')}{2}\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx'=0$$

in light of the constraint on $\Theta_\alpha^{(p)}$ given in Eq. (45).

We follow a similar procedure for the summations involving the difference $x-x'$. For the summation on $\sin(m\pi/L)(x-x')$, we have

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\sum_{m=1}^{\infty}\sin\frac{m\pi(x-x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$=\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\frac{1}{2}\cot\frac{\pi(x-x')}{2L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

in the interval $-\pi<(\pi/L)(x-x')<+\pi$, which includes the cases $0<(\pi/L)(x-x')<+\pi$ and $-\pi<(\pi/L)(x-x')<0$. The integral on the right hand side contains a singular contribution of the form $L/(x-x')$ which is extracted using the following expansion, Gradsheteyn and Ryzhik ([45], p. 43),

$$\frac{1}{2}\cot\frac{\pi(x-x')}{2L}=\frac{1}{\pi}\frac{L}{x-x'}+\frac{1}{2\pi}\frac{x-x'}{L}\sum_{n=1}^{\infty}\frac{1}{\left(\dfrac{x-x'}{2L}\right)^2-n^2}$$

For the special case $x-x'=0$, we trivially obtain

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\sum_{m=1}^{\infty}\sin\frac{m\pi(x-x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'=0$$

For the remaining summation on $\cos(m\pi/L)(x-x')$, we need to consider three cases. In the interval $0<\frac{\pi}{L}(x-x')<+\pi$ we have

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\sum_{m=1}^{\infty}\cos\frac{m\pi(x-x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$=\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\frac{d}{dx}\sum_{m=1}^{\infty}\frac{L}{m\pi}\sin\frac{m\pi(x-x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$=\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}}\frac{d}{dx}\left(\frac{L-(x-x')}{2}\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx'=0$$

in light of the constraint on $\Theta_\alpha^{(p)}$ given in Eq. (45). We obtain the same result in the interval $-\pi<(\pi/L)(x-x')<0$,

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \sum_{m=1}^{\infty} \cos\frac{m\pi(x-x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$= -\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \frac{d}{dx}\sum_{m=1}^{\infty}\frac{L}{m\pi}\sin\frac{m\pi(-x+x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$= -\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \frac{d}{dx}\left(\frac{L-(-x+x')}{2}\right)\theta_{(\cdot)\alpha}^{(p)}(x')dx' = 0$$

The case $x-x'=0$ requires special attention because the sum on $\cos(m\pi/L)(x-x')$ becomes infinite. In fact, the sum on $\cos(m\pi/L)(x-x')$ behaves like the Dirac delta function since in a small region around $-x_\triangle/2<x_0=x-x'<+x_\triangle/2$ with $x_\triangle>0$,

$$\frac{1}{L}\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \sum_{m=1}^{\infty}\cos\frac{m\pi(x-x')}{L}d(x-x')$$

$$= \frac{1}{L}\lim_{x_\triangle\to 0}\int_{-x_\triangle/2}^{+x_\triangle/2}\sum_{m=1}^{\infty}\cos\frac{m\pi x_0}{L}dx_0$$

$$= \frac{1}{L}\lim_{x_\triangle\to 0}\sum_{m=1}^{\infty}\frac{\sin\dfrac{m\pi x_\triangle}{2L}}{\dfrac{m\pi}{2L}} = \frac{2}{\pi}\lim_{x_\triangle\to 0}\frac{\pi-\dfrac{\pi x_\triangle}{2L}}{2} = 1$$

Therefore, the integral involving the sum $\cos(m\pi/L)(x-x')$ for the case $x-x'=0$ becomes

$$\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \sum_{m=1}^{\infty}\cos\frac{m\pi(x-x')}{L}\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

$$= \int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} L\cdot\delta(x-x')\theta_{(\cdot)\alpha}^{(p)}(x')dx'$$

Appendix D: Constants for Singular Integral Equations

The diagonal and modal matrices $\mathbf{\Lambda}_\alpha$ and $\mathbf{R}_\alpha^*$ appearing in Eq. (68) have the following form

$$\mathbf{\Lambda}_\alpha = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\Lambda_\alpha & 0 \\ 0 & 0 & +\Lambda_\alpha \end{bmatrix}, \quad \mathbf{R}_\alpha^* = \begin{bmatrix} 0 & 1 & 1 \\ -R_{21} & -iR_{22} & iR_{22} \\ 1 & -iR_{33} & iR_{33} \end{bmatrix}$$

where

$$\Lambda_\alpha = \sqrt{\frac{(B_{13}^{*\alpha})^2 B_{22}^{*\alpha} - (B_{12}^{*\alpha})^2 B_{33}^{*\alpha}}{B_{11}^{*\alpha}[(B_{23}^{*\alpha})^2 - B_{22}^{*\alpha}B_{33}^{*\alpha}]}}$$

and

$$R_{21} = \frac{B_{13}^{*\alpha}}{B_{12}^{*\alpha}}, \quad R_{22} = \frac{1}{\Lambda_\alpha}\frac{B_{13}^{*\alpha}B_{23}^{*\alpha} - B_{12}^{*\alpha}B_{33}^{*\alpha}}{B_{22}^{*\alpha}B_{33}^{*\alpha} - (B_{23}^{*\alpha})^2},$$

$$R_{33} = \frac{1}{\Lambda_\alpha}\frac{B_{12}^{*\alpha}B_{23}^{*\alpha} - B_{13}^{*\alpha}B_{22}^{*\alpha}}{B_{22}^{*\alpha}B_{33}^{*\alpha} - (B_{23}^{*\alpha})^2}$$

References

[1] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Groningen, The Netherlands.

[2] Muskhelishvili, N. I., 1953, *Singular Integral Equations*, Noordhoff, Groningen, The Netherlands.

[3] Sneddon, I. N., 1951, *Fourier Transforms*, McGraw-Hill, New York.

[4] Sneddon, I. N., 1961, *Fourier Series*, Routledge and Paul, London.

[5] Westergaard, H. M., 1939, "Bearing Pressures and Cracks," J. Appl. Mech., **66**, pp. A49–A53.

[6] Koiter, W. T., 1959, "An Infinite Row of Collinear Cracks in an Infinite Elastic Sheet," Ing.-Arch., **28**, pp. 168–173.

[7] Sneddon, I. N., and Srivastava, K. N., 1965, "The Stress in the Vicinity of an Infinite Row of Collinear Cracks in an Elastic Body," Proc. R. Soc. Edinburgh, Sect. A: Math. Phys. Sci., **67**, pp. 39–49.

[8] Lowengrub, M., 1966, "A Two-dimensional Crack Problem," Int. J. Eng. Sci.. **4**, pp. 289–299.

[9] Sneddon, I. N., and Lowengrub, M., 1969, *Crack Problems in the Classical Theory of Elasticity*, Wiley, New York.

[10] Erdogan, F., 1962, "On the Stress Distribution in Plates with Collinear Cuts Under Arbitrary Loads," *Proc. of the Fourth U.S. National Congress. of Appl. Mech.*, Vol. 1, pp. 547–553.

[11] Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," J. Appl. Mech., **32**, pp. 418–423.

[12] Delameter, W. R., Herrmann, G., and Barnett, D. M., 1975, "Weakening of an Elastic Solid by a Rectangular Array of Cracks," J. Appl. Mech., **42**, pp. 74–80; Errata: J. Appl. Mech., **44**, p. 190.

[13] Isida, M., 1976, "Analysis of Stress Intensity Factors for Plates Containing Randomly Distributed Cracks," Trans. Jpn. Soc. Mech. Eng., **35**, pp. 1815–1822.

[14] Chen, Y. Z., 1984, "General Case of Multiple Crack Problems in an Infinite Plate," Eng. Fract. Mech., **20**, pp. 295–297.

[15] Horii, H., and Nemat-Nasser, S., 1985, "Elastic Fields of Interacting Inhomogeneities," Int. J. Solids Struct., **21**, pp. 731–745.

[16] Kachanov, M., 1985, "A Simple Technique of Stress Analysis in Elastic Solids with Many Cracks," Int. J. Fract., **28**, pp. R11–R19.

[17] Kachanov, M., 1987, "Elastic Solids with Many Cracks: A Simple Method of Analysis," Int. J. Solids Struct., **23**, pp. 23–43.

[18] Mauge, C., and Kachanov, M., 1994, "Anisotropic Materials with Interacting Arbitrarily Oriented Cracks—Stress Intensity Factors and Crack-Microcrack Interactions," Int. J. Fract., **65**, pp. 115–139.

[19] Du, Y. J., and Aydin, A., 1991, "Interaction of Multiple Cracks and Formulation of Echelon Crack Arrays," Int. J. Numer. Analyt. Meth. Geomech., **15**, pp. 205–218.

[20] Binienda, W. K., 1994, "Stress Intensity Factors for Fully Interacting Cracks in a Multicrack Solid," J. Offshore Mech. Arct. Eng., **116**, pp. 56–63.

[21] Shbeeb, N. I., Binienda, W. K., and Kreider, K. I., 1999, "Analysis of the Driving Forces for Multiple Cracks in an Infinite Nonhomogeneous Plate, Part I: Theoretical Analysis & Part II: Numerical Solutions," J. Appl. Mech., **66**, pp. 492–506.

[22] Erdogan, F., and Gupta, G. D., 1971, "The Stress Analysis of Multi-layered Composites with a Flaw," Int. J. Solids Struct., **7**, pp. 39–61.

[23] Erdogan, F., 1971, "Layered Composites with an Interface Flaw," Int. J. Solids Struct., **7**, pp. 1089–1107.

[24] Chatterjee, S. N., Pindera, M-J., Pipes, R. B., and Dick, B., 1982, "Composite Defect Significance," MSC TFR 1312/1108, NADC Report No. 81034-60, Spring House, PA.

[25] Chatterjee, S. N., 1987, "Three- and Two-Dimensional Stress Fields Near Delaminations in Laminated Composite Plates," Int. J. Solids Struct., **23**, pp. 1535–1549.

[26] Pindera, M-J., 1991, "Local/Global Stiffness Matrix Formulation for Composite Materials and Structures," Composites Eng., **1**, pp. 69–83.

[27] Choi, H. J., and Thangjitham, S., 1994, "The Interlaminar Crack-tip Response in a Fiber-Reinforced Composite Laminate," Int. J. Fract., **66**, pp. 121–138.

[28] Isida, M., 1971, "Effect of Width and Length on Stress Intensity Factors of Internally Cracked Plates Under Various Boundary Conditions," Int. J. Fract. Mech., **7**, pp. 301–316.

[29] Chatterjee, S. N., 1979, "On Interlaminar Defects in Laminated Composites," *Modern Developments in Composite Materials and Structures*, edited by J. R. Vinson, ASME, New York, pp. 1–15.

[30] Chen, Y. Z., 1988, "Multiple Crack Problems for Finite Plate with Arbitrary Contour Configuration," Eng. Fract. Mech., **31**, pp. 289–295.

[31] Zhan, S., Wang, T., and Han, X., 1999, "Analysis of Two-dimensional Finite Solids with Microcracks," Int. J. Solids Struct., **36**, pp. 3735–3753.

[32] Seelig, T., Rafiee, S., and Gross, G., 2000, "A Simple Method for the Investigation of Elastic Bodies of Finite Size Containing Many Cracks," Int. J. Eng. Sci., **38**, pp. 1459–1472.

[33] Erdogan, F., Gupta, G. D., and Cook, T. S., 1973, "Numerical Solution of Singular Integral Equations," *Methods of Analysis and Solutions of Crack Problems*, edited by G. C. Sih, Netherlands, Noordhoff, Leyden, pp. 368–425.

[34] Kim, J-H., and Paulino, G. H., 2005, "Consistent Formulations of the Interaction Integral Method for Fracture of Functionally Graded Materials," J. Appl. Mech., **72**, pp. 351–364.

[35] Lekhnitskii, S. G., 1981, *Theory of Elasticity of an Anisotropic Body*, Mir, Moscow (English translation).

[36] Bufler, H., 1971, "Theory of Elasticity of a Multilayered Medium," J. Elast., **1**, pp. 125–143.

[37] Rowe, R. K., and Booker, J. R., 1982, "Finite Layer Analysis of Nonhomoge-

neous Soils," J. Eng. Mech., **108**, pp. 115–132.
[38] Pindera, M-J., and Lane, M. S., 1993, "Frictionless Contact of Layered Half Planes: Part I - Analysis," J. Appl. Mech., **60**, pp. 633–639.
[39] Urquhart, E. E., and Pindera, M-J., 1994, "Incipient Separation Between a Frictionless Flat Punch and an Anisotropic Multilayered Half Plane," Int. J. Solids Struct., **31**, pp. 2445–2461.
[40] Bansal, Y., and Pindera, M-J., 2003, "Efficient Reformulation of the Thermoelastic Higher-Order Theory for FGMs," J. Therm. Stresses, **26**, pp. 1055–1092.
[41] Bansal, Y., and Pindera, M-J., 2005, "A Second Look at the Higher-Order Theory for Periodic Multiphase Materials," J. Appl. Mech., **72**, pp. 177–195.
[42] Wang, S. S., and Choi, I., 1983, "The Interface Crack Between Dissimilar Anisotropic Composite Materials," J. Appl. Mech., **50**, pp. 169–178.
[43] Wang, S. S., and Choi, I., 1983, "The Interface Crack Behavior in Dissimilar Anisotropic Composites Under Mixed-Mode Loading," J. Appl. Mech., **50**, pp. 179–183.
[44] Ting, T. C. T., 1996, *Anisotropic Elasticity (Theory and Applications)*, Oxford University Press, Oxford.
[45] Gradsheteyn, I. S., and Ryzhik, I. M., 2000, *Table of Integrals, Series and Products*, 6th ed., Academic, New York.

Plane Analysis of Finite Multilayered Media With Multiple Aligned Cracks—Part II: Numerical Results

Linfeng Chen

Marek-Jerzy Pindera

Civil Engineering Department,
University of Virginia,
Charlottesville, VA 22904-4742

In Part I of this paper, elasticity solutions were developed for finite multilayered domains, weakened by aligned cracks, that are in a state of generalized plane deformation under two types of end constraints. In Part II we address computational aspects of the developed solution methodology that must be implemented numerically, and present new fundamental results that are relevant to modern technologically important applications involving defect criticality of multilayers. The computational aspects include discussion of the various parameters that influence the accuracy with which numerical results are generated and subsequent verification by a comparison with previously reported results in the limit, as the in-plane dimensions become very large and layer anisotropy vanishes. The present solution quantifies the thus far undocumented effects of finite dimensions, crack location, and material anisotropy due to a unidirectional fiber-reinforced layer's orientation on Mode I, II, and III stress intensity factors in composite multilayers with single and multiple interacting cracks under different loading and boundary conditions. These effects may have significant impact on defect criticality of advanced multilayered structures when cracks are in close proximity to vertical and horizontal boundaries.
[DOI: 10.1115/1.2201889]

1 Synopsis

In Part I of this paper, an exact solution was presented to the plane elasticity problem of a multilayer containing horizontal cracks under two types of end-face boundary constraints. The multilayer has finite dimensions in the analysis plane and is infinitely long in the out-of-plane direction. The end constraints mimic horizontal and vertical pins at the multilayer's left and right faces. The number of layers is arbitrary as is the number of horizontal cracks on any interface separating two adjacent layers. The individual layers can be (transversely) isotropic, orthotropic, or monoclinic, as is typical of advanced composite laminated plates. Monoclinic layers are obtained by rotating a transversely isotropic or orthotropic ply by an angle about the axis perpendicular to the layered configuration.

Following the local/global stiffness matrix formulation and solution methodology for composite structures; cf. Pindera [1], and specifying the traction vector $\mathbf{T}_\alpha^+(x)$ on the bottom face of the pth crack on the cracked αth interface (Fig. 2 in Part I), the problem was reduced to the determination of the unknown crack-opening displacement functions $\mathbf{\Theta}_\alpha^{(p)}$ for each crack governed by the following singular integral equations valid in the interval $c_\alpha^{(p)} < x < d_\alpha^{(p)}$ (Eq. (1) of Part I),

$$\mathbf{T}_\alpha^+(x) = \bar{\mathbf{A}}_\alpha^* \mathbf{\Theta}_\alpha^{(p)}(x) + \frac{1}{\pi}\int_{c_\alpha^{(p)}}^{d_\alpha^{(p)}} \bar{\mathbf{B}}_\alpha^* \frac{\mathbf{\Theta}_\alpha^{(p)}(x')}{x'-x}dx' + \frac{1}{\pi}\sum_{\beta=2}^{n}\sum_{q=1}^{Q^{(\beta)}}\int_{c_\beta^{(q)}}^{d_\beta^{(q)}} \bar{\mathbf{K}}_{\alpha\beta}(x,x')\mathbf{\Theta}_\beta^{(q)}(x')dx' + \mathbf{F}_\alpha(x) \quad (1)$$

In the above, n in the summation limit is the total number of layers (assuming that in the most general case all interfaces are cracked), $Q^{(\beta)}$ is the number of cracks on the βth interface, $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ are regular Fredholm kernels, $\mathbf{F}_\alpha(x)$ is the specified external load vector, and $\bar{\mathbf{A}}_\alpha^*$, $\bar{\mathbf{B}}_\alpha^*$ are constant square matrices whose elements depend only on the material properties of adjacent layers. Specifically, from Eq. (59) of Part I repeated here for convenience, these matrices are

$$\bar{\mathbf{A}}_\alpha^* = \begin{bmatrix} 0 & -B_{12}^{*\alpha} & -B_{13}^{*\alpha} \\ B_{12}^{*\alpha} & 0 & 0 \\ B_{13}^{*\alpha} & 0 & 0 \end{bmatrix}, \quad \bar{\mathbf{B}}_\alpha^* = \begin{bmatrix} B_{11}^{*\alpha} & 0 & 0 \\ 0 & B_{22}^{*\alpha} & B_{23}^{*\alpha} \\ 0 & B_{23}^{*\alpha} & B_{33}^{*\alpha} \end{bmatrix} \quad (2)$$

where $B_{ij}^{*\alpha}$'s are given in terms of the asymptotic values of the elements of the local stiffness matrices of adjacent layers, namely $\mathbf{B}_\alpha^* = \mathbf{K}_{11}^{*,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}$ with $[\mathbf{K}^{*,\alpha}]^{-1} = [\mathbf{K}_{22}^{*,\alpha-1} + \mathbf{K}_{11}^{*,\alpha}]^{-1}$. Here, we consider horizontal cracks situated between adjacent layers with the same elastic moduli so that $\bar{\mathbf{A}}_\alpha^* = \mathbf{0}$, thereby avoiding unnecessary complications associated with the physically inadmissible, crack-tip oscillatory stress behavior that occurs when the adjacent plies have different elastic moduli, cf. Ting [2] for an excellent survey of the various approaches to overcome this difficulty.

The unknown crack-opening density functions for the normalized pth crack in the interval $|t_\alpha^{(p)}| \leq 1$ along the αth interface are approximated by a series of Chebyshev polynomials $T_j(t_\alpha^{(p)})$ of the first kind, with the associated influence coefficients $\mathbf{C}_{j(\alpha)}^{(p)}$ multiplied by the weight function $\boldsymbol{\omega}_\alpha(t_\alpha^{(p)}) = (1-t_\alpha^{(p)})^{-\frac{1}{2}}(1+t_\alpha^{(p)})^{-1/2}$

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 18, 2005; final manuscript received April 1, 2006. Review conducted by Y. Huang. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–SantF 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.

$$\boldsymbol{\Theta}_{\alpha}^{(p)}(t_{\alpha}^{(p)}) = \frac{(d_{\alpha}^{(p)} - c_{\alpha}^{(p)})}{2}(1 - t_{\alpha}^{(p)})^{-1/2}(1 + t_{\alpha}^{(p)})^{-1/2}\sum_{j=0}^{N_j} \frac{(2j)!}{2^{2j}(j\,!)^2}\mathbf{C}_{j(\alpha)}^{(p)} T_j(t_{\alpha}^{(p)}) \tag{3}$$

Using a collocation technique developed by Erdogan et al. [3] based on a Chebyshev polynomial orthogonality identity, the solution of the above singular integral equations is reduced to a system of algebraic equations in the unknown influence coefficients (see Eq. (76) of Part I),

$$\frac{\pi}{2}\theta_j^{(-1/2,-1/2)}\mathbf{C}_{(j+1)(\alpha)}^{(p)} + \sum_{\beta=2}^{n}\sum_{q=1}^{Q^{(\beta)}}\sum_{k=0}^{N_j}\frac{(d_{\beta}^{(q)} - c_{\beta}^{(q)})}{2}\mathbf{D}_{jk(\alpha\beta)}^{(pq)}\mathbf{C}_{k(\beta)}^{(q)} = \mathbf{G}_{j(\alpha)}^{I(p)} + \mathbf{G}_{j(\alpha)}^{II(p)} \tag{4}$$

for each $j=0,1,2,\ldots,N_j$. The constant matrices and vectors $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$, $\mathbf{G}_{j(\alpha)}^{I(p)}$, $\mathbf{G}_{j(\alpha)}^{II(p)}$ are given below

$$\mathbf{D}_{jk(\alpha\beta)}^{(pq)} = \frac{(2j)!}{2^{2j}(j\,!)^2}\cdot\frac{(2k)!}{2^{2k}(k\,!)^2} \times \int_{-1}^{+1}\left[\int_{-1}^{+1}\mathbf{h}_{\alpha\beta}(t_{\alpha}^{(p)},\tau_{\beta}^{(q)})T_j(t_{\alpha}^{(p)})\hat{\boldsymbol{\omega}}_{\alpha}(t_{\alpha}^{(p)})dt_{\alpha}^{(p)}\right] \times T_k(\tau_{\beta}^{(q)})\boldsymbol{\omega}_{\beta}(\tau_{\beta}^{(q)})d\tau_{\beta}^{(q)}$$

$$\mathbf{G}_{j(\alpha)}^{I(p)} = -\frac{(2j)!}{2^{2j}(j\,!)^2}\int_{-1}^{+1}\bar{\mathbf{B}}_{\alpha}^{*-1}\mathbf{F}_{\alpha}(t_{\alpha}^{(p)})T_j(t_{\alpha}^{(p)})\hat{\boldsymbol{\omega}}_{\alpha}(t_{\alpha}^{(p)})dt_{\alpha}^{(p)}$$

$$\mathbf{G}_{j(\alpha)}^{II(p)} = \frac{(2j)!}{2^{2j}(j\,!)^2}\int_{-1}^{+1}\bar{\mathbf{B}}_{\alpha}^{*-1}\mathbf{T}_{\alpha}^{+}(t_{\alpha}^{(p)})T_j(t_{\alpha}^{(p)})\hat{\boldsymbol{\omega}}_{\alpha}(t_{\alpha}^{(p)})dt_{\alpha}^{(p)} \tag{5}$$

where $\hat{\boldsymbol{\omega}}_{\alpha}(t_{\alpha}^{(p)}) = (1-t_{\alpha}^{(p)})^{1/2}(1+t_{\alpha}^{(p)})^{1/2}$ and $\mathbf{h}_{\alpha\beta}(x,x') = \bar{\mathbf{B}}_{\alpha}^{*-1}\bar{\mathbf{K}}_{\alpha\beta}(x,x')$.

Part II of the paper is focused on the verification of this solution and a presentation of new fundamental results involving crack interactions in finite multilayer geometries with different crack arrays and material anisotropies. These results have been generated for multilayers with end-face boundary conditions that mimic horizontal pins. In light of the numerical implementation of the developed solution, the computational approach is first discussed vis-a-vis solution's convergence in the following section. The third section is devoted to the investigation of previously undocumented effects of laminate dimensions, crack location, and layer anisotropy on the stress intensity factors of single and multiple horizontal cracks situated at different elevations under internal and external loading, including cracks vertically stacked in a single column as well as cracks diagonally arranged in an echelon array. The solution's verification is carried out by comparison with known results in the limit as the in-plane dimensions become very large for anisotropic layers oriented in a manner that produces isotropic behavior in the analysis plane. Finally, limitations of this solution and future prospects are discussed in the last section before a summary and conclusions.

2 Computational Approach and Convergence

An accurate solution of the algebraic system of equations for the unknown coefficients $\mathbf{C}_{j(\alpha)}^{(p)}$ of the crack-opening displacement representation, Eq. (3), requires accurate evaluation of the coefficients $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$, $\mathbf{G}_{j(\alpha)}^{I(p)}$, and $\mathbf{G}_{j(\alpha)}^{II(p)}$ that must be done numerically. The accuracy with which $\mathbf{G}_{j(\alpha)}^{I(p)}$ and $\mathbf{G}_{j(\alpha)}^{II(p)}$ are calculated depends on the convergence behavior of the elements in the Fourier series domain that form the external loading vector $\mathbf{F}_{\alpha}(x)$ and the internal traction $\mathbf{T}_{\alpha}^{+}(x)$ applied to the crack faces, the latter being typically constant. The convergence behavior of the $\mathbf{F}_{\alpha}(x)$ elements depends on the smoothness of the applied loading. Single integrals are employed to determine $\mathbf{G}_{j(\alpha)}^{I(p)}$ and $\mathbf{G}_{j(\alpha)}^{II(p)}$ for the specified $\mathbf{T}_{\alpha}^{+}(x)$ and $\mathbf{F}_{\alpha}(x)$, which require much less computational effort than in the calculation of $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$ discussed below.

The accuracy with which the coefficients $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$ are determined depends on the following three items: evaluation of the kernel functions $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ appearing in the associated integrals; representation of the crack-opening displacements; and integration of double integrals. The first two items involve the use of a sufficient number of terms in the Fourier series representation of the displacement fields in the individual layers, and two additional summations involving the function $\Omega(x,x')$ [or $\Omega(x',x)$] and the crack-opening displacement approximations based on the Chebyshev polynomials. Specifically, the Fredholm kernels $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ that are directly related to the functions $\mathbf{h}_{\alpha\beta}(x,x')$ in the kernels associated with $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$ involve the first two summations, as seen by listing typical elements of these Fredholm kernels below for horizontally pinned ends,

$$\bar{K}_{11}^{\alpha\beta}(x,x') = \frac{2\pi}{L}\left[\kappa_{11}^{\alpha\beta} + \left(\Omega(x,x') + \frac{L}{2\pi}\frac{1-\delta_{pq}}{x'-x}\right)\delta_{\alpha\beta}B_{11}^{*\alpha}\right]$$

$$\bar{K}_{12}^{\alpha\beta}(x,x') = \frac{2\pi}{L}\kappa_{12}^{\alpha\beta}$$

where

$$\kappa_{11}^{\alpha\beta}(x,x') = \sum_{m=1}^{\infty}A_{11}^{m(\alpha\beta)}\cos\left(\frac{m\pi x}{L}\right)\sin\left(\frac{m\pi x'}{L}\right)$$

$$\kappa_{12}^{\alpha\beta}(x,x') = \sum_{m=1}^{\infty}(-A_{12}^{m(\alpha\beta)})\cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{m\pi x'}{L}\right)$$

The elements $A_{ij}^{m(\alpha\beta)}$ given by

$$\mathbf{A}_{\alpha\beta}^{m} = \mathbf{K}_{11}^{m,\alpha}\mathbf{G}_{\alpha\beta}^{m} + \mathbf{K}_{12}^{m,\alpha}\mathbf{G}_{(\alpha+1)\beta}^{m} + \bar{\mathbf{K}}_{11}^{m,\alpha}[\mathbf{K}^{*,\alpha}]^{-1}\mathbf{K}_{22}^{*,\alpha-1}\delta_{\alpha\beta} - \mathbf{K}_{12}^{m,\alpha}[\mathbf{K}^{*,\alpha+1}]^{-1}\mathbf{K}_{11}^{*,\alpha+1}\delta_{(\alpha+1)\beta} \tag{6}$$

where $\mathbf{G}_{\alpha\beta}^{m}$ are the appropriate $(\alpha\beta)$ elements of the inverse of the global stiffness matrix $\mathbf{K}^{m}$ multiplied by the matrix $\mathbf{L}^{m}$ that defines the locations of cracked interfaces, i.e., $\mathbf{G}^{m}=[\mathbf{K}^{m}]^{-1}\mathbf{L}^{m}$, see Eq. (38) of Part I, are obtained by inverting the global stiffness matrix for each harmonic m as many times as is necessary to obtain convergence. The function $\Omega(x,x')$ defined on the square $0<x/L,x'/L<1$, which is reproduced below for convenience

$$\Omega(x,x') = \frac{1}{4}\cot\frac{\pi(x+x')}{2L} - \frac{x-x'}{4\pi L}\sum_{n=1}^{N_\Omega}\frac{1}{\left(\frac{x-x'}{2L}\right)^2 - n^2} \tag{7}$$

also contains a summation that must be carried out in a convergent fashion. These two summations are discussed below in more detail.

The convergence behavior of $A_{ij}^{m(\alpha\beta)}$ that contribute to the kernels $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ is established by determining the value of $m=N_m$ at which these elements become negligibly small, which in turn produces integrable kernels. For a single row of collinear cracks $\alpha=\beta$ so that the elements $A_{ij}^{m(\alpha\beta)}$ reduce to $A_{ij}^{m(\alpha\alpha)}$, while for vertically offset cracks both $\alpha=\beta$ and $\alpha\neq\beta$ are present so that all combinations $A_{ij}^{m(\alpha\beta)}$ must be investigated. Under special loading conditions, the form of $A_{ij}^{m(\alpha\beta)}$ is known in advance. For instance, $A_{31}^{m(\alpha\beta)}=A_{32}^{m(\alpha\beta)}=A_{13}^{m(\alpha\beta)}=A_{23}^{m(\alpha\beta)}=0$ if in-plane and out-of-

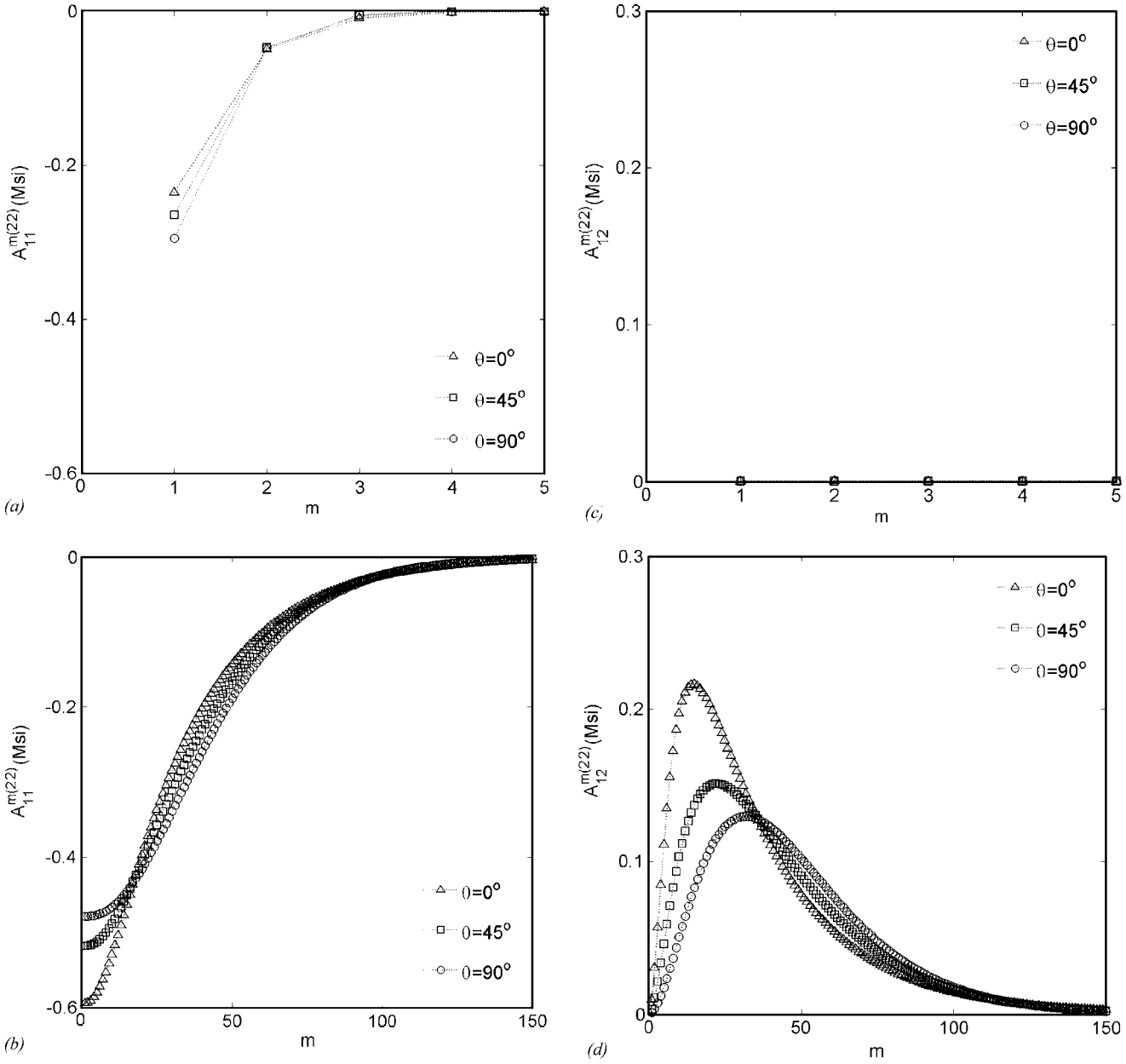


Fig. 1 Convergence of the $A_{11}^{m(\alpha\beta)}$ and $A_{12}^{m(\alpha\beta)}$ elements with the harmonic number m for a two-layer unidirectional graphite/epoxy laminate with three fiber rotations about the z axis and increasingly thinner top layers characterized by $\mu=h_1/H=0.5, 0.01$ (top and bottom)

plane tractions are decoupled at the cracked interface. Similarly, $A_{ii}^{m(\alpha\beta)} >> A_{ij}^{m(\alpha\beta)}$ $(i,j=1,2,3)$ if normal traction is much more dominant than shear tractions at the cracked interface.

The rates of convergence of the elements $A_{ij}^{m(\alpha\beta)}$ depend on the ratio of the layer thickness to laminate length and the material anisotropy of the layers adjacent to cracks, as suggested by Eq. (6). Obviously, crack size and collinear crack location along a given interface have no effect on N_m. This is demonstrated in Fig. 1 for a square plate divided into two layers with decreasing ratios of the top layer to plate thickness, namely $h_1/H=0.5$, 0.01. The material used to demonstrate the effect of anisotropy in this case was unidirectional graphite/epoxy. This material was used throughout the entire paper as discussed in the following section. The effect of anisotropy was demonstrated for the following fiber rotations about the z axis: $\theta=0$ deg, 45 deg, 90 deg (see Fig. 1 of Part I). As explained in Sec. 3, these orientations produce orthotropic, monoclinic, and isotropic elastic properties in the analysis plane $x-z$.

The effects of layer thickness and material anisotropy on the convergence of the elements $A_{11}^{m(\alpha\beta)}$ and $A_{12}^{m(\alpha\beta)}$ with increasing harmonic number m are demonstrated in Fig. 1, where increasingly larger values of N_m are required for convergence as the top layer thickness becomes increasingly small. As observed, material anisotropy becomes increasingly less important in the convergence behavior as m increases. Further, coupling between the opening and in-plane sliding crack opening modes represented by $A_{12}^{m(\alpha\beta)}$ vanishes when the crack interface is in the middle of the plate $(h_1/H=0.5)$, and increases with decreasing distance from the top surface. Coupling between the opening and out-of-plane slid-

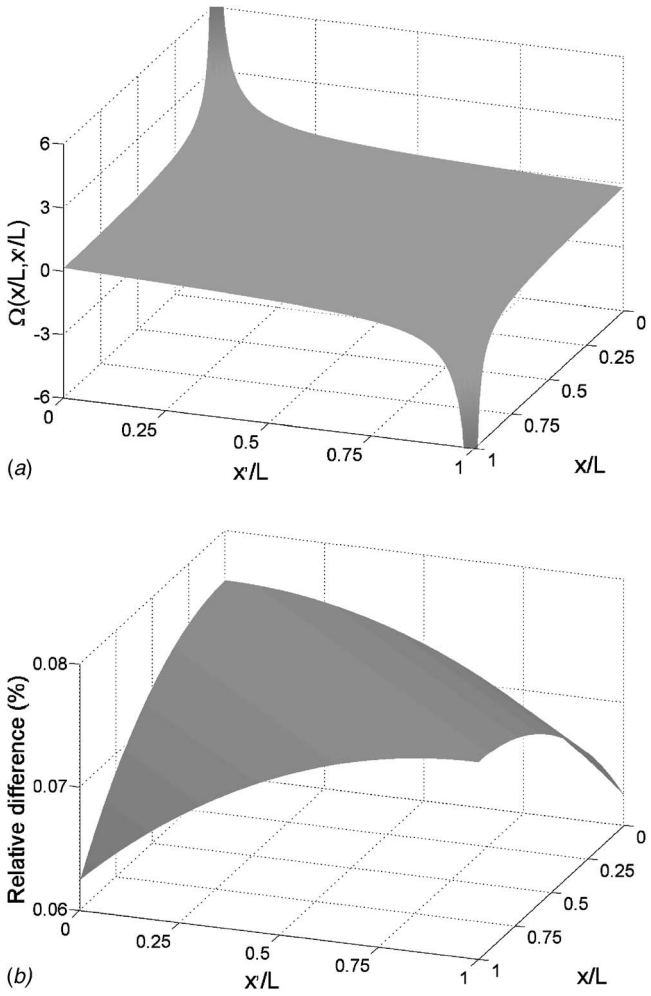


Fig. 2 Surface plot of $\Omega(x,x')$ in the nondimensionalized space

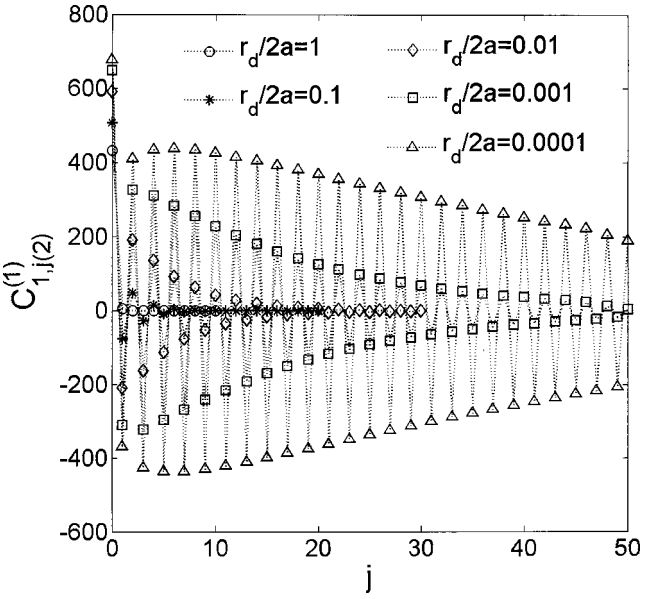


Fig. 3 Convergence of Chebyshev polynomial coefficients used in approximating the crack opening displacements of two collinear cracks with different separation distances in an infinite isotropic plate under pure Mode I loading

ing crack opening modes represented by $A_{13}^{m(\alpha\beta)}$ (not shown) vanishes for the orthotropic and isotropic configurations. For the monoclinic configuration, it exhibits the same trend as $A_{12}^{m(\alpha\beta)}$ with decreasing top layer thickness but is much weaker. The corresponding effects on the convergence of the elements $A_{22}^{m(\alpha\beta)}$ and $A_{33}^{m(\alpha\beta)}$ exhibit similar trends and therefore are not shown.

The second summation that contributes to $\bar{\mathbf{K}}_{\alpha\beta}(x,x')$ involves the calculation of $\Omega(x,x')$, which can be graphically represented by a surface. The surface of $\Omega(x,x')$ is plotted in Fig. 2(*a*) over the domain $0<x/L,x'/L<1$ for $N_\Omega=400$, which peaks when $(x/L,x'/L)\rightarrow(0,0)$ or $(1,1)$. Since the elements of the kernel $\bar{\mathbf{K}}_{\alpha\beta}$ in Eq. (1) are defined in the interval $c_\alpha^{(p)}<x<d_\alpha^{(p)}$, the function $\Omega(x,x')$ controls the extent of crack-tip interaction with the vertical boundaries. In particular, this interaction is strong when the crack is situated close to the left or right boundary such that the crack-tip and $\Omega(x,x')$ domains are characterized by $x/L,x'/L\rightarrow 0$ or $x/L,x'/L\rightarrow 1$. The crack-vertical boundary interaction controlled by $\Omega(x,x')$ for horizontally pinned end constraints is the same as collinear crack-crack interaction if one imagines another fictitious crack symmetrically located on the other side of the same vertical boundary due to the problem's symmetry and intrinsic periodicity of the half-range Fourier expansion. The presence of symmetry reduces a two-crack interaction problem to a single crack-vertical boundary interaction problem with less than half the required computational effort if both material and loading symmetry are preserved. The same reasoning can be applied in the reduction of any finite-length layered component possessing plane(s) of symmetry along its length in the x direction to the analysis of the smallest periodic unsymmetric length segment. Figure 2(*b*) illustrates that the maximum relative difference between the values of $\Omega(x,x')$ as the summation limit N_Ω increases from 400 to 800 is only 0.076%, which demonstrates that $\Omega(x,x')$ converges at $N_\Omega=400$, regardless of the type of interaction.

The third summation that affects the solution's accuracy involves Chebyshev polynomials, multiplied by unknown coefficients, which represent the actual crack shape upon superposition. The number of these polynomials necessary to accurately capture the crack-opening displacement is determined by studying the rates of decay of the influence coefficients $\mathbf{C}_{j(\alpha)}^{(p)}$ for each crack and application, and then truncating the number of Chebyshev polynomials accordingly. A small N_j is enough when the crack shape is regular while more terms are needed in the case of a crack distorted, for instance, by an extremely strong collinear crack interaction, by the presence of vertically offset cracks nearby, or by the proximity to vertical or horizontal boundaries. When the shape of the crack opening displacement is not too distorted, numerical results indicate that $N_j=10$ is sufficient for most applications. In the case of a strong collinear crack interaction, however (which is an example considered in Sec. 3), substantially more terms are required. This is illustrated in Fig. 3, which demonstrates the rate of convergence of the influence coefficients $\mathbf{C}_{j(\alpha)}^{(p)}$ for different numbers of Chebyshev polynomials used to approximate the crack opening displacements of two collinear cracks in an infinite isotropic plate as a function of the normalized separation distance $r_d/2a$. As observed, the number of Chebyshev polynomials required to accurately approximate the crack opening displacement rapidly increases with decreasing crack separation distance. At large $r_d/2a$ values, the Chebyshev polynomial coefficients quickly die out, requiring only a few Chebyshev polynomials to yield a good crack opening displacement approximation (as $r_d/2a\rightarrow\infty$, the number of terms goes to 1 and we recover the single noninteracting crack solution). At small $r_d/2a$ values, a large number is required. The results for $r_d/2a$

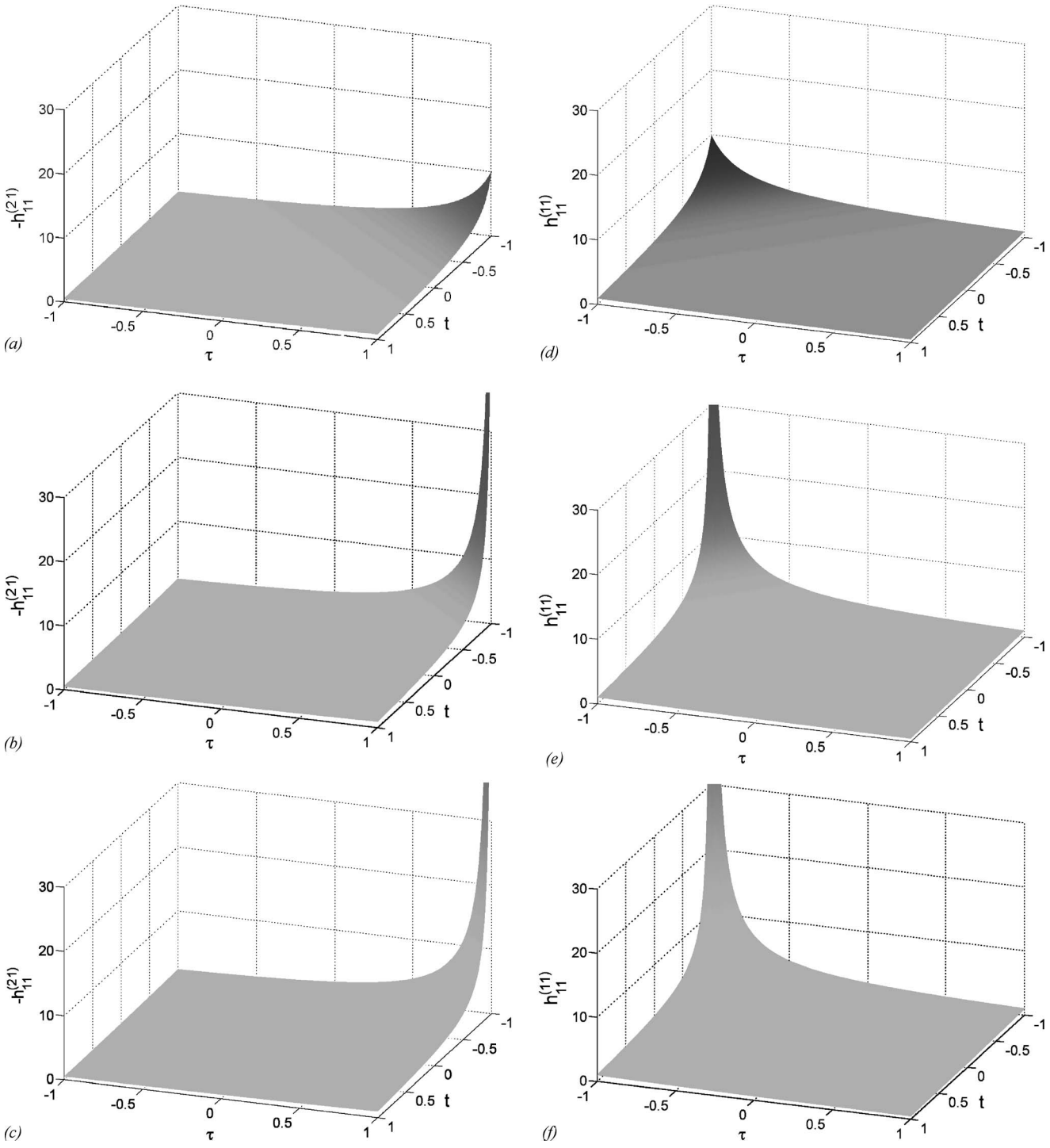


Fig. 4 Surface plots of the $h_{11}^{(pq)}(x, x')$ elements for the three normalized separation distances $r_d/2a=0.1$, 0.01, 0.001 (top, middle and bottom): (*a*)-(*c*) two collinear crack interaction; (*d*)-(*f*) single crack interaction with vertical boundary

$=0.001$ were generated using 50 Chebyshev polynomials. As observed, a substantially greater number is required to obtain an accurate solution when $r_d/2a=0.0001$.

The most time consuming aspect of setting up Eq. (4) is the calculation of the double integrals in $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$. This calculation must be carried out carefully in the presence of strong crack interaction, especially interaction involving adjacent collinear cracks situated very close to each other. For instance, the integrand $\mathbf{h}_{\alpha\beta}(t_\alpha^{(p)}, \tau_\beta^{(q)})$ increases dramatically in the vicinity of two collinear inner crack tips, where strong crack interaction occurs. Such behavior is observed in Fig. 4(*a*)–4(*c*) in the case of two collinear cracks centrally positioned in a large unidirectional graphite/epoxy plate subjected to normal loading with the fiber rotation of 90 deg about the z axis, which produces isotropic elastic material parameters in the $x-z$ plane. In this figure, surface plots of the element $h_{11}^{(\alpha\beta)}$ for $(p,q)=(2,1)$ are given for the normalized separation distances $r_d/2a=0.1$, 0.01, 0.001, showing a rapid increase in the magnitude of this element in the vicinity of the normalized coordinates $(-1,+1)$. These coordinates represent interaction of the left corner of the right crack with the right

corner of the left crack. Similar behavior is observed for the remaining diagonal elements $h_{22}^{(\alpha\beta)}$, $h_{33}^{(\alpha\beta)}$, except for the sign reversal. In the extreme situation when two collinear cracks almost coalesce, $\mathbf{h}_{\alpha\beta}(t_\alpha^{(p)}, \tau_\beta^{(q)})$ becomes nearly singular and direct evaluation of $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$ cannot be performed accurately. Similar distributions are observed in the presence of a vertical boundary halfway between the two cracks, as observed in Fig. 4(*d*)–4(*f*). In this case, $(p,q)=(1,1)$ and the rapid increase of $h_{11}^{(\alpha\beta)}$ with decreasing distance from the vertical boundary left of the crack occurs in the vicinity of the coordinates $(-1,-1)$. This crack-vertical boundary interaction is due to the contribution from $\Omega(x,x')$ in the expressions for the Fredholm kernels, in contrast with the crack-crack interaction due to the term $(L/2\pi)(1-\delta_{pq})/(x'-x)$.

We mention that some special properties of $\mathbf{h}_{\alpha\beta}(t_\alpha^{(p)}, \tau_\beta^{(q)})$ can be deduced in advance of performing the actual numerical integration in order to increase the computational algorithm's efficiency. For instance, in the case of isotropic and orthotropic materials, decoupling of the in-plane and out-of-plane crack opening displacement modes under single mode loading leads to $h_{13}^{(\alpha\beta)}=h_{23}^{(\alpha\beta)}=h_{31}^{(\alpha\beta)}=h_{32}^{(\alpha\beta)}=0$. Also, in those situations when the normal tractions are dominant, $h_{12}^{(\alpha\beta)}\approx 0$ and $h_{21}^{(\alpha\beta)}\approx 0$. In the results shown in Fig. 4, all these conditions are satisfied due to the layer's isotropy, crack's symmetric location, and applied normal loading that eliminates the presence of shear modes along the crack plane.

3 Numerical Results

Numerical results for a single crack, two interacting cracks and multiple cracks in homogeneous and layered configurations are presented to demonstrate the effect of finite dimensions and layer anisotropy on the stress intensity factors. The effect of material anisotropy is demonstrated by rotating a unidirectional graphite/epoxy composite layer, with the fiber direction initially coincident with the x axis, by the angle θ about the z axis. Figure 5 illustrates the variation of the elastic stiffness matrix coefficients $\bar{C}_{ij}(\theta)$, with the rotation angle calculated using the transformation equations given by Eq. (4) in Part I of this paper. The homogenized or macroscopic engineering moduli of this composite used to calculate the elastic stiffness matrix elements C_{ij} in the transformation equations were generated using the micromechanics model called FVDAM (formerly HFGMC), Bansal and Pindera [4,5]. Table 1 lists the elastic moduli of graphite fibers and epoxy matrix used in the calculations, which were performed based on the unit cell containing 65% fibers by volume arranged in a hexagonal array shown in Fig. 6. As discussed by Nye [6], such a unit cell produces homogenized transversely isotropic properties for a unidirectional composite, which indeed are correctly predicted by the employed micromechanics model, Table 2. The plane of isotropy is the x_2-x_3 plane which is orthogonal to the fiber axis denoted by x_1 in Fig. 6. Therefore, the initial fiber orientation of θ=0#x; with the fiber axis coincident with the global x axis produces an apparently orthotropic material in the analysis plane $x-z$, with distinct Young's moduli E_{xx}, E_{zz}, distinct Poisson's ratios ν_{xz}, ν_{zx}, and shear modulus G_{xz}, which is not related to E_{zz} and ν_{xz} through the transverse isotropy relations. At θ=90 deg, the homogenized moduli of the unidirectional composite become isotropic in the analysis plane, with the isotropic relations $E_{xx}=E_{zz}$, $\nu_{xz}=\nu_{zx}$, and $G_{xz}=E_{xx}/2(1+\nu_{xz})$. At both limits of the fiber rotation angle θ =0 deg and 90 deg about the z axis, the in-plane and out-of-plane responses are uncoupled, i.e., $B_{23}^*=0$, as seen in Fig. 7, which illustrates the variation of the $\bar{\mathbf{B}}_\alpha^*$ elements with θ. Therefore, in the limit as the layers' dimensions become very large with respect to the crack length, the present solution for θ=90 deg is compared with solutions available in the literature for single and multiple crack problems on infinite isotropic domains for verification purposes, noting that $B_{11}^*=B_{22}^*$ at this rotation angle; Fig. 7.

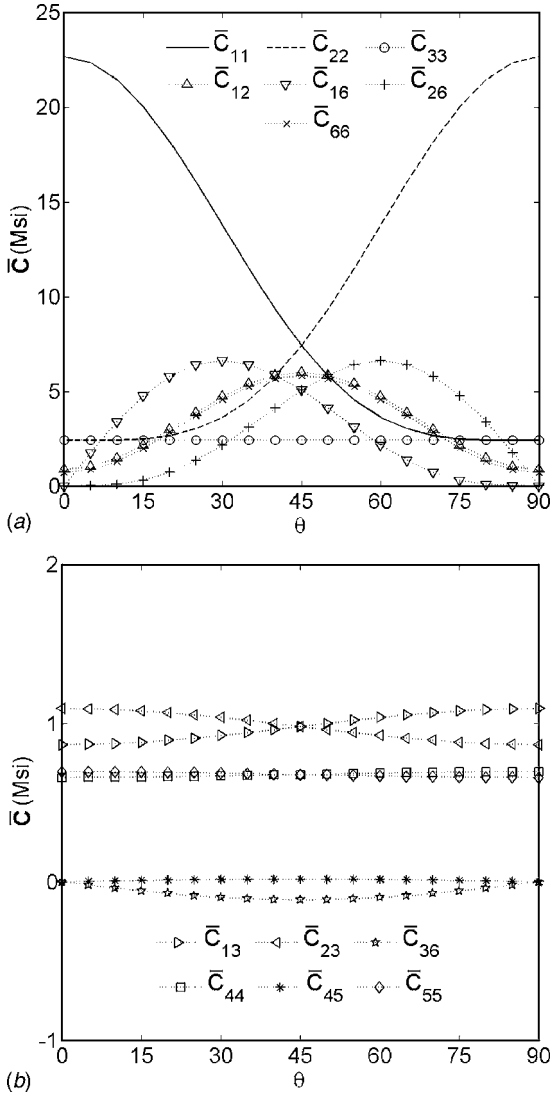


Fig. 5 Effective moduli of the graphite/epoxy unidirectional composite as a function of the fiber rotation angle θ about the *z* axis

We start with a single crack problem, progress to two interacting cracks that are either collinear or vertically offset, and then consider multiple horizontal cracks that are either vertically stacked in a single column or diagonally stacked in an echelon array in spatially homogeneous and layered graphite/epoxy laminates. In all examples, the overall length of a layer or a layered laminate is L and its height is H. The layer (or laminate) height to length ratio is denoted by $\rho=H/L$ and the layer (or laminate) to crack length ratio is denoted by $\eta=L/2a$. Another geometric parameter that represents the top layer to laminate thickness ratio is denoted by $\mu=h_1/H$. The effect of these parameters as well as

Table 1 Elastic properties of the constituent phases (graphite fibers and epoxy matrix) used in the FVDAM calculations. The graphite fiber is transversely isotropic in the x_2-x_3 plane orthogonal to the fiber axis denoted by x_1, while the epoxy matrix is fully isotropic.

Material	E_{11} (Msi)	E_{22} (Msi)	G_{12} (Msi)	ν_{12}	ν_{23}
Graphite fiber	33.800	3.350	1.300	0.200	0.400
Epoxy matrix	0.776	0.776	0.287	0.350	0.350

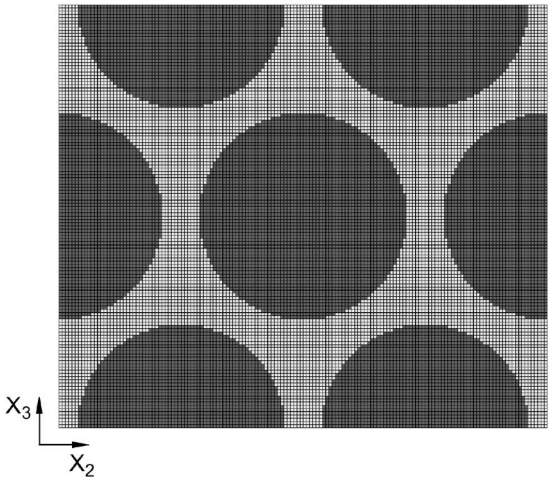


Fig. 6 Repeating unit cell of a transversely isotropic graphite/epoxy unidirectional composite with 65% fiber content used in the FVDAM calculation of the homogenized elastic properties. The hexagonal fiber arrangement in the 2–3 plane ensures transverse isotropy in this plane.

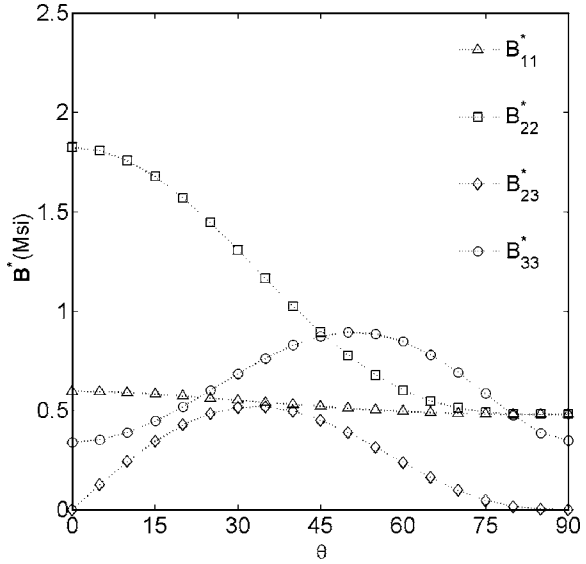


Fig. 7 Elements of the matrix $\bar{B}^*_\alpha$ of the graphite/epoxy unidirectional composite as a function of the fiber rotation angle θ about the z axis

layer anisotropy due to the layer's rotation about the z axis is investigated for the considered single and double crack problems involving spatially homogeneous configurations in the first two subsections. In the third subsection involving vertically stacked cracks, the effects of rotation angle and manner of external boundary condition application are first examined for spatially uniform laminates. The second part of the third subsection demonstrates the effect of orientational grading on stress intensity factors of diagonally stacked cracks embedded in layers with different elastic parameters due to gradual layer rotation.

3.1 Finite Layer With a Single Crack. We first consider a single crack of length $2a$ centrally positioned in a square layer in the analysis plane, $\rho=1$, $\mu=0.5$, Fig. 8, and investigate the effect of material anisotropy due to the layer rotation by the angle θ about the z axis on Mode I stress intensity factor for three layer-to-crack length ratios $\eta=1.2$, 5, 100. These ratios control the extent of crack interaction with the layer's boundaries, with $\eta\to\infty$ representing a single crack embedded in an infinite medium.

Figure 9 presents the normalized Mode I stress intensity factor K_I/K_I^∞ due to loading by internal pressure of magnitude $p=100$ MPa applied to the crack faces as a function of the rotation angle θ for the three ratios η. The normalizing stress intensity factor $K_I^\infty=\pi p/\sqrt{2}$ represents the value obtained from the infinite plate problem for an isotropic medium. As observed, the layer's anisotropy has virtually no effect on the Mode I stress intensity factor, even for very low η ratios that result in a substantial interaction of the crack tip with the layer's vertical boundaries, producing $K_I/K_I^\infty\simeq 1.7$ when $\eta=1.2$. As $\eta\to\infty$, the regular kernel contributions vanish, and the solution for the crack opening displacement reduces to that for a crack embedded in an infinite homogeneous anisotropic plane. Since the considered problem is symmetric with respect to the geometry and loading, Mode II and III stress intensity factors are zero. The absence of these crack-opening displacement modes, together with little variation in the coefficient B^*_{11} with θ seen in Fig. 7, produces virtually no variation in K_I with the rotation angle. This will not be the case in the presence of all three crack-opening displacement modes, as will be seen in the sequel.

Next, we demonstrate the effects of the two geometric ratios ρ and η on the Mode I stress intensity factor for the $\theta=90$ deg layer orientation when the crack is centrally positioned. At this orientation, the layer's in-plane response is isotropic, and thus our results can be directly compared to available solutions. Figure 10 presents the normalized stress intensity factor K_I/K_I^∞ as a function of the ratio η for the ratios $\rho=0.1$, 0.2, 0.5, 1.0 that represent rectangular layers elongated along the horizontal direction that become increasingly more square. Results for $\rho>1.0$ produce results comparable to the square plate results, and thus are not included here. As expected, the normalized stress intensity factor K_I/K_I^∞ decreases with increasing η at different rates that depend on the layer's aspect ratio ρ, approaching 1.0 as $\eta\to\infty$. As $\eta\to 1.0$, the normalized stress intensity factor K_I/K_I^∞ increases rapidly as the layer becomes progressively thinner, demonstrating the combined influence of the horizontal and vertical boundaries on crack tip interaction.

The last example demonstrates the effect of layer anisotropy and crack's proximity to the layer's upper surface of a square plate for two η ratios chosen to quantify the degree of crack tip interaction with vertical and horizontal boundaries. The crack proximity to the horizontal boundary is controlled by the ratios $\mu=0.5$,

Table 2 Elastic properties of the unidirectional graphite/epoxy composite with the fiber volume fraction 0.65 predicted by the FVDAM model. Note that the model correctly predicts that the hexagonal fiber array produces homogenized elastic moduli characteristic of a transversely isotropic composite with $E_{22}=E_{33}$, $G_{12}=G_{13}$, $\nu_{12}=\nu_{13}$ and $G_{23}=E_{22}/2(1+\nu_{23})$.

Material	E_{11} (Msi)	E_{22} (Msi)	G_{12} (Msi)	G_{23} (Msi)	ν_{12}	ν_{23}
Graphite/epoxy	22.244	1.906	0.696	0.659	0.247	0.446

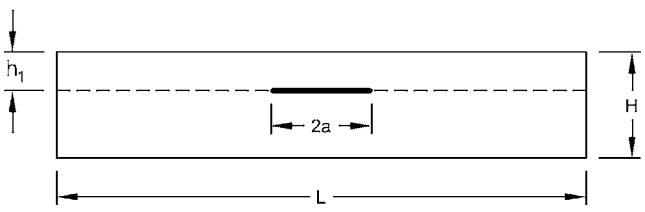


Fig. 8 Geometry of the single-crack problem, showing the geometric parameters used in the numerical study

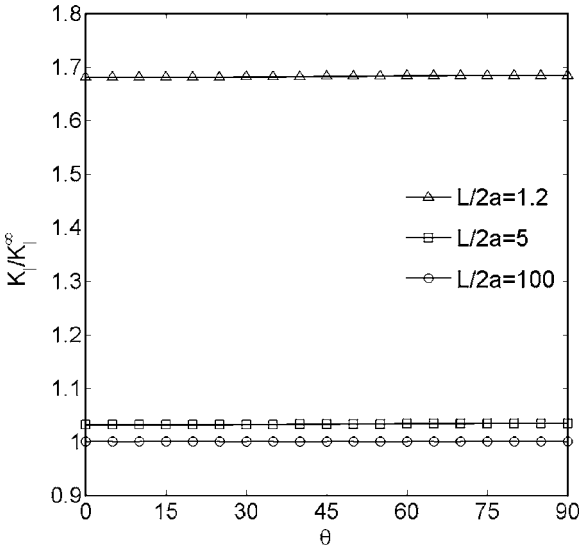


Fig. 9 Normalized Mode I stress intensity factor K_I/K_I^∞ as a function of the fiber rotation angle θ for a square plate with a centrally positioned crack of increasingly larger crack lengths characterized by the ratios $\eta=L/2a=100$, 5, 1.2

0.1, 0.01, whose decreasing magnitudes denote decreasing distances to the layer's top surface. The first ratio is included for reference and represents a centrally positioned crack considered in the first example. It is only for this ratio that Mode II and III stress intensity factors vanish for the applied internal pressure loading. The stress intensity factors have been generated as a function of θ for the above μ ratios and the two ratios $\eta=100$, 1.2, which produce minimal and substantial crack tip interaction with the vertical boundaries, respectively.

The results for the three sets of stress intensity factors, normalized by the Mode I stress intensity factor K_I^∞ for the infinite isotropic plate problem, are shown in Fig. 11. Examining the results for the $\eta=100$ case, we conclude that layer anisotropy has only a noticeable influence when the crack is very close to the top surface. Even in this case, the increase in the Mode I stress intensity factor is just on the order of 5%, as the fiber orientation increases from $\theta=0$ deg to 90 deg. The influence of the layer anisotropy on the Mode II and III stress intensity factors for cracks situated very close to the top surface is more substantial, however these stress intensity factors are a small fraction of the Mode I factor. In particular, the Mode II stress intensity factor is on the order of 2.3%–3.6% of Mode I, depending on the rotation angle θ. The Mode III factor is an order of magnitude lower. The influence of layer anisotropy increases dramatically when the crack length increases relative to the layer length, producing substantial crack tip interaction with vertical boundaries. The results for the $\eta=1.2$ case are included in Fig. 11 and demonstrate not only the substantial influence of layer anisotropy on the three stress intensity factors, but also the possibility that maximum values of these factors can occur at rotation angles other than $\theta=0$ deg or 90 deg, depending on the crack's proximity to the layer's top surface. Further, it is observed that for cracks very close to the layer's top surface, Mode II stress intensity factor dominates while Mode III stress intensity factor is a significant percentage of the Mode I value. These values are substantially greater relative to the corresponding Mode I result for the isotropic infinite plate problem.

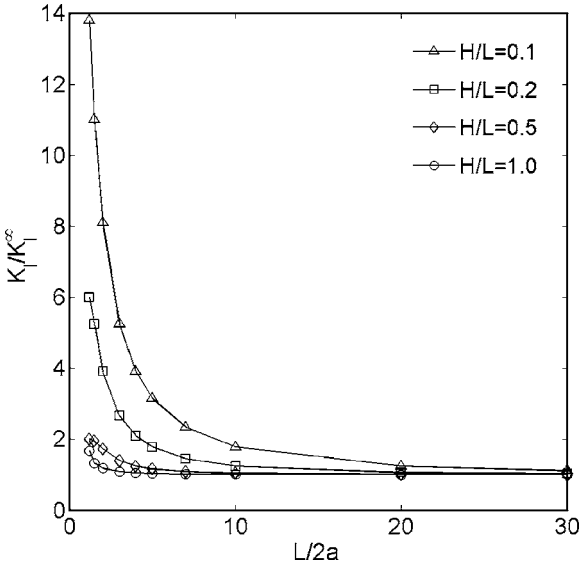


Fig. 10 Normalized Mode I stress intensity factor K_I/K_I^∞ as a function of the ratio $\eta=L/2a$ for a rectangular plate with a centrally positioned crack and the rotation angle $\theta=90$ deg characterized by the ratios $\rho=H/L=1.0$, 0.5, 0.2, 0.1

3.2 Finite Layer With Two Interacting Cracks. In this subsection, we consider a layer with two parallel cracks of equal length $2a$, the offset angle ϕ, and the distance of r_d between the inner crack tips, Fig. 12. The applied loading is uniform internal pressure p applied to the crack faces.

3.2.1 Collinear Cracks: $\phi=0$ deg. We first examine the collinear crack case by setting the offset angle $\phi=0$ deg, and investigate the effect of rotation angle θ on the stress intensity factors at the inner and outer crack tips for different crack distances relative to the top surface defined by the ratios $\mu=0.5$, 0.1, 0.01. Since the layer dimensions L and H have a significant influence on the stress intensity factors, they are set large enough to enable a comparison with the infinite solution for an isotropic medium when $\theta=90$ deg and $\mu=0.5$. In order to mimic the infinite medium, we set $\rho=1$ and $\eta=100$.

Figure 13 illustrates the effect of rotation angle θ on Mode I, II and III stress intensity factors at the inner and outer crack tips of the right crack for the three μ ratios when the normalized horizontal separation distance $r_d/2a=0.1$ between the two cracks produces a moderate level of stress interaction. When the cracks are far from the top surface, the crack interaction produces a 50% increase at the inner crack tip in Mode I stress intensity factor at $\theta=0$ deg relative to the single crack case, Fig. 11. This increase becomes 60% for this orientation at the closest distance to the top surface. As also observed in the single crack case, it is only very close to the top surface when the rotation angle θ has a noticeable effect on Mode I, II, and III stress intensity factors at both crack tips that increase in magnitude with the rotation angle, with Mode II and III actually reversing sign. The crack interaction produces a greater increase in the Mode I stress intensity factor with a rotation angle relative to the single crack case. On the other hand, the relative increase in Mode II stress intensity factor with rotation angle does not change significantly relative to the single crack case. The maximum value of Mode III stress intensity factor also occurs in the same θ range, but increases by 100% relative to the single crack case. The stress intensity factors at the outer crack tips are also magnified by the crack interaction relative to the single crack case, and dependence on the rotation angle is also observed very close to the top surface. Interestingly, the magnitudes of the Mode II and III stress intensity factors in this case are further increased relative to the inner crack tip, in contrast with the expected decrease in Mode I stress intensity factor, but the sign remains the same.

Next, we compare the results of our solution when $\theta=90$ deg with the results of an exact elasticity solution for an isotropic infinite medium, Erdogan [7], as a function of the $r_d/2a$ ratio that

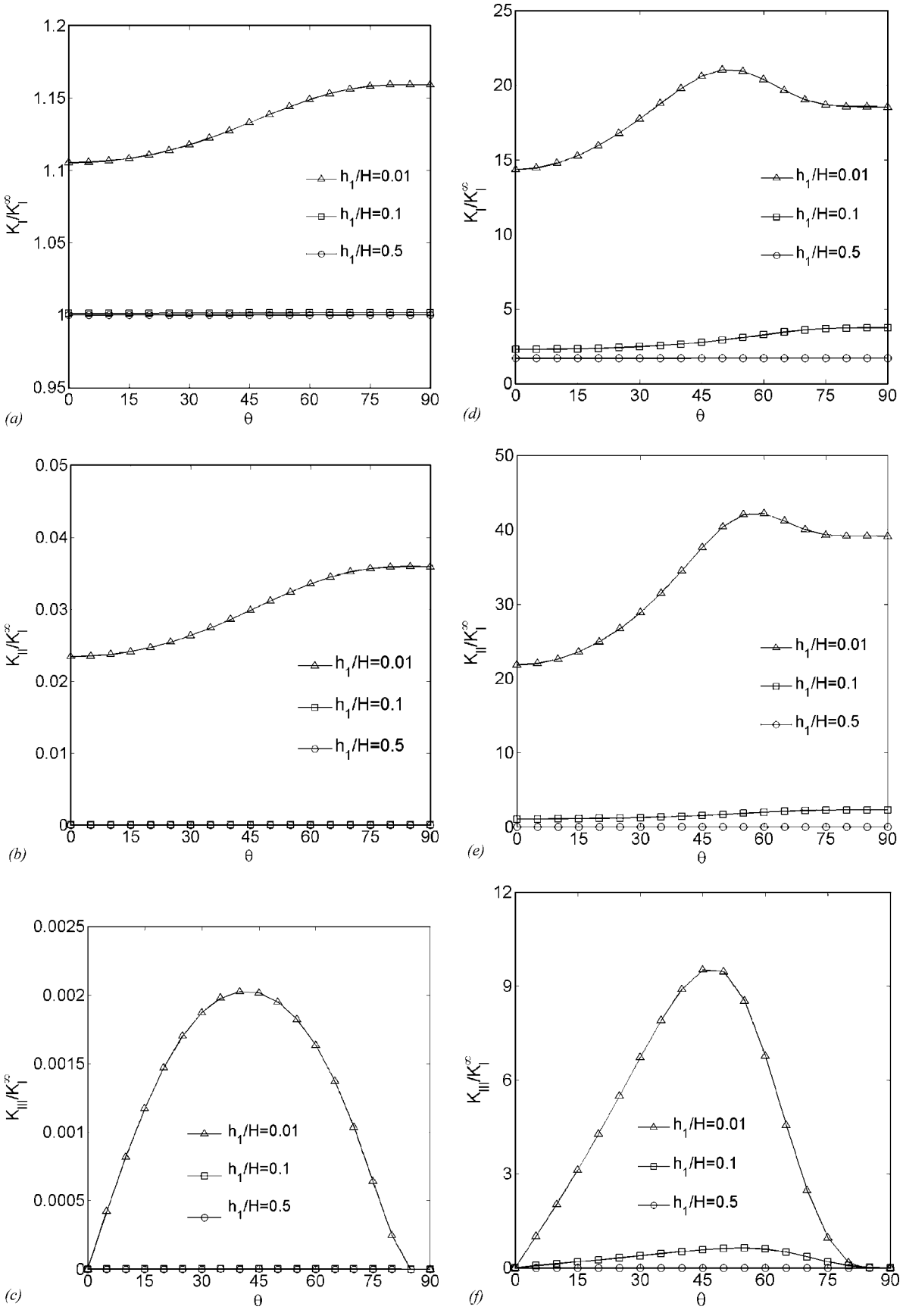


Fig. 11 Normalized Mode I, II and III stress intensity factors K_I/K_I^∞, K_{II}/K_I^∞, K_{III}/K_I^∞ as a function of the fiber rotation angle θ for a square plate with a centrally positioned crack increasingly closer to the upper surface, demonstrating the effect of the crack tip interaction with vertical boundaries controlled by the η ratio for (*a*)-(*c*) $\eta = L/2a = 100$; (*d*)-(*f*) $\eta = L/2a = 1.2$

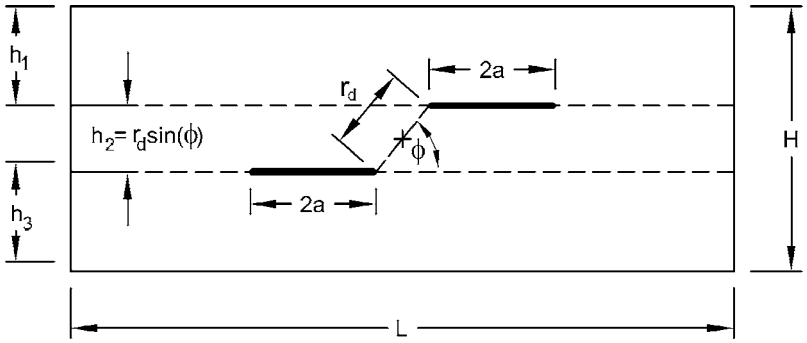


Fig. 12 Geometry of the interacting two-crack problem

varies from 1.0 to 0.001, representing different magnitudes of crack interactions. In this case, the crack is situated in the middle of the plate, $\mu=0.5$, and since it is loaded by internal pressure the only nonzero stress intensity factor is the Mode I factor. Figure 14 compares the variation of the Mode I stress intensity factor with the $r_d/2a$ ratio at both the inner and outer crack tips obtained using the two solutions. As observed, the differences cannot be discerned at this graphical resolution. An examination of the actual numerical values, Table 3, reveals the greatest difference at the smallest employed ratio $r_d/2a=0.001$, which produces an almost seven-fold Mode I stress intensity factor magnification relative to a single crack in an infinite isotropic medium. This difference is approximately 3.8% at the inner crack tip and 0.54% at the outer tip. Fifty-one Chebyshev polynomials were employed to approximate the crack-opening displacement field for this case, as also illustrated in Fig. 3 and included in the table. At $r_d/2a=0.01$, the differences between the two solutions decrease to 0.45% and 0.0% at the inner and outer crack tips, respectively. Also included in the table is the comparison for $r_d/2a=0.0001$. At such a small separation distance, Fig. 3 indicates that the number of Chebyshev polynomials required for an accurate solution is substantially greater, which explains the reported difference of almost 15% at the inner crack tips between Erdogan's exact and present solutions. At the outer crack tip the difference reduces to 1.87%

3.2.2 Offset Cracks: $\phi=45$ deg. Next, we examine stress intensity factors as a function of the rotation angle θ for the two horizontal cracks offset by the angle $\phi=45$ deg, with the inner crack tips separated by the distance $r_d/2a=0.1$, 0.2. The plate aspect ratio is initially set at $\rho=1$, and the results are generated for the relative crack length ratio $\eta=100$ to simulate an infinite medium for a direct comparison with the collinear crack case. Figure 15(*a*)–15(*c*) illustrates the variation of stress intensity factors at the inner crack tips with θ for the two normalized separation distances. In contrast with the two collinear cracks situated in the middle of a large plate, the offset angle ϕ also produces Mode II and III stress fields at the inner crack tips absent in the collinear case due to symmetry. Both Mode I and II stress intensity factors increase monotonically with the rotation angle while Mode III factor attains a maximum magnitude between $\theta=30$ deg and 45 deg. Unlike the former case, the relative variation of Mode I stress intensity factor with θ is not insignificant, while the relative variations in Mode II and III stress intensity factors with θ are substantially greater. The offset angle also produces an increase in the Mode I stress intensity factor at $\theta=0$ deg relative to that for the collinear cracks at the same normalized separation distance. In the case of Mode I and II stress intensity factors, decreasing the normalized separation distance increases the magnitude of these factors without a substantial change in the manner in which they vary with the rotation angle. At the separation distance $r_d/2a=0.2$, the Mode II stress intensity factor is initially negative and then becomes positive beyond $\theta=20$ deg, indicating a change in the relative shearing of the crack faces with increasing θ.

The effect of the layer aspect ratio for the normalized separation distance $r_d/2a=0.1$ and the layer-to-crack length ratio $\eta=5$ when the crack-boundary interaction becomes important is illustrated in Fig. 15(*d*)–15(*f*) for decreasing relative layer thicknesses characterized by $\rho=1$, 0.5, 0.2. In this case, decreasing the ratio ρ produces an increase in the stress intensity factors accompanied by a substantial change in the manner of variation with the rotation angle θ. In particular, decreasing the relative layer thickness increases the difference between the stress intensity factors of the $\theta=0$ deg and 90 deg configurations for the Mode I and II crack opening displacements, with the concomitant increase in the Mode III stress intensity factor between $\theta=30$ deg and 45 deg.

Finally, the present solution for $\theta=90$ deg and $\eta=100$ is compared with the results reported by Isida [8] and Binienda [9] for the offset angles of $\phi=45$ deg, 38.7 deg and the relative inner crack tip separation distances of $r_d/2a=0.707$ and 2.121 based on these authors' solutions of multiple interacting cracks in an infinite isotropic plate. This comparison shown in Table 4 illustrates very good agreement.

3.3 Multilayer With Cracks at Different Elevations. In this subsection, we consider two multilayered configurations with horizontal cracks at equally spaced elevations; Fig. 16. In the first configuration shown in Fig. 16(*a*), the equally spaced cracks are stacked in a single vertical column, whereas in the second configuration shown in Fig. 16(*b*) the cracks are vertically offset in a diagonal or echelon array. In both cases, we calculate Mode I, II, and III stress intensity factors at the tips of cracks positioned at different distances from the top and bottom surfaces.

When the applied loading on the top and bottom surfaces of the configuration with vertically stacked cracks takes the form of uniform vertical displacements, this problem reduces to a single crack in a repeating unit cell subjected to periodic boundary conditions when the configuration is homogeneous. Equally spaced cracks in a single vertical column in an infinite isotropic plate have been investigated by Isida [10] and Horii and Nemat-Nasser [11], thereby providing a means of testing the predictive capability of our model for the fiber rotation angle $\theta=90$ deg. The results generated for a layer with $\theta=90$ deg rotation angle as a function of the ratio η under periodic boundary conditions for the normalized vertical crack separation distance $d/2a=2$ are shown in Table 5 to demonstrate that in the limit as $\eta\to\infty$ our results converge to those of Isida [10] and Horii and Nemat-Nasser [11]. When loading takes the form of uniform normal pressure applied on the top and bottom surfaces, the stress intensity factors will vary with the distance from both surfaces, approaching the value generated under periodic boundary conditions for a configuration containing a sufficiently large number of cracks. This type of problem is of interest in the micromechanics of heterogeneous materials, where the calculation of homogenized properties due to material heterogeneity, including cracks and porosities, must be performed on

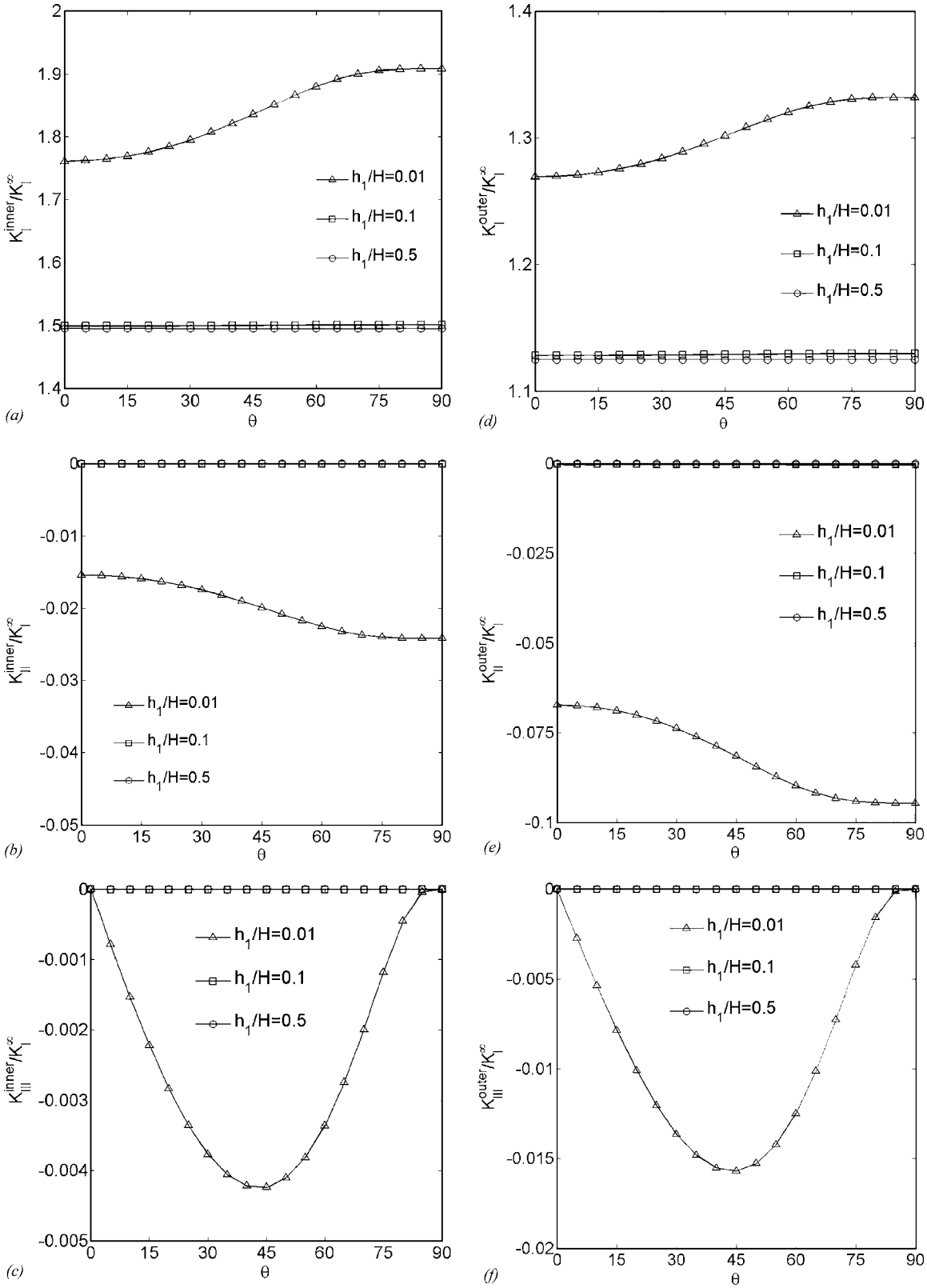


Fig. 13 Normalized Mode I, II and III stress intensity factors K_I/K_I^{∞}, K_{II}/K_I^{∞}, K_{III}/K_I^{∞} at the inner and outer tips of two collinear cracks separated by the normalized distance $r_d/2a=0.1$ at increasingly closer normalized distances $\mu = h_1/H$ to the upper surface of a large square plate ($\eta=L/2a=100$) as a function of the fiber rotation angle θ

material subvolumes containing sufficient microstructural details, to make the result independent of the applied boundary conditions. This is the first problem considered in this subsection.

The echelon crack problem has been investigated by Du and Aydin [12] in the context of an infinite isotropic plate subjected to farfield uniform normal traction. Herein, we investigate this problem under uniform vertical displacement boundary conditions in order to determine the influence of the rotation angle θ on the

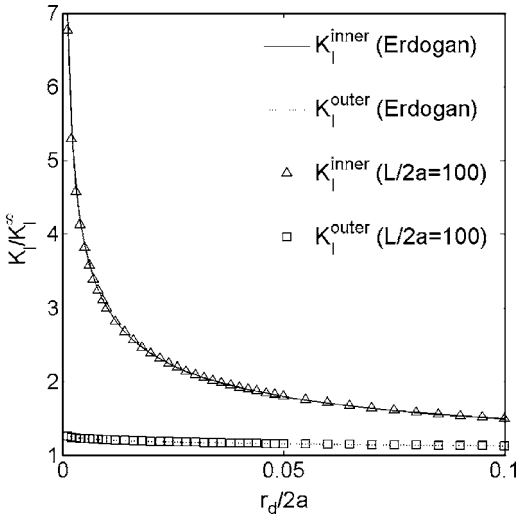


Fig. 14 Normalized Mode I stress intensity factor K_I/K_I^{∞} at the inner and outer tips of two collinear cracks centrally positioned ($\mu=h_1/H=0.5$) in a large square plate with the rotation angle $\theta=90$ deg as a function of the normalized distance $r_d/2a$

local crack-tip stress intensity factors at different crack elevations. Under this type of boundary condition, the finite multilayered configuration responds like a unit cell in a periodic array containing an infinite number of repeating rows of diagonally positioned horizontal cracks. Therefore, any variation in the local stress intensity factors with vertical elevation can be attributed to the actual crack distribution and the resulting local crack-crack and crack-horizontal boundary interaction caused by periodicity rather than the manner of load application investigated in the first problem.

In the above problems, each crack is embedded between adjacent layers of the same properties. Therefore, a plate with n vertically stacked or offset horizontal cracks will consist of $2n$ layers. In the first case, all unidirectional graphite/epoxy layers have the same orientation that produce either orthotropic ($\theta=0$ deg) or isotropic ($\theta=90$ deg) elastic properties in the analysis plane. We note that only the surface layers produce a strong dependence of the stress intensity factors on θ under traction loading, with the interior dependence vanishing for the considered loading and geometry. In the second case, we also investigate the effect of gradually changing the fiber rotation angle for each pair of adjacent layers containing a single crack, effectively producing an orientationally graded layered composite plate. In both cases, the normalized vertical crack spacing is $d/2a=0.5$.

Table 3 A comparison of the Mode I stress intensity factors for two collinear cracks in an infinite isotropic plate as a function of the separation distance $r_d/2a$ generated by the present solution with the results reported by Erdogan [7]. N_j denotes the highest order of Chebyshev polynomials used in approximating the crack opening displacement.

		Erdogan [7]		Present solution (%)	
$r_d/2a$	N_j	K_I^{inner}	K_I^{outer}	K_I^{inner}	K_I^{outer}
0.0001	50	11.2457	1.3347	12.8750 [14.49]	1.3098 [−1.87]
0.0010	50	7.0224	1.2576	6.7582 [−3.76]	1.2508 [−0.54]
0.0050	44	3.8411	1.2242	3.8093 [−0.83]	1.2235 [−0.06]
0.0100	40	3.0048	1.2058	2.9907 [−0.45]	1.2057 [0.00]
0.0500	30	1.7950	1.1510	1.7898 [−0.29]	1.1508 [−0.02]
0.1000	20	1.4914	1.1220	1.4896 [−0.12]	1.1222 [0.02]
0.5000	10	1.1125	1.0517	1.1121 [−0.04]	1.0517 [0.00]
1.0000	10	1.0480	1.0280	1.0480 [0.00]	1.0281 [0.00]

3.3.1 Vertically Stacked Cracks. We consider two multilayers with the orientations $\theta=0$ deg, 90 deg containing vertical columns of 19 cracks characterized by the layer-to-crack length ratio $\eta=10$, and first generate baseline values for the stress intensity factors by subjecting the two configurations to a uniform vertical displacement on the top surface with the bottom surface constrained. As expected, the Mode I stress intensity factors for all cracks are the same and equal to the normalized value of $K_I/K_I^{per}=1.0$, with the Mode II and III factors identically zero. Identical results are obtained for a single horizontal crack as was verified.

Next, the loading on the top surface is changed to uniform normal traction with the bottom surface constrained. In light of symmetry, this is equivalent to a multilayer with 38 vertically stacked cracks loaded by uniform normal tractions at both surfaces. Figure 17 presents a comparison of the normalized stress intensity factors K_I/K_I^{per} and K_{II}/K_I^{per} at the tips of cracks at different distances from the top surface of the multilayers with $\theta=0$ deg and 90 deg orientations under this loading. As observed, under uniform traction loading Mode I stress intensity factors for cracks in the immediate vicinity of the plate's top surface are substantially higher than the value obtained under uniform displacement boundary conditions, and only approach this value for cracks far removed from the top surface. It is only at the tenth crack that the Mode I stress intensity factor becomes approximately the same as that under periodic boundary conditions. Mode II stress intensity factors are also significant near the top surface, but in contrast with the former case quickly decrease to zero with increasing distance from the top surface. The Mode II stress intensity factor for the third crack from the top surface is already acceptably small. The rotation angle $\theta=90$ deg, which produces isotropic elastic moduli in the $x-z$ plane, results in the largest stress intensity factors for each crack relative to other rotation angles, including $\theta=0$ deg. The differences, however, are limited to the first few top surface layers, with the differences vanishing in the interior.

The above results can be used to generate the effective stiffness elements C_{13}, C_{23}, C_{33} of materials weakened by horizontal cracks under uniform loading by ε_{33} when $\varepsilon_{11}=\varepsilon_{22}=0$ (as is the case in this example). For the considered example, the crack array produces reduced orthotropic elastic stiffness elements that can be employed to calculate the corresponding Young's moduli and Poisson's ratios. As the present solution is exact, such results can be employed as baseline results for comparison with other methods for different arrays of horizontal cracks in periodic arrays.

3.3.2 Vertically Offset Cracks in an Echelon Array. Finally, we determine the stress intensity factors for cracks in multilayers with a diagonal array of horizontal cracks subjected to a uniform vertical displacement on the top surface and constrained on the bottom. The presence of vertically offset cracks gives rise to all three crack opening displacement modes at each crack tip. The presented results are normalized by Mode I stress intensity factor at the left crack tip of the tenth or middle crack when this crack is embedded in a layer with $\theta=90$ deg.

Figure 18(a)–18(c) presents normalized stress intensity factors for each crack in the echelon array, starting with the first crack and progressing downward, for the seven spatially uniform rotation angles $\theta=0$ deg, 15 deg, 30 deg, 45 deg, 60 deg, 75 deg, 90 deg. In contrast with the results for vertically stacked cracks under uniform normal traction loading, Mode I, II, and III stress intensity factors increase with increasing distance from the top surface, with the maximum occurring at the tenth crack, and subsequently decrease. The stress intensity variation with vertical distance is substantial for all three modes. In the case of Mode I and II stress intensity factors, the results are bounded by the stress intensity factors for the 0 deg and 90 deg rotation angles, with the isotropic layer orientation producing largest stress intensity factor at each

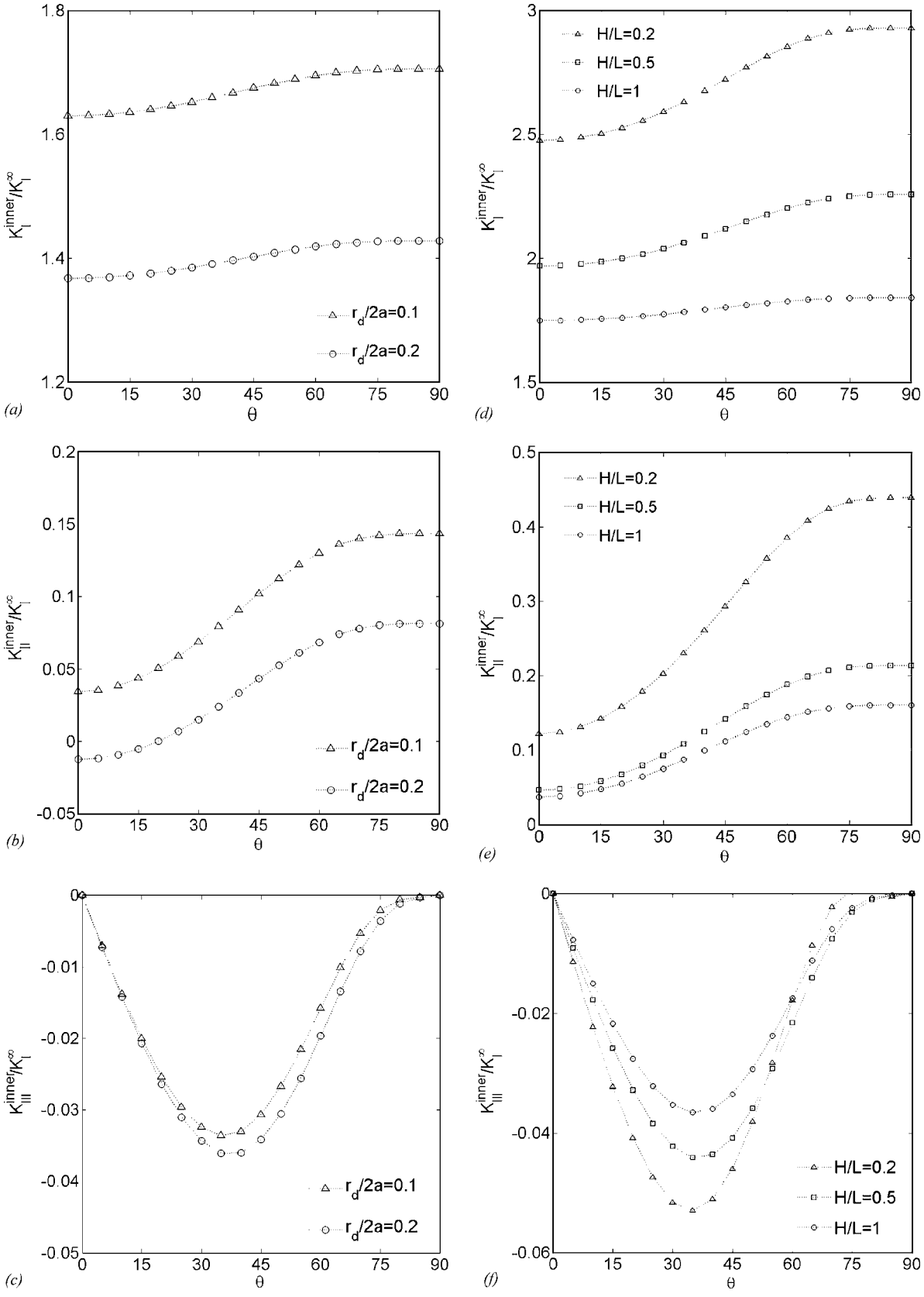


Fig. 15 Normalized stress intensity factors K_I/K_I^{∞}, K_{II}/K_I^{∞} and K_{III}/K_I^{∞} at the inner tips of two horizontal cracks that are offset by $\phi=45$ deg in the middle of a plate as a function of the fiber rotation angle θ, illustrating the effects of: (*a*)-(*c*) normalized separation distance $r_d/2a=0.1, 0.2$ for a large square plate with $\eta=L/2a=100$; (*d*)-(*e*) plate aspect ratio $\rho=H/L=0.2$, 0.5, 1.0 for the separation distance $r_d/2a=0.1$ and a smaller plate with ($\eta=L/2a=5$)

Table 4 A comparison of the Mode I stress intensity factors at the inner tips of two cracks offset by 45 deg and 38.7 deg in an infinite isotropic plate as a function of the separation distance $r_d/2a$ generated by the present solution with the results reported by Isida [8] and Binienda [9].

Offset angle ϕ	$r_d/2a$	Isida [8]	Binienda [9]	Present solution
45 deg	0.707	1.12	1.1254	1.1239
45 deg	1.414	1.04	1.0489	1.0488
45 deg	2.121	1.02	1.0254	1.0256
38.7 deg	0.640	1.13	1.1317	1.1302
38.7 deg	1.281	1.05	1.0551	1.0547
38.7 deg	1.921	1.03	1.0302	1.0303

Table 5 Convergence of the Mode I stress intensity factors for a single column of vertically stacked cracks in a finite-length isotropic plate subjected to periodic boundary conditions to the results of Isida [10] and Horii and Nemat-Nasser [11] with an increasing layer to crack length ratio. The vertical separation distance normalized by the crack length is $d/2a=2$.

$L/2a$	Isida [10]	Horii & Nemat-Nasser [11]	Present solution
10	–	–	0.8430
100	–	–	0.7979
1000	–	–	0.7962
∞	0.78	0.7896	–

rotation angle for Mode I and the orthotropic layer orientation for Mode II. The largest stress intensity factor for Mode III also occurs at the tenth crack for the 45 deg rotation angle. The asymmetry in the results, which is particularly evident in the case of Mode II stress intensity factor, is due to the diagonal array architecture and the top and bottom surface crack interaction under the applied boundary conditions, with reversed results at the right crack tips with increasing distance from the top surface (not shown).

Figure 18(*d*)–18(*f*) presents the corresponding results when each pair of layers containing a single crack is orientationally graded from $\theta=0$ deg at the top surface to $\theta=90$ deg at the bottom surface using 5 deg increments. Included in the figure are the results for the 0 deg and 90 deg rotation angles. These rotation angles provide upper and lower bounds between which the Mode I and II stress intensity factors of the orientationally graded configurations vary with increasing distance from the top surface. In the case of Mode III, coupling between in-plane and out-of-plane stress fields vanishes for orthotropic and isotropic configurations. As suggested by the spatially uniform rotation angle cases presented in Fig. 18(*a*)–18(*c*), the maximum value of Mode III stress intensity factor occurs at the 10th crack for the $\theta=45$ deg rotation angle.

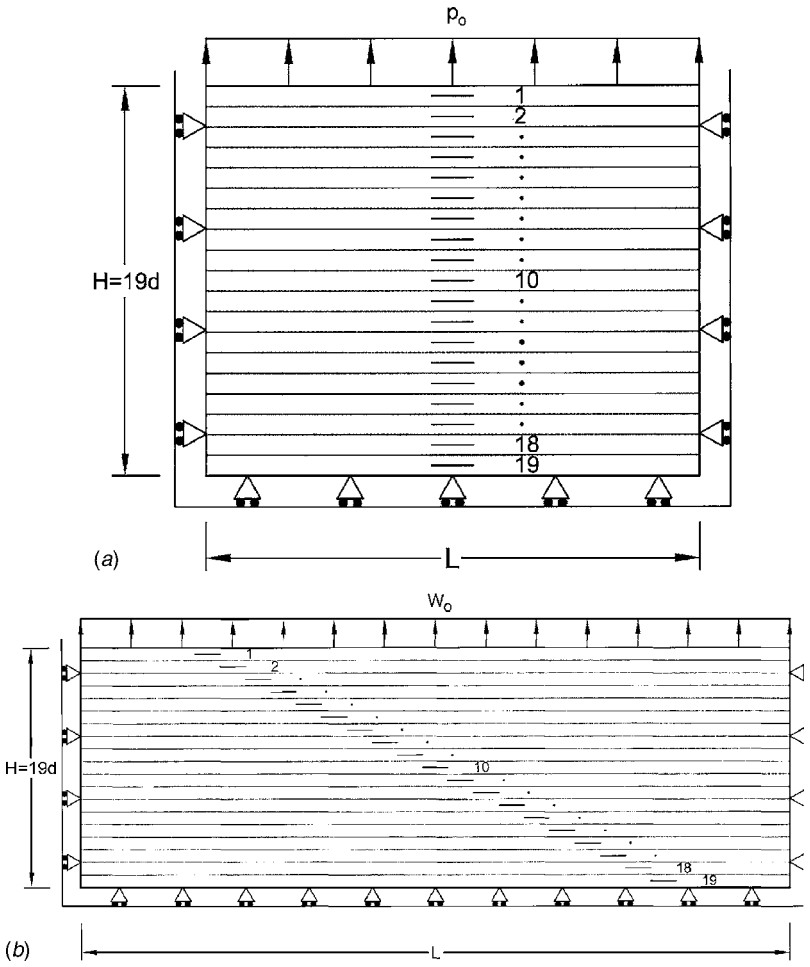


Fig. 16 Geometry of the layered configurations with vertically situated multiple cracks: (*a*) vertically stacked cracks in single column; (*b*) diagonally stacked cracks in an echelon array

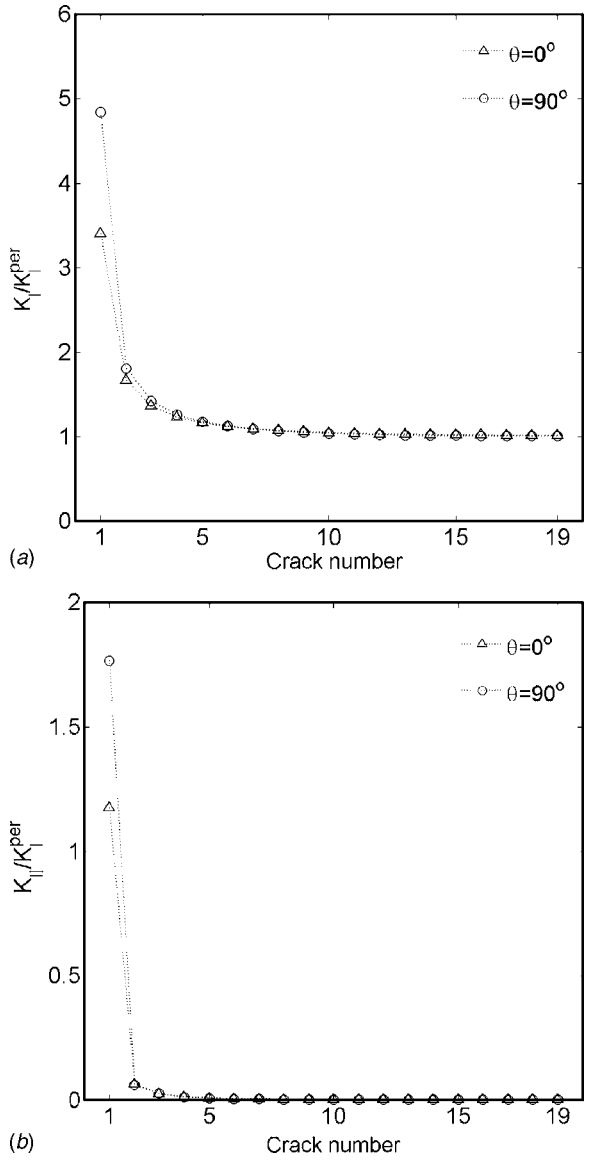


Fig. 17 Normalized stress intensity factors K_I/K_I^{per} and K_{II}/K_I^{per} at the tips of vertically stacked cracks at different distances from the top surface of rectangular layers with the fiber rotation angles θ=0 deg, 90 deg subjected to a uniform external pressure. The normalized vertical crack spacing is $d/2a$=0.5 and the crack-to-layer length ratio is $\eta=L/2a$=10.

4 Limitations and Future Prospects

The local/global stiffness matrix approach provides a systematic framework for constructing the governing system of equations for the crack opening displacement of arbitrarily situated horizontal cracks in multilayers of finite dimensions in the analysis plane. This facilitates the solution's implementation and potentially improves its efficiency due to reduction in the multilayer's global stiffness matrix that arises through efficient application of interfacial continuity conditions. However, in the presence of cracks the number of layers has little effect on the execution time. For instance, the execution time for a single crack embedded in an artificially layered structure comprised of 100 layers differs little from the execution time for the same crack embedded in a two-layer configuration. The execution time increases substantially with an increasing number of cracks. Multiple cracks with very strong interactions require substantial computational effort. An 11-fold loop is used to evaluate the $\mathbf{D}_{jk(\alpha\beta)}^{(pq)}$ coefficients in the final system of algebraic equations for the crack opening displacement functions, which consumes most of the execution time. This calculation is expensive, even for a row of weakly interacting collinear cracks, which requires only a 7-fold loop with a smaller number of Chebyshev polynomials for each crack. Doubling the number of cracked interfaces or cracks along the same interface quadruples, at the minimum, the execution time. Current computational capabilities make it prohibitive and unrealistic to run a case involving several hundred strongly interacting cracks, although a problem involving 100 moderately interacting cracks is practical.

There are several ways to improve the solution's computational efficiency. First, the double integrals associated with the kernels $\mathbf{h}_{\alpha\beta}$ can be transformed into single integrals using hypergeometric functions, as was done by Chatterjee et al. [13]. Second, the number of Chebyshev polynomials can be chosen selectively for each crack according to the intensity of crack interaction, instead of the same number used for all cracks, as was done in the present computational approach. However, in extreme situations, where the normalized separation distance between cracks is very small, $r_d/2a < 0.001$, the bounded Fredholm kernels become very large at the corners of the domain boundaries. This, in turn, makes an accurate integration of the Fredholm kernels very difficult and also requires a very large number of Chebyshev polynomials to accurately approximate the crack opening displacement. In such cases, an asymptotic approach is required. Another approach that shows promise involves partial homogenization of the layered media outside the region of interest, where accurate estimates of stress intensity factors are required, as described by Chen et al. [14] in the context of a multilayer indented by a flat rigid punch.

5 Summary and Conclusions

The numerical results for the stress intensity factors in finite anisotropic multilayers presented in Part II of this two-part paper have demonstrated the developed solution's accuracy upon a comparison with known results generated in the limit, as the multilayer's dimensions approached infinity and as the anisotropy in the analysis plane was reduced to isotropy through an appropriate coordinate transformation. More importantly, new fundamental results have been generated that demonstrate the previously undocumented combined effects of finite dimensions and material anisotropy on Mode I, II, and III stress intensity factors in multilayers composed of differently oriented, unidirectional composite plies that produce monoclinic elastic properties in the analysis plane. These results are technologically important in light of the renewed interest in large-scale laminated composite structures for aircraft applications, in particular, and the ongoing activities in the implementation of multilayer architectures into emerging technologies. Defect criticality is an important consideration in such designs that can now be more efficiently analyzed and designed using the developed elasticity-based solution.

In particular, under Mode I loading of an homogeneous monoclinic layer with a centrally situated crack, only the corresponding crack opening displacement is present and the rotation of the layer's principal material coordinate system does not affect the resulting stress intensity factor. Decreasing the layer to crack length ratio increases the Mode I stress intensity factor uniformly for all rotation angles. All three crack opening displacement modes are present when the crack is situated close to the horizontal boundaries and the dependence on the layer's orientation increases significantly with a decreasing layer to crack length ratio. In the close proximity of the upper surface, Mode II crack opening displacement dominates when strong interaction with vertical boundary occurs, and substantial dependence on the ply orientation is observed for all three modes. In such situations, Mode III is a significant percentage of Modes I and II and thus cannot be neglected.

For collinear cracks subjected to normal pressure and situated in the interior of homogeneous layers with large dimensions, only

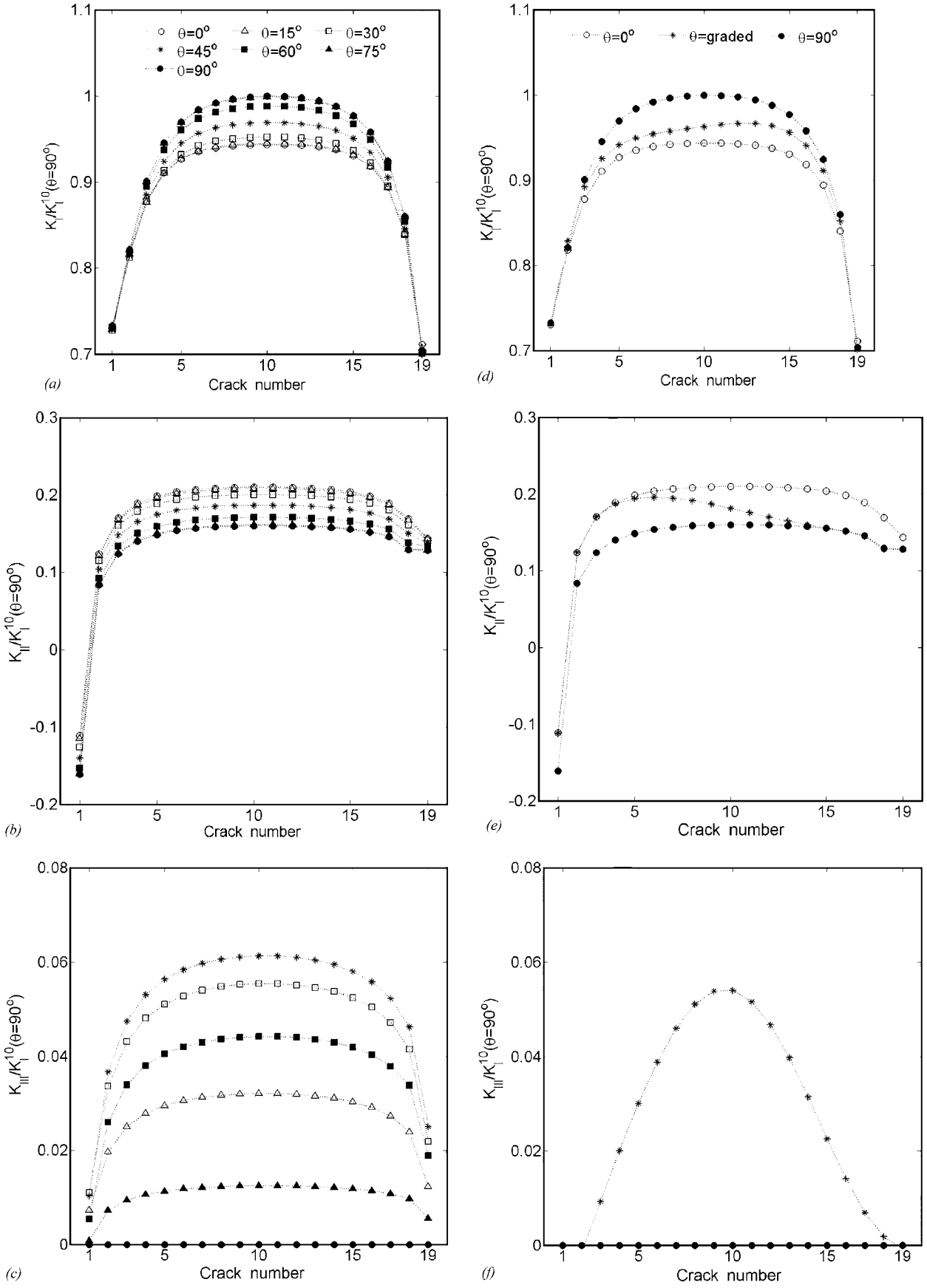


Fig. 18 Normalized stress intensity factors K_I/K_I^{10}, K_{II}/K_I^{10} and K_{III}/K_I^{10} at the tips of diagonally stacked cracks at different distances from the top surface of rectangular layers subjected to a uniform vertical displacement as a function of the fiber rotation angle θ: (*a*)-(*c*) spatially uniform θ; (*d*)-(*f*) functionally graded θ. The normalized vertical crack spacing is $d/2a=0.5$, and the crack-to-layer length ratio is $\eta=L/2a=30$.

Mode I is present, which also does not depend on the layer's degree of anisotropy induced by rotation. Bringing two collinear cracks closer to the top surface produces coupling of the three modes observed in the single crack case, with the concomitant increases in stress intensity factors and dependence on layer's orientation caused by the inner crack tip interaction. Coupling of the three modes also occurs for vertically offset cracks subjected to normal pressure in the interior, which becomes significant with decreasing height to length aspect ratio of the layer.

The present solution's capability of multiple crack analysis makes possible the investigation of fundamental problems in the micromechanics of heterogeneous materials with damage, such as the effect of boundary conditions on the stress intensity factors of cracks situated at different distances from a multilayer's surface. The boundary layer effect is important in multilayers with vertically stacked cracks subjected to homogeneous normal tractions and may propagate far into the multilayer's interior. However, the impact of layer orientation appears to be limited to the surface' immediate vicinity. Vertically offsetting multiple cracks to produce an echelon array in a multilayer subjected to homogeneous displacement boundary conditions results in substantial dependence of the stress intensity factors on layer orientation in the interior, in contrast with the vertically stacked cracks. This dependence can be controlled by orientational grading of the layers containing individual cracks.

The present solution quantifies the heretofore undocumented impact of finite dimensions and layer anisotropy due to a unidirectional fiber-reinforced layer's orientation in composite multilayers under different loading and boundary conditions. These effects may play a critical role in defect criticality of advanced multilayered structures when cracks are situated close to vertical and horizontal boundaries, as demonstrated herein.

Acknowledgment

The first author thanks the Civil Engineering Department at the University of Virginia and the Engineering Material Concepts, LLC for providing partial support in the course of this investigation. The authors also thank Dr. Yogesh Bansal for generating Fig. 6 and the homogenized elastic results in Table 2.

References

[1] Pindera, M-J., 1991, "Local/Global Stiffness Matrix Formulation for Composite Materials and Structures," Composites Eng., **1**, pp. 69–83.
[2] Ting, T. C. T., 1996, *Anisotropic Elasticity (Theory and Applications)*, Oxford University Press, Oxford.
[3] Erdogan, F., Gupta, G. D., and Cook, T. S., 1973. "Numerical Solution of Singular Integral Equations," *Methods of Analysis and Solutions of Crack Problems*, edited by G. C. Sih, Netherlands, Noordhoff, Leyden, pp. 368–425.
[4] Bansal, Y., and Pindera, M-J., 2005, "A Second Look at the Higher-Order Theory for Periodic Multiphase Materials," J. Appl. Mech., **72**, pp. 177–195.
[5] Bansal, Y., and Pindera, M-J., 2006, "Finite-Volume Direct Averaging Micromechanics of Heterogeneous Materials with Elastic-Plastic Phases," Int. J. Plast., **22**, pp. 775–825.
[6] Nye, J. F., 1985, *Physical Properties of Crystals*, Clarendom Press, Oxford.
[7] Erdogan, F., 1962, "On the Stress Distribution in Plates with Collinear cuts under arbitrary loads," *Proc. of the Fourth U.S. National Congress. of Appl. Mech.*, Vol. 1, pp. 547–553.
[8] Isida, M., 1976, "Analysis of Stress Intensity Factors for Plates Containing Randomly Distributed Cracks," Trans. Jpn. Soc. Mech. Eng., **35**, pp. 1815–1822.
[9] Binienda, W. K., 1994, "Stress Intensity Factors for Fully Interacting Cracks in a Multicrack Solid," J. Offshore Mech. Arct. Eng., **116**, pp. 56–63.
[10] Isida, M., 1973, "Method of Laurent Series Expansion for Internal Crack Problems," *Methods of Analysis and Solutions of Crack Problems*, edited by G. C. Sih, Leyden, Netherlands, Noordhoff, pp. 56–130.
[11] Horii, H., and Nemat-Nasser, S., 1985, "Elastic Fields of Interacting Inhomogeneities," Int. J. Solids Struct., **21**, pp. 731–745.
[12] Du, Y. J., and Aydin, A., 1991, "Interaction of Multiple Cracks and Formulation of Echelon Crack Arrays," Int. J. Numer. Analyt. Meth. Geomech., **15**, pp. 205–218.
[13] Chatterjee, S. N., Pindera, M-J., Pipes, R. B., and Dick, B., 1982, "Composite Defect Significance," MSC TFR 1312/1108 (NADC Report No. 81034-60), Spring House, PA.
[14] Chen, L., Urquhart, E. E., and Pindera, M-J., 2005, "Microstructural Effects in Multilayers with Large Moduli Contrast Loaded by Flat Punch," AIAA J., **43**, pp. 962–973.

Journal of Applied Mechanics

Technical Briefs

The True Linearization Coefficients for Nonlinear Systems Under Parametric White Noise Excitations

Giovanni Falsone
Dipartmento di Ingegneria Civile (DIC),
Università di Messina,
Salita Sperone, 31, Vill. S. Agata,
198166 Messina, Italy

In this paper some properties of the stochastic linearization method applied to nonlinear systems excited by parametric Gaussian white noises are discussed. In particular, it is shown that the linearized quantities, obtained by the author in another paper by linearizing the coefficients of the Ito differential rule related to the original system, show the same properties found by Kozin with reference to nonlinear system excited by external white noises. The first property is that these coefficients are the true linearized quantities, in the sense that their exact values are able to give the first two statistical moments of the true response. The second property is that, in the stationary case and in the field of the parameter estimation theory, they represent the maximum likelihood estimates of the linear model quantities fitting the original nonlinear response. [DOI: 10.1115/1.1940665]

1 Introduction

Stochastic linearization technique is the most versatile method for analysis of general nonlinear systems and structures under random excitations. In almost 50 years since its first virtually simultaneous presentations by Booton [1] and Kazakov [2], it has been widely applied in the study of various nonlinear systems for which it is not possible to obtain an exact solution. For example the monograph by Roberts and Spanos [3] and many review papers written in these last years may confirm this success.

The basic idea of the stochastic linearization technique is to replace the original nonlinear system by a linear one in such a way that the difference between the two systems is minimized in some statistical sense. In this way, the quantities appearing in the linearized system involve unknown response statistics which, in the case of Gaussian excitations, are evaluated approximating the response as a Gaussian process.

In the case of nonlinear systems excited by purely external Gaussian excitations, it was shown that, if the error between the original and the linearized systems is evaluated on the motion equations, the stochastic linearization method gives the same results as the Gaussian closure method [4]. Moreover, Kozin evidenced that, in the case of purely external white noise excitations, the linearized quantities obtained by the stochastic linearization applied in this sense represent the *true* linearization quantities [5]. But, as a matter of fact, this concept was first introduced by Caughey [6], even if he did not use the term *true*, and then it was discussed by Crandall [7], too. This concept evidences that, if the exact value of the statistical quantities appearing in the linearized quantities is known, the linearized quantities are able to give the exact first two statistical moments of the response. At last, Kozin himself showed another important property of these linearized quantities in the field of the parameter estimation theory [8]: *if the response of a nonlinear system is observed and is fitted by a linear model, then, in the stationary case, the maximum likelihood estimates of the linear quantities converge to the quantities obtained by the stochastic linearization*. This is an important result from both a theoretical and a practical point of view.

When parametric type excitations act on the nonlinear systems, the applications of the stochastic linearization technique are based on different approaches leading to different results. For example, Chang and Young linearized the motion equations [9], while Wu linearized the Ito differential equation [10]. Falsone [11] showed that only the linearization on the coefficients of the Ito differential rule can originate a stochastic linearization technique able to give the same results of the Gaussian closure method. Moreover he found the corresponding linearized quantities.

The aim of the present paper is to verify if the linearized quantities obtained in [11] for parametric white noise excitations are characterized by the same two important properties revealed by Kozin in the case of purely external white noise excitation. This means that the answers will be given to the following two questions: (1) are the linearized quantities found in [11] *true*, in the Kozin sense?, (2) are they the maximum likelihood estimates of the linear model quantities fitting the original nonlinear response?

2 The True Linearization Coefficients

Let us consider a nonlinear system excited by a parametric Gaussian white noise excitation that, in terms of state variable coordinates, is governed by the following differential equation:

$$\dot{\mathbf{X}}(t) = \mathbf{f}[\mathbf{X}(t)] + \mathbf{g}[\mathbf{X}(t)]W(t) \tag{1}$$

where $W(t)$ is such that $E[W(t)]=0$; $E[W(t+\tau)W(t)]=q\delta(\tau)$, $E[\cdot]$ being the mean operator, $\delta(\cdot)$ the Dirac delta function and q the white noise intensity.

Let us suppose using the stochastic equivalent linearization for finding the approximate solution of Eq. (1). In a previous work [11] it was shown that in this framework a fundamental role is played by the coefficients of the Ito differential rule; this rule can be written as

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 20, 2003; final manuscript received December 2, 2004. Assoc. Editor: M. P. Mignolet.

$$d\phi[\mathbf{X}(t),t] = \frac{\partial \phi[\mathbf{X}(t),t]}{\partial t}dt + (\nabla_{\mathbf{X}}^{T}\phi[\mathbf{X}(t),t])d\mathbf{X}(t) + \frac{1}{2}(\nabla_{\mathbf{x}}^{T[2]}\phi[\mathbf{X}(t),t])\mathbf{g}^{[2]}[\mathbf{X}(t)]qdt \tag{2}$$

where $\phi[\mathbf{X}(t),t]$ is a vector function differentiable with respect to t and twice differentiable with respect to $\mathbf{X}$, $\nabla_{\mathbf{x}}$ is the vector operator whose ith entry is $\partial/\partial X_i$, $d\mathbf{X}(t)$ is the first member of the Ito differential equation taking into account the Wong–Zakai correction term and the exponent into the square brackets indicates the power exponent made following the rule of the Kronecker algebra [12], that is:

$$\mathbf{C}^{[2]} = \mathbf{C} \otimes \mathbf{C} \tag{3}$$

$\mathbf{C}$ being a generic matrix and the symbol $\otimes$ indicating the Kronecker product.

In the above-cited work it was shown that the stochastic equivalent linearization gives the same results as the Gaussian closure method, for parametric excitation, too, only if the linearization is made on the coefficients of the Ito differential rule. In this way, it was found that the nonlinear system must be replaced by a linear one such that the corresponding Ito equation and Ito differential rule are, respectively,

$$d\mathbf{Y}(t) = \mathbf{A}(t)\mathbf{Y}(t)dt + \mathbf{a}(t)dt + \mathbf{b}(t)dL(t) \tag{4}$$

$$d\phi[\mathbf{Y}(t),t] = \frac{\partial \phi[\mathbf{Y}(t),t]}{\partial t}dt + (\nabla_{\mathbf{Y}}^{T}\phi[\mathbf{Y}(t),t])d\mathbf{Y}(t) + \frac{1}{2}(\nabla_{\mathbf{Y}}^{T[2]}\phi[\mathbf{Y}(t),t])\bar{\mathbf{b}}^{[2]}(t)qdt \tag{5}$$

where $L(t)$ is the Wiener process related to $W(t)$ and the linearized terms are:

$$\mathbf{A}(t) = \{E[\mathbf{m}[\mathbf{Y}(t)]\mathbf{Y}^{T}(t)] - E[\mathbf{m}[\mathbf{Y}(t)]]E[\mathbf{Y}(t)]^{T}\}\Sigma_{\mathbf{Y}}^{-1}(t)$$

$$\mathbf{a}(t) = E[\mathbf{m}[\mathbf{Y}(t)]] - \mathbf{A}(t)E[\mathbf{Y}(t)]$$

$$\mathbf{b}(t) = E[\mathbf{g}[\mathbf{Y}(t)]]$$

$$\bar{\mathbf{b}}^{[2]}(t) = E[\mathbf{g}^{[2]}[\mathbf{Y}(t)]] \tag{6}$$

$\Sigma_{\mathbf{Y}}(t) = E[\mathbf{Y}(t)\mathbf{Y}^{T}(t)] - E[\mathbf{Y}(t)]E[\mathbf{Y}(t)]^{T}$ being the covariance matrix of $\mathbf{Y}(t)$ and $\mathbf{m}[\mathbf{Y}(t)]$ the corresponding drift term. As made in [11], even here it is important to note that $\bar{\mathbf{b}}^{[2]}(t) \neq \mathbf{b}^{[2]}(t)$.

Now we want to show that the above-considered linearized terms are the true linear coefficients of the nonlinear system excited by a parametric white noise, where the significance of the adjective *true* is the same as that given by Kozin in the case of external excitation [5], that is: we replace the vector $\mathbf{Y}(t)$ by the vector $\mathbf{X}(t)$ into Eq. (6), thinking of knowing the moments appearing in them; then, if we evaluate the first two moments of the correspondent linearized system, they are identical to the first two moments of the true nonlinear system.

In fact, starting from the Ito differential rule related to the original nonlinear system (1), it is not difficult to verify that the differential equations governing the first two moments of the true response $\mathbf{X}(t)$ are:

$$\dot{E}[\mathbf{X}(t)] = E[\mathbf{m}[\mathbf{X}(t)]]$$

$$\dot{E}[\mathbf{X}^{[2]}(t)] = E[\mathbf{m}[\mathbf{X}(t)] \otimes \mathbf{X}(t)] + E[\mathbf{X}(t) \otimes \mathbf{m}[\mathbf{X}(t)]] + E[\mathbf{g}^{[2]}[\mathbf{X}(t)]]q \tag{7}$$

Starting from Eqs. (4) and (5), it is possible to find the differential equations governing the first two moments of the linearized system, which are:

$$\dot{E}[\mathbf{Y}(t)] = \mathbf{A}(t)E[\mathbf{Y}(t)] + \mathbf{a}(t)$$

$$\dot{E}[\mathbf{Y}^{[2]}(t)] = [\mathbf{A}(t) \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}(t)]E[\mathbf{Y}^{[2]}(t)] + \mathbf{a}(t) \otimes E[\mathbf{Y}(t)] + E[\mathbf{Y}(t)] \otimes \mathbf{a}(t)\bar{\mathbf{b}}^{[2]}(t)q \tag{8}$$

$\mathbf{I}$ being the identity matrix having the same dimension of $\mathbf{X}(t)$. If the expression of $\mathbf{a}(t)$ given in the second equation of Eq. (6) is replaced in the first equation of Eq. (8), we easily obtain:

$$\dot{E}[\mathbf{Y}(t)] = E[\mathbf{m}[\mathbf{Y}(t)]] \tag{9}$$

which has the same form of the first equation of Eq. (7). This means that if we know the true value of $E[\mathbf{m}[\mathbf{Y}(t)]]$, that is $E[\mathbf{m}[\mathbf{X}(t)]]$, the first of the linearized equations (8) is able to give the exact value of the true response first-order moments.

If the expression of $\mathbf{a}(t)$ and $\bar{\mathbf{b}}^{[2]}(t)$ given in the second and fourth equation of Eq. (6) are replaced into the second equation of Eq. (8), then this last one becomes:

$$\dot{E}[\mathbf{Y}^{[2]}(t)] = [\mathbf{A}(t) \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}(t)][E[\mathbf{Y}^{[2]}(t)] - E[\mathbf{Y}(t)]^{[2]}] + E[\mathbf{m}[\mathbf{Y}(t)]] \otimes E[\mathbf{Y}(t)] + E[\mathbf{Y}(t)] \otimes E[\mathbf{m}[\mathbf{Y}(t)]] + E[\mathbf{g}^{[2]}[\mathbf{Y}(t)]]q \tag{10}$$

It is worth noting that:

$$E[\mathbf{Y}^{[2]}(t)] - E[\mathbf{Y}(t)]^{[2]} = \mathrm{Vec}(\Sigma_{\mathbf{Y}}(t)) \tag{11}$$

Vec(•) meaning vectorialized form of (•), that is an operator giving a vector built by all the columns of the matrix (•), each placed below the preceding [12]. Due to a fundamental property of the Vec(•) operator, that is:

$$\mathrm{Vec}(\mathbf{ABC}) = (\mathbf{C}^{T} \otimes \mathbf{A})\mathrm{Vec}(\mathbf{B}) \tag{12}$$

the first term of the second member of Eq. (10) can be rewritten as follows (with the argument t omitted for simplicity):

$$\begin{aligned}[\mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}]\mathrm{Vec}(\Sigma_{\mathbf{Y}}) &= \mathrm{Vec}(\Sigma_{\mathbf{Y}}\mathbf{A}^{T} + \mathbf{A}\Sigma_{\mathbf{Y}}) \\ &= \mathrm{Vec}(\Sigma_{\mathbf{Y}}\Sigma_{\mathbf{Y}}^{-T}\{E[\mathbf{Y}\mathbf{m}^{T}[\mathbf{Y}]] - E[\mathbf{Y}]E[\mathbf{m}^{T}[\mathbf{Y}]]\} + \{E[\mathbf{m}[\mathbf{Y}]\mathbf{Y}^{T}] \\ &\quad - E[\mathbf{m}[\mathbf{Y}]]E[\mathbf{Y}^{T}]\}\Sigma_{\mathbf{Y}}^{-1}\Sigma_{\mathbf{Y}}) \\ &= E[\mathbf{Y} \otimes \mathbf{m}[\mathbf{Y}]] - E[\mathbf{Y}] \otimes E[\mathbf{m}[\mathbf{Y}]] \\ &\quad + E[\mathbf{m}[\mathbf{Y}] \otimes \mathbf{Y}] - E[\mathbf{m}[\mathbf{Y}]] \otimes E[\mathbf{Y}]\end{aligned} \tag{13}$$

where the first equation of Eq. (6) and the symmetry of the matrix $\Sigma_{\mathbf{Y}}$ and of its inverse have been taken into account. If we now replace Eq. (13) into Eq. (10), the following differential equation is obtained:

$$\dot{E}[\mathbf{Y}^{[2]}(t)] = E[\mathbf{m}[\mathbf{Y}(t)] \otimes \mathbf{Y}(t)] + E[\mathbf{Y}(t) \otimes \mathbf{m}[\mathbf{Y}(t)]] + E[\mathbf{g}^{[2]}[\mathbf{Y}(t)]]q \tag{14}$$

that has the same form of the second equation of Eq. (8). This means that if we know the true values of $E[\mathbf{m}[\mathbf{Y}(t)] \otimes \mathbf{Y}(t)]$, $E[\mathbf{Y}(t) \otimes \mathbf{m}[\mathbf{Y}(t)]]$ and $E[\mathbf{g}^{[2]}[\mathbf{Y}(t)]]$, that are $E[\mathbf{m}[\mathbf{X}(t)] \otimes \mathbf{X}(t)]$, $E[\mathbf{X}(t) \otimes \mathbf{m}[\mathbf{X}(t)]]$ and $E[\mathbf{g}^{[2]}[\mathbf{X}(t)]]$, the second of the linearized equations (8) is able to give the exact value of the true response second-order moments.

The fundamental result obtained by the analysis presented in this section is that the linearized terms given in the first, second, and fourth equations of Eq. (6) are the true linearized terms for nonlinear systems excited by parametric Gaussian white noise excitations, in the same sense that Kozin gave to this adjective. It is worth noting that this result can be obtained only if the linearization of the system is made on the coefficients of the Ito differential

rule. Any other kind of linearization, for example on the motion equation or on the Ito differential equation, is not able to give this result.

3 Parameter Estimation and Stochastic Equivalent Linearization

In this section we will show that an important result, found by Kozin [10] for external Gaussian white noise excitations, can be extended to the case of parametric Gaussian white noise excitations, too. That is, if we are fitting the true nonlinear system by a linear model, in the stationary case, the maximum likelihood estimates of the coefficients for the assumed linear model converge to the stationary values of the true linearized terms obtained in the previous section.

The starting point for showing this important result is considering the Ito equation governing the linear model, that is:

$$d\mathbf{Y}(t) = \bar{\mathbf{A}}\mathbf{Y}(t)dt + \bar{\mathbf{a}}dt + \bar{\mathbf{b}}dL(t) \tag{15}$$

The maximum likelihood estimate of $\bar{\mathbf{a}}$ from observation of $\{\mathbf{Y}(t);t\in[0,T]\}$ can be simply obtained by integrating this last equation into $[0,T]$, that is:

$$\int_0^T d\mathbf{Y}(\tau) = \bar{\mathbf{A}}\int_0^T \mathbf{Y}(\tau)d\tau + \bar{\mathbf{a}}T + \bar{\mathbf{b}}\int_0^T dL(\tau) \tag{16}$$

from which, by taking into account that we are observing the true response $\{\mathbf{X}(t);t\in[0,T]\}$, it is possible to find:

$$\bar{\mathbf{a}} = \frac{1}{T}\int_0^T \mathbf{m}[\mathbf{X}(\tau)]d\tau + \frac{1}{T}\int_0^T \mathbf{g}[\mathbf{X}(\tau)]dL(\tau) - \bar{\mathbf{A}}\frac{1}{T}\int_0^T \mathbf{X}(\tau)d\tau - \bar{\mathbf{b}}\frac{1}{T}\int_0^T dL(\tau) \tag{17}$$

If we assume that a stationary ergodic measure exists for Eq. (1), then each term in Eq. (17) possesses an almost sure limit. As a consequence, the maximum likelihood estimate of the vector $\bar{\mathbf{a}}$ is given as:

$$\bar{\mathbf{a}}_{\mathrm{ML}} = \lim_{T\to\infty}\bar{\mathbf{a}} = E[\mathbf{m}[\mathbf{X}]] - \bar{\mathbf{A}}_{\mathrm{ML}}E[\mathbf{X}] \tag{18}$$

$E[\mathbf{m}[\mathbf{X}]]$ and $E[\mathbf{X}]$ being the stationary values of $E[\mathbf{m}[\mathbf{X}(t)]]$ and $E[\mathbf{X}(t)]$, respectively, while $\bar{\mathbf{A}}_{\mathrm{ML}}$ is the maximum likelihood estimate of the matrix $\bar{\mathbf{A}}$. In writing Eq. (18), account has been taken of the ergodicity of $\mathbf{X}(t)$ and of the fact that the second and the fourth integrals appearing in Eq. (17) possess a zero limit.

In order to find $\bar{\mathbf{A}}_{\mathrm{ML}}$, let us multiply each term of Eq. (15) by $\mathbf{Y}^T(t)$ and let us integrate into $[0,T]$, obtaining:

$$\int_0^T d\mathbf{Y}(\tau)\mathbf{Y}^T(\tau) = \bar{\mathbf{A}}\int_0^T \mathbf{Y}(\tau)\mathbf{Y}^T(\tau)d\tau + \bar{\mathbf{a}}\int_0^T \mathbf{Y}^T(\tau)d\tau + \bar{\mathbf{b}}\int_0^T \mathbf{Y}^T(\tau)dL(\tau) \tag{19}$$

If now $\mathbf{Y}(t)$ is replaced by $\mathbf{X}(t)$, each term of Eq. (18) is divided by T and the $\lim_{T\to\infty}$ is applied, then we obtain:

$$\bar{\mathbf{A}}_{\mathrm{ML}}E[\mathbf{X}\mathbf{X}^T] = E[\mathbf{m}[\mathbf{X}]\mathbf{X}^T] - \bar{\mathbf{a}}_{\mathrm{ML}}E[\mathbf{X}^T] \tag{20}$$

which, by taking into account Eq. (18), allows us to obtain $\bar{\mathbf{A}}_{\mathrm{ML}}$ as follows:

$$\bar{\mathbf{A}}_{\mathrm{ML}} = \Sigma_{\mathbf{X}}^{-1}(E[\mathbf{m}[\mathbf{X}]\mathbf{X}^T] - E[\mathbf{m}[\mathbf{X}]]E[\mathbf{X}^T]) \tag{21}$$

It is important to note that when parametric excitations act on the system, in order to model the true nonlinear system by a linear one, it is fundamental to find the estimate of the maximum likelihood of the term $\bar{\mathbf{b}}^{[2]}$, besides of the terms $\bar{\mathbf{a}}$ and $\bar{\mathbf{A}}$. With this aim in mind, the Ito equation with reference to $\mathbf{Y}^{[2]}(t)$ is written. Then, by taking into account Eq. (15), we write:

$$d\mathbf{Y}^{[2]}(t) = (\bar{\mathbf{A}}\otimes\mathbf{I} + \mathbf{I}\otimes\bar{\mathbf{A}})\mathbf{Y}^{[2]}(t)dt + (\bar{\mathbf{a}}\otimes\mathbf{I} + \mathbf{I}\otimes\bar{\mathbf{a}})\mathbf{Y}(t)dt + (\bar{\mathbf{b}}\otimes\mathbf{I} + \mathbf{I}\otimes\bar{\mathbf{b}})\mathbf{Y}(t)dL(t) + \bar{\mathbf{b}}^{[2]}\mathbf{q}dt \tag{22}$$

Integrating this equation into $[0,T]$, replacing $\mathbf{Y}(t)$ by $\mathbf{X}(t)$, applying the $\lim_{T\to\infty}$ and taking into account Eqs. (18) and (21), lastly, after some algebra, the following result is obtained:

$$(\bar{\mathbf{b}}^{[2]})_{\mathrm{ML}} = E[\mathbf{g}^{[2]}[\mathbf{X}]] \tag{23}$$

The results given in Eqs. (18), (21), and (23) evidence that the maximum likelihood estimates of the terms characterizing the linear system modeling the true nonlinear one are the stationary values of the quantities obtained in the previous section when the stochastic linearization method is applied on the coefficients of the Ito differential rule. Hence, these results can be considered as an extension of the Kozin results to the case of nonlinear systems excited by parametric Gaussian white noise excitations and characterized by a nonzero mean response.

4 Conclusions

In this paper two important properties of the stochastic linearization technique applied to nonlinear systems excited by purely external Gaussian white noises have been extended to the case in which the Gaussian white noise excitations are parametric. These two properties, found by Kozin, are: (1) the quantities obtained by the application of the stochastic linearization are the true linear quantities, in the sense that their exact value is able to give the first two statistical moments of the true response; (2) these quantities are the maximum likelihood stationary estimates of the coefficients characterizing the linear model fitting the original nonlinear system in a parameter estimation analysis.

It has been shown that, in the parametric white noise excitation case, these properties are confirmed only if the stochastic linearization is applied as made by the author in a previous paper, that is by linearizing the coefficients of the Ito differential rule related to the nonlinear system.

References

[1] Booton, R. C., 1953, "The Analysis of Non-linear Control Systems With Random Inputs," *Proceedings of the MRI Symposium on Nonlinear Circuits*, Polytechnic of Brooklyn, pp. 341–344.

[2] Kazakov, I. E., 1956, "Approximate Probability Analysis of the Operational Precision of Essential Non-linear Feedback Control," Autom. Remote Control (Engl. Transl.), **17**, pp. 423–450.

[3] Roberts, J. B., and Spanos, P. D., 1990, *Random Vibration and Statistical Linearization*, Wiley, New York.

[4] Crandall, S. H., 1978, "Heuristic and Equivalent Linearization Techniques for Random Vibration of Nonlinear Oscillators," *Proceedings of the Eighth International Conference on Nonlinear Oscillations*, Prague, Vol. 1, pp. 211–226.

[5] Kozin, F., 1987, "The Method of Statistical Linearization for Non-linear Stochastic Vibrations," *Nonlinear Stochastic Dynamic Engineering Systems*, F. Ziegler and G. I. Schueller, eds., IUTAM Symposium, Igls, pp. 45–55.

[6] Caughey, T. K., 1963, "Equivalent Linearization Techniques," J. Acoust. Soc. Am., **35**, pp. 1706–1711.

[7] Crandall, S. H., "On Statistical Linearization for Nonlinear Oscillators," *Problems of the Asymptotic Theory of Nonlinear Oscillations*, Acad. Sci Ukraine SSR, Naukova Dumka, Kiev, 1979.

[8] Kozin, F., 1987, "Parameter Estimation and Statistical Linearization," Stochastic Structural Mechanics, Y. K. Lin and G. I. Schueller, eds., Springer-Verlag, Berlin, pp. 259–267.

[9] Chang, R. I., and Young, G. E., 1989, "Methods and Gaussian Criterion for Statistical Linearization of Stochastic Parametrically and Externally Excited Nonlinear Systems," J. Appl. Mech., **56**, pp. 179–186.

[10] Wu, W. F., 1987, "Comparison of Gaussian Closure Technique and Equivalent Linearization Method," Probab. Eng. Mech., **2**, pp. 2–8.

[11] Falsone, G., 1992, "Stochastic Linearization of MDOF Systems Under Parametric Excitations," Int. J. Non-Linear Mech., **27**, pp. 1025–1037.

[12] Graham, A., 1981, *Kronecker Products and Matrix Calculus with Applications*, Horwood, Chichester, UK.

Local Contact Compliance Relations at Compaction of Composite Powders

Olle Skrinjar

Per-Lennart Larsson
e-mail: pelle@hallf.kth.se

Bertil Storåkers

KTH Solid Mechanics,
Royal Institute of Technology,
Osquars backe 1,
SE-100 44 Stockholm,
Sweden

Local contact behavior of composite powders has been investigated by using the finite element method. In previous analyses of such problems it has in general been assumed that one of the powder materials is rigid while the other deforms at loading as in such a case self-similarity prevails. This is a very good approximation for ceramic/metallic composites but may not be so when the composite consists of two materials of roughly equal hardness. An approximate compliance formula for describing this feature is proposed showing good agreement with corresponding finite element results for representative cases. [DOI: 10.1115/1.2165240]

1 Introduction

Powder compaction is a production method used for making details with complex material compositions and geometries. The principle of compaction has been known for several thousand years (~3000 B.C.) when the Egyptians used iron powder and compressed it into tools. Today several materials such as metals, alloys, polymers, and ceramics are used to produce compacted parts. One major advantage with this method is that the produced parts are near-net shaped and need little or no machining at all. This is a particularly advantageous feature for ceramics since machining is a difficult task. An additional advantage with composite powders is the possibility to design the material behavior as for example is done with low- and high-alloy steels including different volume fractions of copper, nickel, and molybdenum with ferrous material.

The process of compacting powder can be divided into three stages at least from a mechanical point of view. Filling and packing a container with powder is the first stage. Cold or hot pressing from an initial relative density near 0.60–0.70 vol % powder up to a relative density between 0.80 and 0.90 is the second stage, often called stage I. This stage is characterized of powder surrounded by open pores and each contact site between powders can be treated independently. The final stage (stage II), where additional pressure is applied up to almost full density, is characterized by pores that are sealed off and consequently the material behavior is more or less like a porous solid.

It is obvious that if the mechanical behavior can be analyzed in advance the costly experimental phase can be reduced before serial production of details and parts. There are several ways of analyzing the powder compaction process. Stage II is frequently studied using a macroscopic constitutive equation for porous solids and is often based upon phenomenological assumptions, see, e.g., Refs. [1–3]. Packing and stage I are more often analyzed using a micromechanical approach. With such an approach the local interaction and deformation between individual particles is determined and summarized over the entire compound in order to predict the global behavior. Several writers have addressed this issue over the last decades and a better understanding of the mechanical behavior at compaction has been achieved. Thirty years ago, Wilkinson and Ashby [4] presented a pioneering study where they investigated hot isostatic compaction of powders with power law creep behavior. Following the achievements from theoretical and experimental studies by Fischmeister and co-workers, [5,6], an almost complete theory for isostatic compaction was formulated by Helle et al. [7]. Further progress was made for more general loading situations in Refs. [8–10]. All of these investigations, however, have resulted in equations that are partly phenomenological and further investigations are needed in order to achieve a more consistent micromechanical description of the powder compaction problem.

In a micromechanical approach to model powder compaction, a major issue is the mechanical behavior at contact between two powder particles. This matter was studied in detail by Storåkers et al. [11] and Storåkers [12], based on numerical results from analyses of spherical indentation of solids described constitutively by power law creep, plastic flow theory, and general viscoplasticity [13–16]. With these results at hand Larsson et al. [17] analyzed cold and hot compaction of monolithic powders followed by a more general study of composite powders by Storåkers et al. [18]. The latter study, even though comprehensive, was still based on some simplifying approximations and further investigations were needed. One of these (simplifying) approximations concerns the assumption of affine motion of particles at compaction. In short, affine motion means that the motion of each particle is prescribed by the applied macroscopic strain field. This issue was, among other things, studied by Heyliger and McMeeking [19], Skrinjar and Larsson [20], and Martin et al. [21] using the discrete element method and by Procopio and Zavaliangos [22] using a multiparticle finite element method. These authors showed that initially in the powder compaction process, during packing and stage I, rearrangement, in addition to plastic flow, is an important mechanism of densification. Other mechanisms identified to control the densification are power law creep and diffusional flow. However, these latter mechanisms will be more dominant at hot compaction.

The previously discussed contact law developed by Storåkers et al. [11] and Storåkers [12] and used in Refs. [20,21,23,24], is valid if the two particles indenting each other are constitutively described by the same form of hardening behavior or if one of the particles is assumed rigid. Accordingly, the contact law is valid for compaction of monolithic powders or compaction of for example a metallic material and a ceramic powder (being much harder than the metal powder). However, if the two materials indenting each other are, say, power law materials with different strengths and exponents, the contact law mentioned above will not be applicable as emphasized by, e.g., Storåkers et al. [18]. Composite materials consisting of two or more deformable metals are used today since the capability of tailoring the mechanical properties of the final part is possible and of course desirable. Powder mixtures are the easiest way to provide the amount of alloying materials, e.g., copper, nickel, and/or molybdenum together with ferrous materials even though powders of pre-alloyed materials are also very common.

To the best of the authors' knowledge the accuracy of the above-discussed contact law in a general situation has not yet been investigated and the intention is to examine this issue. The investigation will be performed using the finite element method (FEM) and, in particular, the commercial multi-purpose program ABAQUS [25] will be utilized. An alternative approximate for-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 26, 2005; final manuscript received November 16, 2005. Review conducted by R. M. McMeeking.

mula describing contact in a more general situation at dissimilar powders is proposed and compared to results from straightforward FE-simulations.

2 Similarity Analysis: Exact and Approximate

It was stated above that at contact between two deformable particles of dissimilar materials, the advantage of self-similarity of contact solutions no longer applies in general. It is believed though that such solutions can still serve as useful tools when attempting to determine approximate, but accurate, force-indentation relations also in a more general situation. For this reason, the presentation of the present analysis, with such a relation as a goal, starts off by discussing the most important details of self-similarity solutions.

Indentation of a rigid sphere into a rigid-plastic material was analyzed by Biwa and Storåkers [16] based on plastic flow theory. The material was assumed to be power-law hardening and in uniaxial form given by

$$\sigma = \sigma_0 \epsilon^{1/m} \tag{1}$$

where the strength parameter σ_0 and the power law exponent m are material parameters. In a corresponding multiaxial form von Mises yield condition was used.

It was found that the shape of the pressure distribution becomes stationary and only depends on the power law exponent m. In particular the relation between the contact contour radius, a, and indentation depth, h, becomes self-similar, and reduces to

$$a^2 = 2c^2 R_0 h \tag{2}$$

where the indentation invariant is $c^2 = c^2(m)$ and R_0 is the relative curvature.

From high accuracy finite element calculations, i.e., Ref. [11], the relation

$$c^2 = 1.43 e^{(-0.97/m)} \tag{3}$$

has been specified by Fleck et al. [26].

As a result the force-depth relation may be written as

$$h = \frac{1}{c^2}\left(\frac{F}{\alpha \sigma_0 R_0^2}\right)^{2m/(2m+1)} R_0 \tag{4}$$

where F is the indentation force and $\alpha = 2^{(2m-1)/2m} 3^{(m-1)/m} \pi$. For clarity, it should be mentioned that in all cases to follow, solutions based on Eq. (4) will be referred to as the self-similarity solution.

Remembering that the pressure distribution only depends on the exponent m, it was found by Storåkers et al. [11] that in case of two spheres in contact with different radii, R_1, R_2, and of different strength, σ_1, σ_2, but of the same m, Eqs. (2) and (4) may be generalized for contact of two plastic spheres by introducing

$$\frac{1}{R_0} = \frac{1}{R_1} + \frac{1}{R_2} \tag{5}$$

and

$$\frac{1}{\sigma_0^m} = \frac{1}{\sigma_1^m} + \frac{1}{\sigma_2^m} \tag{6}$$

in conformity with Hertz theory for dissimilar linear elastic solids, and notation according to Fig. 1.

From similarity arguments it is then intuitively clear that the deformed contact region will still be spherical with Hertzian curvatures

$$\frac{1}{R_{01}} = \frac{1}{R_1} - \frac{1}{R}$$

$$\frac{1}{R_{02}} = \frac{1}{R_2} + \frac{1}{R} \tag{7}$$

and to be shown shortly that

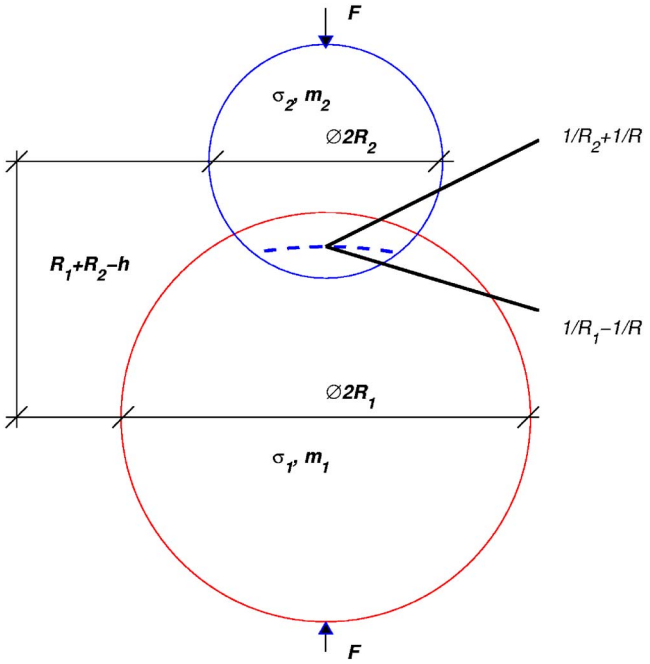


Fig. 1 Schematics of the two particle problem

$$\frac{1}{R} = \left[\frac{1}{R_1} - \left(\frac{\sigma_2}{\sigma_1}\right)^m \frac{1}{R_2}\right] \Bigg/ \left[\left(\frac{\sigma_2}{\sigma_1}\right)^m + 1\right] \tag{8}$$

where R is defined by the curvature change at deformation as in Fig. 1.

If two solids have different power law exponents, m_1, m_2, then the individual pressure distributions will not be conformal and self-similarity and stationarity will be lost. If it is assumed, however, that the deformed contact region will not deviate much from a spherical surface, the combined approach from two dissimilar spheres may by Eqs. (4) to (7) provide an estimate as

$$h = \frac{1}{c_1^2}\left[\frac{F}{\alpha_1 \sigma_1}\left(\frac{1}{R_1} - \frac{1}{R}\right)^2\right]^{2m_1/(2m_1+1)} \left(\frac{1}{R_1} - \frac{1}{R}\right)^{-1} + \frac{1}{c_2^2}\left[\frac{F}{\alpha_2 \sigma_2}\left(\frac{1}{R_2} + \frac{1}{R}\right)^2\right]^{2m_2/(2m_2+1)} \left(\frac{1}{R_2} + \frac{1}{R}\right)^{-1} \tag{9}$$

in obvious notation where R, the curvature due to deformation, may be determined from Eqs. (2) and (4)–(8) as

$$\left[\frac{F}{\alpha_1 \sigma_1}\left(\frac{1}{R_1} - \frac{1}{R}\right)^2\right]^{2m_1/(2m_1+1)} \left(\frac{1}{R_1} - \frac{1}{R}\right)^{-2} = \left[\frac{F}{\alpha_2 \sigma_2}\left(\frac{1}{R_2} + \frac{1}{R}\right)^2\right]^{2m_2/(2m_2+1)} \left(\frac{1}{R_2} + \frac{1}{R}\right)^{-2} \tag{10}$$

The combination of Eqs. (9) and (10) then provides an approximate force-depth relation for dissimilar spheres. In case of monolithic contact, $m_1 = m_2 = m$, the relation is exact and as already forecast Eq. (10) reduces to Eqs. (8) and (9) to Eq. (4) using the composite forms in Eqs. (5) and (6).

3 Numerical Analysis

The present numerical analysis concerns the mechanical behavior at initial and intermediate contact between two spherical particles with different material properties. Both spheres are described constitutively by classical von Mises elastoplasticity. The analysis conducted here is restricted to frictionless normal indentation as has been shown by Carlsson et al. [27] and Larsson and Storåkers [28] that compliance due to friction is almost negligible at normal indentation. The resulting axisymmetric boundary value

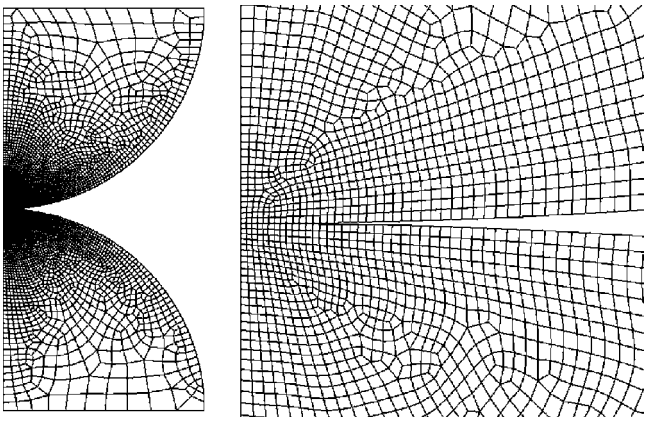


Fig. 2 FE mesh for analysis of contact between two particles with equal radii

problem is solved using the finite element method based on the commercial finite element program ABAQUS [25]. The finite element mesh for the case of two particles with equal radii where all symmetries have been accounted for is shown in Fig. 2 and consists of 6889 four-noded axisymmetric elements and 7466 nodes. In addition, large deformation theory was taken into account in all numerical calculations. The mesh was made denser toward the center of the contact region as can be seen in Fig. 2. In order to verify the accuracy of the results, a mesh with double density was used for selected cases and the results indicate that no further refinement of the mesh is necessary. A comparison with the well-known solution from Hertz elastic contact theory [29] was also conducted for a linear elastic material and the results from the FE model were in excellent agreement with the theoretical solution with a difference of less than 1% for global contact variables. Accordingly, these findings gave some definite confidence in the FE model.

Table 1 Material parameters used in the FE calculations

Material	E (GPa)	σ_y (MPa)	σ_0 (MPa)	m
A	455	50	1044	3
B	455	50	310	5
C	455	50	124	10
D	455	50	⋯	∞

A strict power law material behavior (pertinent to rigid plasticity) according to Eq. (1) is not available in the presently used version of ABAQUS [25]. Instead, the strain-hardening behavior of elasto-plastic materials was modeled using a power law relation according to

$$\sigma = \sigma_y + \sigma_0 \epsilon_p^{1/m} \tag{11}$$

where σ_y is the initial flow stress, σ_0 and m are material parameters introduced above, and ϵ_p is the (effective) plastic strain. In this investigation, materials with hardening exponents $m=3$, 5, 10, and ∞ were analyzed (the limiting case $m=\infty$ corresponding to ideal plasticity). Four representative materials were considered at first with details according to Table 1.

4 Results

This investigation studies the local contact behavior of composite powders and the results presented concern indentation of two spherical particles with different material behavior. Several materials have been investigated and chosen in order to give a representative view of the behavior at contact between dissimilar powders in practical situations. Initially, it was thought advisable to investigate whether or not and at what circumstances the finite element calculations based on the strain-hardening behavior in Eq. (11) could reproduce the self-similarity results in Eq. (4). Accordingly, contact between two spheres of the same material was investigated. This was done initially for power law materials A, B, C, and D in Table 1 and the results are shown in Fig. 3 where the normalized indentation force, $F/(\pi R_0^2 \sigma_0)$, is depicted as function of the normalized indentation depth, h/R_0, and compared to the corresponding self-similarity solution, Eq. (4). The radii of the two spheres are equal here and in the sequel. For the results pertinent to Eq. (4), the indentation invariant c^2 is determined from Eq. (3). Indentation is analyzed up to $0.01 h/R_0$ remembering that stage I compaction is of immediate interest here. Clearly, the agreement between the four sets of results from FE simulations

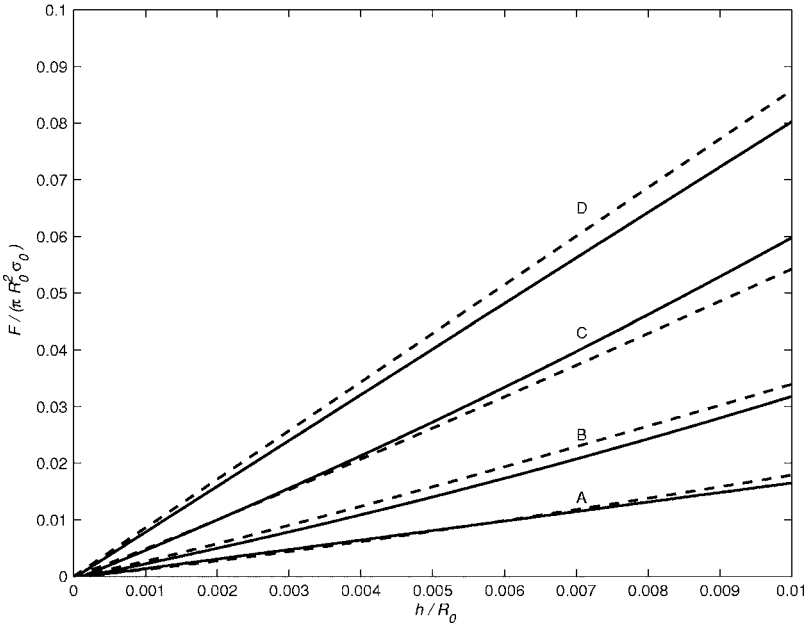


Fig. 3 (—), FE results and (- - -) self-similarity solutions at contact between particles of the same materials, A–D

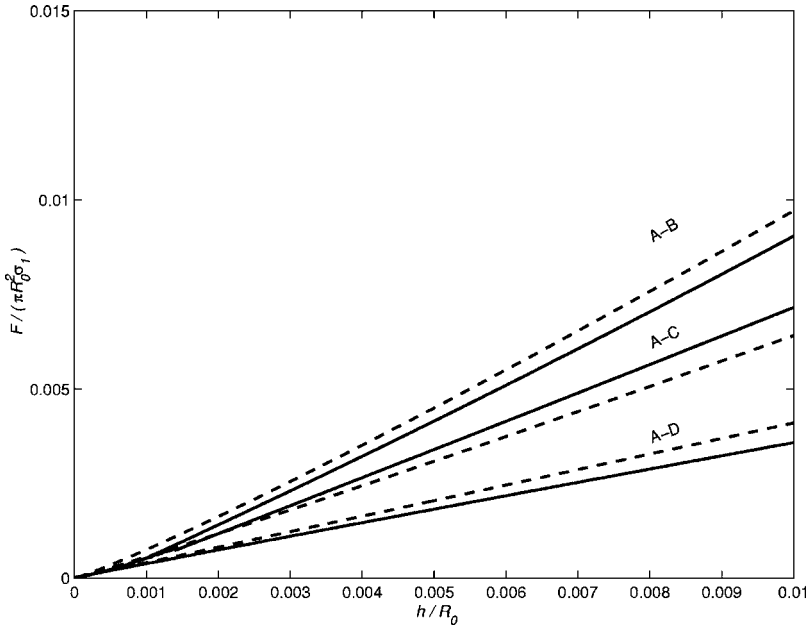


Fig. 4 (—) FE results and (- - -) similarity approximation for material combinations A-B, A-C, A-D, σ_1 is the strength parameter of material A in Table 1

and the self-similarity solution, Eq. (4), is good giving some further confidence as regards the accuracy of the numerical procedure and, indeed, also of the self-similarity solution.

The influence of effects of elasticity and large deformation on the self-similarity solution has recently been investigated in detail by Mesarovic and Fleck [30]. In particular as regards the results from the present perfectly plastic material D, Mesarovic and Fleck [30] found in similar circumstances that the self-similarity solution is relevant for $h/R < 0.015$ and $E/\sigma_0 > 10^3$, say.

With the results in Fig. 3 a foundation is now set to consider contact between two spheres of dissimilar material. In Fig. 4 the normalized indentation force, $F/(\pi R_0^2\sigma_1)$—here and in the following σ_1 is the strength parameter of material A in Table 1—is depicted as a function of the normalized indentation depth, h/R_0, at contact between material A ($m=3$) and materials B–D ($m=5$, 10, and ∞), respectively. The radii of the spheres are equal and the indentation invariant, c^2, is determined from Eq. (3). The results from the similarity approximation pertinent to mixed contact are also depicted in Fig. 4 and as may be seen, the results from the finite element calculations are quite satisfactory with a maximum deviation of less than 9%. When it comes to details of combination A–D, with $m=3$ and $m=\infty$, respectively, separation of the indentation depth, h, into two parts shows that the deformation of the perfectly plastic material in this case is six to seven times as high as that of the strain-hardening material. Accordingly, material A can roughly be considered as rigid indenting a soft particle and true self-similarity will prevail. Such a situation is very common at contact between a strain-hardening and a perfectly plastic particle. Increasing the yield stress of the perfectly plastic material will give essentially the same behavior as above.

Finally, also some results relevant to contact between spheres of different size were carried out. In this case the contact combination A-B was investigated with a size ratio $R_1/R_2=2$. In short, very good agreement was found between results from FE results and the similarity approximation, but for brevity this feature is not shown explicitly here.

5 Conclusions

Contact between dissimilar spheres was studied theoretically and numerically using the finite element method. In particular, force-indentation depth relations were determined with especially micromechanical investigations of composite powder compaction in mind. Based on previous self-similarity solutions for contact between spheres of similar materials, as well as presently performed finite element results, an approximate compliance formula describing contact in a general situation has been derived. Finite element simulations of representative cases have been performed indicating that the proposed formula produces high accuracy predictions.

One particular area where the present results are directly applicable concerns analyses of powder compaction using the discrete element method. In such a case, the approximate formula provides a simple and useful tool for describing the local contact behavior when studying compaction of composite powders.

Acknowledgment

This investigation was supported by the Swedish Research Council, Project No. 621-2001-2731, which is gratefully acknowledged. The writers are indebted to Professor Peter Gudmundson for pertinent and kind advice.

References

[1] Kuhn, H. A., and Downey, C. L., 1971, "Deformation Characteristics and Plasticity Theory of Sintered Powder Materials," Int. J. Powder Metall., **7**, pp. 15–25.
[2] Shima, S., and Oyane, M., 1976, "Plasticity Theory for Powder Compaction," Int. J. Mech. Sci., **18**, pp. 285–291.
[3] Jinka, A. G. K., and Lewis, R. W., 1994, "Finite Element Simulation of Hot Isostatic Pressing of Metal Powders," Comput. Methods Appl. Mech. Eng., **114**, pp. 249–272.
[4] Wilkinson, D. S., and Ashby, M. F., 1975, "Pressure Sintering by Power-Law Creep," Acta Metall., **23**, pp. 1277–1285.
[5] Fischmeister, H. F., Artz, E., and Olsson, L. R., 1978, "Particle Deformation and Sliding During Compression of Spherical Powders, a Study by Quantitative Metallurgy," Powder Metall., **21**, pp. 179–187.
[6] Fischmeister, H. F., and Artz, E., 1982, "Densification of Powders by Particle Deformation," Powder Metall., **26**, pp. 82–88.
[7] Helle, A. S., Easterling, K. E., and Ashby, M. F., 1985, "Hot-Isostatic Pressing Diagrams, New Developments," Acta Metall., **33**, pp. 2163–2174.
[8] Fleck, N. A., Kuhn, L. T., and McMeeking, R. M., 1992, "Yielding of Metal Powder Bonded by Isolated Contacts," J. Mech. Phys. Solids, **40**, pp. 1139–1162.
[9] Kuhn, L. T., and McMeeking, R. M., 1992, "Power Law Creep of Powder Bonded by Isolated Contacts," Int. J. Mech. Sci., **34**, pp. 563–573.

[10] Fleck, N. A., 1995, "On the Cold Compaction of Powders," J. Mech. Phys. Solids, **43**, pp. 1409–1431.
[11] Storåkers, B., Biwa, S., and Larsson, P. -L., 1997, "Similarity Analysis of Inelastic Contact," Int. J. Solids Struct., **34**, pp. 3061–3083.
[12] Storåkers, B., 1997, "Local Contact Behaviour of Viscoplastic Particles," *Proceedings of the IUTAM Symposium on Mechanics of Granular Flow and Powder Compaction*, Kluwer Academic, Amsterdam.
[13] Bower, A. F., Fleck, N. A., Needleman, A., and Ogbonna, N., 1993, "Indentation of a Power Law Creeping Solid," Proc. R. Soc. London, Ser. A, **441**, pp. 97–124.
[14] Storåkers, B., and Larsson, P. -L., 1994, "On Brinell and Boussinesq Indentation of Creeping Solids," J. Mech. Phys. Solids, **42**, pp. 307–332.
[15] Fleck, N. A., Ogbonna, N., and Cocks, A. C. F., 1995, "Transient Creep Analysis of Ball Indentation," Int. J. Mech. Sci., **37**, pp. 1179–1202.
[16] Biwa, S., and Storåkers, B., 1995, "An Analysis of Fully Plastic Brinell Indentation," J. Mech. Phys. Solids, **43**, pp. 1303–1334.
[17] Larsson, P. -L., Biwa, S., and Storåkers, B., 1996, "Analysis of Cold and Hot Isostatic Compaction of Spherical Particles," Acta Mater., **44**, pp. 3655–3666.
[18] Storåkers, B., Fleck, N. A., and McMeeking, R. M., 1999, "The Viscoplastic Compaction of Composite Powders," J. Mech. Phys. Solids, **47**, pp. 785–815.
[19] Heyliger, P. R., and McMeeking, R. M., 2001, "Cold Plastic Compaction of Powders by a Network Model," J. Mech. Phys. Solids, **49**, pp. 2031–2054.
[20] Skrinjar, O., and Larsson, P. -L., 2004, "On Discrete Element Modelling of Compaction of Powders with Size Ratio," Comput. Mater. Sci., **31**, pp. 131–146.
[21] Martin, C. L., Bouvard, D., and Shima, S., 2003, "Study of Particle Rearrangement During Powder Compaction by the Discrete Element Method," J. Mech. Phys. Solids, **51**, pp. 667–693.
[22] Procopio, A. T., and Zavaliangos, A., 2005, "Simulation of Multi-Axial Compaction of Granular Media from Loose to High Relative Densities," J. Mech. Phys. Solids, **47**, pp. 1523–1551.
[23] Skrinjar, O., and Larsson, P. -L., 2004, "Cold Compaction of Composite Powders with Size Ratio," Acta Mater., **52**, pp. 1871–1884.
[24] Martin, C. L., and Bouvard, D., 2003, "Study of the Cold Compaction of Composite Powders by the Discrete Element Method," Acta Mater., **51**, pp. 373–386.
[25] ABAQUS Inc., Pawtucket, RI, *ABAQUS v. 6.4*, 2004.
[26] Fleck, N. A., Storåkers, B., and McMeeking, R. M., 1997, "The Viscoplastic Compaction of Powders," *Proceedings of the IUTAM Symposium on Mechanics of Granular Flow and Powder Compaction*, Kluwer Academic, Amsterdam, The Netherlands.
[27] Carlsson, S., Biwa, S., and Larsson, P. -L., 2000, "On Frictional Effects of Inelastic Contact Between Spherical Bodies," Int. J. Mech. Sci., **42**, pp. 107–128.
[28] Larsson, J., and Storåkers, B., 2002, "On Oblique Contact of Creeping Solids," J. Mech. Phys. Solids, **50**, pp. 2029–2055.
[29] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge.
[30] Mesarovic, S. D., and Fleck, N. A., 2000, "Spherical Indentation of Elastic-Plastic Solids," Proc. R. Soc. London, **455**, pp. 2707–2728.

Localized Bending Waves in a Rib-Reinforced Elastic Orthotropic Plate

M. Belubekyan

K. Ghazaryan

Institute of Mechanics,
National Academy of Sciences of Armenia,
Yerevan, Armenia

P. Marzocca

C. Cormier

Mechanical and Aeronautical Engineering Department,
Clarkson University,
Potsdam, NY 13699-5725

Localized bending waves in a thin elastic orthotropic cantilever plate reinforced by a rigid rib are studied. A condition under which the edge waves can be eliminated is determined. The condition requires that the rib have a certain minimal stiffness. Such waves are time-varying, and have spatially nonuniform bending perturbations that are localized in the proximity of free surface and are quickly decaying to zero. A general solution is presented and the particular case of an isotropic reinforced plate is shown. An inverse method is described for identifying the elastic properties of the rib. [DOI: 10.1115/1.2165242]

1 Introduction

The study of localized bending waves in an elastic isotropic semi-infinite plate was presented first in Ref. [1] and further developed in Refs. [2–4]. The theory of flexural edge waves was given in Ref. [1] using Kirchhoff's plate theory. The existence of a flexural wave, guided by the free edge of a semi- infinite isotropic elastic thin plate, was also demonstrated. Such waves have properties analogous to a Rayleigh wave on an elastic half-space, in that they decay exponentially with distance from the edge. The flexural edge wave is also predicted by Mindlin's plate theory and this prediction agrees with measured data. The propagation of flexural waves guided by the tip of a wedge or the free end of a ridge on a substrate has been studied in some depth [3–6]. Using classical plate theory [3] an explicit expression was obtained for the speed of a flexural edge wave. The dispersion relation for a flexural plate edge wave was found in Ref. [4] by taking the limit of a wedge with zero internal angle. In Ref. [5] the existence and behavior of these waves is discussed in detail and the same wave speed as in Ref. [3] is obtained.

In this paper the plate is modeled as an orthotropic cantilever with one edge being free from mechanical stresses and reinforced with rigid rib, the opposite edge is fixed, while the other two edges of the plate are assumed to be simply supported. The rib is modeled as an elastic beam. The dynamic problem for the elastic bending waves is considered and the necessary and sufficient conditions for the existence of localized bending waves are derived. These conditions depend on elastic plate and rib properties. Some inverse problems of determining the beam stiffness depending on the plate bending vibration frequencies are also considered. Finally, the effect of the localized bending waves and the implications of the proposed solution method are presented and pertinent conclusions are outlined. The paper may provide some insight toward developing a procedure to detect damage in a rib for more complex systems.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 5, 2005; final manuscript received November 21, 2005. Review conducted by K. Ravi-Chandar.

2 Governing Equations and Boundary Conditions

Consider a rectangular elastic plate in a Cartesian reference system (x,y,z), so that the plane (xOy) coincides with the plate middle surface and z is the coordinate along the thickness of a plate, $x\in[0,\infty]$, $y\in[0,b]$, $z\in[-h,h]$, see Fig. 1.

Based on Kirchhoff's hypothesis, the plate bending vibration equation can be written as [7]

$$D_{11}\frac{\partial^4 w}{\partial x^4}+2(D_{11}+2D_{66})\frac{\partial^4 w}{\partial x^2\,\partial y^2}+D_{22}\frac{\partial^4 w}{\partial y^4}+2\rho h\frac{\partial^2 w}{\partial t^2}=0 \quad (1)$$

Here $w(x,y)$ is the plate mid-surface normal displacement, $2h$ is the plate thickness, ρ is the density of the plate material, and D_{11}, D_{12}, D_{22}, D_{66} are physical constants characterizing plate stiffness.

In the case of isotropic plate $D_{11}=D_{22}=(2/3)Eh^3/(1-\upsilon^2)$; $D_{12}=D_{11}\upsilon$; $D_{66}=(1/3)Eh^3(1+\upsilon)$, where E and υ are the elastic modulus and Poisson's ratio, respectively.

The stress couples M_x and H can be expressed as

$$M_x=-D_{11}\frac{\partial^2 w}{\partial x^2}-D_{12}\frac{\partial^2 w}{\partial y^2} \quad (2a)$$

$$H=-2D_{66}\frac{\partial^2 w}{\partial x\,\partial y} \quad (2b)$$

while the normal and generalized normal stress resultants N_x, $\tilde{N}_x$ acting at $x=$const can be cast as

$$N_x=-\frac{\partial}{\partial x}\left[D_{11}\frac{\partial^2 w}{\partial x^2}+(D_{12}+2D_{66})\frac{\partial^2 w}{\partial y^2}\right] \quad (2c)$$

$$\tilde{N}_x=-\frac{\partial}{\partial x}\left[D_{11}\frac{\partial^2 w}{\partial x^2}+(D_{12}+4D_{66})\frac{\partial^2 w}{\partial y^2}\right] \quad (2d)$$

On plate edges, e.g., $y=0,b$, the simply supported boundary condition has been assumed, implying that

$$w=0,\quad \frac{\partial^2 w}{\partial x^2}=0\quad \frac{\partial^2 w}{\partial y^2}=0 \quad (3)$$

The edge $x=0$ is supposed to be free from mechanical stresses and reinforced with a rigid rib, which is modeled as an elastic beam. On this edge, the following boundary conditions are applied [8]

$$M_x=A_0\frac{\partial}{\partial y}\left(\frac{\partial^2 w}{\partial x\,\partial y}\right)\quad \tilde{N}_x=D_0\frac{\partial^4 w}{\partial y^4} \quad (4)$$

Here, A_0 and D_0 are the twist and bending stiffness of the beam, respectively.

Taking into account Eqs. (2) and (4), the boundary conditions at $x=0$ can be written as

$$D_{11}\frac{\partial^2 w}{\partial x^2}+D_{12}\frac{\partial^2 w}{\partial y^2}+A_0\frac{\partial^3 w}{\partial x\,\partial y^2}=0 \quad (5a)$$

$$\frac{\partial}{\partial x}\left[D_{11}\frac{\partial^2 w}{\partial x^2}+(D_{12}+4D_{66})\frac{\partial^2 w}{\partial y^2}\right]+D_0\frac{\partial^4 w}{\partial y^4}=0 \quad (5b)$$

As a limiting case, if the plate is semi-infinite, the attenuation (localization) condition for the out of plane displacement $w(x,y,t)$ as $x\to\infty$ is

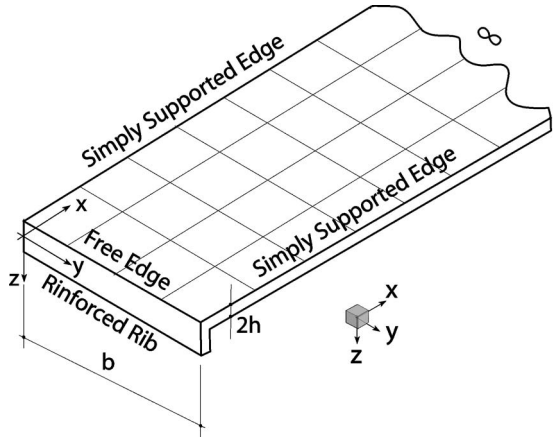


Fig. 1 Model of a cantilever plate with one free and rib reinforced edge

$$\lim_{x\to\infty} w(x,y,t) = 0 \tag{6}$$

The governing equation (1) is solved in conjunction with the boundary conditions, Eqs. (3)–(5), while accounting for the attenuation condition, Eq. (6). This solution, if it exists, will determine the localized bending waves near the edge $x=0$. It is worth noting that the study of localized bending waves in an elastic isotropic semi-infinite plate with a free edge ($M_x=N_x=0$) is given in Ref. [1].

3 Localized Bending Waves Solution for Orthotropic Cantilever Reinforced Plate

We assume the solution of the governing equation (1), so as to satisfy to boundary conditions (3), is of the form

$$w_n(x,y,t) = G_n(x)\sin(\lambda_n y)\exp(i\omega t) \tag{7}$$

where ω is the vibration frequency, $\lambda_n=\pi n/b, n=1,2,3,\ldots$.

The function $G_n(x)$ satisfies the ordinary differential equation

$$\frac{d^4G}{dx^4} - 2(\alpha_2+\alpha_3)\lambda_n^2\frac{d^2G}{dx^2} + \alpha_1\lambda_n^4(1-\eta_n^2)G = 0 \tag{8a}$$

and boundary conditions at $x=0$

$$\frac{d^2G}{dx^2} - \alpha_2\lambda_n^2 G + \beta\lambda_n^2\frac{dG}{dx} = 0 \tag{8b}$$

$$\frac{d^3G}{dx^3} - \lambda_n^2(\alpha_2+4\alpha_3)\frac{dG}{dx} + \gamma\lambda_n^4 G = 0 \tag{8c}$$

where the following dimensionless parameters have been introduced

$$\alpha_1=\frac{D_{22}}{D_{11}} \quad \alpha_2=\frac{D_{12}}{D_{11}} \quad \alpha_3=\frac{2D_{66}}{D_{11}} \quad \eta_n^2=\frac{2\rho h\omega^2}{D_{22}\lambda_n^4}$$

$$\beta=\frac{A_0}{D_{11}} \quad \gamma=\frac{D_0}{D_{11}}$$

It is assumed that Eq. (8) has solutions satisfying the attenuation condition Eq. (6), [9,10]

$$G_n(x) = C_{1n}\exp(-\lambda_n p_1 x) + C_{2n}\exp(-\lambda_n p_2 x) \quad \text{where } (p_1>0, p_2>0) \tag{9}$$

The parameters p_1 and p_2 are as follows

$$p_1 = \sqrt{\alpha_2+\alpha_3+\sqrt{(\alpha_2+\alpha_3)^2-\alpha_1(1-\eta_n^2)}} \tag{10a}$$

$$p_2 = \sqrt{\alpha_2+\alpha_3-\sqrt{(\alpha_2+\alpha_3)^2-\alpha_1(1-\eta_n^2)}} \tag{10b}$$

From Eqs. (10*a*) and (10*b*) the condition for the dimensionless frequency η_n of localized waves is

$$\eta_*^2 < \eta_n^2 < 1 \tag{11a}$$

where

$$\eta_*^2 = 1 - \frac{(\alpha_2+\alpha_3)^2}{\alpha_1} \tag{11b}$$

As is well-known, Refs. [9,11], for any orthotropic material $\alpha_1 > (\alpha_2+\alpha_3)^2$ always, while for any isotropic material $\alpha_1 \equiv (\alpha_2+\alpha_3)^2$ implying that $\eta_* \equiv 0$.

Satisfying the boundary conditions (5), a system of simultaneous algebraic equations has to be solved in terms of the unknown constants C_{n1}, C_{n2}

$$C_{n1}(p_1^2+\beta\lambda_n p_1-\alpha_2) + C_{n2}(p_2^2+\beta\lambda_n p_2-\alpha_2) = 0 \tag{12a}$$

$$C_{n1}[p_1^3-(\alpha_2+2\alpha_3)p_1-\gamma\lambda_n] + C_{n2}[p_2^3-(\alpha_2+2\alpha_3)p_2-\gamma\lambda_n] = 0 \tag{12b}$$

Therefore the dimensionless frequency η_n can be derived from the nontrivial solution of the system (12)

$$(p_2-p_1)K_*(\eta_n) = 0 \tag{13a}$$

where

$$K_*(\eta_n) \equiv [p_1(\eta_n)p_2(\eta_n)]^2 + 2\alpha_3 p_1(\eta_n)p_2(\eta_n) + [(\gamma\lambda_n + \beta\lambda_n p_1(\eta_n)p_2(\eta_n))(p_1(\eta_n)+p_2(\eta_n))] + \gamma\beta\lambda_n^2 - \alpha_2^2 \tag{13b}$$

When $p_1=p_2$ we have $C_{1n}=C_{2n}=0$, and consequently $G_n(x)\equiv 0$

When $p_1\neq p_2$, Eq. (13*a*) implies that

$$K_*(\eta_n) = 0 \quad \text{for } \eta_n \in [\eta_*, 1] \tag{14}$$

On the upper and lower bounds of η_n we have

$$K_*(\eta_n)\big|_{\eta_n=\eta_*} = \alpha_3(3\alpha_3+4\alpha_2) + 2\sqrt{\alpha_3+\alpha_2}[\gamma\lambda_n+\beta\lambda_n(\alpha_3+\alpha_2)] > 0 \tag{15a}$$

$$K_*(\eta_n)\big|_{\eta_n=1} = \gamma\beta\lambda_n^2 + \gamma\lambda_n\sqrt{2(\alpha_2+\alpha_3)} - \alpha_2^2 \tag{15b}$$

The derivative of the function $K_*(\eta_n)$ is found

$$\frac{dK_*}{d\eta_n} = -2\eta_n\left[(p_2-p_1)\times\frac{(\alpha_2+\alpha_3)\alpha_1+\alpha_1\gamma\lambda_n+2\alpha_1\beta\lambda_n p_1 p_2+2\alpha_1\alpha_3(p_1+p_2)}{4p_0p_1p_2}\right] - 2\eta_n\left[\alpha_1+\frac{\alpha_1\beta\lambda_n(p_1+p_2)}{p_1p_2}\right] < 0 \tag{16}$$

where $p_0 = \sqrt{(\alpha_2+\alpha_3)^2-\alpha_1(1-\eta_n^2)}\,\gamma\beta\lambda_n^2 + \gamma\lambda_n\sqrt{2(\alpha_2+\alpha_3)} - \alpha_2^2 > 0.$ (17)

From Eq. (16) it follows that the function $K_*(\eta_n)$ is monotonous decreasing in $\eta_n\in[\eta_*,1]$. Based on Eqs. (14)–(16) and the property of the function $K_*(\eta_n)$ the necessary and sufficient condition of the existence of localized waves is extracted from Eq. (15*b*)

$$\gamma\beta\lambda_n^2 + \gamma\lambda_n\sqrt{2(\alpha_2+\alpha_3)} - \alpha_2^2 < 0 \tag{18}$$

When the rib is absent, i.e., $\gamma=\beta=0$, the condition of Eq. (18) always holds, while the presence of a rigid rib can eventually eliminate the localized wave. Equation (18) also indicates a cutoff wavelength in the y-direction, therefore this theory can only predict waves of short wavelength.

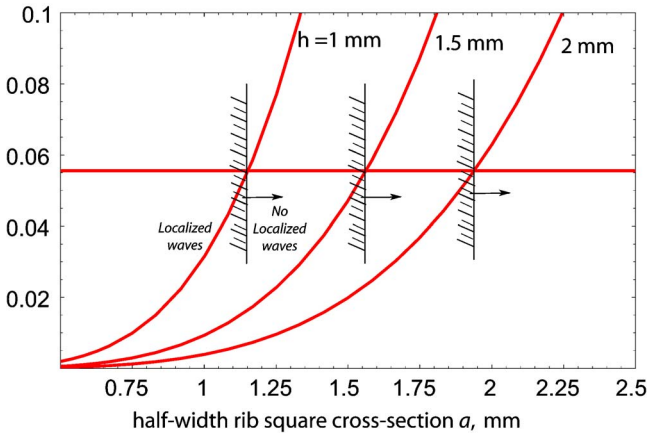


Fig. 2 Condition for the existence of localized bending waves in an isotropic rib reinforced plate, b=100 mm

4 Isotropic Plate With Rectangular Square Cross Section Rib

For an isotropic plate ($\alpha_1=1, \alpha_2=\upsilon, \alpha_3=1-\upsilon$) Eq. (13) can be written as

$$1-\eta_n^2+2(1-\upsilon)\sqrt{1-\eta_n^2}+(\gamma\lambda_n+\beta\lambda_n\sqrt{1-\eta_n^2})(\sqrt{1+\eta_n^2}+\sqrt{1-\eta_n^2})+\gamma\beta\lambda_n^2-\upsilon^2=0 \tag{19}$$

while Eq. (18) changes to

$$\gamma\beta\lambda_n^2+\sqrt{2}\gamma\lambda_n-\upsilon^2<0 \tag{20}$$

Next, we will consider a plate with a rectangular cross section rib beam.

For such a rib the twist and bending stiffness, A_0 and D_0, are

$$A_0\approx\frac{9a^4E_0}{4}\quad D_0=\frac{4a^4E_0}{3} \tag{21}$$

where a is the half-width of a square rib cross section, while E_0 is its elastic modulus. Recasting Eq. (20) in the form

$$\frac{n^2\pi^2a^8}{b^2h^6}\left(\frac{E_0}{E}\right)^2+\frac{\sqrt{2}(1-\upsilon^2)n\pi na^4}{bh^3}\frac{E_0}{E}-(1-\upsilon^2)^2\upsilon^2<0 \tag{22}$$

the localized waves can be eliminated if the condition

$$\frac{\pi a^4}{bh^3}\left(\frac{E_0}{E}\right)>\frac{\sqrt{2}}{2}(1-\upsilon^2)(\sqrt{1+2\upsilon^2}-1) \tag{23}$$

is fulfilled.

In Fig. 2 the condition for the existence of localized bending waves in an isotropic rib reinforced plate for various plate thicknesses is shown. Poission's ratio has been assumed to be 0.3, while b=100 mm. It appears that localized bending waves can be eliminated by selecting the reinforcement appropriately. Larger domains for localized bending waves are expected for larger plate thicknesses.

5 Identification of Rib Elastic Properties

The rib elastic proprieties can be identified via an inverse approach. Based on Eq. (13) for localized vibration we can investigate an inverse problem of identification of rib elastic proprieties using frequencies obtained from experimental data [12,13].

From Eq. (13) it follows that if two frequencies, for example the first and second modal frequencies of the system under consideration, are known then the elastic properties of the rib, i.e., the twist and bending stiffness of the beam (parameters β and γ, respectively) can be determined in the following way.

We define the twist and bending stiffness of the beam as

$$\beta=\frac{A_0}{D_{11}} \tag{24a}$$

$$\gamma=\frac{D_0}{D_{11}} \tag{24b}$$

Assuming that the first two frequencies η_1, η_2 are known, the parameters β and γ can be determined by solving the following simultaneous system of equations

$$(\gamma\lambda_1+\beta\lambda_1p_{11}p_{21})(p_{11}+p_{21})+\gamma\beta\lambda_1^2=\alpha_2^2-(p_{11}p_{21})^2-2\alpha_3p_{11}p_{21} \tag{25a}$$

$$(\gamma\lambda_2+\beta\lambda_2p_{12}p_{22})(p_{12}+p_{22})+\gamma\beta\lambda_2^2=\alpha_2^2-(p_{12}p_{22})^2-2\alpha_3p_{12}p_{22} \tag{25b}$$

where the following notations are used $p_{11}\equiv p_1(\eta_1), p_{21}\equiv p_2(\eta_1), p_{22}\equiv p_2(\eta_2), p_{12}\equiv p_1(\eta_2)$.

6 Conclusions

A study of the localized bending wave in a thin elastic orthotropic cantilever plate reinforced by a rigid rib is presented. A general solution is given and the particular case of an isotropic reinforced plate is analyzed. The bending vibration equation is solved in conjunction with appropriate boundary conditions and an avenue to identify the rib elastic proprieties through an inverse approach is described.

References

[1] Konenkov, Y. W., 1960, "On Rayleigh Type Flexural Waves," Sov. Phys. Acoust., **6**, pp. 122–123.

[2] Ambartsumyan, S. A., and Belubekyan, M. V., 1994, "On Bending Waves Localized Along the Edge of a Plate," Int. J. App. Mech. Translation Prikladnaya Mechanica, **30**, pp. 135–140.

[3] Sinha, B. K., 1974, "Some Remarks on Propagation Characteristics of Ridge Guides for Acoustic Waves at Low Frequencies," J. Acoust. Soc. Am., **56**, pp. 16–18.

[4] McKenna, J., Boyd, G. D., and Thurston, R. N., 1974, "Plate Theory Solution for Guided Flexural Acoustic Waves Along the Tip of a Wedge," IEEE Trans. Sonics Ultrason., **21**, pp. 178–186.

[5] Norris, A. N., Krylov, V. V., and Abrahams, I. D., 1998, "Flexural Edge Waves and Comments on 'A New Bending Wave Solution for the Classical Plate Equation'," J. Acoust. Soc. Am., **104**, pp. 2220–2222.

[6] Thurston, R. N., and McKenna, J., 1974, "Flexural Acoustic Waves Along the Edge of a Plate," IEEE Trans. Sonics Ultrason., **21**, pp. 296–297.

[7] Ambartsumyan, S. A., 1991, *Theory of Anisotropic Plates*, Hemisphere, New York.

[8] Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw Hill, New York.

[9] Norris, A. N., 1994, "Flexural Edge Waves," J. Sound Vib., **171**, pp. 571–573.

[10] Mkrtchyan, H. P., 2003, "Localized Bending Waves in an Elastic Orthotropic Plate," Proceedings of NAS Armenia, Mechanics, 56(4), pp. 66–68.

[11] Sarkisyan, V. S., 1970, "Some Problems of Elasticity Theory of Anisotropy Body," Yerevan State Univ., 443 (in Russian).

[12] Gnuni, V. Ts., and Oganesyan, Z. B., 1991, "Determination of Boundary Conditions of a Circular Ring Plate on Given Frequencies," Proceedings of NAS Armenia, Mechanics, 44(4), pp. 9–16 (in Russian).

[13] Aktyamov, A. M., and Mouftakhov, A. V., 2004, "Identification of Boundary Conditions Using Natural Frequencies," Inverse Probl. Sci. Eng., **12**(4), pp. 393–408.

Spectral Characteristics of the Near-Wall Turbulence in an Unsteady Channel Flow

Sedat F. Tardu
Laboratoire des Ecoulements Géophysiques et Industriels, B.P. 53 X, 38041 Grenoble, France
e-mail: sedat.tardu@hmg.inpg.fr

The modulation characteristics of the turbulent wall shear stress and longitudinal intensities in the inner layer are experimentally investigated in an unsteady channel flow wherein the centerline velocity varies in time in a sinusoidal manner. The fluctuating wall shear stress and velocity signals are temporally filtered and subsequently phase averaged. It is shown that the outer structures corresponding to the low spectrum range have a constant time lag with respect to the centerline velocity modulation. The inner active structures, in particular those with a frequency band containing the mean ejection frequency of the corresponding steady flow dominate the dynamics of the near-wall unsteady turbulence. The structures respond to the imposed shear oscillations in a complex way, depending both on their characteristic scales and the thickness of the oscillating shear zone in which they are embedded. [DOI: 10.1115/1.2166650]

Unsteady turbulent shear flows are encountered in many practical situations in aerohydrodynamics, aeroacoustics or biofluid dynamics. Past research on pulsed pipe or channel flows was first focused on the eventual effects of the forced velocity oscillations on the time-mean flow and second on the modulation characteristics of the oscillating velocity field and the near-wall turbulence. There is now an established consensus that:

- The time-mean flow is unaffected by the imposed unsteadiness even in the presence of large imposed amplitudes and frequencies that may cause reverse flow near the wall.
- In the imposed high-frequency regime, the oscillating shear is confined in the low buffer layer and leads to the coexistence of a purely oscillating viscous Stokes flow with an unaffected time-mean flow.
- The turbulence cannot follow the rapid imposed unsteadiness when the time period becomes comparable with the median time scale of the near-wall turbulence. The turbulent shear stresses become frozen during the oscillation cycle under these circumstances.

Despite significant advances in the understanding and modeling of forced internal wall flows, there is still some lack in understanding the reaction of the fine turbulence structure to imposed time periodical shear [1]. One of the questions that arise concerns the spectral characteristics of unsteady near-wall turbulence. We partly investigated these points in [2] by determining the impact of the unsteadiness on the inner and outer layer structures but only in the low buffer layer. We extend and discuss detailed results in this paper in the entire inner layer.

The experiments were performed in the unsteady water channel described in detail in [3]. The centerline velocity was held constant and equals $\bar{U}_c=17.5$ cm/s. This corresponds to a friction velocity of $\bar{u}_\tau=0.85$ cm/s and a Reynolds number based on the half-height of the channel of $\mathrm{Re}_h=\bar{U}_c h/\nu=8800$. The imposed amplitude was 20% of the centerline velocity throughout the whole study. The imposed frequency in wall units $f^+=f(\nu/\bar{u}_\tau^2)$, where ν is the cinematic viscosity, varied by a factor of 24 from $f^+=2\times10^{-4}$ to $f^+=60\times10^{-4}$. Hereafter, $()^+$ will designate variables normalized by the viscosity and time mean shear velocity. The imposed frequency range investigated here covers $l_s^+=38-7$ in terms of the frequency parameter $l_s^+=\sqrt{1/\pi f^+}$, which is the viscous Stokes length normalized by $l_\nu=\nu/\bar{u}_\tau$. The wall shear stress and the velocity measurements were performed by means of a flush-mounted TSI-1268 W hot film at the wall and a TSI 1276-10 W hot film located in the flow. Further details are provided in Tardu and Vezin [2].

The classical triple decomposition is used. A quantity q is decomposed into a mean $\bar{q}$, an oscillating $\tilde{q}$, and fluctuating q' component. The angle brackets designate the phase average, i.e., $\langle q\rangle=\bar{q}+\tilde{q}$. The modulation characteristics of $\langle q\rangle$ are described by the amplitude $A_{\tilde{q}}$ and phase $\phi_{\tilde{q}}$ of the fundamental mode. The relative amplitude $a_{\tilde{q}}=A_{\tilde{q}}/\bar{q}$ is also introduced for convenience.

In a way similar to Naguib and Wark [4], we use here three digital zero-phase shift filters; namely, Filter 0 with bandpass in wall units $\Delta f_0^+=0-0.0045$, Filter 1 (bandpass $\Delta f_1^+=0.0055-0.022$), and Filter 2 ($\Delta f_1^+=0.0316-0.0482$) to identify outer (Filter 0) and inner (Filters 1 and 2) structures and their characteristics. The filtering is processed through well-designed zero-phase shift 128-point finite impulse response digital filters. The reader is referred to [2] for further important details and related discussions.

The phase shifts of the filtered signals are denoted by $\Delta\phi_i=\phi_{iu'\tilde{u}'}-\phi_{\tilde{u}}$ with $i=0$, 1, and 2 and the corresponding time lags by $\Delta t_i=\Delta\phi_i/2\pi f^+$. The time lag of the modulation of the contribution of the outer structures is remarkably constant in the viscous layer at $y^+<50$ with $\Delta t^+_{0,u'\tilde{u}'}\approx-75$ as shown in Fig. 1. In the low log-layer, however, the behavior of $\Delta t^+_{0,u'\tilde{u}'}$ changes appreciably. The phase shift $\phi_{0,u'\tilde{u}'}-\phi_{\tilde{u}}$ at $y^+=100$ decreases first sharply in the imposed low-frequency range $f^+<0.002$, becomes subsequently constant and joins the line $\Delta t^+_{0,u'\tilde{u}'}\approx-75$ only in the imposed high-frequency regime. Thus, the response time of the outer structures is constant in the viscous layer $y^+<50$ with a repercussion at the wall of $\Delta t^+_{0,\tau'\tau'}\approx-125$ independent of the imposed frequency. The time-lag difference $\Delta t^+_{0,\tau'\tilde{\tau}'}-\Delta t^+_{0,u'\tilde{u}'}\approx-50$ may be expressed as $\Delta t^+=-y_0^+/v_0^+$ where y_0 and v_c stand, respectively, for the distance to the wall of the outer edge of the viscous layer and a characteristic wall normal velocity. One finds $\Delta t^+=-50$ by taking $y_0^+=50$ and $v_c^+=[\sqrt{\overline{v'v'}}/\bar{u}_\tau]_{y^+=y_0^+}\approx1$, i.e., by assuming that the convection velocity is approximately equal to the rms wall normal velocity at the outer edge of the viscous layer. An equivalent assumption could be that the wall normal convection velocity in the viscous layer is about $v_c=\bar{u}_\tau$, as suggested by Eckelmann [5]. The constancy of $\Delta t^+_{0,u'\tilde{u}'}$ at $y^+<50$ points to the difference of the diffusion mechanism governing the passive eddies and the $\langle u'u'\rangle$ modulation. The lack of the diffusion of the outer eddies in the sense we discussed before is in concordance with the idealized inviscid picture of the passive structures.

Figures 2(*a*) and 3 show how the phase shift of the turbulent longitudinal intensity modulation related to the active eddies differs from the passive ones in the entire inner layer. At a given y^+, the phase shift $\phi_{1,u'\tilde{u}'}-\phi_{\tilde{u}}$ decreases linearly until a critical imposed frequency f^+_{cr} beyond which $\phi_{1,u'\tilde{u}'}-\phi_{\tilde{u}}$ is constant, or equivalently the time lag $\Delta t^+_{1,u'\tilde{u}'}=(\phi_{1,u'\tilde{u}'}-\phi_{\tilde{u}})/2\pi f^+$ decreases with f^+ through $\Delta t^+_{1,u'\tilde{u}'}\propto1/f^+$. This is clearly perceptible in Fig. 2(*a*). The critical frequency depends upon y^+. The best physical

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received October 4, 2004; final manuscript received November 29, 2005. Review conducted by B. A. Younis.

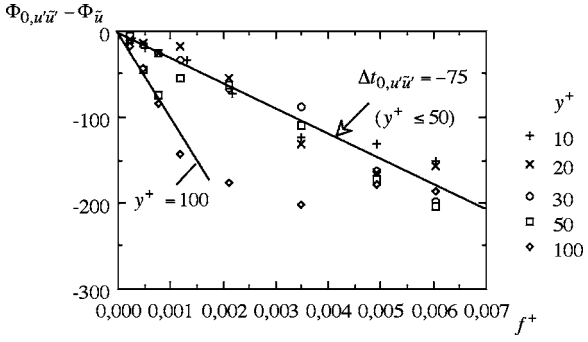


Fig. 1 Phase shift and corresponding time lag of the modulation of the longitudinal turbulent intensity depicted by Filter 0 at several wall normal positions in the inner layer

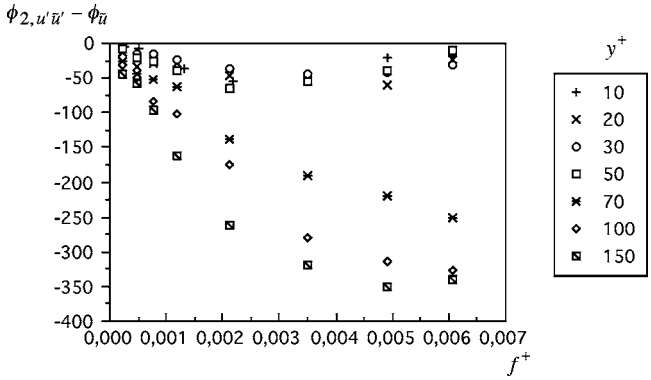


Fig. 3 Phase shift of the inner structures depicted by Filter 2 versus the frequency at different wall normal locations in the inner layer

way to scale it is to use the Stokes length l_s and relate the phenomena to the oscillating shear layer whose thickness is $2l_s$. Figure 2(*b*) shows the distribution of $y^+_{scr}=y^+/l^+_{scr}$ versus y^+. It is seen that $y^+_{scr}\approx 2.5$, which is only slightly larger than the oscillating shear thickness, except in the low buffer region. Consequently, the response time of the structures 1 is constant in the oscillating shear layer and decreases in the plug flow zone wherein $\partial\tilde{u}/\partial y \approx 0$.

The phase shift of the active structures 2 is small and remarkably independent of the imposed frequency in the viscous layer $y^+<50$ (Fig. 3). In the high logarithmic layer $\phi_{2,u'\tilde{u}'}-\phi_{\tilde{u}}$ is closely similar to $\phi_{1,u'\tilde{u}'}-\phi_{\tilde{u}}$ of the structures 1.

The relative amplitudes scaled with local $a_{\tilde{u}}$ are shown in Fig. 4

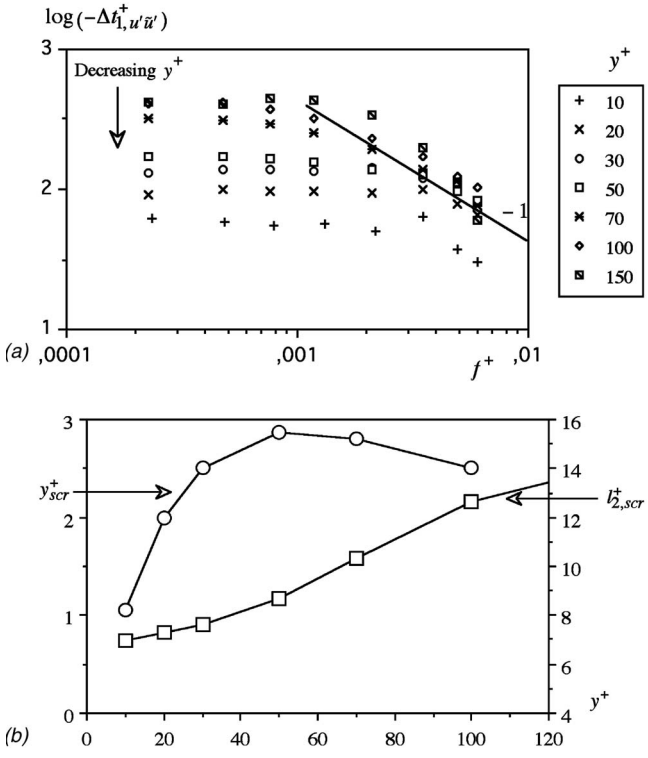


Fig. 2 Modulation characteristics of the inner structures depicted by Filter 1. (*a*) Time lag, (*b*) critical nondimensional frequency and wall normal distance beyond which the phase shift is constant (left) together with the critical Stokes length (right) related to the inner structures 2 (see the text).

versus the imposed frequency at four different y^+ positions. The first striking observation emerging from these results is the sharp decrease of $a_{2,u'\tilde{u}'}/2a_{\tilde{u}}$ from large values of about 1 in the low-frequency regime. The decrease is roughly linear from the quasi-steady limit to some $f^+_{2,cr}$ as shown by broken lines in Fig. 4. The resulting critical Stokes lengths $l^+_{2,scr}=\sqrt{\pi f^+_{2,cr}}$ are reported in Fig. 2(*b*) at the right. It is seen that $l^+_{2,scr}$ increases in the buffer layer until it reaches a plateau region in the log layer. The relative amplitude of the large scale $a_{0,u'\tilde{u}'}$ and inner structures $a_{1,u'\tilde{u}'}$ do not significantly differ from the global response $a_{u'\tilde{u}'}$ in the entire inner layer.

The response of the inner eddies are recapitulated in Fig. 5. The active eddies 2 are closely related to the quasi-streamwise vortices (QSV), which are the major coherent structures in the buffer layer. They contribute mostly to the low- speed streak formation and Reynolds shear stress in steady [4] and unsteady flows [2]. The peculiar behavior of the cutoff in the response of the active structures 2 to imposed unsteadiness can tentatively be explained by the reaction of the QSV to the oscillating shear $\partial\tilde{u}/\partial y$. The QSV are entirely embedded in the oscillating shear zone when $y^+=y^+_s l^+_s \approx 30$, which correspond to the top of the structures. Since $\partial\tilde{u}/\partial y$ is constrained into $y^+_s<2$, this condition implies a critical Stokes length of $l^+_{2,scr}\approx 15$, which corresponds well to the asymptotic limit in Fig. 2(*b*). The majority of the QSV's are in contact with the oscillating shear under this condition and the $\langle u'u'\rangle_2$ modulation extends to the log-layer. The active structures response is fast with small phase shifts $\phi_{2,u'\tilde{u}'}-\phi_{\tilde{u}}$ in this zone $(y^+<50)$ and $\langle u'u'\rangle_2$ diffuses away resulting in larger time lags (Fig. 3). Only smaller structures in their initial stage of development are directly affected by $\partial\tilde{u}/\partial y$ when the latter is constrained into $y^+<15$ as it is shown at left in Fig. 5. Thus, the oscillating shear has to be concentrated sufficiently close to the wall, to stimulate the unsteady reaction of these merely immature QSV and to activate the $\langle u'u'\rangle_2$ modulation in the low buffer layer. This explains the occurrence of a minimum at $l^+_{2,scr}\approx 7$ in Fig. 2(*b*).

The active structures 1 are larger scale eddies that extend beyond the buffer layer. Their time lag is constant in the oscillating shear zone and decreases in the plug layer. Figure 2 suggests that their size may reach 80–100 wall units. The inactive eddies are large-scale motions associated mainly with pressure-strain correlations and turbulent diffusion according to Bradshaw [6]. They do not contribute to the shear stress production. They consequently scale with integral variables in steady flow contrarily to the active eddies scaling with the inner wall variables. The outer

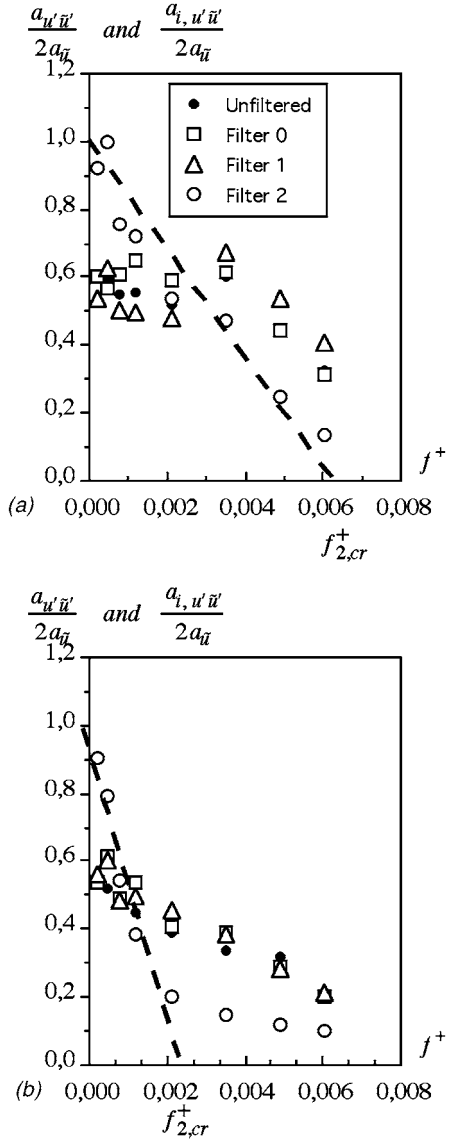


Fig. 4 Relative amplitude of the filtered and unfiltered streamwise fluctuating velocity at (*a*) $y^+=20$, (*b*) $y^+=100$

velocity scale in unsteady flows is the centerline velocity oscillations $\langle U_C \rangle$ and the inner velocity scale is the shear velocity modulation $\langle u_\tau \rangle$. Thus, outer and inner scaling would respectively imply $a_{u'\tilde{u}'} \approx 2a_{u\tilde{c}}$, $\phi_{u'\tilde{u}'} \approx \phi_{u\tilde{c}}$ for the passive structures and $a_{u'\tilde{u}'} \approx 2a_{u\tilde{\tau}} = a_{\tilde{\tau}}$, $\phi_{u'\tilde{u}'} = \phi_{\tilde{\tau}}$ for the inner ones. Neither the amplitudes $a_{i,u'\tilde{u}'}$ nor the phases $\phi_{i,u'\tilde{u}'}$ obey these relationships except in the quasi-steady regime. The structures respond to the imposed shear oscillations in a complex way, depending both on their characteristic scales and on the thickness of the oscillating shear zone in which they are embedded.

The results presented here may be useful in the development of multiple-scale modeling in unsteady flows [7]. Consider to this end the modulation of the kinetic energy equation

$$\frac{\partial \tilde{k}}{\partial t} = \tilde{P} - \tilde{\varepsilon} - \frac{\partial}{\partial y}\left[\frac{1}{\rho} v'\tilde{p}' + v'\tilde{k} - \nu\frac{\partial \tilde{k}}{\partial y}\right] \tag{1}$$

where $\tilde{P}$ and $\tilde{\varepsilon}$ are, respectively, the production and dissipation, the term under the bracket stands for the turbulent and viscous diffusion, and where the second-harmonic production is neglected. We may write $\tilde{k}=\tilde{k}_{\text{Outer}}+\tilde{k}_{\text{Inner}}$ with $\tilde{k}_{\text{Outer}}=\tilde{k}_0$, $\tilde{k}_{\text{Inner}}=\tilde{k}_1+\tilde{k}_2$ since

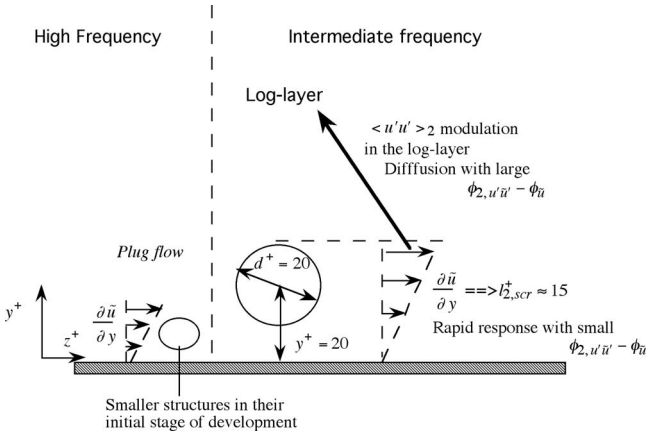


Fig. 5 Schematic representation of the frequency response of structures 2

these components fall into nonoverlapping parts of the spectrum. According to the results presented in this paper and the related discussion we made in the previous sections, the outer passive eddies contribute mainly to the turbulent diffusion. Thus,

$$\frac{\partial \tilde{k}_0}{\partial t} = -\frac{\partial}{\partial y}\left[\frac{1}{\rho} v'\tilde{p}' + v'\tilde{k}\right] \tag{2}$$

The classical multi-scale cascade equations are written as

$$\frac{\partial \tilde{k}_1}{\partial t} = \tilde{P}_1 - \tilde{\varepsilon}_1 + \nu\frac{\partial^2 \tilde{k}_1}{\partial y^2} \tag{3}$$

$$\frac{\partial \tilde{k}_2}{\partial t} = \tilde{P}_2 - \tilde{\varepsilon}_2 + \varepsilon_1 + \nu\frac{\partial^2 \tilde{k}_2}{\partial y^2} - \frac{\partial \tilde{k}_0}{\partial t} \tag{4}$$

where the coupling between the inner structures 1 and 2 takes place through the dissipation and the passive eddies intervene in the turbulent diffusion. The idea here is to couple these equations with the relationships based on the rapid distortion model in a way similar to Mankbadi and Liu [8] and Tardu and Da Costa [1] wherein substantial details can be found. The closure in these models is based on the effective strain parameter $\langle \alpha_{\text{eff}} \rangle$. The structural parameters such as the ratio of the Reynolds shear stress to the kinetic energy are related to $\langle \alpha_{\text{eff}} \rangle$ by $-\langle u'v' \rangle / \langle k \rangle = F(\langle \alpha_{\text{eff}} \rangle)$, where the function F is obtained by the bench data of the steady turbulent flow. The transport equation for $\langle \alpha_{\text{eff}} \rangle$ in its simplest form is

$$\frac{\partial \langle \alpha_{eff} \rangle}{\partial t} = -\frac{\langle \alpha_{eff} \rangle}{T_d(y)} + \frac{\partial \langle u \rangle}{\partial y} \tag{5}$$

where $T_d(y)$ is the rapid distortion time scale. We propose here to combine Eqs. (3) and (4) with each individual transport equation for the effective strain parameters, i.e.,

$$\frac{\partial \langle \alpha_{eff} \rangle_i}{\partial t} = -\frac{\langle \alpha_{eff} \rangle_i}{T_{di}(y)} + \frac{\partial \langle u \rangle}{\partial y} \tag{6}$$

where $i=0-2$ stands for the outer (0) and inner structures (1 and 2) and $T_{di}(y)$ are the corresponding rapid distortion time scales. The latter can easily be estimated from the time lags $\Delta t^+_{i,u'\tilde{u}'}$ investigated in this study. It is believed that this strategy can lead to more efficient modeling of the complex near-wall turbulent unsteady flows.

References

[1] Tardu, S.-F., and Da Costa, P., 2004, "Experiments and Modeling of an Unsteady Turbulent Channel Flow," AIAA J., **43**(1), pp. 140–149.

[2] Tardu, S.-F., and Vezin, P., 2004, "Response of the Streaks, Active, and Passive Eddies in an Unsteady Turbulent Channel Flow," Int. J. Heat Fluid Flow, **25**(6), pp. 925–932.
[3] Tardu, S.-F., Binder, G., and Blackwelder, R.-F., 1994, "Turbulent Channel Flow With Large Amplitude Velocity Oscillations," J. Fluid Mech., **267**, pp. 109–151.
[4] Naguib, A.-M., and Wark, C.-E., 1992, "An Investigation of Wall-Layer Dynamics Using a Combined Temporal Filtering and Correlation Technique," J. Fluid Mech., **243**, pp. 541–560.
[5] Eckelmann, H., 1974, "The Structure of the Viscous Sublayer and the Adjacent Wall Region in a Turbulent Channel Flow," J. Fluid Mech., **65**, pp. 439–459.
[6] Bradshaw, P., 1967, "Inactive Motion and Pressure Fluctuations in Turbulent Boundary Layers," J. Fluid Mech., **30**, pp. 241–258.
[7] Launder, B.-E., 1984, "Second-Moment Closure: Methodology and Practice," in *Turbulence Models and Their Applications*, Editions Eyrolles, Paris, pp. 1–143.
[8] Mankbadi, R.-R., and Liu, J.-T.-L., 1992, "Near Wall Response in Turbulent Shear Flows Subjected to Imposed Unsteadiness," J. Fluid Mech., **238**, pp. 55–71.