ASME
SETTING THE STANDARD

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.
STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or printed in its publications (B7.1, Para. 3).

Journal of Applied Mechanics

Published Bimonthly by ASME

VOLUME 76 • NUMBER 4 • JULY 2009

RESEARCH PAPERS

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

(**Contents continued**)

Journal of Applied Mechanics **Volume 76, Number 4** **JULY 2009**

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

Steady-State Wave Propagation in Multilayered Viscoelastic Media Excited by a Moving Dynamic Distributed Load

Lu Sun

Wenjun Gu

Transportation College,
Southeast University,
Nanjing 210096, China;
Department of Civil Engineering,
Catholic University of America,
Washington, DC 20064

Feiquan Luo

Department of Civil Engineering,
Catholic University of America,
Washington, DC 20064

An analytical solution of steady-state dynamic response of a multilayered viscoelastic medium to a moving distributed load is obtained using a novel approach that combines transfer matrix method with Sun's convolution representation integrated over impulse response function of the layered medium. The layered media under consideration include elastic and viscoelastic media with four different viscoelastic constitutive models, while the moving load is allowed to have a circular spatial distribution, which is more realistic for mimicking tire footprint than a commonly used point load. Efficient numerical algorithms based on fast evaluation of various integral transformations and their inversions are developed and validated through numerical example. [DOI: 10.1115/1.3086586]

Keywords: multilayered medium, viscoelastic, steady-state response, wave equation, integral transform, convolution

1 Introduction

Nowadays highway and airport pavement design is based on static analysis of multilayered elastic media [1]. With the promotion and deployment of high-speed surface transportation, dynamic effects of moving vehicular load on transportation infrastructure become much more significant than ever before [2–8]. To develop a more accurate structure design method for highway, airport, and bridges, dynamics of pavements needs to be investigated thoroughly. A multilayered elastic medium is a commonly used physical model for pavements. The study of dynamic responses of a multilayered medium has been a subject of great interest since 1950s [9]. Thomson [10] and Haskell [11] developed a transfer matrix method to study the propagation of waves in layered elastic media. Thomson–Haskell transfer matrix method is numerically deficient for thicker layers and/or higher frequencies. Kausel and Peek [12] extended the transfer matrix method by proposing a stiffness matrix method, which uses second order approximation to replace transcendent functions in the solution. Due to the adoption of second order approximation, the stiffness matrix method is able to evaluate the integral analytically in the frequency wavenumber domain, allowing more efficient numerical computation. Kennett [13] proposed a numerically more robust method.

Based on Biot's formulation for a multilayered porous medium [14,15], applied Fourier transform to analyze dynamic response of a saturated soil deposited to a moving load [16,17] investigated a similar problem of dynamic response of a poroelastic soil medium under a moving load. The solution is approximately estimated by the superposition of an elastodynamic problem with modified elastic constants and mass density for the whole domain and a diffusion problem for the pore fluid pressure confined to a boundary layer near the free surface of the medium. De Barros and Luco [18] evaluated dynamic displacements of a multilayered medium to a moving line load at a constant velocity using Fourier synthesis of the frequency response. Sun [2] applied integral transform to study the general theory of deterministic and stochastic moving load problem for one-, two-, and three-dimensional applications. Some fundamental results of moving load problem were given by Sun and Greenberg [3]. Lefeuve-Mesgouez et al. [19] took advantage of Helmholtz decomposition and fast Fourier transform and investigated the transmission of ground vibration due to a moving harmonic strip load rigidly attached to the surface of an elastic half-space. A review is given in Ref. [20] on methodologies for moving load induced vibration in half-spaces.

Pak and Guzina [21] studied three-dimensional wave propagation problem of a vertically-heterogeneous elastic half-space with the aid of a displacement-potential representation, Hankel transforms, and Fourier decompositions, and determined the dynamic response of the semi-infinite solid to an arbitrarily distributed buried source. Ditri [22] determined nonuniform stresses in an isotropic half-space from measurements of the dispersion of surface waves. Dieterman and Metrikine [23] studied critical velocities of a harmonic load moving uniformly along an elastic layer. These studies are confined to elastic multilayered media only. Recently, Zhao and Gary [24] presented an original analytical solution of the longitudinal wave propagation in an infinite linear viscoelastic cylindrical bar and its applications to experimental methods of material behavior testing. Benatar et al. [25] conducted theoretical and experimental analysis of longitudinal wave propagation in cylindrical viscoelastic rods. Christensen [26] also studied discussions of some problems concerning loads moving on viscoelastic half-spaces.

Finite element method (FEM) has also been used in the dynamic analysis of layered medium to moving loads [20]. Chatti et al. [27] considered a multilayered elastic medium subjected to dynamic loads for applications in pavement design using finite element analysis. The cumbersome of FEM in this kind of application is twofold. Because the load is moving, one has to remesh the finite element, which is a time consuming process. In addition, no matter how large the medium of investigation is set up in FEM, the moving load will always reach the artificial boundary, causing significant reduction in computational accuracy. Noticing these two shortcomings, Adam et al. [28] adopted both FEM and boundary element method (BEM) for a two-dimensional moving load problem. Sun [2] proposed another alternative for overcoming existing shortcomings of FEM by incorporating nonuniform mess FEM with Sun's convolution representation. More recently, Koh

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 26, 2007; final manuscript received January 26, 2009; published online April 21, 2009. Review conducted by Anthony Waas.

et al. [29] suggested a moving element method for dealing with a two-dimensional moving load problem. A review on numerical methods, primarily, the FEM and BEM, is given by Andersen et al. [30].

The benefit of using FEM is that it can deal with complex geometry and constitutive models of the material. However, for wave propagation in an infinite medium caused by moving loads, the advantages of analytical or semi-analytical methods over FEM are obvious. Analytical and/or semi-analytical methods provide a more explicit representation and physical interpretation of the subject. When extended to more complicated loading scenarios, such as stochastic moving loading with random material properties, (stochastic) FEM is less functional as compared with analytical methods [2]. Furthermore, computational efficiency of analytical and/or semi-analytical methods is much higher than FEM.

In this paper we developed an analytical treatment for studying steady-state responses of a multilayered viscoelastic medium to a moving distributed load. Contributions of this paper are fourfold. First, it proposes a novel approach that combines transfer matrix method with Sun's convolution representation to allow moving load problem to be investigated analytically and computed efficiently. Second, it extends a layered elastic medium to a layered viscoelastic medium with material properties specified by four different viscoelastic models. The consideration of viscosity is crucial for characterization of asphalt concrete pavements and may account for permanent deformation of pavement such as rutting and roughness. Third, in existing studies of moving load problem using analytical methods, the moving load is often assumed to be a point load. In this paper the moving load has a circular spatial distribution, which is more realistic for mimicking tire footprint. Lastly, efficient numerical algorithms based on fast evaluation of various integral transformations and their inversions are developed, presented, and validated in a computational example.

The remainder of this paper is organized as follows. Section 2 briefly introduces Sun's convolution representation for moving source problem in linear systems. Section 3 presents the governing Navier's equation of multilayered viscoelastic media. Section 4 utilizes Fourier transform, Hankel transform, and Laplace transform along with transfer matrix method to solve for impulse response function (IRF) of the multilayered elastic and viscoelastic media. In the derivation of viscoelastic IRF, the elastic-viscoelastic correspondence principle is utilized. The obtained IRFs are explicit solutions of Navier's equation with impulse distributed loading as boundary conditions, which allow efficient numerical algorithms to be carried out. Section 5 incorporates the IRF with Sun's convolution representation to obtain dynamic response of the medium to a variety of moving loads. Section 6 validates the dynamic response of a four-layer elastic medium against known results in the literature, followed by numerical results for steady-state response of a four-layer viscoelastic medium. Section 7 makes concluding remarks.

2 Convolution Representation

The problem under consideration in this paper is the following. In a Cartesian coordinate system $\mathbf{x}=(x,y,z)$, a multilayered linear viscoelastic medium is subjected to a distributed dynamic load, $F(\mathbf{x},t)$, moving on the surface layer of the medium

$$F(\mathbf{x},t)=\frac{H[r_0^2-(x-vt)^2-y^2]}{\pi r_0^2}\delta(z)p(t) \tag{1}$$

where $p(\cdot)$ describes varying amplitudes of the dynamic load, r_0 and v are the radius and the velocity of the moving distributed load, and t is the time variable. In addition, $H(\cdot)$ represents the Heaviside unit step function and $\delta(\cdot)$ represents the Dirac delta function. These are defined, respectively, by

$$H(t-t_0)=\begin{cases}0 & \text{if } t<t_0\\ \frac{1}{2} & \text{if } t=t_0\\ 1 & \text{if } t>t_0\end{cases} \tag{2}$$

and

$$\int_{-\infty}^{\infty}f(x)\delta(x-x_0)dx=f(x_0)\quad\text{and}\quad\int_{-\infty}^{\infty}\delta(x-x_0)dx=1 \tag{3}$$

Here, we are particularly interested in studying dynamic response of this medium caused by such a moving load. A typical method for dealing with a steady-state moving load problem is to apply the Galilean transforms $x'=x-vt$, $y'=y$, and $z'=z$, such that the problem can be reformulated in the transformed moving coordinate system $\mathbf{x}'=(x',y',z')$. This method can be ineffective for transient moving load problem or less efficient for steady-state problem as it involves tedious derivations of the formulas. Since the multilayered viscoelastic medium is a linear system, a more effective method is to formulate the steady-state moving load problem using Sun's convolution representation [2,3]

$$\mathbf{u}(\mathbf{x},t)=\int_0^{\infty}p(t-\vartheta)\mathbf{h}(x-vt+v\vartheta,y,z,\vartheta)d\vartheta \tag{4}$$

where ϑ is an argument variable, $\mathbf{u}(\mathbf{x},t)$ represents the displacement field of the medium, and $\mathbf{h}(\mathbf{x},t)$ is a kernel function or an IRF of the medium, which corresponds to the displacement field of the medium subject to a position-fixed impulse load expressed by

$$F_\delta(\mathbf{x},t)=\frac{H(r_0^2-x^2-y^2)}{\pi r_0^2}\delta(z)\delta(t) \tag{5}$$

The derivation of Sun's convolution representation (4) is based on the superposition principle of a linear system or the Betti–Rayleigh reciprocal theorem in elastodynamics [3,31–36], which is also applicable to buried moving loads. The IRF plays a similar role as the Green's function, except that the former takes the spatial distribution of the load into consideration, while the latter only corresponds to a point impulse load. Advantages of formulating the solution of a moving load problem in terms of Eq. (4) are twofold. First it provides a concise, unified, physical interpretation and decomposition of the dynamics of the medium. Second the dynamic response of the medium subject to a moving load is constructed from the dynamic response of the medium subject to a position-fixed impulse load with axisymmetric spatial distribution as given in Eq. (5), allowing the adoption and utilization of a cylindrical coordinate system (r,θ,Z). This further simplifies the derivation of the solution by reducing the dimension of the problem under consideration from four variables (x,y,z,t) to three variables (r,z,t) because the medium is isotropic and variable θ plays no role in the governing equation for an axisymmetric problem. Therefore, displacement field $\mathbf{u}(\mathbf{x},t)$ on the left-hand-side of Eq. (4) depends on the kernel function or the IRF of the medium, $\mathbf{h}(\cdot)$, which is addressed in Secs. 3–6.

3 Governing Equations

Figure 1 shows a multilayered medium under a moving circular load. For an axisymmetrically distributed impulse load (5), it is convenient to represent the governing equation in a global cylindrical coordinate system (r,θ,Z). In addition to the global cylindrical coordinate system (r,θ,Z), local cylindrical coordinate systems (r_i,θ_i,z_i) are also employed for the ith layer. The relationship between global and local coordinate systems is defined by $r_i=r$, $\theta_i=\theta$, and $z_i=Z-h_i$, in which h_i is the vertical distance from the surface of the multilayered medium to the upper interface of the ith layer. Unless otherwise stated, the subscript i is omitted in the following derivation for the sake of simplicity.

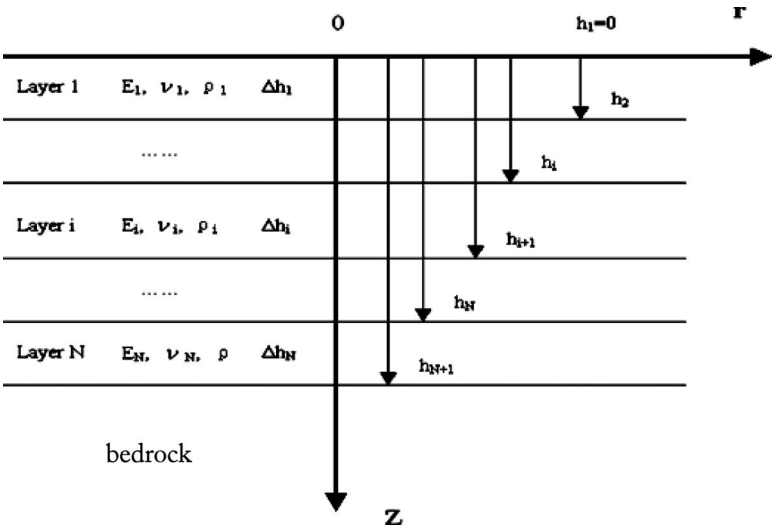


Fig. 1 A multilayered medium resting on bedrock

Motions of a multilayered medium for the ith layer are governed by two wave equations in a cylindrical coordinate system [31–36]

$$(c_d^2 - c_s^2)\frac{\partial \Delta}{\partial r} + c_s^2\left(\nabla^2 u - \frac{u}{r^2}\right) - \ddot{u} = 0 \tag{6a}$$

$$(c_d^2 - c_s^2)\frac{\partial \Delta}{\partial z} + c_s^2 \nabla^2 w - \ddot{w} = 0 \tag{6b}$$

where $u=u(r,z,t)$ and $w=w(r,z,t)$ are displacements of the ith layer along the radial direction r and depth direction z, respectively, and dots in Eqs. (6a) and (6b) stand for differentiation with respect to time t (one and two dots stand for first and second derivatives, respectively). Also, $c_d=\sqrt{(\lambda+2\mu)/\rho}$ and $c_s=\sqrt{\mu/\rho}$ are dilatational and shear wave velocities, respectively. Here λ and μ are two Lame's constants and can be expressed using Young's elastic modulus E and Poisson's ratio ν.

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \tag{7a}$$

$$\mu = \frac{E}{2(1+\nu)} \tag{7b}$$

For linear elastic solids, constitutive equations read

$$\tau_{rz} = \mu\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial r}\right) \tag{8a}$$

$$\sigma_z = \lambda\Delta + 2\mu\frac{\partial w}{\partial z} \tag{8b}$$

where $\Delta=(\partial u/\partial r)+(u/r)+(\partial w/\partial z)$. Realistically, boundary conditions of a multilayer medium as well as a single ith layer should be established at the upper and lower boundaries. For the time being, it is assumed that four boundary conditions are all established at $z=0$ in a local coordinate (the purpose is to simplify the assembly process for multiple layers)

$$\boldsymbol{\tau}_{rz}(r,z,t) = \boldsymbol{\tau}_{rz}(r,z,t)|_{z=0} \quad (r,z) \in S_\sigma \tag{9a}$$

$$\boldsymbol{\sigma}_z(r,z,t) = \boldsymbol{\sigma}_z(r,z,t)|_{z=0} \quad (r,z) \in S_\sigma \tag{9b}$$

$$u(r,z,t) = u(r,z,t)|_{z=0} \quad (r,z) \in s_u \tag{9c}$$

$$w(r,z,t) = w(r,z,t)|_{z=0} \quad (r,z) \in s_u \tag{9d}$$

Here, s_σ and s_u stand for domains where stress and displacement fields are defined. For the steady-state problem under investigation, it is also assumed that the medium be at rest at $t=-\infty$ and no initial stress exists in the medium. That is,

$$u(r,z,-\infty) = \dot{u}(r,z,-\infty) = w(r,z,-\infty) = \dot{w}(r,z,-\infty) = 0 \tag{10}$$

For linear viscoelastic media studied in this paper, their properties can be characterized by mechanical models consisting of springs and dashpots. Figure 2 shows four types of viscoelastic models commonly used for characterizing viscosity in highway pavements: Kelvin model, Maxwell model, Burgers model, and the generalized viscoelastic model [1]. The generalized viscoelastic model contains three other models as special cases. In Fig. 2, σ is the stress, $E_0, E_1, \ldots, E_n$ are Young's elastic moduli corresponding to spring constants, and $T_0, T_1, \ldots, T_n$ are retardation times corresponding to dashpot dampings $\lambda_0, \lambda_1, \ldots, \lambda_n$ via the relationship $T_j=\lambda_j/E_j$. A specific viscoelastic model defines an equivalent modulus E_e through stress-strain relationship $E_e=\sigma/\varepsilon$.

A Kelvin model is a combination of spring and dashpot in parallel, as depicted in Fig. 2(a). Both the spring and the dashpot experience the same strain ε, but the total stress is the sum of the two stresses

$$\sigma = E_1\varepsilon + \lambda_1\frac{\partial \varepsilon}{\partial t} = E_1\left(\varepsilon + T_1\frac{\partial \varepsilon}{\partial t}\right) \tag{11}$$

where retardation time $T_1=\lambda_1/E_1$ and differential operator $D=\partial/\partial t$. Equation (11) can be expressed as

$$\sigma = E_e\varepsilon \tag{12}$$

where equivalent modulus $E_e=E_1+\lambda_1 D=E_1(1+T_1 D)$.

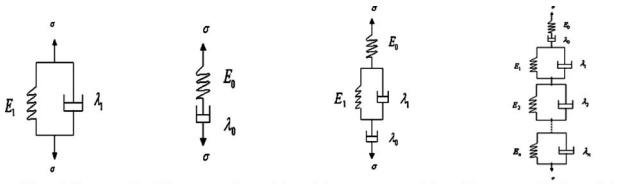


Fig. 2 A schematic plot of four viscoelastic models

A Maxwell model is a combination of spring and dashpot in series, as indicated in Fig. 2(*b*). Both the spring and the dashpot experience the same stress, but the total strain is the sum of two strains

$$\varepsilon = \frac{\sigma}{E_0} + \frac{\sigma}{\lambda_0 \frac{\partial}{\partial t}} = \frac{\sigma}{E_0} + \frac{\sigma}{\lambda_0 D} = \frac{\sigma}{E_0} + \frac{\sigma}{T_0 E_0 D} \tag{13}$$

where retardation time $T_0 = \lambda_0 / E_0$. The equivalent elastic modulus derived from Eq. (13) is

$$E_e = \frac{\sigma}{\varepsilon} = \frac{E_0 T_0 D}{T_0 D + 1} \tag{14}$$

A Burgers model is a combination of Maxwell and Kelvin models in series, as indicated in Fig. 2(*c*). The total strain is composed of three parts:

$$\varepsilon = \frac{\sigma}{E_0} + \frac{\sigma}{E_1 + \lambda_1 \frac{\partial}{\partial t}} + \frac{\sigma}{\lambda_0 \frac{\partial}{\partial t}} = \sigma \left(\frac{T_0 D + 1}{E_0 T_0 D} + \frac{1}{E_1 (T_1 D + 1)} \right) \tag{15}$$

The equivalent elastic modulus derived from Eq. (15) is

$$E_e = \frac{\sigma}{\varepsilon} = \left[\frac{T_0 D + 1}{E_0 T_0 D} + \frac{1}{E_1 (T_1 D + 1)} \right]^{-1} \tag{16}$$

Figure 2(*d*) shows a generalized model that can be used to characterize a wide variety of viscoelastic material. It is a combination of Maxwell and n Kelvin models in series. The total strain is composed of $(n+1)$ parts:

$$\varepsilon = \sigma \frac{T_0 D + 1}{E_0 T_0 D} + \sigma \sum_{i=1}^{n} \frac{1}{E_i (T_i D + 1)} \tag{17}$$

The equivalent elastic modulus derived from Eq. (17) is

$$E_e = \frac{\sigma}{\varepsilon} = \left[\frac{T_0 D + 1}{E_0 T_0 D} + \sum_{i=1}^{n} \frac{1}{E_i (T_i D + 1)} \right]^{-1} \tag{18}$$

4 Impulse Response Functions

4.1 A Single Layer of the Elastic Medium. The solution of elastic media is the basis for that of viscoelastic medium. Therefore, the former is sought first in this subsection, and followed by the latter in Sec. 4.2. Wave equations (6*a*) and (6*b*) are linear partial differential equations, for which integral transformation is an effective method. Since we are interested in the steady-state response of a multilayered medium, we apply Fourier transform with respect to time t and Hankel transform with respect to radius r. Unless otherwise stated, the treatment in this section is with respect to the ith layer only.

We define Fourier transform and its inversion as

$$\hat{f}(q) = \int_{-\infty}^{\infty} f(t) e^{-iqt} dt \tag{19a}$$

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(q) e^{iqt} dq \tag{19b}$$

Applying Fourier transform with respect to time t in Eqs. (6*a*), (6*b*), (8*a*), and (8*b*) leads to

$$(c_d^2 - c_s^2) \frac{\partial \hat{\Delta}}{\partial r} + c_s^2 \left(\nabla^2 \hat{u} - \frac{\hat{u}}{r^2} \right) + q^2 \hat{u} = 0 \tag{20a}$$

$$(c_d^2 - c_s^2) \frac{\partial \hat{\Delta}}{\partial z} + c_s^2 \nabla^2 \hat{w} + q^2 \hat{w} = 0 \tag{20b}$$

$$\hat{\tau}_{rz} = \mu \left(\frac{\partial \hat{u}}{\partial z} + \frac{\partial \hat{w}}{\partial r} \right) \tag{21a}$$

$$\hat{\sigma}_z = \lambda \hat{\Delta} + 2\mu \frac{\partial \hat{w}}{\partial z} \tag{21b}$$

$$\hat{\tau}_{rz}(r,z,q) = \hat{\tau}_{rz}(r,z_b,q) \quad (r,z) \in s_{\boldsymbol{\sigma}} \tag{22a}$$

$$\hat{\sigma}_z(r,z,q) = \hat{\sigma}_z(r,z_b,q) \quad (r,z) \in s_{\boldsymbol{\sigma}} \tag{22b}$$

$$\hat{u}(r,z,q) = \hat{u}(r,z_b,q) \quad (r,z) \in s_u \tag{22c}$$

$$\hat{w}(r,z,q) = \hat{w}(r,z_b,q) \quad (r,z) \in s_u \tag{22d}$$

Define the νth order Hankel transform of $f(r)$ and its inversion as follows:

$$\bar{f}(\xi) = \int_0^{\infty} r f(r) J_\nu(\xi r) dr \tag{23a}$$

$$f(r) = \int_0^{\infty} \xi \bar{f}(\xi) J_\nu(\xi r) d\xi \tag{23b}$$

where J_ν is the Bessel function of the first kind of order ν, and ξ is the radial wave-number corresponding to depth z. Denote $\bar{\hat{u}}(\xi,z,q)$, $\bar{\hat{w}}(\xi,z,q)$, $\bar{\hat{\tau}}_{rz}(\xi,z,q)$, and $\bar{\hat{\sigma}}_z(\xi,z,q)$ the displacement and stress fields after the Fourier and Hankel transforms.

$$\bar{\hat{u}}(\boldsymbol{\xi},z,q) = \int_0^{\infty} r \hat{u}(r,z,q) J_1(\boldsymbol{\xi} r) dr \tag{24a}$$

$$\bar{\hat{w}}(\boldsymbol{\xi},z,q) = \int_0^{\infty} r \hat{w}(r,z,q) J_0(\boldsymbol{\xi} r) dr \tag{24b}$$

$$\bar{\hat{\boldsymbol{\tau}}}_{rz}(\boldsymbol{\xi},z,q) = \int_0^{\infty} r \hat{\boldsymbol{\tau}}_{rz}(r,z,q) J_1(\boldsymbol{\xi} r) dr \tag{24c}$$

$$\bar{\hat{\boldsymbol{\sigma}}}_z(\boldsymbol{\xi},z,q) = \int_0^{\infty} r \hat{\boldsymbol{\sigma}}_z(r,z,q) J_0(\boldsymbol{\xi} r) dr \tag{24d}$$

where the caret variables in Eqs. (24*a*)–(24*d*) have been defined in Eqs. (22*a*)–(22*d*).

Apply Hankel transform of order 1 with respect to radius r in Eqs. (20*a*), (21*a*), (22*a*), and (22*c*). Apply Hankel transform of order 0 with respect to radius r in Eqs. (20*b*), (21*b*), (22*b*), and (22*d*). After these transformations, we have

$$c_s^2 \frac{d^2 \bar{\hat{u}}}{dz^2} - (c_d^2 - c_s^2) \xi \frac{d\bar{\hat{w}}}{dz} + (q^2 - c_d^2 \xi^2) \bar{\hat{u}} = 0 \tag{25a}$$

$$c_d^2 \frac{d^2 \bar{\hat{w}}}{dz^2} + (c_d^2 - c_s^2) \xi \frac{d\bar{\hat{u}}}{dz} + (q^2 - c_s^2 \xi^2) \bar{\hat{w}} = 0 \tag{25b}$$

$$\bar{\hat{\tau}}_{rz} = \mu \left(\frac{d\bar{\hat{u}}}{dz} - \xi \bar{\hat{w}} \right) \tag{26a}$$

$$\bar{\hat{\sigma}}_z = \lambda \xi \bar{\hat{u}} + (\lambda + 2\mu) \frac{d\bar{\hat{w}}}{dz} \tag{26b}$$

with boundary conditions

$$\bar{\hat{\tau}}_{rz}(\xi,z,q) = \bar{\hat{\tau}}_{rz}(\xi,z_b,q) \quad (\xi,z) \in S_\sigma \tag{27a}$$

$$\bar{\hat{\sigma}}_z(\xi,z,q) = \bar{\hat{\sigma}}_z(\xi,z_b,q) \quad (\xi,z) \in S_\sigma \tag{27b}$$

$$\bar{\hat{u}}(\xi,z,q) = \bar{\hat{u}}(\xi,z_b,q) \quad (\xi,z) \in S_u \tag{27c}$$

$$\bar{\hat{w}}(\xi,z,q) = \bar{\hat{w}}(\xi,z_b,q) \quad (\xi,z) \in S_u \tag{27d}$$

It should be noted that Eqs. (25*a*) and (25*b*) are now a set of linear ordinary differential equations (ODEs) with only one independent variable z, for which analytical solution is available.

Multiplying Eq. (25*a*) by $(c_d^2-c_s^2)\xi$ leads to

$$c_s^2(c_d^2-c_s^2)\xi\frac{d^2\bar{\hat{u}}}{dz^2} - (c_d^2-c_s^2)^2\xi^2\frac{d\bar{\hat{w}}}{dz} + (c_d^2-c_s^2)\xi(q^2-c_d^2\xi^2)\bar{\hat{u}} = 0 \tag{28}$$

Differentiating Eq. (25*b*) with respect to z gives

$$c_d^2\frac{d^3\bar{\hat{w}}}{dz^3} + (c_d^2-c_s^2)\xi\frac{d^2\bar{\hat{u}}}{dz^2} + (q^2-c_s^2\xi^2)\frac{d\bar{\hat{w}}}{dz} = 0 \tag{29}$$

Multiplying Eq. (29) by c_s^2 leads to

$$c_s^2c_d^2\frac{d^3\bar{\hat{w}}}{dz^3} + c_s^2(c_d^2-c_s^2)\xi\frac{d^2\bar{\hat{u}}}{dz^2} + c_s^2(q^2-c_s^2\xi^2)\frac{d\bar{\hat{w}}}{dz} = 0 \tag{30}$$

Equation (30) minus Eq. (28) leads to

$$(c_d^2-c_s^2)\xi(q^2-c_d^2\xi^2)\bar{\hat{u}} = c_s^2c_d^2\frac{d^3\bar{\hat{w}}}{dz^3} + [c_s^2(q^2-c_s^2\xi^2) + (c_d^2-c_s^2)^2\xi^2]\frac{d\bar{\hat{w}}}{dz} \tag{31}$$

Differentiating Eq. (31) with respect to z leads to

$$(c_d^2-c_s^2)\boldsymbol{\xi}(q^2-c_d^2\boldsymbol{\xi}^2)\frac{d\bar{\hat{u}}}{dz} = c_s^2c_d^2\frac{d^4\bar{\hat{w}}}{dz^4} + [c_s^2(q^2-c_s^2\boldsymbol{\xi}^2) + (c_d^2-c_s^2)^2\boldsymbol{\xi}^2]\frac{d^2\bar{\hat{w}}}{dz^2} \tag{32}$$

Multiplying Eq. (25*b*) by $(q^2-c_d^2\xi^2)$ leads to

$$c_d^2(q^2-c_d^2\xi^2)\frac{d^2\bar{\hat{w}}}{dz^2} + (c_d^2-c_s^2)\xi(q^2-c_d^2\xi^2)\frac{d\bar{\hat{u}}}{dz} + (q^2-c_s^2\xi^2)(q^2-c_d^2\xi^2)\bar{\hat{w}} = 0 \tag{33}$$

Substituting Eq. (32) into Eq. (33) and performing some algebraic manipulation leads to

$$\frac{d^4\bar{\hat{w}}}{dz^4} + \left[\left(\frac{q^2}{c_s^2}-\boldsymbol{\xi}^2\right) + \left(\frac{q^2}{c_d^2}-\boldsymbol{\xi}^2\right)\right]\frac{d^2\bar{\hat{w}}}{dz^2} + \left(\frac{q^2}{c_s^2}-\boldsymbol{\xi}^2\right)\left(\frac{q^2}{c_d^2}-\boldsymbol{\xi}^2\right)\bar{\hat{w}} = 0 \tag{34}$$

Denoting $A=\boldsymbol{\xi}^2-(q^2/c_s^2)$ and $B=\boldsymbol{\xi}^2-(q^2/c_d^2)$ leads to

$$\frac{d^4\bar{\hat{w}}}{dz^4} + (A+B)\frac{d^2\bar{\hat{w}}}{dz^2} + AB\bar{\hat{w}} = 0 \tag{35}$$

The auxiliary equation of Eq. (35) is

$$r^4 - (A+B)r^2 + AB = 0 \text{ or equivalently } (r^2-A)(r^2-B) = 0 \tag{36}$$

Suppose that four roots of Eq. (36) are r_1, r_2, r_3, and r_4. According to the general solution of linear ODE, the solution of Eq. (35) can be expressed as

Table 1 Four scenarios of four roots of auxiliary Eq. (36)

Scenarios	$r_{1,2}$	$r_{3,4}$	Solution of Eq. (34)
$0>B>A$	$\pm i\sqrt{-A}$	$\pm i\sqrt{-B}$	$\bar{\hat{w}}=c_1e^{i\sqrt{-A}z}+c_2e^{-i\sqrt{-A}z}+c_3e^{i\sqrt{-B}z}+c_4e^{-i\sqrt{-B}z}$
$B>0>A$	$\pm i\sqrt{-A}$	$\pm\sqrt{B}$	$\bar{\hat{w}}=c_1e^{i\sqrt{-A}z}+c_2e^{-i\sqrt{-A}z}+c_3e^{\sqrt{B}z}+c_4e^{-\sqrt{B}z}$
$B>A>0$	$\pm\sqrt{A}$	$\pm\sqrt{B}$	$\bar{\hat{w}}=c_1e^{\sqrt{A}z}+c_2e^{-\sqrt{A}z}+c_3e^{\sqrt{B}z}+c_4e^{-\sqrt{B}z}$

$$\bar{\hat{w}} = c_1e^{r_1z} + c_2e^{r_2z} + c_3e^{r_3z} + c_4e^{r_4z} \tag{37}$$

where c_1, c_2, c_3, and c_4 are some unknown functions of c_s, c_d, ξ, and q and are to be determined from boundary conditions. Since $c_d>c_s$, it follows that $A<B$ and therefore three scenarios exist for these four roots of Eq. (36), which are listed in Table 1.

Using Euler formula $e^{i\Theta}=\cos\Theta+i\sin\Theta$, the solution of Eq. (34) for scenario $0>B>A$ can also be written as

$$\bar{\hat{w}} = C_1\cos(\sqrt{-A}z) + C_2\sin(\sqrt{-A}z) + C_3\cos(\sqrt{-B}z) + C_4\sin(\sqrt{-B}z) \tag{38a}$$

where $C_1=c_1+c_2$, $C_2=(c_1-c_2)i$, $C_3=c_3+c_4$, and $C_4=(c_3-c_4)i$. The solution of Eq. (34) for scenario $0>B>A$ can also be written as

$$\bar{\hat{w}} = C_1\cos(\sqrt{-A}z) + C_2\sin(\sqrt{-A}z) + C_3\cos(i\sqrt{B}z) + C_4\sin(i\sqrt{B}z) \tag{38b}$$

where $C_1=c_1+c_2$, $C_2=(c_1-c_2)i$, $C_3=c_3+c_4$, and $C_4=(c_4-c_3)i$. Similarly, the solution of Eq. (34) for scenario $B>A>0$ can also be written as

$$\bar{\hat{w}} = C_1\cos(i\sqrt{A}z) + C_2\sin(i\sqrt{A}z) + C_3\cos(i\sqrt{B}z) + C_4\sin(i\sqrt{B}z) \tag{38c}$$

where $C_1=c_1+c_2$, $C_2=(c_2-c_1)i$, $C_3=c_3+c_4$, and $C_4=(c_4-c_3)i$. Define $\alpha=\sqrt{(q^2/c_s^2)-\boldsymbol{\xi}^2}$ and $\beta=\sqrt{(q^2/c_d^2)-\boldsymbol{\xi}^2}$. Clearly, if $A<0$ then $\alpha=\sqrt{-A}$, and if $A>0$ then $\alpha=i\sqrt{A}$; if $B<0$ then $\beta=\sqrt{-B}$, and if $B>0$ then $\beta=i\sqrt{B}$. Given this notation Eqs. (38*a*)–(38*c*) can be written in the same concise form as

$$\bar{\hat{w}} = C_1\cos(\alpha z) + C_2\sin(\alpha z) + C_3\cos(\beta z) + C_4\sin(\beta z) \tag{39}$$

It should be noted that C_1, C_2, C_3, C_4, α, and β in Eq. (39) are defined in complex number domain. In the following, we are going to determine C_1, C_2, C_3, and C_4 from boundary conditions.

Differentiating Eq. (39) with respect to z leads to

$$\frac{d\bar{\hat{w}}}{dz} = -C_1\alpha\sin(\alpha z) + C_2\alpha\cos(\alpha z) - C_3\beta\sin(\beta z) + C_4\beta\cos(\beta z) \tag{40}$$

Differentiating Eq. (40) with respect to z leads to

$$\frac{d^2\bar{\hat{w}}}{dz^2} = -C_1\alpha^2\cos(\alpha z) - C_2\alpha^2\sin(\alpha z) - C_3\beta^2\cos(\beta z) - C_4\beta^2\sin(\beta z) \tag{41}$$

Differentiating Eq. (41) with respect to z leads to

$$\frac{d^3\bar{\hat{w}}}{dz^3} = C_1\alpha^3\sin(\alpha z) - C_2\alpha^3\cos(\alpha z) + C_3\beta^3\sin(\beta z) - C_4\beta^3\cos(\beta z) \tag{42}$$

Substituting Eqs. (40) and (42) into Eq. (31) leads to

$$\bar{\tilde{u}} = \frac{c_s^2 c_d^2 (C_1\alpha^3 \sin(\alpha z) - C_2\alpha^3 \cos(\alpha z) + C_3\beta^3 \sin(\beta z) - C_4\beta^3 \cos(\beta z))}{(c_d^2 - c_s^2)\xi(q^2 - c_d^2\xi^2)}$$
$$+ \frac{[c_s^2(q^2 - c_s^2\xi^2) + (c_d^2 - c_s^2)^2\xi^2](-C_1\alpha \sin(\alpha z) + C_2\alpha \cos(\alpha z) - C_3\beta \sin(\beta z) + C_4\beta \cos(\beta z))}{(c_d^2 - c_s^2)\xi(q^2 - c_d^2\xi^2)}$$
$$= \frac{c_s^2 c_d^2\alpha^2 - c_s^2(q^2 - c_s^2\xi^2) - (c_d^2 - c_s^2)^2\xi^2}{(c_d^2 - c_s^2)\xi(q^2 - c_d^2\xi^2)} C_1\alpha \sin(\alpha z) + \frac{-c_s^2 c_d^2\alpha^2 + c_s^2(q^2 - c_s^2\xi^2) + (c_d^2 - c_s^2)^2\xi^2}{(c_d^2 - c_s^2)\xi(q^2 - c_d^2\xi^2)} C_2\alpha \cos(\alpha z)$$
$$+ \frac{c_s^2 c_d^2\beta^2 - c_s^2(q^2 - c_s^2\xi^2) - (c_d^2 - c_s^2)^2\xi^2}{(c_d^2 - c_s^2)\xi(q^2 - c_d^2\xi^2)} C_3\beta \sin(\beta z) + \frac{-c_s^2 c_d^2\beta^2 + c_s^2(q^2 - c_s^2\xi^2) + (c_d^2 - c_s^2)^2\xi^2}{(c_d^2 - c_s^2)\xi(q^2 - c_d^2\xi^2)} C_4\beta \cos(\beta z) = \frac{C_1\alpha}{\xi}\sin(\alpha z) - \frac{C_2\alpha}{\xi}\cos(\alpha z)$$
$$- \frac{c_d^2\xi C_3\beta}{q^2 - c_d^2\xi^2}\sin(\beta z) + \frac{c_d^2\xi C_4\beta}{q^2 - c_d^2\xi^2}\cos(\beta z) = \frac{C_1\alpha}{\xi}\sin(\alpha z) - \frac{C_2\alpha}{\xi}\cos(\alpha z) - \frac{\xi C_3\beta}{\beta^2}\sin(\beta z) + \frac{\xi C_4\beta}{\beta^2}\cos(\beta z) = \frac{C_1\alpha}{\xi}\sin(\alpha z) - \frac{C_2\alpha}{\xi}\cos(\alpha z)$$
$$- \frac{\xi C_3}{\beta}\sin(\beta z) + \frac{\xi C_4}{\beta}\cos(\beta z) \tag{43}$$

Differentiating Eq. (43) with respect to z leads to

$$\frac{d\bar{\tilde{u}}}{dz} = \frac{C_1\alpha^2}{\xi}\cos(\alpha z) + \frac{C_2\alpha^2}{\xi}\sin(\alpha z) - \xi C_3 \cos(\beta z) - \xi C_4 \sin(\beta z) \tag{44}$$

Substituting Eqs. (44) and (39) into Eq. (26*a*) leads to

$$\bar{\tilde{\tau}}_{rz} = \mu\left[\left(\frac{C_1\alpha^2}{\xi}\cos(\alpha z) + \frac{C_2\alpha^2}{\xi}\sin(\alpha z) - \xi C_3 \cos(\beta z) - \xi C_4 \sin(\beta z)\right) - \xi(C_1 \cos(\alpha z) + C_2 \sin(\alpha z) + C_3 \cos(\beta z) + C_4 \sin(\beta z))\right] = \frac{\alpha^2 - \xi^2}{\xi}\mu C_1 \cos(\alpha z) + \frac{\alpha^2 - \xi^2}{\xi}\mu C_2 \sin(\alpha z) - 2\xi\mu C_3 \cos(\beta\beta z) - 2\xi\mu C_4 \sin(\beta\beta z) \tag{45}$$

Substituting Eqs. (43) and (40) into Eq. (26*b*) leads to

$$\bar{\tilde{\sigma}}_z = \lambda\xi\left(\frac{C_1\alpha}{\xi}\sin(\alpha z) - \frac{C_2\alpha}{\xi}\cos(\alpha z) - \frac{\xi C_3}{\beta}\sin(\beta z) + \frac{\xi C_4}{\beta}\cos(\beta z)\right) + (\lambda + 2\mu)(-C_1\alpha \sin(\alpha z) + C_2\alpha \cos(\alpha z) - C_3\beta \sin(\beta z) + C_4\beta \cos(\beta z)) = -2\mu C_1\alpha \sin(\alpha z) + 2\mu C_2\alpha \cos(\alpha z) - \frac{\mu(\alpha^2 - \xi^2)}{\beta} C_3 \sin(\beta z) + \frac{\mu(\alpha^2 - \xi^2)}{\beta} C_4 \cos(\beta z) \tag{46}$$

So far the general solutions of Eqs. (25*a*), (25*b*), (26*a*), and (26*b*) have been given by Eqs. (43), (39), (45), and (46), respectively. Consider a single layer (i.e., the ith layer) of the multilayered medium. In a local coordinate system the upper boundary of the ith layer along the z direction is defined at $z=0$. Suppose that boundary conditions at $z=0$ are given in the Fourier–Hankel transformed domain as $\bar{\tilde{u}}(\boldsymbol{\xi},z,q)|_{z=0}=\bar{\tilde{u}}_{zo}$, $\bar{\tilde{w}}(\boldsymbol{\xi},z,q)|_{z=0}=\bar{\tilde{w}}_{zo}$, $\bar{\tilde{\boldsymbol{\sigma}}}(\boldsymbol{\xi},z,q)|_{z=0}=\bar{\tilde{\boldsymbol{\sigma}}}_{zo}$, and $\bar{\tilde{\tau}}_{rz}(\boldsymbol{\xi},z,q)|_{z=0}=\bar{\tilde{\tau}}_{zo}$. Applying the solution of Eqs. (43), (39), (45), and (46) at $z=0$ and taking into account boundary conditions (27*a*)–(27*d*) lead to

$$\bar{\tilde{u}}_{zo} = -\frac{C_2\alpha}{\xi} + \frac{\xi C_4}{\beta} \tag{47}$$

$$\bar{\tilde{w}}_{zo} = C_1 + C_3 \tag{48}$$

$$\bar{\tilde{\tau}}_{zo} = \frac{\alpha^2 - \xi^2}{\xi}\mu C_1 - 2\xi\mu C_3 \tag{49}$$

$$\bar{\tilde{\sigma}}_{zo} = 2\mu C_2\alpha + \frac{\mu(\alpha^2 - \xi^2)}{\beta} C_4 \tag{50}$$

Solving algebraic equations (47)–(50) gives

$$C_1 = \frac{2\xi^2}{\alpha^2 + \xi^2}\bar{\tilde{w}}_{zo} + \frac{\xi}{\mu(\alpha^2 + \xi^2)}\bar{\tilde{\tau}}_{zo} \tag{51}$$

$$C_2 = -\frac{\xi(\alpha^2 - \xi^2)}{\alpha(\alpha^2 + \xi^2)}\bar{\tilde{u}}_{zo} + \frac{\xi^2}{\mu\alpha(\alpha^2 + \xi^2)}\bar{\tilde{\sigma}}_{zo} \tag{52}$$

$$C_3 = \frac{\alpha^2 - \xi^2}{\alpha^2 + \xi^2}\bar{\tilde{w}}_{zo} - \frac{\xi}{\mu(\alpha^2 + \xi^2)}\bar{\tilde{\tau}}_{zo} \tag{53}$$

$$C_4 = \frac{2\xi}{\alpha^2 + \xi^2}\bar{\tilde{u}}_{zo} + \frac{\beta}{\mu(\alpha^2 + \xi^2)}\bar{\tilde{\sigma}}_{zo} \tag{54}$$

Denoting $\phi=\sqrt{q^2/c_s^2}$ leads to $\alpha^2+\xi^2=(q^2/c_s^2)-\xi^2+\xi^2=\phi^2$. It follows that

$$C_1 = \frac{2\xi^2}{\phi^2}\bar{\tilde{w}}_{zo} + \frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo} \tag{55}$$

$$C_2 = -\frac{\xi(\alpha^2 - \xi^2)}{\alpha\phi^2}\bar{\tilde{u}}_{zo} + \frac{\xi^2}{\mu\alpha\phi^2}\bar{\tilde{\sigma}}_{zo} \tag{56}$$

$$C_3 = \frac{\alpha^2 - \xi^2}{\phi^2}\bar{\tilde{w}}_{zo} - \frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo} \tag{57}$$

$$C_4 = \frac{2\beta\xi}{\phi^2}\bar{\tilde{u}}_{zo} + \frac{\beta}{\mu\phi^2}\bar{\tilde{\sigma}}_{zo} \tag{58}$$

Substituting Eqs. (55)–(58) into Eqs. (43), (39), (45), and (46) leads to

$$\bar{\tilde{u}} = \frac{\left(\frac{2\xi^2}{\phi^2}\bar{\tilde{w}}_{zo} + \frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)\alpha}{\xi}\sin(\alpha z) - \frac{\left(-\frac{\xi(\alpha^2 - \xi^2)}{\alpha\phi^2}\bar{\tilde{u}}_{zo} + \frac{\xi^2}{\mu\alpha\phi^2}\bar{\tilde{\sigma}}_{zo}\right)\alpha}{\xi}\cos(\alpha z) - \frac{\xi\left(\frac{\alpha^2 - \xi^2}{\phi^2}\bar{\tilde{w}}_{zo} - \frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)}{\beta}\sin(\beta z)$$

$$+\frac{\xi\left(\frac{2\beta\xi}{\phi^2}\bar{\tilde{u}}_{zo}+\frac{\beta}{\mu\phi^2}\bar{\tilde{\sigma}}_{zo}\right)}{\beta}\cos(\beta z) \tag{59a}$$

$$\bar{\tilde{w}}=\left(\frac{2\xi^2}{\phi^2}\bar{\tilde{w}}_{zo}+\frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)\cos(\alpha z)+\left(-\frac{\xi(\alpha^2-\xi^2)}{\alpha\phi^2}\bar{\tilde{u}}_{zo}+\frac{\xi^2}{\mu\alpha\phi^2}\bar{\tilde{\sigma}}_{zo}\right)\sin(\alpha z)+\left(\frac{\alpha^2-\xi^2}{\phi^2}\bar{\tilde{w}}_{zo}-\frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)\cos(\beta z)+\left(\frac{2\beta\xi}{\phi^2}\bar{\tilde{u}}_{zo}+\frac{\beta}{\mu\phi^2}\bar{\tilde{\sigma}}_{zo}\right)\sin(\beta z) \tag{59b}$$

$$\bar{\tilde{\tau}}_{rz}=\frac{\alpha^2-\xi^2}{\xi}\mu\left(\frac{2\xi^2}{\phi^2}\bar{\tilde{w}}_{zo}+\frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)\cos(\alpha z)+\frac{\alpha^2-\xi^2}{\xi}\mu\left(-\frac{\xi(\alpha^2-\xi^2)}{\alpha\phi^2}\bar{\tilde{u}}_{zo}+\frac{\xi^2}{\mu\alpha\phi^2}\bar{\tilde{\sigma}}_{zo}\right)\sin(\alpha z)-2\xi\mu\left(\frac{\alpha^2-\xi^2}{\phi^2}\bar{\tilde{w}}_{zo}-\frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)\cos(\beta z)-2\xi\mu\left(\frac{2\beta\xi}{\phi^2}\bar{\tilde{u}}_{zo}+\frac{\beta}{\mu\phi^2}\bar{\tilde{\sigma}}_{zo}\right)\sin(\beta z) \tag{59c}$$

$$\bar{\tilde{\sigma}}_z=-2\mu\left(\frac{2\xi^2}{\phi^2}\bar{\tilde{w}}_{zo}+\frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)\alpha\sin(\alpha z)+2\mu\left(-\frac{\xi(\alpha^2-\xi^2)}{\alpha\phi^2}\bar{\tilde{u}}_{zo}+\frac{\xi^2}{\mu\alpha\phi^2}\bar{\tilde{\sigma}}_{zo}\right)\alpha\cos(\alpha z)-\frac{\mu(\alpha^2-\xi^2)}{\beta}\left(\frac{\alpha^2-\xi^2}{\phi^2}\bar{\tilde{w}}_{zo}-\frac{\xi}{\mu\phi^2}\bar{\tilde{\tau}}_{zo}\right)\sin(\beta z)+\frac{\mu(\alpha^2-\xi^2)}{\beta}\left(\frac{2\beta\xi}{\phi^2}\bar{\tilde{u}}_{zo}+\frac{\beta}{\mu\phi^2}\bar{\tilde{\sigma}}_{zo}\right)\cos(\beta z) \tag{59d}$$

Simplifying the above four expressions results in

$$\bar{\tilde{u}}=\left[\frac{\alpha^2-\xi^2}{\phi^2}\cos(\alpha z)+\frac{2\xi^2}{\phi^2}\cos(\beta z)\right]\bar{\tilde{u}}_{zo}+\left[\frac{2\alpha\xi}{\phi^2}\sin(\alpha z)-\frac{\xi(\alpha^2-\xi^2)}{\beta\phi^2}\sin(\beta z)\right]\bar{\tilde{w}}_{zo}+\left[\frac{\alpha}{\mu\phi^2}\sin(\alpha z)+\frac{\xi^2}{\mu\beta\phi^2}\sin(\beta z)\right]\bar{\tilde{\tau}}_{zo}+\left[-\frac{\xi}{\mu\phi^2}\cos(\alpha z)+\frac{\xi}{\mu\phi^2}\cos(\beta z)\right]\bar{\tilde{\sigma}}_{zo} \tag{60a}$$

$$\bar{\tilde{w}}=\left[\frac{2\beta\xi}{\phi^2}\sin(\beta z)-\frac{\xi(\alpha^2-\xi^2)}{\alpha\phi^2}\sin(\alpha z)\right]\bar{\tilde{u}}_{zo}+\left[\frac{\alpha^2-\xi^2}{\phi^2}\cos(\beta z)+\frac{2\xi^2}{\phi^2}\cos(\alpha z)\right]\bar{\tilde{w}}_{zo}+\frac{1}{\mu}\left[-\frac{\xi}{\phi^2}\cos(\beta z)+\frac{\xi}{\phi^2}\cos(\alpha z)\right]\bar{\tilde{\tau}}_{zo}+\frac{1}{\mu}\left[\frac{\beta}{\phi^2}\sin(\beta z)+\frac{\xi^2}{\alpha\phi^2}\sin(\alpha z)\right]\bar{\tilde{\sigma}}_{zo} \tag{60b}$$

$$\bar{\tilde{\tau}}_{rz}=\mu\left[-\frac{(\alpha^2-\xi^2)^2}{\alpha\phi^2}\sin(\alpha z)-\frac{4\xi^2\beta}{\phi^2}\sin(\beta z)\right]\bar{\tilde{u}}_{zo}+\mu\frac{2\xi(\alpha^2-\xi^2)}{\phi^2}[\cos(\alpha z)-\cos(\beta z)]\bar{\tilde{w}}_{zo}+\left[\frac{\alpha^2-\xi^2}{\phi^2}\cos(\alpha z)+\frac{2\xi^2}{\phi^2}\cos(\beta z)\right]\bar{\tilde{\tau}}_{zo}+\left[\frac{\xi(\alpha^2-\xi^2)}{\alpha\phi^2}\sin(\alpha z)-\frac{2\xi\beta}{\phi^2}\sin(\beta z)\right]\bar{\tilde{\sigma}}_{zo} \tag{60c}$$

$$\bar{\tilde{\sigma}}_z=\mu\frac{2\xi(\alpha^2-\xi^2)}{\phi^2}[\cos(\beta z)-\cos(\alpha z)]\bar{\tilde{u}}_{zo}+\mu\left[-\frac{(\alpha^2-\xi^2)^2}{\beta\phi^2}\sin(\beta z)-\frac{4\xi^2\alpha}{\phi^2}\sin(\alpha z)\right]\bar{\tilde{w}}_{zo}+\left[\frac{\xi(\alpha^2-\xi^2)}{\beta\phi^2}\sin(\beta z)-\frac{2\xi\alpha}{\phi^2}\sin(\alpha z)\right]\bar{\tilde{\tau}}_{zo}+\left[\frac{\alpha^2-\xi^2}{\phi^2}\cos(\beta z)+\frac{2\xi^2}{\phi^2}\cos(\alpha z)\right]\bar{\tilde{\sigma}}_{zo} \tag{60d}$$

The above results can be written more concisely as

$$\bar{\tilde{u}}(\xi,z,q)=\theta_{11}\bar{\tilde{u}}_{zo}+\theta_{12}\bar{\tilde{w}}_{zo}+\theta_{13}\bar{\tilde{\tau}}_{zo}+\theta_{14}\bar{\tilde{\sigma}}_{zo} \tag{61a}$$

$$\bar{\tilde{w}}(\xi,z,q)=\theta_{21}\bar{\tilde{u}}_{zo}+\theta_{22}\bar{\tilde{w}}_{zo}+\theta_{23}\bar{\tilde{\tau}}_{zo}+\theta_{24}\bar{\tilde{\sigma}}_{zo} \tag{61b}$$

$$\bar{\tilde{\tau}}_{rz}(\xi,z,q)=\theta_{31}\bar{\tilde{u}}_{zo}+\theta_{32}\bar{\tilde{w}}_{zo}+\theta_{33}\bar{\tilde{\tau}}_{zo}+\theta_{34}\bar{\tilde{\sigma}}_{zo} \tag{61c}$$

$$\bar{\tilde{\sigma}}_{rz}(\xi,z,q)=\theta_{41}\bar{\tilde{u}}_{zo}+\theta_{42}\bar{\tilde{w}}_{zo}+\theta_{43}\bar{\tilde{\tau}}_{zo}+\theta_{44}\bar{\tilde{\sigma}}_{zo} \tag{61d}$$

or equivalently, for the ith layer

$$\begin{bmatrix}\bar{\tilde{u}}(\boldsymbol{\xi},z,q)\\ \bar{\tilde{w}}(\boldsymbol{\xi},z,q)\\ \bar{\tilde{\boldsymbol{\tau}}}_{rz}(\boldsymbol{\xi},z,q)\\ \bar{\tilde{\boldsymbol{\sigma}}}_{z}(\boldsymbol{\xi},z,q)\end{bmatrix}=[\Phi(z)]\begin{bmatrix}\bar{\tilde{u}}(\boldsymbol{\xi},z,q)|_{z=0}\\ \bar{\tilde{w}}(\boldsymbol{\xi},z,q)|_{z=0}\\ \bar{\tilde{\boldsymbol{\tau}}}_{rz}(\boldsymbol{\xi},z,q)|_{z=0}\\ \bar{\tilde{\boldsymbol{\sigma}}}_{z}(\boldsymbol{\xi},z,q)|_{z=0}\end{bmatrix}\quad\text{and}\quad[\Phi(z)]=[\Phi(\xi,z,q)]=\begin{bmatrix}\theta_{11}\theta_{12}\theta_{13}\theta_{14}\\ \theta_{21}\theta_{22}\theta_{23}\theta_{24}\\ \theta_{31}\theta_{32}\theta_{33}\theta_{34}\\ \theta_{41}\theta_{42}\theta_{43}\theta_{44}\end{bmatrix} \tag{62}$$

where

$$\theta_{11}=\frac{\alpha^2-\xi^2}{\phi^2}\cos(\alpha z)+\frac{2\xi^2}{\phi^2}\cos(\beta z) \tag{63a}$$

$$\theta_{12}=\frac{2\xi\alpha}{\phi^2}\sin(\alpha z)-\frac{\xi(\alpha^2-\xi^2)}{\beta\phi^2}\sin(\beta z) \tag{63b}$$

$$\theta_{13}=\frac{1}{\mu}\left[\frac{\alpha}{\phi^2}\sin(\alpha z)+\frac{\xi^2}{\beta\phi^2}\sin(\beta z)\right] \tag{63c}$$

$$\theta_{14}=\frac{1}{\mu}\left[-\frac{\xi}{\phi^2}\cos(\alpha z)+\frac{\xi}{\phi^2}\cos(\beta z)\right] \tag{63d}$$

$$\theta_{21}=\frac{2\xi\beta}{\phi^2}\sin(\beta z)-\frac{\xi(\alpha^2-\xi^2)}{\alpha\phi^2}\sin(\alpha z) \tag{64a}$$

$$\theta_{22}=\frac{\alpha^2-\xi^2}{\phi^2}\cos(\beta z)+\frac{2\xi^2}{\phi^2}\cos(\alpha z) \tag{64b}$$

$$\theta_{23}=\frac{1}{\mu}\left[-\frac{\xi}{\phi^2}\cos(\beta z)+\frac{\xi}{\phi^2}\cos(\alpha z)\right] \tag{64c}$$

$$\theta_{24}=\frac{1}{\mu}\left[\frac{\beta}{\phi^2}\sin(\beta z)+\frac{\xi^2}{\alpha\phi^2}\sin(\alpha z)\right] \tag{64d}$$

$$\theta_{31}=\mu\left[-\frac{(\alpha^2-\xi^2)^2}{\alpha\phi^2}\sin(\alpha z)-\frac{4\xi^2\beta}{\phi^2}\sin(\beta z)\right] \tag{65a}$$

$$\theta_{32}=\mu\frac{2\xi(\alpha^2-\xi^2)}{\phi^2}[\cos(\alpha z)-\cos(\beta z)] \tag{65b}$$

$$\theta_{33}=\frac{\alpha^2-\xi^2}{\phi^2}\cos(\alpha z)+\frac{2\xi^2}{\phi^2}\cos(\beta z) \tag{65c}$$

$$\theta_{34} = \frac{\xi(\alpha^2 - \xi^2)}{\alpha\phi^2}\sin(\alpha z) - \frac{2\xi\beta}{\phi^2}\sin(\beta z) \tag{65d}$$

$$\theta_{41} = \mu\frac{2\xi(\alpha^2 - \xi^2)}{\phi^2}[\cos(\beta z) - \cos(\alpha z)] \tag{66a}$$

$$\theta_{42} = \mu\left[-\frac{(\alpha^2 - \xi^2)^2}{\beta\phi^2}\sin(\beta z) - \frac{4\xi^2\alpha}{\phi^2}\sin(\alpha z)\right] \tag{66b}$$

$$\theta_{43} = \frac{\xi(\alpha^2 - \xi^2)}{\beta\phi^2}\sin(\beta z) - \frac{2\xi\alpha}{\phi^2}\sin(\alpha z) \tag{66c}$$

$$\theta_{44} = \frac{\alpha^2 - \xi^2}{\phi^2}\cos(\beta z) + \frac{2\xi^2}{\phi^2}\cos(\alpha z) \tag{66d}$$

4.2 Assembly of Responses of Multiple Layers. Since we omitted the subscript i in the foregoing derivations, variable z in Eq. (62) should indeed be z_i. Note that the relationship between a local coordinate system (r_i, z_i, θ_i) of the ith layer and a global coordinate system (r, Z, θ) of the entire multilayered solid is given by $Z = z_i + h_i$. As a result, Eq. (62) can be written in the global coordinate system as

$$\begin{bmatrix} \bar{\bar{u}}(\boldsymbol{\xi}, z, q) \\ \bar{\bar{w}}(\boldsymbol{\xi}, z, q) \\ \bar{\bar{\boldsymbol{\tau}}}_{rz}(\boldsymbol{\xi}, z, q) \\ \bar{\bar{\boldsymbol{\sigma}}}_{z}(\boldsymbol{\xi}, z, q) \end{bmatrix} = [\Phi(\boldsymbol{\xi}, Z - h_i, q)] \begin{bmatrix} \bar{\bar{u}}(\boldsymbol{\xi}, Z, q)|_{Z=h_i} \\ \bar{\bar{w}}(\boldsymbol{\xi}, Z, q)|_{Z=h_i} \\ \bar{\bar{\boldsymbol{\tau}}}_{rz}(\boldsymbol{\xi}, Z, q)|_{Z=h_i} \\ \bar{\bar{\boldsymbol{\sigma}}}_{z}(\boldsymbol{\xi}, Z, q)|_{Z=h_i} \end{bmatrix} \tag{67}$$

The assembly of response of multilayers is implemented through the boundary conditions at interfaces between an upper layer and a lower layer. In this study continuity condition is assumed for response at the interface. That is, stresses and displacements at the bottom of the upper layer, respectively, equal to stresses and displacements at the bottom of the lower layer. Start from the Nth layer (bottom layer in Fig. 1). Following Eq. (62) stress and displacement fields at any location in the Nth layer can be expressed in terms of stress and displacement boundary conditions at the upper boundary of the Nth layer or equivalently the lower boundary of the $(N-1)$th layer. Now, consider the $(N-1)$th layer. Stress and displacement fields at any location in the $(N-1)$th layer (including the lower boundary of the $(N-1)$th layer) can be expressed in terms of stress and displacement boundary conditions at the upper boundary of the $(N-1)$th layer or equivalently the lower boundary of the $(N-2)$th layer. In this way, stress and displacement fields at any location in the Nth and the $(N-1)$th layers can be expressed as functions of the upper boundary conditions of the $(N-1)$th layer. If one keeps assembling multiple layers, eventually, stress and displacement fields at any location in the multilayered medium can be expressed as functions of the upper boundary conditions of the first layer (i.e., top layer).

Indeed, two displacement fields $u(r, z_1, t)|_{z_1=0}$ and $w(r, z_1, t)|_{z_1=0}$ at the upper boundary of a multilayered medium under a vertical load are unknown and thus cannot be used as boundary conditions. For the multilayered medium resting on bedrock as shown in Fig. 1, two stress boundary conditions are given at $z_1 = 0$ and two displacement boundary conditions are at the bottom $z_N = \Delta h_N$ as follows:

$$\boldsymbol{\tau}_{rz}(r, z_1, t) = \boldsymbol{\tau}_{rz}(r, z_1, t)|_{z_1=0} = 0 \quad (r, z) \in S_\sigma \tag{68a}$$

$$\boldsymbol{\sigma}_{z}(r, z_1, t) = \boldsymbol{\sigma}_{z}(r, z_1, t)|_{z_1=0} \quad (r, z) \in S_\sigma \tag{68b}$$

$$u(r, z_N, t) = u(r, z_N, t)|_{z_N=\Delta h_N} = 0 \quad (r, z) \in s_u \tag{68c}$$

$$w(r, z_N, t) = w(r, z_N, t)|_{z_N=\Delta h_N} = 0 \quad (r, z) \in s_u \tag{68d}$$

According to the above-described assembly rule, $u(r, z_N, t)|_{z_N=\Delta h_N}$ and $w(r, z_N, t)|_{z_N=\Delta h_N}$ can be written as functions of $u(r, z_1, t)|_{z_1=0}$, $w(r, z_1, t)|_{z_1=0}$, $\boldsymbol{\tau}_{rz}(r, z_1, t)|_{z_1=0}$, and $\boldsymbol{\sigma}_{z}(r, z_1, t)|_{z_1=0}$. Therefore, the two unknown upper boundary displacement conditions $u(r, z_1, t)|_{z_1=0}$ and $w(r, z_1, t)|_{z_1=0}$ can be solved using two upper boundary stress conditions $\boldsymbol{\tau}_{rz}(r, z_1, t)|_{z_1=0}$ and $\boldsymbol{\sigma}_{z}(r, z_1, t)|_{z_1=0}$ as well as two lower boundary displacement conditions $u(r, z_N, t)|_{z_N=\Delta h_N}$ and $w(r, z_N, t)|_{z_N=\Delta h_N}$. It then follows that all stress and displacement fields in the entire multilayered medium can be solved.

To elaborate the above assembly process, define displacement and stress fields as $\lfloor \bar{\bar{U}}(Z) \rfloor = [\bar{\bar{u}}(\boldsymbol{\xi}, Z, q), \bar{\bar{w}}(\boldsymbol{\xi}, Z, q)]$ and $\lfloor \bar{\bar{T}}(Z) \rfloor = [\bar{\bar{\boldsymbol{\tau}}}_{rz}(\boldsymbol{\xi}, Z, q), \bar{\bar{\boldsymbol{\sigma}}}_{z}(\boldsymbol{\xi}, Z, q)]$, respectively. Equation (67) can be written as

$$[\bar{\bar{U}}(Z), \bar{\bar{T}}(Z)]^T = [\Phi(Z - h_i)][\bar{\bar{U}}(h_i), \bar{\bar{T}}(h_i)]^T \tag{69}$$

where T stands for transpose. For the interface between the ith layer and the $(i-1)$th layer, let $Z = h_i^+$ be the upper boundary belonging to the $(i-1)$th layer, and let $Z = h_i^-$ be the lower boundary belonging to the ith layer. At $Z = h_{i+1}^+$, Eq. (69) becomes

$$[\bar{\bar{U}}(h_{i+1}^+), \bar{\bar{T}}(h_{i+1}^+)]^T = [\Phi(\Delta h_i)][\bar{\bar{U}}(h_i^-), \bar{\bar{T}}(h_i^-)]^T \tag{70}$$

where Δh_i is the thickness of the ith layer, $\Delta h_i = h_{i+1}^+ - h_i$. When a continuous interface condition is assumed, that is, $\lfloor \bar{\bar{U}}(h_i^-) \rfloor = \lfloor \bar{\bar{U}}(h_i^+) \rfloor$ and $\lfloor \bar{\bar{T}}(h_i^-) \rfloor = \lfloor \bar{\bar{T}}(h_i^+) \rfloor$, Eq. (70) can be expanded as follows:

$$\begin{aligned}[\bar{\bar{U}}(h_{i+1}^+), \bar{\bar{T}}(h_{i+1}^+)]^T &= [\Phi(\Delta h_i)][\bar{\bar{U}}(h_i^-), \bar{\bar{T}}(h_i^-)]^T \\ &= [\Phi(\Delta h_i)][\bar{\bar{U}}(h_i^+)^-, \bar{\bar{T}}(h_i^+)]^T \\ &= [\Phi(\Delta h_i)][\Phi(\Delta h_{i-1})][\Phi(\Delta h_{i-2})] \cdots [\Phi(\Delta h_2)] \\ &\quad \times [\Phi(\Delta h_1)][\bar{\bar{U}}(h_1), \bar{\bar{T}}(h_1)]^T \end{aligned} \tag{71}$$

So Eq. (69) becomes

$$\begin{aligned}[\bar{\bar{U}}(Z), \bar{\bar{T}}(Z)]^T &= [\Phi(Z - h_i)][\Phi(\Delta h_{i-1})] \cdots [\Phi(\Delta h_2)][\Phi(\Delta h_1)] \\ &\quad \times [\bar{\bar{U}}(h_1), \bar{\bar{T}}(h_1)]^T \end{aligned} \tag{72}$$

where $h_i \le Z \le h_{i+1}$. Equation (72) means that, for each layer, the transformed displacements and stresses at the bottom interface can be expressed in terms of their counterparts at the upper interface. The bedrock foundation underneath the Nth layer (the $(N+1)$th layer) as indicated by lower boundary conditions in Eqs. (68c) and (68d) can be written in a global coordinate system as $u(r, Z, t)_{Z=h_{N+1}} = 0$ and $w(r, Z, t)_{Z=h_{N+1}} = 0$ or equivalently in the transformed domain

$$\lfloor \bar{\bar{U}}(h_{N+1}) \rfloor = [\bar{\bar{u}}(\boldsymbol{\xi}, h_{N+1}, q), \bar{\bar{w}}(\boldsymbol{\xi}, h_{N+1}, q)] = [0, 0] \tag{73}$$

As a result, displacement and stress fields in the Laplace–Hankel domain can be ultimately represented in terms of stresses (in the Laplace–Hankel domain) on the surface of the multilayered medium and bedrock boundary conditions

$$[\bar{\bar{U}}(Z), \bar{\bar{T}}(Z)]^T = [\Psi(Z)]_{4\times 2}[\bar{\bar{T}}(h_1)]^T_{2\times 1} \tag{74}$$

where

$$[\Psi(Z)] = [\Phi(Z - h_i)][\Phi(\Delta h_{i-1})] \cdots [\Phi(\Delta h_2)][\Phi(\Delta h_1)] \begin{bmatrix} [R_1] \\ I \end{bmatrix}$$

is a 4×2 matrix and $[I] = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ is a 2×2 unit matrix; for $1 \le i \le N$, $[R_i] = -K_{11}^{-1}K_{12}$ is a 2×2 matrix and $[R_{N+1}] = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, in which

$$[K]=\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}=[\Phi(\Delta h_N)][\Phi(\Delta h_{N-1})]\cdots[\Phi(\Delta h_{i+1})][\Phi(\Delta h_i)]$$

in which $[K]$ is a 4×4 matrix, and $[K_{11}]$, $[K_{12}]$, $[K_{21}]$, and $[K_{22}]$ are 2×2 block matrices each; $[\Phi(\Delta h_i)]=\begin{bmatrix} \Theta^i_{11} & \Theta^i_{12} \\ \Theta^i_{21} & \Theta^i_{22} \end{bmatrix}$, where $[\Theta^i_{11}]=\begin{bmatrix} \phi_{11}(\Delta h_i) & \phi_{12}(\Delta h_i) \\ \phi_{21}(\Delta h_i) & \phi_{22}(\Delta h_i) \end{bmatrix}$, $[\Theta^i_{12}]=\begin{bmatrix} \phi_{13}(\Delta h_i) & \phi_{14}(\Delta h_i) \\ \phi_{23}(\Delta h_i) & \phi_{24}(\Delta h_i) \end{bmatrix}$, $[\Theta^i_{21}]=\begin{bmatrix} \phi_{31}(\Delta h_i) & \phi_{32}(\Delta h_i) \\ \phi_{41}(\Delta h_i) & \phi_{42}(\Delta h_i) \end{bmatrix}$, and $[\Theta^i_{22}]=\begin{bmatrix} \phi_{33}(\Delta h_i) & \phi_{34}(\Delta h_i) \\ \phi_{43}(\Delta h_i) & \phi_{44}(\Delta h_i) \end{bmatrix}$.

Applying inverse Fourier transform and inverse Hankel transform of order 1 to $\bar{\tilde{u}}(\xi,Z,q)$ and $\bar{\tilde{\tau}}_{rz}(\xi,Z,q)$, and inverse Fourier transform and inverse Hankel transform of order 0 to $\bar{\tilde{w}}(\xi,Z,q)$ and $\bar{\tilde{\sigma}}_z(\xi,Z,q)$, Eq. (74) results in displacement and stress fields in the space-time domain

$$\begin{Bmatrix} u(r,Z,t) \\ w(r,Z,t) \\ \boldsymbol{\tau}_{rz}(r,Z,t) \\ \boldsymbol{\sigma}_z(r,Z,t) \end{Bmatrix}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\boldsymbol{\xi}\,\mathrm{diag}[J(\xi r)][\Psi(Z)][\bar{\tilde{T}}(h_1)]^T d\boldsymbol{\xi}e^{iqt}dq \tag{75}$$

where

$$[\Psi(Z)]=\begin{bmatrix} \boldsymbol{\psi}_1(\boldsymbol{\xi},Z,q) \\ \boldsymbol{\psi}_2(\boldsymbol{\xi},Z,q) \\ \boldsymbol{\psi}_3(\boldsymbol{\xi},Z,q) \\ \boldsymbol{\psi}_4(\boldsymbol{\xi},Z,q) \end{bmatrix}=\begin{bmatrix} \boldsymbol{\psi}_{11}(\boldsymbol{\xi},Z,q) & \boldsymbol{\psi}_{12}(\boldsymbol{\xi},Z,q) \\ \boldsymbol{\psi}_{21}(\boldsymbol{\xi},Z,q) & \boldsymbol{\psi}_{22}(\boldsymbol{\xi},Z,q) \\ \boldsymbol{\psi}_{31}(\boldsymbol{\xi},Z,q) & \boldsymbol{\psi}_{32}(\boldsymbol{\xi},Z,q) \\ \boldsymbol{\psi}_{41}(\boldsymbol{\xi},Z,q) & \boldsymbol{\psi}_{42}(\boldsymbol{\xi},Z,q) \end{bmatrix}$$

in which $[\psi_1(\xi,Z,q)]$ and $[\psi_2(\xi,Z,q)]$, $[\psi_3(\xi,Z,q)]$, and $[\psi_4(\xi,Z,q)]$ are 1×2 matrices each, and $[J(\boldsymbol{\xi}r)]=\mathrm{diag}[J_1(\boldsymbol{\xi}r)\ J_0(\boldsymbol{\xi}r)\ J_1(\boldsymbol{\xi}r)\ J_0(\boldsymbol{\xi}r)]$. With these notations, Eq. (75) can be written as

$$u(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_1(\xi r)[\psi_1(\xi,Z,q)]\times[\bar{\tilde{\tau}}_{rz}(r,h_1,t),\bar{\tilde{\sigma}}_z(r,h_1,t)]^T d\xi e^{iqt}dq \tag{76a}$$

$$w(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_0(\xi r)[\psi_2(\xi,Z,q)]\times[\bar{\tilde{\tau}}_{rz}(r,h_1,t),\bar{\tilde{\sigma}}_z(r,h_1,t)]^T d\xi e^{iqt}dq \tag{76b}$$

$$\tau_{rz}(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_1(\xi r)[\psi_3(\xi,Z,q)]\times[\bar{\tilde{\tau}}_{rz}(r,h_1,t),\bar{\tilde{\sigma}}_z(r,h_1,t)]^T d\xi e^{iqt}dq \tag{76c}$$

$$\sigma_z(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_0(\xi r)[\psi_4(\xi,Z,q)]\times[\bar{\tilde{\tau}}_{rz}(r,h_1,t),\bar{\tilde{\sigma}}_z(r,h_1,t)]d\xi e^{iqt}dq \tag{76d}$$

In practice, measurement of surface velocity and acceleration can be easier than that of surface displacement. Vertical velocity $\dot{u}(r,Z,t)$ and acceleration $\ddot{u}(r,Z,t)$ of the surface of the multilayered system can be given as first and second derivatives of displacement with respect to time t,

$$\dot{u}(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_1(\xi r)(iq)[\psi_1(\xi,Z,q)]\times[\bar{\tilde{\tau}}_{rz}(r,h_1,t),\bar{\tilde{\sigma}}_z(r,h_1,t)]^T d\xi e^{iqt}dq \tag{77a}$$

$$\ddot{u}(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_1(\xi r)(-q^2)[\psi_1(\xi,Z,q)]\times[\bar{\tilde{\tau}}_{rz}(r,h_1,t),\bar{\tilde{\sigma}}_z(r,h_1,t)]^T d\xi e^{iqt}dq \tag{77b}$$

The IRF, $\mathbf{h}(\mathbf{x},t)=[u(r,Z,t)_\delta, w(r,Z,t)_\delta, \tau_{rz}(r,Z,t)_\delta, \sigma_z(r,Z,t)_\delta]^T$, of a multilayered elastic medium is the solution of Eqs. (6*a*) and (10) given that the exerted dynamic load is expressed by an impulsive load (Eq. (5)), which corresponds to the following boundary condition:

$$\sigma_z(r,h_1,t)=-\frac{1}{\pi r_0^2}\delta(t)H(r_0-r) \tag{78}$$

Here, h_1 is the distance between the surface of the multilayered medium and the upper interface of the surface layer. By definition, $h_1=0$. Clearly, the Fourier–Hankel transform of Eq. (78) can be written as

$$\bar{\tilde{\sigma}}_z(\xi,h_1,q)|_{h_1=0}=-\frac{J_1(r_0\xi)}{\pi r_0\xi} \tag{79}$$

where $J_1(\cdot)$ is the first order Bessel function of the first kind. Because there exists only a normal stress on the surface, elastic solutions (77*a*) and (77*b*) become

$$u(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_1(\xi r)\psi_{12}(\xi,Z,q)\bar{\tilde{\sigma}}_z(r,h_1,t)d\xi e^{iqt}dq \tag{80a}$$

$$w(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_0(\xi r)\psi_{22}(\xi,Z,q)\bar{\tilde{\sigma}}_z(r,h_1,t)d\xi e^{iqt}dq \tag{80b}$$

$$\tau_{rz}(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_1(\xi r)\psi_{32}(\xi,Z,q)\bar{\tilde{\sigma}}_z(r,h_1,t)d\xi e^{iqt}dq \tag{80c}$$

$$\sigma_z(r,Z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty}\xi J_0(\xi r)\psi_{42}(\xi,Z,q)\bar{\tilde{\sigma}}_z(r,h_1,t)d\xi e^{iqt}dq \tag{80d}$$

The IRF $\mathbf{h}(\mathbf{x},t)$ can thus be obtained by substituting Eq. (79) into Eqs. (80*a*)–(80*d*),

$$u(r,Z,t)_\delta=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}J_1(\xi r)J_1(r_0\xi)\psi_{12}(\xi,Z,q)e^{iqt}d\xi dq \tag{81a}$$

$$w(r,Z,t)_\delta=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}J_0(\xi r)J_1(r_0\xi)\psi_{22}(\xi,Z,q)e^{iqt}d\xi dq \tag{81b}$$

$$\tau_{rz}(r,Z,t)_\delta=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}J_1(\xi r)J_1(r_0\xi)\psi_{32}(\xi,Z,q)e^{iqt}d\xi dq \tag{81c}$$

$$\sigma_z(r,Z,t)_\delta=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}J_0(\xi r)J_1(r_0\xi)\psi_{42}(\xi,Z,q)e^{iqt}d\xi dq \tag{81d}$$

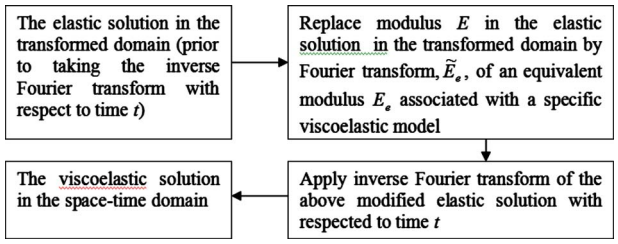


Fig. 3 Procedures for applying the elastic-viscoelastic corresponding principles

4.3 Viscoelastic Media. To obtain the IRF of the multilayered viscoelastic medium, one needs to apply the correspondence principle of elastic-viscoelastic theory [26,37]. As Poisson ratio ν has a relatively small effect on dynamic responses of multilayered media, it is assumed in this study that Poisson ratio ν is a time-independent constant. It follows that the Fourier transform of ν is ν per se. According to the correspondence principle of elastic-viscoelastic theory, the viscoelastic solution of a linear system can be obtained following procedures in Fig. 3.

Equivalent modulus E_e for a specific viscoelastic model is defined through stress-strain relationship $E_e=\sigma/\varepsilon$. For a viscoelastic material, E_e is no longer a constant but a time-dependent function. Because Young's elastic modulus E and Poisson's ratio ν appearing in the elastic solution are in terms of λ and μ, λ and μ as given by Eqs. (7*a*) and (7*b*) also become functions of time. Based on the elastic-viscoelastic correspondence principle, taking Fourier transform $\hat{E}_e$ of an equivalent elastic modulus E_e is equivalent to taking Fourier transform of λ and μ. In other words, to obtain the IRF of the multilayered viscoelastic medium one needs to update λ and μ in elastic solution [Eqs. (81*b*)–(81*d*)] by $\hat{\lambda}$ and $\hat{\mu}$, respectively,

$$\hat{\lambda}=\frac{\hat{E}_e\nu}{(1+\nu)(1-2\nu)} \tag{82a}$$

$$\hat{\mu}=\frac{\hat{E}_e}{2(1+\nu)} \tag{82b}$$

where $\hat{\lambda}$ and $\hat{\mu}$ are Fourier transforms of λ and μ, and Fourier transform of E_e is denoted by $\hat{E}_e$, often called complex modulus. It is noted that in Eqs. (81*b*)–(81*d*) coefficients that involve λ and μ are θ_{ij} $(i,j=1,\ldots,4)$, $\dot{\theta}_{ij}$ $(i=1,2,\text{ and } j=1,\ldots,4)$, α^2, β^2, p^2, c_d, and c_s.

The remainder of this section dedicates to the elaboration of complex modulus $\hat{E}_e$ corresponding to a specific viscoelastic model. Complex modulus, $\hat{E}_e$, of a Kelvin model is given by Fourier transform of E_e given by Eq. (12),

$$\hat{E}_e=E_1(1+T_1\Omega i) \tag{83}$$

where Ω is the frequency of sinusoidal vibratory load $\sigma=\sigma_0 e^{i\Omega t}$ and it corresponds to variable t in the time domain. Following Eq. (14) complex modulus $\hat{E}_e$ of a Maxwell model is defined by

$$\hat{E}_e=\frac{E_0T_0\Omega i}{T_0\Omega i+1} \tag{84}$$

Complex modulus of a Burgers model $\hat{E}_e$ is defined in the Fourier transform of Eq. (16),

$$\hat{E}_e=\left[\frac{E_0T_0\Omega i}{T_0\Omega i+1}+\frac{1}{E_1(1+T_1\Omega i)}\right]^{-1} \tag{85}$$

The complex modulus, $\hat{E}_e$, of the generalized model is defined in the Fourier transform of Eq. (18),

$$\hat{E}_e=\left[\frac{E_0T_0\Omega i}{T_0\Omega i+1}+\sum_{i=1}^{n}\frac{1}{E_i(1+T_i\Omega i)}\right]^{-1} \tag{86}$$

5 Dynamic Response of the Media to Moving Loads

Dynamic response of the multilayered viscoelastic medium to a moving distributed load with varying amplitudes expressed in Eq. (1) can be constructed using Sun's convolution representation by incorporating IRF $\mathbf{h}(\mathbf{x},t)$, as given in Eqs. (81*b*)–(81*d*) with Eq. (4) to give

$$u(x,y,Z,t)=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}p(t-\vartheta)\times J_1[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\times J_1(r_0\xi)\psi_{12}(\xi,Z,q)e^{iq\vartheta}d\vartheta d\xi dq \tag{87a}$$

$$w(x,y,Z,t)=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}p(t-\vartheta)\times J_0[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\times J_1(r_0\xi)\psi_{22}(\xi,Z,q)e^{iq\vartheta}d\vartheta d\xi dq \tag{87b}$$

$$\tau_{rz}(x,y,Z,t)=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}p(t-\vartheta)\times J_1[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\times J_1(r_0\xi)\psi_{32}(\xi,Z,q)e^{iq\vartheta}d\vartheta d\xi dq \tag{87c}$$

$$\sigma_z(x,y,Z,t)=-\frac{1}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}p(t-\vartheta)\times J_0[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\times J_1(r_0\xi)\psi_{42}(\xi,Z,q)e^{iq\vartheta}d\vartheta d\xi dq \tag{87d}$$

where $r=\sqrt{(x-vt+v\vartheta)^2+y^2}$.

The exerted dynamic load is a moving distributed harmonic load given by

$$F_h(\mathbf{x},t)=\frac{H[r_0^2-(x-vt)^2-y^2]}{\pi r_0^2}\delta(z)\sigma_0 e^{i\Omega t} \tag{88}$$

Here, σ_0 is a constant stress. The dynamic response of the medium can be obtained by replacing $p(t-\vartheta)$ in Eqs. (87*a*)–(87*d*) using $\sigma_0 e^{i\Omega(t-\vartheta)}$,

$$u(x,y,Z,t)=-\frac{\sigma_0 e^{i\Omega t}}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}e^{-i(\Omega-q)\vartheta}\times J_1[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]J_1(r_0\xi)\psi_{12}(\xi,Z,q)d\vartheta d\xi dq \tag{89a}$$

$$w(x,y,Z,t)=-\frac{\sigma_0 e^{i\Omega t}}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty}e^{-i(\Omega-q)\vartheta}\times J_0[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]J_1(r_0\xi)\psi_{22}(\xi,Z,q)d\vartheta d\xi dq \tag{89b}$$

$$\tau_{rz}(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-i(\Omega-q)\vartheta} \times J_1[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]J_1(r_0\xi)\psi_{32}(\xi,Z,q)d\vartheta d\xi dq \quad (89c)$$

$$\sigma_z(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{2\pi^2 r_0}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-i(\Omega-q)\vartheta} \times J_0[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]J_1(r_0\xi)\psi_{42}(\xi,Z,q)d\vartheta d\xi dq \quad (89d)$$

Let $x-vt+v\vartheta=|y|R$; the above response field can be rewritten as

$$u(x,y,Z,t) = -\frac{\sigma_0 \exp(i\Omega x/v)}{2\pi^2 r_0 v} \times \int_{-\infty}^{\infty}\int_0^{\infty}\int_{-\infty}^{x-vt/|y|} \exp\{i[(\Omega-q)|y|R+(x-vt)q]/v\}\} \times J_1\left(\frac{\xi}{|y|}\sqrt{1+R^2}\right)J_1(r_0\xi)\psi_{12}(\xi,Z,q)dRd\xi dq \quad (90a)$$

$$w(x,y,Z,t) = -\frac{\sigma_0 \exp(i\Omega x/v)}{2\pi^2 r_0 v} \times \int_{-\infty}^{\infty}\int_0^{\infty}\int_{-\infty}^{x-vt/|y|} \exp\{i[(\Omega-q)|y|R+(x-vt)q]/v\}\} \times J_0\left(\frac{\xi}{|y|}\sqrt{1+R^2}\right)J_1(r_0\xi)\psi_{22}(\xi,Z,q)dRd\xi dq \quad (90b)$$

$$\tau_{rz}(x,y,Z,t) = -\frac{\sigma_0 \exp(i\Omega x/v)}{2\pi^2 r_0 v} \times \int_{-\infty}^{\infty}\int_0^{\infty}\int_{-\infty}^{x-vt/|y|} \exp\{i[(\Omega-q)|y|R+(x-vt)q]/v\}\} \times J_1\left(\frac{\xi}{|y|}\sqrt{1+R^2}\right)J_1(r_0\xi)\psi_{32}(\xi,Z,q)dRd\xi dq \quad (90c)$$

$$\sigma_z(x,y,Z,t) = -\frac{\sigma_0 \exp(i\Omega x/v)}{2\pi^2 r_0 v} \times \int_{-\infty}^{\infty}\int_0^{\infty}\int_{-\infty}^{x-vt/|y|} \exp\{i[(\Omega-q)|y|R+(x-vt)q]/v\}\} \times J_0\left(\frac{\xi}{|y|}\sqrt{1+R^2}\right)J_1(r_0\xi)\psi_{42}(\xi,Z,q)dRd\xi dq \quad (90d)$$

A special case is when the moving harmonic load is a point load. In this case, by taking limit $r_0 \to 0$ to Eqs. (89*a*)–(89*d*) one obtains

$$u(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{4\pi^2}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-i(\Omega-q)\vartheta} \times \xi J_1[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\psi_{12}(\xi,Z,q)d\vartheta d\xi dq \quad (91a)$$

$$w(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{4\pi^2}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-i(\Omega-q)\vartheta} \times \xi J_0[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\psi_{22}(\xi,Z,q)d\vartheta d\xi dq \quad (91b)$$

$$\tau_{rz}(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{4\pi^2}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-i(\Omega-q)\vartheta} \times \xi J_1[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\psi_{32}(\xi,Z,q)d\vartheta d\xi dq \quad (91c)$$

$$\sigma_z(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{4\pi^2}\int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-i(\Omega-q)\vartheta}\xi \times J_0[\xi\sqrt{(x-vt+v\vartheta)^2+y^2}]\psi_{42}(\xi,Z,q)d\vartheta d\xi dq \quad (91d)$$

In the derivation of Eqs. (91*a*)–(91*d*) the following limit is used:

$$\lim_{\varepsilon\to 0}\frac{J_1(\varepsilon)}{\varepsilon} = \frac{1}{2} \quad (92)$$

Another special case occurs when the moving distributed harmonic load is stationary, which corresponds to $v=0$ and leads to

$$u(r,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{\pi r_0}\int_0^{\infty} J_1(\xi r)J_1(r_0\xi)\psi_{12}(\xi,Z,\Omega)d\xi \quad (93a)$$

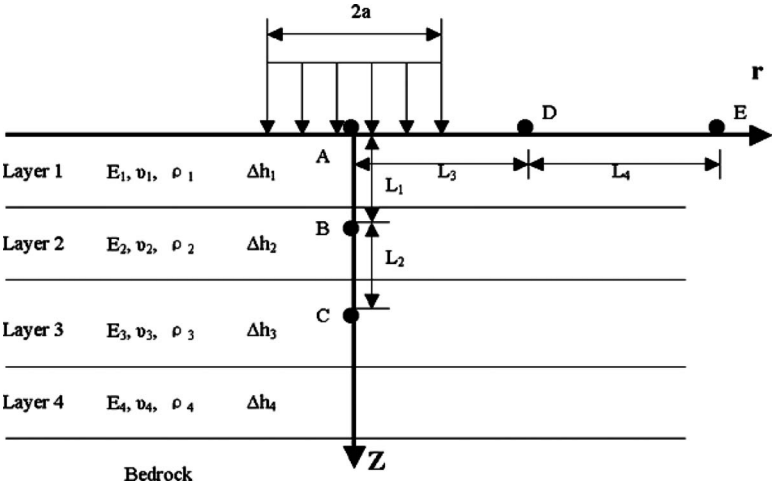


Fig. 4 A four-layer pavement used for numerical study

$$w(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{\pi r_0}\int_0^\infty J_0(\xi r)J_1(r_0\xi)\psi_{22}(\xi,Z,\Omega)d\xi \quad (93b)$$

$$\tau_{rz}(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{\pi r_0}\int_0^\infty J_1(\xi r)J_1(r_0\xi)\psi_{32}(\xi,Z,\Omega)d\xi \quad (93c)$$

$$\sigma_z(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{\pi r_0}\int_0^\infty J_0(\xi r)J_1(r_0\xi)\psi_{42}(\xi,Z,\Omega)d\xi \quad (93d)$$

In the derivation of Eqs. (93*a*)–(93*d*) the following equality and the property of Dirac delta function (3) are used:

$$\int_0^\infty e^{-i(\Omega-q)\vartheta}d\vartheta = 2\pi\delta(q-\Omega) \quad (94)$$

Furthermore, the position-fixed harmonic load is a point load, which corresponds to $r_0=0$. By taking limit $r_0\to 0$ to Eqs. (93*a*)–(93*d*) one obtains

$$u(r,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{2\pi}\int_0^\infty \xi J_1(\xi r)\psi_{12}(\xi,Z,\Omega)d\xi \quad (95a)$$

$$w(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{2\pi}\int_0^\infty \xi J_0(\xi r)\psi_{22}(\xi,Z,\Omega)d\xi \quad (95b)$$

$$\tau_{rz}(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{2\pi}\int_0^\infty \xi J_1(\xi r)\psi_{32}(\xi,Z,\Omega)d\xi \quad (95c)$$

Table 2 Values of parameters used in numerical study

Parameters	Values
E_1	0.1794×10^{10} N/m^2 (0.26 Gpsi)
E_3	0.0138×10^{10} N/m^2 (20 kpsi)
ν_1	0.35
ν_3	0.35
ρ_1	140 pcf (2.2426×10^3 kg/m^3)
ρ_3	125 pcf (2.0023×10^3 kg/m^3)
Δh_1	0.1524 m (6 in.)
Δh_3	0.3048 m (12 in.)
T_0	0.001 s
σ_0	0.6296×10^5 N/m^2 (P=2000 lb)
L_1	0.2 m
L_3	20 m
Ω	16π rad/s
E_2	0.138×10^{10} N/m^2 (0.2 Gpsi)
E_4	0.00552×10^{10} N/m^2 (8 kpsi)
ν_3	0.35
ν_3	0.3
ρ_2	125 pcf (2.0023×10^3 kg/m^3)
ρ_4	110 pcf 1.762×10^3 kg/m^3)
Δh_2	0.2032 m (8 in.)
Δh_4	0.762 m (30 in.)
T_1	0.001 s
a	0.15 m
L_2	0.2 m
L_4	20 m

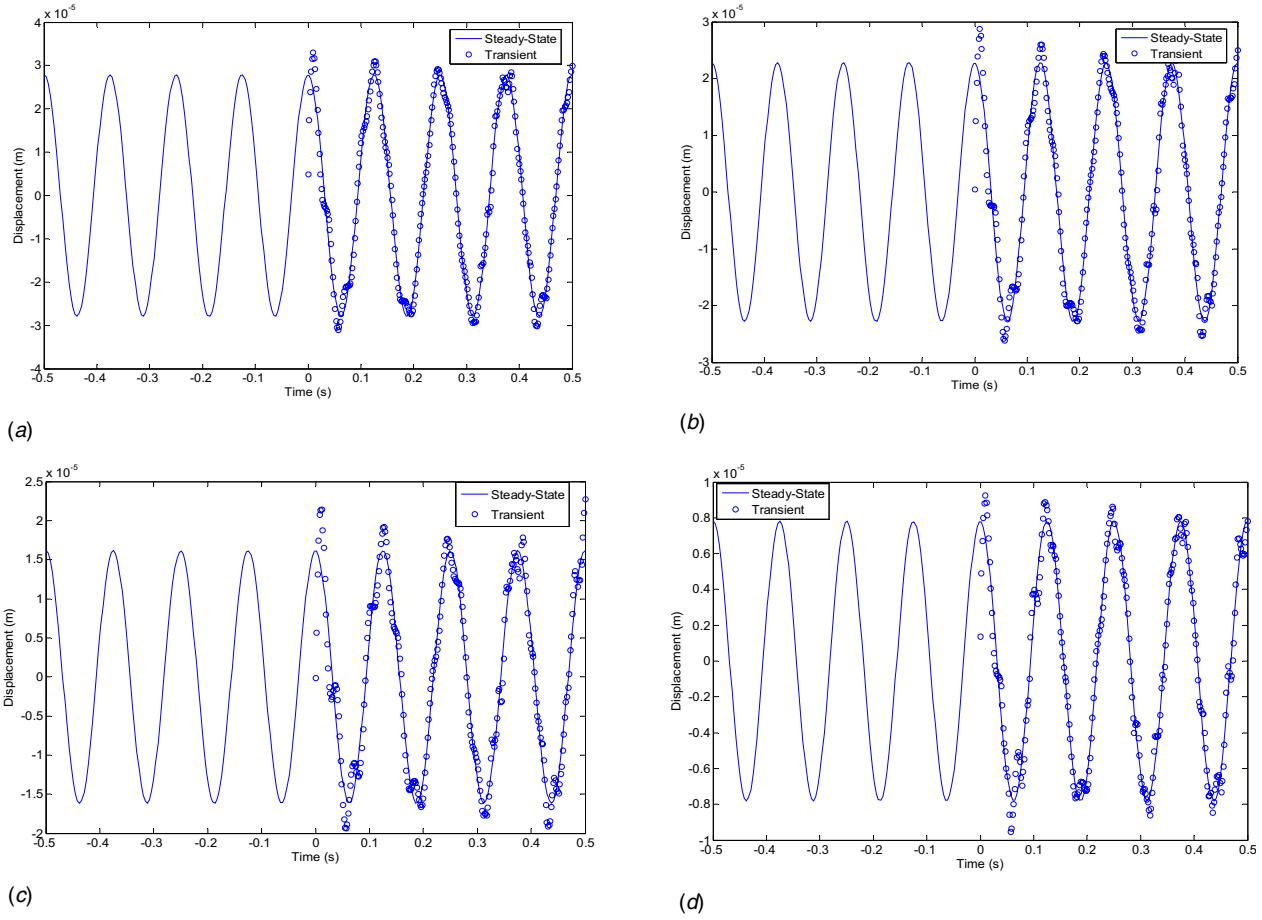
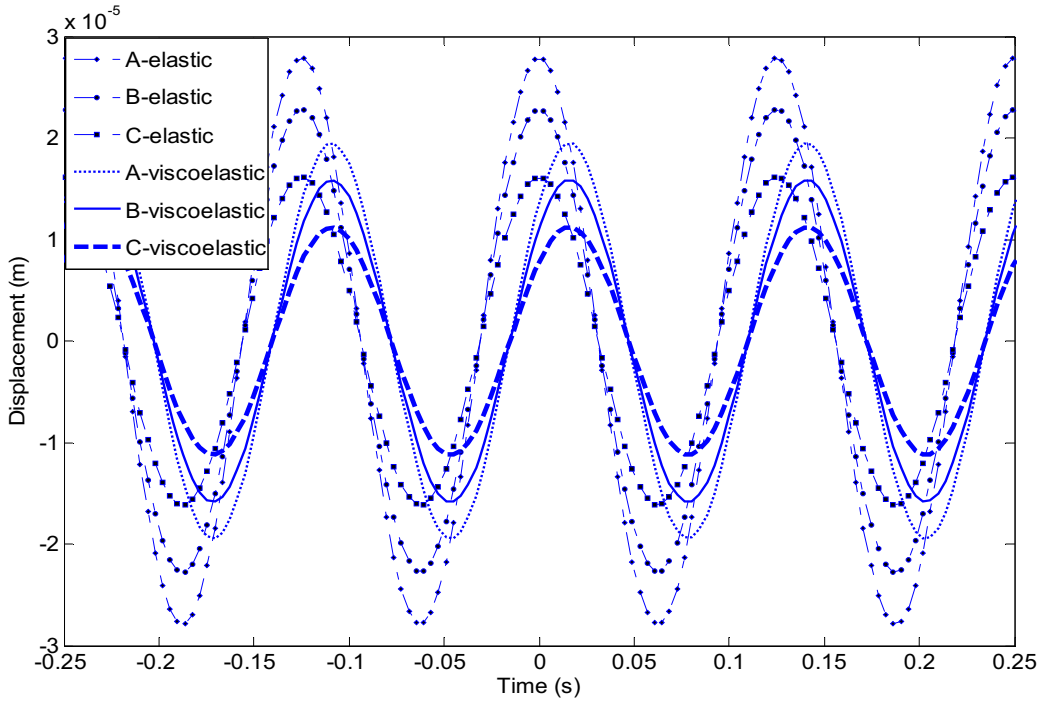
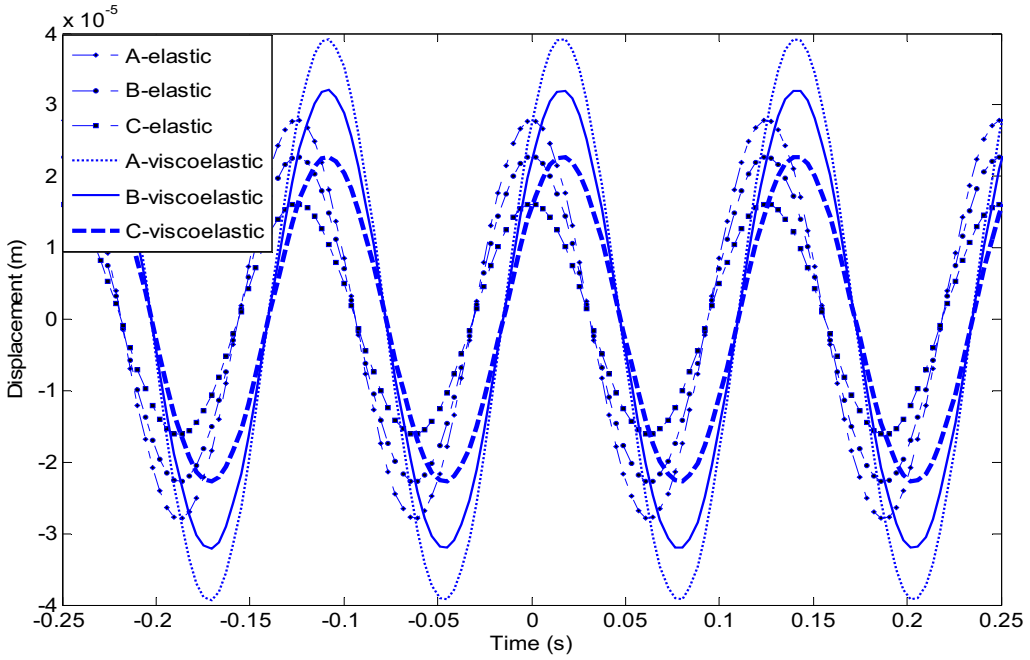


Fig. 5 Vertical displacements and stresses at different locations



(*a*)



(*b*)

Fig. 6 Comparison of elastic and viscoelastic responses

$$\sigma_z(x,y,Z,t) = -\frac{\sigma_0 e^{i\Omega t}}{2\pi}\int_0^\infty \xi J_0(\xi r)\psi_{42}(\xi,Z,\Omega)d\xi \qquad (95d)$$

6 Numerical Computation and Validation

Computation of dynamic response of multilayered viscoelastic media subject to a moving distributed load requires numerical evaluation of Eqs. (87*a*)–(87*d*), which involves the triple integrals in the infinite and semi-infinite intervals. These integrals are highly oscillatory in nature as they contain Bessel functions in the infinite and semi-infinite intervals. Such integrals require a great deal of sophistication in the numerical computation. There are special algorithms available for effective evaluation of such integrals [38]. The algorithms adopted in this paper are the same as those used in Ref. [8], which are then coded and implemented using MATLAB.

Dynamic response of a medium is influenced by a number of factors, such as material properties, number of layers, velocity of

the load, and spatial distribution of the load. We provide a validation study in this section by comparing numerical result of this paper for a position-fixed harmonic load against known results in the literature. A four-layer viscoelastic medium is used in this paper for case study. A typical highway and airport pavement structure consists of four layers: an asphalt concrete surface layer, a sub-base layer, a subgrade layer made of granular materials, and a soil foundation on top of rigid bedrock. A vertical harmonic circular load given by Eq. (96) is applied on the free surface of a four-layer viscoelastic medium on bedrock as in Fig. 4.

$$P(r,0,t) = \sigma_0 e^{i\Omega t} H(r_0 - r) \tag{96}$$

Parameters of each layer and the setting of the problem are given in Table 2 and Fig. 4.

Sun and Luo [8] solved transient wave propagation problem in a multilayered viscoelastic medium subject to an arbitrary dynamic load. The derived solution has been verified using analytical solutions in degenerated condition as well as results obtained from the finite element analysis. For comparison, a multilayered medium is assumed to be stationary at time $t=0$, and then a sudden harmonic circular load corresponding to stress boundary condition $\sigma_z(r,h_1,t)|_{h_1=0} = -\sigma_0 e^{i\Omega t} H(a-r) H(t)$ is exerted. If the formulated solution in this paper is correct, after an initial transition phase, transient response of the medium should become stable and identical to steady-state response of the same medium under Eq. (96). The transient response is obtained using a proven computer program DYNALAYER^T [8], while the steady-state response is obtained using DYNAMOVE developed in this paper. Responses of the medium at four locations are investigated: points A–D, as indicated in Fig. 4.

Figure 5 shows vertical displacement at points A–D computed for both steady-state and transient responses of a multilayered elastic medium. In this figure continuous curves stand for steady-state response, while discrete dots stand for transient response. It can be seen that after a very short duration, say, 0.025 s, the transition phenomenon dies out and the transient response approaches the steady-state response very quickly, and the match between them is almost perfect. It is also observed that amplitudes of displacement response at points A,and B decrease gradually as depth z increases. It is also observed that the amplitude of displacement response at point A is greater than that at point D. Both observations are consistent with the intuition that response gets reduced as the point of interest becomes away (horizontally or vertically) from the load.

To compare steady-state responses of the same four-layered elastic and viscoelastic media under a position-fixed harmonic circular load as studied in Fig. 5, numerical computation is carried out. In this comparison, the surface layer of this four-layer medium is viscoelastic and other layers remain elastic. Two types of viscosity are used: Kelvin model and Maxwell model. Figure 6 shows vertical displacement at points A–C for Kelvin model and for Maxwell model, respectively. In both plots in Fig. 6, there is a significant phase delay of displacement response of the viscoelastic medium compared with that of the elastic medium. Also, amplitudes of displacement response at points A–C decrease gradually. In addition, for a Kelvin viscoelastic model, amplitudes of displacement response of the elastic medium at points A–C are greater than those of the viscoelastic medium. Whereas for a Maxwell viscoelastic model, the situation is reverse, that is, amplitudes of displacement response of the elastic medium at points A–C are less than those of the viscoelastic medium. It should be noted that these two observations are only valid for the parameters used in this numerical study and may not necessarily represent a universal conclusion. It is also found that, in both elastic and viscoelastic multilayered media, the farther away the field point of interest is from the center of the source load, the smaller the amplitude of dynamic displacement response will be. This is consistent with the intuition.

7 Conclusions

In this study analytical solution of steady-state dynamic response of a multilayered viscoelastic medium to a moving distributed load is obtained. Efficient numerical algorithms based on fast evaluation of various integral transformations and their inversions are developed and validated through numerical example.

Acknowledgment

This research is sponsored in part by the National Science Foundation (NSF) under Grant No. CMMI-0408390 and NSF CAREER Award No. CMMI-0644552, by the American Chemical Society Petroleum Research Foundation under Grant No. PRF-44468-G9, by the Changjiang Scholarship of Ministry of Education of China, and by the Huoyingdong Educational Foundation, to which the authors are very grateful. The authors are also thankful to Professor Anthony Waas, Associate Editor, and two anonymous reviewers for their insightful and constructive comments, which greatly help them in improving the content of the original manuscript.

References

[1] Huang, Y. H., 2004, *Pavement Analysis and Design*, Pearson Education, Upper Saddle River, NJ.
[2] Sun, L., 1996, "Theoretical and Experimental Studies on Vehicle Vibration Induced Moving Stochastic Loads and Dynamic Response of Solids and Structures Under Moving Arbitrary Source," Ph.D. thesis, Southeast University, Nanjing, China.
[3] Sun, L., and Greenberg, B. S., 2000, "Dynamic Response of Linear Systems to Moving Stochastic Sources," J. Sound Vib., **229**(4), pp. 957–972.
[4] Sun, L., 2001, "Closed-Form Representation of Beam Response to Moving Line Loads," ASME J. Appl. Mech., **68**(2), pp. 348–350.
[5] Sun, L., 2002, "A Closed-Form Solution of Beam on Viscoelastic Subgrade Subjected to Moving Loads," Comput. Struct., **80**(1), pp. 1–8.
[6] Sun, L., 2003, "An Explicit Representation of Steady State Response of a Beam Resting on an Elastic Foundation to Moving Harmonic Line Loads," Int. J. Numer. Analyt. Meth. Geomech., **27**, pp. 69–84.
[7] Sun, L., 2007, "Steady State Dynamic Response of a Kirchhoff's Slab on a Viscoelastic Kelvin's Foundation to Moving Harmonic Loads," ASME J. Appl. Mech., **74**(6), pp. 1212–1224.
[8] Sun, L., and Luo, F. 2008, "Transient Wave Propagation in Multilayered Viscoelastic Media—Theory, Numerical Computation and Validation," ASME J. Appl. Mech., **75**(3), p. 031007.
[9] Liang, R., and Zeng, S., 2002, "Efficient Dynamic Analysis of Multilayered System During Falling Weight Deflectometer Experiments," J. Transp. Eng., **128**(4), pp. 366–374.
[10] Thomson, W. T., 1950, "Transmission of Elastic Waves Through a Stratified Solid Media," J. Appl. Phys., **21**, pp. 89–93.
[11] Haskell, N. A., 1953, "The Dispersion of Surface Waves on Multilayered Media," Bull. Seismol. Soc. Am., **43**, pp. 17–34.
[12] Kausel, E., and Peek, R., 1982, "Dynamic Loads in the Interior of a Layered Stratum: An Explicit Solution," Bull. Seismol. Soc. Am., **72**, pp. 1459–1481.
[13] Kennett, B. L. N., 1983, *Seismic Wave Propagation in Stratified Media*, Cambridge University Press, Cambridge.
[14] Chapel, F., 1987, "Boundary Element Method Applied to Linear Soil-Structure Interaction on a Heterogeneous Soil," Earthquake Eng. Struct. Dyn., **15**(7), pp. 815–829.
[15] Siddharthan, R., 1993, "Moving Load Responses of Layered Soil. I: Formulation," J. Eng. Mech., **119**(10), pp. 2052–2071.
[16] Theodorakopoulos, D. D., Chassiakos, A. P., and Beskos, D. E., 2004, "Dynamic Effects of Moving Load on a Poroelastic Soil Medium by an Approximate Method," Int. J. Solids Struct., **41**, pp. 1801–1822.
[17] Jin, B., Yue, Z. Q., and Tham, L. G., 2004, "Stresses and Excess Pore Pressure Induced in Saturated Poroelastic Half-Space by Moving Line Load," Soil Dyn. Earthquake Eng., **24**(1), pp. 25–33.
[18] De Barros, F. C. P., and Luco, J. E., 1995, "Stresses and Displacements in a Layered Half-Space for a Moving Line Load," Appl. Math. Comput., **67**, pp. 103–134.
[19] Lefeuve-Mesgouez, G., Peplow, A., and Le Houédec, D., 2000, "Ground Vibration in the Vicinity of a High-Speed Moving Harmonic Strip Load," J. Sound Vib., **231**(5), pp. 1289–1309.
[20] Le Houédec, D., 2001, "Modelling and Analysis of Ground Vibration Problems: A Review," *Civil and Structural Engineering Computing*, Saxe-Coburg Publications, Edinburgh, ROYAUME-UNI, pp. 475–485.
[21] Pak, R. Y. S., and Guzina, B. B., 1995, "Three-Dimensional Wave Propagation Analysis of a Smoothly Heterogeneous Solid," J. Mech. Phys. Solids, **43**(4), pp. 533–551.
[22] Ditri, J. J., 1997, "Determination of Nonuniform Stresses in an Isotropic Elastic Half Space From Measurements of the Dispersion of Surface Waves," J. Mech. Phys. Solids, **45**(1), pp. 51–66.
[23] Dieterman, H. A., and Metrikine, A., 1997, "Critical Velocities of a Harmonic

Load Moving Uniformly Along an Elastic Layer," ASME J. Appl. Mech., **64**(3), pp. 596–600.
[24] Zhao, H., and Gary, G., 1995, "A Three Dimensional Analytical Solution of the Longitudinal Wave Propagation in an Infinite Linear Viscoelastic Cylindrical Bar. Application to Experimental Techniques," J. Mech. Phys. Solids, **43**(8), pp. 1335–1348.
[25] Benatar, A., Rittel, D., and Yarin, A. L., 2003, "Theoretical and Experimental Analysis of Longitudinal Wave Propagation in Cylindrical Viscoelastic Rods," J. Mech. Phys. Solids, **51**(8), pp. 1413–1431.
[26] Christensen, R. M., 2003, *Theory of Viscoelasticity*, 2nd ed., Academic, New York.
[27] Chatti, K., Lysmer, J., and Monismith, C. L., 1994, "Dynamic Finite Element Analysis of Jointed Concrete Pavements," Transp. Res. Rec., **1449**, pp. 79–90.
[28] Adam, M., Pflanz, G., and Schmid, G., 2000, "Two- and Three-Dimensional Modelling of Half-Space and Train-Track Embankment Under Dynamic Loading," Soil Dyn. Earthquake Eng., **19**, pp. 559–573.
[29] Koh, C. G., Chiew, G. H., and Lim, C. C., 2007, "A Numerical Method for Moving Load on Continuum," J. Sound Vib., **300**, pp. 126–138.
[30] Andersen, L., Nielsen, S. R. K., and Krenk, S., 2007, "Numerical Methods for Analysis of Structure and Ground Vibration From Moving Loads," Comput. Struct., **85**(1–2), pp. 43–58.
[31] Eringen, A. C., and Suhubi, E. S., 1975, *Elastodynamics—Linear Theory*, Academic, New York.
[32] Roesset, J. M., and Shao, K. Y., 1985, "Dynamic Interpretation of Dynaflect and Falling Weight Deflectometer Tests," Transp. Res. Rec., **1022**, pp. 7–16.
[33] Schiffman, R. L., 1962, "General Solution of Stresses and Displacements in Layered Elastic Systems," International Conference on the Structural Design of Asphalt Pavement Proceedings, University of Michigan, Ann Arbor, MI.
[34] Magnuson, A. H., Lytton, R. L., and Briggs, R. C., 1991, "Comparison of Computer Predictions and Field Data for Dynamic Analysis of Falling Weight Deflectometer Data," Transp. Res. Rec., **1293**, pp. 61–71.
[35] Shamalta, M., and Metrikine, A. V., 2003, "Analytical Study of the Dynamic Response of an Embedded Railway Track to a Moving Load," Arch. Appl. Mech., **73**(1–2), pp. 131–146.
[36] Al-Khoury, R., Scarpas, A., Kasbergen, C., Blaauwendraad, J., and van Gurp, C., 2001, "Forward and Inverse Models for Parameter Identification of Layered Media," Int. J. Geomech., **1**(4), pp. 441–458.
[37] Tschoegl, N., 1989, *The Phenomenological Theory of Linear Viscoelasticity*, Springer-Verlag, Berlin.
[38] Candel, S. M., 1981, "Simultaneous Calculations of Fourier-Bessel Transforms Up to Order N," J. Comput. Phys., **44**, pp. 243–261.

A Thin Conducting Liquid Film on a Spinning Disk in the Presence of a Magnetic Field: Dynamics and Stability

B. Uma
Center for Risk Studies and Safety,
University of California,
Santa Barbara,
Goleta, CA 93117
e-mail: ubalakrishnan@engr.ucsb.edu

R. Usha
Department of Mathematics,
Indian Institute of Technology Madras,
Chennai 600 036, India
e-mail: ushar@iitm.ac.in

A theoretical analysis of the effects of a magnetic field on the dynamics of a thin nonuniform conducting film of an incompressible viscous fluid on a rotating disk has been considered. A nonlinear evolution equation describing the shape of the film interface has been derived as a function of space and time and has been solved numerically. The temporal evolution of the free surface of the fluid and the rate of retention of the liquid film on the spinning disk have been obtained for different values of Hartmann number M, evaporative mass flux parameter E, and Reynolds number Re. The results show that the relative volume of the fluid retained on the spinning disk is enhanced by the presence of the magnetic field. The stability characteristics of the evolution equation have been examined using linear theory. For both zero and nonzero values of the nondimensional parameter describing the magnetic field, the results show that (a) the infinitesimal disturbances decay for small wave numbers and are transiently stable for larger wave numbers when there is either no mass transfer or there is evaporation from the film surface, and although the magnitude of the disturbance amplitude is larger when the magnetic field is present, it decays to zero earlier than for the case when the magnetic field is absent, and (b) when absorption is present at the film surface, the film exhibits three different domains of stability: disturbances of small wave numbers decay, disturbances of intermediate wave numbers grow transiently, and those of large wave numbers grow exponentially. The range of stable wave numbers increases with increase in Hartmann number. [DOI: 10.1115/1.3086589]

1 Introduction

The flow of a thin film of a viscous fluid over a smooth rotating disk has attracted the attention of several investigations in science and engineering due to its enormous applications in many industrial processes that range from the intensification of heat and mass transfer processes in chemical reactors to powder production in metallurgy. The production of thin films on substrates placed in the grooves of a rotating disk is referred to as "spin coating" in the literature and this technique is employed in coating: very thin and uniform films of photoresist on silicon wafers for integrated circuits, very thin layers of magnetic paint, magnetic storage disks, fabrication of thin uniform layers of plastic scintillator on supporting aluminized Mylar, and so on. In spite of the difficulties in modeling this flow mathematically due to the variation in acceleration along the radius, the flow over a spinning disk has lent itself more naturally to potential technological exploitation due to the possibility of controlling the local accelerations. The final thickness of the film and the uniformity in the thickness are central issues in these applications and they are observed to be influenced by several factors such as the viscosity of the liquid film, different spin-up protocols, heat and mass transfer processes, surface tension effects, and so on.

Since the pioneering study by Emslie et al. [1] on the hydrodynamic analysis of the flow of a Newtonian fluid on a spinning disk, a number of theoretical and experimental studies of the spin-coating process that include modeling of flow over a rotating disk, wave generation in a liquid film moving on the surface of a rotating disk, and stability characteristics of a thin film on a rotating substrate have been reported [2–12]. Reisfeld et al. [12] considered the dynamics and stability of thin liquid films during spin coating with constant rates of evaporation or absorption. They have derived a nonlinear evolution equation describing the shape of the film interface as a function of space and time and examined its stability using linear theory. Their results show that when there is either no mass transfer or there is evaporation from the film surface, infinitesimal disturbances decay for small wave numbers and are transiently stable for larger wave numbers. When absorption is present at the free surface, the film exhibits three different domains of stability: disturbances of small wave numbers decay, disturbances of intermediate wave numbers grow transiently, and those of large wave numbers grow exponentially. The experimental investigations of flow over a spinning disk have considered the measurements of the local maximum or local mean film thickness [13–21] and have determined the characteristics of flow over a spinning disk. These studies have led to modified versions of Nusselt formula that have been developed to fit the observations.

Modeling studies of the flow over a spinning disk have considered the stationary axisymmetric waveless flow in the limit of large Eckman number E, where $E=\nu/\Omega H_c^2$, H_c is a characteristic film thickness [20,22,23], and steady axisymmetric solutions for finite Eckman numbers in the framework of the boundary layer approximation [24–27].

The linear stability characteristics of flow over a spinning disk has been addressed by several investigations that include the studies by Charwat et al. [14] for the case of large Eckman numbers, Sisoev and Shkadov [27] for arbitrary Eckman numbers, and nonaxisymmetric disturbances and Needham and Merkin [28] for large Eckman numbers in the presence of gravity for unsteady flow driven by perturbations applied to the flux. Woods [20] considered a linear stability analysis of spiral perturbations using lu-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 6, 2008; final manuscript received December 9, 2008; published online April 21, 2009. Review conducted by Nesreen Ghaddar.

brication approach. Sisoev et al. [10] extended the methods and the results developed for the falling film problem [22] to model the axisymmetric flow regimes observed to accompany the flow of a thin liquid film over a spinning disk for a wide range of system parameters.

There have also been investigations incorporating the influence of heat transfer in the liquid phase and the thermocapillary force at the liquid-gas interface on the development of flow and thickness of the liquid film on a spinning disk, which arise due to the temperature difference between the disk and the ambient gas and the initial liquid [29–34].

The investigations on flow and dynamics of liquid films on a spinning substrate mentioned above reveal that the rate of film thinning slows down beyond a specific height (depending on the rotational speed) of the film. It has been observed, in general, that the final stage of the film thickness is inversely proportional to the square root of spinning time for τ (spinning time) $\rightarrow\infty$, so that the spinner has to be operated for quite a long time in order to obtain the desired thickness of the film. This is due to the action of the centrifugal force, which pushes the liquid radially outward from the surface of the disk. The film thins progressively and there is an increase in viscous resistance with thinning, which in turn decreases the outward radial velocity. The radial flow practically ceases after a sufficient lapse of time and during this period; the chief mechanism of mass loss is due to evaporation only. A solid skin is formed on the surface layer that puts greater resistance to the remaining liquid for evaporation. Coating defects occur if the convective flux does not completely cease before this skin hardens sufficiently.

In order to avoid coating defects, one must look for mechanisms that may either increase the rate of thinning or offer high resistance to film thinning. It has been shown [34] that a film thickness inversely proportional to spinning time τ could be obtained, for $\tau\rightarrow\infty$, by imposing a specified axisymmetric temperature distribution on the disk. The results indicate that the rate of thinning of the film could be accelerated and a desired thickness of the film before the hardening of the skin could be obtained on the spinning disk. Another way of increasing the rate of film thinning has been noticed by Strong and Middleman [35] and Rehg and Higgins [3], namely, that the rate of film thinning increases due to the shear induced by air flow over the surface of the film

A mechanism that offers high resistance to film thinning has been indicated by Ray and Dandapat [36], who showed that a magnetic field normal to the substrate puts greater resistance on film thinning right from the beginning of rotation. They have further observed [37] that the thermocapillary force has a profound effect in enhancing the thinning rate of the film, even in the presence of a large Hartmann number. The stability is based on the assumption of a small Reynolds number and uniform rotation of the disk. The effects of large Reynolds number, as well as nonuniform rotation, have been examined in the presence of a transverse magnetic field by Dandapat and Layek [38]. They predicted that an initial impulse rotation followed by an accelerated angular velocity until the desired film thickness is obtained would increase the rate of thinning. Furthermore, they have concluded that increasing the angular velocity could decrease the likelihood of hard skinning commonly observed in spin coating. They suggested that a transverse magnetic field would stabilize the liquid flow in the spin-coating process, resulting in suppressed nonuniformities on the film surface. The numerical solution for the development of a thin conducting film on a smooth surface of a spinning disk in the presence of a magnetic field, including inertial effects, has been obtained [39] for different spin-up protocols and it has been shown that the film thickness increases with the increase in Hartmann number, M, and the rate of depletion is more for small M than for large M. The inclusion of a magnetic field in the investigations [40,41] has some significance insofar as a magnetic field is likely to suppress any irregularities on the film surface by exerting a stabilizing influence on the flow. The above investigations have not examined the instabilities that arise during spin coating, in the presence of a transverse magnetic field.

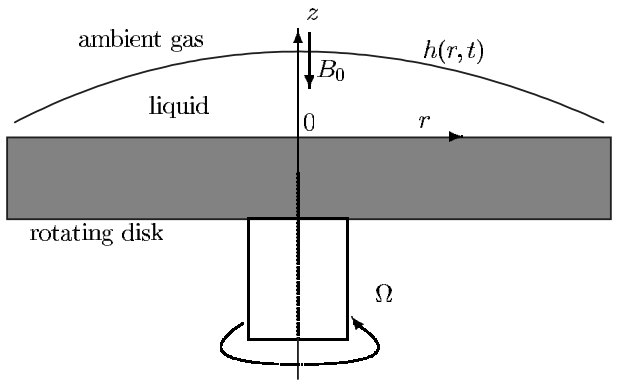


Fig. 1 Schematic representation of film flow on a rotating disk

In view of this, the present study addresses the dynamics and stability characteristics of a viscous conducting film over a spinning disk in the presence of a transverse magnetic field and it extends the investigation by Reisfeld et al. [12]. A model equation describing the temporospatial behavior of the film interface valid for spin-coating applications in the presence of a transverse magnetic field has been derived using a lubrication approximation. A linear stability analysis of this model equation has been considered and the conditions for film instability in several regimes of interest have been derived. The results show that infinitesimal disturbances decay for small wave numbers for different values of Hartmann number considered and that they are transiently stable for larger wave numbers, when there is either no mass transfer or there is evaporation from the film. On the other hand, when absorption is present at the film surface, the film exhibits three different domains of stability: disturbances of small wave numbers decay, those of intermediate wave numbers remain transiently stable, and large wave number disturbances grow exponentially. It has been observed that the range of stable wave numbers increases with increase in Hartmann number M showing that the effect of a magnetic field is to stabilize the film flow system.

2 Mathematical Formulation

An axisymmetric flow of a viscous electrically conducting fluid with constant properties on a rotating disk is considered. A system of cylindrical coordinates (r,θ,z) that rotates with the disk at an angular velocity Ω about the z-axis is used, where r measures the radial distance from the center of the disk, θ is the angle from some fixed radial line in the horizontal plane, and z measures the distance vertically upward from the solid surface of the disk. A uniform magnetic field B_0 acts parallel to the axis of rotation of the disk (Fig. 1).

In this investigation, a flow model that has been widely used in previous magnetohydrodynamics (MHD) boundary layer studies has been employed. First, the externally applied electric field has been arranged to be zero (e.g., with the use of an excellently conducting disk and other suitable wiring [42]). Then it is assumed that within the viscous boundary layer, the electric and magnetic fields **E** and **B** are not significantly different from their values at the disk surface [42], which implies that the induced electric and magnetic fields are negligibly small but not identically zero. In other words, the viscous boundary layer is assumed to be much thinner than the magnetic boundary layer. It is worth mentioning here that such simplifications have been verified for stagnation point flows by Neuringer and McIlroy [43], Rossow [44], and Rathbun [45], who found that the induced electric field and magnetic field can be neglected if the magnetic Prandtl number $\mathrm{Pr}_m=\sigma_e\mu_0\nu\ll 1$, where μ_0 is the magnetic permeability, σ_e is

the electrical conductivity, and ν is the kinematic viscosity (Pr_m can also be regarded as the ratio of fluid viscosity ν to the magnetic viscosity $(\mu_0\sigma_e)^{-1}$). For ionized air $\mathrm{Pr}_m \simeq 10^{-5}-10^{-6}$ and for liquid mercury at room temperature and electrolytes [43,44], $\mathrm{Pr}_m = 10^{-7}$. Hence the induced electric and magnetic fields can be neglected for a wide range of fluids. Sparrow and Cess [42] argued that the same conclusions are also valid for the flow over a spinning disk. In view of this, only the applied magnetic field plays a role and gives rise to the Lorentz force $\mathbf{G}=(G_r,G_\theta,0)$ with components $G_r=-\sigma_e B_0^2 u/\rho$ and $G_\theta=-\sigma_e B_0^2 v/\rho$ in the r and θ increasing directions, where ρ is the density and $\mathbf{v}=(u,v,w)$ is the fluid velocity vector in the cylindrical coordinate system. The liquid-gas interface is located at $z=h(r,t)$, where h is the film thickness. The outward unit normal vector, $\mathbf{n}$, and unit tangent vectors $\mathbf{t}_1$ and $\mathbf{t}_2$ are given by

$$\mathbf{n}=(-h_r,0,1)(1+h_r^2)^{-1/2},\quad \mathbf{t}_1=(1,0,h_r)(1+h_r^2)^{-1/2},\quad \mathbf{t}_2=(0,1,0) \tag{1}$$

Under the above assumptions, the continuity and the Navier–Stokes equation modified to include Lorentz forces due to magnetohydrodynamic interactions governing the unsteady flow of a liquid film on the rotating disk can be expressed as [36,39,42,46]

$$\nabla\cdot\mathbf{v}=0 \tag{2}$$

$$(\mathbf{v}_t+(\mathbf{v}\cdot\nabla)\mathbf{v})=-\frac{1}{\rho}\nabla p+\frac{1}{\rho}\nabla\cdot\boldsymbol{\tau}-[2\boldsymbol{\omega}\times\mathbf{v}+\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})+\mathbf{g}]+\mathbf{G} \tag{3}$$

where p and $\boldsymbol{\tau}$ are the pressure and the viscous stress tensor of the liquid; $\boldsymbol{\omega}=(0,0,\Omega)$ is the angular velocity vector; $\mathbf{g}=(0,0,g)$ is the gravitational vector, and $\mathbf{r}=(r,0,0)$. The boundary conditions are

$$\mathbf{v}=0\quad \text{on}\quad z=0\quad \text{(no-slip condition)} \tag{4}$$

$$\mathbf{n}\cdot\overline{\mathbf{T}}\cdot\mathbf{n}=-\nabla\cdot\mathbf{n}\quad \sigma\quad \text{on}\quad z=h(r,t)$$

$$\text{(balance of normal stress at the free surface)} \tag{5}$$

$$\mathbf{n}\cdot\overline{\mathbf{T}}\cdot\mathbf{t}_1=0\quad \text{on}\quad z=h(r,t)$$

$$\text{(balance of shear stress at the free surface)} \tag{6}$$

$$\mathbf{n}\cdot\overline{\mathbf{T}}\cdot\mathbf{t}_2=0\quad \text{on}\quad z=h(r,t)$$

$$\text{(balance of shear stress at the free surface)} \tag{7}$$

$$\rho(\mathbf{v}-\mathbf{v}^{(i)})\cdot\mathbf{n}=J\quad \text{on}\quad z=h(r,t)$$

$$\text{(kinematic boundary condition at the free surface)} \tag{8}$$

where $\overline{\mathbf{T}}$ is the stress tensor, σ is the coefficient of surface tension, J is the constant evaporative mass flux from the free surface, $\mathbf{v}^{(i)}$ is the velocity of the interface, and $\mathbf{v}^{(i)}\cdot\mathbf{n}=h_t(1+h_r^2)^{-1/2}$. It is assumed that changes in solvent concentration in the film are small and the evaporation from the interface takes place at a constant rate.

Nondimensionalizing the governing equations and the boundary conditions using dimensionless (asterisks) variables

$$h^*=\frac{h}{h_0},\quad z^*=\frac{z}{h_0},\quad r^*=\frac{r}{L},\quad t^*=\frac{t\Omega^2h_0^2}{\nu}$$

$$u^*=\frac{u\nu}{\Omega^2Lh_0^2},\quad v^*=\frac{v\nu^2}{\Omega^3Lh_0^4},\quad w^*=\frac{w\nu}{\Omega^2h_0^3},\quad p^*=\frac{p}{\rho\Omega^2L^2} \tag{9}$$

where h_0, L, and ν are the initial mean film thickness, the disk radius, and the kinematic viscosity, respectively. After scaling, the governing equations and boundary conditions are obtained as (after dropping asterisks)

$$\epsilon\ \mathrm{Re}[u_t+uu_r+wu_z]-\epsilon^2\ \mathrm{Re}^2\frac{v^2}{r}=-p_r+r+2\epsilon\ \mathrm{Re}\ v+\epsilon^2\left[u_{rr}+\frac{u_r}{r}-\frac{u}{r^2}\right]+u_{zz}-M^2u \tag{10}$$

$$\epsilon\ \mathrm{Re}\left[v_t+uv_r+\frac{uv}{r}+wv_z\right]=\epsilon^2\left[v_{rr}+\frac{v_r}{r}-\frac{v}{r^2}\right]+v_{zz}-2u-M^2v \tag{11}$$

$$\epsilon^3\ \mathrm{Re}[w_t+uw_r+ww_z]=-p_z+\epsilon^4\left[w_{rr}+\frac{w_r}{r}\right]+\epsilon^2w_{zz}-\frac{\epsilon\ \mathrm{Re}}{F^2} \tag{12}$$

$$u_r+\frac{u}{r}+w_z=0 \tag{13}$$

$$u=v=w=0\quad \text{on}\quad z=0 \tag{14}$$

$$-p+2\epsilon^2(1+\epsilon^2h_r^2)^{-1}[\epsilon^2u_rh_r^2+w_z-u_zh_r-\epsilon^2w_rh_r]=\frac{\epsilon^3\mathrm{We}}{r(1+\epsilon^2h_r^2)^{3/2}}[h_r+\epsilon^2h_r^3+rh_{rr}]\quad \text{on}\quad z=h(r,t) \tag{15}$$

$$2\epsilon^2h_r(w_z-u_r)+(1-\epsilon^2h_r^2)(u_z+\epsilon^2w_r)=0\quad \text{on}\quad z=h(r,t) \tag{16}$$

$$v_z-\epsilon^2h_r\left(v_r-\frac{v}{r}\right)=0\quad \text{on}\quad z=h(r,t) \tag{17}$$

$$(-h_t-uh_r+w)(1+\epsilon^2h_r^2)^{-1/2}=\frac{2E}{3}\quad \text{on}\quad z=h(r,t) \tag{18}$$

where $\epsilon=h_0/L$ is the aspect ratio, $\mathrm{Re}=U_0h_0/\nu$ is the Reynolds number, $U_0=\Omega^2Lh_0^2/\nu$, $F=\sqrt{U_0^2/gh_0}$ is the Froude number, $\mathrm{We}=\sigma/\rho\Omega^2Lh_0^2$ is the Weber number, M is the Hartmann number given by $M^2=\sigma B_0^2h_0^2/\mu$, and $E=3J/2\rho\epsilon U_0$ is the dimensionless mass flux.

In the present study, evaporation from the film surface is taken to be a constant throughout spinning and it is quite a reasonable approximation as long as the spinning rate is held constant. Meyerhofer [47] was the first to estimate the effect of solvent evaporation on final coating thickness. He added a constant evaporation term, which is effectively the contribution to the interface velocity that is driven by the evaporation process alone to the equation. His arguments revealed that evaporation becomes a dominant process for a critical film thickness and that it takes place in concert with film thinning as the film solidifies. At this stage, the film thickness is a function of evaporation. However, the present study is concerned with the phase of spin coating that occurs shortly after the liquid film is delivered to the disk surface and the liquid film is assumed to be flat and relatively thick (>500 μm) so that the Reynolds number for the flow is appreciable and the film thins due to radial drainage and evaporation. Furthermore, it is to be remarked that such an assumption about evaporation (as a constant) has been considered in the investigation of dynamics and stability of a liquid film on a spinning disk by Reisfeld et al. [12].

Typical values of the physical parameters and the corresponding dimensionless parameters listed in Table 1 of Ref. [12] have been used and are given by

$$h_0 = 5 \times 10^{-2} \text{ cm}, \quad L = 5 \text{ cm}; \quad \Omega = 100 \text{ s}^{-1}, \quad \sigma = 20 \text{ dynes/cm}, \quad \rho = 1.0 \text{ g/cm}^3$$

$$\nu = 1 \text{ cm}^2/\text{s}, \quad J = 1.0 \times 10^{-3} \text{ cm/s}, \quad \epsilon = 0.01, \quad \text{Re} = 6.2, \quad \text{We} = 0.15$$

$$\epsilon^2\text{We} = 1.5 \times 10^{-5}, \quad \text{Fr}^{-2} = 3.1 \times 10^{-3}, \quad E = 1.2 \times 10^{-3} \tag{19}$$

The computations have been performed for different values of Hartmann number.

3 Derivation of Evolution Equation

The dependent variables u, v, w, and p are expanded in powers of ϵ as

$$(u,v,w,p) = \sum_{n=0}^{N} \epsilon^n [u^{(n)}, v^{(n)}, w^{(n)}, p^{(n)}] \tag{20}$$

and are substituted in Eqs. (10)–(17). Equation (20) is based on the assumption that the ratio of the film thickness to the radius of the substrate (rotating disk) is exceedingly small. Such a procedure is analogous to a long-wave analysis and lubrication approximation used by Reisfeld et al. [12] for a liquid film on a spinning disk and that used by Benney [48] and Atherton and Homsy [49] for falling liquid films and those used by Williams and Davis [50] and Burelbach et al. [51] for isothermal and heated thin liquid films. The zeroth order and first order equations are

$$-p_{0r} + r + u_{0zz} - M^2 u_0 = 0$$

$$v_{0zz} - 2u_0 - M^2 v_0 = 0$$

$$-p_{0z} = 0$$

$$u_{0r} + \frac{u_0}{r} + w_{0z} = 0 \tag{21}$$

$$u_0 = 0, \quad v_0 = 0, \quad w_0 = 0 \quad \text{on} \quad z = 0 \tag{22}$$

$$-p_0 = 0, \quad u_{0z} = 0, \quad v_{0z} = 0 \quad \text{on} \quad z = h \tag{23}$$

and

$$\text{Re}[u_{0t} + u_0 u_{0r} + w_0 u_{0z}] = -p_{1r} + 2 \text{ Re } v_0 + u_{1zz} - M^2 u_1$$

$$\text{Re}\left[v_{0t} + u_0 v_{0r} + \frac{u_0 v_0}{r} + w_0 v_{0z}\right] = v_{1zz} - 2u_1 - M^2 v_1$$

$$-p_{1z} - \frac{\text{Re}}{F^2} = 0$$

$$u_{1r} + \frac{u_1}{r} + w_{1z} = 0 \tag{24}$$

$$u_1 = 0, \quad v_1 = 0, \quad w_1 = 0 \quad \text{on} \quad z = 0 \tag{25}$$

$$p_1 = -\epsilon^2 \text{We}\left[h_{rr} + \frac{h_r}{r}\right] \quad \text{on} \quad z = h \tag{26}$$

$$u_{1z} = 0 \quad \text{on} \quad z = h \tag{27}$$

$$v_{1z} = 0 \quad \text{on} \quad z = h \tag{28}$$

The solutions of Eqs. (21)–(28) give the radial and axial velocities and are presented in Appendix. Substituting Eqs. (A1) and (A2) into the kinematic boundary condition at the free surface given by Eq. (18), the evolution equation is obtained as

$$\begin{aligned}
&h_t + \frac{2}{3}E + \frac{2h}{M^2} - \frac{2\sinh Mh}{M^3\cosh Mh} + \frac{rh_r \sinh^2 Mh}{M^2\cosh^2 Mh} \\
&+ \epsilon \text{ Re}\left\{\frac{14\sinh Mh}{3M^7\cosh^3 Mh} + \frac{61\sinh 2Mh}{6M^7\cosh^2 Mh} - \frac{13h}{M^6\cosh^2 Mh}\right. \\
&- \frac{2h}{M^6\cosh^4 Mh} - \frac{10h}{M^6} - \frac{4h^2\sinh Mh}{M^5\cosh^3 Mh} + \frac{4E(\cosh Mh - 1)}{3M^4\cosh^2 Mh} \\
&\left. - \frac{2Eh\sinh Mh}{3M^3\cosh^3 Mh}\right\} + \epsilon \text{ Re}(rh_r)\left\{\frac{12h\sinh Mh}{M^5\cosh^3 Mh} - \frac{5}{M^6}\right. \\
&- \frac{11}{2M^6\cosh^2 Mh} + \frac{21}{2M^6\cosh^4 Mh} + \frac{4h^2}{M^4\cosh^2 Mh} \\
&- \frac{6h^2}{M^4\cosh^4 Mh} + \frac{11h\sinh Mh}{2M^5\cosh^5 Mh} - \frac{2E\sinh Mh}{3M^3\cosh^2 Mh} \\
&\left. + \frac{E\sinh Mh}{M^3\cosh^3 Mh} + \frac{2Eh}{3M^2\cosh^2 Mh} - \frac{Eh}{M^2\cosh^4 Mh}\right\} \\
&+ \epsilon \text{ Re}(r^2 h_{rr})\left\{-\frac{3\sinh^2 Mh}{2M^6\cosh^4 Mh} + \frac{h\sinh Mh}{M^5\cosh^3 Mh}\right. \\
&\left. + \frac{h\sinh Mh}{2M^5\cosh^5 Mh}\right\} + \epsilon \text{ Re}(r^2 h_r^2)\left\{-\frac{11\sinh Mh}{2M^5\cosh^5 Mh}\right. \\
&- \frac{2h}{M^4\cosh^2 Mh} + \frac{h}{M^4\cosh^4 Mh} + \frac{5h}{2M^4\cosh^6 Mh} \\
&\left. + \frac{4\sinh Mh}{M^5\cosh^3 Mh}\right\} + \left\{\frac{\epsilon \text{ Re}}{F^2M^2}\left(h_{rr} + \frac{h_r}{r}\right) - \frac{\epsilon^3\text{We}}{M^2r^3}(r^3 h_{rrrr}\right. \\
&\left. + 2r^2 h_{rrr} - rh_{rr} + h_r)\right\}\left[\frac{\sinh Mh}{M\cosh Mh} - h\right] - \left\{\frac{\epsilon \text{ Re } h_r}{F^2M^2}\right. \\
&\left. - \frac{\epsilon^3\text{We}}{M^2r^2}(r^2 h_{rrr} + rh_{rr} - h_r)\right\}\frac{h_r\sinh^2 Mh}{\cosh^2 Mh} = 0
\end{aligned} \tag{29}$$

In deriving Eq. (29), the term h_t contained within Re term is replaced by its leading order representation in terms of r, h, and its spatial derivatives. It is assumed that E, F^{-2}, and ϵ^2We are $O(1)$ as $\epsilon \to 0$. Equation (29) describes the shape of a thin liquid film interface on a rotating substrate in the presence of a magnetic field. Equation (29) describes the evolution of a thin liquid film on a rotating substrate when centripetal force, inertia, mass transfer from the free surface, magnetic field, gravity, and surface tension are all important.

Although the effect of gravity and surface tension are important in planarization studies [5,42], where the leveling of liquid films on rough rotating disk are investigated, for spin-coating applications Re/F^2 (effect of gravity) and ϵ^2We (effect of surface tension) are very small and the contribution of the terms associated with them in Eq. (29) is exceedingly small ($\simeq 10^{-5}$–10^{-8}) and hence can be neglected for further analysis. The evolution equation for further analysis is given by

$$\begin{aligned}
&h_t + \frac{2}{3}E + \frac{2h}{M^2} - \frac{2\sinh Mh}{M^3\cosh Mh} + \frac{rh_r \sinh^2 Mh}{M^2\cosh^2 Mh} \\
&+ \epsilon \text{ Re}\left\{\frac{14\sinh Mh}{3M^7\cosh^3 Mh} + \frac{61\sinh 2Mh}{6M^7\cosh^2 Mh} - \frac{13h}{3M^6\cosh^2 Mh}\right. \\
&- \frac{2h}{M^6\cosh^4 Mh} - \frac{10h}{M^6} - \frac{4h^2\sinh Mh}{M^5\cosh^3 Mh} + \frac{4E(\cosh Mh - 1)}{3M^4\cosh^2 Mh} \\
&\left. - \frac{2Eh\sinh Mh}{3M^3\cosh^3 Mh}\right\} + \epsilon \text{ Re}(rh_r)\left\{\frac{12h\sinh Mh}{M^5\cosh^3 Mh} - \frac{5}{M^6}\right. \\
&- \frac{11}{2M^6\cosh^2 Mh} + \frac{21}{2M^6\cosh^4 Mh} + \frac{4h^2}{M^4\cosh^2 Mh}
\end{aligned}$$

$$
-\frac{6h^2}{M^4\cosh^4 Mh}+\frac{11h\sinh Mh}{2M^5\cosh^5 Mh}-\frac{2E\sinh Mh}{3M^3\cosh^2 Mh}
+\frac{E\sinh Mh}{M^3\cosh^3 Mh}+\frac{2Eh}{3M^2\cosh^2 Mh}-\frac{Eh}{M^2\cosh^4 Mh}\Bigg\}
+\epsilon\ \mathrm{Re}(r^2h_{rr})\Bigg\{-\frac{3\sinh^2 Mh}{2M^6\cosh^4 Mh}+\frac{h\sinh Mh}{M^5\cosh^3 Mh}
+\frac{h\sinh Mh}{2M^5\cosh^5 Mh}\Bigg\}+\epsilon\ \mathrm{Re}(r^2h_r^2)\Bigg\{-\frac{11\sinh Mh}{2M^5\cosh^5 Mh}
-\frac{2h}{M^4\cosh^2 Mh}+\frac{h}{M^4\cosh^4 Mh}+\frac{5h}{2M^4\cosh^6 Mh}
+\frac{4\sinh Mh}{M^5\cosh^3 Mh}\Bigg\}=0 \tag{30}
$$

It is to be remarked that Eq. (30) reduces to the evolution equation derived by Resifeld et al. [12] when M goes to zero, and to those of Emslie et al. [1], Strong and Middleman [35], and Hwang and Ma [53] when M goes to zero and Re goes to zero. In what follows, the numerical solution of Eq. (30) is obtained by using Crank–Nicolson finite-difference scheme. The boundary conditions at the center of the disk $r=0$ are determined from the symmetry conditions as

$$
\frac{\partial h(r=0,t)}{\partial r}=0 \tag{31}
$$

The boundary condition at the edge of the finite disk cannot be prescribed in a simple mathematical form due to the complicated edge effect, known as the "tea-pot" effect [2,52]; further Reisfeld et al. [12] pointed out that specifying appropriate boundary conditions at the edge of the disk is difficult and that it requires an intimate understanding of the complicated time-dependent film rupture process taking place in that region. For thin films, this effect is regarded as a local phenomenon occurring at the edge and is thought to have negligible influence on the solution away from the edge [2,52]. One of the ways of prescribing the boundary conditions at the edge of the disk suggested by Tu [52], Hwang and Ma [53], and Kim et al. [54] and implemented in their analysis of thin liquid film on a spinning disk is as follows: They have assumed that the disk hypothetically extends beyond the edge as a smooth flat surface and have taken the boundary condition at the edge of the disk as follows: Let the edge of the disk be $r_{\max}$. Then a boundary condition at $r=r_{\max}$ is

$$
\frac{\partial h}{\partial r}(r=r_{\max},t)=0 \tag{32}
$$

The second way of imposing the boundary condition at the edge of the disk is to use the asymptotic properties of Eq. (30) at large values of the radius and use the corresponding results as the condition at large radial distances (that is, at the edge of the disk). The initial condition is furnished by $h(r,0)$, which gives the initial distribution of the height of the free liquid surface.

Thus, the problem of investigating the magnetic field effects on a liquid film flow during spin coating reduces to solving Eq. (30) for $h(r,t)$ subject to the boundary conditions (31) and (32) and the initial conditions. It is of interest to compute the total volume $Q(t)$ of the liquid that remains on the rotating disk and it is given by

$$
Q(t)=2\pi\int_0^{r_{\max}} rh(r,t)dr \tag{33}
$$

4 Numerical Solution and Results

The nonlinear equation (30) is solved numerically by Crank–Nicolson finite-difference scheme by approximating all the spatial derivatives by central differences [53]. The finite domain $(0,r_{\max})$ is uniformly partitioned by points $(r_0,r_1,r_2\ldots r_n=r_{\max})$, where $r_0=0$ and $r_n=r_{\max}=n(\Delta r)$. The finite-difference approximation of Eq. (30) over the discretized domain is obtained and the boundary conditions (31) and (32) are given by

$$
h_{-1}^j=h_1^j
$$

$$
h_n^j=h_{n+1}^j \tag{34}
$$

The last condition makes sense in view of the assumption that the disk extends beyond the edge as a smooth flat surface. The condition implies that beyond the edge of the disk, the film thickness remains the same at all points on the smooth disk. The coefficient matrix $\mathbf{A}$ of the algebraic system

$$
\mathbf{A}(\mathbf{V}^j,\mathbf{V}^{j+1})\mathbf{V}^{j+1}=\mathbf{f}(\mathbf{V}^j,\mathbf{V}^{j+1}) \tag{35}
$$

has a tridiagonal form, where the column vector $\mathbf{V}^{j+1}$ is given by

$$
\mathbf{V}^{j+1}=(h_0^{j+1},h_1^{j+1},\ldots,h_n^{j+1})^T \tag{36}
$$

An iteration procedure is required for every time step. At time step $j+1$, the column vector $\mathbf{V}_m^{j+1}$ is determined by solving the finite-difference equation

$$
\mathbf{A}(\mathbf{V}^j,\mathbf{V}_{m-1}^{j+1})\mathbf{V}_m^{j+1}=\mathbf{f}(\mathbf{V}^j,\mathbf{V}_{m-1}^{j+1}) \tag{37}
$$

where $\mathbf{V}_m^{j+1}$ is the column vector of unknowns at the mth iteration, and $\mathbf{V}_{m-1}^{j+1}$ is the output vector obtained in the previous iteration. For the first iteration, which determines $\mathbf{V}_1^{j+1}$ and $\mathbf{V}_1^{j+1}$, is used for $\mathbf{V}_0^{j+1}$. The numerical algorithm constructed is found to be stable, and the global error of the numerical scheme is $O[(\Delta t)^2,(\Delta r)^2]$ due to the application of Crank–Nicolson scheme. The iterations are repeated until the relative error at each node is less than 10^{-6} for the unknown h_i^{j+1}.

The numerical solution of Eq. (30) subject to boundary conditions (34) and initial condition are obtained using the above computational method for typical parameter values given by Re $=6.2$ and 0.0, $\epsilon=0.01$, $E=-0.0012,0,0.0012$, and $M=2,4,6$ [12,40,41].

Figures 2 and 3 show the results of the numerical computation for the film thickness at different Hartmann numbers for sinusoidal initial distribution (SU: $H_0=1+0.2\cos(\pi r)$) when Re$=6.2$, $E=0$, and $t=0.5$, 1, 5, 10, 20, 30, 40, and 50. It is observed that film thickness decreases with an increase in time for all values of Hartmann number M, and as the magnetic effect increases the rate of depletion decreases and hence more fluid is retained on the spinning disk. The successive surface contours flatten out smoothly, and the nonuniformity in the initial film thickness is effectively reduced as time increases. The smoothing effect produced by the flow is observed as time increases in the case of sinusoidal initial contour. Furthermore, as M increases, more time is required to flatten out smoothly. Also, the results reveal that with much increased spin-coating duration, the magnetic field significantly increases the film flatness in the long run. That is, for a fixed M, as t increases, the film thickness decreases and becomes more flat.

The relative volume Q/Q_0 of the film retained on the rotating substrate at time t, where Q_0 is the initial fluid volume on the disk for the initially sinusoidal film distribution is shown in Fig. 4 for different mass fluxes E and Hartmann numbers M. It is observed that as M increases, the rate of retention of the conducting fluid on the spinning disk increases for all values of E. For a particular Hartmann number M, absorption ($E<0$) retains more fluid on the disk as time increases as compared with evaporation ($E>0$).

Figure 5 presents the relative volume of the film retention for different values of M, when Re$=0$ and $E=0$. Although planarization and thinning of the film are affected by inertia, the changes are not very significant.

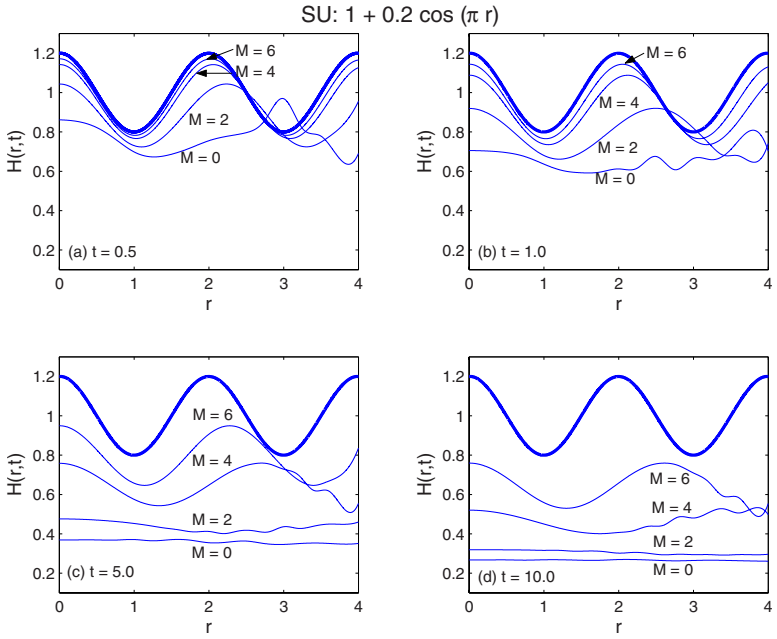


Fig. 2 Successive free surface contours for an initially sinusoidal distribution for Re=6.2, E=0, and ϵ=0.01

Different initial distributions of the height of the film [1] lead to different relative volume (Q/Q_0) of the conducting fluid retained on the spinning disk (Figs. 6 and 7) and this is illustrated by taking Gaussian plus (GP) uniform initial contour $(H_0=k(1+\exp(-r^2)))$ and slowly falling (SF) curve $(H_0=1/(1+r^2)^{1/4})$ as initial distributions of film thickness. The value of k is chosen such that the initial volumes Q_0 of liquid films on the substrate for both the initial distributions are the same. It is observed that slowly falling initial distribution retains more liquid on the spinning substrate than the Gaussian plus initial distribution, and this scenario is more pronounced in the initial times for all values of M considered. As the spinning continues, this trend is observed for larger values of M. The numerical study reveals that the magnetic field imparts a rigidity to the conducting fluid and hence might help in enhancing the rate of retention of the liquid film. In Sec. 5, time-dependent spatially uniform basic state is examined and its stability to long-wave disturbances is analyzed. Such an analysis of the linear stability of a nonstationary problem has been performed by several researchers including Joo et al. [55].

5 Description of Time-Dependent Basic State

The stability characteristics of the nonlinear equation (30) describing the shape of the film thickness as a function of space and time is examined using linear theory. As the film is draining due to centrifugation, the basic state is time dependent and it is assumed

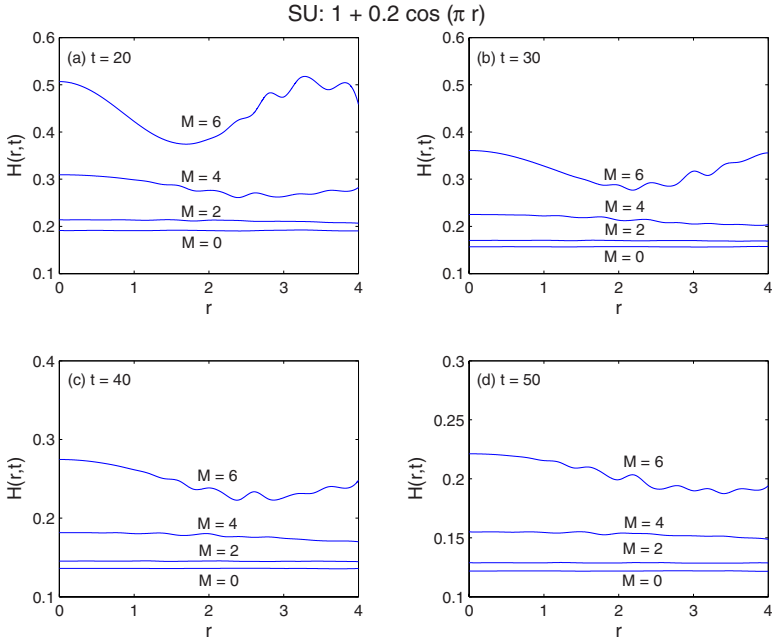


Fig. 3 Successive free surface contours for an initially sinusoidal distribution for Re=6.2, E=0, and ϵ=0.01

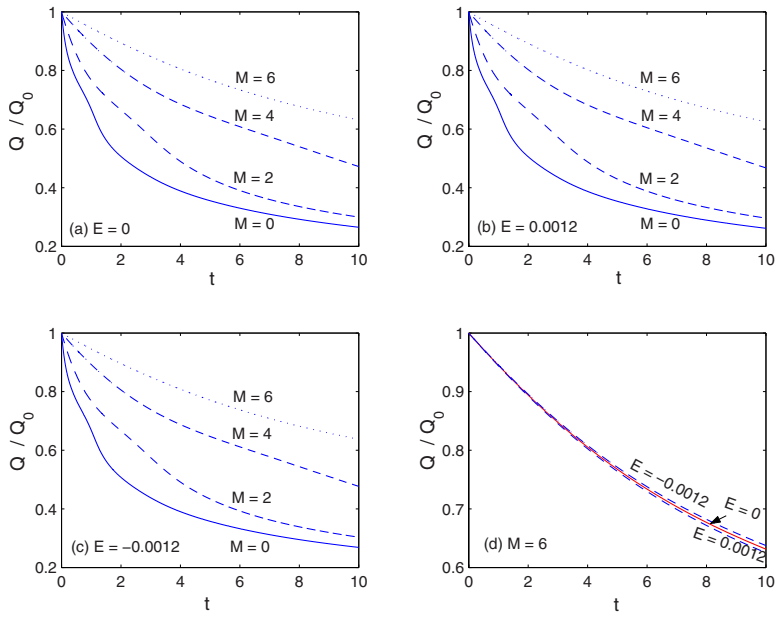


Fig. 4 Relative volume of the lubricant retention for an initially sinusoidal distribution for Re=6.2 and ϵ=0.01

to be flat. The film thickness is independent of the radial position and therefore the dependence on r is removed by using the following transformation [2]:

$$h = \bar{h}(t)$$

$$u = r\bar{u}(z,t)$$

$$v = r\bar{v}(z,t)$$

$$w = \bar{w}(z,t) \tag{38}$$

where quantities with an overbar denote the basic-state quantities. The resulting nonlinear ordinary differential equation governing the basic-state behavior is obtained as

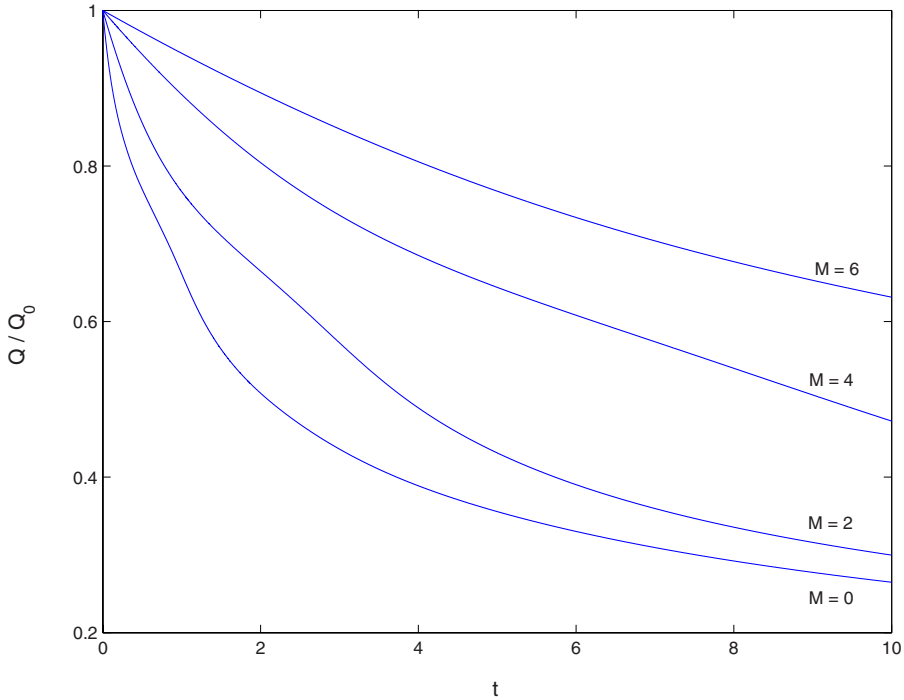


Fig. 5 Relative volume of the lubricant retention for an initially sinusoidal distribution for Re=0.0, *E*=0, and ϵ=0.01

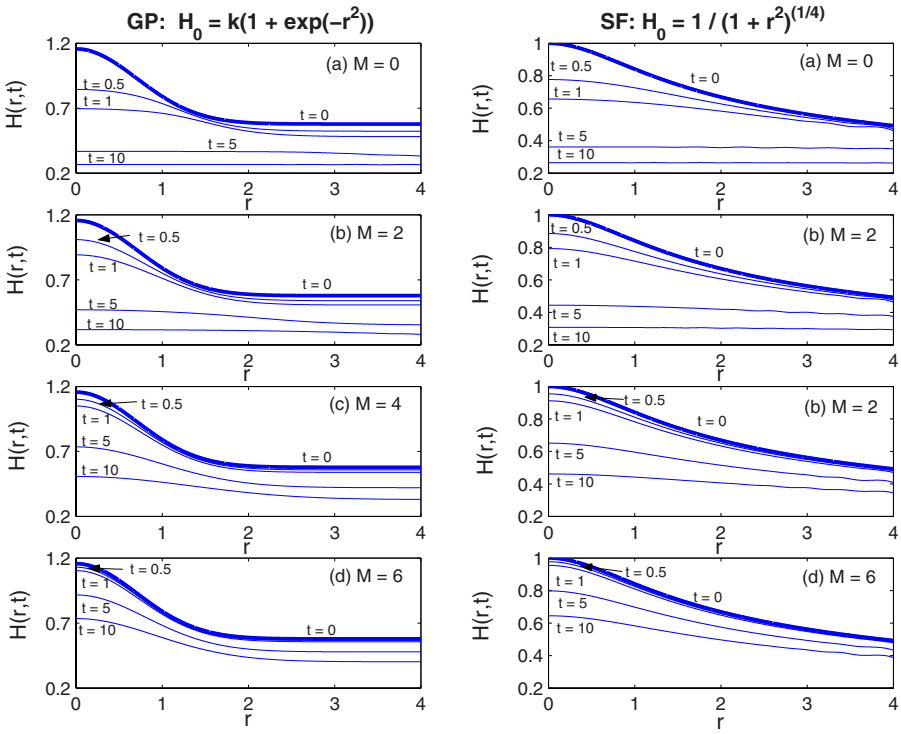


Fig. 6 Successive free surface contours for an initially Gaussian plus and slowly falling distribution for Re=6.2, $E=0$, and $\epsilon=0.01$

$$\bar{h}_t + \frac{2}{3}E + \frac{2\bar{h}}{M^2} - \frac{2\sinh M\bar{h}}{M^3\cosh M\bar{h}} + \epsilon\ \mathrm{Re}\left\{\frac{14\sinh M\bar{h}}{3M^7\cosh^3 M\bar{h}} + \frac{61\sinh 2M\bar{h}}{6M^7\cosh^2 M\bar{h}} - \frac{13\bar{h}}{3M^6\cosh^2 M\bar{h}} - \frac{2\bar{h}}{M^6\cosh^4 M\bar{h}} - \frac{10\bar{h}}{M^6} - \frac{4\bar{h}^2\sinh M\bar{h}}{M^5\cosh^3 M\bar{h}} + \frac{4E(\cosh M\bar{h}-1)}{3M^4\cosh^2 M\bar{h}} - \frac{2E\bar{h}\sinh M\bar{h}}{3M^3\cosh^3 M\bar{h}}\right\} = 0 \tag{39}$$

with

$$\bar{h}(0) = 1 \tag{40}$$

For fixed values of Hartmann number M and evaporation parameter E, Eqs. (39) and (40) are numerically integrated to give the time-dependent basic-state behavior. Figure 8 shows the basic-state film thickness for principal values of the evaporation parameter E and the Hartmann number M.

When the evaporation at the free surface is negligible ($E=0$), then the film thins due to drainage alone. In this case, film thickness decreases monotonically in time and it goes to zero as $t\to\infty$. It takes a longer time to thin initially when the magnetic field is present (Fig. 8(*b*)) and the film thickness tends to zero as $t\to\infty$.

When $E>0$, the fluid film thins both due to evaporation and drainage. In this case, the basic state is a film that thins monotonically to zero thickness in a finite time ($t_d=160(\sim 6.4\ s)$, $M=0$; Fig. 8(*a*)). The presence of the transverse magnetic field delays this and takes a longer time ($t_d=178(\sim 7.12\ s)$, $M=4$); Figs. 8(*b*) and 8(*c*)) for the film to reach zero thickness.

In the case of absorption or condensation to the film ($E<0$), thinning of the film takes place due to absorption, drainage, and inertia. Figure 8(*a*) shows that the basic state that reaches the steady state ($\bar{h}_s=0.106309$, $M=0$). The presence of the magnetic field increases the film thickness ($\bar{h}_s=0.10892$, $M=4$; Fig. 8(*b*)).

The basic state is governed by Eq. (39) with E replaced by $-|E|$. Once the steady state is reached, the position of the basic-state interface $\bar{h}_s$ is determined as the root of

$$\frac{2}{3}|E| - \frac{2\bar{h}_s}{M^2} + \frac{2\sinh M\bar{h}_s}{M^3\cosh M\bar{h}_s} - \epsilon\ \mathrm{Re}\left\{\frac{14\sinh M\bar{h}_s}{3M^7\cosh^3 M\bar{h}_s} + \frac{61\sinh 2M\bar{h}_s}{6M^7\cosh^2 M\bar{h}_s} - \frac{13\bar{h}_s}{3M^6\cosh^2 M\bar{h}_s} - \frac{2\bar{h}_s}{M^6\cosh^4 M\bar{h}_s} - \frac{10\bar{h}_s}{M^6} - \frac{4\bar{h}_s^2\sinh M\bar{h}_s}{M^5\cosh^3 M\bar{h}_s} - \frac{4|E|(\cosh M\bar{h}_s-1)}{3M^4\cosh^2 M\bar{h}_s} + \frac{2|E|\bar{h}_s\sinh M\bar{h}_s}{3M^3\cosh^3 M\bar{h}_s}\right\} = 0 \tag{41}$$

which for small values of ϵ has the approximate solution satisfying

$$\frac{2}{3}|E| - \frac{2\bar{h}_s}{M^2} + \frac{2\sinh M\bar{h}_s}{M^3\cosh M\bar{h}_s} = 0 \tag{42}$$

In the limit $M\to 0$, the solution of Eq. (42) yields

$$\bar{h}_s \sim |E|^{1/3} \tag{43}$$

which agrees with the results of Reisfeld et al. [12].

It is interesting to consider the special case when absorption exactly balances drainage and inertia. In this case, $\bar{h}_s=1$ for all time and this yields a critical value of $E=E_c$, obtained from Eq. (41) by substituting $\bar{h}_s=1$ in Eq. (41). The results show that $E_c=-0.955246605$ when $M=0$, $E_c=-0.3867993513$ when $M=2$, and $E_c=-0.14054192592$ when $M=4$.

The investigation of the stability of the film using linear theory is considered in Sec. 6.

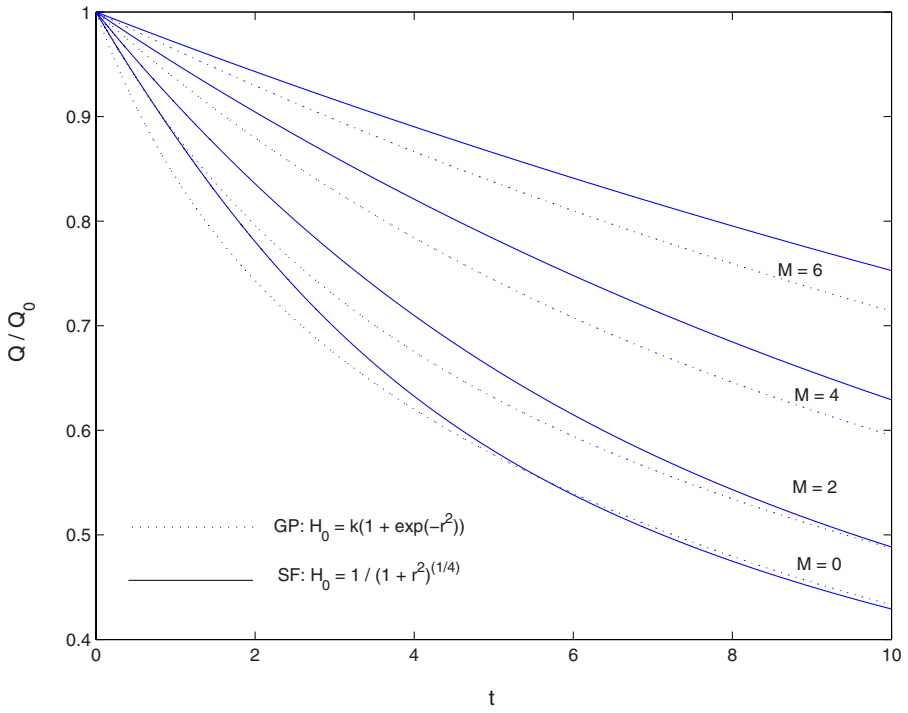


Fig. 7 Relative volume of the lubricant retention for an initially Gaussian plus and slowly falling distribution for Re=6.2, E=0, and ϵ=0.01

6 Linear Stability Analysis and Discussion

The linearized disturbance equation is obtained from Eq. (30) by perturbing the basic state by a small amount $\eta(h=\bar{h}(t)+\delta\eta(r,t))$ and substituting in Eq. (30) and linearizing in disturbance amplitude δ as

$$\eta_t + a_1(t) r^2 \eta_{rr} + a_2(t) r \eta_r + a_3(t)\eta = 0 \tag{44}$$

where

$$a_1(t) = \epsilon\ \mathrm{Re}\left\{ -\frac{3\sinh^2 M\bar{h}}{2M^6\cosh^4 M\bar{h}} + \frac{\bar{h}\sinh M\bar{h}}{M^5\cosh^3 M\bar{h}} + \frac{\bar{h}\sinh M\bar{h}}{2M^5\cosh^5 M\bar{h}} \right\} \tag{45}$$

$$a_2(t) = \frac{\sinh^2 M\bar{h}}{M^2\cosh^2 M\bar{h}} + \epsilon\ \mathrm{Re}\left\{ \frac{12\bar{h}\sinh M\bar{h}}{M^5\cosh^3 M\bar{h}} - \frac{5}{M^6} - \frac{11}{2M^6\cosh^2 M\bar{h}} + \frac{21}{2M^6\cosh^4 M\bar{h}} + \frac{4\bar{h}^2}{M^4\cosh^2 M\bar{h}} - \frac{6\bar{h}^2}{M^4\cosh^4 M\bar{h}} + \frac{11\bar{h}\sinh M\bar{h}}{2M^5\cosh^5 M\bar{h}} - \frac{2E\sinh M\bar{h}}{3M^3\cosh^2 M\bar{h}} + \frac{E\sinh M\bar{h}}{M^3\cosh^3 M\bar{h}} + \frac{2E\bar{h}}{3M^2\cosh^2 M\bar{h}} - \frac{E\bar{h}}{M^2\cosh^4 M\bar{h}} \right\} \tag{46}$$

$$a_3(t) = \frac{2}{M^2} - \frac{2\sinh M\bar{h}}{M^3\cosh M\bar{h}}(\coth M\bar{h} - \tanh M\bar{h}) + \epsilon\ \mathrm{Re}\left\{ \frac{14\sinh M\bar{h}}{3M^7\cosh^3 M\bar{h}}(\coth M\bar{h} - 3\tanh M\bar{h}) - \frac{13}{3M^6\cosh^2 M\bar{h}}(1 - 2\bar{h}\tanh M\bar{h}) + \frac{61\sinh 2M\bar{h}}{6M^7\cosh^2 M\bar{h}}(\coth 2M\bar{h} - 2\tanh M\bar{h}) - \frac{2}{M^6\cosh^4 M\bar{h}}(1 - 4\bar{h}\tanh M\bar{h}) - \frac{10}{M^6} - \frac{4\sinh M\bar{h}}{M^5\cosh^3 M\bar{h}}(2\bar{h} + \bar{h}^2\coth M\bar{h} - 3\bar{h}^2\tanh M\bar{h}) - \frac{4E\tanh M\bar{h}}{3M^4\cosh M\bar{h}} + \frac{8E\tanh M\bar{h}}{3M^4\cosh^2 M\bar{h}} - \frac{2E\sinh M\bar{h}}{3M^3\cosh^3 M\bar{h}}(1 + \bar{h}\coth M\bar{h} - 3\bar{h}\tanh M\bar{h}) \right\} \tag{47}$$

Transforming Eq. (44) into an equation with constant coefficients by the transformation $r=e^{\xi}$, $\xi=\ln r$, gives

$$\eta_t + a_1(t)\eta_{\xi\xi} + [a_2(t) - a_1(t)]\eta_\xi + a_3(t)\eta = 0 \tag{48}$$

Using normal mode analysis and assuming the disturbance quantity η in the form

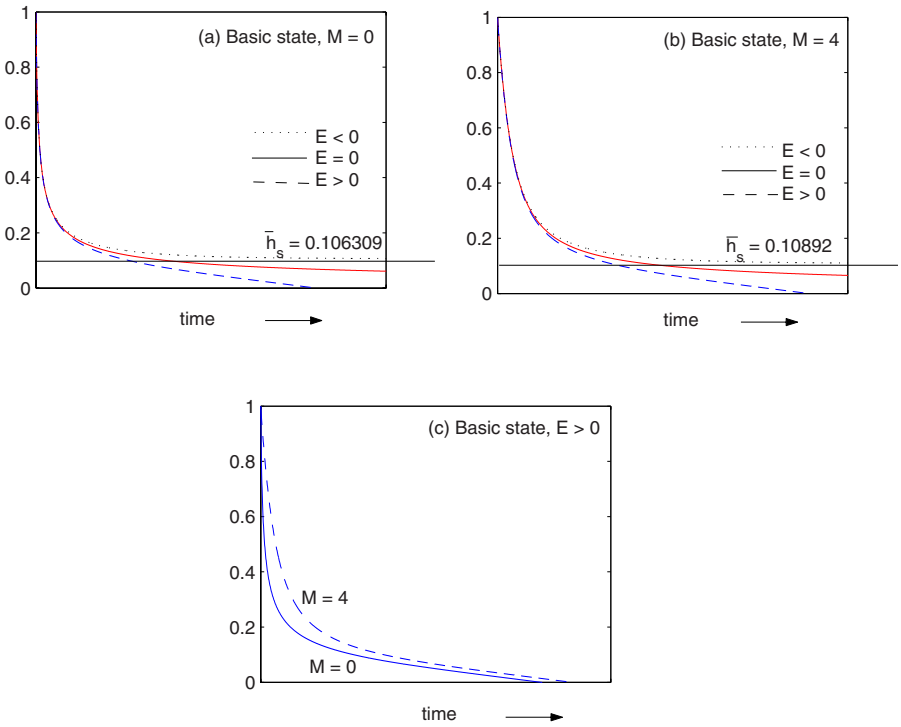


Fig. 8 Basic-state film thickness for principal values of the evaporation parameter and different values of Hartmann number

$$\eta(\xi,t) = H(t)e^{ik_0\xi} \tag{49}$$

where $i=\sqrt{-1}$, the equation for normal mode amplitude $H(t)$ is obtained as

$$\frac{\dot{H}}{H} = a_1(t)\frac{k^2}{\epsilon} - [a_2(t) - a_1(t)]\frac{ik}{\sqrt{\epsilon}} - a_3(t) \tag{50}$$

where k is defined as $k=\sqrt{\epsilon}k_0$. The equation can be rewritten as

$$\frac{d\ln H}{d\bar{h}} = \frac{1}{\bar{h}_t}\left\{a_1(t)\frac{k^2}{\epsilon} - [a_2(t) - a_1(t)]\frac{ik}{\sqrt{\epsilon}} - a_3(t)\right\} \tag{51}$$

In what follows, attention is focused on the region away from the origin where inertial effects are important and where the equations are valid. Equation (51) is then numerically integrated to yield normal mode amplitude by substituting Eq. (39), which governs the behavior of the thinning basic-state film for different values of evaporation parameter E and Hartmann number M.

Equation (50) indicates that $|H|$ will have an extremum point in time and this gives the cutoff wave number k_c. The cutoff wave number k_c increases with an increase in Hartmann number M. The linear stability theory results reduce to those of Reisfeld et al. [12] when Hartmann number $M=0$.

The case when there is no evaporation is considered first. Equation (39) with $E=0$ governs the behavior of the thinning basic-state film in the presence of a transverse magnetic field. The solution $\bar{h}$ of Eq. (39) is substituted into the normal mode equation (50), and Eq. (50) is integrated to yield the normal mode amplitude $H(t)$. Figures 9(*a*) and 9(*b*) present the magnitude of the normal mode amplitude $|H/H_0|$ of the disturbance when Re=6.2, ϵ=0.01, and $M=0,2$, where H_0 is the value of H at $t=0$. It is observed that the amplitude of the disturbance decays to zero for $k^2<k_c^2$ (Fig. 9(*a*)) and stable for lower wave numbers. It is to be noted that $k_c=1.5554$ when $M=0$ and $k_c=10.8184$ for $M=2$. For sufficiently large wave numbers ($k^2>k_c^2$), the disturbance amplitude increases in magnitude to a finite amplitude, but it ultimately decays to zero (i.e., the disturbance is transiently stable [56], Fig. 9(*b*)). Furthermore, it is to be remarked that $|H/H_0|$ grows temporarily in the interval $\bar{h}\in[0,1]$ provided that $k^2>k_c^2$ and decays for $k^2<k_c^2$. Furthermore, the results show that the effect of the magnetic field is to enhance the stability of a thin nonuniform conducting film flow on a spinning disk. Although the amplitude of the disturbance is larger for $M=2$ than $M=0$ when $k^2<k_c^2$, the disturbance decays to zero earlier when $M=2$. When $k^2>k_c^2$, it takes a longer time to attain a maximum finite amplitude when the magnetic field is present. However, the disturbance decays to zero earlier when the magnetic field is present than when it is absent (Fig. 9(*b*)).

When the evaporation from the film surface is present, the basic state of the film is governed by Eq. (39) with $E>0$. A dynamic behavior of the perturbed state is obtained by integrating Eq. (50). In this case, $|H/H_0|$ exhibits both a region in which the disturbance is stable and a region in which it is transiently stable (Figs. 10(*a*) and 10(*b*)). The results show that the disturbance amplitude $|H/H_0|$ grows initially and reaches its maximum amplitude at about $t=2(\sim 0.1$ s) when $M=0$ and at about $t=11(\sim 0.44$ s) when $M=2$ and then decays monotonically. As k_c^2 increases with M, the range of wave numbers where the disturbance remains stable increases with increase with M.

In the case of absorption ($E<0$), the normal mode equation far from steady state position is given by Eq. (50) and (51) with E replaced by $-|E|$. Once the steady state is reached, then the differential equation governing the normal mode amplitude is given by Eq. (50) with $\bar{h}$ replaced by $\bar{h}_s$, which is a solution of Eq. (41). The results show that when the basic state is steady, the solution for normal mode amplitude increases exponentially in time (Fig. 11(*b*)) and the film flow is unstable for $k^2>k_c^2$. If the cutoff wave number is denoted by $k_{c_1}^2$ in this case and the cutoff wave number is denoted by $k_{c_2}^2$ for the case when $E=E_c(\bar{h}_s=1)$, then the results show that for the case of absorption to the film, the disturbance shows three distinct regions of stability, namely, the disturbance is stable for $k^2<k_{c_2}^2$ (Fig. 11(*a*)), is transiently stable for wave num-

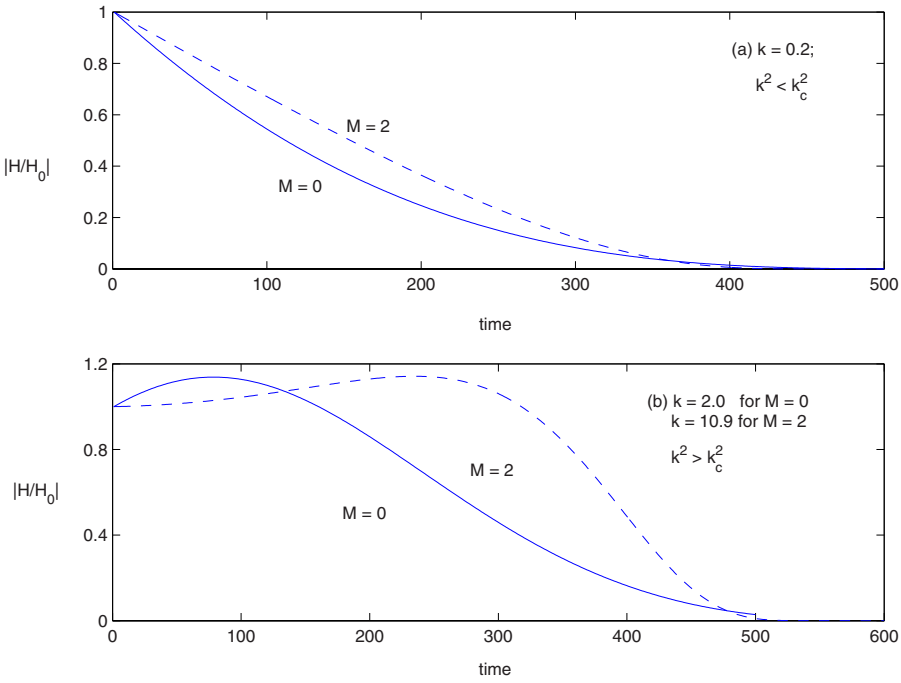


Fig. 9 Normal mode amplitude for a disturbance when Re=6.2, ϵ=0.01, and E=0: (*a*) stable and (*b*) transiently stable

bers in the range $k_{c_2}^2 < k^2 < k_{c_1}^2$ (Fig. 11(*c*)), and is unstable for $k^2 > k_{c_1}^2$ (Fig. 11(*b*)). The disturbance amplitude is larger in the presence of a magnetic field and, since k_c^2 increases with M, the role of the magnetic field is to increase the region of stability. The domains of instability for a rotating liquid film in the presence of transverse magnetic field have been summarized in Fig. 12. It is observed that in the limit $M \to 0$, $k_{c_1}^2$ reduces to $15/(\mathrm{Re}|E|^{4/3})$ and $k_{c_2}^2$ reduces to 15/Re, which agree with the values predicted by Reisfeld et al. [12].

It is interesting to point out here that the following rescaling reduces the number of nondimensional parameters by 1. Equation (41) (when multiplied through by M^3) can be written as

$$\frac{2}{3}|E^*| - 2\bar{h}_s^* + 2\tanh\bar{h}_s^* - \frac{\mathrm{Re}^*}{3}\{14\tanh\bar{h}_s^*\,\mathrm{sech}^2\,\bar{h}_s^* + 61\tanh\bar{h}_s^* - 13\bar{h}_s^*\,\mathrm{sech}^2\,\bar{h}_s^* - 6\bar{h}_s^*\,\mathrm{sech}^4\,\bar{h}_s^* - 30\bar{h}_s^* - 12\bar{h}_s^{*2}\tanh\bar{h}_s^*\,\mathrm{sech}^2\,\bar{h}_s^* - 4|E^*|(\cosh\bar{h}_s^* - 1)\mathrm{sech}^2\,\bar{h}_s^*$$

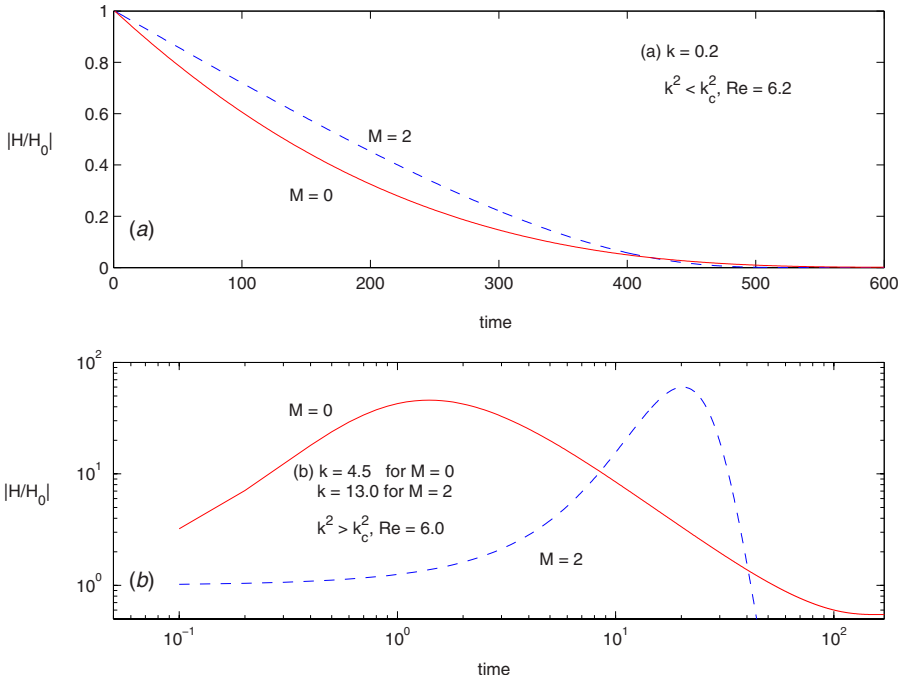


Fig. 10 Normal mode amplitude for a disturbance when E=0.0012 and ϵ=0.01: (*a*) stable and (*b*) transiently stable

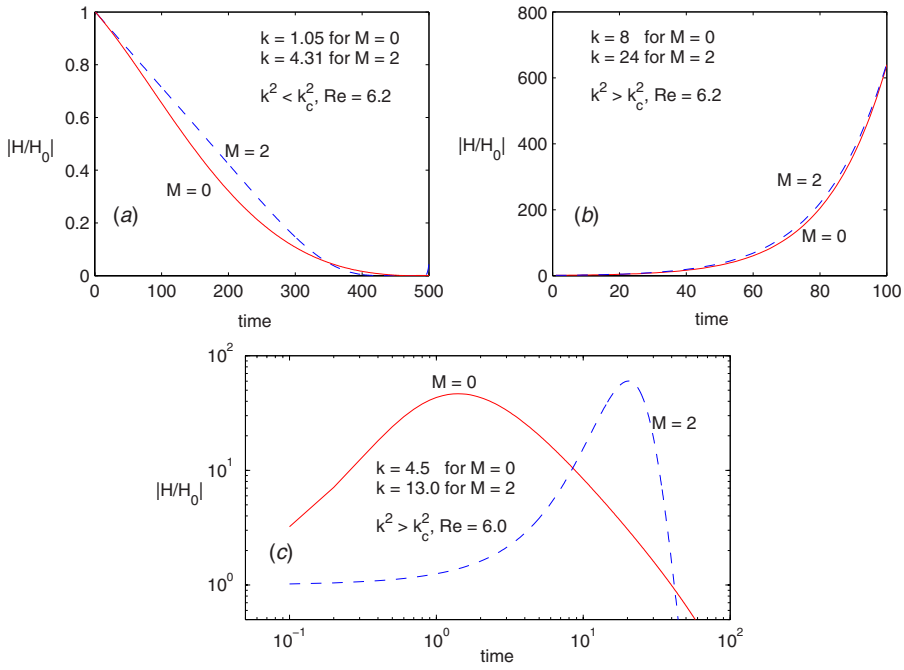


Fig. 11 Normal mode amplitude for a disturbance when $E=-0.0012$ and $\epsilon=0.01$: (*a*) stable, (*b*) unstable, and (*c*) transiently stable

$$+ 2|E^*|\tanh \bar{h}_s^* \operatorname{sech}^2 \bar{h}_s^*\} = 0 \tag{52}$$

or more simply as

$$|E^*| - 3\bar{h}_s^* + 3 \tanh \bar{h}_s^* - \frac{\mathrm{Re}^*}{2}\{2(7 + |E^*| - 6\bar{h}_s^{*2})\tanh \bar{h}_s^* \operatorname{sech}^2 \bar{h}_s^*$$

$$+ 61 \tanh \bar{h}_s^* - 4|E^*|(\cosh \bar{h}_s^* - 1)\operatorname{sech}^2 \bar{h}_s^* - \bar{h}_s^*(6 \operatorname{sech}^4 \bar{h}_s^*$$

$$+ 13 \operatorname{sech}^2 \bar{h}_s^* + 30)\} = 0 \tag{53}$$

involving just the three parameters E^*, $\bar{h}_s^*$, and Re^* defined by

$$E^* = M^3 E, \quad \bar{h}_s^* = M\bar{h}_s, \quad \mathrm{Re}^* = \frac{\epsilon\ \mathrm{Re}}{M^4} \tag{54}$$

rather than original four parameters E, $\bar{h}_s$, Re, and M. In addition, Eq. (53) shows that the solution for $\bar{h}_s^*$ is of the form

$$\bar{h}_s^* = f(|E^*|, \mathrm{Re}^*) \tag{55}$$

(for some function f), so that the solution for $\bar{h}_s$ is of the form

$$\bar{h}_s = f\left(M^3|E|, \frac{\epsilon\ \mathrm{Re}}{M^4}\right) \tag{56}$$

In fact, Eq. (42) and all the other equations can be reduced in a similar sort of way; for example, with the additional rescaling, $h^* = Mh$ and $t^* = M^{-2}t$ even the full evolution equations (29) and (30) can be reduced. It is worth mentioning here that this rescaling clearly shows the dependence on M and reveals interesting information about the structure of the solution.

7 Conclusion

The effects of a magnetic field on the planarization and thinning process of a viscous conducting fluid on a rotating disk and the linear stability of the time-dependent spatially varying basic state of the film flow system have been examined. The evolution equation describing the transient film thickness has been derived on the basis of lubrication theory and small aspect ratio for the film. The effects of magnetic field on the temporal evolution of the free surface of the conducting fluid and the relative volume of the conducting fluid retained on the disk have been studied for different initial distributions of the height of the fluid film. The numerical solutions reveal that the magnetic field has a tendency to impart rigidity to the conducting fluid and hence might help in enhancing the rate of retention of the liquid lubricant.

The linear stability analysis is performed in the region lying away from the origin where inertial effects are important and where the equations are valid. The results indicate that the infinitesimal disturbances decay for small wave numbers and are transiently stable for large wave numbers, for both zero and nonzero values of the Hartmann number, when either there is no mass

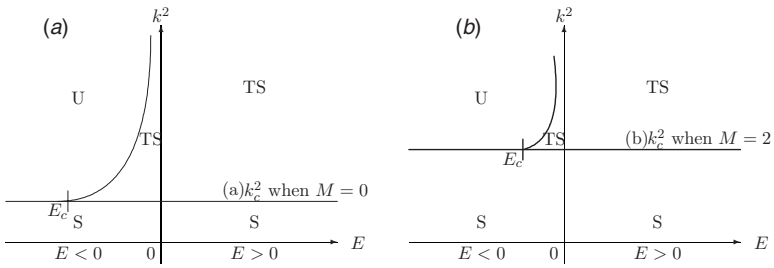


Fig. 12 Domains of instability for (a) $M=0$ and (b) $M=2$. S: stable; TS: transiently stable; U: unstable

transfer or there is evaporation from the film. When there is absorption through the film then the flow system exhibits three different domains of stability. It is observed that in the absence of magnetic field, the localized solutions of the eigenvalue problem and those of the linearized Navier–Stokes equations and boundary conditions describing the axisymmetric wave regimes in a viscous liquid film flow over a spinning disk [10] adequately capture the properties of the full problem. The present study shows that when $M=0$, the solutions of the eigenvalue problem behave analogously and that the disturbance grows to a maximum, then decays, and exhibits a transiently stable state. Furthermore, it is observed that the solution of the eigenvalue problem in the presence of a magnetic field has similar behavior to that described above. This suggests that the corresponding linearized solutions of the Navier–Stokes equations incorporating the influence of a magnetic field would predict the same behavior as the localized solution for the corresponding values of the Hartmann number and this study forms a part of our future work.

Acknowledgment

The authors sincerely thank the referees for their very useful suggestions and comments. Their special thanks to one of the referees for pointing out the dependence of the solution on M and revealing the structure of the solution. The referees' comments have improved the quality and style of the paper. They also thank the editor for his useful suggestions and encouraging remarks.

Appendix

The radial and the axial velocities are obtained by solving Eqs. (21)–(28) and are given by

$$
\begin{aligned}
u = {} & \frac{r}{M^2}\left[1-\frac{\cosh M(z-h)}{\cosh Mh}\right] \\
& + \frac{\epsilon\ \mathrm{Re}\ r}{M^4\cosh^2 Mh}\left\{\frac{\sinh Mh}{M^2\cosh Mh}\sinh Mz + \left(\frac{3}{2M^2}\right.\right. \\
& \left.\left. + \frac{5\cosh^2 Mh}{M^2} + \frac{\cosh 2Mh}{6M^2}\right)\frac{\cosh M(z-h)}{\cosh Mh} + \frac{3h}{M}\sinh Mz \right. \\
& - \frac{3}{2M^2} - \frac{5\cosh^2 Mh}{M^2} - \frac{\cosh 2M(z-h)}{6M^2} - \frac{z\sinh Mz}{M} \\
& \left. - \frac{2}{M}z\sinh M(z-h)\cosh Mh\right\} + \frac{\epsilon\ \mathrm{Re}\ r^2h_r}{M^3}\frac{\sinh Mh}{\cosh^3 Mh}\left\{\right. \\
& - \frac{\cosh M(z-h)}{M^2\cosh Mh} - \frac{\sinh Mh}{2M^2\cosh Mh}\sinh Mz - \frac{h}{2M}\sinh Mz \\
& \left. + \frac{z\sinh Mz}{2M} + \frac{1}{M^2}\right\} + \frac{\epsilon\ \mathrm{Re}\ rh_t}{M\cosh^2 Mh}\left\{-\frac{\sinh Mz}{2M^2}\right. \\
& \left. - \frac{h\sinh Mh}{2M\cosh Mh}\sinh Mz + \frac{z\cosh Mz}{2M}\right\} \\
& + \frac{\epsilon\ \mathrm{Re}\ r}{M^3\cosh Mh}\left\{\frac{z^2}{M}\cosh M(z-h) - \frac{z}{M^2}\sinh M(z-h)\right. \\
& \left. - \frac{hz}{M}\frac{\cosh Mz}{\cosh Mh} + \frac{h^2}{M}\frac{\sinh Mh}{\cosh^2 Mh}\sinh Mz\right\} + \left\{\frac{\epsilon\ \mathrm{Re}\ h_r}{F^2M^2}\right. \\
& \left. - \frac{\epsilon^3\mathrm{We}}{M^2r^2}(r^2h_{rrr} + rh_{rr} - h_r)\right\}\frac{\cosh M(z-h)}{\cosh Mh} \qquad \text{(A1)}
\end{aligned}
$$

$$
\begin{aligned}
w = {} & -\frac{2z}{M^2} + \frac{2}{M^3}\left[\frac{\sinh M(z-h)}{\cosh Mh} + \frac{\sinh Mh}{\cosh Mh}\right] \\
& + \frac{rh_r}{M^2}\frac{[1-\cosh Mz]}{\cosh^2 Mh} - \left[\frac{2\epsilon\ \mathrm{Re}}{M^4\cosh^2 Mh}\right. \\
& \left. - \frac{2\epsilon\ \mathrm{Re}\ rh_r}{M^3}\frac{\sinh Mh}{\cosh^3 Mh}\right]\left\{\left(\frac{3}{2M^2} + \frac{5\cosh^2 Mh}{M^2}\right.\right. \\
& \left.\left. + \frac{\cosh 2Mh}{6M^2}\right)\times\left(\frac{\sinh M(z-h)}{M\cosh Mh} + \frac{\sinh Mh}{M\cosh Mh}\right)\right. \\
& + \frac{\sinh Mh}{M^2\cosh Mh}\left(\frac{\cosh Mz}{M} - \frac{1}{M}\right) + \frac{3h\cosh Mz}{M^2} - \frac{3h}{M^2} \\
& - \frac{3z}{2M^2} - \frac{5z\cosh^2 Mh}{M^2} - \frac{\sinh 2M(z-h)}{12M^3} - \frac{\sinh 2Mh}{12M^3} \\
& - \frac{z\cosh Mz}{M^2} + \frac{\sinh Mz}{M^3} - \frac{2z}{M^2}\cosh M(z-h)\cosh Mh \\
& \left. + \frac{2}{M^3}\sinh M(z-h)\cosh Mh + \frac{2}{M^3}\sinh Mh\cosh Mh\right\} \\
& - \frac{\epsilon\ \mathrm{Re}\ rh_r}{M^3\cosh^2 Mh}\left\{\frac{16}{3M^3}\frac{\sinh 2Mh}{\cosh Mh}(\sinh M(z-h) + \sinh Mh)\right. \\
& - \left(\frac{1}{2M^2} + \frac{5\cosh^2 Mh}{M^2} + \frac{\cosh 2Mh}{6M^2}\right)\left(\frac{\cosh Mz}{M\cosh^2 Mh}\right. \\
& \left.\left. - \frac{1}{M\cosh^2 Mh}\right) + \frac{3}{M^3}(\cosh Mz - 1) - \frac{5z\sinh 2Mh}{M^2}\right. \\
& + \frac{\cosh 2M(z-h)}{6M^3} + \frac{11\cosh 2Mh}{6M^3} + \frac{2z\sinh M(z-2h)}{M^2} \\
& \left. - \frac{2}{M^3}\cosh M(z-2h)\right\} - \left[\frac{3\epsilon\ \mathrm{Re}\ rh_r}{M^3}\frac{\sinh Mh}{\cosh^3 Mh}\right. \\
& \left. + \frac{\epsilon\ \mathrm{Re}\ r^2h_{rr}}{M^3}\frac{\sinh Mh}{\cosh^3 Mh} + \frac{\epsilon\ \mathrm{Re}\ r^2h_r^2}{M^2}\left(\frac{1-2\sinh^2 Mh}{\cosh^4 Mh}\right)\right] \\
& \times\left\{-\frac{\sinh M(z-h)}{M^3\cosh Mh} - \frac{\sinh Mh}{2M^3\cosh Mh}\right. \\
& - \frac{\sinh Mh}{2M^3\cosh Mh}\cosh Mz - \frac{h\cosh Mz}{2M^2} + \frac{h}{2M^2} + \frac{z\cosh Mz}{2M^2} \\
& \left. - \frac{\sinh Mz}{2M^3} + \frac{z}{M^2}\right\} - \frac{\epsilon\ \mathrm{Re}\ r^2h_r^2}{M^2}\frac{\sinh Mh}{\cosh^3 Mh}\left\{\frac{(\cosh Mz - 1)}{2M^3\cosh^2 Mh}\right. \\
& \left. - \frac{(\cosh Mz - 1)}{2M^3}\right\} - \left[\frac{2\epsilon\ \mathrm{Re}\ h_t}{M\cosh^2 Mh} + \frac{\epsilon\ \mathrm{Re}\ rh_{rt}}{M\cosh^2 Mh}\right. \\
& \left. - 2\epsilon\ \mathrm{Re}\ rh_rh_t\frac{\sinh Mh}{\cosh^3 Mh}\right]\times\left\{-\frac{(\cosh Mz - 1)}{M^3}\right. \\
& \left. - \frac{h\sinh Mh}{2M^2\cosh Mh}(\cosh Mz - 1) + \frac{z\sinh Mz}{2M^2}\right\} - \frac{\epsilon\ \mathrm{Re}\ rh_rh_t}{\cosh^2 Mh} \\
& \left\{-\frac{\sinh Mh}{2M^3\cosh Mh}(\cosh Mz - 1) - \frac{h(\cosh Mz - 1)}{2M^2\cosh^2 Mh}\right\} \\
& - \left[\frac{2\epsilon\ \mathrm{Re}}{M^3\cosh Mh} - \frac{\epsilon\ \mathrm{Re}\ rh_r}{M^2}\frac{\sinh Mh}{\cosh^2 Mh}\right]\left\{\frac{z^2}{M^2}\sinh M(z-h)\right. \\
& - \frac{3z}{M^3}\cosh M(z-h) + \frac{3}{M^4}(\sinh M(z-h) + \sinh Mh) \\
& - \frac{hz\sinh Mz}{M^2\cosh Mh} + \frac{h(\cosh Mz - 1)}{M^3\cosh Mh} + \frac{h^2\sinh Mh}{M^2\cosh^2 Mh}(\cosh Mz \\
& \left. - 1)\right\} - \frac{\epsilon\ \mathrm{Re}\ rh_r}{M^2\cosh Mh}\left\{-\frac{z^2}{M^2}\cosh M(z-h) + \frac{3z}{M^3}\sinh M(z\right. \\
& \left. - h) - \frac{3}{M^4}[\cosh M(z-h) - \cosh Mh] - \frac{z\sinh Mz}{M^3\cosh Mh}\right.
\end{aligned}
$$

$$+\frac{(\cosh Mz-1)}{M^4\cosh Mh}+\frac{h\sinh Mh}{M^2\cosh^2 Mh}z\sinh Mz$$
$$+\frac{h\sinh Mh}{M^3\cosh^2 Mh}(\cosh Mz-1)+\frac{h^2(1-\sinh^2 Mh)}{M^2\cosh^3 Mh}(\cosh Mz-1)\Bigg\}+\Bigg\{\frac{\epsilon\ \mathrm{Re}\ h_r}{F^2M^2}-\frac{\epsilon^3\mathrm{We}}{M^2r^2}(r^2h_{rrr}+rh_{rr}-h_r)\Bigg\}\frac{h_r\ (\cosh Mz-1))}{\cosh^2 Mh}-\Bigg\{\frac{\epsilon\ \mathrm{Re}}{F^2M^2}\left(h_{rr}+\frac{h_r}{r}\right)-\frac{\epsilon^3\mathrm{We}}{M^2r^3}(r^3h_{rrrr}+2r^2h_{rrr}-rh_{rr}+h_r)\Bigg\}\times\left[\frac{(\sinh M(z-h)+\sinh Mh)}{M\cosh Mh}-z\right] \tag{A2}$$

References

[1] Emslie, A. G., Bonner, F. D., and Peck, L. G., 1958, "Flow of a Viscous Liquid on a Rotating Disk," J. Appl. Phys., **29**, pp. 858–862.
[2] Higgins, B. G., 1986, "Film Flow on a Rotating Disk," Phys. Fluids, **29**, pp. 3522–3529.
[3] Rehg, T., and Higgins, B. G., 1988, "The Effect of Inertia and Interfacial Shear on Film Flow Over a Rotating Disc," Phys. Fluids, **31**, pp. 1360–1371.
[4] Wang, C. Y., Watson, L. T., and Alexander, K. A., 1991, "Spinning of a Liquid Film From an Accelerating Disk," IMA J. Appl. Math., **46**, pp. 201–210.
[5] Stillwagon, L. E., and Larson, R. G., 1990, "Levelling of Thin Film Over Uneven Substrates During Spin Coating," Phys. Fluids A, **2**, pp. 1937–1944.
[6] Kitamura, A., 2000, "Asymptotic Solution for Film Flow on a Rotating Disk," Phys. Fluids, **12**, pp. 2141–2144.
[7] Moriarty, J. A., Schwartz, L. W., and Tuck, E. O., 1991, "Unsteady Spreading of Thin Liquid Films With Small Surface Tension," Phys. Fluids A, **3**, pp. 733–742.
[8] McKinley, I. S., Wilson, S. K., and Duffy, B. R., 1999, "Spin Coating and Air-Jet Blowing of Thin Viscous Drops," Phys. Fluids, **11**, pp. 30–47.
[9] Melo, F., Joanny, J. F., and Fauve, S., 1989, "Fingering Instability of Spinning Drops," Phys. Rev. Lett., **63**, pp. 1958–1961.
[10] Sisoev, G. M., Matar, O. K., and Lawrence, C. J., 2003, "Axisymmetric Wave Regimes in Viscous Liquid Film Flow Over a Spinning Disk," J. Fluid Mech., **495**, pp. 385–411.
[11] Wu, L., 2005, "Surface Wave Propagation of Thin Liquid Films on a Rotating and Non-Rotating Disk," Phys. Rev. E, **72**, pp. 016313.
[12] Reisfeld, B., Bankoff, S. G., and Davis, S. H., 1991, "The Dynamics and Stability of Thin Liquid Films During Spin Coating: I. Films With Constant Rates of Evaporation or Absorption," J. Appl. Phys., **70**, pp. 5258–5266.
[13] Espig, H., and Hoyle, R., 1965, "Waves in a Thin Liquid Layer on a Rotating Disk," J. Fluid Mech., **22**, pp. 671–677.
[14] Charwat, A. F., Kelly, R. E., and Gazley, C., 1972, "The Flow and Stability of Thin Liquid Films on a Rotating Disk," J. Fluid Mech., **53**, pp. 227–255.
[15] Matsumoto, S., Saito, K., and Takashima, Y., 1974, "The Thickness of a Viscous Liquid Film on a Rotating Disk," J. Chem. Eng. Jpn., **6**, pp. 503–506.
[16] Miyasaka, Y., 1974, "On the Flow of a Viscous Free Boundary Jet on a Rotating Disk, II. Comparison of Experimental Results With Calculated Values by Means of Film Thickness," Bull. JSME, **17**, pp. 1469–1475.
[17] Butuzov, A. I., and Puhovoi, I. I., 1976, "On Regimes of Liquid Film Flows Over a Rotating Surface," J. Eng. Phys., **31**, pp. 217–224.
[18] Rifert, V. G., Barabash, P. A., and Muzhilko, A. A., 1982, "Stochastic Analysis of Wave Surface Structure of Liquid Film Flowing Under Centrifugal Forces," Izv. Vyssh. Uchebn. Zaved., Energetika, **8**, pp. 62–66.
[19] Thomas, S., Faghri, A., and Hankey, W., 1991, "Experimental Analysis and Flow Visualization of a Thin Liquid Film on a Stationary and Rotating Disk," ASME Trans. J. Fluids Eng., **113**, pp. 73–80.
[20] Woods, W. P., 1995, "The Hydrodynamics of Thin Liquid Films Flowing Over a Rotating Disc," Ph.D. thesis, University of Newcastle upon Tyne, UK.
[21] Leneweit, G., Roesner, K. G., and Koehler, R., 1999, "Surface Instabilities of Thin Liquid Film Flow on a Rotating Disk," Exp. Fluids, **26**, pp. 75–85.
[22] Shkadov, V. Ya., 1973, "Some Methods and Problems of the Theory of Hydrodynamic Stability," *Scientific Proceedings 25*, Institute of Mechanics of Lomonosov, Moscow State University, Moscow.
[23] Rauscher, J. W., Kelly, R. E., and Cole, J. D., 1973, "An Asymptotic Solution for the Laminar Flow of Thin Film on a Rotating Disk," ASME J. Appl. Mech., **40**, pp. 43–47.
[24] Dorfman, L. A., 1967, "Flow and Heat Transfer in a Viscous Liquid Layer on a Spinning Disc," J. Eng. Phys., **12**, pp. 309–316.
[25] Miyasaka, Y., 1974, "On the Flow of a Viscous Free Boundary Jet on a Rotating Disk, I. Theoretical Analysis," Bull. JSME, **17**, pp. 1461–1468.
[26] Sisoev, G. M., Taldrik, A. F., and Shkadov, V. Ya., 1986, "Flow of a Viscous Liquid Film on the Surface of a Rotating Disc," J. Eng. Phys., **51**, pp. 1171–1174.
[27] Sisoev, G. M., and Shkadov, V. Ya., 1990, "Helical Waves in a Liquid Film on a Rotating Disc," J. Eng. Phys., **58**, pp. 573–577.
[28] Needham, D. J., and Merkin, J. H., 1987, "The Development of Nonlinear Waves on the Surface of a Horizontally Rotating Thin Liquid Film," J. Fluid Mech., **184**, pp. 357–379.
[29] Dandapat, B. S., and Ray, P. C., 1990, "Film Cooling on a Rotating Disk," Int. J. Non-Linear Mech., **25**, pp. 569–582.
[30] Usha, R., and Ravindran, R., 2004, "Analysis of Cooling of a Conducting Fluid Film of Non-Uniform Thickness on a Rotating Disk," Int. J. Non-Linear Mech., **39**, pp. 153–164.
[31] Rehg, T. J., 1992, "Spin Coating of Monodisperse Colloidal Suspensions: Evidence of Evaporative Convection," Ph.D. thesis, University of California, Davis.
[32] Usha, R., Ravindran, R., and Uma, B., 2005, "Dynamics and Stability of a Thin Liquid Film on a Heated Rotating Disk—Film With Variable Viscosity," Phys. Fluids, **17**, p. 102103.
[33] Wu, L., 2006, "Spin Coating of Thin Liquid Films on an Axisymmetrically Heated Disk," Phys. Fluids, **18**, pp. 063602.
[34] Ddandapat, B. S., and Ray, P. C., 1993, "Flow of a Thin Liquid Film Over a Cold/Hot Rotating Disk," Int. J. Non-Linear Mech., **28**, pp. 489–501.
[35] Strong, L., and Middleman, S., 1989, "Lubricant Retention on a Spinning Disk," AIChE J., **35**, pp. 1753–1756.
[36] Ray, P. C., and Dandapat, B. S., 1994, "Flow of Thin Liquid Film on a Rotating Disc in the Presence of a Transverse Magnetic Field," Q. J. Mech. Appl. Math., **47**, pp. 297–304.
[37] Dandapat, B. S., and Ray, P. C., 1998, "Effect of Thermocapillarity on the Production of Conducting Thin Film in the Presence of a Transverse Magnetic Field," Z. Angew. Math. Mech., **78**, pp. 635–640.
[38] Dandapat, B. S., and Layek, G. C., 1999, "Spin Coating in the Presence of a Transverse Magnetic Field and Non-Uniform Rotation: A Numerical Study," J. Phys. D, **32**, pp. 2483–2491.
[39] Usha, R., and Götz, T., 2001, "Spinning of a Liquid Film Flow on a Rotating Disk in the Presence of a Magnetic Field—A Numerical Solution," Acta Mech., **30**, pp. 1–15.
[40] Usha, R., and Uma, B., 2001, "Flow of a Thin Liquid Film Over a Rough Rotating Disk in the Presence of Transverse Magnetic Field," Z. Angew. Math. Phys., **52**, pp. 793–809.
[41] Usha, R., and Uma, B., 2002, "The Role of Induced Air Shear on the Development of a Conducting Fluid Film Over a Rough Spinning Disk in the Presence of a Transverse Magnetic Field," Z. Angew. Math. Mech., **82**, pp. 211–216.
[42] Sparrow, E. M., and Cess, R. D., 1962, "Magnetohydrodynamic Flow and Heat Transfer About a Rotating Disk," J. Appl. Mech., **29**, pp. 181–187.
[43] Neuringer, J. L., and McIlroy, W., 1958, "Incompressible Two-Dimensional Stagnation-Point Flow of an Electrically Conducting Viscous Fluid in the Presence of a Magnetic Field," J. Aeronaut. Sci., **25**, pp. 194–198.
[44] Rossow, V. J., 1958, "Magnetohydrodynamic Analysis of Heat Transfer Near a Stagnation Point," J. Aeronaut. Sci., **25**, pp. 234–235.
[45] Rathbun, A. S., 1961, "On the Flow of an Electrically Conducting Fluid Toward a Stagnation Point in the Presence of a Magnetic Field," Ph.D. thesis, University of Pittsburgh, Pittsburgh.
[46] Kumari, M., and Nath, G., 2004, "Unsteady MHD Film Flow Over a Rotating Infinite Disk," Int. J. Eng. Sci., **42**, pp. 1099–1117.
[47] Meyerhofer, D., 1998, "Characteristics of Resist Films Produced by Spinning," J. Appl. Phys., **47**, pp. 3993–3997.
[48] Benney, D. J., 1966, "Long Waves on Liquid Film," J. Math. Phys., **45**, pp. 150–155.
[49] Atherton, R. W., and Homsy, G. M., 1976, "On the Derivation of Evolution Equations for Interfacial Waves," Chem. Eng. Commun., **2**, pp. 57–77.
[50] Williams, M. B., and Davis, S. H., 1982, "Nonlinear Theory of Film Rupture," J. Colloid Interface Sci., **90**, pp. 220–228.
[51] Burelbach, J. P., Bankoff, S. G., and Davis, S. H., 1988, "Nonlinear Stability of Evaporating/Condensing Liquid Films," J. Fluid Mech., **195**, pp. 463–494.
[52] Tu, Y., 1983, "Depletion and Retention of Fluid on a Rotating Disk," ASME J. Lubr. Technol., **105**, pp. 625–629.
[53] Hwang, J. H., and Ma, F., 1989, "On the Flow of a Thin Liquid Film Over a Rough Rotating Disk," J. Appl. Phys., **66**, pp. 388–394.
[54] Kim, J. S., Kim, S., and Ma, F., 1991, "On the Flow of a Thin Liquid Film Over a Rotating Disk," J. Appl. Phys., **69**, pp. 2593–2601.
[55] Joo, S. W., Bankoff, S. G., and Davis, S. H., 1991, "Long-Wave Instabilities of Heated Falling Films: Two-Dimensional Theory of Uniform Layers," J. Fluid Mech., **230**, pp. 117–146.
[56] Davis, S. H., 1976, "Stability of Time-Periodic Flows," Annu. Rev. Fluid Mech., **8**, pp. 57–74.

Micromechanical Interpretation of the Dissipation Associated With Mode I Propagation of Microcracks in Brittle Materials

Bernhard Pichler[1]
Laboratory for Materials and Structures (LMSGC),
Ecole Nationale des Ponts et Chaussées (ENPC),
6 et 8, Avenue Blaise Pascal,
F-77455 Marne-la-Vallée, France
e-mail: bernhard.pichler@lmsgc.enpc.fr

Luc Dormieux
Laboratory for Materials and Structures (LMSGC),
Ecole Nationale des Ponts et Chaussées (ENPC),
6 et 8, Avenue Blaise Pascal,
F-77455 Marne-la-Vallée, France
e-mail: luc.dormieux@lmsgc.enpc.fr

This paper deals with the dissipation associated with quasistatic microcracking of brittle materials exhibiting softening behavior. For this purpose an elastodamaging cohesive zone model is used, in which cohesive tractions decrease (during crack propagation) with increasing displacement discontinuities. Constant cohesive tractions are included in the model as a limiting special case. Considering a representative volume element containing a dilute distribution of many parallel microcracks, we quantify energy dissipation associated with mode I microcrack propagation. This is done in the framework of thermodynamics, without restricting assumptions on the size of the cohesive zones. Model predictions are compared with exact solutions, which are accessible for constant cohesive tractions. The proposed model reliably predicts both onset of crack propagation and the dissipation during microcracking. It is shown that the energy release rate is virtually equal to the area under the softening curve, if the microscopic tensile strength is at least twice as large as the macroscopic tensile strength. This result justifies approaches relying on the concept of constant energy release rate, such as those frequently used in the engineering practice. [DOI: 10.1115/1.3086594]

Keywords: cohesive zone model, damage, dissipation, microcrack propagation brittle materials, material softening, elastodamaging cohesive zone, micromechanics, theorem of minimum potential energy, thermodynamics, cohesive zone size, energy release rate, Lebesgue's dominated convergence theorem

1 Introduction

Continuum micromechanics provides a promising theoretical framework for studying microcracking in brittle media such as cementitious materials or rock. The important role that micromechanical stiffness estimates play in describing the damaging influence of microcracks on the macroscopic (apparent) stiffness of the studied material can be seen in Refs. [1–3] and the references therein. Recently, such stiffness estimates were combined with crack propagation criteria from linear elastic fracture mechanics, see Refs. [4,5]. These combined fracture-micromechanics models predict that stable microcrack propagation is associated with macroscopic strain softening (i.e., with decreasing macrostresses, but increasing macrostrains), as is observed in macroscopic laboratory testing of concrete and sandstone under uniaxial tension, see, e.g., Ref. [6]. It is emphasized that the two combined fracture-micromechanics models predict (in agreement with experimental observations) that microcracking is associated with energy dissipation. Notably, in both approaches (i) the microcracks are represented as flat penny-shaped pores (i.e., as classical noncohesive cracks), and (ii) linear elastic behavior of the solid matrix is taken into account. These aspects raise need for theoretical improvement, since (i) there are infinite stresses at the edges of noncohesive cracks, and since (ii) both models are not able to describe the thermodynamical source of dissipation, but owe an explanation of *how* dissipation occurs.

Cohesive zone models resolve the problem of the stress singularities at the edges of classical cracks. Nowadays, a great variety of such models are described in the open literature, representing extensions of the independent pioneering works of Dugdale [7] and of Barenblatt [8,9], respectively. Considering quasistatic microcracking of brittle materials exhibiting softening behavior, an elastodamaging cohesive zone model was recently proposed by the authors [10]. Making use of the theorem of minimum potential energy, the size of the cohesive zone was estimated based on a class of kinematically admissible displacement fields, considering a representative volume element (RVE) containing a dilute distribution of many parallel noninteracting microcracks.

Herein, we use the elastodamaging cohesive zone model of Pichler and Dormieux [10] in order to explain how dissipation occurs during quasistatic microcracking of brittle materials exhibiting softening behavior. We aim at quantifying the amount of dissipated energy, in the classical framework of thermodynamics. Thereby, we *neither* restrict our considerations to cases where the cohesive zone is small, as compared with the crack radius, *nor* do we consider that the cohesive zone size is constant during crack propagation; both of which are assumptions frequently used in the pertinent literature. The outcome of this paper is meant to deliver justifications (i) for micromechanical stiffness estimates for microcracked media relying on classical noncohesive cracks, see Refs. [1–3], and (ii) for approaches relying on the concept of critical (dissipative) energy release rate. Achieving these aims would open the door for further enhancing the micromechanical description of microcracking processes in media such as cementitious materials and rock.

This paper is organized as follows. In Sec. 2, we briefly describe the findings published in Ref. [10], which are prerequisites for the original contributions of this paper. The essential starting points for the present work comprise the envisaged constitutive behavior of the cohesive zones, as well as the estimates of the

[1]On leave from Institute for Mechanics of Materials and Structures, Vienna University of Technology (TU Wien), Karlsplatz 13/202, A-1040 Vienna, Austria.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 22, 2007; final manuscript received September 15, 2008; published online April 22, 2009. Review conducted by Zhigang Suo.

cohesive zone size. Section 3 contains original contributions. In the framework of thermodynamics, we calculate the dissipation of a single cohesive crack propagating in mode I. First, we deal with constant cohesive tractions, then we extend our analysis toward consideration of softening behavior in the cohesive zones. Finally, we study the dissipation within an RVE containing a dilute distribution of many propagating cohesive microcracks, which allows for deriving an improved dilute stiffness estimate of the microcracked RVE. Section 4 deals with the validation of the model described in Secs. 2 and 3. We compare model predictions with accessible exact solutions, referring to the special case of constant cohesive tractions. We list model predictions for (i) the size of the cohesive zone, (ii) the relation between remote tensile stress and the crack opening at the inner edges of the cohesive zones, and (iii) the dissipation during crack propagation. Model predictions are compared with corresponding exact solutions. With regard to point (iii), this comparison requires an exact expression for the dissipation of a propagating elastodamaging cohesive crack with constant cohesive tractions, which is so far not available in literature. The derivation of this exact dissipation expression complements the original contributions of this paper. Section 5 contains a final discussion, conclusions, and an outlook to future research work. There, we also incorporate the new results into the much wider context of the combined fracture-micromechanics models of Dormieux et al. [4] and Pichler et al. [5], and we also point out how both the scientific community, as well as practical engineers, can gain benefit from the work described herein.

2 Fundamentals of the Proposed Approach

2.1 Elastodamaging Cohesive Zones: Constitutive Softening Law. Herein, we study quasistatic, i.e., *time-independent*, microcracking[2] such as that observed in displacement-controlled tensile tests on cementitious materials or on rock. Therefore, we do not make use of time-dependent cohesive zone models designed to account for dynamic crack growth, such as, e.g., the appealing approach of Glennie [11] introducing the magnitude of cohesive tractions to be loading-rate dependent, or the one of Siegmund and Needleman [12] modeling *elastoviscoplastic* cohesive behavior. Due to the brittle nature of cementitious materials and of rock observed under tensile loading, ductile effects are unlikely to occur and, hence, elastoplastic cohesive zone models, such as the ones by Tvergaard and Hutchinson [13], do not appear to be suitable in the present context. With regard to analyses of macrocracking of structures made of cementitious materials or of rock, a *softening-type* cohesive behavior is commonly introduced to reproduce results from structural testing. Accordingly, cohesive tractions decrease with increasing displacement discontinuities, see the pioneering approach of Hillerborg et al. [14], which inspired a large number of follow-up works, such as, e.g., Refs. [15–17]. In accordance with these well-accepted macroscopic approaches, we consider here cohesive softening behavior on a lower scale of observation. Consequently, we do not consider brittle microcracking with cohesive hardening behavior, in which cohesive tractions *increase* with increasing displacement discontinuities, see, e.g., results from molecular dynamics analyses of Yamakov et al. [18] dealing with brittle microcracking along the $\Sigma 99$ grain boundary in aluminum under isotropic tension. Nevertheless, such nanomechanical analyses raise the hope that molecular dynamics simulations will soon provide insight into cohesive zone behavior of media as complex as cementitious materials or rock.

In accordance with our focus on quasistatic microcracking of brittle materials exhibiting softening behavior, we consider the following elastodamaging cohesive zone model. We envision ahead of every microcrack a zone in which some of the bonds between formerly adjacent material particles are already broken, whereas other bonds are still intact. To model this three-dimensional damaged domain ahead of the microcracks, we introduce a two-dimensional interface, referred to as cohesive zone. Suitable for studying mode I propagation of microcracks, we consider normal tractions T_n and normal displacement discontinuities $[\![\xi_n]\!]$ in the cohesive zone, both of which normal with respect to the two-dimensional interface. The constitutive relation between $[\![\xi_n]\!]$ and T_n is given by the elastic law

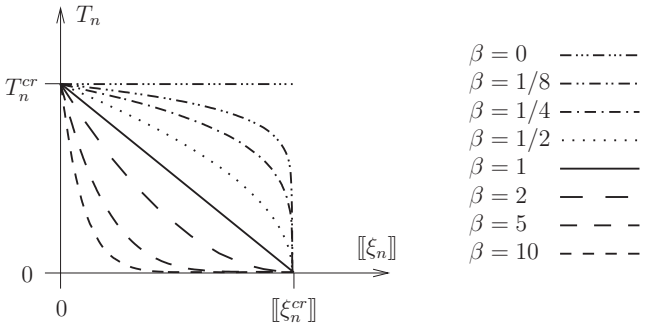


Fig. 1 Elastodamaging cohesive zone law (defined in Eqs. (1) and (2)), relating cohesive normal tractions T_n to the normal displacement discontinuities $[\![\xi_n]\!]$ (also referred to as "softening curve"); plotted for different values of the power law exponent β

$$T_n = K([\![\xi_n]\!])[\![\xi_n]\!] \tag{1}$$

In Eq. (1), K denotes the elastic stiffness of the cohesive zone, which depends on the displacement discontinuity $[\![\xi_n]\!]$. The latter represents the damage variable, motivating the expression "elastodamaging" cohesive zone model. With increasing displacement discontinuity, i.e., with increasing damage, $K([\![\xi_n]\!])$ decreases according to the power law

$$K([\![\xi_n]\!]) = \frac{T_n^{\mathrm{cr}}}{[\![\xi_n]\!]}\left(1 - \frac{[\![\xi_n]\!]}{[\![\xi_n^{\mathrm{cr}}]\!]}\right)^{\beta} \tag{2}$$

where T_n^{cr}, $[\![\xi_n^{\mathrm{cr}}]\!]$, and $\beta \geq 0$ are material constants denoting the maximum bearable (*critical*) tensile microstress, the (critical) displacement discontinuity at which no tensile traction can be transmitted across the interface anymore, and a dimensionless constant governing the shape of the T_n-$[\![\xi_n]\!]$ relationship (see Fig. 1), respectively.

In order to provide more insight into the cohesive zone stiffness (Eq. (2)), it is useful to consider a material point P, which lies originally ahead of a cohesive zone, being part of the (sane) elastic material. The chosen approach implies that P becomes part of the cohesive zone, once the stress tensor component σ_{nn} at P becomes equal to T_n. Then, since progressive microcracking is assumed, $[\![\xi_n]\!]$ at P monotonously increases, i.e., the current value of $[\![\xi_n]\!]$ at P is—at the same time—also the maximum displacement discontinuity that P has experienced so far

$$[\![\xi_n]\!] = \max_t [\![\xi_n]\!] \tag{3}$$

This justifies introduction of $[\![\xi_n]\!]$ as the damage variable in Eq. (2). Unloading (resulting in closure of the cohesive zone) and reloading at constant damage (resulting in re-opening of the cohesive zone) can be accounted for simply by replacing $[\![\xi_n]\!]$ in Eq. (2) by $\max_t [\![\xi_n]\!]$, resulting in a linear elastic relationship between T_n and $[\![\xi_n]\!]$ reading as

[2]Microcracking is used as a synonym for progressive microcrack propagation.

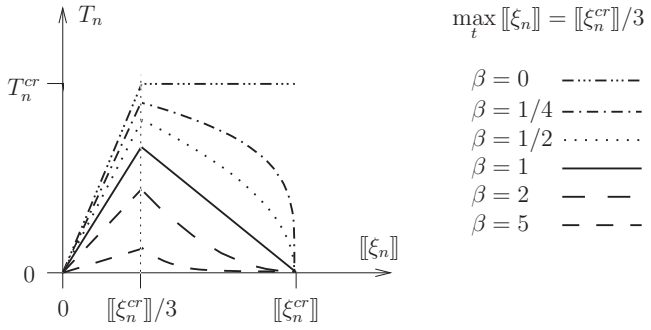


Fig. 2 Evaluation of elastodamaging cohesive zone law (4), for different values of the power law exponent β, and for $\max_t[\![\xi_n]\!] = [\![\xi_n^{cr}]\!]/3$

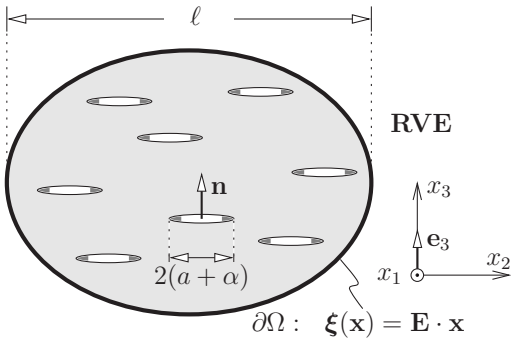


Fig. 3 Two-dimensional illustration of a three-dimensional representative volume element comprising an elastic matrix and parallel cohesive microcracks of identical size and orientation

$$T_n = K(\max_t[\![\xi_n]\!])[\![\xi_n]\!], \quad K(\max_t[\![\xi_n]\!]) = \frac{T_n^{\mathrm{cr}}}{\max_t[\![\xi_n]\!]}\left(1 - \frac{\max_t[\![\xi_n]\!]}{[\![\xi_n^{\mathrm{cr}}]\!]}\right)^{\beta} \tag{4}$$

Equation (4) implies that the elastic cohesive zone stiffness—which becomes "visible" upon unloading—depends on T_n^{cr}, $[\![\xi_n^{\mathrm{cr}}]\!]$, β, and $\max_t[\![\xi_n]\!]$, see (see Fig. 2). Since the loading/unloading branches in Fig. 2 (see diagram curves in the abscissa range from zero to $[\![\xi_n^{\mathrm{cr}}]\!]/3$) are *linear elastic*, they are related to a reversible process (no dissipation). This is a significant difference to the dissipative *hardening* branches described in Ref. [18]. Anyway, unloading-reloading cycles are not the present focus, but microcracking analyses will be carried out, for which the cohesive zone model formulation provided in Eqs. (1) and (2) is to be used.

Linear softening curves, such as those used in, e.g., Refs. [14,17], and quadratic softening curves, as used in, e.g., Ref. [19], are included in Eq. (2) as the special cases $\beta=1$ and $\beta=2$, respectively. While the cited works prove special cases of the proposed model to be useful in research and in the engineering practice, the aim of this paper is to deal with any value of β being equal to or greater than zero. Notably, the power law (2) is appealing, since it allows for studying an infinite amount of qualitatively different T_n-$[\![\xi_n]\!]$ relationships, based on three material constants only. A rather simple bilinear T_n-$[\![\xi_n]\!]$ relationship, for instance, would already require the introduction of four material constants.

2.2 RVE-Related Estimation of the Cohesive Zone Size. Herein, we consider a representative volume element containing an elastic matrix and many parallel cohesive microcracks. The cracks are assumed to be significantly smaller than the characteristic dimensions of the RVE (separation of scales), and to be arranged disordered, rendering the macroscopic behavior of the RVE transversely isotropic. In more detail, we consider N cracks per unit volume, with normal vectors pointing in the x_3 direction ($\mathbf{n}=\mathbf{e}_3$). The cracks are represented as circular slits of radius $a+\alpha$, where α denotes the width of the cohesive zone, see Figs. 3 and 4. For the sake of simplicity, we assume a dilute distribution of cohesive microcracks. Accordingly we do not account for crack interaction, but we assume that the strain field in the vicinity of a crack within the RVE can be suitably approximated by the strain field in the vicinity of a single crack embedded in an infinite matrix, which is subjected at infinity to the same Hashin boundary conditions as the RVE, described next: The RVE is subjected, on its boundary $\partial\Omega$, to displacement vectors $\boldsymbol{\xi}(\mathbf{x})$ corresponding to uniform macroscopic strains $\mathbf{E}$ reading as [20]

$$\boldsymbol{\xi}(\mathbf{x}) = \mathbf{E}\cdot\mathbf{x} \quad \text{with} \quad \mathbf{x} \in \partial\Omega \tag{5}$$

In Eq. (5), $\boldsymbol{\xi}$ and $\mathbf{x}$ denote the displacement vector and the position vector, respectively, and $\mathbf{E}$ stands for the tensor of "prescribed" macroscopic strains. In accordance with the assumed cracking mode I, we consider macroscopic strains in the form

$$\mathbf{E} = E_{33}\mathbf{e}_3 \otimes \mathbf{e}_3 + E_{\ell at}(\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2) \tag{6}$$

With regard to the theoretical basis for the estimation of the cohesive zone size, it is worth emphasizing that we have restricted our considerations to stable and time-independent microcracking, such as observed in displacement-controlled macroscopic laboratory tests on concrete or sandstone [6]. Such tests can be stopped at any arbitrary time instant, even in the softening regime [21]. If this is done, the stresses within the specimen will be in equilibrium with the current macroscopic loading and *no* irreversible processes will take place since the cracks will remain stationary at their current sizes (no increase in damage). The situation obtained after stopping the tension test is—assessed from the viewpoint of the proposed model—a purely elastic equilibrium configuration. Accordingly, the theorem of minimum potential energy can be used to estimate the size of the cohesive zone for that configuration. The process of microcracking, in turn, can be understood as the transition from one elastic equilibrium configuration to another. Since the amount of damage increases during this transition, it is a dissipative process, and its analysis requires a framework of thermodynamics, as will be considered in Sec. 3.

In the framework of the theorem of minimum potential energy, kinematically admissible displacement fields are considered, i.e., continuous and piecewise continuously differentiable displacement fields satisfying the displacement boundary conditions (herein Eq. (5)). Such displacements fields are used to calculate the elastic energy stored in the considered structure. The theorem states, that the elastic energy referring to the *displacement solution* of the underlying elasticity problem is a lower bound for the elastic energy referring to *any* other kinematically admissible displacement field, see, e.g., Ref. [22]. Commonly, *classes* of kinematically admissible displacement fields are considered, which are characterized by one or more optimization variables. Calculating the elastic energy, and minimizing it with respect to the optimization variable(s), yields an "optimal" kinematically admissible displacement field.

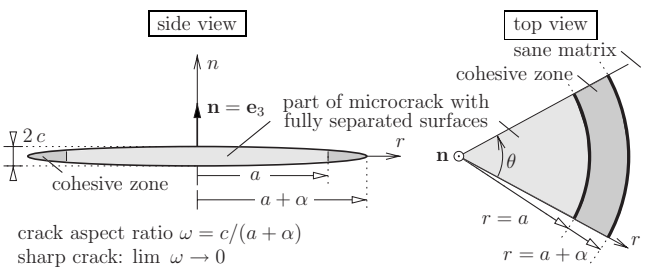


Fig. 4 Microcrack comprising an annulus-shaped cohesive zone

Herein we consider the class of kinematically admissible displacement fields analyzed in Sec. 5 of Ref. [10], introducing the size of the cohesive zone as an optimization variable. This class of displacement fields is equal to the displacement solution of a *fictitious* RVE (subjected to the boundary conditions (5)), which comprises an elastic matrix and a dilute distribution of N sharp penny-shaped parallel ($\mathbf{n}=\mathbf{e}_3$) microcracks per unit volume, with radius $a+\alpha$, but *without* cohesive tractions. The macroscopic stiffness estimate of such an RVE reads as [23,24]

$$\mathbb{C}_{\text{hom}}^{\text{dil}} = \mathbb{C}_m : \left[\mathbb{I} - \frac{4\pi}{3} N (a+\alpha)^3 \mathbb{T}\right] \tag{7}$$

In Eq. (7), $\mathbb{C}_m$ is the isotropic elastic stiffness tensor of the matrix, $\mathbb{I}$ is the symmetric fourth-order unity tensor, and the fourth-order tensor $\mathbb{T}$ is defined as

$$\mathbb{T} = \lim_{\omega \to 0} \omega [\mathbb{I} - \mathbb{S}_c^m(\omega)]^{-1} \tag{8}$$

In Eq. (8), ω is the crack aspect ratio (see also Fig. 4), and $\mathbb{S}_c^m$ is Eshelby's tensor. The nonvanishing components of $\mathbb{T}$ read as [23,24]

$$T_{3311} = T_{3322} = \frac{4\nu_m(1-\nu_m)}{(1-2\nu_m)\pi}, \quad T_{3333} = \frac{4(1-\nu_m)^2}{(1-2\nu_m)\pi},$$

$$T_{1313} = T_{2323} = \frac{2(1-\nu_m)}{(2-\nu_m)\pi} \tag{9}$$

with the symmetries $T_{ijkl}=T_{jikl}=T_{ijlk}$. ν_m denotes Poisson's ratio of the matrix. The crack opening displacements related to $\mathbb{C}_{\text{hom}}^{\text{dil}}$ represent the kinematically admissible displacement discontinuities and read as

$$[\![\xi_n]\!] = 2\lambda \sqrt{(a+\alpha)^2 - r^2} \tag{10}$$

where λ is a macroscopic loading parameter defined as

$$\lambda = E_{33} T_{3333} + 2E_{\ell at} T_{3311} \tag{11}$$

Remarkably, Eq. (10) is identical to the classical solution for the crack opening displacement of a single penny-shaped crack embedded in an infinite, elastic, and isotropic matrix, derived in the framework of linear elastic fracture mechanics, see, e.g., Ref. [25].

Calculating the potential energy of the *real* RVE, based on the described class of kinematically admissible displacement fields, and minimizing it with respect to the optimization variable "cohesive zone size," delivers the following equation for the optimal cohesive zone size [10]:

$$\frac{2T_n^{\text{cr}}}{1+\beta}\left[[\![\xi_n^{\text{cr}}]\!] - \frac{([\![\xi_n^{\text{cr}}]\!] - 2\lambda\sqrt{(2a+\alpha)\alpha})^{1+\beta}}{[\![\xi_n^{\text{cr}}]\!]^{\beta}}\right] - \frac{E_m(a+\alpha)\lambda^2\pi}{2(1-\nu_m^2)} = 0 \tag{12}$$

During crack propagation $[\![\xi_n]\!]$ is equal to $[\![\xi_n^{\text{cr}}]\!]$ at $r=a$. Accordingly, Eq. (10) is specified for $r=a$, the resulting expression is set equal to $[\![\xi_n^{\text{cr}}]\!]$, and the obtained equation is solved for the macroscopic loading parameter λ, which yields

$$\lambda = \frac{[\![\xi_n^{\text{cr}}]\!]}{2\sqrt{(2a+\alpha)\alpha}} \tag{13}$$

Inserting Eq. (13) into Eq. (12) and solving the resulting equation for α yields optimal estimates of the cohesive zone size *during crack propagation* as

$$\alpha = a(\chi - 1 + \sqrt{\chi^2 + 1}) \tag{14}$$

with

$$\chi = \frac{\pi E_m [\![\xi_n^{\text{cr}}]\!](1+\beta)}{32 T_n^{\text{cr}} a (1-\nu_m^2)} \tag{15}$$

It is emphasized that the displacement discontinuity function (10) and the cohesive zone size (Eq. (14)) converge toward exact solutions in the limit $\beta \to \infty$. For any positive value of β smaller than infinity, $[\![\xi_n]\!]$ and α are best possible estimates within the chosen class of kinematically admissible displacement fields.

3 Dissipation of Elastodamaging Cohesive Microcracks

3.1 Thermodynamics Background. Herein, we deal with the rate of energy, which is dissipated by cohesive cracks propagating through the considered RVE. "Dissipation" denotes the transformation of mechanical energy into other types of energy, which cannot be retransformed into efficient mechanical work. In the classical framework of thermodynamics, the dissipation is equal to the difference between the rate of work of the external forces (i.e., the rate of work expended on the considered system by its environment) and the rate of elastic internal energy, see, e.g., Ref. [26].

Let us first consider a system, which consists of the matrix of the considered RVE only, i.e., we cut the matrix free from all the cohesive cracks. Therefore, the action expended by the cohesive cracks on the matrix, i.e., the cohesive tractions, become part of the external forces of the investigated system. Since the behavior of the matrix is considered herein as linear elastic, the rate of work of the external forces (the ones acting on the boundary of the RVE and the ones acting at the locations of the cut-away cohesive zones) is always equal to the rate of elastic internal energy. Consequently, no dissipation occurs in the matrix, but all dissipative effects within the RVE obviously occur in the elastodamaging cohesive cracks.

Let us now consider a system, which consists of one single cohesive crack, cut free from the considered RVE. The external forces of this system are given by the action of the matrix on the crack, i.e., by the cohesive tractions T_n. Consequently, the rate of work expended on the crack by its environment reads as

$$\dot{W}_{\text{crack}}^{\text{ex}} = \int_{r=a}^{a+\alpha} T_n [\![\dot{\xi}_n]\!] 2r\pi dr \tag{16}$$

In Eq. (16), $[\![\dot{\xi}_n]\!]$ denotes the rate of the displacement discontinuity. The elastic internal energy stored in the cohesive zone can be expressed as

$$W_{\text{crack}}^{\text{in}} = \int_{r=a}^{a+\alpha} \psi 2r\pi dr \tag{17}$$

where ψ is the elastic free energy per unit area of the cohesive zone. The rate of $W_{\text{crack}}^{\text{in}}$ follows from deriving Eq. (17) with respect to time

$$\dot{W}_{\text{crack}}^{\text{in}} = \frac{d}{dt}\int_{r=a}^{a+\alpha} \psi 2r\pi dr = \int_{r=a}^{a+\alpha} \dot{\psi} 2r\pi dr \tag{18}$$

In Eq. (18), the temporal derivatives of the integration bounds yield vanishing contributions because during crack propagation $\psi=0$, both at the inner and at the outer edge of the cohesive zone; since at $r=a: T_n=0$, and since at $r=a+\alpha: [\![\xi_n]\!]=0$. The rate of dissipated energy of one cohesive crack is obtained from combining Eqs. (16) and (18) as

$$\dot{\mathcal{D}} = \dot{W}_{\text{crack}}^{\text{ex}} - \dot{W}_{\text{crack}}^{\text{in}} = \int_{r=a}^{a+\alpha} (T_n [\![\dot{\xi}_n]\!] - \dot{\psi}) 2r\pi dr \tag{19}$$

Equation (19) is valid for any constitutive cohesive zone behavior. Considering elastodamaging behavior, with the displacement dis-

continuity as the only damage variable, the elastic free energy density reads as

$$\psi = \tfrac{1}{2}[\![\xi_n]\!]K([\![\xi_n]\!])[\![\xi_n]\!] \tag{20}$$

Insertion of Eq. (20) into Eq. (19), and considering Eq. (1), yields the rate of dissipation during propagation of one *elastodamaging* cohesive crack as

$$\begin{aligned}\dot{\mathcal{D}}_{\text{crack}} &= -\int_{r=a}^{a+\alpha} \frac{1}{2}[\![\xi_n]\!]\dot{K}([\![\xi_n]\!])[\![\xi_n]\!]2r\pi dr \\ &= -\pi\int_{r=a}^{a+\alpha} [\![\xi_n]\!]^2 \frac{dK([\![\xi_n]\!])}{d[\![\xi_n]\!]}[\![\dot{\xi}_n]\!]rdr\end{aligned} \tag{21}$$

After the second equal sign in Eq. (21), it was taken into account that K depends only on $[\![\xi_n]\!]$. Specifying Eq. (21) for Eq. (2), finally yields the rate of dissipation during propagation of one microcrack with elastodamaging behavior, described by the *power law* (2)

$$\begin{aligned}\dot{\mathcal{D}}_{\text{crack}} &= T_n^{\text{cr}}\pi\int_{r=a}^{a+\alpha}\left(1-\frac{[\![\xi_n]\!]}{[\![\xi_n^{\text{cr}}]\!]}\right)^{\beta}[\![\dot{\xi}_n]\!]rdr \\ &+ T_n^{\text{cr}}\pi\int_{r=a}^{a+\alpha}\beta\frac{[\![\xi_n]\!]}{[\![\xi_n^{\text{cr}}]\!]}\left(1-\frac{[\xi_n]}{[\xi_n^{\text{cr}}]}\right)^{\beta-1}[\![\dot{\xi}_n]\!]rdr\end{aligned} \tag{22}$$

In the sequel, Eq. (22) will be evaluated based on the optimal kinematically admissible displacement discontinuity (Eq. (10)). This will be done first for constant cohesive tractions ($\beta=0$), see Sec. 3.3, while in Sec. 3.4 we consider softening cohesive behavior ($\beta>0$). It is recalled that estimation of the size of the cohesive zone was done on the level of an RVE. Therefore information on the RVE is also present in the subsequent single crack-related analyses. For the sake of model validation, Eq. (22) will be also evaluated for the exact shape of the displacement discontinuity, which is available for the special case of constant cohesive tractions ($\beta=0$), see Sec. 4.2.

3.2 Displacement Discontinuity (Eq. (10)) and Its Rate During Crack Propagation. The displacement discontinuity during crack propagation is obtained by inserting λ from Eq. (13) into Eq. (10)

$$[\![\xi_n]\!] = [\![\xi_n^{\text{cr}}]\!]\sqrt{\frac{(a+\alpha)^2-r^2}{(2a+\alpha)\alpha}} \tag{23}$$

Equation (23) satisfies the condition for crack propagation, i.e., that $[\![\xi_n]\!]$ is equal to $[\![\xi_n^{\text{cr}}]\!]$ at the inner edge of the cohesive zone (at $r=a$). With regard to the time derivative $[\![\dot{\xi}_n]\!]$ appearing in Eq. (22), we introduce the (increasing) crack radius a as the parameter for crack propagation, i.e., we set

$$[\![\dot{\xi}_n]\!] = \frac{d[\![\xi_n]\!]}{da}\dot{a} = \left(\frac{\partial[\![\xi_n]\!]}{\partial a}+\frac{\partial[\![\xi_n]\!]}{\partial\alpha}\frac{d\alpha}{da}\right)\dot{a} \tag{24}$$

In Eq. (24), the fact that α depends, through Eqs. (14) and (15), on the crack radius, i.e., that the size of the cohesive zone changes during crack propagation, was accounted for. The derivative of α with respect to a follows from Eqs. (14) and (15) as

$$\frac{d\alpha}{da} = \sqrt{\frac{1}{\chi^2+1}}-1 \tag{25}$$

The rate of displacement discontinuity during crack propagation is obtained by inserting both Eqs. (23) and (25) into Eq. (24) as

$$[\![\dot{\xi}_n]\!] = \frac{(r^2-a^2)(a+\alpha)+a\sqrt{\chi^2+1}[(a+\alpha)^2-r^2]}{\sqrt{(a+\alpha)^2-r^2}\sqrt{(2a+\alpha)^3\alpha^3}\sqrt{\chi^2+1}}[\![\xi_n^{\text{cr}}]\!]\dot{a} \tag{26}$$

For clarity of the subsequent calculations, $[\![\xi_n]\!]$ from Eq. (10) and $[\![\dot{\xi}_n]\!]$ from Eq. (26) are expressed as

$$[\![\xi_n]\!] = [\![\xi_n^{\text{cr}}]\!]\Xi(\rho) \quad \text{and} \quad [\![\dot{\xi}_n]\!] = \Lambda(\rho)\frac{[\![\xi_n^{\text{cr}}]\!]}{a}\dot{a} \tag{27}$$

with the dimensionless functions

$$\begin{aligned}\Xi(\rho) &= \sqrt{\frac{(1+\kappa)^2-\rho^2}{(2+\kappa)\kappa}} \quad \text{and} \\ \Lambda(\rho) &= \frac{(\rho^2-1^2)(1+\kappa)+1\sqrt{\chi^2+1}[(1+\kappa)^2-\rho^2]}{\sqrt{(1+\kappa)^2-\rho^2}\sqrt{(2+\kappa)^3\kappa^3}\sqrt{\chi^2+1}}\end{aligned} \tag{28}$$

and the dimensionless parameters

$$\rho = \frac{r}{a} \quad \text{and} \quad \kappa = \frac{\alpha}{a} = \chi - 1 + \sqrt{\chi^2+1} \tag{29}$$

3.3 Single Crack-Related Dissipation Considering Constant Cohesive Tractions. Specifying the rate of dissipation (22) for constant cohesive tractions, i.e., for $\beta=0$, and introducing the dimensionless parameter ρ yield

$$\begin{aligned}\dot{\mathcal{D}}_{\text{crack}} &= T_n^{\text{cr}}a^2\pi\int_{\rho=1}^{1+\kappa}[\![\dot{\xi}_n]\!]\rho d\rho + T_n^{\text{cr}}a^2\pi\lim_{\beta\to 0}\left[\int_{\rho=1}^{1+\kappa}\beta\left(1\right.\right. \\ &\left.\left.-\frac{[\![\xi_n]\!]}{[\![\xi_n^{\text{cr}}]\!]}\right)^{\beta-1}\frac{[\![\xi_n]\!]}{[\![\xi_n^{\text{cr}}]\!]}[\![\dot{\xi}_n]\!]\rho d\rho\right]\end{aligned} \tag{30}$$

Herein, we specify Eq. (30) for $[\![\xi_n]\!]$ and $[\![\dot{\xi}_n]\!]$ from Eqs. (27) and (28). Integration of the first term into the right-hand side of Eq. (22), denoted subsequently as $\dot{\mathcal{D}}_1$, is straightforward and yields

$$\dot{\mathcal{D}}_1 = T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]a\pi\dot{a}\int_{\rho=1}^{1+\kappa}\Lambda(\rho)\rho d\rho = T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]a\pi\dot{a}\left(1+\frac{2\chi}{3\sqrt{\chi^2+1}}\right) \tag{31}$$

Evaluation of the second term on the right-hand side of Eq. (22), denoted subsequently as $\dot{\mathcal{D}}_2$, deserves special attention

$$\begin{aligned}\dot{\mathcal{D}}_2 &= T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]a\pi\dot{a}\mathcal{L} \quad \text{with} \quad \mathcal{L} = \lim_{\beta\to 0}\left[\int_{\rho=1}^{1+\kappa}\beta(1\right. \\ &\left.-\Xi(\rho))^{\beta-1}\Xi(\rho)\Lambda(\rho)\rho d\rho\right]\end{aligned} \tag{32}$$

Though β tends to zero, $\mathcal{L}$ is not equal to zero since $(1-\Xi(\rho))^{\beta-1}\Xi(\rho)\Lambda(\rho)\rho$ tends to infinity at $\rho=1$ (i.e., at $r=a$). Based on Lebesgue's dominated convergence theorem, see, e.g., Ref. [27], it is shown in the Appendix that

$$\mathcal{L} = 1 \tag{33}$$

The rate of dissipated energy during propagation of a cohesive crack with constant cohesive tractions follows from combining Eqs. (31)–(33) as

$$\dot{\mathcal{D}}_{\text{crack}} = T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]2a\pi\dot{a}\left(1+\frac{\chi}{3\sqrt{\chi^2+1}}\right) \tag{34}$$

Dividing Eq. (34) by the rate of newly produced crack area

$$\dot{A} = \frac{d}{dt}(a^2\pi) = 2a\pi\dot{a} \tag{35}$$

yields the dissipated energy per newly produced crack area (commonly referred to as energy release rate)

$$\mathcal{G}_{\text{crack}} = \frac{\dot{\mathcal{D}}_{\text{crack}}}{\dot{A}} = T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]\left(1 + \frac{\chi}{3\sqrt{\chi^2+1}}\right) \tag{36}$$

3.4 Single Crack-Related Dissipation Considering Cohesive Softening Behavior. Herein, we consider the dissipation associated with propagation of a cohesive crack exhibiting softening behavior in terms of decreasing cohesive tractions with increasing displacement discontinuities. The rate of energy dissipated during crack propagation follows from inserting both Eqs. (23) and (26) into Eq. (22), and from solving the integrals for values of β greater than zero as

$$\dot{\mathcal{D}}_{\text{crack}} = \frac{T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]}{1+\beta}2a\pi\dot{a}\left(1 + \frac{\chi}{\sqrt{\chi^2+1}}\frac{(2+5\beta+\beta^2)}{(2+\beta)(3+\beta)}\right) \tag{37}$$

Remarkably, evaluation of Eq. (37) for the limit $\beta \to 0$ delivers the result obtained in Eq. (34), validating the derivations of Sec. 3.3 and of the Appendix. Dividing Eq. (37) by the rate of newly produced crack area (Eq. (35)), yields the energy release rate $\mathcal{G}_{\text{crack}}$ as

$$\mathcal{G}_{\text{crack}} = \frac{\dot{\mathcal{D}}_{\text{crack}}}{\dot{A}} = \frac{T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]}{1+\beta}\left(1 + \frac{\chi}{\sqrt{\chi^2+1}}\frac{(2+5\beta+\beta^2)}{(2+\beta)(3+\beta)}\right) \tag{38}$$

Equation (38) represents $\mathcal{G}_{\text{crack}}$ as a function of the dimensionless parameter χ, see Eq. (15), and of the three material constants T_n^{cr}, $[\![\xi_n^{\text{cr}}]\!]$, and β, characterizing the cohesive softening law, see Eqs. (1) and (2). In other words, Eq. (38) allows for the desired physical interpretation of the dissipative effects associated with mode I propagation of elastodamaging cohesive microcracks.

If $\chi \ll 1$, then the size of the cohesive zone is (i) almost constant during crack propagation, and it is (ii) significantly smaller than the crack radius, as follows from specification of Eq. (14) for $\chi \ll 1$:

$$\alpha \approx a\chi = \frac{\pi E_m[\![\xi_n^{\text{cr}}]\!](1+\beta)}{32T_n^{\text{cr}}(1-\nu_m^2)} \ll a \tag{39}$$

In turn, Eq. (38) reveals that if $\chi \ll 1$, the energy release rate $\mathcal{G}_{\text{crack}} \approx T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]/(1+\beta)$. The latter expression is equal to area under the related T_n-$[\![\xi_n]\!]$ diagram depicted in Fig. 1. This result is consistent with the traditional energy approach of dissipation based on the concept of critical (dissipative) energy release rate, i.e., it justifies approaches in which the toughness is *introduced* as the area under the softening curve, see, e.g., Ref. [16].[3]

If $\chi \not\ll 1$, then the size of the cohesive zone depends on the crack radius (i.e., α evolves during crack propagation), see Eqs. (14) and (15). In such cases, the energy release rate is larger than the area under the related T_n-$[\![\xi_n]\!]$ diagram, see Eq. (38).

3.5 Dissipation Within an RVE Comprising Many Propagating Cohesive Microcracks. Considering that all cracks within the RVE introduced in Sec. 2.2 are propagating, the rate of energy dissipated within the RVE is equal to the rate of energy dissipated per propagating crack, multiplied by the number of cracks propagating through the RVE, i.e.,

$$\dot{\mathcal{D}}_{\text{RVE}} = \dot{\mathcal{D}}_{\text{crack}} N\Omega = \frac{T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]}{1+\beta}2a\pi\dot{a}\left(1 + \frac{\chi}{\sqrt{\chi^2+1}}\frac{(2+5\beta+\beta^2)}{(2+\beta)(3+\beta)}\right)N\Omega \tag{40}$$

In Eq. (40), Ω denotes the volume of the RVE, and N is the number of cracks per unit volume. $\dot{\mathcal{D}}_{\text{RVE}}$ can be calculated alternatively by a macroscopic reasoning: The rate of energy dissipated within the RVE is equal to the difference between the rate of work expended on the RVE by its environment (i.e., by the external forces acting on the boundary of the RVE) and the rate of elastic internal energy stored within the RVE (both in the matrix and in the cohesive zones). The related mathematical formulation reads, by analogy to Eq. (21), as

$$\dot{\mathcal{D}}_{\text{RVE}} = \dot{W}_{\text{RVE}}^{\text{ex}} - \dot{W}_{\text{RVE}}^{\text{in}} = -\frac{\Omega}{2}\mathbf{E}:\dot{\mathbb{C}}_{\text{hom}}:\mathbf{E} \tag{41}$$

In Eq. (41), $\dot{\mathbb{C}}_{\text{hom}}$ is the rate of the homogenized macroscopic stiffness tensor. Introducing again the (increasing) crack radius a as the parameter for crack propagation, Eq. (41) reads as

$$\dot{\mathcal{D}}_{\text{RVE}} = -\dot{a}\frac{\Omega}{2}\mathbf{E}:\frac{d}{da}(\mathbb{C}_{\text{hom}}^{\text{idil}}):\mathbf{E} \quad \text{with} \quad \frac{d}{da} = \frac{\partial}{\partial a} + \frac{\partial}{\partial \alpha}\frac{d\alpha}{da} \tag{42}$$

In Eq. (42), it was taken into account that we will estimate $\mathbb{C}_{\text{hom}}$ by $\mathbb{C}_{\text{hom}}^{\text{idil}}$, an *improved dilute* estimate, which accounts explicitly for the cohesive zones. By setting Eq. (40) equal to Eq. (42), $\mathbb{C}_{\text{hom}}^{\text{idil}}$ can be identified as

$$\mathbb{C}_{\text{hom}}^{\text{idil}} = \mathbb{C}_m:\left[\mathbb{I} - \frac{4\pi}{3}N(a+\alpha)^3\left(1 - \frac{3\kappa^2(2+\kappa)^2}{2\chi(2+\beta)(3+\beta)(1+\kappa)^3}\right)\mathbb{T}\right] \tag{43}$$

where $\kappa(\chi)$ was defined in Eq. (29). The improved stiffness estimate (Eq. (43)) differs from the classical one (Eq. (7)) by the second term in the parenthesis of Eq. (43). This additional term accounts for the cohesive zones. It decreases with increasing β, and it decreases with decreasing χ.

4 Model Validation

This section aims at validation of the model developments described so far. In this context, it is emphasized that the displacement discontinuity function (10) and the cohesive zone size (Eq. (14)), which were the basis for all derivations presented up to this point, do not represent exact solutions, but best possible estimates within the chosen class of kinematically admissible displacement fields, see Sec. 2.2. The exact shape of the displacement discontinuity and the exact size of the cohesive zone are reported in the open literature for the special case of constant cohesive tractions ($\beta=0$), see Refs. [29,30]. Subsequently, these exact solutions will be used to assess the predictive capabilities of the proposed model, namely, by comparing model predictions specified for $\beta=0$ with exact solutions. This comparison will comprise the following three quantities:

1. the size of the cohesive zone
2. the relationship between remote tensile stresses and the opening of the cohesive crack at the inner edge of the cohesive zone
3. the energy dissipated during propagation of a cohesive crack

Sections 4.1 and 4.2 contain model predictions and related exact solutions, respectively. In Sec. 4.3, the comparison between model predictions and exact solutions is made. Conclusions are drawn in Sec. 5.

[3]To show that the energy released is equal to the area under the softening curve, one can also use the *J*-integral concept, which goes back to the pioneering work of Rice [28]. Employing this concept in the present context would involve integrals leading around the cohesive zone. Hence, the *J*-integral approach would not allow for direct identification of the source of dissipation. The aim of the submitted manuscript, however, is to explain how dissipation occurs. This provided the motivation to calculate the dissipation at the very domain where it occurs, namely, inside the cohesive zone.

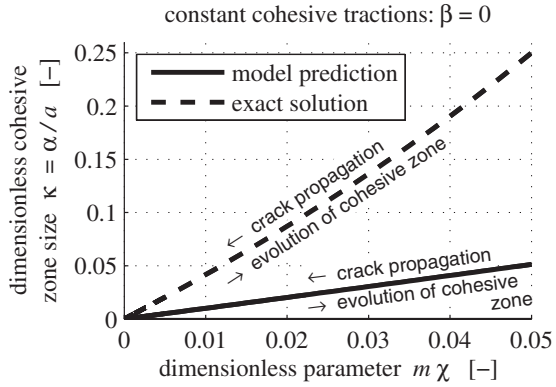


Fig. 5 Comparison of model prediction (Eq. (46)) with exact solution (55); size of the cohesive zone during its evolution and during crack propagation, see Table 1

4.1 Model Predictions for Constant Cohesive Tractions ($\beta=0$). The evolution of the cohesive zone size is considered in the following comprehensive manner. We consider an RVE comprising many parallel microcracks of radius a, initially without cohesive zones, and we deal with a monotonous increase in macroscopic strains referring to macroscopic tensile loading. The related behavior of the cohesive zones is characterized by two phases: (i) growth of the cohesive zones ahead of the stationary (nonpropagating) cracks (as detailed below), and (ii) the behavior of fully grown cohesive zones during crack propagation. Notably, up to this point we exclusively dealt with phase (ii). As for the description of phase (i), we parametrize the growth of α with the dimensionless parameter $m \in [0,1]$, defined as

$$[\![\xi_n]\!]_{r=a} = m[\![\xi_n^{\mathrm{cr}}]\!] \tag{44}$$

In Eq. (44), $[\![\xi_n]\!]_{r=a}$ denotes the displacement discontinuity at the inner edge of the cohesive zone (i.e., at $r=a$). In the considered context, $m=0$ corresponds to the "initial configuration," i.e., to a nonexistent cohesive zone, while $m=1$ corresponds to a "fully grown" cohesive zone, in the sense that the displacement discontinuity at the inner edge of the cohesive zone is equal to the critical value $[\![\xi_n^{\mathrm{cr}}]\!]$, see Eq. (44). The macroscopic loading parameter λ can be expressed as a function of parameter m, by specification of Eq. (10) for $r=a$ and by consideration of Eq. (44)

$$\lambda = \frac{m[\![\xi_n^{\mathrm{cr}}]\!]}{2\sqrt{(2a+\alpha)\alpha}} \tag{45}$$

The size of the cohesive zone with constant cohesive tractions, expressed as a function of m, is obtained from specifying Eq. (12) for $\beta=0$, for λ from Eq. (45), and from solving thereafter for α. The result reads in dimensionless form as

$$\frac{\alpha}{a} = (m\chi) - 1 + \sqrt{(m\chi)^2+1} \tag{46}$$

When dealing with growth of the cohesive zones ahead of stationary (nonpropagating) cracks, χ is constant, while α from Eq. (46) and λ from Eq. (45) can be interpreted as parameter forms, with m as the parameter increasing from 0 up to 1, see the solid line in Fig. 5 and Table 1. Reaching $m=1$ in the context of monotonous macroscopic load increase corresponds to onset of crack propagation, i.e., a change from phase (i) to phase (ii). Accordingly, the crack radius a will start to increase, while at $r=a$ (where a is at any time the current value of the crack radius) the displacement discontinuities of the cohesive zones remain equal to $[\![\xi_n^{\mathrm{cr}}]\!]$. Hence, when dealing with crack propagation, m is constant and equal to 1. In addition, χ from Eq. (15), α from Eq. (14), and λ from Eq. (13) apply, and they can be interpreted as parameter forms, with the increasing crack radius a as the parameter, see Table 1. Remarkably, since the crack radius appears in the denominator in the right-hand side of Eq. (15), χ decreases during crack propagation. Consequently, the dimensionless cohesive zone size, α/a, decreases during crack propagation, see the solid line in Fig. 5.

The relationship between the remote tensile stresses and the opening of the cohesive crack at the inner edge of the cohesive zone is studied for isotropic remote tension

$$\boldsymbol{\Sigma} = \Sigma_m^{\infty}\mathbb{1} \tag{47}$$

where $\mathbb{1}$ denotes the second-order unity tensor. The corresponding principal strain components read as

$$E_{33} = E_{\ell at} = \frac{\Sigma_m^{\infty}(1-2\nu_m)}{E_m} \tag{48}$$

Inserting Eq. (48) into Eq. (11) allows for expressing the macroscopic loading parameter λ as a function of Σ_m^{∞}

$$\lambda = \frac{4\Sigma_m^{\infty}(1-\nu_m^2)}{\pi E_m} \tag{49}$$

Insertion of Eqs. (49) and (46) into Eq. (45), solving for Σ_m^{∞}, and dividing the resulting expression by T_n^{cr} yield the model proposed relationship between $\Sigma_m^{\infty}/T_n^{\mathrm{cr}}$ as a function of the dimensionless parameter $m\chi$

$$\frac{\Sigma_m^{\infty}}{T_n^{\mathrm{cr}}} = \frac{2\sqrt{2(m\chi)}}{\sqrt{(m\chi)+\sqrt{(m\chi)^2+1}}} = 2\sqrt{2(m\chi)} - \sqrt{2(m\chi)^3} + \mathcal{O}(\sqrt{(m\chi)^5}) \tag{50}$$

Σ_m^{∞} increases in phase (i), while it decreases in phase (ii), see the solid line in Fig. 6 and Table 1.

The energy release rate (Eq. (36)) is divided by $T_n^{\mathrm{cr}}[\![\xi_n^{\mathrm{cr}}]\!]$ to obtain a dimensionless expression

Table 1 Two phases of cohesive zone behavior encountered during monotonous load increase in a RVE containing microcracks with initially nonexistent cohesive zones ($\beta=0$)

	Phase (i): growth of cohesive zone	Phase (ii): crack propagation
$m = \dfrac{[\![\xi_n]\!]_{r=a}}{[\![\xi_n^{\mathrm{cr}}]\!]}$	Increasing from 0 to 1	Constant and equal to 1
Crack radius: a	Constant and equal to initial value a_{ini}	Increasing, starting from the initial value a_{ini}
$\chi = \dfrac{\pi E_m [\![\xi_n^{\mathrm{cr}}]\!]}{32 T_n^{\mathrm{cr}} a(1-\nu_m^2)}$	Constant, with $a=a_{\mathrm{ini}}$	Decreasing with increasing a

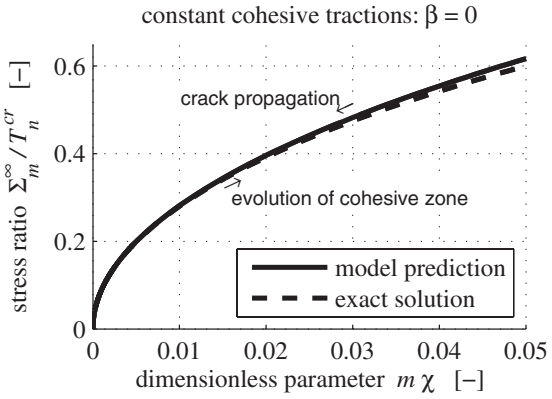


Fig. 6 Comparison of model prediction (Eq. (50)) with exact solution (56): relationship between remote stresses and crack opening at the inner edge of the cohesive zone (consider: $m = [\![\xi_n]\!]_{r=a}/[\![\xi_n^{cr}]\!]$), see Table 1

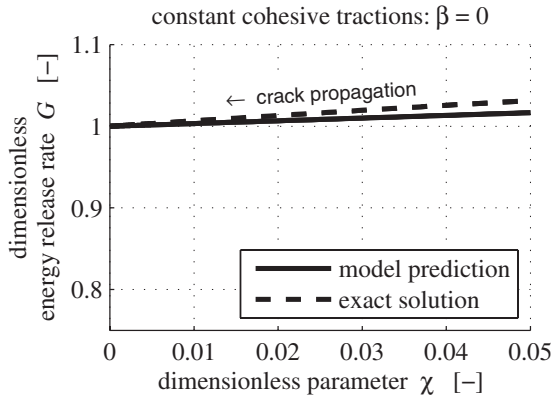


Fig. 7 Comparison of model prediction (Eq. (51)) with exact solution (66): dimensionless energy release rate during crack propagation, see Table 1; to render the curves distinguishable, the ordinate shows the interval [0.75,1.10] only

$$G = \frac{\mathcal{G}_{\text{crack}}}{T_n^{\text{cr}}[\![\xi_n^{\text{cr}}]\!]} = 1 + \frac{\chi}{3\sqrt{\chi^2+1}} = 1 + \frac{1}{3}\chi + \mathcal{O}(\chi^3) \tag{51}$$

According to Eq. (51), the single crack-related energy release rate slightly decreases during crack propagation, see the solid line in Fig. 7.

4.2 Exact Solutions for Constant Cohesive Tractions. Chen and Keer [29] considered an elasticity problem comprising an infinite linear elastic matrix containing one circular crack of radius $a+\alpha$. The solid is subjected, at infinity, to the remote stresses (Eq. (47)). In the annular region of the crack between $r=a$ and $r=a+\alpha$, normal tractions T_n^{cr} are prescribed both on the lower and on the upper crack surface. To ensure the stresses at the crack tip to be finite, crack radius a, annular zone width α, remote mean stress Σ_m^∞, and prescribed normal tractions T_n^{cr} are related through [29]

$$\frac{a}{a+\alpha} = \sqrt{1 - \left(\frac{\Sigma_m^\infty}{T_n^{\text{cr}}}\right)^2} \tag{52}$$

The separation of the crack faces in the annular zone reads as [29]

$$[\![\xi_n]\!] = 2\lambda\sqrt{(a+\alpha)^2 - r^2} - \frac{8(1-\nu_m^2)}{\pi}\frac{T_n^{\text{cr}}}{E_m}\int_{s=r}^{a+\alpha}\sqrt{\frac{s^2-a^2}{s^2-r^2}}ds,$$
$$a \le r \le a+\alpha \tag{53}$$

with λ from Eq. (49). The solution for the displacement discontinuity at the inner edge of the annular zone follows from specification of Eq. (53) for $r=a$ as

$$[\![\xi_n]\!]_{r=a} = \frac{8(1-\nu_m^2)}{\pi E_m} a T_n^{\text{cr}}\left[1 - \sqrt{1 - \left(\frac{\Sigma_m^\infty}{T_n^{\text{cr}}}\right)^2}\right] \tag{54}$$

An expression for the annular zone width, which is directly comparable with Eq. (46), is obtained from replacing the square root expression in Eq. (54) by the left-hand side of Eq. (52), from replacing $[\![\xi_n]\!]_{r=a}$ in Eq. (54) by the right-hand side of Eq. (44), and from solving the resulting expression for α/a as

$$\frac{\alpha}{a} = \frac{4(m\chi)}{1-4(m\chi)} \tag{55}$$

see the dashed line in Fig. 5. An expression, which is directly comparable with Eq. (50), is obtained from solving Eq. (54) for $\Sigma_m^\infty / T_n^{\text{cr}}$

$$\frac{\Sigma_m^\infty}{T_n^{\text{cr}}} = 2\sqrt{2(m\chi)(1-2(m\chi))} = 2\sqrt{2(m\chi)} - 2\sqrt{2(m\chi)^3} + \mathcal{O}(\sqrt{(m\chi)^5}) \tag{56}$$

see the dashed line in Fig. 6.

An expression for the energy dissipated during stable propagation of a cohesive crack with elastodamaging cohesive behavior is not available in the literature. Notably Chen and Keer [29] considered an elasticity problem in which the cohesive tractions were *prescribed* as part of the boundary conditions of the matrix, whereby nothing was said about the source for these tractions. Herein, we interpret Chen and Keer's cohesive tractions as the result of an elastodamaging cohesive behavior defined in Eqs. (1) and (2), with $\beta=0$. To calculate the related rate of dissipation associated with crack propagation, we subsequently specify the exact displacement discontinuity function (53) for crack propagation, we calculate its rate, and we insert both expressions into Eq. (30).

The shape of the displacement discontinuity in the cohesive zone during propagation of a crack with constant cohesive tractions follows from specification of Eq. (53) for $r=a$, from setting the resulting expression equal to $[\![\xi_n^{\text{cr}}]\!]$, from solving the obtained equation for the macroscopic loading parameter λ, and from reinserting λ into Eq. (53) as

$$[\![\xi_n]\!] = [\![\xi_n^{\text{cr}}]\!]\left[\left(1 + \frac{8(1-\nu_m^2)T_n^{\text{cr}}\alpha}{\pi E_m [\![\xi_n^{\text{cr}}]\!]}\right)\sqrt{\frac{(a+\alpha)^2 - r^2}{\alpha(2a+\alpha)}} - \frac{8(1-\nu_m^2)T_n^{\text{cr}}}{\pi E_m [\![\xi_n^{\text{cr}}]\!]}\int_{s=r}^{a+\alpha}\sqrt{\frac{s^2-a^2}{s^2-r^2}}ds\right] \tag{57}$$

The rate of displacement discontinuity is obtained by inserting Eq. (57) into Eq. (24). In the present context of the exact solution, the derivative of α with respect to a is obtained under consideration of Eq. (15) from Eq. (55) as

$$\frac{d\alpha}{da} = -\frac{\alpha^2}{a^2} \tag{58}$$

In order to calculate the rate of dissipation, $[\![\xi_n]\!]$ and $[\![\dot{\xi}_n]\!]$ are expressed as defined in Eq. (27). In the present context of the exact solution, the dimensionless functions $\Xi(\rho)$ and $\Lambda(\rho)$ read as

$$\Xi(\rho) = \left(1 + \frac{\kappa}{4\chi}\right)\sqrt{\frac{(1+\kappa)^2 - \rho^2}{\kappa(2+\kappa)}} - \frac{1}{4\chi}\int_{\tau=\rho}^{1+\kappa}\sqrt{\frac{\tau^2-1}{\tau^2-\rho^2}}d\tau \tag{59}$$

and

$$\Lambda(\rho)=\frac{(1+\kappa)^2+(1-\kappa-\kappa^2)\rho^2-\kappa(1+\kappa)(3+\kappa-\rho^2)(4\chi)^{-1}}{(2+\kappa)\sqrt{(1+\kappa)^2-1}\sqrt{(1+\kappa)^2-\rho^2}}+\frac{1}{4\chi}\int_{\tau=\rho}^{1+\kappa}\frac{1}{\sqrt{\tau^2-1}\sqrt{\tau^2-\rho^2}}d\tau \tag{60}$$

Now, we specify the rate of dissipation (Eq. (30)) for $[\![\xi_n]\!]$ and $[\![\dot{\xi}_n]\!]$ from Eqs. (27), (59), and (60). Integration of the first term on the right-hand side of Eq. (30), denoted as $\dot{\mathcal{D}}_1$, is done under consideration of the elliptic integral

$$\int_{\rho=1}^{1+\kappa}\left(\int_{\tau=\rho}^{1+\kappa}\frac{1}{\sqrt{\tau^2-1}\sqrt{\tau^2-\rho^2}}d\tau\right)\rho d\rho=\kappa \tag{61}$$

and yields

$$\dot{\mathcal{D}}_1=T_n^{\mathrm{cr}}[\![\xi_n^{\mathrm{cr}}]\!]a\pi\dot{a}\int_{\rho=1}^{1+\kappa}\Lambda(\rho)\rho d\rho=T_n^{\mathrm{cr}}[\![\xi_n^{\mathrm{cr}}]\!]a\pi\dot{a}\left(1+\frac{4\chi(1-8\chi)}{3(1+4\chi)^2}\right) \tag{62}$$

Again, evaluation of the second term on the right-hand side of Eq. (22), denoted as $\dot{\mathcal{D}}_2$, deserves special attention

$$\dot{\mathcal{D}}_2=T_n^{\mathrm{cr}}[\![\xi_n^{\mathrm{cr}}]\!]a\pi\dot{a}\mathcal{L}\quad\text{with}\quad\mathcal{L}=\lim_{\beta\to 0}\left[\int_{\rho=1}^{1+\kappa}\beta(1-\Xi(\rho))^{\beta-1}\Xi(\rho)\Lambda(\rho)\rho d\rho\right] \tag{63}$$

Though β tends to zero, $\mathcal{L}$ is not equal to zero since again $(1-\Xi(\rho))^{\beta-1}\Xi(\rho)\Lambda(\rho)\rho$ tends to infinity at $\rho=1$ (i.e., at $r=a$). Remarkably, the exact expressions for $\Xi(\rho)$ and $\Lambda(\rho)$ from Eqs. (59) and (60) differ from the optimal kinematically admissible expressions (Eq. (28)). Nonetheless, all qualitative and quantitative properties referred to in the Appendix are the same. In other words, although the exact displacement discontinuity and its rate differ from the related functions in Sec. 3, all descriptions in the Appendix are valid in both cases, and hence, the same result is obtained, namely

$$\mathcal{L}=1 \tag{64}$$

The exact rate of dissipated energy during propagation of an elastodamaging cohesive crack with constant cohesive tractions follows from combining Eqs. (62)–(64) as

$$\dot{\mathcal{D}}_{\mathrm{crack}}=T_n^{\mathrm{cr}}[\![\xi_n^{\mathrm{cr}}]\!]2a\pi\dot{a}\left(1+\frac{2\chi(1-8\chi)}{3(1+4\chi)^2}\right) \tag{65}$$

Dividing Eq. (65) by the rate of newly produced crack area (Eq. (35)) and by $T_n^{\mathrm{cr}}[\![\xi_n^{\mathrm{cr}}]\!]$ yields a dimensionless expression for the energy release rate, which is directly comparable with Eq. (51), as

$$G=\frac{\dot{\mathcal{D}}_{\mathrm{crack}}}{\dot{A}T_n^{\mathrm{cr}}[\![\xi_n^{\mathrm{cr}}]\!]}=1+\frac{2\chi(1-8\chi)}{3(1+4\chi)^2}=1+\frac{2}{3}\chi+\mathcal{O}(\chi^3) \tag{66}$$

4.3 Comparison of Model Predictions With Exact Solutions. Model predicted cohesive zone sizes significantly underestimate the exact solutions in case of constant cohesive tractions. The related estimates (Eqs. (14) and (46)) cannot be recommended for values of β smaller than 3, as was already concluded in Ref. [10]. If better estimates of the cohesive zone size are required, a more sophisticated class of kinematically admissible displacement fields and related estimates can be found in Ref. [10].

Model predictions regarding the relationship between remote tensile stresses and the opening of the cohesive crack at the inner edge of the cohesive zone are satisfactory. The model prediction (Eq. (50)) reproduces the exact solution (Eq. (56)) quite well, since the leading terms of the power series expansions in Eqs. (50) and (56) are identical. In more detail, model predictions differ from exact solutions by higher-order terms only, e.g., by less than 1.81% relative error if

$$\chi<0.034 \tag{67}$$

i.e., if

$$\frac{\max(\Sigma_m^{\infty})}{T_n^{\mathrm{cr}}}<0.5 \tag{68}$$

In Eq. (68), $\max(\Sigma_m^{\infty})$ denotes the remote tensile stress at onset of microcrack propagation. Since the tensile microstrength T_n^{cr} of brittle materials, such as concrete or rock, is expected to be significantly larger (e.g., more than twice as large) than bearable macroscopic tensile stresses (see, e.g., Ref. [31]), (Eq. (68)), and hence, Eq. (67) are also regarded as being fulfilled.

Model predictions regarding the energy release rate during propagation of cracks with constant cohesive tractions are also satisfactory. The model prediction (Eq. (51)) reproduces the exact solution (Eq. (66)) quite well, since again the leading terms of the power series expansions in Eqs. (51) and (66) are identical. Model predictions differs from the exact solutions by higher-order terms only, e.g., by less than 0.49% relative error if Eq. (67) is satisfied, i.e., if the tensile microstrength is at least twice as large as the remote tensile stresses at onset of crack propagation.

5 Discussion, Conclusions, and Outlook

This paper dealt with the dissipation associated with quasistatic microcracking of brittle materials exhibiting softening behavior. For this purpose the elastodamaging cohesive zone model of Pichler and Dormieux [10] was used. From the satisfactory model performance illustrated in Fig. 6, it can be concluded that the proposed approach allows for reliable prediction of the macroscopic load intensity, at which the crack opening at the inner edge of the cohesive zone becomes equal to the critical separation $[\![\xi_n^{\mathrm{cr}}]\!]$. Since this is the criterion for onset of crack propagation, Fig. 6 highlights that the proposed approach allows for reliable prediction of onset of microcrack propagation. The again satisfactory model performance illustrated in Fig. 7, in turn, highlights that the proposed approach allows for reliable prediction of the dissipation during microcrack propagation.

Remarkably, the model predictions were derived based on a class of kinematically admissible displacement fields corresponding to microcracks without cohesive stresses (characterized by $\beta\to\infty$). This means that in the limit $\beta\to\infty$, all model predictions converge toward exact solutions. Hence, the reliability of model predictions is expected to increase with increasing power law exponent β. In this context, the accuracy of model predictions illustrated in Figs. 6 and 7 is remarkable, since these illustrations refer to *constant* cohesive tractions ($\beta=0$). Hence, one may conclude that model predictions for values of $\beta>0$ are even more reliable than the already satisfactory results of Figs. 6 and 7. Accordingly, both the energy release rate (Eq. (36)) and the improved stiffness estimate (Eq. (43)) can be viewed as being corroborated for any power law exponent from zero to infinity.

For materials whose microscopic tensile strength is at least more than twice as large as the macroscopic tensile strength, it was shown that the dimensionless parameter χ, defined in Eq. (15), is very small compared with 1. For $\chi\ll 1$, the second term in the parenthesis of Eq. (43) is very small compared with 1. This way, the improved stiffness estimate (Eq. (43)) degenerates to the classical one (Eq. (7)), which indicates that the cohesive zones do not significantly contribute to the macroscopic stiffness. This result justifies the use of stiffness estimates based on classical noncohesive cracks, see Refs. [1–3] and the references therein.

Results obtained for the energy release rate during propagation of a cohesive crack are of interest both for practical engineers, as

well as for the scientific community. It was shown that in a wide class of practical applications, the energy release rate during crack propagation is—up to terms of higher order—indeed equal to the area under the T_n-$[\![\xi_n]\!]$ diagram, commonly referred to as the softening curve. This is remarkable, since throughout this paper, we neither restricted our considerations to cases where the cohesive zone is small, as compared with the crack radius, nor did we consider that the cohesive zone size is constant during crack propagation, which are both assumptions frequently used in the pertinent literature. This way, the obtained results justify approaches relying on the classical concept of constant (dissipative) energy release rate. This concept is frequently implemented into finite element codes used in the engineering practice, introducing the toughness of the material as the area under the softening curve, independent of the size of the cohesive zone. With regard to the ongoing scientific discussion on how to describe microcracking in the framework of homogenization techniques, the presented results provide a strong argument for combined fracture-micromechanics models, such as those proposed in Refs. [4,5], respectively. This opens the door for further enhancing the micromechanical description of microcracking processes in media such as cementitious materials and rock.

Notably, the proposed model is based on three material constants: the tensile microstrength T_n^{cr}, the displacement discontinuity at which no forces can be transmitted across the cohesive zone anymore $[\![\xi_n^{\mathrm{cr}}]\!]$, and the power law exponent β, see Eq. (2). Their identification for a specific brittle softening material remains a challenging task. Up to the knowledge of the authors, these are nowadays, unfortunately, inaccessible by state-of-the-art material testing. Hence, the material constants introduced herein need to be back analyzed from macroscopic material experiments. In this context, transmission electron microscopy (TEM) using an atomic force microscope (AFM) is a promising experimental tool, which might soon be able to provide access to the material constants. Alternatively, recent results of molecular dynamics simulations, such as, e.g., Ref. [18], raise the hope that the aforementioned material constants will soon be accessible for a wide class of brittle materials, including cementitious materials and rock.

While nanomechanical simulations will hopefully provide soon the material constants of the cohesive microcracks, this paper contains first steps toward a simple and computationally cheap realization of the successive upscaling step, i.e., the link from the microscale (the level of one cohesive microcrack) to the macroscale (the level of a representative volume element containing many cohesive microcracks). In case of parallel noninteracting cohesive cracks, this second scale transition is provided by the derived improved dilute estimate. From a scientific viewpoint, it would be desirable to extend both the estimation of the cohesive zone size and the estimation of the macroscopic stiffness toward consideration of interacting cohesive microcracks. This, however, was not the topic of this paper, but it provides motivation for future research.

Acknowledgment

The authors are grateful for fruitful discussions with Christian Hellmich, Vienna University of Technology, Austria, and Patrick de Buhan, Ecole Nationale des Ponts et Chaussées, Marne-la-Vallée, France.

Nomenclature

List of Symbols and Abbreviations

a = crack radius
$\dot{a}$ = rate of a
$\dot{A}$ = rate of crack area
c = half opening at the center of a microcrack (Fig. 4)
$\mathbb{C}_m$ = elastic isotropic stiffness tensor of the matrix (fourth-order tensor)
$\mathbb{C}_{\mathrm{hom}}$ = homogenized elastic stiffness of the RVE (fourth-order tensor)
$\dot{\mathbb{C}}_{\mathrm{hom}}$ = rate of $\mathbb{C}_{\mathrm{hom}}$
$\mathbb{C}_{\mathrm{hom}}^{\mathrm{dil}}$ = dilute estimate of $\mathbb{C}_{\mathrm{hom}}$
$\mathbb{C}_{\mathrm{hom}}^{\mathrm{idil}}$ = improved dilute estimate of $\mathbb{C}_{\mathrm{hom}}$
$\dot{\mathcal{D}}_{\mathrm{crack}}$ = rate of energy dissipated due to propagation of one microcrack
$\dot{\mathcal{D}}_1$ = first term on the right-hand side of Eq. (30)
$\dot{\mathcal{D}}_2$ = second term on the right-hand side of Eq. (30)
$\dot{\mathcal{D}}_{\mathrm{RVE}}$ = rate of dissipation of an RVE with many propagating microcracks
$\mathbf{e}_i$ = base vector of Cartesian coordinate frame
E_m = Young's modulus of the isotropic matrix
$\mathbf{E}$ = macroscopic strains of the RVE (second-order tensor)
E_{33} = normal strain component of $\mathbf{E}$, see Eq. (6)
$E_{\ell at}$ = normal strain component of $\mathbf{E}$, see Eq. (6)
F = dimensionless function, see Eq. (A4)
$\mathcal{G}_{\mathrm{crack}}$ = energy release rate during propagation of one crack
G = dimensionless expression of $\mathcal{G}_{\mathrm{crack}}$, see Eqs. (51) and (66), respectively
$\mathbb{1}$ = second-order unity tensor
$\mathbb{I}$ = symmetric fourth-order unity tensor
i, j, k, and l = indices: $i,j,k,l \in \{1,2,3\}$
$K([\![\xi_n]\!])$ = elastodamaging stiffness of cohesive zone; $[\![\xi_n]\!]$ =damage variable
$\dot{K}([\![\xi_n]\!])$ = rate of $K([\![\xi_n]\!])$
$\mathcal{L}$ = dimensionless function, see Eqs. (32) and (63)
m = dimensionless crack opening at the inner edge of the cohesive zone, see Eq. (44)
N = number of cracks per unit volume of the RVE
$\mathbf{n}$ = unit normal vector
r = radial coordinate of local (crack related) base frame (Fig. 4)
$\mathbb{S}_c^m$ = Eshelby tensor (fourth-order tensor)
s = integration parameter
t = time
$\mathbb{T}$ = fourth-order tensor defined in Eq. (8)
T_{ijkl} = components of $\mathbb{T}$
T_n = cohesive tensile normal tractions
T_n^{cr} = tensile microstrength, maximum value of T_n
$\mathbf{x}$ = position vector
$\dot{W}_{\mathrm{crack}}^{\mathrm{ex}}$ = rate of work expended on a crack by its environment
$W_{\mathrm{crack}}^{\mathrm{in}}$ = elastic internal energy stored in a cohesive crack
$\dot{W}_{\mathrm{crack}}^{\mathrm{in}}$ = rate of $W_{\mathrm{crack}}^{\mathrm{in}}$
$\dot{W}_{\mathrm{RVE}}^{\mathrm{ex}}$ = rate of work expended on the RVE by its environment
$\dot{W}_{\mathrm{RVE}}^{\mathrm{in}}$ = rate of elastic internal energy stored in the RVE
α = cohesive zone size (=width of annular cohesive zone)
β = dimensionless material constant (power law exponent related to the shape of the T_n-$[\![\xi_n]\!]$ diagram, see Eq. (2))
κ = dimensionless cohesive zone size, see Eq. (29)
Λ = dimensionless function related to $[\![\dot{\xi}_n]\!]$, see Eq. (27)
λ = macroscopic loading parameter defined in Eq. (11)

ν_m = Poisson's ratio of the isotropic matrix
Ξ = dimensionless function related to $[\![\xi_n]\!]$, see Eq. (27)
$\boldsymbol{\xi}$ = displacement vector prescribed at the boundary of the RVE
$[\![\xi_n]\!]$ = normal displacement discontinuity
$\max_t[\![\xi_n]\!]$ = maximum normal displacement discontinuity experienced by a material point belonging to a cohesive zone=damage variable in Eq. (4)
$[\![\dot{\xi}_n]\!]$ = rate of $[\![\xi_n]\!]$
$[\![\xi_n]\!]_{r=a}$ = $[\![\xi_n]\!]$ at $r=a$ (i.e., at the inner edge of the cohesive zone)
$[\![\xi_n^{\mathrm{cr}}]\!]$ = value of $[\![\xi_n]\!]$ at which all bonds between initially adjacent material points are broken such that no tensile forces can be transmitted across the cohesive zone anymore
ρ = dimensionless radial coordinate, see Eq. (29)
$\boldsymbol{\Sigma}$ = macroscopic stresses of the RVE (second-order tensor)
Σ_m^{∞} = normal stress components of $\boldsymbol{\Sigma}$
τ = dimensionless integration parameter
Φ = dimensionless function, see Eq. (70)
χ = dimensionless parameter composed of material constants, see Eq. (15)
Ψ = dimensionless function, see Eq. (70)
ψ = elastic free energy per unit area of the cohesive zone
$\dot{\psi}$ = rate of ψ
ω = crack aspect ratio, see Fig. 4
Ω = volume of the RVE
$\partial\Omega$ = surface of the RVE

Appendix: Evaluation of the Function $\mathcal{L}$ Based on Lebesgue's Dominated Convergence Theorem

This appendix deals with evaluation of the dimensionless function

$$\mathcal{L}=\lim_{\beta\to 0}\left[\int_{\rho=1}^{1+\kappa}\beta(1-\Xi(\rho))^{\beta-1}\Xi(\rho)\Lambda(\rho)\rho d\rho\right]\neq 0 \quad \text{(A1)}$$

The structure of the integrand in Eq. (A1) provides the motivation to introduce the following two dimensionless functions:

$$\Phi(\rho)=1-\Xi(\rho) \quad \text{and} \quad \Psi(\rho)=\Xi(\rho)\Lambda(\rho)\rho \quad \text{(A2)}$$

such that $\mathcal{L}$ reads as

$$\mathcal{L}=\lim_{\beta\to 0}\int_{\rho=1}^{1+\kappa}\beta\Phi^{\beta-1}\Psi d\rho \quad \text{(A3)}$$

The integrand in Eq. (A3) is further modified such that partial integration can be simply applied. For this purpose, it is useful to introduce the function

$$F(\rho,\beta)=\Phi^{\beta} \quad \text{(A4)}$$

Deriving F partially with respect to ρ (partial derivatives with respect to ρ are indicated subsequently by a prime), and rearranging terms yields $\beta\Phi^{\beta-1}=F'/\Phi'$. Inserting this relationship into Eq. (A3) and integrating by parts yield

$$\mathcal{L}=\lim_{\beta\to 0}\int_{\rho=1}^{1+\kappa}F'\frac{\Psi}{\Phi'}d\rho=\lim_{\beta\to 0}\left[\Phi^{\beta}\left.\frac{\Psi}{\Phi'}\right|_1^{1+\kappa}+\int_{\rho=1}^{1+\kappa}\Phi^{\beta}\left(-\frac{\Psi}{\Phi'}\right)'d\rho\right] \quad \text{(A5)}$$

Accounting in Eq. (A5) for

$$\Phi(1)=0 \quad \text{and} \quad \Phi(1+\kappa)=1 \quad \text{(A6)}$$

yields

$$\mathcal{L}=\lim_{\beta\to 0}1^{\beta}\left.\frac{\Psi}{\Phi'}\right|_1^{1+\kappa}+\lim_{\beta\to 0}\int_{\rho=1}^{1+\kappa}\Phi^{\beta}\left(-\frac{\Psi}{\Phi'}\right)'d\rho \quad \text{(A7)}$$

For further simplification of Eq. (A7), it would be desirable to exchange the sequence of integration and taking the limit $\beta\to 0$. For this purpose, characteristics of the integrand are analyzed in more detail. Notably

$$0\leq\Phi^{\beta}\leq 1 \quad \text{for all} \quad \beta\geq 0 \quad \text{and} \quad \text{for} \quad 1\leq\rho\leq 1+\kappa \quad \text{(A8)}$$

For the functions of interest herein, it can be shown numerically that

$$0<\left(-\frac{\Psi}{\Phi'}\right)' \quad \text{for} \quad 1\leq\rho\leq 1+\kappa \quad \text{(A9)}$$

From Eqs. (A8) and (A9) follows that

$$0\leq\Phi^{\beta}\left(-\frac{\Psi}{\Phi'}\right)'\leq\left(-\frac{\Psi}{\Phi'}\right)' \quad \text{for all} \quad \beta\geq 0 \quad \text{and} \quad \text{for} \quad 1\leq\rho\leq 1+\kappa \quad \text{(A10)}$$

In other words, the integrand in Eq. (A7), $\Phi^{\beta}(-\Psi/\Phi')'$, is dominated by $(-\Psi/\Phi')'$, in the entire cohesive zone (i.e., in the entire interval of integration) and for any β greater or equal than zero. Since the latter term is integrable, i.e., since

$$\int_{\rho=1}^{1+\kappa}\left(-\frac{\Psi}{\Phi'}\right)'d\rho=\left.-\frac{\Psi}{\Phi'}\right|_1^{1+\kappa}=0-(-1)=1<\infty \quad \text{(A11)}$$

Lebesgue's dominated convergence theorem applies to the second term on the right-hand side of Eq. (A7), permitting us to exchange the sequence of integration and taking the limit $\beta\to 0$, see, e.g., Ref. [27]

$$\mathcal{L}=\left.\frac{\Psi}{\Phi'}\right|_1^{1+\kappa}+\int_{\rho=1}^{1+\kappa}\lim_{\beta\to 0}\Phi^{\beta}\left(-\frac{\Psi}{\Phi'}\right)'d\rho \quad \text{(A12)}$$

Considering Eq. (A8) it follows that $\lim_{\beta\to 0}\Phi^{\beta}=1$. Inserting this result into Eq. (A12) finally yields

$$\mathcal{L}=\left.\frac{\Psi}{\Phi'}\right|_1^{1+\kappa}+\int_{\rho=1}^{1+\kappa}\left(-\frac{\Psi}{\Phi'}\right)'d\rho=\left.\frac{\Psi}{\Phi'}\right|_1^{1+\kappa}-\left.\frac{\Psi}{\Phi'}\right|_1^{1+\kappa}=-\left.\frac{\Psi}{\Phi'}\right|_1=\lim_{\rho\to 0}\frac{\Psi}{\Phi'}=1 \quad \text{(A13)}$$

References

[1] Mura, T., 1987, *Micromechanics of Defects in Solids*, Martinus Nijhoff, Dordrecht.
[2] Nemat-Nasser, S., and Hori, M., 1999, *Micromechanics: Overall Properties of Heterogeneous Materials*, 2nd ed., North-Holland, Amsterdam.
[3] Dormieux, L., Kondo, D., and Ulm, F.-J., 2006, *Microporomechanics*, Wiley, New York.
[4] Dormieux, L., Kondo, D., and Ulm, F.-J., 2006, "A Micromechanical Analysis of Damage Propagation in Fluid-Saturated Cracked Media," C. R. Mec., **334**(7), pp. 440–446.
[5] Pichler, B., Hellmich, Ch., and Mang, H. A., 2007, "A Combined Fracture-Micromechanics Model for Tensile Strain-Softening in Brittle Materials, Based on Propagation of Interacting Microcracks," Int. J. Numer. Analyt. Meth. Geomech., **31**(2), pp. 111–132.
[6] van Vliet, M. R. A., and van Mier, J. G. M., 2000, "Experimental Investigation of Size Effect in Concrete and Sandstone Under Uniaxial Tension," Eng. Fract. Mech., **65**(2–3), pp. 165–188.
[7] Dugdale, D. S., 1960, "Yielding of Steel Sheets Containing Slits," J. Mech. Phys. Solids, **8**(2), pp. 100–104.
[8] Barenblatt, G. I., 1959, "The Formation of Equilibrium Cracks During Brittle Fracture. General Ideas and Hypotheses. Axially-Symmetric Cracks," J. Appl. Math. Mech., **23**(3), pp. 622–636.

[9] Barenblatt, G. I., 1962, "The Mathematical Theory of Equilibrium Cracks in Brittle Fracture," *Advances in Applied Mechanics*, Vol. 7, Academic, London, pp. 55–129.
[10] Pichler, B., and Dormieux, L., 2007, "Cohesive Zone Size of Microcracks in Brittle Materials," Eur. J. Mech. A/Solids, **26**(6), pp. 956–968.
[11] Glennie, E. B., 1971, "A Strain-Rate Dependent Crack Model," J. Mech. Phys. Solids, **19**(5), pp. 255–272.
[12] Siegmund, T., and Needleman, A., 1997, "A Numerical Study of Dynamic Crack Growth in Elastic-Viscoplastic Solids," Int. J. Solids Struct., **34**(7), pp. 769–787.
[13] Tvergaard, V., and Hutchinson, J. W., 1992, "The Relation Between Crack Growth Resistance and Fracture Process Parameters in Elastic-Plastic Solids," J. Mech. Phys. Solids, **40**(6), pp. 1377–1397.
[14] Hillerborg, A., Modeér, M., and Petersson, P.-E., 1976, "Analysis of Crack Formation and Crack Growth in Concrete by Means of Fracture Mechanics and Finite Elements," Cem. Concr. Res., **6**(6), pp. 773–782.
[15] Elices, M., Guinea, G. V., Gómez, J., and Planas, J., 2002, "The Cohesive Zone Model: Advantages, Limitations and Challenges," Eng. Fract. Mech., **69**(2), pp. 137–163.
[16] Elices, M., and Planas, J., 1996, "Fracture Mechanics Parameters of Concrete: An Overview," Adv. Cem. Based Mater., **4**(3–4), pp. 116–127.
[17] Camacho, G. T., and Ortiz, M., 1996, "Computational Modelling of Impact Damage in Brittle Materials," Int. J. Solids Struct., **33**(20–22), pp. 2899–2938.
[18] Yamakov, V., Saether, E., Phillips, D. R., and Glaessgen, E. H., 2006, "Molecular-Dynamics Simulation-Based Cohesive Zone Representation of Intergranular Fracture Processes in Aluminum," J. Mech. Phys. Solids, **54**(9), pp. 1899–1928.
[19] Bjerke, T. W., and Lambros, J., 2003, "Theoretical Development and Experimental Validation of a Thermally Dissipative Cohesive Zone Model for Dynamic Fracture of Amorphous Polymers," J. Mech. Phys. Solids, **51**(6), pp. 1147–1170.
[20] Hashin, Z., 1983, "Analysis of Composite Materials: A Survey," ASME J. Appl. Mech., **50**(3), pp. 481–505.
[21] van Mier, J. G. M., and van Vliet, M. R. A., 2002, "Uniaxial Tension Tests for the Determination of Fracture Parameters of Concrete: State of the Art," Eng. Fract. Mech., **69**(2), pp. 235–247.
[22] Salençon, J., 2001, *Handbook of Continuum Mechanics*, Springer-Verlag, Berlin.
[23] Dormieux, L., and Kondo, D., 2004, "Approche micromécanique du couplage perméabilité–endommagement," C. R. Mec., **332**(2), pp. 135–140.
[24] Dormieux, L., and Kondo, D., 2005, "Poroelasticity and Damage Theory for Saturated Cracked Media," *Applied Micromechanics of Porous Media* (CISM Course and Lecture Notes), L. Dormieux and F.-J. Ulm, eds., Springer, Wien, Vol. 480, pp. 153–186.
[25] Sneddon, I. N., and Lowengrub, M., 1970, *Crack Problems in the Classical Theory of Elasticity*, Wiley, New York.
[26] Gurtin, M. E., 1979, "Thermodynamics and the Cohesive Zone in Fracture," Z. Angew. Math. Phys., **30**(6), pp. 991–1003.
[27] Bartle, R. G., 1995, *The Elements of Integration and Lebesgue Measure*, Wiley, New York.
[28] Rice, J. R., 1968, "A Path-Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks," ASME J. Appl. Mech., **35**, pp. 379–386.
[29] Chen, W. R., and Keer, L. M., 1993, "Mixed-Mode Fatigue Crack Propagation of Penny-Shaped Cracks," ASME J. Eng. Mater. Technol., **115**(4), pp. 365–372.
[30] Li, S., and Wang, G., 2004, "On Damage Theory of a Cohesive Medium," Int. J. Eng. Sci., **42**(8–9), pp. 861–885.
[31] Leblond, J.-B., 2003, *Mécanique de la rupture fragile et ductile*, Hermes, Paris, in French.

Dynamic Stability Analysis of Stiffened Shell Panels With Cutouts

S. N. Patel[1]
Research Student
e-mail: shuvendu_patel@yahoo.co.in

P. K. Datta
Professor
e-mail: pkdatta@aero.iitkgp.ernet.in

Department of Aerospace Engineering,
Indian Institute of Technology Kharagpur,
Kharagpur,
721 302 West Bengal, India

A. H. Sheikh
Associate Professor
School of Civil, Environmental and Mining Engineering,
University of Adelaide,
Adelaide,
South Australia 5005, Australia
e-mail: ahsheikh@civeng.adelaide.edu.au

A finite element dynamic instability analysis of stiffened shell panels with cutout subjected to uniform in-plane harmonic edge loading along the two opposite edges is presented in this paper. The eight-noded isoparametric degenerated shell element and a compatible three-noded curved beam element are used to model the shell panels and the stiffeners, respectively. As the usual formulation of degenerated beam element is found to overestimate the torsional rigidity, an attempt has been made to reformulate it in an efficient manner. Moreover the new formulation for the beam element requires five degrees of freedom per node as that of shell element. Bolotin method is applied to analyze the dynamic instability regions. Numerical results of convergence studies are presented and comparison is made with the published results from literature. The effects of various parameters such as shell geometry, radius of curvature, cutout size, stiffening scheme, and dynamic load factors are considered in dynamic instability analysis of stiffened shell panels with cutout. The free vibration and static stability (buckling) results are also presented. With the consideration of radius of curvatures the panels reduce from deep shell case to shallow shell case and finally become flat plate. [DOI: 10.1115/1.3086595]

Keywords: finite element method, perforated stiffened shell panels, degenerated curved beam element, degenerated shell element, deep shell, shallow shell, cutouts, dynamic instability and in-plane exciting load

1 Introduction

The use of shell panels is common in many activities of aerospace, mechanical, civil, and marine engineering structures. The panels are often attached with a suitable stiffened structural form to enhance the specific strength/stiffness to weight ratio of the structure. The geometrical discontinuities such as cutouts are inevitable in these structures to facilitate different purposes. The stress distribution in these structures becomes nonuniform due to the presence of stiffeners and cutouts. During the service period these structures often experience the exciting in-plane forces. Due to these pulsating in-plane forces the stiffened panels may lead to the condition of dynamic instability where these structures vibrate in the transverse direction and the amplitude of vibration increases without bound. This condition is also called parametric resonance. This phenomenon is entirely different from the usual resonance of forced vibration. In forced vibration when the frequency of the transverse forcing system matches with the natural frequency of the structure, resonance occurs. Thus the resonance phenomenon in forced vibration problem is relatively simple since the structure loses stability at constant frequencies of the transverse loads. On the other hand the instability in case of parametric resonance occurs over a range of frequencies of the in-plane force rather than a single value. Again parametric resonance of a structure may occur at load level much less than the static buckling load while the static instability of the structure sets in at the static buckling load values. Thus a structural component designed to withstand static buckling load may easily fail in an environment having periodic in-plane loading. So a designer is ought to consider the parametric resonance aspect while dealing a structure subjected to dynamic loading atmosphere.

Many researchers in the past have undertaken the work on dynamic instability of structures. Bolotin [1] presented the general theory of dynamic stability of various elastic systems and discussed the peculiarities of the phenomena of instability. Extensive bibliographies of earlier works on these problems are given in review papers of Evan-Iwanowski [2], Ibrahim [3], and Simitses [4]. Hutt and Salam [5] first used the finite element method for the dynamic instability analysis of plates subjected to uniform loading. Most of the investigators [6–9] studied the dynamic instability behavior of closed cylindrical shell with a simply supported boundary condition using an analytical approach. The dynamic instability of conical shell was studied by Ng et al. [10] using generalized differential quadrature method. However, the studies of dynamic instability for shell panels are few in literature. Sahu and Datta [11] studied the dynamic instability of doubly curved panels with nonuniform edge loadings. The dynamic stability of uniaxially loaded cylindrical panels with transverse shear effect was studied by Ng et al. [12]. However, for the stiffened structure, Troitsky [13] extensively reviewed the literature pertaining to rectangular stiffened plates for static, dynamic, and stability analyses, which are based on orthotropic plate idealization. Studies on free vibration of stiffened plates and shells are available in literature [14–20]. Timoshenko and Gere [21] presented numerical tables for buckling load of rectangular plates stiffened by longitudinal and transverse ribs. The dynamic stability analyses of stiffened plates and shells are also very few in literature. Thomas and Abbas [22] presented the vibration characteristics and dynamic stability of stiffened plates. The parametric resonance of stiffened rectangular plates is investigated by Duffield and Willems [23]. Merrit and Willems [24] investigated the dynamic instability behavior for skew stiffened plates. The dynamic stability of radially stiffened annular plates with radial stiffener subjected to in-plane force is investigated by Mermertas and Belek [25]. Liao and Cheng [26] gave some results for dynamic instability of stiffened composite plates and shells with uniform in-plane forces. Shrivastava et al. [27,28] investigated the dynamic instability of isotropic stiffened plates with uniform and nonuniform edge loadings. For

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 17, 2007; final manuscript received January 2, 2009; published online April 22, 2009. Review conducted by Sanjay Govindjee.

stiffened/unstiffened structures with cutout the available literature regarding vibration, buckling, and dynamic instability analyses is relatively few. Sahu and Datta [29] investigated the effect of dynamic instability characteristics on doubly curved panels with opening. Srivastava et al. [30] studied the effect of square cutouts and stiffening scheme on dynamic instability of stiffened plates using finite element method. The earlier contributions do not cover all aspects of dynamic instability behavior of stiffened shell panels with cutout. The present work gives a contribution to study the dynamic instability behavior of the stiffened shell panels with cutout, but experimental works in large scale are still to be performed to understand the phenomena completely.

In order to model a shell panel without any significant approximation related to the representation of arbitrary shell geometry, structural deformation, and other associated aspects, the isoparametric 3D degenerated shell element [31,32] having eight nodes is used. For the stiffeners, a compatible three-noded isoparametric curved beam element is used. The beam element is always placed along the edge of shell elements and this is intentionally not placed within the shell element in order to avoid the problem of stress jump within the shell element. The basic concept underlying in the formulation of degenerated shell element [31] has been extended to derive beam/stiffener elements having any arbitrary curve geometry suitable for use in two- or three-dimensional problems [33,34]. Unfortunately the 3D degenerated beam element based on the above formulation [33,34] has some problem in torsional mode since it overestimates torsonal rigidity [34]. The problem becomes more severe in case of stiffeners having a narrow cross section like blade stiffener. Keeping this aspect in view, the stiffener element is reformulated where the above mentioned problem has been eliminated by using torsion correction factor. In order to achieve that the stiffener bending in the plane of the shell surface is neglected. This should not affect the solution accuracy since deformation of the stiffener in that plane will be very small due to high in-plane rigidity of the shell skin. Moreover, the new formulation has the advantage that it requires five degrees of freedom per node while it is six in case of existing formulation [33,34]. Actually the stiffener element will directly share the five nodal unknowns of the shell element.

In the present study the dynamic instability analyses are carried out for different types of stiffened shell panels such as flat plate, cylindrical shell panel, spherical shell panel, and parabolic hyperboloid shell panel with different cutout sizes at the center. The effects of various parameters such as shell geometry, radius of curvature, cutout size, stiffening scheme, and dynamic load factors are considered in dynamic instability analysis of stiffened shell panels with cutout. Both deep and shallow shell panels are considered in the present investigation.

2 Theory and Formulation

It has been mentioned in Sec. 1 that the stiffened panel structure is modeled by finite element technique. In order to have a better representation, the shell skin and stiffeners are modeled as discrete/separate elements. The formulation of these elements is presented below.

2.1 Shell Element. The formulation of the shell element is based on the basic concept of Ahmad [31], where the three-dimensional solid element used to model the shell is degenerated with the help of certain extractions obtained from the consideration that one of the dimensions across the shell thickness is sufficiently small compared with other dimensions (Fig. 1). The detailed derivation of this element is available in literature [31,32].

2.1.1 Elastic Stiffness Matrix. The elastic stiffness matrix of an element can be derived easily and it is expressed as

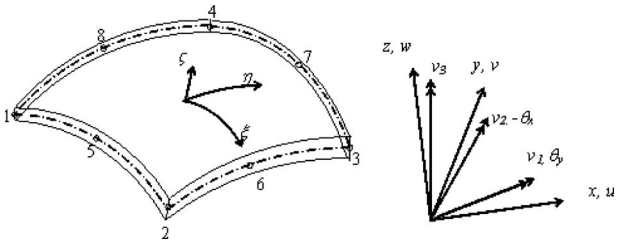


Fig. 1 Eight-noded quadrilateral degenerated shell element

$$[k_e] = \int [B]^T[D'][B]dxdydz = \int [B]^T[D'][B]|J|d\xi d\eta d\varsigma \quad (1)$$

where $|J|$ is the determinant of the Jacobian matrix $[J]$.

2.1.2 Mass Matrix. The consistent mass matrix has been adopted in the present study. Following the usual techniques, it can be derived and expressed as

$$[m_e] = \int \rho[N_D]^T[N_D]dxdydz = \int \rho[N_D]^T[N_D]|J|d\xi d\eta d\zeta \quad (2)$$

where ρ is the material density.

2.1.3 Geometric Stiffness Matrix. The geometric stiffness matrix of the element is expressed as

$$[k_g] = \int [B_G]^T[\tau][B_G]dv = \int [B_G]^T[\tau][B_G]|J|d\xi d\eta d\zeta$$

where

$$[\tau] = \begin{bmatrix} [\tau^1] & [0] & [0] \\ [0] & [\tau^1] & [0] \\ [0] & [0] & [\tau^1] \end{bmatrix} \quad (3)$$

The submatrix $[\tau^1]$ within the initial stress matrix $[\tau]$ in the above equation is

$$[\tau^1] = \begin{bmatrix} \sigma'_x & \tau'_{xy} & \tau'_{xz} \\ \tau'_{xy} & \sigma'_y & \tau'_{yz} \\ \tau'_{xz} & \tau'_{yz} & 0 \end{bmatrix} \quad (4)$$

2.2 Stiffener Element. The derivation of the stiffener element is based on the basic concept used to derive the shell element. In this case the stiffener element modeled with three-dimensional solid element is degenerated with the help of certain extractions obtained from the consideration that the dimension across the stiffener depth as well as breadth is small compared with that along the length. The stiffener element follows an edge of a shell element where the parameters of three nodes lying on that shell element edge are used to express the geometry and deformation of the stiffener utilizing compatibility between shell and stiffeners. It helps to eliminate the involvement of additional degrees of freedom for the modeling of stiffeners. The stiffener element having any arbitrary curved geometry is mapped into a regular domain in ξ-η-ζ coordinate system where all these coordinates vary from -1 to $+1$. Again ξ is taken along the stiffener axis while η and ζ are taken along the width and depth directions, respectively. It has been found that the vectors $\bar{v}_{1i}$, $\bar{v}_{2i}$, and $\bar{v}_{3i}$ are quite useful for the representation of geometry and deformation of the shell element. For the stiffener element a similar set of vectors $\bar{v}^s_{1i}$, $\bar{v}^s_{2i}$, and $\bar{v}^s_{3i}$ is used and these may be obtained from those of the shell elements ($\bar{v}_{1i}$, $\bar{v}_{2i}$, and $\bar{v}_{3i}$) as

$$\bar{v}_1^s = \bar{v}_1 \cos\theta_s + \bar{v}_2 \sin\theta_s, \quad \bar{v}_2^s = -\bar{v}_1 \sin\theta_s + \bar{v}_2 \cos\theta_s, \quad \text{and} \quad \bar{v}_3^s = \bar{v}_3 \tag{5}$$

where $(\bar{v}_{1i}^s - \bar{v}_{2i}^s)$ is oriented at an angle of θ_s with respect to $(\bar{v}_{1i} - \bar{v}_{2i})$ and $\bar{v}_{1i}^s$ follows the stiffener axis.

With these vectors, the coordinate at any point within the stiffener may be expressed in terms of coordinates (x_i, y_i, z_i) of those three nodes of the corresponding shell element edge as

$$\begin{Bmatrix} x \\ y \\ z \end{Bmatrix} = \sum_{i=1}^{3} N_{si} \begin{Bmatrix} x_i \\ y_i \\ z_i \end{Bmatrix} + \sum_{i=1}^{3} N_{si}\left(\frac{\zeta d_s}{2} + e\right) \begin{Bmatrix} l_{3i}^s \\ m_{3i}^s \\ n_{3i}^s \end{Bmatrix} + \sum_{i=1}^{3} N_{si}\left(\frac{\eta b_s}{2}\right) \times \begin{Bmatrix} l_{2i}^s \\ m_{2i}^s \\ n_{2i}^s \end{Bmatrix} \tag{6}$$

where b_s is stiffener width, d_s is its depth, and e is the eccentricity (distance of the stiffener axis from the shell midsurface). The expressions of the quadratic shape functions N_{si} along ξ are as follows:

$$N_{s1} = \xi(\xi - 1)/2, \quad N_{s2} = 1 - \xi^2, \quad N_{s3} = \xi(\xi + 1)/2 \tag{7}$$

Now considering the deformation of the stiffener element, the present formulation differs from the usual one [33,34] where six degrees of freedom are generally taken to represent the biaxial bending apart from torsion and axial deformations. In the present study the bending of the stiffener in the tangential plane of the shell is not considered. This has helped to eliminate the involvement of the sixth degrees of freedom θ_z like that of a shell element. Moreover the usual formulation [33,34] overestimates the torsional rigidity and it cannot be corrected simply with some correction factor since it got mixed with other terms. The present formulation facilitates to treat it nicely where a torsion correction factor is introduced. Actually this is the primary object for the reformulation of the stiffener element. Based on this the displacement components at any point within the stiffener may be expressed as

$$\begin{Bmatrix} u \\ v \\ w \end{Bmatrix} = \sum_{i=1}^{3} N_{si}[T_{vi}] \begin{Bmatrix} u_i \\ v_i \\ w_i \end{Bmatrix} - \sum_{i=1}^{3} N_{si}\left(\frac{\zeta d_s}{2} + e\right) \begin{bmatrix} l_{1i} & l_{2i} \\ m_{1i} & m_{2i} \\ n_{1i} & n_{2i} \end{bmatrix} \begin{Bmatrix} \theta_{xi} \\ \theta_{yi} \end{Bmatrix} = [N_{Ds}]\{\delta_s\} \tag{8}$$

where $\{\delta_s\} = [u_1\ v_1\ w_1\ \theta_{x1}\ \theta_{y1}\ u_2\ v_2\ \ldots\ \theta_{y3}]^T$ and the matrix $[T_{vi}]$ is used to make the component of translational displacement along $\bar{v}_{2i}^s$ at shell midplane zero since the bending of the stiffener in the tangential plane of the shell is not considered. Its effect should be insignificant since bending deformation in this mode will be very small due to high in-plane rigidity of the shell skin. Moreover the flexural rigidity of stiffener in this mode is usually found to be small.

The matrix $[T_{vi}]$ used in the above equation may be expressed with the help of $\bar{v}_{1i}^s$ and $\bar{v}_{2i}^s$ as

$$[T_{vi}] = \begin{bmatrix} l_{1i}^s & 0 & l_{3i}^s \\ m_{1i}^s & 0 & m_{3i}^s \\ n_{1i}^s & 0 & n_{3i}^s \end{bmatrix} \begin{bmatrix} l_{1i}^s & m_{1i}^s & n_{1i}^s \\ l_{2i}^s & m_{2i}^s & n_{2i}^s \\ l_{3i}^s & m_{3i}^s & n_{3i}^s \end{bmatrix} \tag{9}$$

Similar to shell element, the stress and strain components at any point within the stiffener element are taken in a local axis system $(x'\text{-}y'\text{-}z')$ corresponding to $\bar{v}_{1i}^s$, $\bar{v}_{2i}^s$, and $\bar{v}_{3i}^s$. The relationship between them may be expressed as

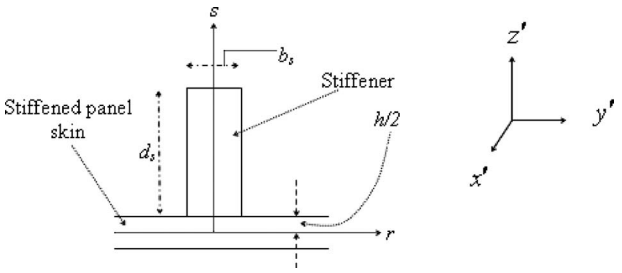


Fig. 2 Stiffener attached to the panel surface

$$\begin{Bmatrix} \sigma_{x'} \\ \tau_{x'z'} \\ \tau_{x'y'} \end{Bmatrix} = \begin{bmatrix} \bar{Q}_{1m} & 0 & 0 \\ 0 & \beta_s \bar{Q}_{5m} & 0 \\ 0 & 0 & \beta_t \bar{Q}_{6m} \end{bmatrix} \begin{Bmatrix} \varepsilon_{x'} \\ \gamma_{x'z'} \\ \gamma_{x'y'} \end{Bmatrix} \quad \text{or} \quad \{\sigma'\} = [D_s']\{\varepsilon'\} \tag{10}$$

where β_s is the shear correction factor, which is taken as 5/6. The torsion correction factor β_t and other rigidity parameters in the rigidity matrix $[D_s']$ are presented below. The stress-strain relationship of the stiffener in the axis system $(x'\text{-}r\text{-}s)$ as shown in Fig. 2 may be written as

$$\begin{Bmatrix} \sigma_{x'} \\ \sigma_r \\ \tau_{x'r} \\ \tau_{x's} \\ \tau_{rs} \end{Bmatrix} = \begin{bmatrix} \frac{E}{1-\nu^2} & \frac{\nu E}{1-\nu^2} & 0 & 0 & 0 \\ \frac{\nu E}{1-\nu^2} & \frac{E}{1-\nu^2} & 0 & 0 & 0 \\ 0 & 0 & G & 0 & 0 \\ 0 & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & G \end{bmatrix} \begin{Bmatrix} \varepsilon_{x'} \\ \varepsilon_r \\ \gamma_{x'r} \\ \gamma_{x's} \\ \gamma_{rs} \end{Bmatrix} \tag{11}$$

The rigidity parameters $[D_s']$ of Eq. (10) may be obtained from Eq. (11); utilizing the stress-free conditions $(\sigma_r = 0)$ and $(\tau_{rs} = 0)$, the rigidity parameters will be

$$\bar{Q}_{1m} = \frac{E}{1-\nu^2} \quad \text{and} \quad \bar{Q}_{5m} = \bar{Q}_{6m} = G \tag{12}$$

The torsion correction factor β_t may be written as

$$\beta_t = \frac{3kb_s^2 \bar{Q}_{5m}\left(\frac{h}{2} - \left(d_s + \frac{h}{2}\right)\right)}{\bar{Q}_{6m}\left(\left(\frac{h}{2}\right)^3 - \left(d_s + \frac{h}{2}\right)^3\right)} = \frac{3kb_s^2\left(\frac{h}{2} - \left(d_s + \frac{h}{2}\right)\right)}{\left(\left(\frac{h}{2}\right)^3 - \left(d_s + \frac{h}{2}\right)^3\right)} \tag{13}$$

where k is the factor to get torsional constant of an isotropic beam having rectangular section, which is a function of d_s/b_s ratio of the cross section. The torsional constant (k) for the rectangular section (depth$=d_s$ and thickness (width)$=b_s$) can be calculated from the torsional equation of rectangular cross section given by Timoshenko and Goodier [35] as

$$k = \frac{1}{3}\left[1 - \frac{192}{\pi^5}\left(\frac{b_s}{d_s}\right) \sum_{n=1,3,5,\ldots}^{\infty} \frac{1}{n^5} \tanh\frac{n\pi(d_s/2)}{b_s}\right] \tag{14}$$

Now Eqs. (5)–(10) may be used to derive the elastic stiffness matrix $[k_{es}]$, mass matrix $[m_{es}]$, and geometric stiffness matrix $[k_{gs}]$ of a stiffener element following the procedure used for shell element, and these matrices may be expressed as follows:

$$[k_{es}] = \int [B_s]^T [D_s'][B_s] dx dy dz = \int [B_s]^T [D_s'][B_s] |J| d\xi d\eta d\zeta \tag{15}$$

$$[m_{es}] = \int \rho[N_{Ds}]^T[N_{Ds}]dxdydz = \int \rho[N_{Ds}]^T[N_{Ds}]|J|d\xi d\eta d\zeta \tag{16}$$

$$[k_{gs}] = \int [B_{Gs}]^T[\tau_s][B_{Gs}]dxdydz = \int [B_{Gs}]^T[\tau_s][B_{Gs}]|J|d\xi d\eta d\zeta \tag{17}$$

where $[B_s]$ and $[B_{Gs}]$ are analogous to $[B]$ and $[B_G]$ of the shell element, respectively. The initial stress matrix $[\tau_s]$ looks identical to that of the shell element but its submatrix $[\tau_s^1]$ for the stiffener element is

$$[\tau_s^1] = \begin{bmatrix} \sigma'_{xs} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{18}$$

The elastic stiffness matrix, mass matrix, and geometric stiffness matrix are computed for all the shell elements and stiffener elements of the entire structure, and these matrices are accordingly assembled together to form the corresponding global matrices $[K]$, $[M]$, and $[K_G]$ of the stiffened shell panels with cutout. The skyline storage algorithm is used to keep these three big size matrices in single array. Both deep shell and shallow shell panels can be analyzed with these elements.

2.3 Governing Equations. With the stiffness matrix $[K]$, mass matrix $[M]$, and geometric stiffness matrix $[K_G]$ of the total structure (panel skin and stiffeners, obtained in the Secs. 2.1 and 2.2 and adding those, respectively), the equation of motion of the structure can be written as

$$[M]\{\ddot{q}\} + [[K] - P[K_G]]\{q\} = \{0\} \tag{19}$$

This is a general equation and it can be reduced as a special case to get the governing equations for buckling, vibration, and dynamic stability problems as follows.

2.3.1 Buckling

$$[[K] - P_{cr}[K_G]]\{q\} = \{0\} \tag{20}$$

where P_{cr} is the critical load of buckling.

2.3.2 Vibration

$$[[K] - P[K_G]]\{q\} - \omega^2[M]\{q\} = \{0\} \tag{21}$$

where ω is the vibration frequency of the structure subjected to in-plane load and it becomes the natural frequency of vibration if P attains zero value. The frequency of vibration becomes zero when P attains critical value.

2.3.3 Dynamic Stability. Equation (19) can also be used to solve the dynamic stability problem. Let the in-plane load P be periodic and may be expressed as

$$P(t) = P_s + P_t \cos \Omega t \tag{22}$$

where Ω is the frequency of excitation, P_s is the static component of P, and P_t is the amplitude of its dynamic component, which may be expressed in terms of static buckling load P_{cr} as follows:

$$P_s = \alpha P_{cr}, \quad P_t = \beta P_{cr} \tag{23}$$

where α and β may be defined as static and dynamic load factors, respectively. Now Eq. (19) can be written as

$$[M]\{\ddot{q}\} + [[K] - \alpha P_{cr}[K_G] - \beta P_{cr}[K_G]\cos \Omega t]\{q\} = 0 \tag{24}$$

The above equation represents a system of second order differential equation with periodic coefficient, which is basically the Mathieu–Hill equation. The boundaries of dynamic instability regions can be found by the periodic solutions having periods of T and $2T$, where $T = 2\pi/\Omega$. The range of primary instability region with period of $2T$ is of practical importance [1] where the solution can be achieved by expressing $\{q\}$ in the form of the trigonometric series as

$$\{q\} = \sum_{k=1,3,5}^{\infty} \left[\{a_k\}\sin\frac{k\Omega t}{2} + \{b_k\}\cos\frac{k\Omega t}{2} \right] \tag{25}$$

After substitution of the above equation (25) in Eq. (24) and taking the first term of the series, the quantities associated with $\sin \Omega t/2$ and $\cos \Omega t/2$ are separated out and processed accordingly to eliminate the time dependent component, and it leads to

$$\left[[K] - \alpha P_{cr}[K_G] \pm \frac{1}{2}\beta P_{cr}[K_G] - \frac{\Omega^2}{4}[M] \right]\{q_{ab}\} = 0 \tag{26}$$

where $\{q_{ab}\}$ is either $\{a_k\}$ or $\{b_k\}$ depending on the use of plus or minus of the dynamic in-plane load component, respectively. It is basically an eigenvalue problem and it can be solved for known value of α, β, and P_{cr}. The two frequencies corresponding to plus and minus will indicate the boundaries of the dynamic instability region.

2.4 Method of Solution. In the present investigation the different analyses for the laminated composite stiffened plate are implemented by the finite element computer program written in FORTRAN-90. The equations are solved using the technique proposed by Corr and Jennings [36] where the matrices $[K]$, $[M]$, and $[K_G]$ are stored in a single array according to skyline storage algorithm. In all the cases, the stiffness matrix $[K]$ is factorized according to Cholesky's decomposition technique. With this, the solution for displacement is simply obtained by its forward elimination and backward substitution techniques. These displacement components are used to find out the stress field in the Gauss points. It is necessary because the stress field in the structure is not uniform due the nonuniform nature of in-plane load, boundary conditions, and presence of stiffeners. These stresses are used to calculate the geometric stiffness matrices. The solution of Eqs. (20), (21), and (26) goes through a number of operations [36]. Moreover Eq. (26) requires a number of iterations to get the solution since it comes under the category of eigenvalue problem. The simultaneous iteration eigensolver technique developed by Corr and Jennings [36] is utilized to extract the eigenvalues. In such cases, the solution of eigenvector and eigenvalue is more than one where the different solutions correspond to different modes of vibration or different modes of buckling. The mode with the lowest eigenvalue is quite important and it is known as fundamental mode. In the present investigation the dynamic instability regions are plotted for the first mode of vibration.

3 Results and Discussions

The convergence and validation of the proposed finite element model are presented, taking various examples from literature. The problem under investigation is explained and the numerical results of the considered problem are discussed in this section.

3.1 Convergence and Validation. The convergence and accuracy of the proposed method are first established by comparing the results of various problems with those of earlier investigators' available in literature. The problems taken are described below.

3.1.1 Free Vibration of a Simply Supported Stiffened Plate. The problem of a simply supported stiffened plate as shown in Fig. 3 is used to carry out the convergence study. In the edges along the side a the displacements u, w, and θ_x are restricted and v and θ_y are allowed. In the edges along the side b the displacements v, w, and θ_y are restricted and u and θ_x are allowed. All these five displacements are the same as the displacement used in the finite element formulation. The whole structure is modeled with different mesh sizes for the free vibration analysis. The de-

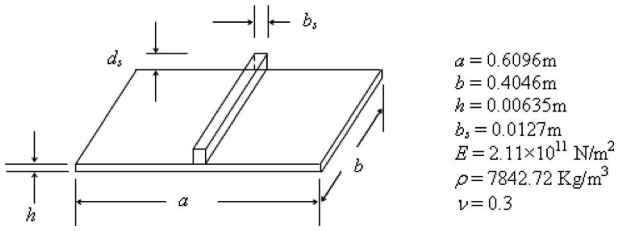


Fig. 3 Rectangular stiffened plate

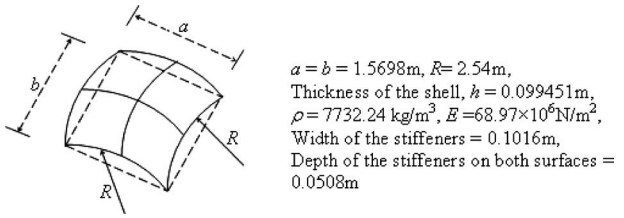


Fig. 4 Stiffened spherical shell panel having a square base

tails of the stiffened plate are given in Fig. 3. The results for the first two natural frequencies are shown in Table 1. It shows a rapid convergence with mesh refinement.

Again a mesh size of 8×8 is found to be sufficient to attain the convergence. For the purpose of validation, the frequencies are compared with the results reported by Long [37] using composite beam-plate method and with finite element results of Thomas and Abbas [22]. The results are in good agreement.

3.1.2 Free Vibration of a Spherical Shell Panel of Square Base Having Two Stiffeners Along Its Two Center Lines. The simply supported stiffened spherical shell panel as shown in Fig. 4 is used to investigate the free vibration analysis. The details of the shell and stiffeners are given in Fig. 4 where the stiffeners are taken to be symmetric with respect to the shell midsurface. The whole structure is modeled with an 8×8 mesh for the analysis. The first five natural frequencies are compared with the finite element results of Nayak and Bandopadhyay [17], Samanta and Mukhopadhyay [19], and Prusty [38] in Table 2. The results obtained in these studies [17,19,38] are found to be in good agreement compared with the present results.

3.1.3 Buckling of a Rectangular Plate With a Central Stiffener Under Uniaxial Load. The problem of the rectangular stiffened plate (Fig. 5) having simply supported boundary conditions at its four edges is investigated for different plate aspect ratios (a/b) and stiffener parameters $(\delta=A_s/bh,\ \gamma=EI_s/bD)$. The plate thickness ratio (a/h) and isotropic plate and stiffener material (ν) are taken as 100 and 0.3, respectively. The dimension a is varied keeping b as constant to get different values of aspect ratio (a/b). The stiffened plate is modeled with an 8×8 mesh for the analysis in all cases. The nondimensional critical buckling stress parameter $k=\sigma_{cr}/(\pi^2D/b^2h)$ obtained in the present analysis is presented in Table 3 with the analytical solution of Timoshenko and Gere [21] and finite element solution of Mukhopadhyay [16]. The table shows that the agreement between the results is very good. As Timoshenko and Gere [21] did not consider the effect of stiffener eccentricity and torsional rigidity, these parameters are taken as zero in the present problem.

3.1.4 Buckling of a Rectangular Plate With Central Cutouts. The problem of simply supported plate ($a=b=0.508$ m and $h=0.00254$ m, Fig. 6) having square cutout ($ca=cb$) is investigated for different cutout ratios (ca/a). The material properties of this plate, made up of Ti–6Al–4V titanium alloy [39] are $E=11.04\times10^{10}$ N/m^2, $G=4.28\times10^{10}$ N/m^2, and $\nu=0.31$. The whole structure is modeled with a 10×10 mesh for the buckling analysis. The contributions of the elements coming under the cutout are not included in forming the global stiffness $[K]$ and global geometric stiffness $[K_G]$ matrices. The results obtained in the present method are compared with the structural performance and resizing (SPAR) finite element results of Ko [39] in Table 4. The results are in good agreement.

3.1.5 Dynamic Instability of Simply Supported Square Plate With Central Cutouts. The dynamic instability analysis is carried out for a simply supported plate ($a=b=0.5$ m and $h=0.005$ m, Fig. 6) having square ($ca=cb$) cutout ratios of $ca/a=0.2$ and $ca/a=0.8$. The material properties of the plate are $E=70\times10^9$ N/m^2, $G=26.92\times10^9$ N/m^2, $\nu=0.3$, and $\rho=2800$ kg/m^3. The whole structure is modeled with a 10×10 mesh for the dynamic instability analysis. The first mode buckling load of this plate without cutout is taken as the reference load to plot the dynamic instability regions of the plate for both cutout sizes. The static load factor (α) is 0.0. The nondimensional excitation frequencies $(\bar{\Omega}=\Omega a^2\sqrt{\rho h/D})$ obtained in the present analysis are plotted in Fig. 7 along with the finite element results of Sahu and Datta [29]. The results are matching well.

3.1.6 Dynamic Instability of Stiffened Square Plate Simply Supported at All the Four Sides. The square plate ($a=b=0.6$ m and $h=0.00633$ m) is attached with a stiffener ($b_s=0.0127$ m and

Table 1 Natural frequency (Hz) of a simply supported rectangular stiffened plate

	Mode 1			Mode 2		
	d_s (m)			d_s (m)		
Analysis	0.0	0.0254	0.0508	0.0	0.0254	0.0508
Present (2×2)	197.750	381.348	703.148	1419.06	1949.09	1572.55
Present (4×4)	137.525	270.400	305.941	279.151	293.429	348.586
Present (6×6)	136.901	263.784	291.624	264.098	279.250	325.867
Present (8×8)	136.846	262.932	290.413	263.255	278.285	323.185
Present (10×10)	136.857	262.765	290.181	263.109	278.102	322.692
Present (12×12)	136.855	262.718	290.110	263.053	278.046	322.552
Present (14×14)	136.854	262.702	290.082	263.047	278.023	322.500
Thomas and Abbas[a]	136.5	259.1	293.8	265.5	280.9	332.4
Long[b]		267.1	273.8		280.3	331.8

[a]Reference [22].
[b]Reference [37].

Table 2 Natural frequency (rad/s) of a simply supported spherical shell panel with two stiffeners along the two central lines

Mode No.	Present (8×8)	Nayak and Bandopadhyay[a] El-8[b] (8×8)	Nayak and Bandopadhyay[a] El-9[c] (8×8)	Samanta and Mukhopadhyay[d]	Prusty[e] (16×16)
1	40.2300	40.26	40.26	41.70	40.81
2	69.0282	70.98	70.97	74.11	72.13
3	69.1158	70.98	70.97	74.36	72.13
4	91.2984	96.06	96.06	99.18	92.92
5	105.680			104.94	105.69

[a]Reference [17].
[b]Eight-noded element.
[c]Nine-noded element.
[d]Reference [19].
[e]Reference [38].

d_s=0.0222 m) on the bottom surface along its one center line. The load is acting at the edges uniformly in the direction of the stiffener. All the sides are simply supported. The nondimensional excitation frequency ($\bar{\Omega}=\Omega a^2\sqrt{\rho h/D}$) is obtained, taking first mode buckling load of the stiffened plate. The static load factor (α) is kept as 0.2. The results of the upper and lower bounding frequencies of the exciting force are presented in Table 5 along with the finite element results taken from the graph (Fig. 10) of Srivastava et al. [28]. It shows that the results are in good agreement.

3.2 Problem Under Investigation. It is a doubly curved stiffened shell panel with square cutout located at the center of the panel subjected to uniform edge loading (Fig. 8) along x-direction. The panel is attached with one stiffener along x-direction in the center line. Four other stiffeners are also attached to the four sides of the opening; the length of each stiffener is equal to the size of the opening. The projection of the cutout in xy-plane is square (ca/cb=1.0) in all cases and located at the center of the panels. The geometrical properties of the stiffened panel are $a=b$ =0.5 m, h (thickness of the skin) =0.005 m, b_s (stiffener width) =0.005 m, and d_s (stiffener depth) =0.02 m; and the material properties of the panel skin and stiffeners are $E=70\times10^9$ N/m^2, $G=26.92\times10^9$ N/m^2, ν=0.3, and ρ=2800 kg/m^3. The stiffeners are attached to the bottom surface in all cases. All four sides of the stiffened panels are simply supported in prebuckling, buckling, free vibration, and dynamic instability analysis. The equation of surface for flat panel, cylindrical shell panel, and spherical shell panel is defined as

$$z=\sqrt{R_x^2-x^2}+\left(\sqrt{R_y^2-\left(y-\frac{b}{2}\right)^2}-\sqrt{R_y^2-\left(\frac{b}{2}\right)^2}\right) \quad (27)$$

where x varies from $-a/2$ to $+a/2$ and y varies from 0 to b.

The flat surface is obtained by taking $R_x=R_y=0.5\times10^6$ m, which is sufficiently large in comparison to a and b. The cylindrical surface is obtained by taking $R_x=0.5\times10^6$ m. To get the cylindrical surface with different radii of curvature, R_y is varied and it varies from 0.2625 m to 0.5×10^6 m. When R_y becomes 0.5×10^6 m the cylindrical panels become flat panels. The spherical shell panel is obtained by varying both R_x and R_y. Both radii of curvatures vary from 0.2625 m to 0.5×10^6 m and the ratio of radii of curvatures R_x and R_y is 1.0 in all cases. When both radii of curvature become 0.5×10^6 m, the spherical panels also reduce to the case of flat panel like the cylindrical panel. The surface of parabolic hyperboloid is defined as

$$z=2\times\sqrt{R_x^2-\left(\frac{a}{2}\right)^2}-\sqrt{R_x^2-x^2}+\left(\sqrt{R_y^2-\left(y-\frac{b}{2}\right)^2}-\sqrt{R_y^2-\left(\frac{b}{2}\right)^2}\right) \quad (28)$$

In this case R_x varies from -0.2625 m to -0.5×10^6 m and R_y varies from 0.2625 m to 0.5×10^6 m with the ratio of R_x and R_y as -1.0 in all cases. When $(-R_x)=R_y=0.5\times10^6$ m, this panel also reduces to flat panel. The basic shapes of the panels obtained from the above equations (27) and (28) are shown in Fig. 9.

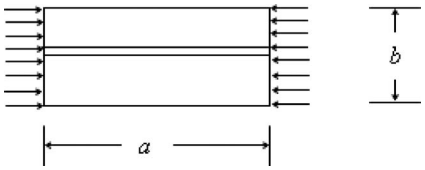


Fig. 5 Simply supported stiffened rectangular plate under uniaxial compression

Table 3 Buckling load parameter (k) for a simply supported rectangular stiffened plate under uniaxial compression

a/b	γ=10, δ=0.05 Present (8×8)	Timoshenko and Gere[a]	Mukhopadhyay[b]	γ=5, δ=0.2 Present	Timoshenko and Gere[a]	Mukhopadhyay[b]
0.6	16.403	16.5		16.463	16.5	
0.8	16.686	16.8		12.729	13.0	
1.0	15.908	16.0	15.91	9.612	9.72	9.65
2.0	10.110	10.2	10.16	6.244	6.24	6.24
3.0	11.867	12.0	11.94	6.512	6.53	6.48
4.0	10.135	10.2		6.264	6.24	

[a]Reference [21].
[b]Reference [16].

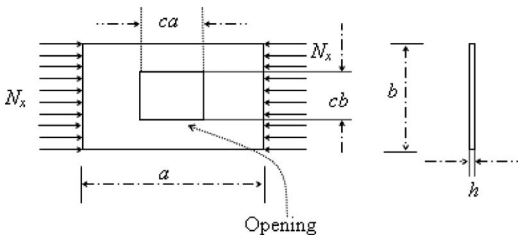


Fig. 6 Perforated plate under uniaxial compression

The results of free vibration frequencies and excitation frequencies of the load obtained in rad/s and the critical buckling stress obtained in N/m^2 are presented in nondimensional form as follows:

(i) nondimensional natural frequency: $\varpi = \omega b^2\sqrt{\rho h/D}$

(ii) nondimensional buckling load: $\bar{P}_{cr} = P_{cr}b/D$, where $P_{cr} = \sigma_{cr}(b \cdot h + b_s \cdot d_s)$

(iii) nondimensional bounding excitation frequency: $\bar{\Omega} = \Omega b^2\sqrt{\rho h/D}$

3.3 Numerical Results. The numerical results taking various parameters of the investigation are presented in three groups as free vibration, buckling, and dynamic instability. In the analysis the whole structure is divided with a 10×10 mesh.

3.3.1 Free Vibration. The free vibration analysis of all four (plate, cylindrical, spherical, and parabolic hyperboloid) types of stiffened panels is carried out considering the effect of radius of curvature and different cutout sizes. First to study the effect of radius of curvature on the free vibration characteristics of the panel, each of the four panels is divided in four types considering the attachment of stiffener and cutout ($ca/a = 0.4$). Those are the following.

(1) Type (i): unstiffened without cutout
(2) Type (ii): unstiffened with cutout ($ca/a = 0.4$)
(3) Type (iii): stiffened without cutout
(4) Type (iv): stiffened with stiffened cutout ($ca/a = 0.4$)

The nondimensional natural frequencies are presented in Table 6. Then to find out the effect of cutout size on the free vibration behavior of the panels, each of the four panels is divided into two types as un_c (unstiffened panels with opening) and st_c (stiffened panel with stiffened opening). In this analysis three radii of curvature to span ratio as 0.525, 1.0, and 10.0 are taken. The cutout ratios (ca/a) taken are 0.0, 0.2, 0.4, 0.6, and 0.8. The results are presented in Table 7.

It is observed from Table 6 that the frequency of the unstiffened cylindrical shell panel (Type (i)) gradually decreases with the increase in radius of curvature. This is because the stiffness of the panel gradually reduces with the increase in radius of curvature. The same panel when a cutout (of ratio $ca/a = 0.4$) is made (Type (ii)) the free vibration behavior changes significantly. The frequency of vibration increases with the increase in radius of curvature up to some point and then decreases with the increase in

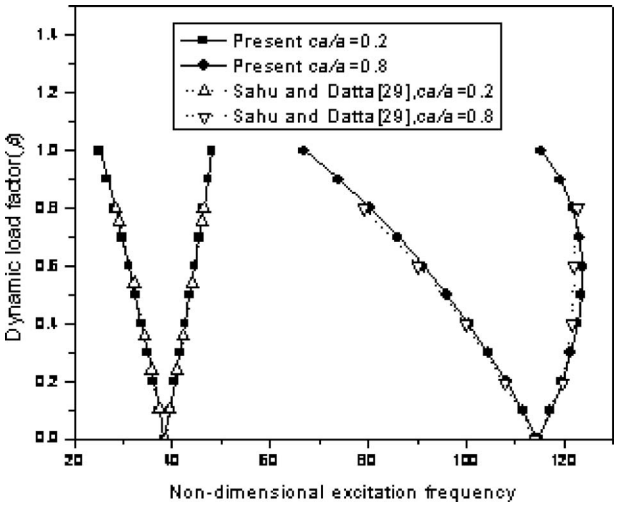


Fig. 7 Dynamic instability regions of plate with square cutout under uniaxial compression

radius of curvature. It is maximum for a particular value of radius of curvature. If the radius of curvature is less or more than this value the frequency decreases. With a cutout in the structure, both the stiffness and the mass reduce. The frequency of vibration depends on both the stiffness and mass of the structure. The behavior of panel (Type (ii)) is attributed by the combined effect of stiffness and mass of the structure. If a stiffener is attached to this panel (Type (iii)), the natural frequency of vibration gradually decreases with the increase in radius of curvature as (Type (i)). But with a low radius of curvature the frequency of the unstiffened cylindrical panel (Type (i)) is more than that of the stiffened cylindrical (Type (iii)) panel. The stiffener does not increase the frequency rather it decreases the frequency at a lower radius of curvature. When a stiffener is attached to the structure both the stiffness and the mass of the structure increase. At a low radius of curvature the cylindrical panel is at a higher stiffness. So due to the stiffener the net stiffness of the structure does not increase more in comparison to the increase in the mass. It causes the reduction in frequency of vibration. This may be increased with the increase in number of very thin stiffeners.

At a higher radius of curvature the frequency of the unstiffened panel is less than that of the stiffened panel. In the case of the stiffened cylindrical shell panel with stiffened opening (Type (iv)), the frequency of vibration increases with the increase in radius of curvature up to some point and then decreases with the increase in radius of curvature as that of Type (ii). In spherical shell panel for all four types (Types (i)–(iv)), the frequency of vibration increases with the increase in radius of curvature up to some point and then decreases with the increase in radius of curvature. Again for low radius of curvature the frequency of the unstiffened spherical panel (Type (i)) is more than that of the stiffened spherical (Type (iii)) panel like the cylindrical panel. The frequency of vibration of the spherical shell panel stiffened/unstiffened without cutout in low radius of curvature is very high in comparison to cylindrical, parabolic hyperboloid, and flat panel (plate). But the frequency of

Table 4 Buckling load N_x (lb/in.) of the simply supported square under uniaxial compression

	Buckling load N_x (lb/in.)				
Analysis	$ca/a = 0.0$	$ca/a = 0.2$	$ca/a = 0.4$	$ca/a = 0.6$	$ca/a = 0.8$
Present (10×10)	146.157	127.551	110.424	103.786	104.190
Ko[a]	145.734	125.244	109.416	102.880	

[a]Reference [39].

Table 5 Nondimensional excitation frequency of a square simply supported stiffened plate under uniaxial compression

	Present (8×8)		Srivastava et al.[a]	
β	$\bar{\Omega}^l$	$\bar{\Omega}^u$	$\bar{\Omega}^l$	$\bar{\Omega}^u$
0	56.49	56.49	56.37	56.37
0.2	52.85	59.91	52.83	59.76
0.4	48.94	63.15	48.89	62.99
0.6	44.68	66.22	44.56	66.06
0.8	39.96	69.16	39.92	68.89
1	34.61	71.98		
1.2	28.26	74.68		

[l]*denotes lower boundary of the instability zone and* [u] *denotes upper boundary of the instability zone.*
[a]Reference [28].

spherical shell panel with low radius of curvature reduces drastically when there is a cutout in the panel. In parabolic hyperboloid shell panel for all four types (Types (i)–(iv)), the frequency decreases gradually with the increase in radius of curvature.

From Table 7, it is observed that the frequency of vibration of cylindrical shell panel (un_c) with low radius of curvature reduces gradually with high rate up to cutout ratio (ca/a) of 0.6 and it increases slightly at $ca/a=0.8$. When an opening is made in this panel and four stiffeners (size=cutout size) are attached to the four sides of the opening, then the free vibration behavior of the panel (st_c) changes. The frequency is more at $ca/a=0.2$ than at $ca/a=0.0$. The frequency reduces up to $ca/a=0.6$ and then again increases. At a higher radius of curvature the natural frequency decreases at $ca/a=0.2$ and after that it gradually increases. The spherical and parabolic hyperboloid shell panels behave the same as the cylindrical shell panel at a higher radius of curvature (radius of curvature to span ratio=10.0). The plate also behaves similarly. In the spherical panel at a low radius of curvature (radius of curvature to span ratio=0.525) the frequency of vibration reduces gradually with a high rate up to cutout ratio (ca/a) of 0.6 and it increases slightly at $ca/a=0.8$, but at a radius of curvature to span ratio=1.0, the frequency decreases gradually with the increase in cutout size both for stiffened and unstiffened cases. In parabolic hyperboloid shell panel, in all cases, the frequency increases gradually with the increase in cutout size.

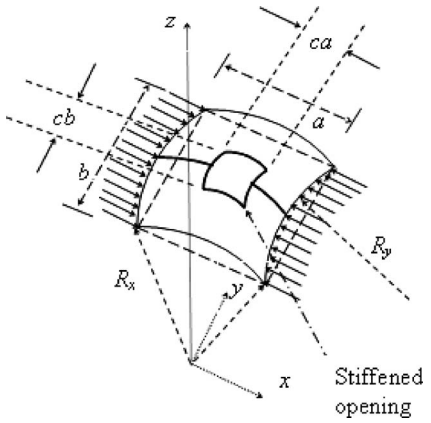


Fig. 8 Perforated stiffened doubly curved panel with stiffened central opening under uniaxial compression

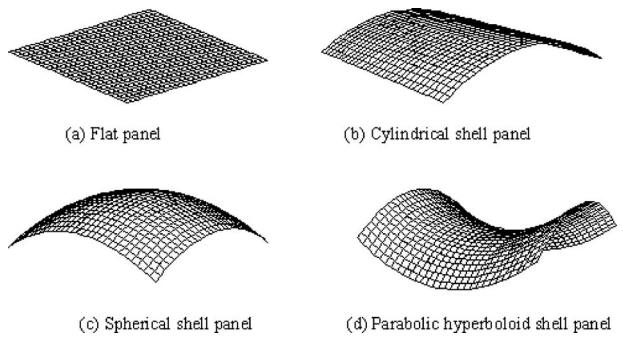


Fig. 9 Geometry of different shell panels

3.3.2 Static Stability (Buckling). The static stability (buckling) analyses of the panels are carried out similar to the free vibration analysis presented above.

The nondimensional buckling load of the panels for different radius of curvature to span ratios is presented in Table 8 and for different cutout sizes are presented in Table 9. It is observed from Table 8 that the buckling load of cylindrical shell panel for all four types (Types (i)–(iv)) gradually decreases with the increase in radius of curvature. The buckling load of the cylindrical panel with cutout is less than that of the panel without cutout. At a lower radius of curvature the decrease rate is much higher compared with that at a higher radius of curvature. Again the buckling load of the stiffened panel (Type (iii)) is higher than the unstiffened cylindrical shell panel (Type (i)). Comparing the stiffened and unstiffened cylindrical shell panels, the rate of increase in buckling load of the stiffened panel with respect to that of the unstiffened panel increases with the increase in radius of curvature. In case of spherical shell panel, the buckling load increases with the increase in radius of curvature up to some point and then decreases with the increase in radius of curvature for Types (i), (iii), and (iv). It is maximum for a particular value of radius of curvature. In Type (ii) the buckling load gradually decreases with the increase in radius of curvature. At a lower radius of curvature the buckling load of spherical shell panel without cutout is less than that of the cylindrical shell panel. The buckling load of the stiffened shell panel (Types (iii) and (iv)) is always greater than the unstiffened panel (Types (i) and (ii)), respectively. The buckling load of the stiffened spherical panel with stiffened opening (Type (iv)) is greater than the stiffened panel (Type (iii)) at a low radius of curvature to span ratio of 0.525 and 0.55. In parabolic hyperboloid shell the buckling load for all four types (Types (i)–(iv)) gradually decreases with the increase in radius of curvature. The buckling load of the stiffened parabolic hyperboloid shell panel (Types (iii) and (iv)) is always greater than the unstiffened panel (Types (i) and (ii)), respectively. At a low radius of curvature, the buckling load of panel with cutout (Types (ii) and (iv)) is greater than the panel without cutout (Types (i) and (iii)), respectively.

The buckling load of the unstiffened cylindrical (for all radii of curvature) and spherical shell panels (at a low radius of curvature to span ratio of 0.525 and 1.0) decreases with the increase in cutout size. The buckling load of the unstiffened spherical shell panel at a radius of curvature to span ratio of 10.0 is highest at ca/a ratio of 0.2 and then gradually reduces with the increase in cutout size. The buckling load of the same panel without cutout and with cutout of ca/a ratio more than 0.2 is less than the value at ca/a ratio of 0.2. The buckling load of the unstiffened parabolic hyperboloid shell panel at a radius of curvature to span ratio of 0.525 is more at small cutout ($ca/a=0.2$) than that of the panel without cutout and gradually decreases with the increase in cutout size. For stiffened cylindrical panels (at a low radius of curvature to span ratio of 0.525 and 1.0), the buckling load decreases with the increase in cutout size. At a radius of curvature to span ratio of 10.0, the buckling load decreases up to some point with the increase in the cutout size and then increases with the increase in the cutout size. For stiffened spherical panels (at a low radius of curvature to span ratio of 0.525 and 10.0), the buckling load decreases up to some point with the increase with the cutout size and then increases with the increase in the cutout size. At a radius of curvature to span ratio of 1.0, the buckling load decreases gradu-

Table 6 Natural frequencies (nondimensional) of the panels for different radii of curvature to span ratio

Geometry	Panel types	Radius of curvature to span ratio 0.525	0.55	0.6	0.8	1.0	2.0	5.0	10.0	10^6
	i	122.58	116.63	109.81	101.12	84.50	59.04	38.47	25.73	19.73
	ii	72.85	73.62	74.79	76.67	75.78	51.41	37.39	26.13	20.81
	iii	118.34	112.67	106.20	98.23	84.35	59.16	42.15	32.64	29.43
Cylindrical shell panel	iv	119.02	122.90	128.20	95.75	80.84	60.49	38.44	31.07	29.18
	i	415.72	436.28	444.63	387.62	321.96	164.67	68.82	38.46	19.73
	ii	52.17	56.06	60.82	72.26	78.70	85.86	66.58	37.98	20.81
	iii	407.54	426.05	433.99	375.50	305.27	157.42	68.61	42.26	29.43
Spherical shell panel	iv	115.00	131.12	155.01	204.17	216.37	132.65	61.08	38.81	29.18
	i	51.00	45.43	37.45	24.77	21.40	19.53	19.66	19.71	19.73
	ii	60.11	54.67	45.81	29.46	24.30	20.82	20.78	20.80	20.81
Parabolic hyperboloid	iii	75.57	64.29	48.83	29.94	27.78	28.20	28.88	19.14	29.43
shell panel	iv	90.82	79.44	62.23	37.07	31.76	29.04	29.06	29.12	29.18
	i					19.73				
	ii					20.81				
	iii					29.43				
Plate	iv					29.18				

ally with the increase in cutout size. For stiffened parabolic hyperboloid shell panel at a radius of curvature to span ratio of 0.525 and 1.0 the buckling load increases up to some point with the increase in the cutout size and then decreases. But at a radius of curvature to span ratio of 10.0 the behavior is just opposite.

3.3.3 Dynamic Instability. After obtaining the free vibration and static buckling characteristics it is now pertinent to study the dynamic instability of stiffened shell panels with cutout. The results of the dynamic instability analysis of the considered problem are presented in this section with critical discussion. The effects of various parameters such as shell geometry, radius of curvature, and cutout size are taken for the analysis. In the analysis the whole structure is divided with a 10×10 mesh. The static load factor (α) is kept as 0.4 in all cases. The dynamic load factor (β) is varied from 0 to 1.0. The buckling stress of the stiffened spherical shell panel (radius of curvature to span ratio of 0.2625) without cutout is taken as the reference load, so as to plot the instability zones.

3.3.3.1 Effect of radius of curvature. The effect of radius of curvature of the stiffened shell panel on their dynamic instability characteristics is discussed in this section. The cylindrical, spherical, and parabolic hyperboloid stiffened shell panels without cutout (Type (iii), described in Sec. 3.3.1) are chosen for this

Table 7 Natural frequencies (nondimensional) of the panels for different cutout ratios

Geometry	Cutout size (ca/a)	Radius of curvature to span ratio 0.525 un_c	0.525 st_c	1.0 un_c	1.0 st_c	10.0 un_c	10.0 st_c
	0.0	122.58	118.34	84.50	84.35	25.73	32.64
	0.2	109.46	121.42	83.71	83.48	25.23	28.19
	0.4	72.85	119.02	75.78	80.40	26.13	31.07
	0.6	52.82	109.50	57.02	76.86	32.29	39.86
Cylindrical shell panel	0.8	54.71	120.93	73.96	84.91	59.20	62.88
	0.0	415.72	407.54	321.96	305.27	38.46	42.26
	0.2	152.88	306.00	201.19	240.96	37.92	38.26
	0.4	52.17	115.00	78.70	216.37	37.98	38.81
	0.6	45.17	83.99	48.98	107.09	42.25	43.49
Spherical shell panel	0.8	59.01	135.18	46.10	75.17	65.19	61.20
	0.0	51.00	75.57	21.40	27.78	19.71	29.14
	0.2	53.46	78.61	21.62	25.78	19.23	24.37
	0.4	60.11	90.82	24.30	31.76	20.80	29.12
	0.6	67.37	121.50	31.79	43.74	28.35	40.74
Parabolic hyperboloid shell panel	0.8	71.55	173.81	44.55	69.52	56.72	61.20
	0.0		19.73^{un_c}			29.43^{st_c}	
	0.2		19.23^{un_c}			24.47^{st_c}	
	0.4		20.81^{un_c}			29.18^{st_c}	
	0.6		28.41^{un_c}			40.83^{st_c}	
Plate	0.8		57.37^{un_c}			67.27^{st_c}	

$^{un\_c}$ denotes unstiffened panels with opening and $^{st\_c}$ denotes stiffened panel with stiffened opening.

Table 8 Buckling load (nondimensional) of the panels for different radii of curvature to span ratio

Geometry	Panel types	Radius of curvature to span ratio 0.525	0.55	0.6	0.8	1.0	2.0	5.0	10.0	10^6
	i	1657.5	1485.5	1246.6	817.2	631.6	309.1	130.7	77.9	53.6
	ii	126.9	126.1	122.4	105.5	93.6	67.9	57.2	47.1	43.6
	iii	1788.5	1597.3	1361.7	919.0	716.4	355.1	198.7	133.0	125.4
Cylindrical shell panel	iv	770.2	738.1	651.7	413.2	311.6	179.4	127.2	99.6	98.0
	i	794.2	975.8	1108.3	1007.0	851.0	490.0	237.8	137.5	53.6
	ii	256.8	227.6	192.3	152.9	145.9	142.8	131.5	113.9	43.6
	iii	1261.7	1513.5	1630.6	1337.8	1102.7	670.9	407.5	288.6	125.4
Spherical shell panel	iv	1318.7	1534.4	1535.2	1257.8	1023.5	588.1	323.7	179.8	98.0
	i	110.0	99.4	80.8	52.4	48.1	50.8	53.2	53.5	53.6
	ii	114.0	100.4	78.4	46.0	40.7	44.4	44.7	44.0	43.6
	iii	398.3	284.2	150.4	53.3	52.5	77.1	103.0	113.4	125.4
Parabolic hyperboloid shell panel	iv	514.8	424.3	299.6	234.4	312.8	261.0	129.1	109.6	98.0
	i					53.6				
	ii					43.6				
	iii					125.4				
Plate	iv					98.0				

analysis. The dynamic instability regions (DIRs) of cylindrical, spherical, and parabolic hyperboloid panels for different radii of curvature to span ratio are plotted in Figs. 10–12, respectively.

In the stiffened cylindrical (Fig. 10) and parabolic hyperboloid (Fig. 11) shell panels the dynamic instability region shifts to the lower frequency zone with a wider width with the increase in the radius of curvature. It indicates that these structures become dynamically less stable due to the increase in radius of curvature. In the case of stiffened spherical shell panel the instability region (Fig. 12) shifts to the higher frequency zone with increase in radius of curvature up to some value of radius of curvature and then instability region shifts to the lower frequency zone with further increase in radius of curvature.

3.3.3.2 Effect of cutout size. The cylindrical, spherical, and parabolic hyperboloid stiffened shell panels with radius of curvature to span ratio of 0.525, and plate with cutout (ratios of ca/a =0.0, 0.2, 0.4, 0.6, and 0.8) are taken for the analysis.

The DIRs of the cylindrical, spherical, parabolic hyperboloid, and flat panels for different cutout ratios are plotted in Figs. 13–16, respectively. The dynamic instability region of the stiffened cylindrical shell panel shifts to the higher frequency zone when the panel changes from no cutout to cutout ratio of 0.2 (Fig. 13). When the cutout increases further the instability region shifts to the lower frequency zone with a wider width. In the case of the stiffened spherical shell panel the instability region (Fig. 14) shifts

Table 9 Buckling load (nondimensional) of the panels for different cutout ratios

Geometry	Cutout size (ca/a)	Radius of curvature to span ratio 0.525 un_c	0.525 st_c	1.0 un_c	1.0 st_c	10.0 un_c	10.0 st_c
	0.0	1657.54	1788.51	631.66	716.42	77.97	133.08
	0.2	473.53	927.75	292.81	445.25	66.02	98.53
	0.4	126.94	770.26	93.61	311.60	47.15	99.69
	0.6	44.04	301.15	34.00	148.16	36.48	113.02
Cylindrical shell panel	0.8	21.26	109.73	21.79	80.79	36.86	152.55
	0.0	794.20	1261.77	815.04	1102.73	137.51	288.62
	0.2	752.87	1260.48	533.14	1091.52	145.45	153.47
	0.4	256.87	1318.73	145.93	1023.56	113.96	179.80
	0.6	387.25	942.94	77.86	530.32	62.16	153.15
Spherical shell panel	0.8	168.20	1203.50	55.29	266.99	44.309	136.56
	0.0	110.09	398.31	48.16	52.56	53.56	113.44
	0.2	114.91	469.38	43.42	78.96	48.16	63.96
	0.4	114.03	514.88	40.75	312.88	44.00	109.62
	0.6	99.76	447.59	35.76	161.68	42.06	148.73
Parabolic hyperboloid shell panel	0.8	59.17	356.61	23.50	65.01	40.65	136.56
	0.0		53.63^{un_c}				125.45^{st_c}
	0.2		48.26^{un_c}				86.93^{st_c}
	0.4		43.61^{un_c}				98.04^{st_c}
	0.6		41.17^{un_c}				115.19^{st_c}
Plate	0.8		40.49^{un_c}				146.07^{st_c}

$^{un\_c}$ denotes unstiffened panels with opening and $^{st\_c}$ denotes stiffened panel with stiffened opening

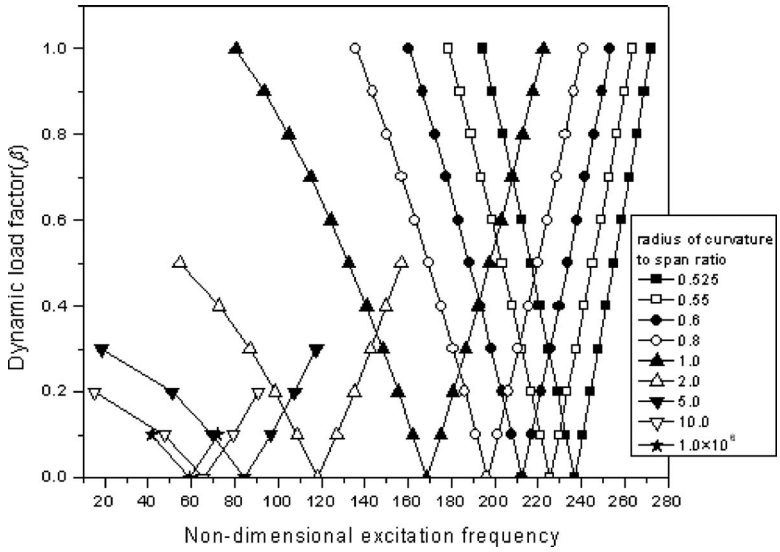


Fig. 10 Dynamic instability regions of stiffened cylindrical shell panel without cutout for various radii of curvature to span ratio

to the lower frequency zone with increase in cutout size. The instability zone shifts to the higher frequency zone when the cutout ratio changes from 0.6 to 0.8. In the stiffened parabolic hyperboloid shell panel with cutout (Fig. 15) the instability region gradually shifts to the higher frequency zone with the increase in cutout size. The dynamic instability region of the stiffened plate (Fig. 16) shifts to the lower frequency zone when the stiffened plate changes from no cutout to cutout ratio of 0.2. When the cutout increases further the instability region shifts to the higher frequency zone. All these behaviors are attributed to the combined effect of stiffness and mass of the structures.

4 Conclusions

The conclusions of the present investigation are as follows.

1. The radius of curvature has a significant effect on the free vibration and buckling behavior of the stiffened shell panels with and without cutout.
2. The natural frequency of the stiffened cylindrical shell panel without cutout is maximum when the radius of curvature is minimum. When there is an opening in the panel this behavior changes.
3. In the stiffened spherical shell panel with/without cutout the free vibration frequency is maximum for an optimal radius of curvature.
4. In the parabolic hyperboloid shell panel the frequency decreases gradually with the increase in radius of curvature.
5. The buckling behavior of the stiffened/unstiffened shell panels with/without cutout is also affected significantly by the radius of curvature of the panel.
6. The presence of cutout needs careful attention for the design of these structures for vibration, buckling, and dynamic stability.
7. In the stiffened cylindrical and parabolic hyperboloid shell panels the dynamic instability region shifts to the lower frequency zone with a wider width with the increase in the

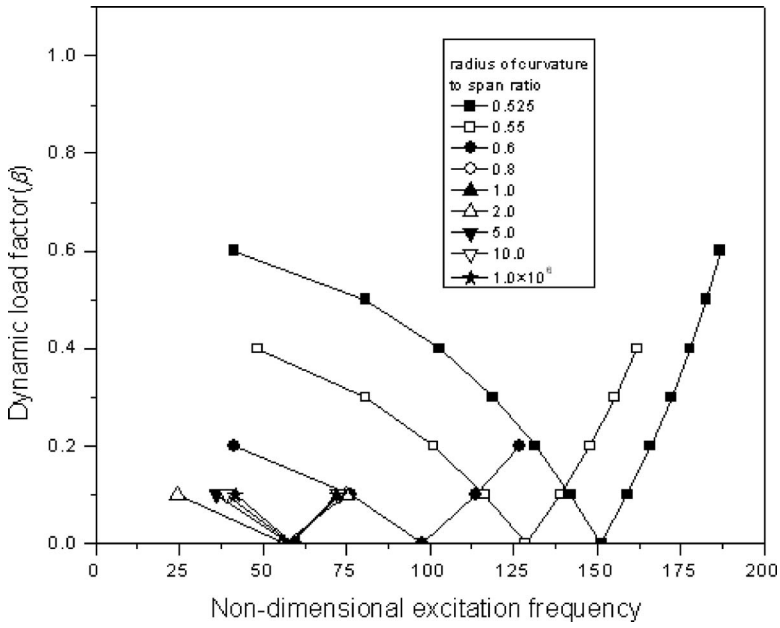


Fig. 11 Same as Fig. 10 but for stiffened parabolic hyperboloid panel

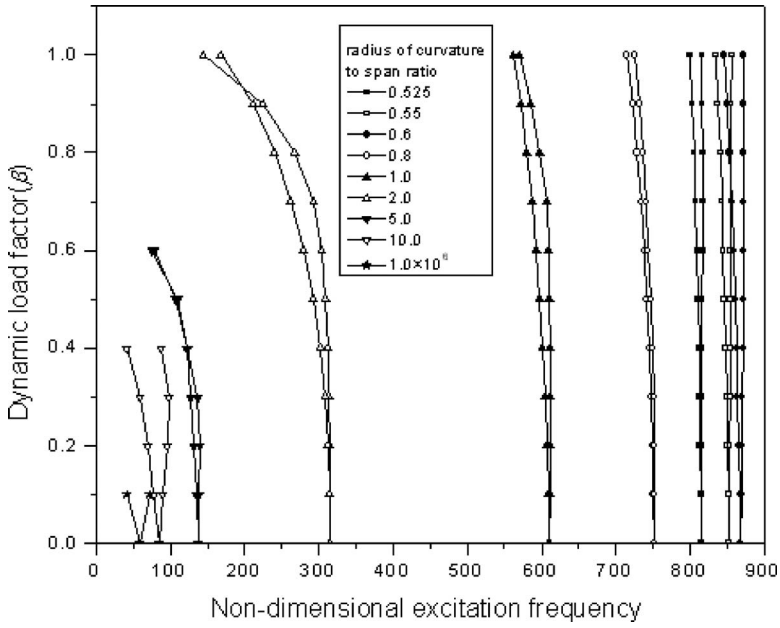


Fig. 12 Same as Fig. 10 but for stiffened spherical shell panel

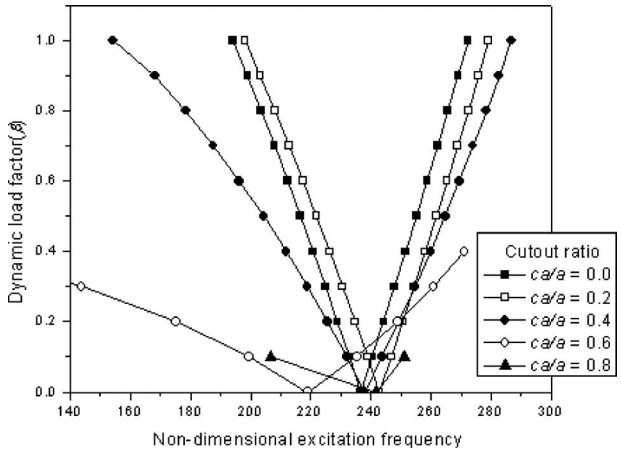


Fig. 13 Dynamic instability regions of stiffened cylindrical shell panel for various cutout ratios

radius of curvature. It indicates that these structures become dynamically less stable due to the increase in radius of curvature.

8. In the case of the stiffened spherical shell panel the instability region (Fig. 12) shifts to the higher frequency zone with increase in radius of curvature up to some value of radius of curvature and then the instability region shifts to the lower frequency zone with further increase in radius of curvature.
9. The dynamic instability region of the stiffened cylindrical shell panel shifts to the higher frequency zone when the panel changes from no cutout to cutout ratio of 0.2. When the cutout increases further the instability region shifts to the lower frequency zone with a wider width. In the case of the stiffened spherical shell panel the instability region shifts to the lower frequency zone with increase in cutout size. The instability zone shifts to the higher frequency zone when the cutout ratio changes from 0.6 to 0.8.
10. In the stiffened parabolic hyperboloid shell panel with cutout the instability region gradually shifts to the higher frequency zone with the increase in cutout size.

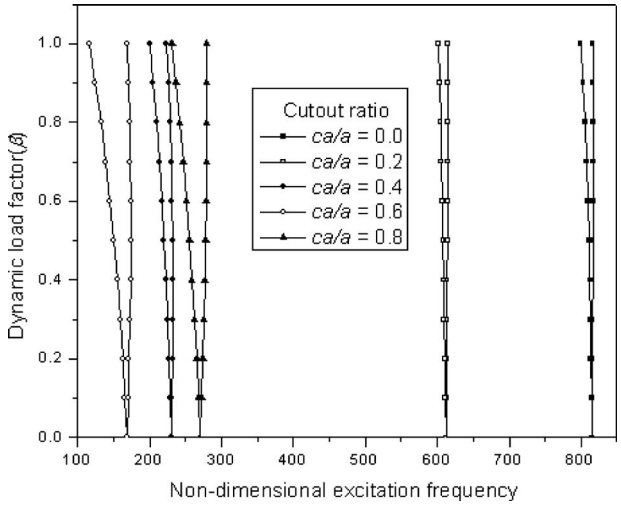


Fig. 14 Same as Fig. 13 but for stiffened spherical shell panel

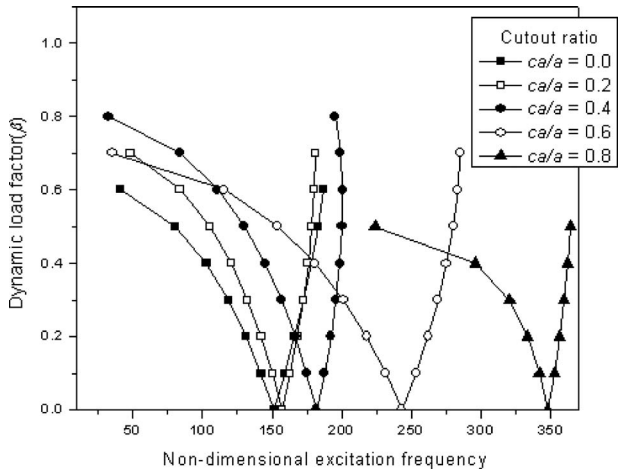


Fig. 15 Same as Fig. 13 but for stiffened parabolic hyperboloid panel

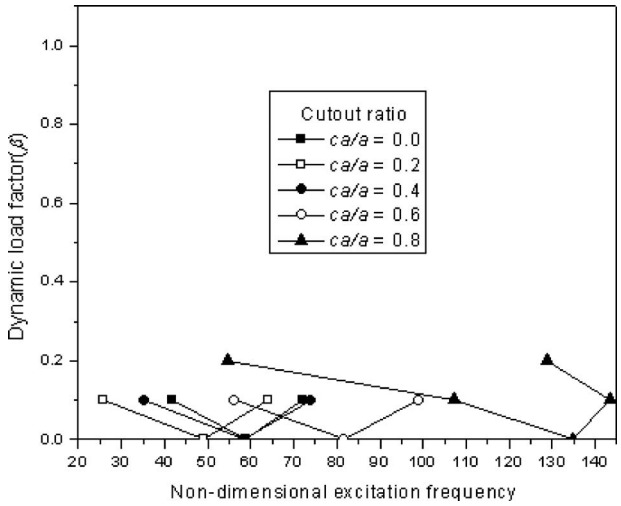


Fig. 16 Same as Fig. 13 but for stiffened plate

11. The dynamic instability region of the stiffened plate shifts to the lower frequency zone when the stiffened plate changes from no cutout to cutout ratio of 0.2. When the cutout increases further the instability region shifts to the higher frequency zone with a higher width.

References

[1] Bolotin, V. V., 1964, *The Dynamic Stability of Elastic Systems*, Holden-day, Inc., San Francisco, California, USA.

[2] Evan-Iwanowski, R. M., 1965, "On the Parametric Response of Structures," Appl. Mech. Rev., **18**(9), pp. 699–702.

[3] Ibrahim, R. A., 1978, "Parametric Vibration Part III. Current Problems (1)," Shock Vib. Dig., **10**(3), pp. 41–57.

[4] Simitses, G. J., 1987, "Instability of Dynamically Loaded Structures," Appl. Mech. Rev., **40**(10), pp. 1403–1408.

[5] Hutt, J. M., and Salam, A. E., 1971, "Dynamic Stability of Plates by Finite Element Method," J. Eng. Mech., **97**, pp. 897–899.

[6] Lam, K. Y., and Ng, T. Y., 1997, "Dynamic Stability of Cylindrical Shells Subjected to Conservative Periodic Axial Loads Using Different Shell Theories," J. Sound Vib., **207**, pp. 497–520.

[7] Lam, K. Y., and Ng, T. Y., 1999, "Parametric Resonance of Cylindrical Shells by Different Shell Theories," AIAA J., **37**, pp. 137–140.

[8] Ng, T. Y., Lam, K. Y., and Reddy, J. N., 1998, "Parametric Resonance of a Rotating Cylindrical Shell Subjected to Periodic Axial Load," J. Sound Vib., **214**, pp. 513–529.

[9] Ng, T. Y., Lam, K. Y., and Reddy, J. N., 1998, "Dynamic Stability of Cross-Ply Laminated Composite Cylindrical Shells," Int. J. Mech. Sci., **40**(8), pp. 805–823.

[10] Ng, T. Y., Hua, L. I., Lam, K. Y., and Loy, C. T., 1999, "Parametric Instability of Conical Shell by Generalized Differential Quadrature Method," Int. J. Numer. Methods Eng., **44**, pp. 819–837.

[11] Sahu, S. K., and Datta, P. K., 2001, "Parametric Instability of Doubly Curved Panels Subjected to Non-Uniform Harmonic Loading," J. Sound Vib., **240**(1), pp. 117–129.

[12] Ng, T. Y., Lam, K. Y., and Reddy, J. N., 1999, "Dynamic Stability of Cylindrical Panels With Transverse Shear Effect," Int. J. Solids Struct., **36**, pp. 3483–3496.

[13] Troitsky, M. S., 1976, *Stiffened Plates Bending Stability and Vibration*, Elsevier Scientific, Amsterdam.

[14] Aksu, G., 1982, "Free Vibration of Stiffened Plates Including the Effect of In Plane Inertia," Trans. ASME, J. Appl. Mech., **49**, pp. 206–212.

[15] Mukherjee, A., and Mukhopadhyay, M., 1988, "Finite Element Free Vibration of Eccentrically Stiffened Plates," Comput. Struct., **30**(6), pp. 1303–1317.

[16] Mukhopadhyay, M., 1989, "Vibration and Stability Analysis of Stiffened Plates by Semi-Analytic Finite Difference Method, Part I: Consideration of Bending Displacements Only," J. Sound Vib., **130**(1), pp. 27–39.

[17] Nayak, A. N., and Bandyopadhyay, J. N., 2002, "On Free Vibration of Stiffened Shallow Shells," J. Sound Vib., **255**(2), pp. 357–382.

[18] Olson, M. D., and Hazell, C. R., 1977, "Vibration Studies of Some Integral Rib-Stiffened Plates," J. Sound Vib., **50**, pp. 43–61.

[19] Samanta, A., and Mukhopadhyay, M., 2004, "Free Vibration of Stiffened Shells by the Finite Element Technique," Eur. J. Mech. A/Solids, **23**, pp. 159–179.

[20] Zeng, H., and Bert, C. W., 2001, "A Differential Quadrature Analysis of Vibration for Rectangular Stiffened Plates," J. Sound Vib., **241**(2), pp. 247–252.

[21] Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw-Hill Kogakusha, Tokyo, Japan.

[22] Thomas, J., and Abbas, B. H. A., 1983, "Vibration Characteristics and Dynamic Stability of Stiffened Plates," Structure, Structural Dynamic and Material Conference, Collection of Technical Papers, Part 2, AIAA, pp. 277–285.

[23] Duffield, R. C., and Willems, N., 1972, "Parametric Resonance of Stiffened Rectangular Plates," ASME J. Appl. Mech., **39**, pp. 217–226.

[24] Merrit, R. G., and Willems, N., 1973, "Parametric Resonance of Skew Stiffened Plates," ASME J. Appl. Mech., **40**, pp. 439–444.

[25] Mermertas, M., and Belek, H. T., 1991, "Dynamic Stability of Radially Stiffened Annular Plates," Comput. Struct., **40**(3), pp. 651–657.

[26] Liao, C. L., and Cheng, C. R., 1994, "Dynamic Stability of Stiffened Laminated Composite Plates and Shells Subjected to In Plane Pulsating Forces," J. Sound Vib., **174**(3), pp. 335–351.

[27] Srivastava, A. K. L., Datta, P. K., and Sheikh, A. H., 2002, "Vibration and Dynamic Stability of Stiffened Plates Subjected to In-Plane Harmonic Edge Loading," Int. J. Struct. Stab. Dyn., **2**(2), pp. 185–206.

[28] Srivastava, A. K. L., Datta, P. K., and Sheikh, A. H., 2003, "Dynamic Instability of Stiffened Plates Subjected to Non-Uniform In-Plane Edge Loading," J. Sound Vib., **262**, pp. 1171–1189.

[29] Sahu, S. K., and Datta, P. K., 2002, "Dynamic Stability of Curved Panels With Cutouts," J. Sound Vib., **251**(4), pp. 683–696.

[30] Srivastava, A. K. L., Datta, P. K., and Sheikh, A. H., 2003, "Dynamic Instability of Stiffened Plates With Cutout Subjected to In-Plane Uniform Edge Loadings," Int. J. Struct. Stab. Dyn., **3**(3), pp. 391–403.

[31] Ahmad, S., 1970, "Analysis of Thick and Thin Shell Structures by Curved Finite Elements," Int. J. Numer. Methods Eng., **2**, pp. 419–451.

[32] Zienkiewicz, O. C., 1977, *The Finite Element Method*, Tata McGraw-Hill, New Delhi.

[33] Bathe, K. J., 1996, *Finite Element Procedure*, Prentice-Hall of India Private Limited, New Delhi.

[34] Ferguson, G. H., and Clark, R. D., 1979, "A Variable Thickness Curved Beam and Shell Stiffener With Shear Deformation," Int. J. Numer. Methods Eng., **14**, pp. 581–592.

[35] Timoshenko, S. P., and Goodier, J. M., 1951, *Theory of Elasticity*, McGraw-Hill Kogakusha, Tokyo, Japan.

[36] Corr, R. B., and Jennings, A., 1976, "A Simultaneous Iteration Algorithm for Symmetric Eigenvalue Problems," Int. J. Numer. Methods Eng., **10**, pp. 647–663.

[37] Long, B. R., 1968, "Vibration of Eccentrically Stiffened Plates," Shock and Vib. Bulletin, **38**, pp. 45–53.

[38] Prusty, B. G., 2001, "Static, Dynamic, Buckling and Failure Analysis of Composite Stiffened Shell Structures, A Finite Element Approach," Ph.D. thesis, Indian Institute of Technology, Kharagpur, India.

[39] Ko, W. L., 1998, "Mechanical and Thermal Buckling of Rectangular Plates with Different Central Cutouts," NASA/TM, Technical Paper No. 206542.

Refined Wave-Based Control Applied to Nonlinear, Bending, and Slewing Flexible Systems

William J. O'Connor
School of Electrical Electronic and Mechanical Engineering,
University College Dublin,
Dublin 4, Ireland
e-mail: william.oconnor@ucd.ie

Alessandro Fumagalli
Dipartimento di Ingegneria Aerospaziale,
Politecnico di Milano,
20156 Milan, Italy
e-mail: alessandro.fumagalli@polimi.it

The problem of controlling the position and vibration of open-chain flexible structures undergoing fast maneuvers is of wide interest. In this work, the general flexible structure is actuated by a single actuator at one end, which, depending on the case of interest, is capable of rotating, translating, or simultaneously translating and rotating the root of the flexible system. The goal is to control the motion of the entire flexible system from rest to rest. This needs a simultaneous synthesis of position control and active vibration damping. A new strategy is presented based on further developments of wave-based control. As before it views the actuator motion as simultaneously launching and absorbing mechanical waves into and out of the system. But a new simple method of resolving the actuator motion into two waves is presented. By measuring the elastic forces exchanged at the interface between the actuator and the rest of the system, a returning displacement wave can be resolved. This is then added to a set, launch wave to determine the actuator motion. Typically the launch wave is set to reach half the target displacement, and the addition of the return wave absorbs the vibration while simultaneously moving the system the second half of the target displacement, neatly achieving the two goals in one controlled motion. To date wave-based control has been applied to lumped, second-order, longitudinally vibrating systems. The refined method avoids a difficulty that previously arose in some contexts, thereby making wave-based control even more generic. It can easily control nonlinear elastic systems, laterally bending systems (in the sense of Euler–Bernoulli beams), and slewing systems where lateral translation and system rotation are strongly coupled. Numerical simulation results are presented for controlled, rest-to-rest maneuvers of representative flexible structures, all controlled using the same (linear) algorithm. The first case is control of a string of rigid bodies interconnected by nonlinear springs. The second problem is the rotational control of a very flexible one-link planar manipulator. Finally, in an extension of the previous system, the actuator both translates and rotates, slewing the flexible system to a target lateral displacement and a target rotation angle simultaneously. The strategy is found to be remarkably effective with many advantages. It seamlessly integrates position and vibration control. It is rapid, robust, energy efficient, and computationally light. It requires little sensing, little knowledge of the flexible system dynamics, and copes well with nonideal actuator behavior. It is generic and easily handles a wide variety of flexible systems. It can get the entire system to stop dead exactly at target with little vibration in transit.
[DOI: 10.1115/1.3086434]

Keywords: flexible structures, robotics, space structures, chain multibody structures, dynamics, active vibration damping, position control, flexible joints, wave-based control

1 Introduction

This paper presents a new method for motion control of flexible, open-chain, multibody systems, such as those arising in robotics and in space structures. Typically an actuator, located at one end of the flexible system, attempts to reposition the entire system and thereby its unconstrained tip at the other end. The proposed method interprets the motion of the structure as two-way mechanical waves, which enter and leave the structure at the actuator. The strategy is based on understanding and measuring this two-way motion at the actuator-system interface, and then controlling it by suitable motion of the actuator. This line of attack leads to control strategies that prove surprisingly simple, robust, and effective. The strategy will be outlined and developed for simple flexible systems composed of a series of lumped masses interconnected by springs. It will be shown that the same controller works very effectively for a larger class of flexible systems, including nonlinear systems and continuous (distributed) systems such as slewing continuum flexible beams (undergoing both rotation and lateral translation).

For an overview of other approaches used to date for this important problem, see, for example, Ref. [1]. For the most part the extensive literature in this area has focused on issues such as deriving and solving system equations [2–5], developing dynamic models [6,7], system identification [8,9], modal analysis and control [10,11], various kinds of optimization [12–15], various forms of command or input shaping [16–20], sliding mode control [21,22], and structure embedded actuators [23]. Wave ideas, different from the one here proposed, have also been applied to control the vibration (but not the net displacement) of flexible systems [24–26].

The cited references are only a few examples of the very extensive literature available on the topic, indicating both the

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 22, 2007; final manuscript received December 30, 2008; published online April 22, 2009. Review conducted by Wei-Chau Xie.

importance and the complexity of the problem. Reference [27] (page 165) observes that

> to date a general solution to the control problem [of flexible structures] has yet to be found. One important reason is that computationally efficient (real-time) mathematical methods do not exist for solving the extremely complex sets of partial differential equations and incorporating the associated boundary conditions that most accurately model flexible structures.

The authors here propose that this long-desired general solution is provided by the wave-based strategy. It has already been applied with great success to rectilinear strings of masses and linear springs [28–34] and to gantry crane systems [35,36]. In this work, the approach is extended to deal with a larger class of flexible systems, all controlled by the same algorithm and configuration.

In Secs. 2–6, wave analysis and control will be developed focusing on the simplest flexible system, a chain of rigid masses interconnected by linear springs. The developed theory (and the related control scheme) is then applied to much more complex systems with absolutely no changes. As confirmed by the tests, the proposed method works very effectively on a large class of flexible systems.

2 Literature Review and Key Ideas of Wave-Based Approach

Classical control strategies, such as proportional-integral-derivative, linear quadratic, Lyapunov-based, and H_∞ methods, frequently give poor results when applied to controlling systems with significant flexibility. At best they lead to control parameters that work for specific maneuvers of specific systems: Any system change or maneuver change then requires a new controller design. One reason is that these general approaches are almost blind to the peculiar dynamic features that characterize flexible systems, especially delays between actuator motion and tip motion. These methods tend to fight against these dynamics rather than exploit them. Because of these difficulties, a series of special techniques has been developed specifically for flexible systems, including posicast, input shaping, sliding mode, input filtering, special trajectory planning, and bang-bang optimal control (derived from Pontryagins minimum principle). However, it can be argued that such special techniques go too far in the opposite direction from the standard control strategies in that they require exact knowledge of the flexible system dynamics and are not very robust to system changes, or to modeling uncertainties, or to real actuator behavior, or (in some cases) to changes in the desired maneuver.

The "blindness" of the classical strategies has been replaced by a need for very detailed vision of the system along with precision in maneuver planning and execution.

The wave-based approach by contrast avoids both problems. It identifies and exploits general features common to a wide variety of flexible systems, yet it does not require detailed knowledge of the flexible system, nor advanced knowledge of the maneuver, nor unrealistically ideal performance by the actuator. Instead of struggling with the flexibility, it collaborates with it, and exploits it, to achieve precise control in a natural way. This leads to many benefits, including robustness, rapid motion, vibration control, and energy efficiency.

Energy and momentum enter and leave the flexible system at the interface between it and the actuator. The motion within the system propagates simultaneously in two directions, from the actuator to the free-end, and back again, albeit in ways that are faltering, complex, dispersing, and highly dynamic. Rest-to-rest motion corresponds to getting the energy and momentum into, and then out of, the system in just the right way to ensure that, when all the energy has been extracted and the entire system has come to rest again, it is, at that instant, exactly at target.

The actuator interacts directly only with the part of the system dynamics to which it is directly connected, that is, the interface.

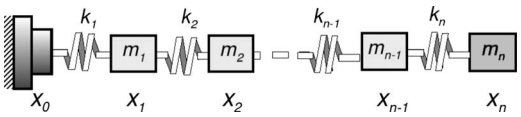


Fig. 1 A simple flexible system: rigid bodies connected by linear springs

Its interaction with the rest of the system, and with the tip, is indirect and delayed, mediated by these local interface dynamics. The two-way component motions, one "outgoing" and one "returning," therefore need to be defined and measured only at the interface, where all the controlling is to be done (and indeed where it can only be done) over time.

It is contended here that the control problem therefore comes down to proper control of the interface between the actuator and the flexible system, which in turn involves (a) understanding and interpreting the interface dynamics, (b) making a few key measurements at this interface, and then (c) managing the two-way energy/momentum flow through the interface by suitable motion of the actuator. This interface control is precisely what is needed, being both necessary and sufficient. To the extent that other control strategies are successful, they must also manage and control the actuator-system interface, even if not as consciously and deliberately as in the present approach.

In other words, only limited aspects of the system dynamics are pertinent to the control problem, and these should be seen from the actuator's perspective, acting on and through the interface dynamics. The actuator is seen as launching a "wave" into the system and responding to a similar returning wave, and these two actions can happen simultaneously. The term wave here includes a net (dc) component as well as a vibratory one.

The key idea of wave-based control is to consider the actuator as launching mechanical waves into the system. The launch wave is set to reach half the target displacement. The return wave is measured and added to this to give the total actuator motion. The adding of the return wave to the actuator input has two crucial effects. First it rapidly dampens all vibration in the real system, becoming steady only when the real system vibrations have ceased. Second, at steady state, the dc value of the return wave will be identical to that of the launch component ($\frac{1}{2}$ reference), and together they will sum to the reference exactly. The returning wave component reveals to the controller all the system information it needs to achieve superb control of both vibration and position in a format that suits its purposes exactly.

3 Wave-Based Modeling

In the previous works on wave-based control, the dynamics of the flexible system were resolved into two "kinematical" components. This approach will be briefly presented here as it forms the basis of the developments carried out in this work.

Consider first perhaps the simplest flexible system: a string of rigid bodies interconnected by linear springs, as depicted in Fig. 1. A rectilinear displacement actuator is connected to one end.

The dynamics of such a system can be described by the superposition of two semi-infinite systems, i.e., two systems that extend forward to infinity, without any terminating boundary. Physically an infinite system implies that physical quantities such as a movement or a force travel along the system in a single direction and never come back.

Since the real system is finite, a single infinite system is inadequate to model its dynamics. Rather, a second infinite system models the returning component of the physical quantity moving through the real system. The dynamics of the real finite system is given by the superposition of the dynamics of the two semi-infinite system, together with suitable boundary conditions.

This modeling approach leads to the scheme depicted in Fig. 2. The blocks G_1, G_2, …, G_n represent the *outgoing* wave transfer

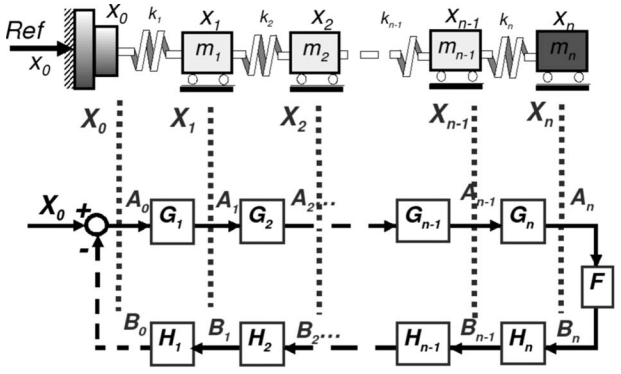


Fig. 2 A simple flexible system modeled as the superposition of two semi-infinite systems using WTFs with $X_i = A_i + B_i$

functions (WTFs), while the blocks H_1, H_2, …, H_n are *returning* WTFs. These transfer functions G_i are chosen to give the value of the physical quantity flowing through the ith station as a function of the value flowing over the $(i-1)$th station, going from left to right (i.e., the outgoing component). In the same way, H_i is assigned to give the value of the physical quantity flowing through the ith station as a function of the value flowing through the $(i+1)$th station, from right to left (i.e., the returning component).

3.1 Position-Position Wave Transfer Function. In previous wave-based modeling and control work, the physical quantity of interest has been the displacement of the masses. For example, with reference to the simple flexible system depicted in Fig. 2, G_i gives the position of the ith mass as a function of the position of $(i-1)$th mass. This choice leads to expressions for the outgoing and returning transfer functions, G and H, which are easily implemented and computationally light. To maintain the simplicity and the effectiveness of the approach, the developments carried out in this work will make use of these *position to position* wave transfer functions.

Consider the case of a uniform system where all the springs are characterized by the same stiffness and all the masses are equal. If the position of the ith body is indicated by $x_i(t)$, the equation of motion of the mass itself is

$$m\ddot{x}_i(t) = k(x_{i-1}(t) - 2x_i(t) + x_{i+1}(t)) \tag{1}$$

where m is the mass of the body and k is the stiffness of the springs. Assuming all initial conditions are zero and letting $\omega_n = \sqrt{k/m}$, the Laplace transformation of the previous equation yields

$$s^2 X_i(s) = \omega_n^2 (X_{i-1}(s) - 2X_i(s) + X_{i+1}(s)) \tag{2}$$

We want to derive the transfer function relating the movement of the ith body as a function of $(i-1)$th one. So we look for an expression like

$$X_i(s) = G_i(s) X_{i-1}(s) \tag{3}$$

If the system is uniform, note that

$$G_{i+1}(s) = G_i(s) = G_{i-1}(s) = G(s) \;\forall\; i \tag{4}$$

so

$$X_{i+1}(s) = G(s) X_i(s) = G^2(s) X_{i-1}(s) \tag{5}$$

Substituting Eqs. (3) and (5) into Eq. (2) yields a quadratic expression for $G(s)$:

$$X_{i-1}(s)[s^2 G(s) - \omega_n^2 (1 - 2G(s) + G^2(s))] = 0 \tag{6}$$

Thus,

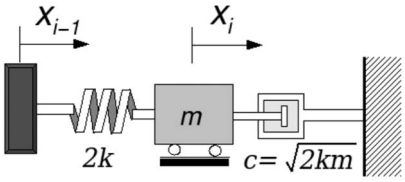


Fig. 3 Analog approximation of the wave transfer function of Eq. (7), with X_{i-1} as the input, X_i the output: cf. Eq. (3)

$$G(s) = \frac{1}{2\omega_n^2}[(s^2 + 2\omega_n^2) \pm s\sqrt{s^2 + 4\omega_n^2}] \tag{7}$$

Only the solution for $G(s)$ with the negative sign before the radical is *causal*, remaining finite with large s and having a phase lag at all frequencies. This solution corresponds to a displacement wave moving from left to right through the system.

$H(s)$, the returning wave transfer function, relates the displacement of the ith body, in the second semi-infinite system, to the displacement of the $(i+1)$th one. Following the same steps taken to derive $G(s)$, the expression for $H(s)$, if the system is uniform, turns out to be the same of the expression for $G(s)$. For more details, see Ref. [28].

This transfer function is not easy to implement in the time domain, involving irrational functions of s in the Laplace domain. As shown in Ref. [37], however, it can be approximated in the time domain by an analog system, composed of a spring, $2k$, a mass, m, and a grounded damper, $c = \sqrt{2km}$, as shown in Fig. 3. Thus, the approximating "position to position" WTF, relating $X_i(s)$ to $X_{i-1}(s)$, turns out to be

$$G(s) = \frac{\omega^2}{s^2 + \omega s + \omega^2} \tag{8}$$

where $\omega = \sqrt{2k/m}$ and $\xi = \sqrt{2km}$.

4 Position-Position Control

The ability to resolve the movement of the flexible system into an outgoing component, traveling from the actuator to the free-end, and a returning motion, coming back from the free-end to the actuator, leads to very effective control schemes.

The idea is to do this resolving at the interface between the actuator and the rest of the flexible system and move the actuator so as to *absorb* the returning motion component while simultaneously launching the outgoing motion. So, from a control point of view, one is interested only in the behavior of the first part of the system, between the actuator and the first body, denoted *station* 0 and *station* 1, respectively. This situation is sketched in Fig. 4.

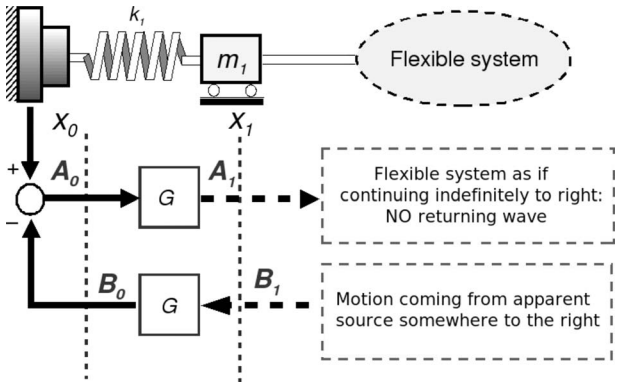


Fig. 4 Position-position control: only the first part of the system is of interest

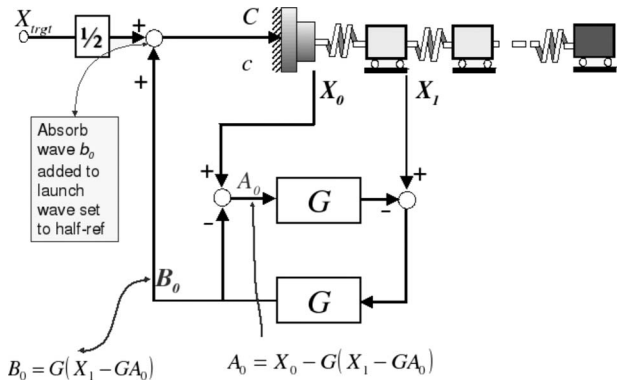


Fig. 5 Position-position control scheme

By measuring the position of station 1 ($x_1(t)$) as well as the position of station 0 ($x_0(t)$) the outgoing and returning wave can be resolved. Defining $a(t)$ and $b(t)$ as the outgoing and returnings, respectively, working in the Laplace domain gives

$$X_0(s) = A_0(s) + B_0(s) \tag{9}$$

$$X_1(s) = A_1(s) + B_1(s) \tag{10}$$

The capital letters represent the Laplace transform of the time domain quantities indicated in small letters. The actual position of the bodies is given by the sum of the outgoing and returning components, according to the wave-based model.

Since

$$A_1(s) = G(s)A_0(s) \tag{11}$$

and

$$B_1(s) = G^{-1}(s)B_0(s) \tag{12}$$

by substituting these into Eq. (10), $B_0(s)$ can be computed as

$$B_0(s) = G(s)(X_1(s) - G(s)A_0(s)) \tag{13}$$

and substituting into Eq. (9),

$$A_0(s) = X_0(s) - G(s)(X_1(s) - G(s)A_0(s)) \tag{14}$$

By measuring $x_0(t)$ and $x_1(t)$, i.e., the time domain values of $X_0(s)$ and $X_1(s)$, the two wave components can be determined from Eqs. (13) and (14). In practice, this can be implemented using the arrangement in Fig. 5, with $G(s)$ approximated by damped second-order systems. The returning component, $b_0(t)$, identified in this way, is then added to the set, launch component, to make up the total actuator motion $x_0(t)$, as shown in Fig. 5.

The control arrangement can be considered as equivalent to opening the return loop of Fig. 2, that is canceling the dashed returning arrow in the block scheme of the figure. Thus, in the wave model, the motion input to the system is considered to travel around the one-way loop before leaving the system, never to return. In practice this is achieved by measuring the return wave and continuously moving the actuator to absorb it, thereby actively damping vibrations while allowing a precise net displacement. Because the steady-state loop gain is unity, the final steady-state value of $b_0(t)$ will equal that of $a_0(t)$, so if the latter is set to half the target displacement, the system will settle at target.

5 Rotating Flexible Systems

The approach described above works very effectively, without modification, for a wide variety of cases, uniform or nonuniform, translational or axially-rotational, long or short, damped (internally) or undamped, stiff or flexible, lumped or distributed, or mixed. It can even be made to cope with nonlinear spring behavior [38].

Regarding rotational systems, a distinction must be made. For axially-rotating systems such as gear-wheel shafts, modeled as rotational inertial masses interconnected by torsional springs rotating about a common axis, the equations of motion are second order in time and space and identical in form to those of rectilinear strings of masses and springs considered above. A quite different case arises, however, in robotics or in slewing of space structures, for example, where the lumped system consists of a string of inertial elements interconnected by revolute joints with torsional springs. The axes of rotation of the elements, and of the entire system, are then perpendicular to the plane of motion, and in general there is both translation and rotation of each inertial element. The equations of motion are now fourth order in space, as seen, for example, in the continuous case in the form of the Euler–Bernoulli beam equation.

When the same wave-based control system was applied to the first case (axially-rotating systems), not surprisingly it was found to work exactly as in the rectilinear displacement case. Then, when applied to the slewing case, wave-based control again appeared to work for pure rotation, that is, when the axis of the rotation of the actuator was fixed in space. When, however, the actuator combined two controlled motions, translation and rotation, a partially concealed problem became clear.

A combined slewing maneuver might specify a net target lateral displacement of the system as well as a net target rotation. A physical example might be a torsional actuator mounted on a trolley, which itself could move on a rail, with both motions controlled, or a controlled antenna array on a satellite whose motion is also controlled. When an adapted version of wave-based control was applied to this case, it was still successful in the sense that it still controlled vibration and brought the system to rest at a new position and angle. The problem was that the final translation and rotation were not predictable in general when both motions happened simultaneously. Typically there were errors in the final steady-state translation displacement and in the rotation angle.

Various ad hoc solutions were found to this problem, but a full discussion will not be presented here. It seemed that the waves associated with translation and with rotation became mixed in a way that was hard to predict so that waves launched by rotation of the actuator could then be absorbed by translation of the actuator, and vice-versa, so that the half-plus-half strategy was compromised by this mixing. A given motion, say, at station 1, which was a combination of translation and rotation, could not be uniquely associated with a certain translation and rotation of station 0. This ambiguity in resolving the wave components led to inconsistencies when the controller attempted to absorb returning waves by simultaneous translation and rotation. Eventually the problem prompted a fairly radical rethink and reformulation of wave-based control in general, which not only solved this particular problem but made wave-based control even more generic.

6 WBC With Cross-Over Wave Transfer Functions

As presented above, and in previous papers, the wave transfer functions described the relationship between the motion (or the force) in adjacent points in the system. As is generally the case in physics, the wave variable was of one kind (e.g., displacement), and the WTFs were dimensionless. The new idea was to allow the WTFs to cross over between variables, with new "cross-over WTFs" (xWTFs) defined as the transfer function relating, for example, displacement at one point, with, for example, force at another point (or vice-versa) when the system is made one-way, that is, quasi-infinite. The new xWTFs are no longer dimensionless.

This new concept, important in itself, in turn allowed a second step, which also has powerful consequences and was here decisive. With cross-over WTFs, the two related variables can now refer to the same location in the system. In particular, for wave-based control, focusing all the wave analysis on a single point eliminates problems associated with the mixing (or ambiguity)

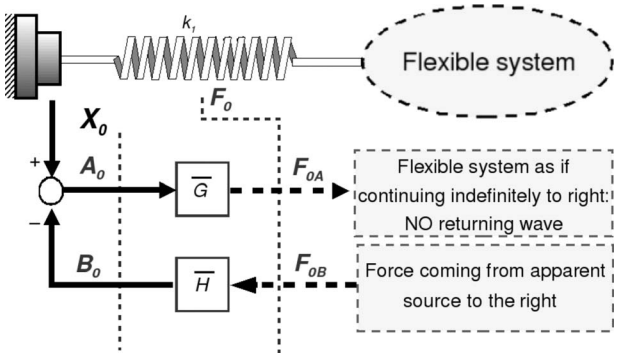


Fig. 6 Wave-based control using cross-over WTFs: only the interface between the actuator and the flexible system is of interest

between translation and rotation when going between points. The chosen location is the interface between the controlling actuator and the attached flexible system.

The idea is to modify the arrangement of Fig. 4 in the way shown in Fig. 6 with two *x*WTFs, the first relating force to displacement and the second displacement to force. This allows the force at the actuator-system interface to be resolved into outgoing and returning components:

$$F_0(s) = F_{0A}(s) + F_{0B}(s) \tag{15}$$

where the subscript 0 refers to station 0, that is, the interface, while the subscripts A and B identify, respectively, the outgoing and returning components.

For a uniform rectilinear system of masses and springs, as in Fig. 1, F_0 is the force in the first spring given by

$$F_0(s) = k[X_1(s) - X_0(s)] = k[(A_1(s) + B_1(s)) - (A_0(s) + B_0(s))]$$
$$= F_{0A} + F_{0B} \tag{16}$$

where $A_i(s)$ and $B_i(s)$ are, respectively, the outgoing and returning components of the displacement of the ith station (Eqs. (9) and (10)) and

$$F_{0A}(s) = k(A_1(s) - A_0(s)) \tag{17}$$

$$F_{0B}(s) = k(B_1(s) - B_0(s)) \tag{18}$$

In this view, the outgoing wave transfer function, $\bar{G}(s)$ in Fig. 6, relates the outgoing elastic force at the interface (F_{0A} in Fig. 6) with the outgoing displacement of station 0 within an assumed semi-infinite system.

Formally, it is

$$F_{0A}(s) = \bar{G}(s)A_0(s) \tag{19}$$

So, using Eqs. (11) and (17),

$$\bar{G}(s)A_0(s) = k(A_1(s) - A_0(s)) \tag{20}$$

$$= k(G(s)A_0 - A_0(s)) \tag{21}$$

giving

$$\bar{G}(s) = k(G(s) - 1) \tag{22}$$

The returning cross-over wave transfer function must relate the returning component of the motion B_0 to the returning component of the elastic force, F_{0B}, both at the actuator interface. Thus

$$B_0(s) = \bar{H}(s)F_{0B}(s) \tag{23}$$

which, using Eqs. (12) and (18), becomes

$$B_0(s) = \bar{H}(s)[k(B_1(s) - B_0(s))] \tag{24}$$

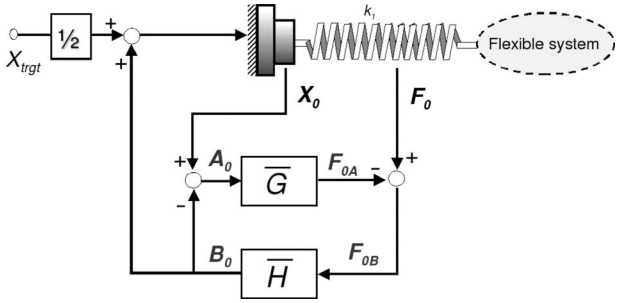


Fig. 7 Wave-based control using cross-over WTFs: only the interface between the actuator and the flexible system is of interest

$$= \bar{H}(s)[k(G(s)^{-1}B_0(s) - B_0(s))] \tag{25}$$

and dividing by $B_0(s)$ yields

$$\bar{H}(s) = \frac{1}{k(G^{-1}(s) - 1)} = \frac{G(s)}{k(1 - G(s))} \tag{26}$$

Thus by measuring both the position and force at the actuator, outgoing and returning displacement and force waves can be identified. As in previous work, the returning component of the motion is added to the set movement of the actuator, thereby absorbing it, damping vibrations, and moving the system by a further steady-state net displacement equal to that of the outgoing component. Thus, again, the actuator motion is set to half the reference input plus the returning wave $b_0(t)$ (that is, $B_0(s)$ in the time domain). The resulting control scheme is depicted in Fig. 7.

7 Control Applications and Testing

The new control scheme using colocated cross-over WTFs has been extensively tested in numerical simulations by applying it to many kinds of flexible system control. In this section, three quite different applications are reported.

The first deals with a nonlinear rectilinear elastic system comprising two masses (the first representing the actuator and the second the tip mass) interconnected by a nonlinear (hardening) spring. The second applies the same control to rotating a very flexible distributed robot arm modeled as a beam. Finally the same distributed one-link arm is simultaneously both translated and rotated from rest to rest by a single actuator. In all cases the control scheme performs very well, with rapid transient response, active vibration damping during motion, short settling time, and zero steady-state error.

7.1 Nonlinear Lumped System. In the first application, the system to be controlled consists of two masses, each of 1 kg, interconnected by a nonlinear spring whose reaction force is $F = k_1(x_1 - x_0) + k_2(x_1 - x_0)^3$, where x_0 and x_1 are the displacements of the actuator and end mass.

The cross-over *x*WTF $\bar{G}$, from motion to force, is modeled by a system as in Fig. 3, with the actuator position, X_{i-1}, as the input, and the force in the spring as the output.

The *x*WTF $\bar{H}$, from position to force, is modeled and implemented as in Fig. 8.

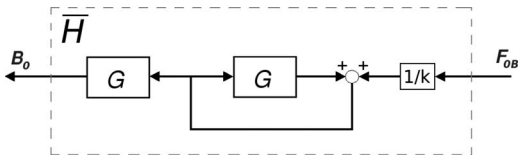


Fig. 8 Block scheme implementing Eqs. (26) and (23), which becomes the *x*WTF $\bar{H}$ in Fig. 7

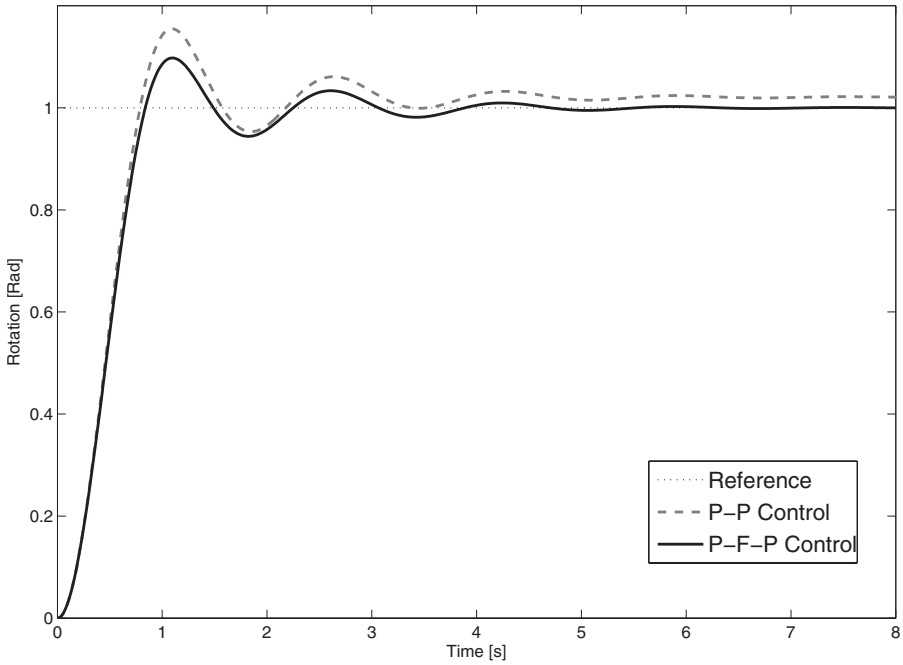


Fig. 9 Tip mass response to a step input using standard and cross-over WTFs

Figure 9 shows the response of the end mass to a unit step reference, using standard wave-based control (WBC) (scheme as in Fig. 5) and the colocated xWTF control (scheme as in Fig. 7). As evident from the picture, the standard scheme fails when applied to the nonlinear system. While it still effectively dampens the vibrations, it leaves the system with a small steady-state position error. This effect was considered in Ref. [38] where two other solutions were proposed to modify the control scheme to deal with the nonlinearity. By contrast, the new scheme avoids this kind of error without any modification. The old scheme requires measuring waves going from the actuator to the mass and back again, inappropriately assuming linear superposition within the nonlinear spring. In the new scheme the entire wave measurement and resolving takes place at the actuator end, by the colocation there of the position and force measurements. Thus the nonlinear spring dynamics are avoided in measuring the waves, and superposition within the spring is not invoked. Provided the integral of the spring force returns to zero, which is must if the initial and final momenta are zero, the final steady-state position error will be zero.

7.2 One-Link Planar Manipulator. The second test case is control of a planar one-link flexible manipulator. A rotational actuator, placed at the root of the flexible beam, is used to move the system in the horizontal plane. The target is a rest-to-rest rotation of the entire flexible system.

The flexible system is now modeled as a continuous (distributed) system rather than a lumped one. A powerful multibody dynamics simulation package, MBDYN, was used, which accurately models beam bending effects using the finite volume method. It also captures nonlinear behavior associated with large deformation, whether geometrical or material (elastic or elastoplastic behavior). See Ref. [39] for further details on the beam model and Ref. [40] (and reference therein) for an overview of the software.

The beam is assumed to be of aluminum (mass density $\rho = 2700\ \text{kg/m}^3$, Young modulus $E = 70{,}000$ MPa, and Poisson ratio $\nu = 0.33$). The geometric characteristics are $t = 3$ mm, $h = 0.1$ m, and $L = 1$ m, where t is the thickness of the beam, h the height, and L the length of the beam itself.

The same control configuration as in the lumped case was used, with exactly the same way of modeling the xWTFs, that is, using a lumped second-order system, with dynamics described by Eq. (8). For the control scheme implementation, only the elastic force at the interface between the actuator and the rest of the system is required. So, only the elastic force at the root of the beam needs to be measured, for example, using a strain gauge, as depicted in Fig. 10.

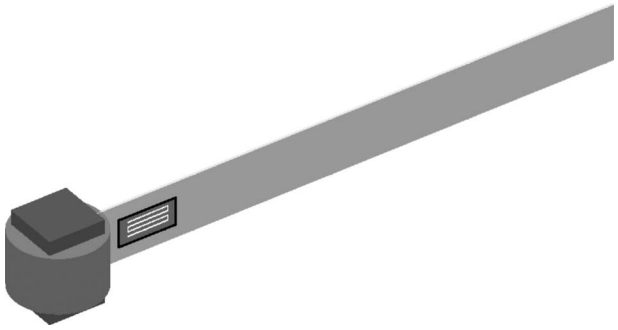


Fig. 10 The one-link planar manipulator with the required strain gauge

Since the actuator can apply only rotations, the outgoing and returning waves are defined in terms of rotations, θ. F_0 in Eq. (16) is a moment rather than a force, and, in this particular case, will be the resultant elastic moment acting on the actuator.[1] The xWTF $\bar{G}(s)$ in Eq. (19) transforms from angular rotation to moment while the xWTF $\bar{H}(s)$ in Eq. (23) goes from moment to angle.

Despite the more complex nature of the system, the same simple transfer functions of Eqs. (22), (26), and (8) are found to give very good performance. The values of the parameters of Eq. (8) are not critical: for any reasonable choice the steady-state error will be zero and stability is guaranteed. As will be seen, by adjusting these parameters the performance of the control system can be fine-tuned further to achieve a classical trade-off between improved rise-time, overshoot, and settling time.

[1] If the strain gauge is not placed exactly at the root of the beam, the moment of interface forces about the actuator axis will have additional contributions from the shear force and even from the axial force, multiplied by their respective moment arms.

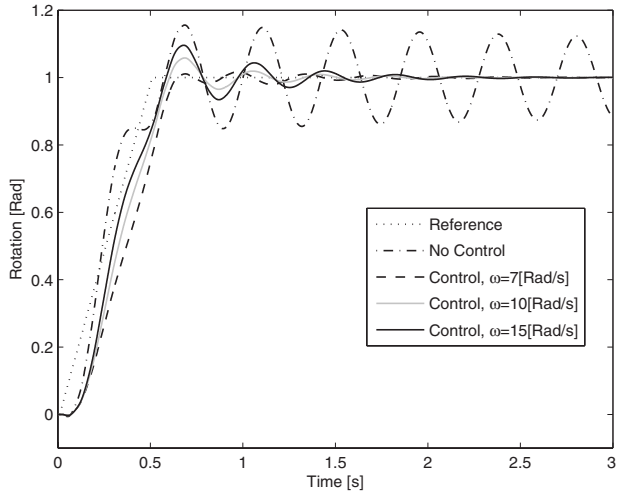


Fig. 11 Beam tip rotations for a very fast rest-to-rest maneuver. The reference input is to rotate the system by 1 rad in 0.5 s, following a ramp. Responses are shown for no control and with *G*(*s*) (Eq. (8)) in the controller tuned to different natural frequencies.

Figure 11 shows the rest-to-rest rotation of the system tip with and without control. The system is rotated by 1 rad in 0.5 s, with a ramp input. Four curves are shown. One is the uncontrolled case in which the vibrations are very large. The other three curves are the response of the system under wave-based control using *x*WTFs.

By varying the natural frequency ω of the transfer function $G(s)$ (Eq. (8)), the response can be made faster (by increasing ω) but less damped, or slower (by decreasing ω) and more damped. Eventually, by an appropriate choice of ω, the overshoot can be completely avoided.

Figure 12 shows the launching and the returning waves a_0 and b_0, whose sum gives the actuator movement θ_0. The launch wave is set to be half the reference input, while the returning wave is resolved by the control scheme and added to the launch wave to give the movement of the actuator. Again, it is worth noting that, at steady state, the two waves assume the same dc value of half

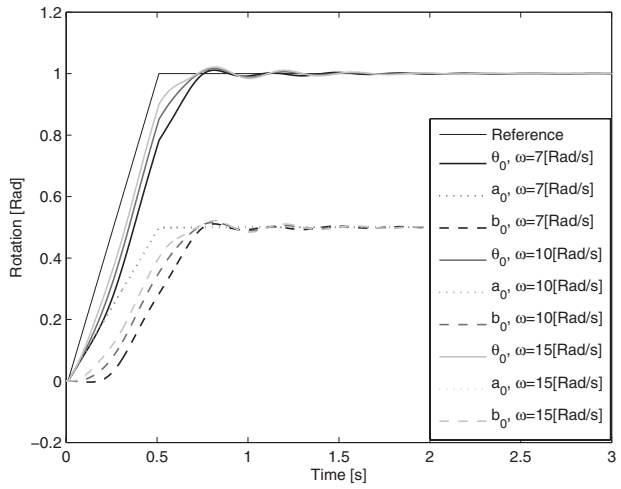


Fig. 12 Actuator rotations for a very fast rest-to-rest maneuver. The total actuator motion is the sum of the launch wave a_0 and the measured returning wave b_0. The curves refer to cases with no control and with controls tuned at different natural frequencies for *G*(*s*) (Eq. (8)).

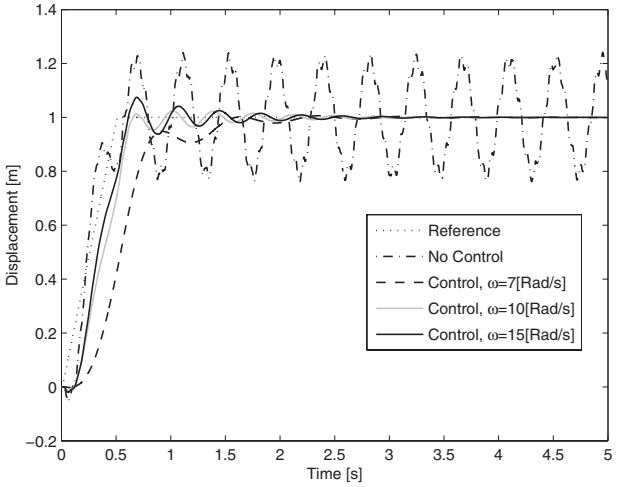


Fig. 13 Tip translation absorbing translational (shear) waves

the target. The final angle accuracy is limited only by that of the actuator angle sensor.

If the launch wave a_0 is a ramp, or constant velocity, then during the transit the returning wave also quickly settles to a ramp of equal velocity. After a short while the system then moves at a constant angular velocity throughout, as if rigid, with vanishing vibration. The system energy is then all kinetic, the entire strain energy having been released. For most applications this rigidlike behavior for the midtransit phase would be considered ideal. For short maneuver times, there may not be time for this effect to be noticed, but even without it, the system will still arrive exactly at target and quickly settle there.

7.3 Slewing One-Link Manipulator. Now consider the case of the same distributed flexible arm but controlled by an actuator capable of both rotation and translation. The translation component is taken to be in the *y* direction, perpendicular to both the rotation axis of the actuator and the initial direction of the axis of the beam in Fig. 10.

7.3.1 Pure Lateral Translation. First consider the case of pure lateral translation, with an actuator capable of only lateral motion without rotation. An example would be a flexible system cantilevered from a trolley on rails, with the trolley moving along the rails at right angles to the flexible system. The flexible system is clamped to the trolley, the motion of which constitutes the actuating input.

Because the actuator can supply only lateral motion, both the launching and absorbing must be achieved by this lateral motion. When a wave-based strategy similar to that for pure rotation is applied, it is found that the system does rapidly move to the correct new position, again with an elegantly controlled, almost vibrationless motion.

When, as here, the actuator is moved laterally at a fixed angle, it not only launches a translation wave but also implicitly launches a rotational (bending) wave, needed to keep the interface angle constant. The actuator not only pushes the system sideways but also supplies an implicit "twist" necessary "just to keep it straight:" in other words, applying a shear action necessarily implies also applying a bending moment. This double launch action propagates through the system and must then be absorbed in return to the actuator. Although the actuator can only translate, it is still able to affect the "double" absorption. The control scheme uses the actuator's lateral or *y*-position and the shear, elastic force in the same *y* direction, both measured at the actuator-system interface, to resolve the movement into outgoing and returning component motions.

Figure 13 shows the tip response, while Fig. 14 shows the actuator motion with its two component motions, a_0 and b_0. Again,

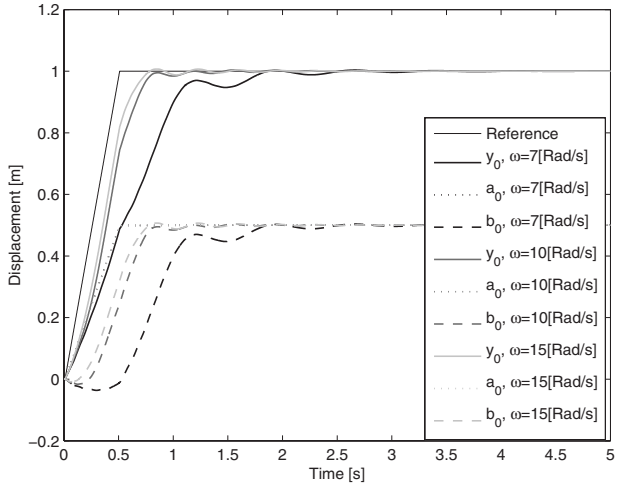


Fig. 14 Actuator translation absorbing translational (shear) waves

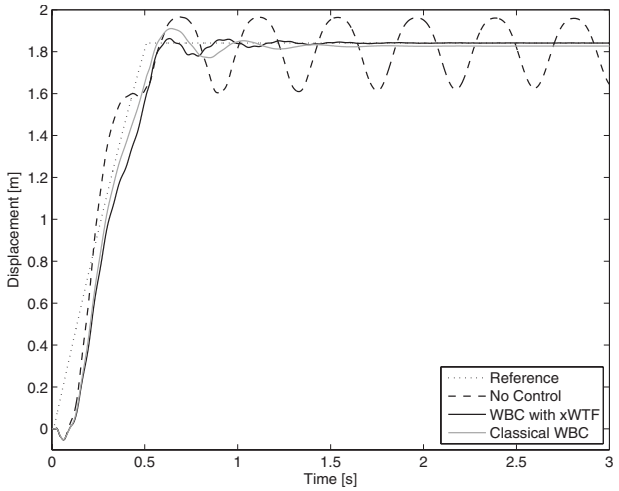


Fig. 15 *y* movement of the system tip under combined translation and rotation controlled by absorbing both translational and rotational waves simultaneously

by changing the natural frequency of the model of the transfer function $G(s)$, the transient of the response can be fine-tuned within an already well-behaved response.

7.4 Combined Angular Rotation and Lateral Displacement. When the actuator both translates and rotates, the launching again involves both kinds of motion, whether explicitly or implicitly. The absorbing can also be done with translation alone, or with rotation alone, or with both acting simultaneously. Each case was investigated with the traditional wave-based control (involving measurements at the actuator and at a second point within the flexible system). Again in all cases the system settled, but at a steady-state position that, in general, was hard to predict. The final absorbed components of each motion typically were no longer half the launched components, indicating a mixing between the translation and rotation component waves. However, when the modified wave-based control was implemented using *x*WTFs and making all the measurements at the interface between the actuator and the system, the control system was found to work perfectly. The desired lateral translation and the desired rotation could be treated as if they were independent motions, each with their own control systems as before (as in Secs. 7.2 and 7.3.1), even when both were active simultaneously, using the same simple control technique for each. Thus the *y*-translation control system uses the interface translation and interface (elastic) force in the same *y*-direction to do the launch and absorbing, while the rotation control uses the interface angle and bending moment. Because of the net rotation, in general the orientation of the beam axis will not be perpendicular to the *y*-displacement direction, so the elastic force in the *y*-direction will be a combination of the relevant components of the shear and axial forces acting at the interface at every instant.

Figures 15 and 16 show the response for a translation of the actuator of 1 m combined with a rotation of 1 rad. Both standard WBC and colocated *x*WTF WBC have been used. While both quickly dampen system vibrations, the standard WBC leaves the system in a wrong final position, whereas WBC with *x*WTFs has zero steady-state errors.

One way to view the improvement is as follows. Classical WBC involved making measurements at two physically separated points, say, stations 0 and 1. But for flexing vibrating systems that both translate and rotate, the interconnecting two-way motion is not uniquely defined. A given rotation and translation combination at station 1 at any instant can arise due to different possible translation and rotation combinations at station 0. So the outgoing and returning waves have an inherent ambiguity. On the other hand, by using *x*WTFs, all the measurements and wave resolution can be done at a single point, which is also the point at which energy and momentum enter and leave the flexible system, and where the control is done. The addition of the returning wave motion guarantees energy absorption, while the conservation of momentum (linear and angular) ensures that the net associated absorbing displacements will equal the launch displacements. Thus the system must settle and do so exactly at target.

8 Final Comments

In all the results presented in this paper, there is scope for some further improvement, in particular, in the form of the launch wave chosen by the controller before the addition of the return wave. For example, the ideas outlined in Ref. [32], called wave-echo control, can be used to modify the launch wave in real time to cause the tip to stop dead on arrival at target, in a time-reversed, mirror-image of the tip's initial motion. Also, the way of modeling the WTF $G(s)$ could be further developed to some advantage. But these further refinements are not considered here.

Although the actuator has here been assumed to be ideal, it does not have to be. Because the wave measurements are taken at the actuator-system interface, they come after the actuator dynamics. Thus the control strategy based on these measurements is not

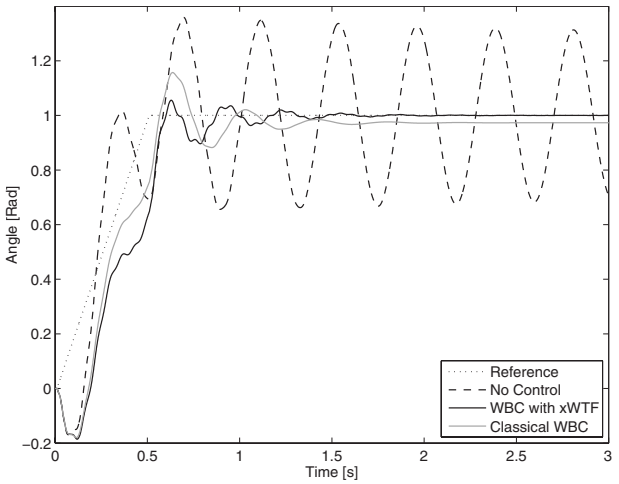


Fig. 16 Rotation of the system tip under combined translation and rotation controlled by absorbing both translational and rotational waves simultaneously

strongly inhibited when the assumption of ideal actuator performance is relaxed. Again, for brevity, this issue is not elaborated upon here.

In developing the analysis and applying the control strategy, at no point was it necessary to make the "small angle" assumption frequently invoked in analyzing laterally vibrating systems. In fact, the translations and rotations can be arbitrarily large, whether in the transient motions or in the steady-state displacements, without loss of accuracy and performance.

As has been seen, a control strategy that achieves rest-to-rest maneuvers of a flexible system using a single actuator must insert energy and linear and angular momenta into the system and then extract them again in such a way that when they have all returned to zero, the system is exactly at target. If so, any other strategy must be doing just this (perhaps even unconsciously). Wave-based control claims to do this consciously and deliberately.

9 Conclusion

New wave-based control strategies have been presented that can get a generic flexible system to move from rest in one position to rest exactly at a new target position, very rapidly, with little vibration in transit, and short settling times. The strategy has been extensively tested on computer simulated systems, involving both lumped and distributed flexible systems, both linear and nonlinear.

The computational load is light, so real-time control is very feasible. Few measurements are needed and these are taken at the actuator. There are no restrictions to small displacements and the approach easily copes with arbitrarily large, linear or angular, transient or steady-state motions.

Most of the problems associated with other approaches seem to have been avoided. For example, no accurate system model is needed, nor system identification, nor modal analysis (whether with exact, or assumed modes, or truncated modes). There is no need for precise switching times, as in bang-bang control, or for ideal actuator behavior. There is no separation of the motion into rigid-body and flexible modes, nor subsequent concerns about coupling between these two artificially separated motions. Mode spillover does not arise. The chatter issues of sliding mode control are also avoided, and actuator commands are smooth. Neither is there any conflict between position control and active vibration damping, these two functions having been seamlessly integrated into one motion.

Any of the many attractive features of the new control strategy could be considered a significant point of merit. To find them all in one method seems exceptional.

References

[1] Meirovitch, L., 1990, *Dynamics and Control of Structures*, Wiley, New York, Chap. 9.

[2] Bhat, S. P., and Miu, D. K., 1990, "Precise Point-to-Point Positioning Control of Flexible Structures," ASME J. Dyn. Syst., Meas., Control, **112**(4), pp. 667–674.

[3] Bhat, S. P., and Miu, D. K., 2000, "Discrete Time Point-to-Point Control of Flexible Structures," Proceedings of the American Control Conference, Jun., pp. 2433–2437.

[4] Jayasuriya, S., and Choura, S., 1991, "On the Finite Settling Time and Residual Vibration Control of Flexible Structures," J. Sound Vib., **148**(1), pp. 117–136.

[5] Vincent, T. L., Joshi, S. P., and Yeong, C. L., 1989, "Positioning and Active Damping of Spring-Mass Systems," ASME J. Dyn. Syst., Meas., Control, **111**, pp. 592–599.

[6] Kelkar, A. G., 1998, "On the Dynamics and Control of Flexible Multi-Body Non-Linear Space Systems," Transactions of the IMechE International Conference on Multi-Body Dynamics, London, UK, Dec., pp. 343–352.

[7] Feliu, J. J., Feliu, V., and Cerrada, C., 1999, "Load Adaptive Control of Single-Link Flexible Arms Based on a New Modelling Technique," IEEE Trans. Robot. Autom., **15**, pp. 793–804.

[8] Abduljabbar, Z., ElMadany, M. M., and Al-Dokhiel, H. D., 1993, "Controller Design of a One-Link Flexible Robot Arm," Comput. Struct., **49**(1), pp. 117–126.

[9] Lee, T. H., Ge, S. S., and Wang, Z. P., 2001, "Adaptive Robust Control Design for Multi-Link Flexible Robots," Mechatronics, **11**(8), pp. 951–967.

[10] Yang, H., Krishnan, H., and Ang, M. H., Jr., 1998, "A Modal Feedback Control Law for Vibration Control of Multi-Link Flexible Robots," *Proceedings of the American Control Conference*, Philadelphia, PA, June, pp. 1821–1832.

[11] Meirovitch, L., and Shenar, J., 1984, "Control of Large Flexible Spacecraft by the Independent Modal-Space Control Method," Technical Report.

[12] Meirovitch, L., and Kwak, M. K., 1990, "Control of Spacecraft With Multi-Targeted Flexible Antennas," J. Astronaut. Sci., **38**, pp. 189–99.

[13] Pao, L. Y., and La-orpacharapan, C., 2004, "Shaped Time-Optimal Feedback Controllers for Flexible Structures," ASME J. Dyn. Syst., Meas., Control, **126**, pp. 173–186.

[14] Pao, L. Y., 1996, "Minimum-Time Control Characteristics of Flexible Structures," J. Guid. Control Dyn., **19**, pp. 123–129.

[15] Tanaka, N., and Kikushima, Y., 1999, "Optimal Vibration Feedback Control of an Euler-Bernoulli Beam: Toward Realization of the Active Sink Method," ASME J. Vibr. Acoust., **121**, pp. 174–182.

[16] Singhose, W. E., Singer, N. C., and Seering, W. P., 1994, "Design and Implementation of Time-Optimal Negative Input Shapes," *Proceedings of the 1994 International Mechanical Engineering Congress and Exposition*, ASME Dynamic Systems and Control Division, Chicago, IL, pp. 151–157.

[17] Cutforth, C. F., and Pao, L. Y., 2004, "Adaptive Input Shaping for Maneuvering Flexible Structures," Automatica, **40**(4), pp. 685–693.

[18] Mimmi, G., and Pennacchi, P., 2001, "Pre-Shaping Motion Input for a Rotating Flexible Link," Int. J. Solids Struct., **38**, pp. 2009–2023.

[19] Mohamed, Z., and Tokhi, M. O., 2004, "Command Shaping Techniques for Vibration Control of a Flexible Robot Manipulator," Mechatronics, **14**(1), pp. 69–90.

[20] Yamada, I., and Nakagawa, M., 1985, "Reduction of Residual Vibration in Position Control Mechanisms," ASME J. Vib., Acoust., Stress, Reliab. Des., **107**, pp. 47–52.

[21] Chen, X., and Fukuda, T., 2001, "Robust Sliding-Mode Tip Position Control for Flexible Arms," IEEE Trans. Ind. Electron., **48**, pp. 1048–1056.

[22] Chen, X., Zhai, G., and Kano, H., 2004, "Robust Sliding Mode Control for Flexible Arms," Proceedings of the Modelling, Identification, and Control, Grindelwald, Switzerland.

[23] Song, G., and Agrawal, B. N., 2001, "Vibration Suppression of Flexible Spacecraft During Attitude Control," Acta Astronaut., **49**(2), pp. 73–83.

[24] Mizutani, K., Yatomi, C., and Inoue, K., 1996, "Active Vibration Control for Flexible Structures Using a Wave-Absorbing Control Method," JSME Int. J., Ser. C, **39**(2), pp. 188–193.

[25] Saigo, M., Tanaka, N., and Tani, K., 1998, "An Approach to Vibration Control of Multiple-Pendulum System by Wave Absorption," ASME J. Vibr. Acoust., **120**, pp. 524–533.

[26] Saigo, M., Tanaka, N., and Nam, D. H., 2004, "Torsional Vibration Suppression by Wave-Absorption Control With Imaginary System," J. Sound Vib., **270**, pp. 657–672.

[27] Robinett, R. D., III, Dohrmann, C. R., Eisler, G. R., Feddema, J. T., Parker, G. G., Wilson, D. G., and Stokes, D., 2002, *Flexible Robot Dynamics and Controls*, Plenum, New York.

[28] O'Connor, W. J., and Lang, D., 1998, "Position Control of Flexible Robot Arms Using Mechanical Waves," ASME J. Dyn. Syst., Meas., Control, **120**(3), pp. 334–339.

[29] O'Connor, W. J., and Hu, C., 2002, "A Simple, Effective Position Control Strategy for Flexible Systems," *International Federation of Automatic Control, Proceedings of the Second IFAC Conference on Mechatronic Systems*, Berkeley, CA, pp. 153–158.

[30] O'Connor, W. J., 2005, "Excellent Control of Flexible Systems via Control of the Actuator-System Interface," *Proceedings of the CDC-ECC'05, 44th IEEE Conference on Decision and Control and European Control Conference ECC* Seville, Spain, Dec. 12–15.

[31] O'Connor, W. J., 2005, "Wave-Based Modelling and Control of Lumped, Multibody Flexible Systems," *Proceedings of the International Conference on Advances in Computational Multibody Dynamics, ECCOMAS Thematic Conference Multibody Dynamics*, Madrid, Spain, Jun. 21–24.

[32] O'Connor, W. J., 2006, "Wave-Echo Control of Lumped Flexible Systems," J. Sound Vib., **298**, pp. 1001–1018.

[33] O'Connor, W. J., 2008, "Wave-Based Analysis and Control of Lumped Flexible System Dynamics," *Informatics in Control Automation and Robotics*, (Lecture Notes in Electrical Engineering), Vol. 15, Springer, Berlin, pp. 25–34.

[34] O'Connor, W. J., 2007, "Wave-Based Analysis and Control of Lump-Modeled Flexible Robots," IEEE Trans. Robot., **23**, pp. 342–352.

[35] O'Connor, W. J., and Hu, C., 2004, "A Simple, Effective Position Control Strategy for Flexible Systems," *Proceedings of the ACC04*, Boston, MA, pp. 153–158.

[36] O'Connor, W. J., 2003, "A Gantry Crane Problem Solved," ASME J. Dyn. Syst., Meas., Control, **125**(4), pp. 569–576.

[37] McKewon, D. J., and O'Connor, W. J., 2007, "Wave-Based Control—Implementation and Comparison," *Proceedings of the IEEE American Control Conference*, Jul., pp. 4209–4214.

[38] O'Connor, W. J., Ramos, F., McKewon, D. J., and Feliu, V., 2008, "Wave-Based Control of Nonlinear Flexible Mechanical Systems," Nonlinear Dyn. (published online).

[39] Ghiringhelli, G. L., Masarati, P., and Mantegazza, P., 2000, "Multibody Implementation of Finite Volume c(0) Beams," AIAA J., **1**, pp. 131–138.

[40] Masarati, P., and Mantegazza, P., 2000, "Mbdyn-Multibody Dynamics," http://www.aero.polimi.it/~mbdyn/.

Gottfried Spelsberg-Korspeter
e-mail: speko@dyn.tu-darmstadt.de

Daniel Hochlenert
e-mail: hochlenert@dyn.tu-darmstadt.de

Oleg N. Kirillov[1]
e-mail: kirillov@dyn.tu-darmstadt.de

Peter Hagedorn
e-mail: peter.hagedorn@dyn.tu-darmstadt.de

Department of Mechanical Engineering,
Dynamics and Vibrations Group,
Technische Universität Darmstadt,
Hochschulstrasse 1,
64289 Darmstadt, Germany

In- and Out-of-Plane Vibrations of a Rotating Plate With Frictional Contact: Investigations on Squeal Phenomena

Rotating plates are used as a main component in various applications. Their vibrations are mainly unwanted and interfere with the functioning of the complete system. The present paper investigates the coupling of disk (in-plane) and plate (out-of-plane) vibrations of a rotating annular Kirchhoff plate in the presence of a distributed frictional loading on its surface. The boundary value problem is derived from the basics of the theory of elasticity using Kirchhoff's assumptions. This results in precise information about the coupling between the disk and the plate vibrations under the action of frictional forces. At the same time we obtain a new model, which is efficient for analytical treatment. Approximations to the stability boundaries of the system are calculated using a perturbation approach. In the last part of the paper nonlinearities are introduced leading to limit cycles due to self-excited vibrations. [DOI: 10.1115/1.3112734]

1 Introduction

Rotating plates occur in many technical applications such as computer hard drives, turbines, saws, clutches, and disk brakes. In some of these applications squeal phenomena arise due to instabilities of the system, of which the plate plays an essential role. Those effects are usually unwanted. In many cases the instabilities and therefore the squeal phenomena are due to self-excited vibrations originating from frictional loads acting on the surface of the rotating plate. The present paper deals with rotating plates in contact with idealized friction pads.

Various questions related to rotating plates have been extensively investigated in literature. Probably among the first publications in the field were the papers by Lamb and Southwell [1,2], giving the equations of motion of a rotating Kirchhoff plate taking into account the effect of centrifugal stiffening due to membrane stresses. Another approach is to derive the equations of motion for plates using the corresponding kinematic assumptions in the basic equations of the theory of elasticity (see, e.g., Ref. [3]). For spinning disks, different kinds of modeling strategies are reviewed and compared in Refs. [4,5].

It is well known that the linear disk and plate equations are decoupled. Whereas a derivation from the linear deformation gradient yields linear equations, disk and plate equations are coupled through nonlinear terms when the nonlinear deformation gradient is used [3]. If the plate interacts with other mechanical systems, coupling between disk and plate vibrations is possible even when using a geometrically linearized theory. In our paper, the plate interacts with idealized friction pads, which may lead to such a type of coupling. Especially concerning the problem of brake squeal, where only small deformations occur and therefore the geometrically linearized theory seems to be appropriate, the influence of the disk vibrations to the plate vibrations has not been fully investigated, although experimental results seem to indicate this type of coupling [6]. An overview directly related to the present paper, i.e., friction induced instabilities in rotating disks, is given by Mottershead [7].

Ono et al. [8] considered a rotating plate in frictional contact with a transverse dynamical system as a model of a computer hard drive. Centrifugal and aeroelastic effects were taken into account, as required by typical rotational speeds of hard drives. In the context of brake squeal Ouyang and Mottershead [9] investigated a stationary plate loaded by a rotating friction couple. These results explain the instabilities occurring in a rotating plate, provided the effect of centrifugal stiffening can be neglected. It has, however, to be said that the model contains slight inconsistencies regarding the modeling of the frictional contact, since normal and friction forces are not perpendicular at all times in this analysis. A broad overview of models related to the squeal of disk brakes is given in Ref. [10], and an overview about the physical phenomena in friction induced vibrations is given in Ref. [11].

Nevertheless, none of the papers cited above comment on the possible coupling between disk (in-plane) and plate (out-of-plane) vibrations, in general, and due to contact forces originating from a frictional contact, in particular. The paper by Tseng and Wickert [12] aims in this direction by investigating a rotating plate loaded by a given shear stress acting like a follower force on the surface of the plate. Following this approach it is possible to calculate the membrane in-plane stresses in a first step and then introduce them into the plate equations in a second step. Since the forces acting on the plate are modeled as a given shear stress their magnitude does not depend on the transverse displacement of the plate, which excludes additional coupling effects of the in- and out-of-plane vibrations.

In the present paper we propose a model of a disk brake, which takes into account in- and out-of-plane motions of the brake rotor, and its interaction with friction pads. Using Kirchhoff's kinematical assumptions for the rotating continuum in interaction with friction pads we derive a new boundary value problem from the basics of the theory of elasticity using the variational approach. The consistent contact formulation based on Coulomb's law of friction given by Hochlenert et al. [13] is retained and extended to distributed contact. Concentrating on the coupling effect of the pads, which has not been investigated in literature, we prefer to work with the linear deformation gradient. For completeness we show in Sec. 4.4 that working with the nonlinear deformation

[1]Visiting from the Institute of Mechanics, Moscow State Lomonosov University, Michurinskii Prospect 1, 119192 Moscow, Russia.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 15, 2007; final manuscript received February 3, 2009; published online April 22, 2009. Review conducted by Arvind Raman.

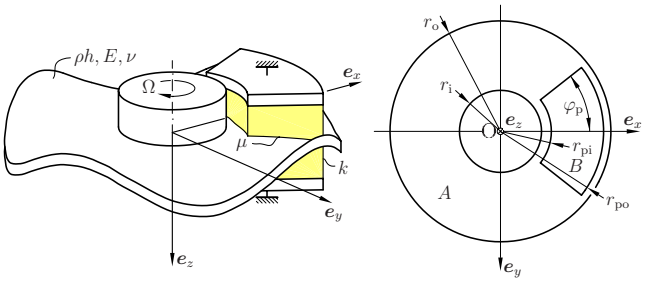


Fig. 1 Kirchhoff plate in distributed frictional contact

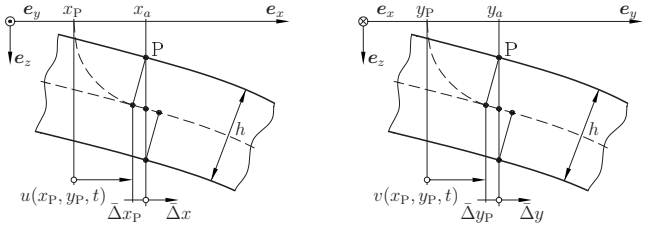


Fig. 3 Contact kinematics

gradient essentially introduces the effect of centrifugal stiffening into the model but does not yield an additional coupling in the linearized equations. Neglecting in-plane degrees of freedom of the pads we investigate whether there is a coupling induced by the out-of-plane degrees of freedom of the pads. It is established that the linearized equations for the in- and out-of-plane motions are decoupled and thus can be solved independently. This justifies the discretization approach described in Ref. [13]. The mathematical model presented is efficient to handle even analytically using perturbation techniques and at the same time it is rather flexible allowing for various extensions.

A perturbation approach developed in Refs. [14,15] is used to approximate the stability boundaries of the present model. In this context the terms originating from the frictional load are considered as perturbations to the problem. The equations of motion of the unperturbed problem corresponding to the out-of-plane vibrations can be written as infinitely many uncoupled two-dimensional systems, making it possible to derive approximations to the stability boundaries by considering these reduced systems only. New explicit perturbation formulas are derived, which prove to be an efficient tool in the stability analysis.

In the last part of the paper a reduced nonlinear model with parameters corresponding to a standard disk brake is studied. It is shown that depending on the parameters, the system can get unstable either via a sub- or a supercritical Hopf bifurcation. The latter one serves as an explanation for the nonlinear effects observed in laboratory experiments with a standard disk brake that to the extent of the authors' knowledge so far have not been investigated.

2 Derivation of the Mathematical Model

Consider a rotating Kirchhoff plate in frictional contact with idealized brake pads, as shown in Fig. 1. The pads are composed of massless pins in contact with the plate at a single point. In the following we will refer to this as a pointwise elastic model. For simplicity the pins are assumed to have identical characteristics. The model, however, can easily be extended to take into account a nonuniform distribution of the load among the pins. In Fig. 2 the plate is in contact with a single pair of pins. For the derivation of the equations of motion we have to consider the kinematics of the problem and the forcing terms arising due to the pads.

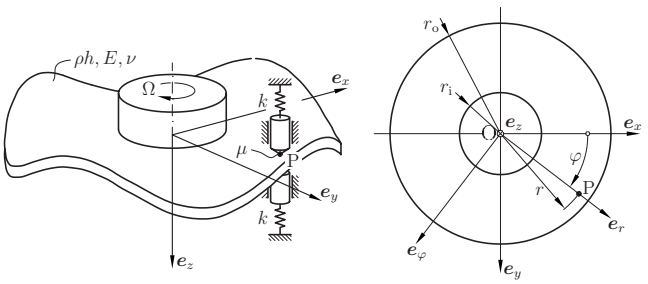


Fig. 2 Kirchhoff plate in pointwise frictional contact

2.1 Kinematics. As usual in Kirchhoff plate theory we assume the following.

- Transverse normal stress can be neglected, i.e., $\sigma_z=0$.
- Material points located on a normal to the neutral plane in the undeformed configuration and in the deformed configuration will be located on one and the same normal to the deformed neutral plane; moreover, the segment formed by these normals is inextensible.

A point on the neutral plane of the plate with the position vector $\boldsymbol{p}_{M0}=x\boldsymbol{e}_x+y\boldsymbol{e}_y$ in the undeformed configuration experiences the displacement

$$\boldsymbol{u}_M = u(x,y,t)\boldsymbol{e}_x + v(x,y,t)\boldsymbol{e}_y + w(x,y,t)\boldsymbol{e}_z \tag{1}$$

and its position vector in the deformed configuration is given by

$$\boldsymbol{p}_M(x,y,t) = (x+u(x,y,t))\boldsymbol{e}_x + (y+v(x,y,t))\boldsymbol{e}_y + w(x,y,t)\boldsymbol{e}_z \tag{2}$$

According to the Kirchhoff assumption the unit normal vector at a point of the neutral plane is given by

$$\boldsymbol{e}_\nabla(x,y,t) = \frac{\dfrac{\partial \boldsymbol{p}_M}{\partial x}\times\dfrac{\partial \boldsymbol{p}_M}{\partial y}}{\left\|\dfrac{\partial \boldsymbol{p}_M}{\partial x}\times\dfrac{\partial \boldsymbol{p}_M}{\partial y}\right\|} \tag{3}$$

and hence a point on the plate surface is given by

$$\boldsymbol{p}(x,y,t) = (x+u(x,y,t))\boldsymbol{e}_x + (y+v(x,y,t))\boldsymbol{e}_y + w(x,y,t)\boldsymbol{e}_z - \frac{h}{2}\boldsymbol{e}_\nabla(x,y,t) \tag{4}$$

The displacement vector of an arbitrary point of the plate is therefore

$$\boldsymbol{u}(x,y,z,t) = \boldsymbol{u}_M(x,y,t) + z(\boldsymbol{e}_\nabla(x,y,t) - \boldsymbol{e}_z) \tag{5}$$

which can be expanded in a Taylor series with respect to u, v, and w to arbitrary order. The linearized expression reads

$$\boldsymbol{u}(x,y,z,t) = (u - zw_{,x})\boldsymbol{e}_x + (v - zw_{,y})\boldsymbol{e}_y + w\boldsymbol{e}_z \tag{6}$$

When differentiating Eq. (2) with respect to time we have to consider $\dot{x}=-\Omega y$ and $\dot{y}=\Omega x$ due to the rotation of the plate, which will be important for the inertia terms arising in the equations of motion.

When considering the point on the surface of the plate currently in contact with a certain pin belonging to the pad we have to calculate its positions x_P and y_P as shown in Fig. 3. From geometrical considerations it is seen that

$$\bar{\Delta}x_P = -\frac{h}{2}\sin(\arctan w_{,x}(x_a+\bar{\Delta}x_P, y_a+\bar{\Delta}y_P, t))$$

$$= -\frac{h}{2}\frac{w_{,x}(x_a+\bar{\Delta}x_P, y_a+\bar{\Delta}y_P, t)}{\sqrt{1+w_{,x}(x_a+\bar{\Delta}x_P, y_a+\bar{\Delta}y_P, t)^2}}$$

$$\bar{\Delta}y_P = -\frac{h}{2}\sin(\arctan w_{,y}(x_a + \bar{\Delta}x_P, y_a + \bar{\Delta}y_P, t))$$

$$= -\frac{h}{2}\frac{w_{,y}(x_a + \bar{\Delta}x_P, y_a + \bar{\Delta}y_P, t)}{\sqrt{1 + w_{,y}(x_a + \bar{\Delta}x_P, y_a + \bar{\Delta}y_P, t)^2}}$$

where $w_{,x}$ means differentiation of the function w with respect to x. These are fixed point equations of the type

$$(\bar{\Delta}x_P^{k+1}, \bar{\Delta}y_P^{k+1})^T = \boldsymbol{g}(\bar{\Delta}x_P^k, \bar{\Delta}x_P^k) \quad \text{with} \quad \bar{\Delta}x_P^0 = 0, \quad \bar{\Delta}y_P^0 = 0 \tag{7}$$

Since $w_{,x}$ and $w_{,y}$ are small compared with unity, the Banach fixed point theorem is applicable and Eq. (7) can be solved to arbitrary precision.

We can now calculate $x_P = x_a + \Delta x_P$ and $y_P = y_a + \Delta y_P$ where

$$\Delta x_P = \bar{\Delta}x_P + u(x_P, y_P, t) \tag{8}$$

$$= -\frac{h}{2}\frac{w_{,x}(x_P, y_P, t)}{\sqrt{1 + w_{,x}(x_P, y_P, t)^2}} + u(x_P, y_P, t) \tag{9}$$

$$\Delta y_P = \bar{\Delta}y_P + v(x_P, y_P, t) \tag{10}$$

$$= -\frac{h}{2}\frac{w_{,y}(x_P, y_P, t)}{\sqrt{1 + w_{,y}(x_P, y_P, t)^2}} + v(x_P, y_P, t) \tag{11}$$

are again fixed point equations that can be solved to arbitrary precision. In the linear case, however, we have

$$\Delta x_P = -\frac{h}{2}w_{,x}(x_a, y_a, t) + u(x_a, y_a, t) + o(u, w) \tag{12}$$

$$\Delta y_P = -\frac{h}{2}w_{,y}(x_a, y_a, t) + v(x_a, y_a, t) + o(v, w) \tag{13}$$

Similar relations hold for the lower contact point. Since Δx_P and Δy_P appear only in the argument of the functions u and w, in the geometrically linearized equations they do not make a difference, namely,

$$w(a + \Delta x_P, y, t) = w(a, y, t) + o(u, w) \tag{14}$$

2.2 Contact Forces. We now investigate the contact forces acting between one of the pins of the pad and the plate (see Fig. 4). With a pad formed by infinitely many distributed pins, these forces will be substituted by stresses. The normal force is given by

$$N_P = -N_{\bar{P}} = N_P \boldsymbol{e}_\nabla|_P$$

and the friction force is

$$\boldsymbol{R}_P = -\boldsymbol{R}_{\bar{P}} = R_P \frac{\boldsymbol{v}_{\bar{P}} - \boldsymbol{v}_P}{|\boldsymbol{v}_{\bar{P}} - \boldsymbol{v}_P|}$$

which means that its direction is opposite to the relative velocity between the pin and the contact point on the plate. Using Coulomb's law of friction

$$R_P = \mu N_P \tag{15}$$

and the force balance at the pin in $\boldsymbol{e}_z$-direction

$$(\boldsymbol{N}_{\bar{P}} + \boldsymbol{R}_{\bar{P}}) \cdot \boldsymbol{e}_z + N_0 - k(z_{\bar{P}} + h/2) = 0 \tag{16}$$

makes it possible to calculate $\boldsymbol{N}_{\bar{P}}$ and $\boldsymbol{R}_{\bar{P}}$.

The resulting contact forces can be obtained by integrating over all pins, i.e., the area B of material points in contact with the pins (see Fig. 1). Since the segments normal to the neutral plane stay rigid, according to the kinematical assumptions, loads on the surface of the plate can be replaced with an equivalent force

$$\boldsymbol{F} = F_x \boldsymbol{e}_x + F_y \boldsymbol{e}_y + F_z \boldsymbol{e}_z \tag{17}$$

and a torque

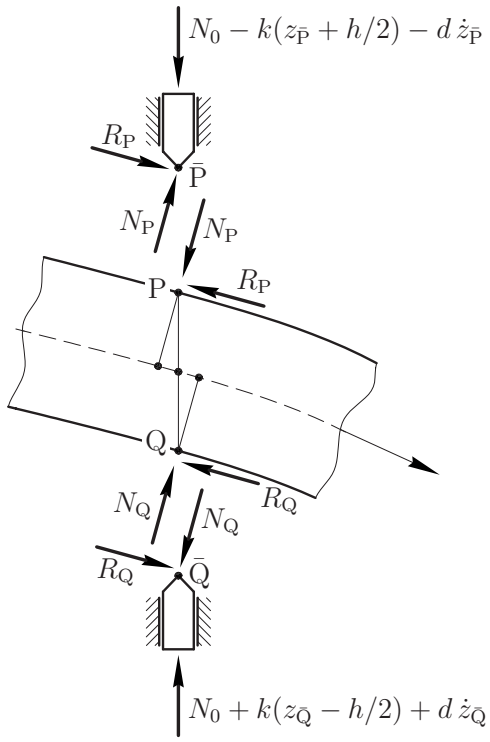


Fig. 4 Contact forces

$$\boldsymbol{T} = T_x \boldsymbol{e}_x + T_y \boldsymbol{e}_y \tag{18}$$

both acting on the neutral plane of the plate.

3 Boundary Value Problem for the Rotating Plate With Pads

3.1 Principle of Virtual Work. In Kirchhoff plate theory, the inertia of the plate is assumed to be concentrated in the neutral plane of the plate. Hence the principle of virtual work can be stated as

$$\int_V \left(\rho \frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{p}_M \cdot \delta\boldsymbol{p}_M + \sigma_{xx}\delta e_{xx} + \ \ldots\ + \sigma_{xz}\delta e_{xz}\right)\mathrm{d}V$$

$$= \int_A (\boldsymbol{F} \cdot \delta\boldsymbol{p}_M + T_x \delta w_{,y} - T_y \delta w_{,x})\mathrm{d}A \tag{19}$$

where A is the area of the surface of the plate, see Fig. 1. The terms occurring in Eq. (19) will be discussed in more detail below. The acceleration vector can be calculated from Eq. (2) by simple differentiation, noting that $\dot{x} = -\Omega y$ and $\dot{y} = \Omega x$, due to the rotation of the plate. Substituting the displacement vector (6) into the well known strain-displacement relations one obtains the strains e_{ij}. To keep expressions simple we first work with e_{ij} linearized with respect to u, v, and w. The stresses in Eq. (19) are calculated from the stress-strain relations for linear isotropic material with the plane stress assumption $\sigma_{zz} = 0$.

3.2 Derivation of a Boundary Value Problem. In order to derive the equations of motion from Eq. (19) we have to apply a variant of Gauß theorem

$$\int_V f\delta w_{,x}\mathrm{d}V = -\int_V f_{,x}\delta w \mathrm{d}V + \int_S f\delta w n_x \mathrm{d}S \tag{20}$$

several times, where n_x denotes the x-measure number of the normal vector $\boldsymbol{n}$ on S. Similar relations hold for the y- and the

z-direction. Carrying out the variations in Eq. (19) we arrive at expressions

$$\int_V f_x \delta u \mathrm{d}V + \int_S (\boldsymbol{V}^u \cdot \boldsymbol{n}) \delta u \mathrm{d}S + \int_S (\boldsymbol{M}_x^u \cdot \boldsymbol{n}) \delta u_{,x} \mathrm{d}S + \int_S (\boldsymbol{M}_y^u \cdot \boldsymbol{n}) \delta u_{,y} \mathrm{d}S + \int_V f_y \delta v \mathrm{d}V + \int_S (\boldsymbol{V}^v \cdot \boldsymbol{n}) \delta v \mathrm{d}S + \int_S (\boldsymbol{M}_x^v \cdot \boldsymbol{n}) \delta v_{,x} \mathrm{d}S + \int_S (\boldsymbol{M}_y^v \cdot \boldsymbol{n}) \delta v_{,y} \mathrm{d}S + \int_V f_z \delta w \mathrm{d}V + \int_S (\boldsymbol{V}^w \cdot \boldsymbol{n}) \delta w \mathrm{d}S + \int_S (\boldsymbol{M}_x^w \cdot \boldsymbol{n}) \delta w_{,x} \mathrm{d}S + \int_S (\boldsymbol{M}_y^w \cdot \boldsymbol{n}) \delta w_{,y} \mathrm{d}S = 0 \tag{21}$$

from which the boundary value problems for u, v, and w follow by applying the main theorem of variational calculus. It is interesting to note that the boundary value problem is by no means unique. This can be easily seen by noting that terms of the form

$$\int_V f \delta w_{,xy} \mathrm{d}V \tag{22}$$

occur. Depending on whether integration by parts is performed first for x or y, different boundary conditions arise. In fact, there are infinitely many possible boundary value problems that can be derived from Eq. (19). In order to preserve symmetries of the problem, we will treat terms of the form (22) in the following way:

$$\int_V f \delta w_{,xy} \mathrm{d}V = \frac{1}{2} \int_V f \delta w_{,xy} \mathrm{d}V + \frac{1}{2} \int_V f \delta w_{,xy} \mathrm{d}V = \frac{1}{2} \left(- \int_V f_{,x} \delta w_{,y} \mathrm{d}V + \int_S f \delta w_{,y} n_x \mathrm{d}S \right) + \frac{1}{2} \left(- \int_V f_{,y} \delta w_{,x} \mathrm{d}V + \int_S f \delta w_{,x} n_y \mathrm{d}S \right) \tag{23}$$

where a second integration by parts yields

$$\int_V f \delta w_{,xy} \mathrm{d}V = \frac{1}{2} \left(\int_V f_{,xy} \delta w \mathrm{d}V - \int_S f_{,x} \delta w n_y \mathrm{d}S + \int_S f \delta w_{,y} n_x \mathrm{d}S \right) + \frac{1}{2} \left(\int_V f_{,yx} \delta w \mathrm{d}V - \int_S f_{,y} \delta w n_x \mathrm{d}S + \int_S f \delta w_{,x} n_y \mathrm{d}S \right) \tag{24}$$

Choosing this symmetric approach and using the common abbreviation $D = Eh^3/12(1-\nu^2)$, the equations of motion follow as

$$\rho h u_{,tt} + 2\rho h \Omega (x u_{,ty} - y u_{,tx}) - \frac{6D}{h^2}((1-\nu) u_{,yy} + (1+\nu) v_{,xy} + 2u_{,xx}) - \rho h \Omega^2 (x - x^2 u_{,yy} + x u_{,x} + 2xy u_{,xy} + y u_{,y} - y^2 u_{,xx}) = F_x \tag{25a}$$

$$\rho h v_{,tt} + 2\rho h \Omega (x u_{,ty} - y v_{,tx}) - \frac{6D}{h^2}(1+\nu)(u_{,xy} + (1-\nu) v_{,xx} + 2v_{,yy}) - \rho h \Omega^2 (y - y^2 v_{,xx} + y v_{,y} + 2xy v_{,xy} + y u_{,y} - x^2 v_{,yy} + x v_{,x}) = F_y \tag{25b}$$

$$\rho h w_{,tt} + 2\rho h \Omega (x w_{,ty} - y w_{,tx}) + D(w_{,xxxx} + 2w_{,xxyy} + w_{,yyyy}) - \rho h \Omega^2 (y w_{,y} - x^2 w_{,yy} + x w_{,x} + 2xy w_{,xy} - y^2 w_{,xx}) = F_z - T_{x,y} + T_{y,x} \tag{26}$$

with boundary conditions following from

$$\delta u_S \int_{-h/2}^{h/2} (\sigma_{xx} n_x + \sigma_{xy} n_y) \mathrm{d}z = \delta u_S \left(\frac{12D}{h^2} (u_{,x} + \nu v_{,y}) n_x + \frac{6D(1-\nu)}{h^2} (u_{,y} + v_{,x}) n_y \right) = 0 \tag{27a}$$

$$\delta v_S \int_{-h/2}^{h/2} (\sigma_{xy} n_x + \sigma_{yy} n_y) \mathrm{d}z = \delta v_S \left(\frac{6D(1-\nu)}{h^2} (u_{,y} + v_{,x}) n_x + \frac{12D}{h^2} (\nu u_{,x} + v_{,y}) n_y \right) = 0 \tag{27b}$$

and

$$\delta w_S \int_{-h/2}^{h/2} \left(\left(\frac{1}{h} T_y + z\sigma_{xy,y} + z\sigma_{xx,x} \right) n_x + \left(-\frac{1}{h} T_x + z\sigma_{yy,y} + z\sigma_{xy,x} \right) n_y \right) \mathrm{d}z = [D(w_{,xyy} + w_{,xxx}) - T_y] n_x + [D(w_{,yyy} + w_{,xxy}) + T_x] n_y = 0 \tag{28a}$$

$$\delta w_{S,x} \int_{-h/2}^{h/2} (z\sigma_{xx} n_x + z\sigma_{xy} n_y) \mathrm{d}z + \delta w_{S,y} \int_{-h/2}^{h/2} (z\sigma_{xy} n_x + z\sigma_{yy} n_y) \mathrm{d}z = \delta w_{S,x} (w_{,xx} + \nu w_{,yy}) n_x + \delta w_{S,y} (w_{,yy} + \nu w_{,xx}) n_y = 0 \tag{28b}$$

By inspection of the boundary value problems (26) and (28) of the plate (out-of-plane) and the boundary value problems (25) and (27) of the disk (in-plane) it can be seen that the coupling of the plate and disk equations depends on F_x, F_y, F_z and T_x, T_y only, when using the linear deformation gradient. In the linear case for the pointwise elastic model with no in-plane degree of freedom of the pad we have

$$F_i = F_i(u, u_{,j}, u_{,t}, u_{,jt}, v, v_{,j}, v_{,t}, v_{,jt}), \quad i,j \in \{x,y\} \tag{29a}$$

$$F_z = F_z(w, w_{,j}, w_{,t}, w_{,jt}), \quad j \in \{x,y\} \tag{29b}$$

$$T_i = T_i(w, w_{,j}, w_{,ij}, w_{,t}, w_{,jt}), \quad i,j \in \{x,y\} \tag{29c}$$

and therefore the plate and disk equations decouple. If second order terms are considered in the deformation gradient, we obtain the strain-displacement relations

$$e_{xx} = u_{,x} - z w_{,xx} + \tfrac{1}{2}(u_{,x}^2 + v_{,x}^2 + w_{,x}^2 + z^2 w_{,x,y} + z^2 w_{,xx}^2) + z w_{,x} u_{,xx} + z w_{,y} v_{,xx} \tag{30a}$$

$$e_{yy} = v_{,y} - z w_{,yy} + \tfrac{1}{2}(u_{,y}^2 + v_{,y}^2 + w_{,y}^2 + z^2 w_{,x,y} + z^2 w_{,yy}^2) + z w_{,x} u_{,yy} + z w_{,y} v_{,yy} \tag{30b}$$

$$e_{zz} = 0 \tag{30c}$$

$$e_{xy} = \tfrac{1}{2} u_{,y} + \tfrac{1}{2} v_{,x} - z w_{,xy} + \tfrac{1}{2}(u_{,y} u_{,x} + v_{,y} v_{,x} + w_{,y} w_{,x} + z^2 (w_{,yy} w_{,xy} + w_{,xy} w_{,xx})) + z(w_{,x} u_{,xy} + w_{,y} v_{,xy}) \tag{30d,e}$$

$$e_{yz} = 0 \tag{30f}$$

$$e_{xz} = 0 \tag{30g}$$

and it is clear that the quadratic terms containing w, u, and v will yield a coupling of the disk and plate equations after taking the variation of e_{ij} [3]. We note, however, that due to the integration of the equations over z the coupling occurs only through nonlinear terms. A second fact to be noted is that, when linearizing the equations about a prestressed configuration, i.e., $\sigma_{ij}=\sigma_{ij}^0+\Delta\sigma_{ij}$, the prestress terms only enter the equations of motion if the nonlinear deformation gradient is used. With the linear deformation gradient terms of the form $\sigma_{ij}^0\delta e_{ij}$ vanish in the process of integrating by parts. This is why the effect of centrifugal stiffening [8] does not appear even when the equations of motion are linearized about a prestressed configuration. Since we are particularly interested in small deformations of the plate we continue working with the linearized strain relations and discuss the influence of centrifugal stiffening in Sec. 4.4.

3.3 Transformation to Polar Coordinates. We will now transform our equations into polar coordinates, i.e., we will express the functions u, v, and w and the components of the normal vector $\boldsymbol{n}$ in polar coordinates. The function $w(x,y,t)$ and its derivatives can be written in polar coordinates using

$$\frac{\partial^{a+b}}{\partial x^a\,\partial y^b}w(x,y,t) = \frac{\partial^{a+b}}{\partial x^a\,\partial y^b}\tilde{w}(r(x,y),\varphi(x,y),t) \tag{31}$$

and noting that $r(x,y)=\sqrt{x^2+y^2}$ and $\varphi(x,y)=\arctan(y/x)$. After carrying out the differentiations we set $x=r\cos\varphi$ and $y=r\sin\varphi$. For u and v we proceed similarly by writing

$$u(x,y,t) = \tilde{u}(r,\varphi,t)\cos\varphi - \tilde{v}(r,\varphi,t)\sin\varphi \tag{32a}$$

$$v(x,y,t) = \tilde{v}(r,\varphi,t)\cos\varphi + \tilde{u}(r,\varphi,t)\sin\varphi \tag{32b}$$

which means that we define a new displacement vector for points on the neutral plane

$$\boldsymbol{u}_M = \tilde{u}(r,\varphi,t)\boldsymbol{e}_r + \tilde{v}(r,\varphi,t)\boldsymbol{e}_\varphi + \tilde{w}(r,\varphi,t)\boldsymbol{e}_z \tag{33}$$

In the following we omit the tilde in $\tilde{u}(r,\varphi,t)$, $\tilde{v}(r,\varphi,t)$, and $\tilde{w}(r,\varphi,t)$. There should be no confusion between $u(x,y,t)$, $v(x,y,t)$, $w(x,y,t)$ and $\tilde{u}(r,\varphi,t)$, $\tilde{v}(r,\varphi,t)$, $\tilde{w}(r,\varphi,t)$ since they can be distinguished by their arguments and by the context.

3.3.1 Contact Forces. The contact forces and torques contain contributions of the upper and the lower pad, which read

$$F_r^P = \kappa(r,\varphi)\left(-N_0 w_{,r} - h\mu N_0\left(\frac{w_{,tr}}{2r\Omega} + \frac{w_{,r\varphi}}{2r} - \frac{w_{,\varphi}}{2r^2}\right) + \mu N_0\frac{\Omega(v-u_{,\varphi})-u_{,t}}{r\Omega}\right) \tag{34a}$$

$$F_r^Q = \kappa(r,\varphi)\left(N_0 w_{,r} + h\mu N_0\left(\frac{w_{,tr}}{2r\Omega} + \frac{w_{,r\varphi}}{2r} - \frac{w_{,\varphi}}{2r^2}\right) + \mu N_0\frac{\Omega(v-u_{,\varphi})-u_{,t}}{r\Omega}\right) \tag{34b}$$

$$F_\varphi^P = \kappa(r,\varphi)\left(-\mu N_0 + k\mu w - \frac{N_0}{r}(1+\mu^2)w_{,\varphi}\right) \tag{34c}$$

$$F_\varphi^Q = \kappa(r,\varphi)\left(-\mu N_0 - k\mu w + \frac{N_0}{r}(1+\mu^2)w_{,\varphi}\right) \tag{34d}$$

$$F_z^P = \kappa(r,\varphi)(N_0 - kw) \tag{34e}$$

$$F_z^Q = \kappa(r,\varphi)(-N_0 - kw) \tag{34f}$$

and the contact torques are obtained from

$$\boldsymbol{T}^P = \frac{h}{2}\boldsymbol{e}_\nabla \times (F_r^P\boldsymbol{e}_r + F_\varphi^P\boldsymbol{e}_\varphi + F_z^P\boldsymbol{e}_z) \tag{35a}$$

$$\boldsymbol{T}^Q = -\frac{h}{2}\boldsymbol{e}_\nabla \times (F_r^Q\boldsymbol{e}_r + F_\varphi^Q\boldsymbol{e}_\varphi + F_z^Q\boldsymbol{e}_z) \tag{35b}$$

where $\kappa(r,\varphi)$ is a weight function describing the area of the pads. It can be chosen to be a continuous function or, for example, as $\kappa(r,\varphi)=1$ if $(r,\varphi)\in B$, i.e., if (r,φ) are in the domain of the pad (denoted by B to be distinguished from the whole area of the plate surface denoted earlier by A) and $\kappa(r,\varphi)=0$ otherwise. In principle it would be possible to assume that all pad parameters depend on r and φ, to introduce more generality into the model. No realistic data are, however, currently available on such distributions from experiments, and for the sake of simplicity we do not proceed further in that direction. In the linear case the normal vector reads

$$\boldsymbol{e}_\nabla = -w_{,r}\boldsymbol{e}_r - \frac{w_{,\varphi}}{r}\boldsymbol{e}_\varphi + \boldsymbol{e}_z \tag{36}$$

and it follows that

$$F_r = F_r^P + F_r^Q = \kappa(r,\varphi)2\mu N_0\frac{\Omega(v-u_{,\varphi})-u_{,t}}{r\Omega} \tag{37a}$$

$$F_\varphi = F_\varphi^P + F_\varphi^Q = -\kappa(r,\varphi)2\mu N_0 \tag{37b}$$

$$F_z = F_z^P + F_z^Q = -\kappa(r,\varphi)2kw \tag{37c}$$

$$T_r = T_r^P + T_r^Q = \kappa(r,\varphi)\left(hk\mu w - \frac{hN_0\mu^2 w_{,\varphi}}{r}\right) \tag{37d}$$

$$T_\varphi = T_\varphi^P + T_\varphi^Q = \kappa(r,\varphi)h^2N_0\mu\left(\frac{w_{,rt}}{2r\Omega} - \frac{w_{,\varphi}}{2r^2} + \frac{w_{,r\varphi}}{2r}\right) \tag{37e}$$

where many terms from the upper and lower pads cancel, so that the in-plane forces F_r and F_φ depend on u and v and their derivatives only, whereas F_z, T_r, and T_φ only depend on w and its derivatives. This is the justification for Eq. (29), which explains the decoupling of disk and plate equations, as discussed in Sec. 3.2. It is due to the assumption that the pads do not have a degree of freedom in the in-plane direction. The dependence of F_r and F_φ on u and v comes from the relative velocity of the contact points on the pad and the plate. If we would give the upper and the lower pad independent degrees of freedom in the in-plane direction, say, $u_{\bar{P}}$ and $v_{\bar{P}}$ for the upper pad, they would appear in the relative velocity. Therefore if the in-plane degrees of freedom of the upper and lower pads are independent, they will appear in F_r, F_φ and T_r, T_φ and thus couple the disk and plate equations. Note that in practice in-plane vibrations are possible and were detected [6]. The analysis shows that they can either be explained through nonlinear coupling of disk and plate equations or that a linear coupling can arise through in-plane degrees of freedom of the pads, which also occur in practice.

3.3.2 Plate Equations. Transforming the plate equations into polar coordinates we obtain

$$\rho h w_{,tt} + 2h\rho\Omega w_{,t\varphi} + \Omega^2\rho h w_{,\varphi\varphi} + \frac{D}{r^4}(4w_{,\varphi\varphi} + w_{,\varphi\varphi\varphi\varphi}) + \frac{D}{r^3}(w_{,r} + 2w_{,r\varphi\varphi}) + \frac{D}{r^2}(2w_{,rr\varphi\varphi} - w_{,rr}) + 2\frac{D}{r}w_{,rrr} + Dw_{,rrrr} = F_z - \frac{1}{r}T_{r,\varphi} + T_{\varphi,r} \tag{38}$$

where the transformed natural boundary operators read

$$\boldsymbol{V}^w = V_r^w\boldsymbol{e}_r + V_\varphi^w\boldsymbol{e}_\varphi \tag{39a}$$

$$V_r^w = -M_\varphi + \frac{2D}{r^3}w_{,\varphi\varphi} + \frac{D}{r^2}(w_{,r} - w_{,r\varphi\varphi}) - \frac{D}{r}w_{,rr} - Dw_{,rrr} \tag{39b}$$

$$V_\varphi^w = M_r - \frac{D}{r^3}w_{,\varphi\varphi\varphi} - \frac{D}{r^2}w_{,r\varphi} - \frac{D}{r}w_{,rr\varphi} \tag{39c}$$

and

$$\boldsymbol{M} = M_r\boldsymbol{e}_r + M_\varphi\boldsymbol{e}_\varphi \tag{39d}$$

$$M_r = -\frac{D}{r^2}(1-\nu)w_{,\varphi} + \frac{D}{r}(1-\nu)w_{,r\varphi} \tag{39e}$$

$$M_\varphi = \frac{D\nu}{r^2}w_{,\varphi\varphi} + \frac{D\nu}{r}w_{,r} + Dw_{,rr} \tag{39f}$$

For the annular plate, clamped at the inner and free at the outer radius, we obtain the geometric boundary conditions $w|_{r=r_i}=0$ and $w_{,r}|_{r=r_i}=0$, and the natural boundary conditions $V_r|_{r=r_o}=0$ and $M_r|_{r=r_o}=0$.

3.3.3 Disk Equations. The boundary value problem for the disk in polar coordinates reads

$$\rho h\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\ddot{u}\\ \ddot{v}\end{bmatrix} + \begin{bmatrix}\kappa 2\dfrac{\mu N_0}{r\Omega} + 2\Omega\rho h\dfrac{\partial}{\partial\varphi} & -2\Omega\rho h\\ 2\Omega\rho h & 2\Omega\rho h\dfrac{\partial}{\partial\varphi}\end{bmatrix}\begin{bmatrix}\dot{u}\\ \dot{v}\end{bmatrix} + \begin{bmatrix}L_{11} & L_{12}\\ L_{21} & L_{22}\end{bmatrix}\begin{bmatrix}u\\ v\end{bmatrix} = \begin{bmatrix}\rho h\Omega^2 r\\ -\kappa 2\mu N_0\end{bmatrix} \tag{40}$$

(where as before $\kappa=\kappa(r,\varphi)$) with the linear operators

$$L_{11} = -\frac{Eh}{1-\nu^2}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2} + \frac{1-\nu}{2r^2}\frac{\partial^2}{\partial\varphi^2}\right) - \rho h\Omega^2\left(1 - \frac{\partial^2}{\partial\varphi^2}\right) + \kappa 2\frac{\mu N_0}{r}\frac{\partial}{\partial\varphi} \tag{41a}$$

$$L_{12} = -\frac{Eh}{1-\nu^2}\left(\frac{1+\nu}{2}\frac{1}{r}\frac{\partial^2}{\partial r\,\partial\varphi} - \frac{3-\nu}{2}\frac{1}{r^2}\frac{\partial}{\partial\varphi}\right) - 2\rho h\Omega^2\frac{\partial}{\partial\varphi} - \kappa 2\frac{\mu N_0}{r} \tag{41b}$$

$$L_{21} = -\frac{Eh}{1-\nu^2}\left(\frac{1+\nu}{2}\frac{1}{r}\frac{\partial^2}{\partial r\,\partial\varphi} + \frac{3-\nu}{2}\frac{1}{r^2}\frac{\partial}{\partial\varphi}\right) + 2\rho h\Omega^2\frac{\partial}{\partial\varphi} \tag{41c}$$

$$L_{22} = -\frac{Eh}{1-\nu^2}\left(\frac{1-\nu}{2}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2}\right) - \rho h\Omega^2\left(1 - \frac{\partial^2}{\partial\varphi^2}\right) \tag{41d}$$

and the natural boundary operators

$$\boldsymbol{V}^{u,v} = V_r^{u,v}\boldsymbol{e}_r + V_\varphi^{u,v}\boldsymbol{e}_\varphi \tag{42a}$$

$$V_r^u = \frac{1}{1-\nu^2}\left[u_{,r} + \nu\left(\frac{u}{r} + \frac{1}{r}v_{,\varphi}\right)\right] \tag{42b}$$

$$V_r^v = \frac{1}{2(1+\nu)}\left[\frac{1}{r}u_{,\varphi} + v_{,r} - \frac{v}{r}\right] \tag{42c}$$

$$V_\varphi^v = \frac{1}{2(1+\nu)}\left[\frac{1}{r}u_{,\varphi} + v_{,r} - \frac{v}{r}\right] \tag{42d}$$

$$V_\varphi^v = \frac{1}{1-\nu^2}\left[\frac{u}{r} + \frac{1}{r}v_{,\varphi} + \nu u_{,r}\right] \tag{42e}$$

For the annular disk clamped at the inner radius r_i and free at the outer radius r_o we have the geometric boundary conditions $u|_{r=r_i}=v|_{r=r_i}=0$ and the natural boundary conditions $V_r^u|_{r=r_o}=V_r^v|_{r=r_o}=0$. For $\mu=0$, that is, neglecting F_r and F_φ, the boundary value problem coincides with the one derived in Ref. [16]. Equations (40) and (42) form a linear inhomogeneous boundary value problem of which the solution is given by the general solution of the homogeneous boundary value problem plus a particular solution. Its stability can be studied investigating the stability of the trivial solution of the homogeneous boundary value problem.

3.4 Comparison to the Results Obtained Previously. In comparison to the results obtained in Ref. [13] using the Ritz discretization approach, two major differences shall be discussed. The first one is that in the continuous approach we use surface contact between disk and pads. The reason for this is that the Kirchhoff plate cannot deal with finite torques applied at points. A finite torque $\hat{\boldsymbol{T}}=\hat{T}_x\boldsymbol{e}_x+\hat{T}_y\boldsymbol{e}_y$ applied at a single point (x_a,y_a) of the plate appears in the principle of virtual work as a term

$$\delta W = \hat{T}_x\delta w_y(x_a,y_a,t) - \hat{T}_y\delta w_x(x_a,y_a,t) \tag{43}$$

After carrying out the variations there is no further term containing $\delta w_x(x_a,y_a,t)$ or $\delta w_y(x_a,y_a,t)$, since in δe_{ij} no term of the form $\delta w_{,xyy}$ or $\delta w_{,xxy}$ arises. Therefore one would conclude from the main theorem of variational calculus that $\hat{\boldsymbol{T}}=\boldsymbol{0}$, which is a contradiction. The Kirchhoff plate can therefore not resist finite torques applied at points. It is, however, very well able to resist finite transverse forces applied at points. Their contribution to the virtual work is

$$\delta W = \hat{F}_z\delta w(x_a,y_a,t) \tag{44}$$

Further terms containing $\delta w(x_a,y_a,t)$ arise from integration by parts of the term (22), yielding conditions for corner forces of different sections of the plate.

The second fact to be noted is that we used a strategy for linearization different from the one used in Ref. [13]. Whereas in Ref. [13] energy expressions were expanded in a Taylor series up to second order, to be sure to obtain the complete linear equations, in the boundary value problem derived in the present case only the linear contact forces appear, because a purely geometric linearization has been performed. The terms Δx_P and Δy_P therefore do not enter the equations of motion as explained in Sec. 2.1 with Eq. (14).

Another difference is that we got rid of the kinematic assumption that points on the neutral plane can only move transversely in the rotating frame, allowing for in-plane deformations of the plate. Nevertheless we again stress that this does not yield a term depending on u and v in the linearized contact forces, which would mean plate and disk equations cannot be solved independently.

Summarizing the previous results, we note that the boundary value problems of the plate and disk are decoupled as long as no in-plane motions of the pads are considered in the model. Furthermore a centrifugal stiffening cannot arise in the equations of motion when working with the linear deformation gradient. We have also recalled the well known fact from Kirchhoff theory that the plate is not capable to withstand finite torques (see, e.g., Ref. [17]).

4 Perturbation Analysis

We now interpret the terms arising from the pads as perturbations, i.e., in Eqs. (38) and (39) and Eqs. (40) and (42) we replace k with γk, and N_0 with δN_0. For the following calculations we introduce the dimensionless time $\bar{t}=t/r_o\sqrt{E/\rho}$ and the radius $\bar{r}=r/r_o$ yielding the dimensionless parameters

$$\bar{\Omega} = \Omega r_o \sqrt{\frac{\rho}{E}}, \quad \bar{r}_i = \frac{r_i}{r_o}, \quad \bar{r}_{pi} = \frac{r_{pi}}{r_o}, \quad \bar{r}_{po} = \frac{r_{po}}{r_o} \tag{45}$$

$$\bar{h} = \frac{h}{r_o}, \quad \bar{k} = k\frac{r_o}{E}, \quad \bar{N}_0 = \frac{N_0}{E}, \quad \bar{D} = \frac{\bar{h}^3}{12(1-\nu^2)} \tag{46}$$

To simplify notation, in the following the bars are omitted and derivatives are to be understood as derivatives with respect to the dimensionless variables. It will be shown in the sequel that by separating time from the equations, the boundary value problems for the plate and the disk can be written in the form

$$\boldsymbol{L}(\boldsymbol{w}) = \boldsymbol{L}^0(\boldsymbol{w}) + \gamma \boldsymbol{L}^{1\gamma}(\boldsymbol{w}) + \delta \boldsymbol{L}^{1\delta}(\boldsymbol{w}) = 0 \tag{47a}$$

$$\boldsymbol{U}_i(\boldsymbol{w}) = \boldsymbol{U}_i^0(\boldsymbol{w}) = 0 \tag{47b}$$

where we denote by $\boldsymbol{L}$ the operator matrix corresponding to the eigenvalue problem and $\boldsymbol{U}_i$ are operator matrices corresponding to the boundary conditions. In the plate problem we set $w(r,\varphi,t) = \bar{w}(r,\varphi)e^{\lambda t}$, where we again skip the bar for notational simplicity. The boundary value problem can be stated as

$$L^0(w) = \lambda^2 w + 2\lambda\Omega w_{,\varphi} + \Omega^2 w_{,\varphi\varphi} + \frac{D}{r^4}(4w_{,\varphi\varphi} + w_{,\varphi\varphi\varphi\varphi}) + \frac{D}{r^3}(w_{,r} + 2w_{,r\varphi\varphi}) + \frac{D}{r^2}(2w_{,rr\varphi\varphi} - w_{,rr}) + 2\frac{D}{r}w_{,rrr} + Dw_{,rrrr} \tag{48a}$$

$$L^{1\gamma}(w) = \kappa(r,\varphi)2kw - \left(\kappa(r,\varphi)\frac{hk\mu w}{r}\right)_{,\varphi} \tag{48b}$$

$$L^{1\delta}(w) = -\left(\kappa(r,\varphi)h^2 N_0 \mu\left(\lambda\frac{w_{,rt}}{2r\Omega} - \frac{w_{,\varphi}}{2r^2} + \frac{w_{,r\varphi}}{2r}\right)\right)_{,r} \tag{48c}$$

with boundary conditions

$$U_1 = w(r_i,\varphi,t) = 0, \quad U_2 = w_{,r}(r_i,\varphi,t) = 0 \tag{49}$$

$$U_3 = V_r^w|_{r=r_o} = 0, \quad U_4 = M_r|_{r=r_o} = 0 \tag{50}$$

where the operators for the differential equation and the boundary conditions act on the scalar function w and are therefore not represented by bold characters. In case $\kappa(r,\varphi)$ is chosen as a discontinuous function, we note that $L^{1\gamma}$ and $L^{1\delta}$ contain transition terms. In the perturbation formulas derived later, they can easily be eliminated by integration by parts. There are no perturbations in the boundary condition (47*b*); the boundary conditions of the perturbed problem therefore coincide with those of the unperturbed problem.

For the disk we perform the ansatz of separation of variables

$$\begin{bmatrix} u(r,\varphi,t) \\ v(r,\varphi,t) \end{bmatrix} = \begin{bmatrix} \bar{u}(r,\varphi) \\ \bar{v}(r,\varphi) \end{bmatrix} e^{\lambda t} = \boldsymbol{w} e^{\lambda t} \tag{51}$$

for simplicity skipping the bar for u and v; this yields the operators

$$\boldsymbol{L}^0(\boldsymbol{w}) = \left[\lambda^2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \lambda\begin{bmatrix} 2\Omega\dfrac{\partial}{\partial\varphi} & -2\Omega \\ 2\Omega & 2\Omega\dfrac{\partial}{\partial\varphi} \end{bmatrix} + \begin{bmatrix} L_{11}^0 & L_{12}^0 \\ L_{21}^0 & L_{22}^0 \end{bmatrix}\right]\begin{bmatrix} u \\ v \end{bmatrix} \tag{52a}$$

$$\boldsymbol{L}^{1\gamma}(\boldsymbol{w}) = \boldsymbol{0} \tag{52b}$$

$$\boldsymbol{L}^{1\delta}(\boldsymbol{w}) = \kappa(r,\varphi)\left[\lambda\begin{bmatrix} \dfrac{2\mu N_0}{r\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \dfrac{2\mu N_0}{r} & -\dfrac{2\mu N_0}{r} \\ 0 & 0 \end{bmatrix}\right]\begin{bmatrix} u \\ v \end{bmatrix} \tag{52c}$$

where

$$L_{11}^0 = -\frac{1}{1-\nu^2}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2} + \frac{1-\nu}{2r^2}\frac{\partial^2}{\partial\varphi^2}\right) - \Omega^2\left(1 - \frac{\partial^2}{\partial\varphi^2}\right) \tag{53a}$$

$$L_{12}^0 = -\frac{1}{1-\nu^2}\left(\frac{1+\nu}{2}\frac{1}{r}\frac{\partial^2}{\partial r\,\partial\varphi} - \frac{3-\nu}{2}\frac{1}{r^2}\frac{\partial}{\partial\varphi}\right) - 2\Omega^2\frac{\partial}{\partial\varphi} \tag{53b}$$

$$L_{21}^0 = -\frac{1}{1-\nu^2}\left(\frac{1+\nu}{2}\frac{1}{r}\frac{\partial^2}{\partial r\,\partial\varphi} + \frac{3-\nu}{2}\frac{1}{r^2}\frac{\partial}{\partial\varphi}\right) + 2\Omega^2\frac{\partial}{\partial\varphi} \tag{53c}$$

$$L_{22}^0 = -\frac{1}{1-\nu^2}\left(\frac{1-\nu}{2}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2}\right) - \Omega^2\left(1 - \frac{\partial^2}{\partial\varphi^2}\right) \tag{53d}$$

and the natural boundary conditions read

$$\boldsymbol{U}_1(\boldsymbol{w}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} u(r,t) \\ v(r,t) \end{bmatrix}\Bigg|_{r=r_i} = \boldsymbol{0} \tag{54a}$$

$$\boldsymbol{U}_2(\boldsymbol{w}) = \begin{bmatrix} \dfrac{\partial}{\partial r} + \nu\dfrac{1}{r} & \dfrac{1}{r}\dfrac{\partial}{\partial\varphi} \\ \dfrac{1}{r}\dfrac{\partial}{\partial\varphi} & \dfrac{\partial}{\partial r} - \dfrac{1}{r} \end{bmatrix}\begin{bmatrix} u(r,t) \\ v(r,t) \end{bmatrix}\Bigg|_{r=r_o} = \boldsymbol{0} \tag{54b}$$

We assume that the parameters δ and γ are smooth functions of a parameter ε. This corresponds to a variation along a smooth curve parametrized by ε in the parameter space [14,15]. It is possible to expand $\delta(\varepsilon)$ and $\gamma(\varepsilon)$ around $\varepsilon=0$ assuming that $\delta(0)=\gamma(0)=0$, for example,

$$\gamma(\varepsilon) = \gamma_{,\varepsilon}(0)\varepsilon + \cdots = \gamma_1\varepsilon + \cdots \tag{55}$$

$$\delta(\varepsilon) = \delta_{,\varepsilon}(0)\varepsilon + \cdots = \delta_1\varepsilon + \cdots \tag{56}$$

Assuming this kind of perturbation we write the perturbed boundary value problem up to first order in ε as

$$\boldsymbol{L}(\boldsymbol{w}) + \varepsilon\boldsymbol{L}^{1\varepsilon}(w) = 0 \tag{57}$$

$$\boldsymbol{U}_i(\boldsymbol{w}) = 0 \tag{58}$$

where $\boldsymbol{L}^{1\varepsilon} = \gamma_1\boldsymbol{L}^{1\gamma} + \delta_1\boldsymbol{L}^{1\delta}$.

According to Refs. [18,15,14], for a simple eigenvalue λ_0 and the corresponding eigenfunction $\mathbf{u}$, we set

$$\boldsymbol{w} = \boldsymbol{u} + \varepsilon\boldsymbol{w}_1^\varepsilon + \cdots \tag{59}$$

$$\lambda = \lambda_0 + \varepsilon\lambda_1^\varepsilon + \cdots = \lambda_0 + \varepsilon(\gamma_1\lambda_1^\gamma + \delta_1\lambda_1^\delta) + \cdots \tag{60}$$

Substitution into Eq. (57), taking the scalar product with the eigenfunction $\boldsymbol{v}$ of the adjoint of the unperturbed problem, and collecting the terms linear in ε yield

$$\lambda_1^\varepsilon = -\frac{\langle\boldsymbol{L}^{1\varepsilon}(\boldsymbol{u}),\boldsymbol{v}\rangle}{\left\langle\dfrac{\partial\boldsymbol{L}^0}{\partial\lambda}(\boldsymbol{u}),\boldsymbol{v}\right\rangle} \tag{61}$$

where $\langle\boldsymbol{u},\boldsymbol{v}\rangle = \int_A \boldsymbol{v}^*\boldsymbol{u}\,dA$, the asterisk denotes complex conjugate transpose.

For a semisimple eigenvalue λ_0 of the unperturbed problem, a double eigenvalue with two linearly independent (matrix) eigenfunctions, with corresponding eigenfunctions $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ substituting $\boldsymbol{u}=\alpha_1\boldsymbol{u}_1+\alpha_2\boldsymbol{u}_2$ in Eq. (59) and collecting terms linear in ε yields [15]

$$\boldsymbol{L}^{1\varepsilon}(\alpha_1\boldsymbol{u}_1+\alpha_2\boldsymbol{u}_2)+\lambda_1^{\varepsilon}\frac{\partial \boldsymbol{L}^0}{\partial\lambda_0}(\alpha_1\boldsymbol{u}_1+\alpha_2\boldsymbol{u}_2)=\boldsymbol{0} \tag{62}$$

The scalar product with the corresponding two eigenfunctions $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ of the adjoint of the unperturbed problem yields

$$\begin{bmatrix} \langle \boldsymbol{L}^{1\varepsilon}(\boldsymbol{u}_1),\boldsymbol{v}_1\rangle+\lambda_1^{\varepsilon}\left\langle \frac{\partial \boldsymbol{L}^0}{\partial\lambda}(\boldsymbol{u}_1),\boldsymbol{v}_1\right\rangle & \langle \boldsymbol{L}^{1\varepsilon}(\boldsymbol{u}_1),\boldsymbol{v}_2\rangle+\lambda_1^{\varepsilon}\left\langle \frac{\partial \boldsymbol{L}^0}{\partial\lambda}(\boldsymbol{u}_1),\boldsymbol{v}_2\right\rangle \\ \langle \boldsymbol{L}^{1\varepsilon}(\boldsymbol{u}_2),\boldsymbol{v}_1\rangle+\lambda_1^{\varepsilon}\left\langle \frac{\partial \boldsymbol{L}^0}{\partial\lambda}(\boldsymbol{u}_2),\boldsymbol{v}_1\right\rangle & \langle \boldsymbol{L}^{1\varepsilon}(\boldsymbol{u}_2),\boldsymbol{v}_2\rangle+\lambda_1^{\varepsilon}\left\langle \frac{\partial \boldsymbol{L}^0}{\partial\lambda}(\boldsymbol{u}_2),\boldsymbol{v}_2\right\rangle \end{bmatrix}\begin{bmatrix}\alpha_1\\ \alpha_2\end{bmatrix}=\boldsymbol{0} \tag{63}$$

For nontrivial solutions in α_1 and α_2, the determinant of the matrix has to vanish, which yields λ_1^{ε}. An analogous formula describing the splitting of a semisimple eigenvalue of multiplicity r

$$\det\left[\langle \boldsymbol{L}^{1\varepsilon}\boldsymbol{u}_k,\boldsymbol{v}_j\rangle+\lambda_1^{\varepsilon}\left\langle \frac{\partial \boldsymbol{L}^0}{\partial\lambda}\boldsymbol{u}_k,\boldsymbol{v}_j\right\rangle\right]=0,\quad j,k=1,\ldots,r \tag{64}$$

generalizes the result derived for scalar differential operators in Ref. [15] to the case of operator matrices. We note that the approach based on the perturbation theory of multiple eigenvalues [19] differs from that used in Ref. [20] and allows one to study the phenomena of veering and overlapping of eigenvalue branches with the use of operator derivatives and eigenvectors of a multiple eigenvalue calculated only for the values of parameters corresponding to coalescence of eigenvalues.

4.1 Spectrum of the Unperturbed Problem

4.1.1 Plate. We investigate the unperturbed boundary eigenvalue problem of the plate

$$L^0(w)=0 \tag{65}$$

$$U_i(w)=0,\quad i=1,\ldots,4 \tag{66}$$

and expand w in terms of the eigenfunctions of the corresponding nonrotating plate, that is, $\Omega=0$,

$$w(r,\varphi,t)=\sum_{i=0}^{\infty}\left(W_i^c(r,\varphi)q_i^c(t)+W_i^s(r,\varphi)q_i^s(t)\right) \tag{67}$$

The boundary eigenvalue problem obtained after separation of time for the nonrotating annular plate in dimensionless form reads

$$\nabla^4 W(r,\varphi)=\frac{h}{D}\omega_{mn}^2 W(r,\varphi) \tag{68}$$

$$W(r_i,\varphi)=0,\quad W_{,r}(r_i,\varphi)=0 \tag{69}$$

$$M_r(r_o,\varphi)=0,\quad V_r(r_o,\varphi)=0 \tag{70}$$

where V_r and M_r are given in Eqs. (39*b*) and (39*e*), respectively.

According to Ref. [21], the eigenfunctions are

$$W_i^c(r,\varphi)=R_{mn}(r)\cos m\varphi,\quad m=0,\ldots,\infty,\quad n=1,\ldots,\infty$$

$$W_i^s(r,\varphi)=R_{mn}(r)\sin m\varphi,\quad m=\infty,\ldots,0,\quad n=\infty,\ldots,1$$

where the $R_{mn}(r)=C_1I(\beta_{mn}r)+C_2J(\beta_{mn}r)+C_3Y(\beta_{mn}r)+C_4K(\beta_{mn}r)$ are linear combinations of Bessel and modified Bessel functions, and the constants $C_1,\ldots,C_4$ are determined from the linear equation after substitution of the ansatz into the boundary conditions (69) and (70). The quantities $\beta_{mn}=\sqrt[4]{(h/D)\omega_{mn}^2}$ are the roots of the characteristic equation resulting from equating the determinant of the matrix of the linear system to zero. The numbers (m,n) denote the number of nodal diameters and nodal circles of the plate. Since the plate is clamped at the inner radius, numeration of nodes n starts from $n=1$. The functions $W_i^c(r,\varphi)$ corresponding to $m=0$ are umbrella modes of the plate. The eigenfunctions for the nonrotating annular plate are known to form a complete orthogonal base for the problem of the rotating plate. We now project the problem of the rotating plate onto the eigenfunctions of the nonrotating plate, i.e., we write

$$\int_A L^0(w)W_j^c\mathrm{d}A=\int_A L^0\left(\sum_{i=0}^{\infty}\left(W_i^cq_i^c+W_i^sq_i^s\right)\right)W_i^c\mathrm{d}A=0 \tag{71}$$

$$\int_A L^0(w)W_j^s\mathrm{d}A=\int_A L^0\left(\sum_{i=0}^{\infty}\left(W_i^cq_i^c+W_i^sq_i^s\right)\right)W_i^s\mathrm{d}A=0 \tag{72}$$

Equations (71) and (72) form an infinite dimensional matrix equation of the type

$$(\boldsymbol{M}\lambda^2+\boldsymbol{G}\lambda+\boldsymbol{K})\boldsymbol{q}=\boldsymbol{0},\quad \boldsymbol{q}=[q_1^c,\ldots,q_\infty^c,q_\infty^s,\ldots,q_1^s]^T$$

Using this and the orthogonality relations for the trigonometric functions sin and cos and the orthogonality of the eigenfunctions of the nonrotating plate we find that the matrices of the unperturbed problem read

$$\boldsymbol{M}=\mathrm{diag}(M_{mn}) \tag{73}$$

$$\boldsymbol{G}=\mathrm{antidiag}(2mM_{mn}\Omega,-2mM_{mn}\Omega) \tag{74}$$

$$\boldsymbol{K}=\mathrm{diag}(\omega_{mn}^2M_{mn}-m^2M_{mn}\Omega^2) \tag{75}$$

where $M_{mn}=h\pi\int_{r_i}^{r_o}R_{mn}^2(r)r\mathrm{d}r$. Consequently, the equations of motion of the unperturbed problem decouple into infinitely many pairs of two coupled equations of the form

$$\begin{bmatrix}1&0\\0&1\end{bmatrix}\ddot{\boldsymbol{q}}+\begin{bmatrix}0&-2m\Omega\\2m\Omega&0\end{bmatrix}\dot{\boldsymbol{q}}+\begin{bmatrix}\omega_{mn}^2-m^2\Omega^2&0\\0&\omega_{mn}^2-m^2\Omega^2\end{bmatrix},$$

$$\boldsymbol{q}=0,\quad m>0 \tag{76}$$

where we divided it by M_{mn}. The spectrum of the unperturbed problem, namely, the eigenvalues of Eq. (76), can be calculated analytically as

$$\lambda^{1\pm}=\pm i(\omega_{mn}+m\Omega) \tag{77a}$$

$$\lambda^{2\pm}=\pm i(\omega_{mn}-m\Omega) \tag{77b}$$

and form the geometric structure called a spectral mesh in Ref. [22] depicted in Fig. 5 for the dimensionless parameters

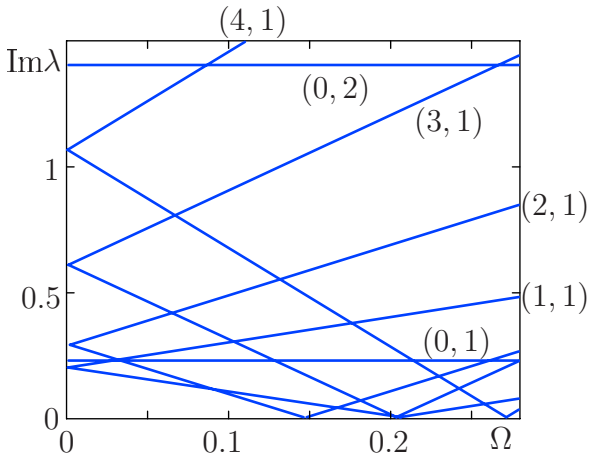


Fig. 5 Spectrum of the out-of-plane vibrations of the unperturbed problem (plate) for r_i=0.154 (spectral mesh)

$$h = 0.1605, \quad D = 3.7858 \times 10^{-4}, \quad r_i = 0.154, \quad \nu = 0.3 \tag{78}$$

which are based on the parameters for a disk brake used in Ref. [13]. They correspond to a brake rotor represented by an equivalent Kirchhoff plate with r_o=0.162 m, E=4.16×10^{10} N/m, and ρ=4846 kg/m^3. These parameters were identified in Ref. [23] at the test rig. For numerical calculations we use the parameters of the corresponding disk brake throughout the paper. We emphasize at this point that for the problem of brake squeal only angular velocities up to 4000 rpm are physically relevant, which corresponds to a dimensionless $\Omega = 2.3 \times 10^{-2}$. In other applications much larger dimensionless angular velocities arise.

The horizontal lines in Fig. 5 for m=0 are umbrella modes of the plate. The spectrum of the reduced two-dimensional system corresponding to $m > 0$ is simple, except at the critical velocities

$$\Omega_1^c = 0, \quad \lambda_1^{c\pm} = \pm i\omega_{mn}$$

$$\Omega_2^c = \frac{1}{m}\omega_{mn}, \quad \lambda_2^{c\pm} = 0$$

where it is seen to be semisimple by calculating the corresponding eigenvectors. Note that additional crossings of eigenvalues occurring in the spectrum are semisimple as well, which follows from the decoupling of the matrices. Qualitatively it agrees with other results obtained in literature especially with the observation made in Ref. [24].

4.1.2 Disk. The spectrum for the unperturbed problem of the disk was calculated by Chen and Jhu in Ref. [16] using Lame potentials. Since it is numerically difficult to solve the characteristic equation, in this paper we prefer to calculate the spectrum using a Ritz discretization approach. The results obtained with the Ritz method have been validated by comparison to the results in Refs. [25,16]. The energy expressions from Eq. (19) corresponding to the disk equations are discretized using

$$u(r,\varphi) = \sum_{m=0}^{\hat{m}} \sum_{n=1}^{\hat{n}} \left[\sin\left(\frac{n\pi(r-r_i)}{2(r_o-r_i)} \right) \cos m\varphi \right] q_{mn}^{uc}(t) + \sum_{m=1}^{\hat{m}} \sum_{n=1}^{\hat{n}} \left[\sin\left(\frac{n\pi(r-r_i)}{2(r_o-r_i)} \right) \sin m\varphi \right] q_{mn}^{us}(t) \tag{79}$$

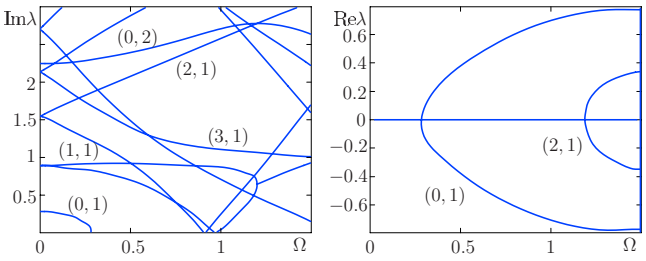


Fig. 6 Spectrum of the in-plane vibrations of the unperturbed problem (disk) for r_i=0.153

$$v(r,\varphi) = \sum_{m=0}^{\hat{m}} \sum_{n=1}^{\hat{n}} \left[\sin\left(\frac{n\pi(r-r_i)}{2(r_o-r_i)} \right) \cos m\varphi \right] q_{mn}^{vc}(t) + \sum_{m=1}^{\hat{m}} \sum_{n=1}^{\hat{n}} \left[\sin\left(\frac{n\pi(r-r_i)}{2(r_o-r_i)} \right) \sin m\varphi \right] q_{mn}^{vs}(t) \tag{80}$$

which can be seen to form a base for the problem, since they satisfy the geometric boundary conditions and since the trigonometric functions are known to be complete. Carrying out the variations in Eq. (19) over the functions q_{mn}^{ij}, we obtain the equations of motion

$$\boldsymbol{M}\ddot{\boldsymbol{q}} + \boldsymbol{G}\dot{\boldsymbol{q}} + \boldsymbol{K}\boldsymbol{q} = \boldsymbol{0} \tag{81}$$

of the unperturbed problem, where $\boldsymbol{M}$ and $\boldsymbol{K}$ are symmetric and $\boldsymbol{G}$ is skew-symmetric. The corresponding eigenvalue problem reads

$$[\boldsymbol{M}\lambda^2 + \boldsymbol{G}\lambda + \boldsymbol{K}] \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = \boldsymbol{0} \tag{82}$$

where

$$\boldsymbol{u} = [q_{01}^{uc}, q_{01}^{us}, \ldots, q_{\hat{m}\hat{n}}^{uc}, q_{\hat{m}\hat{n}}^{us}]^T \tag{83}$$

$$\boldsymbol{v} = [q_{01}^{vc}, q_{01}^{vs}, \ldots, q_{\hat{m}\hat{n}}^{vc}, q_{\hat{m}\hat{n}}^{vs}]^T \tag{84}$$

and yields the spectrum depicted in Fig. 6.

At Ω=0.28 the (0,1) eigenform of the system loses stability by divergence (we use again the notation (m,n), where as for the plate m denotes the number of nodal diameters and n denotes the number of nodal circles). The system exhibits flutter at Ω=1.2.

4.2 Perturbation Formulas. For the following perturbation formulas we introduce the dimensionless parameters

$$\bar{r}_{pi} = 0.5988, \quad \bar{r}_{po} = 0.9444, \quad \varphi_p = 0.7$$

$$\bar{k} = 0.0093, \quad \bar{N}_0 = 6.4669 \times 10^{-6}, \quad \mu = 0.6$$

corresponding to a real disk brake [26,13]. We again skip the bar in the following for notational convenience. The weight function corresponding to the pads is chosen such that $\kappa(r,\varphi)=1$ if $(r,\varphi) \in B$, i.e., if (r,φ) are in the domain of the pad (denoted by B to be distinguished from the whole area of the plate surface denoted earlier by A) and $\kappa(r,\varphi)=0$ otherwise.

4.2.1 Plate. Expanding the perturbed eigenfunctions of the plate in terms of the eigenfunctions of the nonrotating plate and projection onto themselves yields

$$[\boldsymbol{M}\lambda^2 + (\boldsymbol{G} + \gamma\Delta\boldsymbol{G}^{\gamma} + \delta\Delta\boldsymbol{G}^{\delta})\lambda + \boldsymbol{K} + \gamma\Delta\boldsymbol{K}^{\gamma} + \delta\Delta\boldsymbol{K}^{\delta}]\boldsymbol{u} = 0 \tag{85}$$

where $\boldsymbol{M}$, $\boldsymbol{G}$, and $\boldsymbol{K}$ have been derived in Sec. 4.1.1 and

$$\Delta G_{ij}^{\gamma} = 0 \tag{86a}$$

$$\Delta G_{ij}^{\delta} = \int_B \frac{h^2 N_0 \mu}{2r\Omega} W^i_{,r} W^j_{,r} \mathrm{d}B \tag{86b}$$

$$\Delta K_{ij}^{\gamma} = \int_B \left[2kW^i W^j - \frac{hk\mu}{r} W^i_{,\varphi} W^j \right] \mathrm{d}B \tag{86c}$$

$$\Delta K_{ij}^{\delta} = \int_B \left[\mu^2 \frac{hN_0}{r^2} W^i_{,\varphi} W^j_{,\varphi} - \frac{h^2 N_0 \mu}{2r^2} W^i_{,r} W^j_{,\varphi} + \frac{h^2 N_0 \mu}{2r} W^i_{,r} W^j_{,r\varphi} \right] \mathrm{d}B \tag{86d}$$

are the perturbation matrices. Note that we have used integration by parts in some of the integrals and thereby eliminated transition terms originating from the discontinuity of the weight function $\kappa(r,\varphi)$. The matrices (86) are similar to the expressions one would have obtained from the Ritz discretization used in Ref. [13] taking into account an arbitrary number of shape functions. However, due to the reasons discussed in Sec. 3.4, the expressions obtained from the Ritz discretization contain a few more terms in ΔK_{ij}^{γ} and ΔK_{ij}^{δ}, which read

$$\Delta K_{ij}^{\gamma} = \int_B \left[2kW^i W^j - \frac{hk\mu}{r} W^i_{,\varphi} W^j \right] \mathrm{d}B \tag{87}$$

$$\Delta K_{ij}^{\delta} = \int_B \left[(1+\mu^2) \frac{hN_0}{r^2} W^i_{,\varphi} W^j_{,\varphi} - \frac{h^2 N_0 \mu}{2r^3} W^i_{,\varphi\varphi} W^j_{,\varphi} + \frac{h^2 N_0 \mu}{2r^2} (W^i_{,\varphi} W^j_{,r} - W^i_{,r} W^j_{,\varphi}) + hN_0 W^i_{,r} W^j_{,r} + \frac{h^2 N_0 \mu}{2r} (W^i_{,r} W^j_{,r\varphi} - W^i_{,\varphi r} W^j_{,r}) \right] \mathrm{d}B \tag{88}$$

It can be seen numerically that the additional terms are so small that they have practically no influence on the results.

The perturbation matrices are split into a symmetric and a skew-symmetric part

$$\boldsymbol{\Delta G}^{\gamma,\delta} = \boldsymbol{G}^{\gamma,\delta} + \boldsymbol{D}^{\gamma,\delta}, \quad \boldsymbol{G}^{\gamma,\delta} = -(\boldsymbol{G}^{\gamma,\delta})^T, \quad \boldsymbol{D}^{\gamma,\delta} = (\boldsymbol{D}^{\gamma,\delta})^T$$

$$\boldsymbol{\Delta K}^{\gamma,\delta} = \boldsymbol{K}^{\gamma,\delta} + \boldsymbol{N}^{\gamma,\delta}, \quad \boldsymbol{K}^{\gamma,\delta} = (\boldsymbol{K}^{\gamma,\delta})^T, \quad \boldsymbol{N}^{\gamma,\delta} = -(\boldsymbol{N}^{\gamma,\delta})^T$$

Using formula (61) for simple eigenvalues, we arrive at the expression

$$\lambda_1^{\gamma} = \frac{-\lambda_0 \bar{\boldsymbol{u}}^T \boldsymbol{\Delta G}^{\gamma} \boldsymbol{u} - \bar{\boldsymbol{u}}^T \boldsymbol{\Delta K}^{\gamma} \boldsymbol{u}}{2\lambda_0 \bar{\boldsymbol{u}}^T \boldsymbol{M} \boldsymbol{u} + \bar{\boldsymbol{u}}^T \boldsymbol{G} \boldsymbol{u}} = \frac{\omega^2 \bar{\boldsymbol{u}}^T \boldsymbol{\Delta G}^{\gamma} \boldsymbol{u} - i\omega \bar{\boldsymbol{u}}^T \boldsymbol{\Delta K}^{\gamma} \boldsymbol{u}}{\bar{\boldsymbol{u}}^T (-\boldsymbol{M}\omega^2 - \boldsymbol{K}) \boldsymbol{u}} = \frac{\omega^2 (\boldsymbol{a}^T \boldsymbol{D}^{\gamma} \boldsymbol{a} + \boldsymbol{b}^T \boldsymbol{D}^{\gamma} \boldsymbol{b}) + 2\omega \boldsymbol{a}^T \boldsymbol{N}^{\gamma} \boldsymbol{b} + i[2\omega^2 \boldsymbol{a}^T \boldsymbol{G}^{\gamma} \boldsymbol{b} + \omega(\boldsymbol{a}^T \boldsymbol{K}^{\gamma} \boldsymbol{a} + \boldsymbol{b}^T \boldsymbol{K}^{\gamma} \boldsymbol{b})]}{\boldsymbol{a}^T (-\boldsymbol{M}\omega^2 - \boldsymbol{K}) \boldsymbol{a} + \boldsymbol{b}^T (-\boldsymbol{M}\omega^2 - \boldsymbol{K}) \boldsymbol{b}} \tag{89}$$

where a and b are the real and imaginary parts of the eigenvector $\boldsymbol{u}_0$. To calculate λ_1^{δ} just replace γ with δ. From Eq. (89) we observe that for small γ and δ the stability behavior is determined by $\boldsymbol{D}^{\gamma,\delta}$ and $\boldsymbol{N}^{\gamma,\delta}$ only. Due to the decoupling of the equations for the unperturbed problem, only the 2×2 matrices having the same position in the perturbation matrices as the blocks corresponding to the eigenvalue of the unperturbed problem are relevant for the first term in the expansion of λ. The zeros in the corresponding eigenvector of the unperturbed problem suppress the influence of other terms. For the semisimple eigenvalues, more terms are relevant for the splitting of the double eigenvalue. From Eq. (63) it follows that λ_1^{ε} is determined from

$$\det \begin{bmatrix} c_1 \lambda_1^{\varepsilon} + a & b \\ c & c_2 \lambda_1^{\varepsilon} + d \end{bmatrix} = 0 \tag{90}$$

where

$$c_1 = 2\lambda_0 \bar{\boldsymbol{u}}_1^T \boldsymbol{M} \boldsymbol{u}_1 + \bar{\boldsymbol{u}}_1^T \boldsymbol{G} \boldsymbol{u}_1 \tag{91a}$$

$$c_2 = 2\lambda_0 \bar{\boldsymbol{u}}_2^T \boldsymbol{M} \boldsymbol{u}_2 + \bar{\boldsymbol{u}}_2^T \boldsymbol{G} \boldsymbol{u}_2 \tag{91b}$$

$$a = -\lambda_0 \bar{\boldsymbol{u}}_1^T \boldsymbol{\Delta G}^{\varepsilon} \boldsymbol{u}_1 - \bar{\boldsymbol{u}}_1^T \boldsymbol{\Delta K}^{\varepsilon} \boldsymbol{u}_1 \tag{91c}$$

$$b = -\lambda_0 \bar{\boldsymbol{u}}_1^T \boldsymbol{\Delta G}^{\varepsilon} \boldsymbol{u}_2 - \bar{\boldsymbol{u}}_1^T \boldsymbol{\Delta K}^{\varepsilon} \boldsymbol{u}_2 \tag{91d}$$

$$c = -\lambda_0 \bar{\boldsymbol{u}}_2^T \boldsymbol{\Delta G}^{\varepsilon} \boldsymbol{u}_1 - \bar{\boldsymbol{u}}_2^T \boldsymbol{\Delta K}^{\varepsilon} \boldsymbol{u}_1 \tag{91e}$$

$$d = -\lambda_0 \bar{\boldsymbol{u}}_2^T \boldsymbol{\Delta G}^{\varepsilon} \boldsymbol{u}_2 - \bar{\boldsymbol{u}}_2^T \boldsymbol{\Delta K}^{\varepsilon} \boldsymbol{u}_2 \tag{91f}$$

We obtain

$$\lambda_1^{\varepsilon} = \frac{-(dc_1 + ac_2) \pm \sqrt{(dc_1 - ac_2)^2 + 4c_1 c_2 cb}}{2c_1 c_2} \tag{92}$$

which shows that in the generic case, when the expression under the square root does not vanish, avoided crossings of imaginary parts and imperfect merging of modes occur as depicted in Fig. 7, where we used the $\delta=3$ and $\gamma=3$ for the pads. Since

$$\frac{\|\Delta \boldsymbol{K}^{\gamma}\|_F}{\|\boldsymbol{K}_0\|_F} = 0,\ 0491, \quad \frac{\|\Delta \boldsymbol{G}^{\gamma}\|_F}{\|\boldsymbol{K}_0\|_F} = 0, \quad \frac{\|\Delta \boldsymbol{K}^{\delta}\|_F}{\|\boldsymbol{K}_0\|_F} = 1.618 \times 10^{-6}$$

$$\frac{\|\Delta \boldsymbol{G}^{\delta}\|_F}{\|\boldsymbol{K}_0\|_F} = \frac{1.306 \times 10^{-6}}{\Omega}, \quad \boldsymbol{K}_0 = \boldsymbol{K}|_{\Omega=0}$$

where $\|\cdot\|_F$ stands for the Frobenius norm, even the perturbations $\gamma\Delta\boldsymbol{K}^{\gamma} + \delta\Delta\boldsymbol{K}^{\delta}$ and $\gamma\Delta\boldsymbol{G}^{\gamma} + \delta\Delta\boldsymbol{G}^{\delta}$ can be considered small for speed ranges not too close to $\Omega=0$ (cf. also Fig. 11).

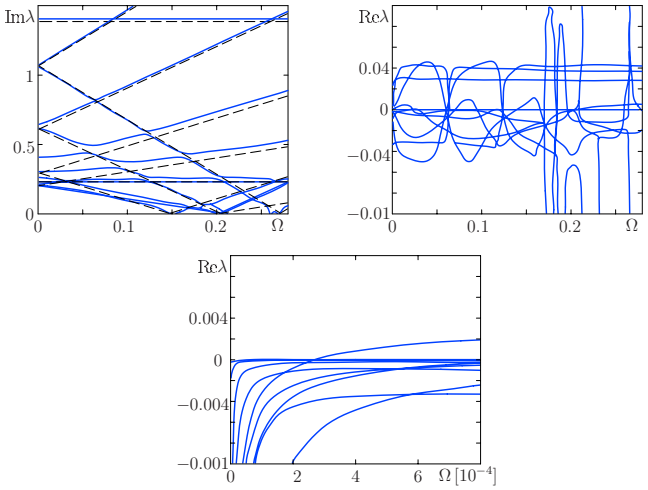


Fig. 7 Spectrum of the perturbed plate problem for $\gamma=\delta=3$. Dashed lines: unperturbed problem; lower plot: zoom for small Ω.

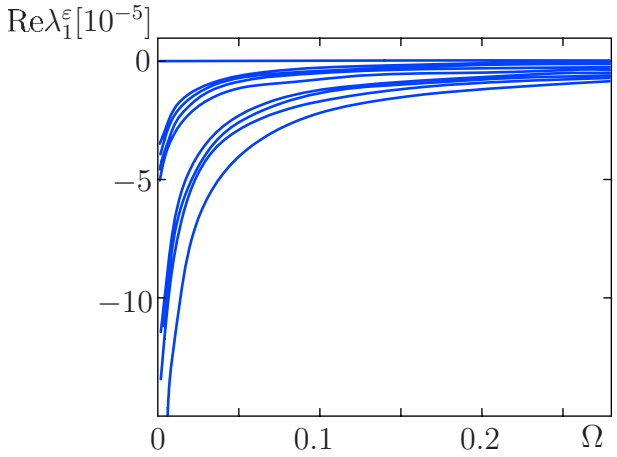


Fig. 8 λ_1^ε for the perturbed disk problem for r_i=0.153

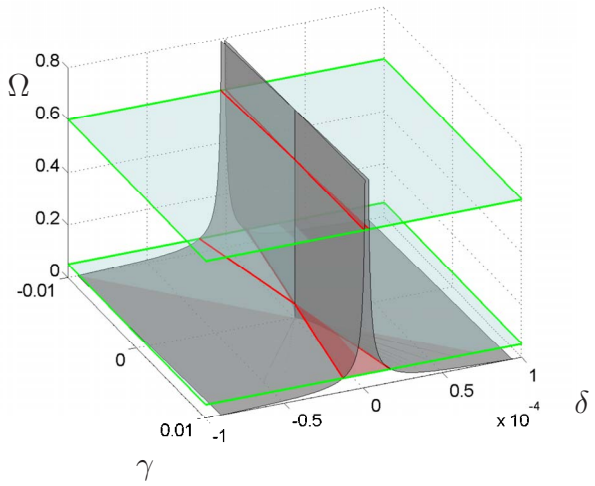


Fig. 9 Three-dimensional stability boundaries

Generic pictures for avoided crossings have been derived in Ref. [27] for the matrix case. In our case they look similar, except for the fact that we have a vanishing real part for all eigenvalues in the unperturbed problem. We see that the combined action of the dissipative and nonconservative positional forces moves some of the eigenvalues to the right side of the complex plane for the values of Ω not exceeding the first critical speed. For the parameters considered in Fig. 7 the system gets unstable at a dimensionless $\Omega=2.7\times10^{-4}$, which for the physical disk brake studied corresponds to an angular velocity of about 46 rpm. Therefore, the application of the brake pads to the disk can cause flutter instability in the subcritical range, which is typical for squeal.

4.2.2 Disk. According to Eq. (61), for simple eigenvalues of the unperturbed problem the first term in the expansion (60) reads

$$\lambda_1^\varepsilon = \frac{\displaystyle\int_B \left[\frac{2\mu N_0 h}{r\Omega}\bar{u}u + \frac{2\mu N_0 h}{r}\bar{u}u - \frac{2\mu N_0 h}{r}\bar{u}v\right]\mathrm{d}B}{\displaystyle\int_A [2\lambda(u\bar{u}+v\bar{v}) + 2\Omega[\bar{u}(u_{,\varphi}-v)+\bar{v}(u+v_{,\varphi})]]\mathrm{d}A} \tag{93}$$

Substituting the eigenfunctions obtained from the Ritz method $u=\boldsymbol{U}\boldsymbol{u}$ and $v=\boldsymbol{V}\boldsymbol{v}$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are row matrices containing the shape functions for u and v, and $[\boldsymbol{u}^T,\boldsymbol{v}^T]^T$ is the eigenvector corresponding to Eq. (82), yields

$$\lambda_1^\varepsilon = \frac{-\lambda_0\begin{bmatrix}\bar{\boldsymbol{u}}^T & \bar{\boldsymbol{v}}^T\end{bmatrix}\boldsymbol{\Delta G}^\delta\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{v}\end{bmatrix} - \begin{bmatrix}\bar{\boldsymbol{u}}^T & \bar{\boldsymbol{v}}^T\end{bmatrix}\boldsymbol{\Delta K}^\delta\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{v}\end{bmatrix}}{2\lambda_0\begin{bmatrix}\bar{\boldsymbol{u}}^T & \bar{\boldsymbol{v}}^T\end{bmatrix}\boldsymbol{M}\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{v}\end{bmatrix} + \begin{bmatrix}\bar{\boldsymbol{u}}^T & \bar{\boldsymbol{v}}^T\end{bmatrix}\boldsymbol{G}\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{v}\end{bmatrix}} \tag{94}$$

which coincides with the perturbation formulas one would have obtained by perturbing the discretized eigenvalue problem (82). The first correction terms from Eq. (60) for simple eigenvalues of the perturbed disk problem are shown in Fig. 8 for the real part.

We observe that in the range below the first critical speed $\mathrm{Re}(\lambda_1^\varepsilon)$ is strictly negative. For $\Omega\to0$ the real parts become infinite, which can be explained by the term proportional to $1/\Omega$ in the damping matrix. For the semisimple eigenvalues we expect the same behavior, since the eigenvalues of a matrix polynomial depend continuously on the matrix entries.

4.3 Stability Boundaries. In Sec. 4.2.1 we derived formulas for the change in simple and semisimple eigenvalues occurring in the spectrum of the unperturbed plate problem caused by small changes in the parameters. For a fixed rotational speed of the plate, the stable region in the parameter plane γ,δ is given by those areas where all eigenvalues of the problem have a nonpositive real part. For each simple purely imaginary eigenvalue of the unperturbed problem λ_j there is a stable region, which in the first approximation is the half plane

$$\gamma\,\mathrm{Re}(\lambda_1^{j\gamma}) + \delta\,\mathrm{Re}(\lambda_1^{j\delta}) \le 0, \quad \forall\, j \tag{95}$$

At a fixed rotational speed Ω, where all eigenvalues are simple, the stability region in the γ,δ plane is given by the intersection of the half-planes defined in Eq. (95). Depending on the parameters, it can be a sector limited by an angle, a line (for $\mathrm{Re}(\lambda_1^{1\gamma})/\mathrm{Re}(\lambda_1^{1\delta})=\cdots=\mathrm{Re}(\lambda_1^{n\gamma})/\mathrm{Re}(\lambda_1^{n\delta})$) or just the point $\gamma=\delta=0$. Approximations to the stability boundary therefore coincide with the lines

$$\gamma = -\frac{\mathrm{Re}(\lambda_1^{1\delta})}{\mathrm{Re}(\lambda_1^{1\gamma})}\delta = -\frac{\omega(\boldsymbol{a}^T\boldsymbol{D}^\delta\boldsymbol{a}+\boldsymbol{b}^T\boldsymbol{D}^\delta\boldsymbol{b})+2\boldsymbol{a}^T\boldsymbol{N}^\delta\boldsymbol{b}}{\omega(\boldsymbol{a}^T\boldsymbol{D}^\gamma\boldsymbol{a}+\boldsymbol{b}^T\mathbf{D}^\gamma\boldsymbol{b})+2\boldsymbol{a}^T\boldsymbol{N}^\gamma\boldsymbol{b}}\delta \tag{96}$$

Again, recall $\boldsymbol{u}=\boldsymbol{a}+i\boldsymbol{b}$. For the two-dimensional system of the plate corresponding to its (3,1) mode, the stability boundaries are depicted in Fig. 9, where only the areas with positive γ and δ are physically meaningful.

From Fig. 9 one expects that for $\Omega\to0$ the stable region coincides with the half plane $\Omega=0,\delta>0$ and for $\Omega\to\infty$ the stable region is a line. It is not difficult to verify this analytically.

Consider the case $\Omega\to0$. In this case $d_{11}^\delta\to\infty$ and consequently in Eq. (96) the numerator tends to infinity since always two eigenvectors of the unperturbed system with nonvanishing first component of $\boldsymbol{a}$ or $\boldsymbol{b}$ can be found (the eigenvectors of a purely gyroscopic system span the solution space provided that $\boldsymbol{M}$ and $\boldsymbol{K}$ are positive definite). The stability boundary is therefore given by $\delta=0$.

Now consider the case $\Omega\to\infty$. We will show that in the limit the eigenvectors corresponding to the pairs of eigenvalues (76) coincide and Eq. (96) yields the same expression for each eigenvalue. We note that for increasing Ω the relative difference between λ^1 and λ^2 tends to zero, i.e.,

$$\lim_{\Omega\to\infty}\Delta\lambda = \lim_{\Omega\to\infty}\frac{\lambda^1-\lambda^2}{\lambda^1} = 0 \tag{97}$$

as we can see from Eq. (76). Consider the matrix $\boldsymbol{A}$ describing Eq. (76) written as a first order system with eigenvectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ corresponding to λ^1 and λ^2, i.e.,

$$\boldsymbol{A}\boldsymbol{v}_1 = \lambda^1\boldsymbol{v}_1 \tag{98}$$

$$\boldsymbol{A}\boldsymbol{v}_2 = \lambda^2\boldsymbol{v}_2 \tag{99}$$

Using Eq. (97) we obtain

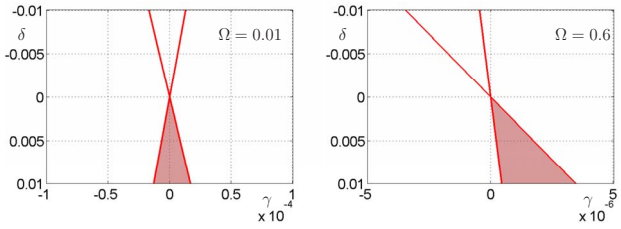


Fig. 10 Stability boundaries in the sub- and supercritical ranges

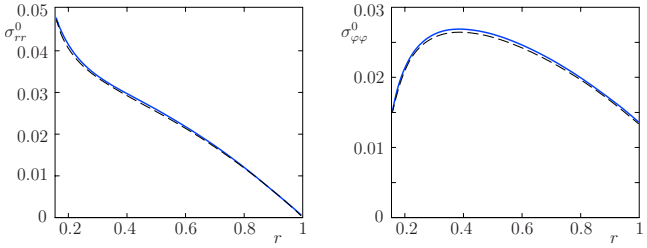


Fig. 12 Prestress due to rotation of the disk ($\Omega=0.28$)

$$\boldsymbol{A}(\boldsymbol{v}_1-\boldsymbol{v}_2)=\lambda^1(\boldsymbol{v}_1-\boldsymbol{v}_2)+\Delta\lambda\boldsymbol{v}_2 \tag{100}$$

$$\boldsymbol{A}(\boldsymbol{v}_2-\boldsymbol{v}_1)=\lambda^2\left(\boldsymbol{v}_2-\frac{1}{1-\Delta\lambda}\boldsymbol{v}_1\right) \tag{101}$$

which by addition and taking the limit $\Delta\lambda\to 0$ yields

$$(\lambda^1-\lambda^2)(\boldsymbol{v}_1-\boldsymbol{v}_2)=0 \tag{102}$$

and hence $\boldsymbol{v}_1=\boldsymbol{v}_2$.

From the above reasoning taking into account that the stable lines for different pairs of eigenvalues have different slopes we conclude that the system definitely becomes unstable for some Ω.

A second fact to be observed from Fig. 10 is that in the supercritical range damping can have a destabilizing effect, which agrees with the results of Ref. [28]. In view of the Thompson–Tait–Chetaev theorem [29] this effect is not surprising, although this theorem is not directly applicable, since the forces originating from the prestress N_0 are not purely dissipative.

We stress that the approximations for the stability boundaries are only valid for small perturbations. In Fig. 11 they are compared with numerical calculations of the eigenvalues at particular points in the parameter space. In the vicinity of the unperturbed problem they are seen to give good approximations.

4.4 The Effect of Centrifugal Stiffening in the Plate Equations. In this section we discuss how the effect of centrifugal stiffening enters into the plate vibrations and compare the results with the ones obtained in Secs. 4.2 and 4.3. Using the assumption that $u_{,x}$, $u_{,y}$, $v_{,x}$, and $v_{,y}$ occurring in Eq. (29) are of the order of magnitude of $w_{,x}^2$ and $w_{,y}^2$ and neglecting terms containing z^2 since the plate is thin [3] we obtain

$$e_{xx}=u_{,x}-zw_{,xx}+\tfrac{1}{2}w_{,x}^2 \tag{103a}$$

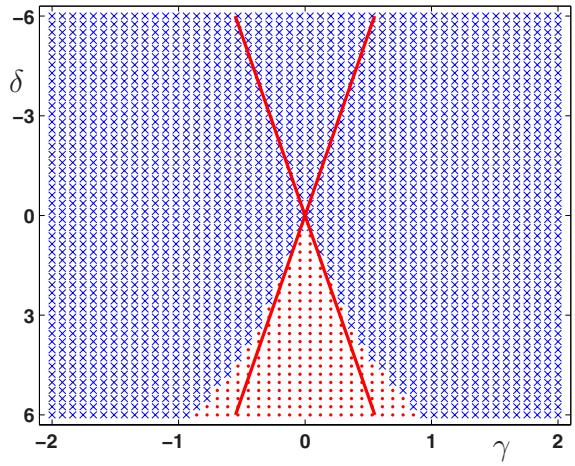


Fig. 11 Comparison of analytical and numerical results for $\Omega=10^{-3}$ (dot: stable; cross: unstable)

$$e_{yy}=v_{,y}-zw_{,yy}+\tfrac{1}{2}w_{,y}^2 \tag{103b}$$

$$e_{zz}=0 \tag{103c}$$

$$e_{xy}=\tfrac{1}{2}u_{,y}+\tfrac{1}{2}v_{,x}-zw_{,xy}+\tfrac{1}{2}w_{,y}w_{,x}, \tag{103d}$$

$$e_{yz}=0 \tag{103e}$$

$$e_{xz}=0 \tag{103f}$$

Deriving the equations of motion from the principle of virtual work under consideration of the prestress of the disk due to the rotation and transformation into polar coordinates after linearization yields

$$\rho h w_{,tt}+2h\rho\Omega w_{,t\varphi}+\Omega^2\rho h w_{,\varphi\varphi}-\frac{1}{r}\frac{\partial}{\partial r}(rh\sigma_{rr}^0 w_{,r})-\frac{h\sigma_{\varphi\varphi}^0}{r^2}w_{,\varphi\varphi}$$

$$+D\nabla^4 w=F_z-\frac{1}{r}T_{r,\varphi}+T_{\varphi,r} \tag{104}$$

where the boundary conditions are as previously given by Eq. (38) and σ_{rr}^0 and $\sigma_{\varphi\varphi}^0$ are the prestresses of the disk in radial and circumferential directions, which will be calculated in the sequel. We remark that due to the rotational symmetry of the plate $\sigma_{r\varphi}^0=0$.

4.4.1 Calculation of the Prestress. The prestress originating from the rotation is calculated by finding a stationary solution for Eq. (40) with $\kappa=0$ from which the strain can be calculated. The prestress can then be found from the stress-strain relations. Due to the symmetry of the disk it is clear that σ_{rr}^0 and $\sigma_{\varphi\varphi}^0$ depend on r only and that $v=0$. In dimensionless form the resulting differential equation reads

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+(1-\nu^2)\Omega^2-\frac{1}{r^2}\right)u=-(1-\nu^2)\Omega^2 r \tag{105}$$

with boundary conditions

$$u|_{r_i}=0,\quad \left(\frac{\partial}{\partial r}+\nu\frac{1}{r}\right)u|_{r=r_o}=0 \tag{106}$$

The boundary value problem for the ordinary differential equation (105) with boundary condition (106) coincides with the one obtained in Ref. [30] by linearization of the boundary value problem obtained using the nonlinear deformation gradient. The general solution of Eq. (105) is given by

$$u(r)=C_1J_1(r\sqrt{\Omega^2(1-\nu^2)})+C_2Y_1(r\sqrt{\Omega^2(1-\nu^2)})-r \tag{107}$$

as derived in Ref. [30] where J_1 and Y_1 are the well known Bessel functions of the first and second kinds. The constants C_1 and C_2 are obtained by adjusting Eq. (107) to the boundary condition (106). For the dimensionless angular velocity $\Omega=0.28$ the resulting $\sigma_{rr}^0(r)$ and $\sigma_{\phi\phi}^0(r)$ are plotted in Fig. 12 as solid lines.

For small angular velocities Ω the term proportional to Ω^2 in Eq. (105) can be neglected. The corresponding result was derived in Ref. [31] using first order theory, where also the solution for the

simplified boundary value problem for the corresponding ordinary differential equation is given. The prestress can then be calculated from the stress-strain relations as

$$\sigma_{rr}^0(r) = \left(a_1 + \frac{a_2}{r^2} - \frac{3+\nu}{8}r^2\right)\Omega^2 \tag{108}$$

$$\sigma_{\varphi\varphi}^0(r) = \left(a_1 - \frac{a_2}{r^2} - \frac{1+3\nu}{8}r^2\right)\Omega^2 \tag{109}$$

where

$$a_1 = \left(\frac{1+\nu}{8}\right)\frac{(3+\nu)+(1-\nu)\left(\frac{r_i}{r_o}\right)^4}{(1-\nu)\left(\frac{r_i}{r_o}\right)^2+(1+\nu)} \tag{110}$$

$$a_2 = \left(\frac{1-\nu}{8}\right)\frac{(3+\nu)\left(\frac{r_i}{r_o}\right)^2+(1+\nu)\left(\frac{r_i}{r_o}\right)^4}{(1-\nu)\left(\frac{r_i}{r_o}\right)^2+(1+\nu)} \tag{111}$$

for the boundary condition (106). The expression for prestress given by Eq. (108) is easier to handle in numerical calculations since Ω appears polynomially and not as the argument of a Bessel function as in Eq. (105). The corresponding prestress is plotted in Fig. 12 by the dashed lines for $\Omega=0.28$. Since it is seen from Fig. 12 that for the rotational speeds considered in this paper Eq. (108) is a very good approximation, all numerical calculations for the plate are carried out with these equations. For larger Ω Eq. (108) is no longer valid and it cannot be used to determine static instabilities as is done in Ref. [30] using Eq. (107).

4.4.2 Results From the Perturbation Analysis. Performing a perturbation analysis on the prestressed plate analogous to the investigations performed in Secs. 4.2 and 4.3 previous sections one at first has to calculate the spectrum of the unperturbed problem. The only differences in the unperturbed boundary value problems are the terms proportional to $\sigma_{rr}^0(r)$ and $\sigma_{\phi\phi}^0(r)$. Expanding the eigenfunctions of the unperturbed problem in terms of eigenfunctions of the nonrotating plate as in Eq. (76) and projecting onto them one sees by similar argument as in Sec. 4.1 that equations corresponding to different nodal diameters m will decouple. Due to the dependence of $\sigma_{rr}^0(r)$ and $\sigma_{\phi\phi}^0(r)$ on r a further decoupling does not occur.

The spectrum of the unperturbed problem resulting from the Ritz approach using the ten lowest eigenmodes of the nonrotating plate is given in Fig. 13 by the solid lines.

The dashed lines show the spectrum neglecting the influence of the prestress and one can observe the stiffening effect since the eigenfrequencies increase. The eigenvalues for the perturbed problem using the same shape functions in a Ritz approach for the perturbed problem are shown in Fig. 14.

The perturbation formulas are similar to Eq. (89) but larger matrices and eigenvectors have to be used since equations corresponding to the same nodal diameter in the unperturbed problem are coupled. We see that there is no qualitative difference to the graphs obtained neglecting the prestress. The same holds true for the stability boundaries; therefore they are not redrawn for the prestressed case.

5 Nonlinear Analysis

When dealing with linear models, in case of an instability of the trivial solution, the amplitudes of the vibration become infinite. In practice this does not happen, since nonlinearities limit the amplitudes. In this section we therefore introduce a nonlinear stiffness characteristic in the pads and investigate the discretized equations of motion using the continuation method. We concentrate on the disk brake model with parameters given in Sec. 4.2. Since it is

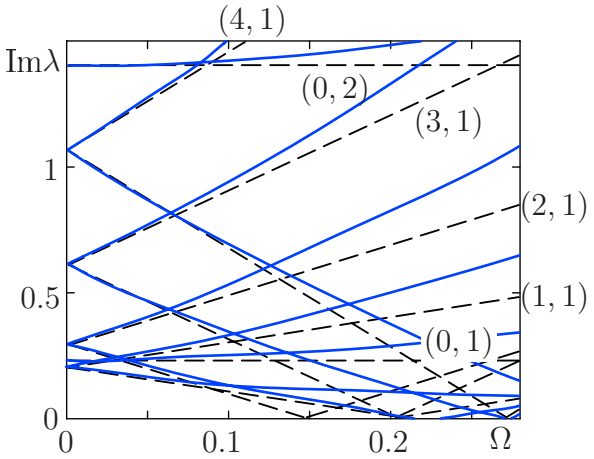


Fig. 13 Spectrum of the out-of-plane vibrations of the unperturbed problem (prestressed plate) for r_i=0.154 (spectral mesh)

observed in the laboratory that the squealing brake rotor almost vibrates in the first pair of eigenmodes of the nonrotating plate, we limit the expansion of the solution to these modes, i.e.,

$$w(r,\varphi,t) = R_{mn}(r)(q_1(t)\cos m\varphi + q_2(t)\sin m\varphi) \tag{112}$$

where $R_{mn}(r)$ is the radial component of an eigenfunction of the corresponding nonrotating plate. As before, the quantities m and n denote the number of nodal diameters and nodal circles, respectively. Note that all boundary conditions of the rotating plate are satisfied by the shape functions. The derivation of the equations of motion made clear that nonlinear kinematics, i.e., using the nonlinear deformation gradient, yields equations that are highly coupled. Since we expect the amplitude of the disk to be very small, we stick to the assumption of linear kinematics and assume that nonlinearities only enter the system through a nonlinear pad stiffness. Considering this nonlinearity of the pads, the force balance at the upper brake pad (16) now reads

$$(\boldsymbol{N}_{\bar{P}} + \boldsymbol{R}_{\bar{P}}) \cdot \boldsymbol{e}_z + N_0 - F_{\bar{P}} = 0 \tag{113}$$

where

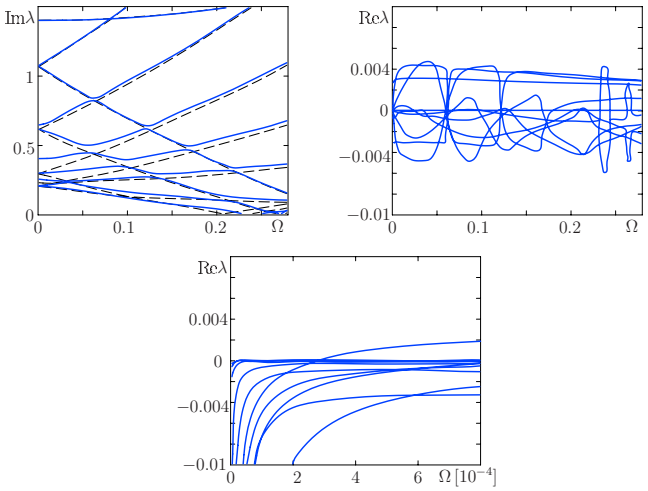


Fig. 14 Spectrum of the perturbed plate problem for $\gamma=\delta=3$. Dashed lines: unperturbed problem; lower plot: zoom for small Ω.

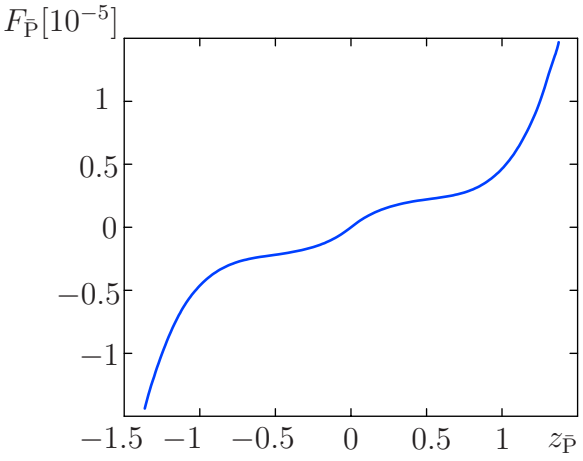


Fig. 15 Stiffness characteristic

$$F_{\bar{P}} = k_1\left(z_{\bar{P}} + \frac{h}{2}\right) + \mathrm{sgn}\left(z_{\bar{P}} + \frac{h}{2}\right)k_2\left(z_{\bar{P}} + \frac{h}{2}\right)^2 + k_3\left(z_{\bar{P}} + \frac{h}{2}\right)^3 + k_5\left(z_{\bar{P}} + \frac{h}{2}\right)^5 \tag{114}$$

and the sign function ensures that the stiffness characteristic is point-symmetric with respect to the prestressed configuration.

Very little is known about the stiffness and damping parameters of the pads; therefore the minimal model can only be investigated qualitatively. For a bifurcation analysis using AUTO [32] the stiffness and damping parameters are varied in a physically plausible range. Depending on the parameters we obtain a subcritical or a supercritical Hopf bifurcation when increasing the rotational speed of the disk.

Using the dimensionless stiffness parameters

$$k_1 = 0.0093, \quad k_2 = -1.30 \times 10^5, \quad k_3 = 4.69 \times 10^{11}, \quad k_5 = 2.21 \times 10^{22} \tag{115}$$

we obtain the stiffness characteristic shown in Fig. 15 and the bifurcation diagram in Fig. 16.

The phenomenon of a subcritical Hopf bifurcation has been observed in the laboratory. When accelerating the disk, squeal arises only above a certain angular velocity, but when the squealing brake is decelerated squeal is observed almost until the disk comes to a complete rest. However, we remark that this characteristic, i.e., the range of speed, is much smaller in the present model than one would have expected from laboratory experiments. This is most likely due to the fact that a minimal model is investigated and adjacent parts of the disk brake are not taken into account. Nevertheless the characteristic effect can be observed in the phenomenological model (cf. Fig. 16).

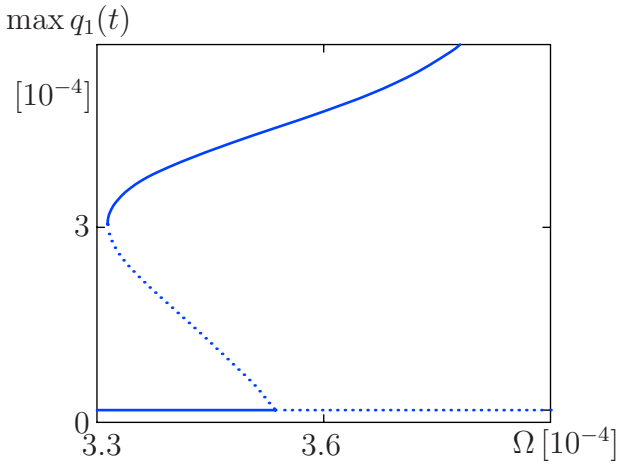


Fig. 16 Bifurcation diagram

6 Conclusion

This paper investigates the stretching and bending of an annular Kirchhoff plate in frictional contact with idealized distributed pads. The model of the pads is rather flexible to be employed in a variety of problems including clutches. The equations of motion are derived from the basics of the theory of elasticity and coupling between disk and plate equations is carefully examined. Under the assumptions used in this paper, the disk and the plate equations are decoupled in the linear case. Using perturbation techniques, it is shown that at least in the subcritical range for parameters of a brake disk the plate equations determine the stability behavior of the system. Subcritical flutter instability is observed caused by the combined action of dissipative and nonconservative positional forces. Approximations to the stability boundaries of the system are calculated from the perturbation formulas. The interpretation of the problem as a perturbation problem provides an effective tool for the stability analysis.

Finally, a nonlinear analysis of the discretized plate equations is performed to investigate the influences of the nonlinearities originating from the friction material. Observations of nonlinearities made at a squealing disk brake at the test rig, which have not been investigated in literature, can be verified qualitatively from the theoretical model using the continuation method.

Although plate and disk equations decouple in the linear case, i.e., the stability behavior is determined by the plate equations only, they are coupled through nonlinear terms. An instability in the plate equations is an instability of the complete system since the disk equations are excited through nonlinear terms. Consequently the squealing brake will show in- and out-of-plane vibrations as seen from experiments.

Acknowledgment

The work was partly supported by the Alexander von Humboldt-Foundation and DFG Grant No. HA 1060/43-1.

References

[1] Lamb, H., and Southwell, R., 1921, "The Vibrations of a Spinning Disk," Proc. R. Soc. London, Ser. A, **99**, pp. 272–280.
[2] Southwell, R., 1922, "On the Free Transverse Vibrations of a Uniform Circular Disc Clamped at Its Centre; and on the Effects of Rotation," Proc. R. Soc. London, Ser. A, **101**, pp. 133–153.
[3] Washizu, K., 1974, *Variational Methods in Elasticity and Plasticity*, Pergamon, New York.
[4] Baddour, N., and Zu, J., 2001, "A Revisit of Spinning Disk Models. Part I: Derivation of Equations of Motion," Appl. Math. Model., **25**, pp. 541–559.
[5] Baddour, N., and Zu, J., 2001, "A Revisit of Spinning Disk Models. Part II: Linear Transverse Vibrations," Appl. Math. Model., **25**, pp. 561–578.
[6] Moser, F., Fischer, M., and Rumold, W., 2002, "Dreidimensionale Messung von Betriebsschwingformen quietschender Scheibenbremsen," VDI-Ber., **1736**, pp. 71–83.
[7] Mottershead, J., 1998, "Vibration- and Friction-Induced Instability in Disks," Shock Vib. Dig., **30**(1), pp. 14–31.
[8] Ono, H., Chen, J., and Bogy, D. B., 1991, "Stability Analysis for the Head-Disk Interface in a Flexible Disk Drive," ASME J. Appl. Mech., **58**, pp. 1005–1014.
[9] Ouyang, H., and Mottershed, J. E., 2004, "Dynamic Instability of an Elastic Disk Under the Action of a Rotating Friction Couple," ASME J. Appl. Mech., **71**, pp. 753–758.
[10] Kinkaid, N. M., O'Reilly, O. M., and Papadopoulos, P., 2003, "Automotive Disc Brake Squeal," J. Sound Vib., **267**, pp. 105–166.
[11] Ibrahim, R. A., 1994, "Friction-Induced Vibration, Chatter, Squeal and Chaos, Part I: Mechanics of Contact and Friction," Appl. Mech. Rev., **47**(7), pp. 227–253.
[12] Tseng, J. G., and Wickert, J. A., 1998, "Nonconservative Stability of a Friction Loaded Disk," ASME J. Vibr. Acoust., **120**, pp. 922–929.
[13] Hochlenert, D., Spelsberg-Korspeter, G., and Hagedorn, P., 2007, "Friction Induced Vibrations in Moving Continua and Their Application to Brake Squeal," ASME J. Appl. Mech., **74**, pp. 542–549.
[14] Kirillov, O. N., and Seyranian, A. P., 2004, "Collapse of the Keldysh Chains

and Stability of Continuous Nonconservative Systems," SIAM J. Appl. Math., **64**(4), pp. 1383–1407.
[15] Seyranian, A. P., and Kliem, W., 2001, "Bifurcations of Eigenvalues of Gyroscopic Systems With Parameters Near Stability Boundaries," ASME J. Appl. Mech., **68**, pp. 199–205.
[16] Chen, J. S., and Jhu, J. L., 1996, "On the In-Plane Vibration and Stability of a Spinning Annular Disk," J. Sound Vib., **195**(4), pp. 585–593.
[17] Hagedorn, P., 1986, "Eine Bemerkung zur Momentenimpedanz bei Plattenschwingungen in der Substrukturtechnik," ZAMP, **37**, pp. 293–303.
[18] Vishik, M. I., and Lyusternik, L. A., 1960, "The Solution of Some Perturbation Problems for Matrices and Selfadjoint or Non-Selfadjoint Differential Equations I," Russ. Math. Surveys, **15**, pp. 1–73.
[19] Vishik, M. I., and Lyusternik, L. A., 1960, "The Asymptotic Behaviour of Solutions of Linear Differential Equations With Large or Quickly Changing Coefficients and Boundary Conditions," Russ. Math. Surveys, **15**, pp. 23–91.
[20] Vidoli, S., and Vestroni, F., 2005, "Veering Phenomena in Systems With Gyroscopic Coupling," ASME J. Appl. Mech., **72**, pp. 641–647.
[21] Leissa, A., 1969, "Vibrations of Plates," NASA, Report No. SP-160.
[22] Günther, U., and Kirillov, O. N., 2006, "Krein Space Related Perturbation Theory for MHD α^2-Dynamos and Resonant Unfolding of Diabolical Points," J. Phys. A, **39**, pp. 10057–10076.
[23] Jearsiripongkul, T., 2005, "Squeal in Floating Caliper Disk Brakes: A Mathematical Model," Ph.D. thesis, Technische Universität Darmstadt, Darmstadt, Germany.
[24] Renshaw, A. A., and Mote, C. D., 1992, "Absence of One Nodal Diameter Critical Speed Modes in an Axisymmetric Rotating Disk," Trans. ASME, J. Appl. Mech., **59**, pp. 687–688.
[25] Chen, J.-S., and Jhu, J.-L., 1996, "In-Plane Response of a Rotating Annular Disk Under Fixed Concentrated Edge Loads," Int. J. Mech. Sci., **38**(12), pp. 1285–1293.
[26] Hochlenert, D., 2006, "Selbsterregte Schwingungen in Scheibenbremsen: Mathematische Modellbildung und aktive Unterdrückung von Bremsenquietschen," Ph.D. thesis, Technische Universität Darmstadt, Darmstadt, Germany.
[27] Kirillov, O., Mailybaev, A., and Seyranian, A., 2005, "Unfolding of Eigenvalue Surfaces Near a Diabolic Point Due to a Complex Perturbation," J. Phys. A, **38**(24), pp. 5531–5546.
[28] Kirillov, O. N., 2007, "Destabilization Paradox Due to Breaking the Hamiltonian and Reversible Symmetry," Int. J. Non-Linear Mech., **42**(1), pp. 71–87.
[29] Merkin, D. R., 1997, *Introduction to the Theory of Stability*, Springer, New York.
[30] Seemann, W., and Wauer, J., 1988, "Vibration of High Speed Disk Rotors," *Rotating Machinery Dynamics Proceedings of the Second International Symposium on Transport Phenomena and Dynamics of Rotating Machinery (ISROMAC-2)*.
[31] Timoshenko, S., and Goodier, J. N., 1951, *Theory of Elasticity*, McGraw-Hill, New York.
[32] Doedel, E., Pfaffenroth, R., Champneys, A., Fairgrieve, T., Kuznetsov, B. S., and Wang, X., 2001, "AUTO 2000: Continuation and Bifurcation Software for Ordinary Differential Equations (With HomCont)," California Institute of Technology, Technical Report.

Jinyu Zhu
e-mail: j7zhu@engmail.uwaterloo.ca

W.-C. Xie

Department of Civil and Environmental Engineering,
Faculty of Engineering,
University of Waterloo,
Waterloo, ON, N2L 3G1, Canada

Ronald M. C. So

X. Q. Wang

Department of Mechanical Engineering,
The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong

Parametric Resonance of a Two Degrees-of-Freedom System Induced by Bounded Noise

The dynamic stability of a two degrees-of-freedom system under bounded noise excitation with a narrowband characteristic is studied through the determination of moment Lyapunov exponents. The partial differential eigenvalue problem governing the moment Lyapunov exponent is established. For weak noise excitations, a singular perturbation method is employed to obtain second-order expansions of the moment Lyapunov exponents and Lyapunov exponents, which are shown to be in good agreement with those obtained using Monte Carlo simulation. The different cases when the system is in subharmonic resonance, combination additive resonance, and combined resonance in the absence of noise, respectively, are considered. The effects of noise and frequency detuning on the parametric resonance are investigated. [DOI: 10.1115/1.2999427]

Keywords: parametric resonance, moment Lyapunov exponents, bounded noise

1 Introduction

The equations of motion for many engineering problems are of the general form

$$\ddot{q}_i(t) + 2\beta_i\dot{q}_i(t) + \omega_i^2 q(t) + \zeta(t)\omega_i\sum_{j=1}^{2} k_{ij}q_j(t) = 0, \quad i = 1,2 \quad (1)$$

where q_i's are the generalized coordinates, β_i and ω_i are the ith damping constant and the circular natural frequency, respectively, and $\zeta(t)$ is a stochastic process describing the excitation.

The sample or almost-sure stability of the trivial solutions of system (1) is determined by the largest Lyapunov exponent, which characterizes the average exponential rate of growth of the solutions of system (1) for large t and is defined as

$$\lambda_{q(t)} = \lim_{t\to\infty}\frac{1}{t}\log\|\mathbf{q}(t)\| \quad (2)$$

where $\mathbf{q}(t)=\{q_1,\dot{q}_1,q_2,\dot{q}_2\}^T$ and $\|\mathbf{q}\|=(\mathbf{q}^T\mathbf{q})^{1/2}$ is the Euclidean norm. If the largest Lyapunov exponent is negative, the trivial solution of system (1) is stable with probability 1; otherwise, it is unstable.

On the other hand, the stability of the pth moment $\mathbf{E}[\|\mathbf{q}\|^p]$ of the solutions of system (1) is governed by the pth moment Lyapunov exponent defined by

$$\Lambda_{q(t)}(p) = \lim_{t\to\infty}\frac{1}{t}\log\,\mathbf{E}[\|\mathbf{q}\|^p] \quad (3)$$

where $\mathbf{E}[\cdot]$ denotes the expected value. If $\Lambda_{q(t)}(p)$ is negative, then the pth moment is stable; otherwise, it is unstable almost surely. The moment Lyapunov exponents are important in obtaining a complete picture of the dynamic stability of the trivial solution of system (1).

The relationship between the sample stability and the moment stability was formulated by Arnold [1]. The pth moment Lyapunov exponent $\Lambda_{q(t)}(p)$ is a convex analytic function in p that passes through the origin, and the slope at the origin is equal to the largest Lyapunov exponent $\lambda_{q(t)}$, i.e.,

$$\lambda_{q(t)} = \lim_{p\to 0}\frac{\Lambda_{q(t)}(p)}{p} \quad (4)$$

Equation (1) represents a number of practical flow-induced vibration problems encountered in aerospace, power, and structural engineering. Namachchivaya and Vedula [2] used a similar set of equations to study the stability of a downstream cylinder in the wake of upstream cylinder arrays. The two degrees-of-freedom represent the cylinder motions in the lift and the drag directions, respectively. Their interaction would induce unstable motions at certain reduced velocities, which are related to ω_i and k_{ij}. Moreover, when the approach flow is turbulent, the coefficients of interaction were modeled by real noise processes, and the stabilization effect observed experimentally was explained. For a civil structure such as a bridge deck, the two degrees-of-freedom usually represent the bending motion and the torsional motion [3,4]. Again, the aerodynamic interaction between the two motions gives rise to unstable motion under certain conditions. The effect of flow turbulence was also addressed, and it was shown that the turbulence has a stabilizing or destabilizing effect, depending on whether it increases or decreases the mean critical wind velocity. A similar aeroelastic coupling between the translational mode and the tortional mode can also be found in the flutter of an airfoil [5]. The flutter speed was found to be decreased by flow turbulence, especially its longitudinal component, and this change in the flutter point is mainly due to fluid stiffness.

A systematic study of moment Lyapunov exponents is presented by Arnold et al. [6] for linear Itô systems and by Arnold et al. [7] for linear stochastic systems under real noise excitations. Xie [8,9] obtained weak noise expansions of the moment Lyapunov exponent, the Lyapunov exponent, and the stability index of a two-dimensional system under real noise excitation and bounded noise excitation in terms of the small fluctuation parameter. In a recent paper, Zhu et al. [10] modeled the vortex shedding force of a cylinder in a cross-flow as a bounded noise process and found that parametric instability occurred in the lock-in region.

For a two degrees-of-freedom system, Namachchivaya and Vedula [2] obtained an asymptotic approximation of the moment Lyapunov exponent and the Lyapunov exponent with one critical mode and another asymptotically stable mode driven by a real noise of small intensity. They showed that the system can be stabilized by real noise. Namachchivaya and Van Roessel [11] studied two coupled oscillators driven by real noise. They set up the eigenvalue problem using the perturbation and stochastic averag-

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 25, 2008; final manuscript received September 16, 2008; published online April 22, 2009. Review conducted by Robert M. McMeeking.

ing methods, respectively, and obtained an approximation for the moment Lyapunov exponents of the two degrees-of-freedom system. Furthermore, Namachchivaya and Van Roessel [12] determined the moment Lyapunov exponent of two coupled oscillators with commensurable frequencies excited by a real noise using a perturbation method. Li and Lin [4] studied a two degrees-of-freedom system excited by the bounded noise process. They used the stochastic averaging method to determine the Lyapunov exponent of the system and thus determined the boundary of stability of the system.

A systematic presentation of the theory of random dynamical systems and a comprehensive list of references are provided by Arnold [13]. The theory and techniques of studying the stability of stochastic systems are presented in Ref. [14].

In this paper, a two degrees-of-freedom system under bounded noise excitation is considered. Depending on the central frequency of the bounded noise, various types of parametric resonance may occur in the system. The partial differential eigenvalue problem governing the moment Lyapunov exponent is established. For weak noise excitations, a singular perturbation method is employed to obtain second-order expansions of the moment Lyapunov exponents. The effects of the bounded noise and the detuning frequency on the resonance are investigated. The present paper explores the moment stability as well as the sample stability of the system and can be considered as an extension of Ref. [4], which studied the sample stability of the two degrees-of-freedom system through the stochastic averaging method. The stochastic stability of the system, which is in subharmonic resonance, combination additive resonance, and combined resonance when the excitation is a harmonic function, respectively, is studied. The numerical results for both the Lyapunov exponents and moment Lyapunov exponents are obtained by Monte Carlo simulation and compared with the analytical results to validate the analytical approach.

2 Formulation

Consider the following two degrees-of-freedom system

$$\ddot{q}_1 + 2\varepsilon\beta_1\dot{q}_1 + \omega_1^2 q_1 + \varepsilon\omega_1(k_{11}q_1 + k_{12}q_2)\zeta(t) = 0 \qquad (5a)$$

$$\ddot{q}_2 + 2\varepsilon\beta_2\dot{q}_2 + \omega_2^2 q_2 + \varepsilon\omega_2(k_{21}q_1 + k_{22}q_2)\zeta(t) = 0 \qquad (5b)$$

where $\zeta(t)$ is the excitation process, and ε is a small parameter introduced to make the analytical analysis more convenient.

2.1 Deterministic Excitation. If the excitation $\zeta(t)$ is deterministic and harmonic, e.g., $\zeta(t)=\cos \nu t$, system (5) can be in parametric resonance depending on the frequency ν. When the central frequency ν is not in the vicinities of $2\omega_i$ or $|\omega_1 \pm \omega_2|$, the system is stable and there is no resonance. If $\nu=2\omega_i$, the ith mode is excited and the system is in subharmonic resonance in the ith mode. The dynamic stability behavior of the system is the same as that of a single degree-of-freedom system in the first-order approximation. If the excitation frequency ν is in the vicinities of the linear combinations of two natural frequencies, i.e., $\nu=|\omega_1 \pm \omega_2|$, both modes are excited and the system is in combination resonance. See, e.g., Ref. [14], for details on parametric resonance in multiple degrees-of-freedom systems.

2.2 Stochastic Excitation. In most practical applications, the excitation $\zeta(t)$ has to be described by a random process and in many cases a narrowband process. To consider the effect of noise on parametric resonance, consider the excitation $\zeta(t)$ as a narrowband process, modeled by a bounded noise

$$\zeta(t) = \cos\ \eta(t) = \cos[\nu t + \varepsilon^{1/2}\sigma W(t) + \psi]$$

in which $W(t)$ is the standard Wiener process, and ψ is a uniformly distributed random number in $(0, 2\pi)$ that makes $\zeta(t)$ a stationary process.

Using the transformation $q_1=x_1$, $\dot{q}_1=\omega_1 x_2$, $q_2=x_3$, and $\dot{q}_2=\omega_2 x_4$, Eqs. (5*a*) and (5*b*) can be written as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \varepsilon\zeta(t)\mathbf{B}\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^4 \qquad (6)$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & \omega_1 & 0 & 0 \\ -\omega_1 & -2\varepsilon\beta_1 & 0 & 0 \\ 0 & 0 & 0 & \omega_2 \\ 0 & 0 & -\omega_2 & -2\varepsilon\beta_2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -k_{11} & 0 & -k_{12} & 0 \\ 0 & 0 & 0 & 0 \\ -k_{21} & 0 & -k_{22} & 0 \end{bmatrix}$$

Applying the transformation

$$x_1 = e^{\rho}\cos\phi_1\cos\theta, \quad x_3 = e^{\rho}\cos\phi_2\sin\theta$$
$$x_2 = -e^{\rho}\sin\phi_1\cos\theta, \quad x_4 = -e^{\rho}\sin\phi_2\sin\theta \qquad (7)$$

one can obtain the following set of equations for the logarithm of amplitude ρ ($\rho=\log\|\mathbf{x}\|$), phase variables (ϕ_1, ϕ_2, θ), and noise process η

$$\dot{\rho} = \sum_{j=0}^{1} \varepsilon^j q_j(\phi_1, \phi_2, \theta, \eta) = m_\rho, \quad \dot{\theta} = \sum_{j=0}^{1} \varepsilon^j s_j(\phi_1, \phi_2, \theta, \eta) = m_\theta$$
$$\dot{\phi}_i = \sum_{j=0}^{1} \varepsilon^j h_{ij}(\phi_1, \phi_2, \theta, \eta) = m_{\phi_i}, \quad d\eta = \nu dt + \varepsilon^{1/2}\sigma dW(t) \qquad (8)$$

where

$$q_0(\phi_1, \phi_2, \theta, \eta) = 0$$

$$\begin{aligned} q_1(\phi_1, \phi_2, \theta, \eta) = {} & \tfrac{1}{4}\{(k_{11}\sin 2\phi_1 + k_{22}\sin 2\phi_2) + (k_{11}\sin 2\phi_1 \\ & - k_{22}\sin 2\phi_2)\cos 2\theta + [(k_{12}+k_{21})\sin\phi^+ + (k_{12} \\ & - k_{21})\sin\phi^-]\sin 2\theta\}\cos\eta(t) - \tfrac{1}{2}[\beta_1(1-\cos 2\phi_1) \\ & + \beta_2(1-\cos 2\phi_2)] - \tfrac{1}{2}[\beta_1(1-\cos 2\phi_1) - \beta_2(1 \\ & - \cos 2\phi_2)]\cos 2\theta \end{aligned}$$

$$h_{10}(\phi_1, \phi_2, \theta, \eta) = \omega_1$$

$$\begin{aligned} h_{11}(\phi_1, \phi_2, \theta, \eta) = {} & \tfrac{1}{2}[k_{11}(1+\cos 2\phi_1) + k_{12}(\cos\phi^+ \\ & + \cos\phi^-)\tan\theta]\cos\eta(t) - \beta_1\sin 2\phi_1 \end{aligned}$$

$$h_{20}(\phi_1, \phi_2, \theta, \eta) = \omega_2$$

$$\begin{aligned} h_{21}(\phi_1, \phi_2, \theta, \eta) = {} & \tfrac{1}{2}[k_{22}(1+\cos 2\phi_2) + k_{21}(\cos\phi^+ \\ & + \cos\phi^-)\cot\theta]\cos\eta(t) - \beta_2\sin 2\phi_2 \end{aligned}$$

$$s_0(\phi_1, \phi_2, \theta, \eta) = 0$$

$$\begin{aligned} s_1(\phi_1, \phi_2, \theta, \eta) = {} & \tfrac{1}{4}\{(k_{21}-k_{12})\sin\phi^+ - (k_{12}+k_{21})\sin\phi^- \\ & + (k_{22}\sin 2\phi_2 - k_{11}\sin 2\phi_1)\sin 2\theta + [(k_{12} \\ & + k_{21})\sin\phi^+ + (k_{12}-k_{21})\sin\phi^-]\cos 2\theta\}\cos\eta(t) \\ & + \tfrac{1}{2}[\beta_1(1-\cos 2\phi_1) - \beta_2(1-\cos 2\phi_2)]\sin 2\theta \end{aligned}$$

In the above expressions, $\phi^{\pm}=\phi_1 \pm \phi_2$.

The pth norm of $\mathbf{x}$ can be written as $P=\|\mathbf{x}\|^p=(e^{\rho})^p=e^{p\rho}$. The Itô equation for P can be derived using the Itô's lemma

$$dP = \left(m_\rho\frac{\partial}{\partial\rho} + m_{\phi_1}\frac{\partial}{\partial\phi_1} + m_{\phi_2}\frac{\partial}{\partial\phi_2} + m_\theta\frac{\partial}{\partial\theta}\right)P dt = p m_\rho P dt \qquad (9)$$

Applying a linear stochastic transformation

$$S = T(\eta, \phi_1, \phi_2, \theta)P, \quad P = T^{-1}(\eta, \phi_1, \phi_2, \theta)S$$

$$-\infty < \eta < +\infty, \quad 0 \le \phi_{1,2} < 2\pi, \quad 0 \le \theta < \frac{\pi}{2}$$

the Itô equation for the transformed pth norm process S can also be derived using Itô's lemma

$$\mathrm{d}S = \left(\tfrac{1}{2}\varepsilon\sigma^2 T_{\eta\eta} + \nu T_\eta + m_{\phi_1} T_{\phi_1} + m_{\phi_2} T_{\phi_2} + m_\theta T_\theta + m_P T\right) P \mathrm{d}t + \varepsilon^{1/2}\sigma T_\eta P \mathrm{d}W \tag{10}$$

For bounded and nonsingular transformation $T(\eta, \phi_1, \phi_2, \theta)$, both processes P and S are expected to have the same stability behavior. Therefore, $T(\eta, \phi_1, \phi_2, \theta)$ is chosen so that the drift term of the Itô differential equation (10) is independent of the noise process $\eta(t)$, the phase processes ϕ_1, ϕ_2, and θ, so that

$$\mathrm{d}S = \Lambda S \mathrm{d}t + \varepsilon^{1/2}\sigma T_\eta T^{-1} S \mathrm{d}W \tag{11}$$

Comparing Eqs. (10) and (11), it is seen that such a transformation $T(\eta, \phi_1, \phi_2, \theta)$ is given by the following equation:

$$\tfrac{1}{2}\varepsilon\sigma^2 T_{\eta\eta} + \nu T_\eta + m_{\phi_1} T_{\phi_1} + m_{\phi_2} T_{\phi_2} + m_\theta T_\theta + m_P T = \Lambda T,$$

$$-\infty < \eta < +\infty, \quad 0 \le \phi_{1,2} < 2\pi, \quad 0 \le \theta < \frac{\pi}{2} \tag{12}$$

in which $T(\eta, \phi_1, \phi_2, \theta)$ is a periodic function in ϕ_1 and ϕ_2 of period 2π and is bounded when $\eta \to \pm\infty$. Equation (12) defines an eigenvalue problem of a second-order differential operator with Λ being the eigenvalue and $T(\eta, \phi_1, \phi_2, \theta)$ the associated eigenfunction. From Eq. (12), the eigenvalue Λ is seen to be the Lyapunov exponent of the pth moment of system (6), i.e., $\Lambda = \Lambda_x(p)$.

3 Weak Noise Expansions of the Moment Lyapunov Exponent

3.1 Singular Perturbation Expansion. For weak noise excitation, i.e., $\varepsilon = o(1)$, perturbation methods can be applied to solve the partial differential eigenvalue problem (12) for the perturbative expansions of the moment Lyapunov exponent $\Lambda(p)$. Since the small parameter ε appears as a coefficient of the term $T_{\eta\eta}$, a method of singular perturbation (see, e.g., Ref. [15]) must be applied.

Denote the frequency $\nu = \nu_0 + \varepsilon\Delta$, where ν_0 is the central frequency and Δ is the detuning parameter. Applying the stretching transformation

$$\eta = \varepsilon^{1/2}\xi + (\alpha_1\phi_1 + \alpha_2\phi_2), \quad \xi = \varepsilon^{-1/2}[\eta - (\alpha_1\phi_1 + \alpha_2\phi_2)]$$

in which α_1 and α_2 are constants, one has

$$\frac{\partial T}{\partial \eta} = \frac{\partial T}{\partial \xi}\frac{\partial \xi}{\partial \eta} = \varepsilon^{-1/2} T_\xi, \quad \frac{\partial^2 T}{\partial \eta^2} = \varepsilon^{-1} T_{\xi\xi}$$

$$\frac{\partial T}{\partial \phi_1} = \frac{\partial T}{\partial \phi_1} + \frac{\partial T}{\partial \xi}\frac{\partial \xi}{\partial \phi_1} = T_{\phi_1} - \varepsilon^{-1/2}\alpha_1 T_\xi$$

$$\frac{\partial T}{\partial \phi_2} = \frac{\partial T}{\partial \phi_2} + \frac{\partial T}{\partial \xi}\frac{\partial \xi}{\partial \phi_2} = T_{\phi_2} - \varepsilon^{-1/2}\alpha_2 T_\xi$$

Thus, Eq. (12) becomes

$$\tfrac{1}{2}\sigma^2 T_{\xi\xi} + \omega_1 T_{\phi_1} + \omega_2 T_{\phi_2} + \varepsilon^{-1/2}[\nu_0 - (\alpha_1\omega_1 + \alpha_2\omega_2)]T_\xi + \varepsilon^{1/2}(\Delta - \alpha_1\bar{h}_{11} - \alpha_2\bar{h}_{21})T_\xi + \varepsilon(\bar{h}_{11}T_{\phi_1} + \bar{h}_{21}T_{\phi_2} + \bar{s}_1 T_\theta + p\bar{q}_1 T) = \Lambda T, \quad -\infty < \xi < +\infty, \quad 0 \le \phi_{1,2} < 2\pi, \quad 0 \le \theta < \frac{\pi}{2} \tag{13}$$

where

$$\begin{aligned}
\bar{q}_1(\phi_1, \phi_2, \theta, \xi) &= \tfrac{1}{4}\{(k_{11}\sin 2\phi_1 + k_{22}\sin 2\phi_2) + (k_{11}\sin 2\phi_1 - k_{22}\sin 2\phi_2)\cos 2\theta + [(k_{12} + k_{21})\sin\phi^+ + (k_{12} - k_{21})\sin\phi^-]\sin 2\theta\}\cos(\varepsilon^{1/2}\xi + \alpha_1\phi_1 + \alpha_2\phi_2) \\
&\quad - \tfrac{1}{2}[\beta_1(1 - \cos 2\phi_1) + \beta_2(1 - \cos 2\phi_2)] \\
&\quad - \tfrac{1}{2}[\beta_1(1 - \cos 2\phi_1) - \beta_2(1 - \cos 2\phi_2)]\cos 2\theta
\end{aligned}$$

$$\begin{aligned}
\bar{h}_{11}(\phi_1, \phi_2, \theta, \xi) &= \tfrac{1}{2}[k_{11}(1 + \cos 2\phi_1) + k_{12}(\cos\phi^+ + \cos\phi^-)\tan\theta] \\
&\quad \times \cos(\varepsilon^{1/2}\xi + \alpha_1\phi_1 + \alpha_2\phi_2) - \beta_1 \sin 2\phi_1
\end{aligned}$$

$$\begin{aligned}
\bar{h}_{21}(\phi_1, \phi_2, \theta, \xi) &= \tfrac{1}{2}[k_{22}(1 + \cos 2\phi_2) + k_{21}(\cos\phi^+ + \cos\phi^-)\cot\theta] \\
&\quad \times \cos(\varepsilon^{1/2}\xi + \alpha_1\phi_1 + \alpha_2\phi_2) \\
&\quad - \beta_2 \sin 2\phi_2 \bar{s}_1(\phi_1, \phi_2, \theta, \xi) \\
&= \tfrac{1}{4}\{(k_{21} - k_{12})\sin\phi^+ - (k_{12} + k_{21})\sin\phi^- \\
&\quad + (k_{22}\sin 2\phi_2 - k_{11}\sin 2\phi_1)\sin 2\theta \\
&\quad + [(k_{12} + k_{21})\sin\phi^+ + (k_{12} - k_{21})\sin\phi^-]\cos 2\theta\} \\
&\quad \times \cos(\varepsilon^{1/2}\xi + \alpha_1\phi_1 + \alpha_2\phi_2) + \tfrac{1}{2}[\beta_1(1 - \cos 2\phi_1) \\
&\quad - \beta_2(1 - \cos 2\phi_2)]\sin 2\theta
\end{aligned}$$

If $\nu_0 = \alpha_1\omega_1 + \alpha_2\omega_2$, then Eq. (13) can be reduced to

$$\tfrac{1}{2}\sigma^2 T_{\xi\xi} + \omega_1 T_{\phi_1} + \omega_2 T_{\phi_2} + \varepsilon^{1/2}(\Delta - \alpha_1\bar{h}_{11} - \alpha_2\bar{h}_{21})T_\xi + \varepsilon(\bar{h}_{11}T_{\phi_1} + \bar{h}_{21}T_{\phi_2} + \bar{s}_1 T_\theta + p\bar{q}_1 T) = \Lambda T, \quad -\infty < \xi < +\infty, \quad 0 \le \phi_{1,2} < 2\pi, \quad 0 \le \theta < \frac{\pi}{2} \tag{14}$$

in which the eigenfunction T is treated as a function of ξ, ϕ_1, ϕ_2, θ, and ε. Denoting $z = \varepsilon^{1/2}\xi$, the eigenfunction $T(\xi, \phi_1, \phi_2, \theta, \varepsilon)$ becomes $Y(\xi, z, \phi_1, \phi_2, \theta)$. It can be shown that

$$T_\xi = Y_\xi + \varepsilon^{1/2} Y_z, \quad T_{\xi\xi} = Y_{\xi\xi} + 2\varepsilon^{1/2} Y_{\xi z} + \varepsilon Y_{zz} \tag{15}$$

Substituting Eq. (15) into Eq. (14) leads to

$$\mathcal{L}(p)Y = \Lambda(p)Y, \quad \mathcal{L}(p)Y = \mathcal{L}_0 Y + \varepsilon^{1/2}\mathcal{L}_1 Y + \varepsilon\mathcal{L}_2 Y \tag{16}$$

where

$$\mathcal{L}_0 Y = \tfrac{1}{2}\sigma^2 Y_{\xi\xi} + \omega_1 Y_{\phi_1} + \omega_2 Y_{\phi_2}$$

$$\mathcal{L}_1 Y = \sigma^2 Y_{\xi z} + (\Delta - \alpha_1\bar{h}_{11} - \alpha_2\bar{h}_{21})Y_\xi$$

$$\mathcal{L}_2 Y = \tfrac{1}{2}\sigma^2 Y_{zz} + (\Delta - \alpha_1\bar{h}_{11} - \alpha_2\bar{h}_{21})Y_z + (\bar{h}_{11}Y_{\phi_1} + \bar{h}_{21}Y_{\phi_2} + \bar{s}_1 Y_\theta + p\bar{q}_1 Y)$$

Expanding the eigenvalue $\Lambda_{q(t)}(p)$ and the eigenfunction $Y(\xi, z, \varphi)$ yields

$$\Lambda_{q(t)}(p) = \sum_{n=0}^{\infty} \varepsilon^{n/2}\Lambda_n, \quad Y(\xi, z, \phi_1, \phi_2, \theta) = \sum_{n=0}^{\infty} \varepsilon^{n/2} Y_n(\xi, z, \phi_1, \phi_2, \theta) \tag{17}$$

where $Y_n(\xi, z, \phi_1, \phi_2, \theta)$ are periodic functions in ϕ_1 and ϕ_2 of period 2π. Substituting Eq. (17) into Eq. (16), expanding, and equating terms of the same orders of $\varepsilon^{1/2}$ yield

$$\mathrm{O}(1): \ \mathcal{L}_0 Y_0 = \Lambda_0 Y_0$$

$$\mathrm{O}(\varepsilon^{1/2}): \ \mathcal{L}_0 Y_1 + \mathcal{L}_1 Y_0 = \Lambda_0 Y_1 + \Lambda_1 Y_0 \tag{18}$$

$$O(\varepsilon^1):\ \mathcal{L}_0 Y_2 + \mathcal{L}_1 Y_1 + \mathcal{L}_2 Y_0 = \sum_{i=0}^{2} \Lambda_i Y_{n-i}$$

$$O(\varepsilon^{n/2}):\ \mathcal{L}_0 Y_n + \mathcal{L}_1 Y_{n-1} + \mathcal{L}_2 Y_{n-2} = \sum_{i=0}^{n} \Lambda_i Y_{n-i}, \quad n = 3,4,\ldots$$

3.2 Zeroth-Order Perturbation. The zeroth-order perturbation equation is $\mathcal{L}_0 Y_0 = \Lambda_0 Y_0$, or

$$\frac{\sigma^2}{2}\frac{\partial^2 Y_0}{\partial \xi^2} + \omega_1 \frac{\partial Y_0}{\partial \phi_1} + \omega_2 \frac{\partial Y_0}{\partial \phi_2} = \Lambda_0 Y_0 \tag{19}$$

Since the moment Lyapunov exponent $\Lambda_{q(t)}(p)$ passes through the origin, i.e.,

$$\Lambda_{q(t)}(0) = \Lambda_0(0) + \varepsilon^{1/2}\Lambda_1(0) + \varepsilon\Lambda_2(0) + \cdots = 0$$

one obtains $\Lambda_0(0) = \Lambda_1(0) = \Lambda_2(0) = \cdots = 0$. Because Eq. (19) does not contain p explicitly, $\Lambda_0(0) = 0$ implies $\Lambda_0(p) = 0$. Applying the method of separation of variables and letting

$$Y_0(\xi, z, \phi_1, \phi_2, \theta) = X_0(\xi) Z_0(z, \theta) \Phi_{01}(\phi_1) \Phi_{02}(\phi_2)$$

Equation (19) becomes

$$\frac{\Phi_{01}'}{\Phi_{01}} = a_1$$

$$\frac{\Phi_{02}'}{\Phi_{02}} = a_2$$

$$\frac{\sigma^2}{2}\frac{\ddot{X}_0}{X_0} = -(a_1 + a_2) = \kappa$$

Solving the $\Phi_{01}(\phi_1)$ equation yields $\Phi_{01}(\phi_1) = Ce^{a_1\phi_1}$. For $\Phi_{01}(\phi_1)$ to be a periodic function of period 2π, the constant $a_1 = 0$ and hence $\Phi_{01}(\phi_1) = C_1$. Similarly, $a_2 = 0$ and $\Phi_{02}(\phi_2) = C_2$. The $X_0(\xi)$ equation results in $X_0(\xi) = D_0 + D_1\xi$. For $X_0(\xi)$ to be a bounded function as $\xi \to \pm\infty$, it is required that $D_1 = 0$ and hence $X_0(\xi) = D_0$. The zeroth-order perturbation of the eigenfunction is therefore $Y_0(\xi, z, \phi_1, \phi_2, \theta) = Z_0(z, \theta)$, where $Z_0(z, \theta)$ is a function to be determined.

The adjoint equation of Eq. (19) is

$$\frac{\sigma^2}{2}\frac{\partial^2 Y_0^*}{\partial \xi^2} - \omega_1 \frac{\partial Y_0^*}{\partial \phi_1} - \omega_2 \frac{\partial Y_0^*}{\partial \phi_2} = 0 \tag{19'}$$

Employing the method of separation of variables with

$$Y_0^*(\xi, z, \phi_1, \phi_2, \theta) = X_0^*(\xi) Z_0^*(z, \theta) \Phi_{01}^*(\phi_1) \Phi_{02}^*(\phi_2)$$

it is easy to show that

$$\Phi_{01}^*(\phi_1) = \frac{1}{2\pi}, \quad 0 \le \phi_1 < 2\pi$$

$$\Phi_{02}^*(\phi_2) = \frac{1}{2\pi}, \quad 0 \le \phi_2 < 2\pi$$

$$X_0^*(\xi) = \text{constant}, \quad -\infty < \xi < +\infty$$

$$Y_0^*(\xi, z, \phi_1, \phi_2, \theta) = Z_0^*(z, \theta) \tag{20}$$

3.3 First-Order Perturbation. Since $\Lambda_0 = 0$, the first-order perturbation equation becomes

$$\mathcal{L}_0 Y_1 = \Lambda_1 Y_0 - \mathcal{L}_1 Y_0 \tag{21}$$

From the Fredholm alternative, for Eq. (21) to have nonzero solutions, it is required that

$$(\Lambda_1 Y_0 - \mathcal{L}_1 Y_0, Y_0^*) = 0 \tag{22}$$

where $Y_0^*(\xi, z, \phi_1, \phi_2, \theta)$ is the solution (20) of the adjoint Eq. (19′) of the first-order perturbation equation, and (f, g) denotes the inner product of functions $f(\cdot)$ and $g(\cdot\cdot)$ defined as

$$(f, g) = \int_{z=-\infty}^{+\infty} \int_{\xi=-\infty}^{+\infty} \int_{\phi_1=0}^{2\pi} \int_{\phi_2=0}^{2\pi} \int_{\theta=0}^{\frac{\pi}{2}} f(\cdot) g(\cdot\cdot) \mathrm{d}\theta \mathrm{d}\phi_2 \mathrm{d}\phi_1 \mathrm{d}\xi \mathrm{d}z$$

Since $Y_0(\xi, z, \phi_1, \phi_2, \theta) = Z_0(z, \theta)$, which leads to $\mathcal{L}_1 Y_0 = 0$, Eq. (22) results in $\Lambda_1(p) = 0$. Equation (21) then becomes $\mathcal{L}_0 Y_1 = 0$. Following the same procedure as in Sec. 3.2, it is easy to show that $Y_1(\xi, z, \phi_1, \phi_2, \theta) = Z_1(z, \theta)$.

3.4 Second-Order Perturbation. Since $\Lambda_0 = \Lambda_1 = 0$, $\mathcal{L}_1 Y_1 = 0$, the second-order perturbation equation becomes

$$\mathcal{L}_0 Y_2 = \Lambda_2 Y_0 - \mathcal{L}_2 Y_0 \tag{23}$$

From the Fredholm alternative, for Eq. (23) to have nontrivial solutions, it is required that

$$(\Lambda_2 Y_0 - \mathcal{L}_2 Y_0, Y_0^*) = 0$$

which can be written as

$$\int_{z=-\infty}^{+\infty} \int_{\theta=0}^{\frac{\pi}{2}} Z_0^*(z, \theta) \left\{ \int_{\phi_1=0}^{2\pi} \int_{\phi_2=0}^{2\pi} \left[\frac{\sigma^2}{2}\frac{\partial^2 Z_0}{\partial z^2} + (\Delta - \alpha_1 \hat{h}_{11} - \alpha_2 \hat{h}_{21})\frac{\partial Z_0}{\partial z} + \hat{s}_1 \frac{\partial Z_0}{\partial \theta} + (p\hat{q}_1 - \Lambda_2) Z_0 \right] \mathrm{d}\phi_2 \mathrm{d}\phi_1 \right\} \mathrm{d}\theta \mathrm{d}z = 0 \tag{24}$$

Because Eq. (24) is valid for an arbitrary function $Z_0^*(z, \theta)$, it results in

$$\int_{\phi_1=0}^{2\pi} \int_{\phi_2=0}^{2\pi} \left[\frac{\sigma^2}{2}\frac{\partial^2 Z_0}{\partial z^2} + (\Delta - \alpha_1 \hat{h}_{11} - \alpha_2 \hat{h}_{21})\frac{\partial Z_0}{\partial z} + \hat{s}_1 \frac{\partial Z_0}{\partial \theta} + (p\hat{q}_1 - \Lambda_2) Z_0 \right] \mathrm{d}\phi_2 \mathrm{d}\phi_1 = 0$$

which, after performing the integration, leads to

$$L(p) Z_0 = \Lambda_2 Z_0 \tag{25}$$

where

$$L(p) Z_0 = \frac{\sigma^2}{2}\frac{\partial^2 Z_0}{\partial z^2} + (\Delta - \alpha_1 \hat{h}_{11} - \alpha_2 \hat{h}_{21})\frac{\partial Z_0}{\partial z} + \hat{s}_1 \frac{\partial Z_0}{\partial \theta} + p\hat{q}_1 Z_0$$

$$\hat{h}_{11} = \int_{\phi_1=0}^{2\pi} \int_{\phi_2=0}^{2\pi} \bar{h}_{11} \mathrm{d}\phi_2 \mathrm{d}\phi_1, \quad \hat{h}_{21} = \int_{\phi_1=0}^{2\pi} \int_{\phi_2=0}^{2\pi} \bar{h}_{21} \mathrm{d}\phi_2 \mathrm{d}\phi_1$$

$$\hat{s}_1 = \int_{\phi_1=0}^{2\pi} \int_{\phi_2=0}^{2\pi} \bar{s}_1 \mathrm{d}\phi_2 \mathrm{d}\phi_1, \quad \hat{p}_1 = \int_{\phi_1=0}^{2\pi} \int_{\phi_2=0}^{2\pi} \bar{p}_1 \mathrm{d}\phi_2 \mathrm{d}\phi_1$$

Hence, the second-order perturbation of the moment Lyapunov exponent Λ_2 is the eigenvalue of a second-order partial differential eigenvalue problem (25) with the function $Z_0(z, \theta)$ being the associated eigenfunction.

Equation (25) can be solved using a double Fourier series. Since θ is defined on the interval $[0, \pi/2)$, the Fourier series for θ should include $\cos 4n\theta$ and $\sin 4n\theta$, which come from the extension of the interval from $[0, \pi/2)$ to $[0, 2\pi)$ (see, e.g., Ref. [16]). Hence, the eigenfunction $Z_0(z, \theta)$ can be expressed as

$$Z_0(z,\theta) = CC_{00} + \sum_{n=1}^{K} CC_{0n}\cos 4n\theta + \sum_{n=0}^{K}\sum_{m=1}^{K}(CC_{mn}\cos mz + SC_{mn}\sin mz)\cos 4n\theta + \sum_{n=1}^{K} CS_{0n}\sin 4n\theta + \sum_{n=1}^{K}\sum_{m=1}^{K}(CS_{mn}\cos mz + SS_{mn}\sin mz)\sin 4n\theta \quad (26)$$

or, more concisely,

$$Z_0(z,\theta) = \sum_{n=0}^{K}\sum_{m=0}^{K}(CC_{mn}\cos mz + SC_{mn}\sin mz)\cos 4n\theta + \sum_{n=0}^{K}\sum_{m=0}^{K}(CS_{mn}\cos mz + SS_{mn}\sin mz)\sin 4n\theta \quad (26')$$

in which $SC_{0n}=CS_{m0}=SS_{m0}=SS_{0n}=0$, for all values of m and n, all the other CC_{mn}, SC_{mn}, CS_{mn}, and SS_{mn} are constant coefficients to be determined. The Fourier series is truncated to include K sine and cosine terms for the purpose of numerical analysis; when $K \to \infty$, the exact result is obtained.

Substituting Eqs. (26′) into (25), multiplying it by $\cos rz\cos 4s\theta$, $\sin rz\cos 4s\theta$, $\cos rz\sin 4s\theta$, or $\sin rz\sin 4s\theta$, respectively, and integrating with respect to z from 0 to 2π and with respect to θ from 0 to $\pi/2$ yield a set of equations

$$\sum_{m=0}^{K}\sum_{n=0}^{K}\begin{pmatrix} L_{mnrs}^{CCCC} & L_{mnrs}^{SCCC} & L_{mnrs}^{CSCC} & L_{mnrs}^{SSCC} \\ L_{mnrs}^{CCSC} & L_{mnrs}^{SCSC} & L_{mnrs}^{CSSC} & L_{mnrs}^{SSSC} \\ L_{mnrs}^{CCCS} & L_{mnrs}^{SCCS} & L_{mnrs}^{CSCS} & L_{mnrs}^{SSCS} \\ L_{mnrs}^{CCSS} & L_{mnrs}^{SCSS} & L_{mnrs}^{CSSS} & L_{mnrs}^{SSSS} \end{pmatrix}\begin{pmatrix} CC_{mn} \\ SC_{mn} \\ CS_{mn} \\ SS_{mn} \end{pmatrix} = \Lambda_2\begin{pmatrix} CC_{rs} \\ SC_{rs} \\ CS_{rs} \\ SS_{rs} \end{pmatrix}, \quad \begin{matrix} r=0,1,\ldots,K \\ s=0,1,\ldots,K \end{matrix} \quad (27)$$

where

$$L_{mnrs}^{CCCC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\cos 4n\theta)\times\cos rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SCCC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\cos 4n\theta)\times\cos rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{CSCC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\sin 4n\theta)\times\cos rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SSCC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\sin 4n\theta)\times\cos rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{CCSC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\cos 4n\theta)\times\sin rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SCSC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\cos 4n\theta)\times\sin rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{CSSC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\sin 4n\theta)\times\sin rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SSSC} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\sin 4n\theta)\times\sin rz\cos 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{CCCS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\cos 4n\theta)\times\cos rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SCCS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\cos 4n\theta)\times\cos rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{CSCS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\sin 4n\theta)\times\cos rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SSCS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\sin 4n\theta)\times\cos rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{CCSS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\cos 4n\theta)\times\sin rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SCSS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\cos 4n\theta)\times\sin rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{CSSS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\cos mz\sin 4n\theta)\times\sin rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

$$L_{mnrs}^{SSSS} = \int_{\theta=0}^{\frac{\pi}{2}}\int_{z=0}^{2\pi} L(p)(\sin mz\sin 4n\theta)\times\sin rz\sin 4s\theta \mathrm{d}z\mathrm{d}\theta$$

These equations can be further cast into a generalized linear algebraic eigenvalue problem of the form

$$[\mathbf{A} - \Lambda_2^{(K)}\mathbf{B}]\mathbf{X} = \mathbf{0} \quad (28)$$

where the superscript "(K)" signifies that the Fourier series is truncated to include K harmonic terms, $\mathbf{A}$ and $\mathbf{B}$ are matrices of dimension $(2K+1)^2\times(2K+1)^2$, and

$$\mathbf{X} = \{CC_{00}, CC_{01}, \ldots, CC_{KK}; SC_{10}, SC_{11}, \ldots, SC_{KK}; CS_{01}, CS_{02}, \ldots, CS_{KK}; SS_{11}, SS_{12}, \ldots, SS_{KK}\}^T$$

For system (28) to have nontrivial solutions, the determinant of the coefficient matrix must be zero, i.e.,

$$|\mathbf{A} - \Lambda_2^{(K)}\mathbf{B}| = 0$$

By solving this generalized eigenvalue problem, the moment Lyapunov exponent can be determined.

Having obtained an approximate result of the second-order perturbation $\Lambda_2^{(K)}$ of the moment Lyapunov exponent, an approximation of the moment Lyapunov exponent is given by

$$\Lambda_{x(t)}(p) \approx \varepsilon\Lambda_2^{(K)} \quad (29)$$

Using Eq. (4), an approximation of the Lyapunov exponent can be easily obtained

$$\lambda_{x(t)} \approx \varepsilon\lambda_2^{(K)}, \quad \lambda_2^{(K)} = \lim_{p\to 0}\frac{\Lambda_2^{(K)}}{p} \quad (30)$$

4 Parametric Resonances

4.1 Subharmonic Resonance. To study the subharmonic resonance in the first mode, the central frequency of the bounded noise is taken as $\nu_0 = 2\omega_1$. Assuming $\omega_1 \neq \omega_2$ and $2\omega_1 \neq \omega_2$, the coefficients in Eq. (25) are given by

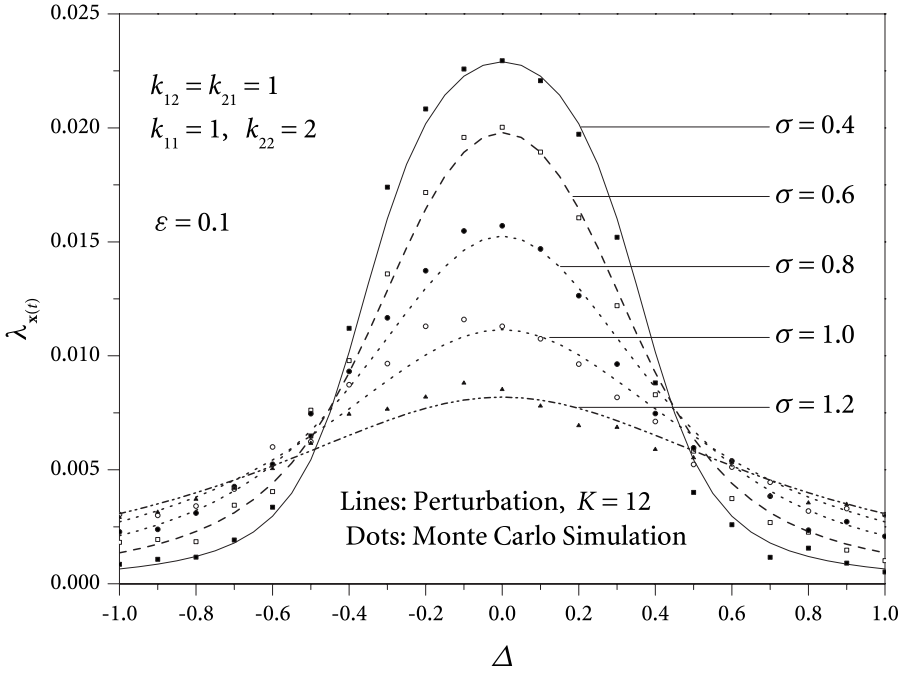


Fig. 1 Lyapunov exponent for $\beta_1=\beta_2=0$ and $\nu_0=2\omega_1$

$$\hat{h}_{11} = \tfrac{1}{4}k_{11}\cos z, \quad \hat{h}_{21} = 0$$

$$\hat{s}_1 = \tfrac{1}{8}k_{11}\sin 2\theta \sin z + \tfrac{1}{2}(\beta_1 - \beta_2)\sin 2\theta$$

$$\hat{q}_1 = -\tfrac{1}{8}k_{11}(1+\cos 2\theta)\sin z - \tfrac{1}{2}[(\beta_1+\beta_2)+(\beta_1-\beta_2)\cos 2\theta]$$

In fact, this is the singular case in which θ approaches zero as time goes to infinity since only one mode of the system is excited. Thus, the coefficients are

$$\hat{h}_{11} = \tfrac{1}{4}k_{11}\cos z, \quad \hat{h}_{21} = 0, \quad \hat{q}_1 = -\tfrac{1}{4}k_{11}\sin z - \beta_1, \quad \hat{s}_1 = 0$$

Hence, the eigenvalue problem is reduced to that of a single degree-of-freedom problem studied by Xie [9]. The eigenvalue problem can be solved using the Fourier series.

$$Z_0(z) = C_0 + \sum_{k=1}^{N}(C_k \cos kz + S_k \sin kz) \tag{31}$$

where C_0, C_k, S_k, $k=1,2,\ldots,N$, are constant coefficients to be determined. Following the same procedure shown in Sec. 3, the moment Lyapunov exponent can be determined. Hence, the Lyapunov exponent can be evaluated from Eq. (4), which is shown in Fig. 1.

For the numerical simulation, the original equation can be discretized using the Euler scheme for iterations $n=0,1,2,\ldots$,

$$x_1^{n+1} = x_1^n + \omega_1 x_2^n \cdot \Delta t$$

$$x_2^{n+1} = x_2^n + [-\omega_1 x_1^n - 2\varepsilon\beta_1 Y_2^n - \varepsilon\cos\ \eta^n(k_{11}x_1^n + k_{12}x_3^n)] \cdot \Delta t$$

$$x_3^{n+1} = x_3^n + \omega_2 x_4^n \cdot \Delta t$$

$$x_4^{n+1} = x_4^n + [-\omega_2 x_3^n - 2\varepsilon\beta_2 x_4^n - \varepsilon\cos\ \eta^n(k_{21}x_1^n + k_{22}x_3^n)] \cdot \Delta t$$

$$\eta^{n+1} = \eta^n + \nu \cdot \Delta t + \varepsilon^{1/2}\sigma \cdot \Delta W^n$$

These equations can be simulated iteratively and the numerical algorithm for determining the Lyapunov exponents [17] can be applied to evaluate $\lambda_{x(t)}$. In the Monte Carlo simulation, the time step is chosen as $\Delta t = 10^{-6}$, the frequencies are $\omega_1=1$ and $\omega_2=4$, and the number of iterations is 2×10^9. A comparison of the Lyapunov exponents $\lambda_{x(t)}$ obtained by perturbation and Monte Carlo simulation, as shown in Fig. 1, reveals that there is an excellent agreement between the two results. The moment Lyapunov exponents Λ_2 are shown in Fig. 2 for $\sigma=1.0$.

4.2 Combination Additive Resonance. To study the combination additive resonance, the central frequency of the bounded noise is taken as $\nu_0=\omega_1+\omega_2$. Assuming $\omega_1 \neq \omega_2$, the coefficients in Eq. (25) are given by

$$\hat{h}_{11} = \tfrac{1}{4}k_{12}\cos z \tan\theta, \quad \hat{h}_{21} = \tfrac{1}{4}k_{21}\cos z \cot\theta$$

$$\hat{s}_1 = -\tfrac{1}{8}[(k_{21}-k_{12})+(k_{12}+k_{21})\cos 2\theta]\sin z + \tfrac{1}{2}(\beta_1-\beta_2)\sin 2\theta$$

$$\hat{q}_1 = -\tfrac{1}{8}(k_{12}+k_{21})\sin 2\theta \sin z - \tfrac{1}{2}[(\beta_1+\beta_2)+(\beta_1-\beta_2)\cos 2\theta]$$

Note that k_{11} and k_{22} do not appear in the above coefficients, which means that they have no influence on the stability of the system in combination additive resonance. However, their influences have to be taken into account if ω_1 and ω_2 are close, i.e., $\omega_1+\omega_2 \approx 2\omega_1$. By a suitable scaling of coordinates, it is always possible to take $k_{12}=\pm k_{21}=k>0$ without the loss of generality. Hence, assume $k_{12}=k_{21}=1$ in the following analysis.

Following the same procedure as in the previous section, the moment Lyapunov exponents and Lyapunov exponents can be determined and are shown in Figs. 3–9.

The Lyapunov exponents obtained by perturbation and Monte Carlo simulation are shown in Fig. 3. The frequencies of the system are taken as $\omega_1=1$ and $\omega_2=4$, and the central frequency of the bounded noise is taken as $\nu_0=\omega_1+\omega_2=5$. From Fig. 3, one can clearly see that the parametric resonance occurs when the frequency detuning Δ is small ($\Delta<0.5$). Figure 4 shows the moment Lyapunov exponents for the undamped system. When the system is undamped, the pth moment Lyapunov exponent is positive for all $p>0$. It is seen that when the frequency detuning Δ is increased the effect of combination resonance is reduced, which is similar to the results of one degree-of-freedom system obtained in Ref. [9].

The numerical results of the moment Lyapunov exponents given by Eq. (29) converge when K is sufficiently large. In this

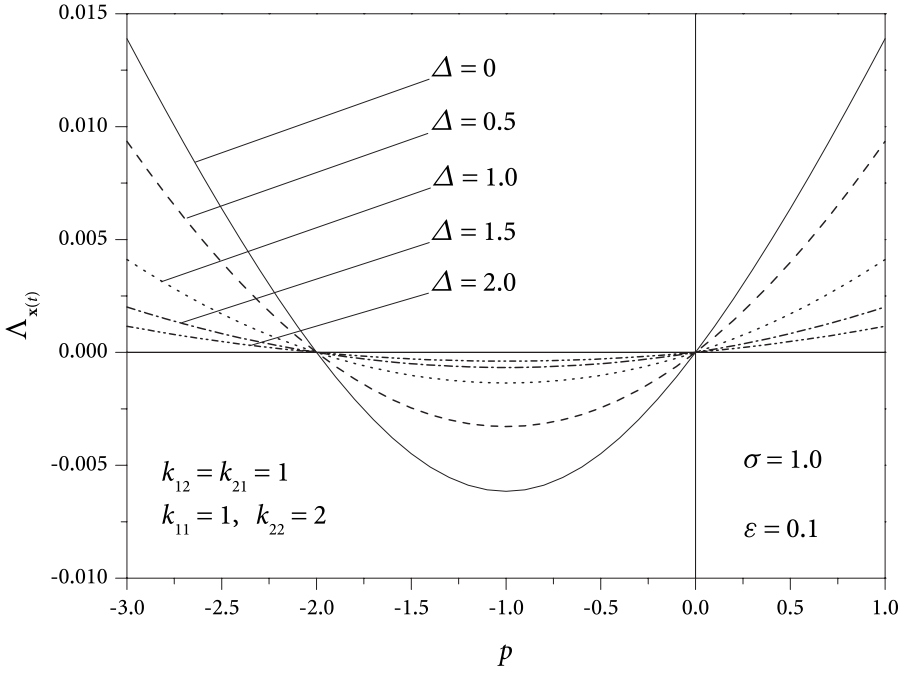


Fig. 2 Moment Lyapunov exponent for $\beta_1=\beta_2=0$, and $\nu_0=2\omega_1$ ($K=12$)

study, $K=20$ yields satisfactory results and is used to calculate the moment Lyapunov exponents for the undamped system. For the damped system, a larger value of K is needed for satisfactory results, as shown in Figs. 5 and 6. It is seen that the rate of convergence varies for different system parameters.

The Monte Carlo simulation procedure proposed by Xie and Huang [18] is applied to determine the pth moment Lyapunov exponents of system (6). The numerical results are presented in Figs. 8 and 9 for $\beta_1=0.1$ and 0.2, respectively, with various values of σ. In the simulation, the sample size is chosen as $S=5000$, the time step is $\Delta t=10^{-5}$, the number of iterations is 2×10^8, the total time of simulation is 2000, and the state vector is normalized after every time period of $t=20$. The approximate analytical results of $\Lambda_{x(t)}(p)$ are also plotted for comparison. The number of harmonic terms in the double Fourier series expansion is chosen as $K=40$. It is seen that the approximate analytical results agree quite well with the numerical results.

By increasing the value of σ, i.e., the bandwidth of the narrow-band excitation modeled by the bounded noise process increases, and the effect of combination resonance is reduced, as reflected in the decrease in the Lyapunov exponent and the moment Lyapunov

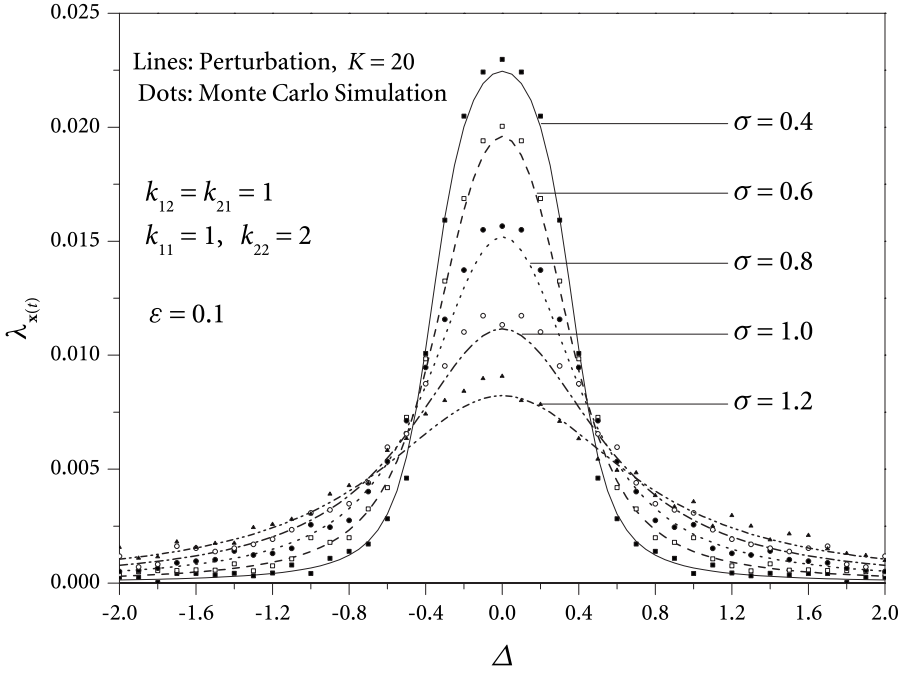


Fig. 3 Lyapunov exponent for $\beta_1=\beta_2=0$, and $\nu_0=\omega_1+\omega_2$

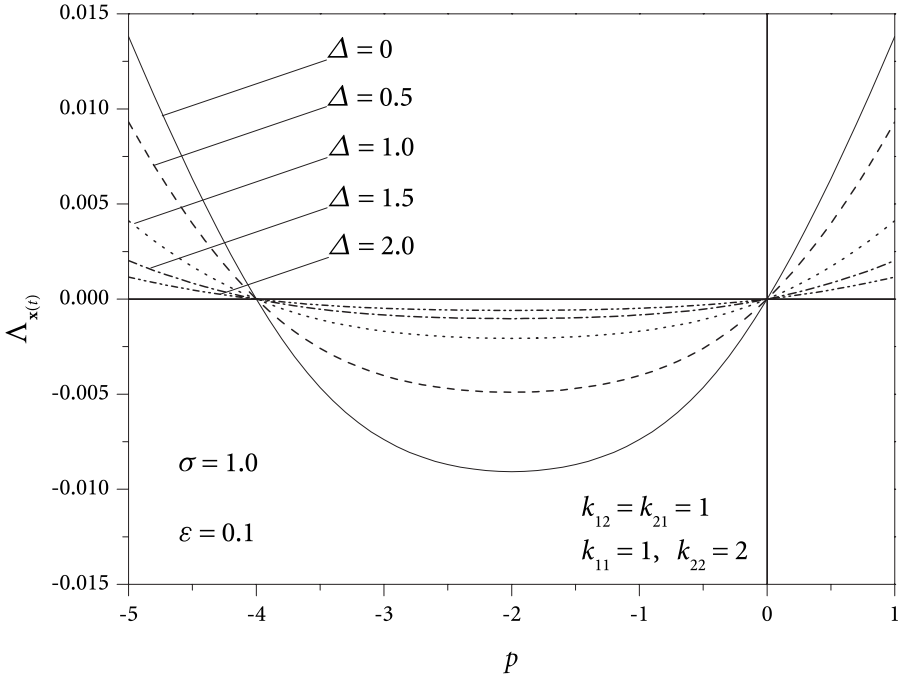


Fig. 4 Moment Lyapunov exponent for $\beta_1=\beta_2=0$ and $\nu_0=\omega_1+\omega_2$ ($K=20$)

exponents, and the increase in the stability index, as seen in Figs. 7–9. Figure 7 shows that Lyapunov exponent becomes negative for all values of Δ when σ is increased to 0.8, meaning that the system is stable almost surely. Furthermore, the system is unstable in the pth moment ($p>0$)) for small values of σ, as shown in Figs. 8 and 9. For instance, Fig. 9 shows that when $\sigma=1.5$ the system is unstable in the pth moment only if $p>7$. The effect of damping can be seen from the comparison of Figs. 8 and 9. At $\sigma=1.5$, the stability index is about $p=2$ for $\beta_1=\beta_2=0.1$ (Fig. 8). If β_1 is increased to 0.2, the stability index is about $p=7$. As expected, increasing the damping always stabilizes the system.

By varying the central frequency ν_0 to $|\omega_1-\omega_2|$ and following the same procedure, one can easily obtain the moment Lyapunov exponents for combination differential resonance.

4.3 Subharmonic and Combination Additive Resonance. Assuming $\omega_1\approx\omega_2$, one has $2\omega_1\approx 2\omega_2\approx\omega_1+\omega_2$. The central frequency of the bounded noise is taken as $\nu_0=\omega_1+\omega_2$. Here, the parameters of the system are $\omega_1=\omega_2=1$ and $\nu_0=2$. Thus, the coefficients in Eq. (25) are given by

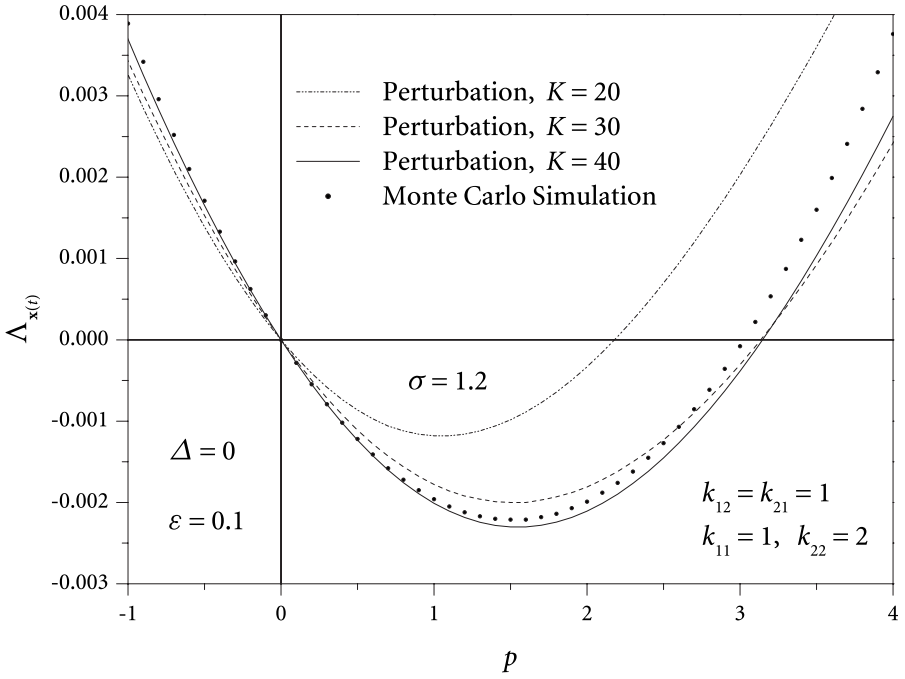


Fig. 5 Moment Lyapunov exponent for $\beta_1=0.2$, $\beta_2=0.1$, and $\nu_0=\omega_1+\omega_2$

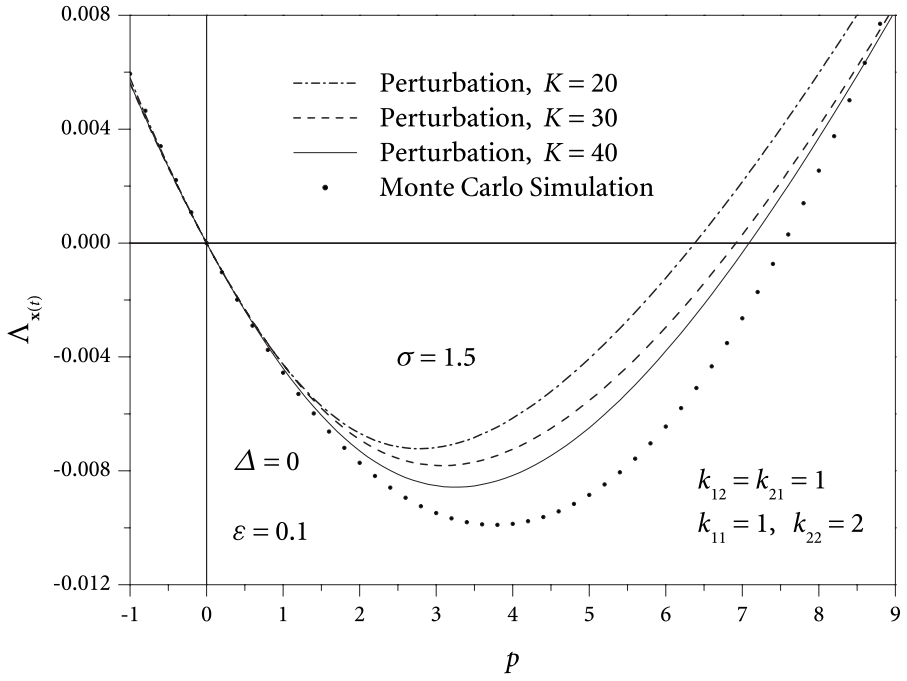


Fig. 6 Moment Lyapunov exponent for $\beta_1=0.2$, $\beta_2=0.1$, and $\nu_0=\omega_1+\omega_2$

$$\hat{h}_{11} = \tfrac{1}{4}(k_{11} + k_{12}\tan\theta)\cos z, \quad \hat{h}_{21} = \tfrac{1}{4}(k_{22} + k_{21}\cot\theta)\cos z$$

$$\hat{s}_1 = -\tfrac{1}{8}[(k_{21} - k_{12}) + (k_{22} - k_{11})\sin 2\theta + (k_{12} + k_{21})\cos 2\theta]\sin z + \tfrac{1}{2}(\beta_1 - \beta_2)\sin 2\theta$$

$$\hat{q}_1 = -\tfrac{1}{8}[(k_{11} + k_{22}) + (k_{11} - k_{22})\cos 2\theta + (k_{12} + k_{21})\sin 2\theta]\sin z - \tfrac{1}{2}[(\beta_1 + \beta_2) + (\beta_1 - \beta_2)\cos 2\theta]$$

The above expressions show that the coefficients are linear combinations of the coefficients for subharmonic and combination additive resonances. Thus, all k_{ij}'s contribute to the resonance of

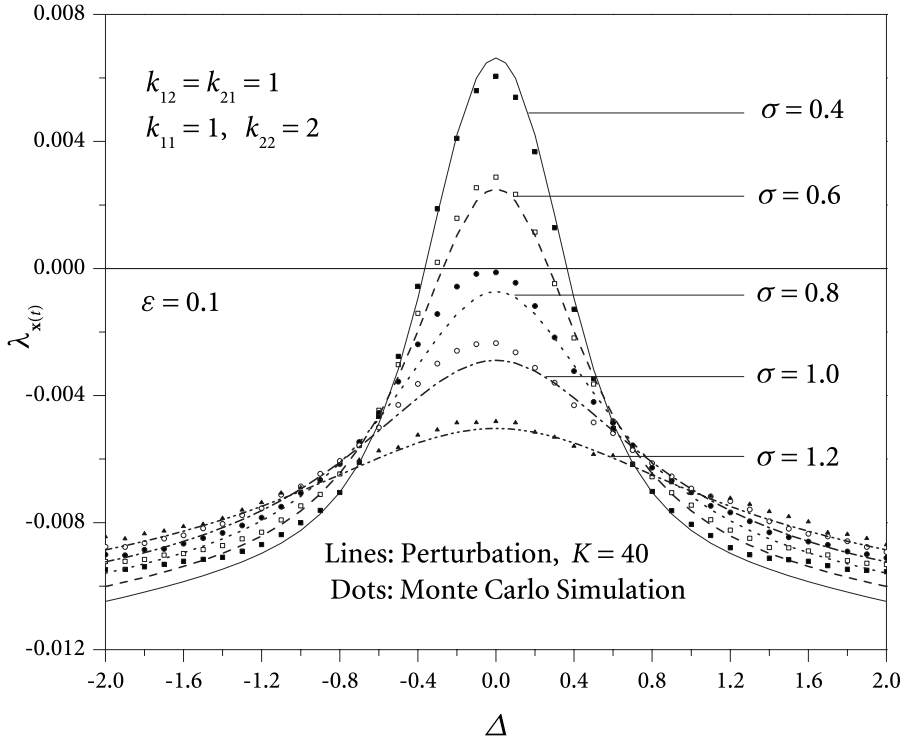


Fig. 7 Lyapunov exponent for $\beta_1=0.2$, $\beta_2=0.1$, and $\nu_0=\omega_1+\omega_2$

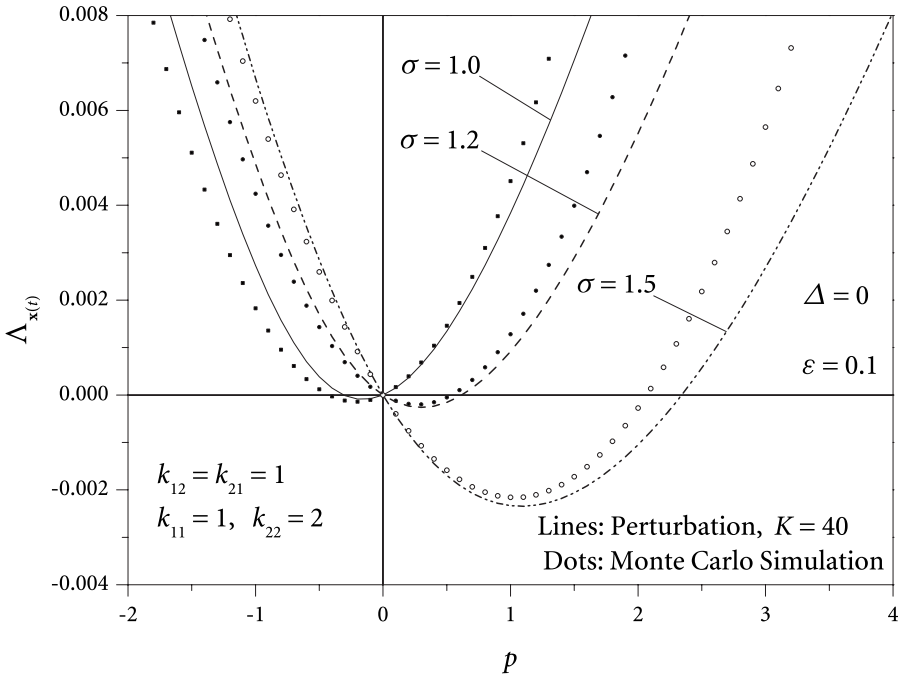


Fig. 8 Moment Lyapunov exponent for $\beta_1 = \beta_2 = 0.1$, and $\nu_0 = \omega_1 + \omega_2$

the system. Similarly, increasing the value of σ results in a more stable system, which can be seen from Figs. 10 and 11. One can also draw the conclusion that larger values of k_{22} yields more significant resonances since larger k_{22} means larger amplitude of the bounded noise. Figure 12 shows that the damping can reduce the resonance significantly and even stabilize the system when the value of σ is relatively large. For example, the Lyapunov exponents for $\sigma = 1.5$ are nearly all negative except those close to the central frequency. The moment Lyapunov exponents for the undamped system with $\sigma = 1$ and various detuning parameters Δ are shown in Fig. 13. The stabilization of damping can also be observed in Fig. 14 in terms of the moment Lyapunov exponents. Consistently, larger values of σ also make the system more stable in the pth moments for $p > 0$.

A comparison of Figs. 1, 3, and 11 indicates that the combined effect of subharmonic and combination additive resonances results in not only more significant resonance in terms of larger Lyapunov exponent than either case separately but also wider resonance region in terms of frequency detuning Δ. This means that the parametric resonance can be triggered more easily. For example, the resonance region for subharmonic resonance is

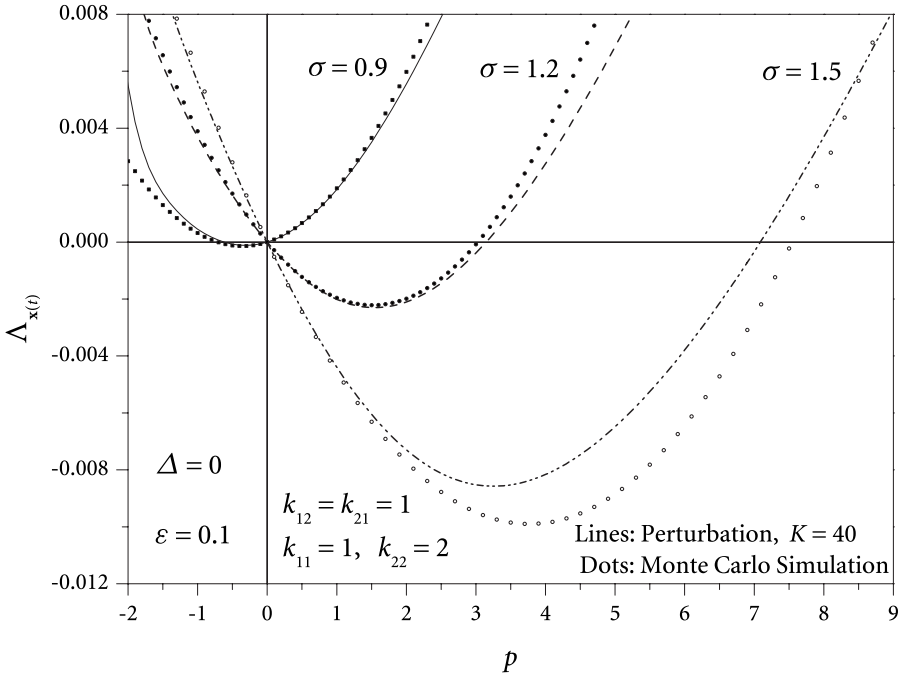


Fig. 9 Moment Lyapunov exponent for $\beta_1 = 0.2$, $\beta_2 = 0.1$, and $\nu_0 = \omega_1 + \omega_2$

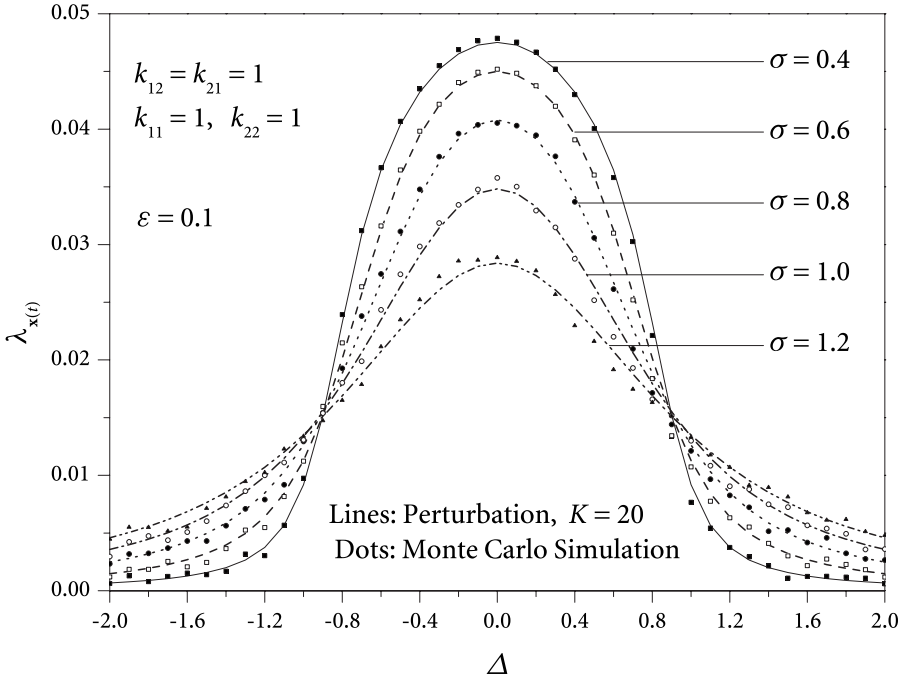


Fig. 10 Lyapunov exponent for $\beta_1=\beta_2=0$, $\omega_1=\omega_2$, and $\nu_0=2\omega_1$

around $|\Delta|<0.5$ from Fig. 1, while the region is expanded to $|\Delta|<1.0$ from Fig. 11 when the effects of subharmonic resonance and combination additive resonance are combined.

As shown above, the coefficients for subharmonic resonance is associated with k_{11} or k_{22} only, while the coefficients for combination additive resonance includes only k_{12} and k_{21}, which represent the coupling between the two degrees-of-freedom. When the subharmonic resonance and the combination additive resonance occur simultaneously, the resonance results have contributions from all k_{ij}. Depending on the coefficients, the convergent rates of the Fourier series are different. For subharmonic resonance, $K=12$ is sufficient to obtain satisfactory results. However, a relatively large K ($K=40$) is needed for the combination additive resonance with damping. Meanwhile, other parameters such as the damping coefficients $\beta_{1,2}$, noise intensity σ, and frequency detuning Δ also influence the rates of convergence (see, e.g., Figs. 5 and 6).

Larger values of k_{ij} yield more significant resonance since they are the amplitudes of the bounded noise. Decreasing the value of σ has the same effect on the stability because the power spectrum of the bounded noise is more narrowbanded, and the effect of

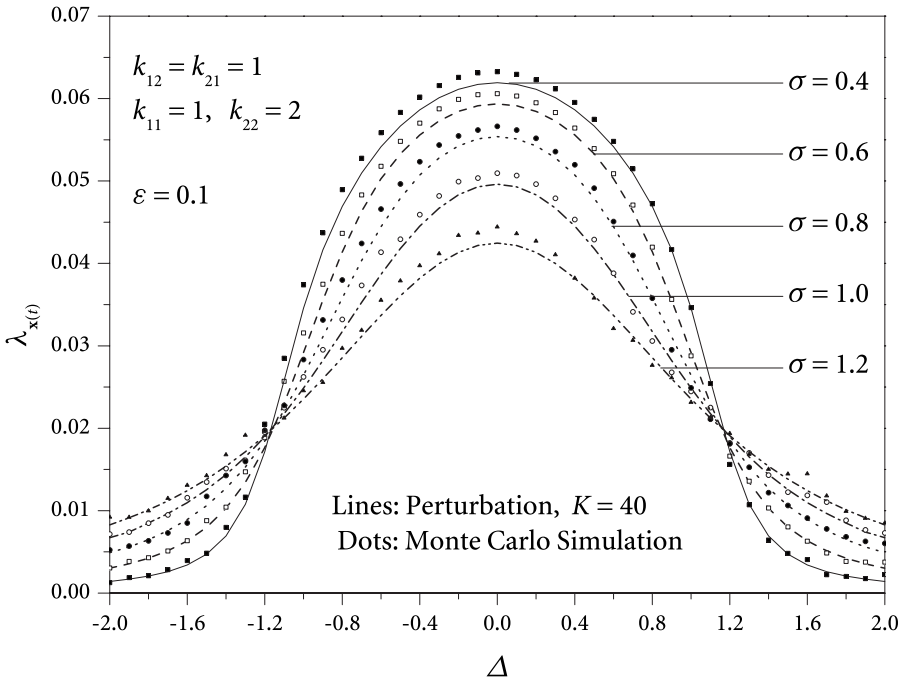


Fig. 11 Lyapunov exponent for $\beta_1=\beta_2=0$, $\omega_1=\omega_2$, and $\nu_0=2\omega_1$

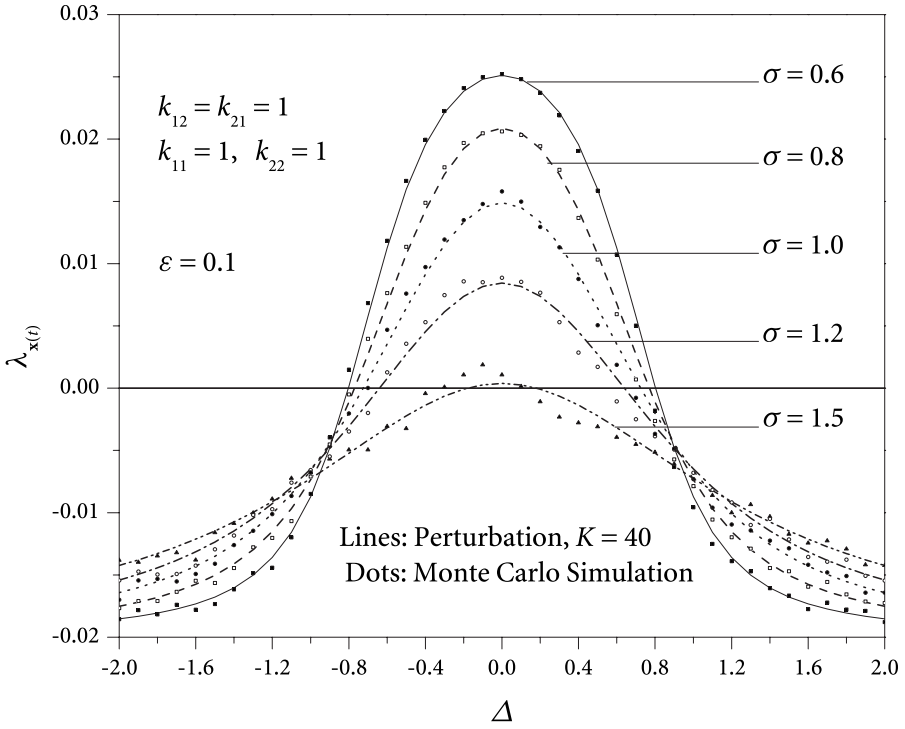


Fig. 12 Lyapunov exponent for $\beta_1=\beta_2=0.2$, $\omega_1=\omega_2$, and $\nu_0=2\omega_1$

parametric resonance is more prominent. The frequency detuning parameter Δ is also a key parameter to the stability of the system. When the central frequency of the bounded noise is offset from the resonance frequencies, the effect of parametric resonance is significantly reduced. The damping always has a stabilizing effect.

5 Conclusion

The dynamic stability of a two degrees-of-freedom system subjected to bounded noise excitation is studied by determining the moment Lyapunov exponents. The partial differential eigenvalue problem governing the moment Lyapunov exponent is established using the theory of stochastic dynamical system. For weak noise excitations, a singular perturbation method is employed to obtain second-order expansions of the moment Lyapunov exponents. A double Fourier series is used to solve the eigenvalue problem. The Lyapunov exponent is then obtained using the relationship between the moment Lyapunov exponent and the Lyapunov exponent. The accuracy of the approximate analytical results are vali-

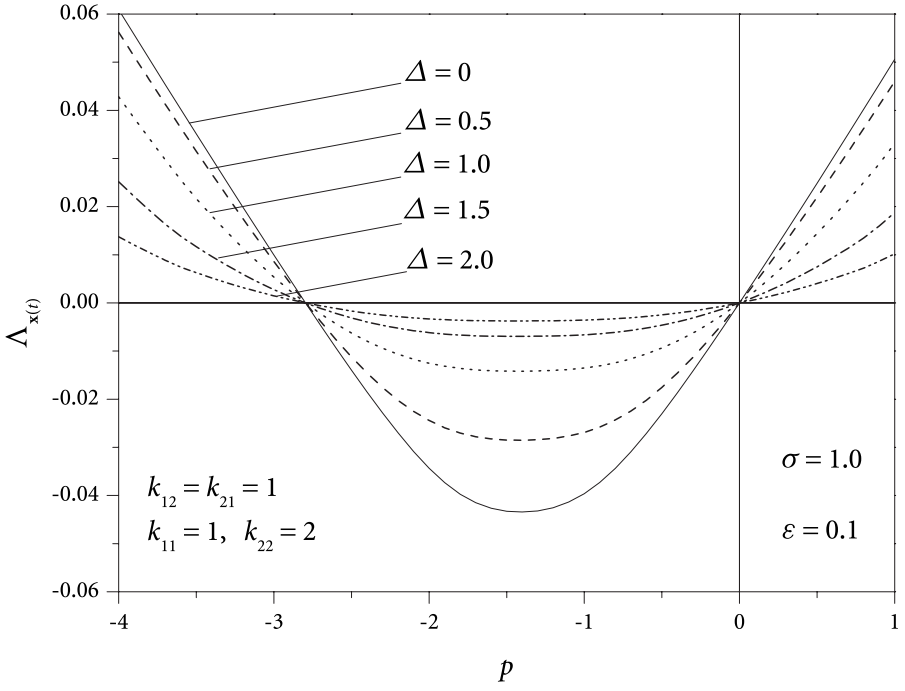


Fig. 13 Moment Lyapunov exponent for $\beta_1=\beta_2=0$, $\omega_1=\omega_2$, and $\nu_0=2\omega_1$ ($K=20$)

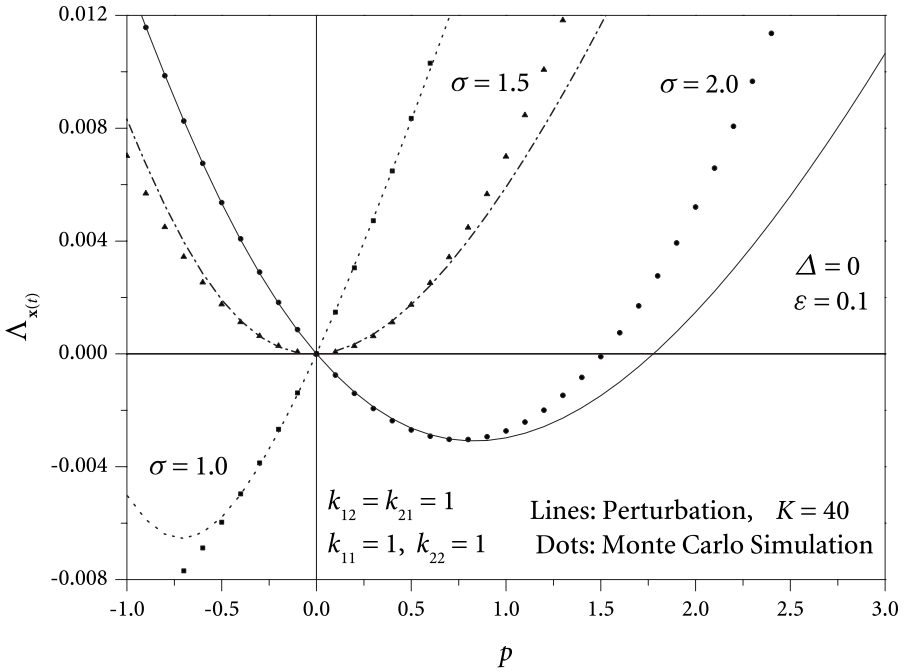


Fig. 14 Moment Lyapunov exponent for $\beta_1 = \beta_2 = 0.2$, $\omega_1 = \omega_2$, and $\nu_0 = 2\omega_1$

dated and assessed by comparing with those obtained using Monte Carlo simulation. It is observed that there is an excellent agreement between the analytical results and the numerical results. The effect of noise on various parametric resonances, such as subharmonic resonance, combination additive resonance, and combined subharmonic and combination additive resonance, is investigated. The effects of parameters on the stability are studied.

Acknowledgment

Financial supports from the Department of Civil and Environmental Engineering, Faculty of Engineering, University of Waterloo, the Natural Sciences and Engineering Research Council of Canada through Grant No. OGPO131355, and Research Grant Council of the Government of the HKSAR under Project Nos. PolyU5307/03E and PolyU5321/04E are gratefully acknowledged.

This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca).

References

[1] Arnold, L., 1984, "A Formula Connecting Sample and Moment Stability of Linear Stochastic Systems," SIAM J. Appl. Math., **44**, pp. 793–802.
[2] Sri Namachchivaya, N., and Vedula, L., 2000, "Stabilization of Linear Systems by Noise: Application to Flow Induced Oscillations," Dyn. Stab. Syst., **15**(2), pp. 185–208.
[3] Lin, Y. K., and Li, Q. C., 1993, "New Stochastic Theory for Bridge Stability in Turbulent Flow," J. Eng. Mech., **119**(1), pp. 113–121.
[4] Li, Q. C., and Lin, Y. K., 1995, "New Stochastic Theory for Bridge Stability in Turbulent Flow: II," J. Eng. Mech., **121**, pp. 102–116.
[5] Poirel, D., and Price, S. J., 2003, "Random Binary (Coalescence) Flutter of a Two-Dimensional Linear Airfoil," J. Fluids Struct., **18**, pp. 23–42.
[6] Arnold, L., Oeljeklaus, E., and Pardoux, E., 1986, "Almost Sure and Moment Stability for Linear Itô Equations," *Lyapunov Exponents* (Lecture Notes in Mathematics Vol. 1186), L. Arnold and V. Wihstutz, eds., Springer-Verlag, Berlin, pp. 85–125.
[7] Arnold, L., Kliemann, W., and Oeljeklaus, E., 1986, "Lyapunov Exponents of Linear Stochastic Systems," *Lyapunov Exponents* (Lecture Notes in Mathematics Vol. 1186), L. Arnold and V. Wihstutz, eds., Springer-Verlag, Berlin, pp. 129–159.
[8] Xie, W.-C., 2001, "Moment Lyapunov Exponents of a Two-Dimensional System Under Real Noise Excitation," J. Sound Vib., **239**(1), pp. 139–155.
[9] Xie, W.-C., 2003, "Moment Lyapunov Exponents of a Two-Dimensional System Under Bounded Noise Parametric Excitation," J. Sound Vib., **263**(3), pp. 593–616.
[10] Zhu, J., Wang, X. Q., Xie, W.-C., and So, R. M. C., 2008, "Flow-Induced Instability Under Bounded Noise Excitation in Cross-Flow," J. Sound Vib., **312**, pp. 476–495.
[11] Sri Namachchivaya, N., and Van Roessel, H. J., 2001, "Moment Lyapunov Exponent and Stochastic Stability of Two Coupled Oscillators Driven by Real Noise," ASME J. Appl. Mech., **68**, pp. 903–914.
[12] Sri Namachchivaya, N., and Van Roessel, H. J., 2004, "Stochastic Stability of Coupled Oscillators in Resonance: A Perturbation Approach," ASME J. Appl. Mech., **71**, pp. 759–768.
[13] Arnold, L., 1998, *Random Dynamical Systems*, Springer-Verlag, Berlin.
[14] Xie, W.-C., 2006, *Dynamic Stability of Structures*, Cambridge University Press, New York.
[15] Zauderer, E., 1989, *Partial Differential Equations of Applied Mathematics*, 2nd ed., Wiley, New York.
[16] Myint-U, T., and Debnath, L., 2007, *Linear Partial Differential Equations for Scientists and Engineers*, 4th ed., Birkhäuser, Boston.
[17] Wolf, A., Swift, J., Swinney, H., and Vastano, A., 1985, "Determining Lyapunov Exponents From a Time Series," Physica D, **16**, pp. 285–317.
[18] Xie, W.-C., and Huang, Q., 2009, "Simulation of Moment Lyapunov Exponents for Linear Homogeneous Stochastic Systems," ASME J. Appl. Mech., **76**, p. 031001.

Delamination Susceptibility of Coatings Under High Thermal Flux

Z. Xue
School of Engineering and Applied Sciences,
Harvard University,
29 Oxford Street,
Cambridge, MA 02138

A. G. Evans
Department of Mechanical Engineering,
University of California,
Santa Barbara, CA 93106

J. W. Hutchinson
School of Engineering and Applied Sciences,
Harvard University,
29 Oxford Street,
Cambridge, MA 02138

Delamination of coatings initiated by small cracks paralleling the free surface is investigated under conditions of high thermal flux associated with a through-thickness temperature gradient. A crack disrupts the heat flow thereby inducing crack tip stress intensities that can become critical. A complete parametric dependence of the energy release rate and mode mix is presented in terms of the ratio of the crack length to its depth below the surface and coefficients characterizing heat transfer across the crack and across the gaseous boundary layer between the surface and the hot gas. Proximity to the surface elevates the local temperature, which in turn, may significantly increase the crack driving force. A detailed assessment reveals that the energy release rates induced by high heat flux are capable of extending subsurface delaminations in thermal barrier coatings, but only when the modulus has been elevated by either calcium-magnesium-alumino-silicate (CMAS) penetration or sintering. Otherwise, the energy release rate remains well below the toughness. [DOI: 10.1115/1.3086590]

1 Introduction

When brittle coatings function in the presence of thermal gradients and high heat flux, they are susceptible to delamination and spalling. The most widely investigated examples are thermal barrier coatings (TBCs) used in turbines for aeropropulsion and power generation. Articles that describe and analyze the mechanisms capable of providing sufficient energy release rate, G_{delam}, to drive delamination have been presented [1–4]. They are in two basic mechanism categories, governed by the sign of the stress in the coating at its surface. (i) When this stress is tensile, sufficient energy release rate of delamination cracks parallel to the surface only arises when a secondary crack perpendicular to the delamination links it to the free surface allowing the stress above the delamination to be released. Similarly, if the delamination connects to a free edge the coating can displace in mixed mode as the delamination extends [1]. The scenario providing the largest G_{delam} is that involving a temperature gradient during operation having sufficient magnitude to induce an appreciable tensile stress at the surface upon cooling to ambient [2,5]. Other mechanisms that generate tension at the surface are not sufficiently potent: these include sintering-induced stresses. (ii) When the surface is in compression, an energy release rate for an edge delamination crack still exists (albeit that it is strictly mode II). In this case, if the delamination extends to a free edge, or is connected to the surface by a wide cracklike gap, the stress in the coating above the delamination can be released as the delamination extends. In the absence of cracks linking to the surface or free edges, the energy above the delamination can be released by buckling, but this requires a large initial delamination to already exist, formed by some (independently specified) mechanism. Another potent mechanism involves rapid heat-up in the presence of a subsurface flaw (Fig. 1). In this scenario, the flaw is thermally insulting, resulting in a temperature difference, ΔT_{flaw}, between its two faces (with compression at the coating surface). This ΔT_{flaw} induces an energy release rate of the cracklike flaw. Earlier estimates [3] implied that G_{delam} was too small to be of concern when temperature boundary conditions pertain at the top and bottom of the coating set by temperatures computed assuming no flaw. However, in most TBC applications, heat transfer boundary conditions apply such that the temperature at the surface above the crack is considerably higher than in the absence of a crack, rendering this mechanism more potent. Such boundary conditions are investigated here. Indeed, in high heat flux tests, dramatic spalling of the coating has been observed during heat-up [6]. The intent of this article is to re-examine rapid heating with the new boundary conditions. The emphasis will be on relatively small flaws (less than the coating thickness) to be compatible with the foregoing observations.

The analysis to be presented regards the coating as elastic (no creep) with isotropic thermal and elastic properties. Insight will be acquired by solving a sequence of increasingly complex thermal problems (Fig. 1), starting with an isolated crack in an infinite body, followed by a crack in a semi-infinite coating, and, finally, a crack in a finite coating on a conducting substrate. The problems differ thermally from that considered previously by virtue of the heat transfer boundary condition between the hot gas and the coating surface. It will become apparent that this situation causes the material above the crack to become hotter than in the absence of the crack, thereby elevating the energy release rate. The analysis will include the presence of an initial residual stress in the coating, at ambient, again to be consistent with the practical situation.

The application of the results will be illustrated for a TBC on a superalloy substrate. The coating will have thickness, $H=1$ mm, consisting of yttria-stabilized zirconia (YSZ) deposited by air plasma spray (APS) with thermal conductivity, $k=1$ W/m K, thermal expansion coefficient, $\alpha=11$ ppm/°C, and in-plane modulus, $E=20$ GPa [5]. For coatings infiltrated by a glasslike substance due to debris ingested into the engine, such as CMAS to be discussed later, the assumption of isotropic thermal and elastic properties is probably justified. On the other hand, elastic and thermal isotropy is obviously an approximation for uninfiltrated plasma spray coatings, but necessary at this stage since details of the anisotropies are not yet available. The heat flux boundary condition at the surface has such a strong effect on the crack driving force that results based on the idealized isotropic coating provide considerable insight. Similarly, detailed knowledge of the fracture anisotropy of the coatings is not yet known although the morphology of the microstructure of plasma spray coatings is expected to

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 6, 2008; final manuscript received January 26, 2009; published online April 22, 2009. Review conducted by Martin Ostoja-Starzewski.

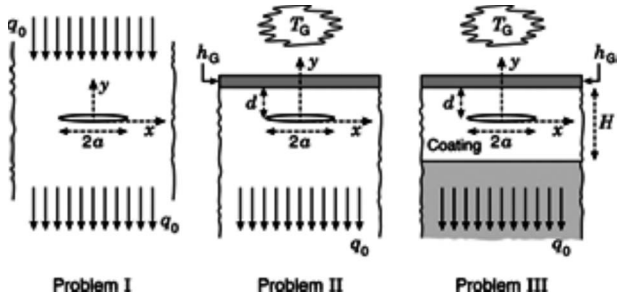


Fig. 1 Three problems analyzed in the paper. The coefficient of heat transfer across the crack is denoted by h_C in each of the three problems. The heat transfer coefficient across the gas boundary layer at the surface of the thermal barrier coating in Problems II and III (shown as a shaded layer) is denoted by h_G. The temperature of the hot gas above the boundary layer is T_G. In all three problems, the vertical heat flux in the absence of the crack is denoted by q_0.

give rise to some anisotropy, which would have some effect on delamination trajectories, as will be remarked on in the summary discussion.

2 Problem Statement

Each of the three problems depicted in Fig. 1 is infinite in extent in the x-direction. Throughout the paper, $T_0(y)$ denotes the temperature distribution in the absence of the crack. Under the steady state thermal conditions considered here, $T_0(y)$ satisfies $\nabla^2 T_0 = d^2T_0/d^2y = 0$. The heat flux through the coating when no crack is present is denoted by

$$q_0 = kdT_0/dy \tag{1}$$

with k as the thermal conductivity.

Denote the heat transfer coefficient governing the thermal conduction across the crack by h_C, such that the downward thermal flux across it is $h_C(T^+ - T^-)$ expressed in terms of the temperatures on its top (+) and bottom (−) faces. The condition

$$k\frac{\partial T}{\partial y} = h_C(T^+ - T^-) \tag{2}$$

must be satisfied at all points with $\partial T/\partial y$ continuous across the crack. For Problems II and III, the role of heat transfer through the gaseous boundary separating the surface of the coating and the hot gas at temperature T_G is taken into account. Denote the heat transfer coefficient for the boundary layer by h_G, such that at any point along the surface with temperature T_{surface} the heat flux into the surface is

$$k\frac{\partial T}{\partial y} = h_G(T_G - T_{\text{surface}}) \tag{3}$$

The steady-state temperature distributions for the three problems in the absence of the crack are as follows. For Problem I, the heat flux, q_0, is specified

$$T_0(y) = T_0(0) + \frac{q_0 y}{k} \tag{4}$$

with $T_0(0)$ having no influence. For Problem II, both q_0 and T_G are specified

$$T_0(y) = T_{0\text{surface}} + \frac{q_0(y-d)}{k}, \quad T_{0\text{surface}} = T_G - \frac{q_0}{h_G} \tag{5}$$

For Problem III, T_G and $T_{\text{substrate}}$ are specified

$$T_0(y) = T_{0\text{surface}}\left(\frac{H-d+y}{H}\right) + T_{0\text{interface}}\left(\frac{d-y}{H}\right) \tag{6}$$

with $T_{0\text{interface}}$ as the temperature at the coating/substrate interface. The heat flux and surface temperature are

$$q_0 = \frac{\mathrm{B}_G}{(\mathrm{B}_G + 1 + \mathrm{B}_G\ H_S/H)}\frac{k(T_G - T_{\text{substrate}})}{H} \tag{7}$$

and

$$T_{0\text{surface}} = T_G - \frac{q_0}{h_G}$$

where the dimensionless Biot number for the gaseous boundary layer is defined as

$$\mathrm{B}_G = \frac{h_G H}{k} \tag{8}$$

A sense for the numerical values of the variables is provided for the TBC example cited in Sec. 1. With $T_G - T_{0\text{surface}} = 400\,^\circ\text{C}$ and $T_{0\text{surface}} - T_{0\text{interface}} = 400\,^\circ\text{C}$ as the temperatures drops across the boundary layer and coating, respectively, the parameters become $h_G = 1\ \text{kW/m}^2\,\text{K}$, $\mathrm{B}_G = 1$, and $q_0 = 0.4\ \text{MW/m}^2$.

Regardless of the constraint at infinity, only two nonzero thermal-stress components result from $T_0(y)$: $\sigma_{xx}(y)$ and $\sigma_{zz}(y)$. These components induce zero energy release rate for a crack parallel to the surface. Similarly, residual stresses parallel to the surface have no effect on the energy release rate of cracks parallel to the surface. *Consequently, only the temperature change,* $\Delta T(x,y)$, *due to the presence of the crack produces stress intensities*, where

$$\Delta T(x,y) = T(x,y) - T_0(y) \tag{9}$$

Steady-state requires $\nabla^2 \Delta T(x,y) = 0$. If the crack does not impede the heat flux ($h_C = \infty$), then $\Delta T(x,y) = 0$ and the energy release rate is zero.

3 Isolated Crack in an Infinite Body

Results are reviewed for the plane strain problem of a crack in an infinite body (Problem I, Fig. 1) subject to a downward-directed vertical heat flux, $q_0 = kdT_0/dy$, remote from the crack. In the limit of no heat transfer across the crack ($h_C = 0$), a closed form solution exists [7,8]. With K_{I} and K_{II} denoting the mode I and II stress intensity factors, the problem is pure mode II ($K_{\text{I}} = 0$), with energy release rate given by

$$G_0(0) = \frac{\pi}{16}\Omega\left(\frac{a}{H}\right)^3, \quad \Omega = \frac{EH^3\alpha^2 q_0^2(1+\nu)}{k^2(1-\nu)} \tag{10}$$

where $2a$ is the crack length, E and ν are Young's modulus and Poisson's ratio of the material, α its coefficient of thermal expansion, and H is a characteristic dimension (later equated to the coating thickness). The rationale for the notation $G_0(0)$ will become apparent; the subscript 0 signifies an isolated crack. The result (Eq. (10)) has the notable features that the energy release rate is sensitive to crack length, $G_0 \sim a^3$, and the heat flux, $G_0 \sim q_0^2$. These strong dependencies often dominate the incidence of delamination, as elaborated later.

If $h_C \neq 0$, the problem is still pure mode II, but the energy release rate now depends on a dimensionless Biot number defined by

$$\mathrm{B}_C^* = \frac{h_C a}{k} \tag{11}$$

This problem does not have a closed form solution. A numerical result [9] for the energy release rate is plotted in Fig. 2. Over the range of Biot numbers plotted ($0 \le \mathrm{B}_G^* \le 2$), an accurate fit to the numerical results is

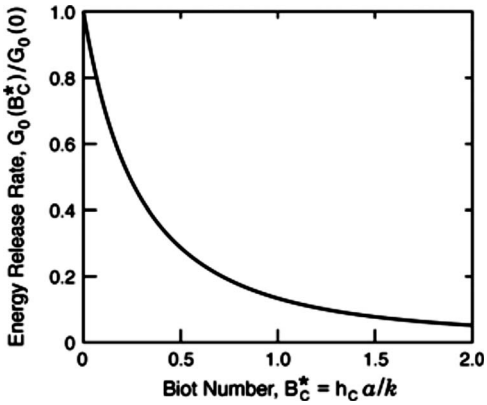


Fig. 2 Effect of heat conduction across an isolated crack on energy release rate [9]. The normalizing value $G_0(0)$ is the energy release rate for the nonconducting crack given by Eq. (10). Pure mode II pertains ($\psi=90$ deg).

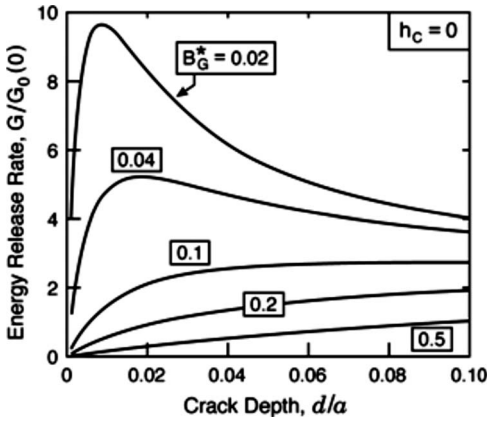


Fig. 3 Amplification of energy release rate for shallow cracks, as predicted by the asymptotic result for the energy release rate of a crack just below the surface as dependent on the Biot number governing heat transfer through the gas boundary layer, $B_G^*=h_G a/k$. The crack length is $2a$ and the depth below the surface is d. The crack is nonconducting ($h_C=0$) and has mode mix $\psi=52.1$ deg. The normalizing value, $G_0(0)$, pertains to the isolated crack experiencing the same overall heat flux given by Eq. (10).

$$G_0(B_C^*) = \frac{G_0(0)}{\left(1 + 3.709\ B_C^* + \left(\frac{\pi}{2}B_C^*\right)^2\right)} \tag{12}$$

The Poisson's ratio dependence is precisely captured through Eq. (10), and the limit for large B_C^*, $G_0(0)/(\pi\ B_C^*/2)^2$, has been obtained by a rigorous perturbation expansion. A closely related result for a crack on an interface will be reported in Sec. 5. The dependence of the energy release rate on crack length is dominated by that implicit in $G_0(0)$.

4 Near-Surface Cracks in a Semi-Infinite Coating

In Problem II depicted in Fig. 1, a crack of length $2a$ lies at a depth d below the surface of the semi-infinite body. The role of heat transfer through the gaseous boundary separating the surface of the coating and the hot burning gas at temperature T_G is taken into account, as discussed in Sec. 1. The temperature distribution in the absence of the crack, $T_0(y)$, in Problem II is given by Eq. (5) with q_0 and T_G specified. In the presence of a crack, $T_0(y)$ is approached remotely from the crack. Two normalizations of the Biot number for the boundary layer will be useful

$$B_G = \frac{h_G H}{k} \quad \text{and} \quad B_G^* = \frac{h_G a}{k} \tag{13}$$

The first normalization employing the layer thickness, H, was introduced in Eq. (7). Even though H does not enter directly in Problem II, B_G obtains for an actual coating, as illustrated by the estimate of B_G obtained below Eq. (8), and then $B_G^*=(a/H)B_G$.

Plane strain conditions governing the changes in stress and strain due to the presence of the crack are assumed. As noted earlier, only the change in temperature ΔT due to the presence of the crack generates stress intensities. Moreover, in Problem II, ΔT decays to zero remotely such that the change in remote stress is not influenced by the constraints on deformation.

4.1 The Nonconducting Shallow Crack Limit ($h_C=0$, $d/a \ll 1$). The only analytical result that has been possible to derive is the asymptotic solution for a shallow nonconducting crack ($h_C=0$, $d/a \ll 1$) (see the Appendix). The result is

$$G = \frac{1}{2}\frac{(1+\nu)}{(1-\nu)}Ed\left(\frac{\alpha q_0 a}{k}\right)^2\left[\frac{\pi}{4} + \frac{1}{B_G^*}\left(1 - \frac{\tanh(\sqrt{B_G^*\ a/d})}{\sqrt{B_G^*\ a/d}}\right)\right]^2 \tag{14}$$

with

$$\psi = \tan^{-1}\left(\frac{K_{II}}{K_I}\right) = 52.1\ \text{deg} \tag{15}$$

Normalizing G by $G_0(0)$ in Eq. (10) for the isolated crack of length $2a$ subject to the same overall heat flux q_0 gives

$$\frac{G}{G_0(0)} = \frac{8}{\pi}\frac{d}{a}\left[\frac{\pi}{4} + \frac{1}{B_G^*}\left(1 - \frac{\tanh(\sqrt{B_G^*\ a/d})}{\sqrt{B_G^*\ a/d}}\right)\right]^2 \tag{16}$$

which is plotted in Fig. 3. The most important observation is the existence of a range of shallow locations for which $G/G_0(0)$ can be significantly in excess of unity if $B_G^* \equiv a h_G/k \leq 0.1$. The implication is that cracks near the surface can become critical at considerably lower overall heat flux than deeper cracks. This feature arises because disruption of the heat flow substantially elevates the surface temperature just above the shallow crack. This elevation increases the compressive stress in the ligament above the crack, which, in turn, increases the energy release rate. In the extreme, the local surface temperature can approach T_G.

4.2 The Nonconducting Near-Surface Crack ($h_C=0$). Finite element thermal-stress analyses for all cases were carried out using ABAQUS/Standard software [10]. The deformation is taken to be plane strain, and the material is represented by the linear elasticity. Utilizing symmetry, only the half of the geometry to the right of the symmetry line ($x=0$) was analyzed. Symmetric boundary conditions were applied on the symmetry line. The distance to right edge is taken to be sufficiently large compared with both the crack length and the total thickness in the y-direction, such that the intensity factors are independent of this distance. The right edge is taken to be traction free. The heat transfer boundary condition on the top surface was specified through ABAQUS's thermal load option *SFILM, while a fixed uniform temperature was applied to the bottom surface. The crack was modeled as two separate surfaces with a small gap. The thermal conduction condition across the crack can be specified through ABAQUS's thermal contact option *GAP CONDUCTANCE. The meshes were devised to give highly accurate results for the energy release rate established by comparison with known results, such as those for the isolated crack in Eqs. (10) and (12). In particular, the result for the isolated crack with partial thermal conductivity in Eq. (12) and Fig. 2 was validated to within a fraction of a percent. A highly refined mesh is laid out on the ligament ahead of the tip. Eight

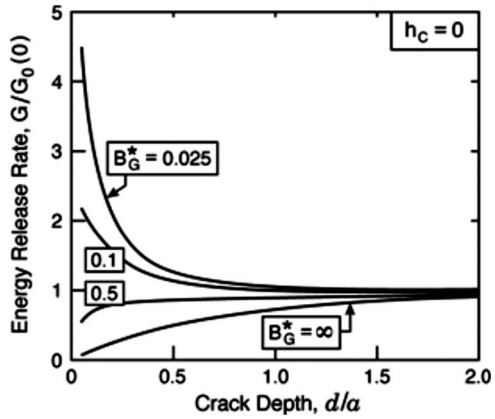


Fig. 4 Normalized energy release rate computed numerically for nonconducting crack of length $2a$ located a distance d below the surface for various Biot numbers, $B_G^* = h_G a/k$, characterizing the gaseous boundary layer at the surface. The normalizing value, $G_0(0)$, pertains to the isolated crack experiencing the same overall heat flux given by Eq. (10), i.e., the limit $d/a \gg 1$.

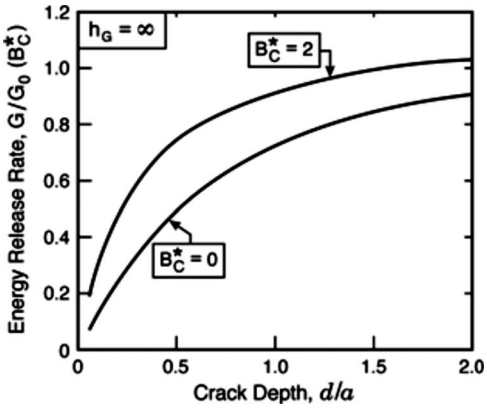


Fig. 6 Normalized energy release rate computed numerically for limit ($h_G = \infty$ and $T_{surface} = T_G$) with no gaseous boundary layer. The Biot number governing heat transfer across the crack is $B_C^* = h_C a/k$, and $G_0(B_C^*)$ is given by Eq. (12). The mode mix, ψ, is plotted in Fig. 5.

node quadrilateral elements with reduced integration were employed for the fully coupled thermal-stress analyses. Such elements exploit the biquadratic shape function for the displacement and a bilinear shape function for the temperature. The crack tip was modeled with a ring of collapsed quadrilateral elements to capture the strain singularity, thus improving the accuracy of the calculations, (see ABAQUS manual for details). At the crack tip, the elemental size is on the order of one-hundredth of the crack depth.

Numerical results for G and ψ computed for $0.05 \leq d/a \leq 2$ and selected B_G^* are plotted in Figs. 4 and 5. They affirm the trends exposed by the asymptotic formulas, Eqs. (14) and (15), and reveal that cracks with $d/a < 1/2$ can experience energy release rates well above those for deep cracks if $B_G^* \leq 0.1$. It appears that the mode mix, ψ, is independent of B_G^*, but it has not been possible to establish this theoretically. The numerical results establish the range of validity of the asymptotic formulas is limited to $d/a \leq 0.1$.

The elevation of the energy release rate due to surface proximity only arises when there is a substantial temperature drop across the gaseous boundary layer (small B_G^*). Conversely, when $B_G^* \to \infty$ such that $T_{surface} = T_G$, the boundary layer is eliminated, and surface proximity reduces the energy release rate. The transition occurs for $B_G^* \cong 0.25$, where G is essentially independent of the crack depth.

For completeness results for the limit $h_G = \infty$, $T_{surface} = T_G$ is plotted in Fig. 6, affirming that proximity to the surface reduces G relative to that for the deep crack.

4.3 The Near-Surface Crack With Combinations of h_C and h_G. Selected results in Fig. 7 illustrate how the thermal conductivities of the boundary layer and the crack interact to determine the energy release rate. The plot quantifies trends that would be expected from the previous plots when either the crack is nonconducting or the boundary layer provides no thermal resistance. Specifically, elevation of the energy release rate due to proximity to the surface depends on both B_C^* and B_G^*. For $B_C^* = 0.2$, appreciable elevation near the surface only occurs if $B_G^* \leq 0.1$. For $B_C^* = 0.5$, it does not occur for any $B_G^* \geq 0.025$ (plot not shown).

5 Cracks in a Coating on a Substrate

When the crack is short ($a/H \ll 1$) and relatively near the surface, the results of Sec. 4 apply. Otherwise, for Problem III, interaction with the substrate must be taken into account. In thermal barrier systems, the conductivity of the metal substrate is typically at least over an order of magnitude greater that of the coating, enabling the temperature along the bottom surface of the sub-

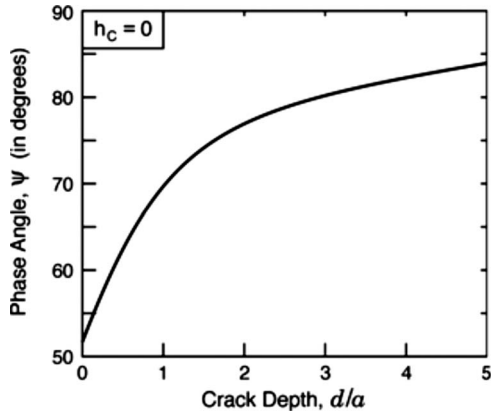


Fig. 5 Mode mix, ψ, computed numerically for nonconducting crack of length $2a$ located a distance d below the surface. The curve applies to all Biot numbers, $B_G^* = h_G a/k$, characterizing the gaseous boundary layer at the surface. This result also applies for Problem II for any combination of B_C^* and B_G^*.

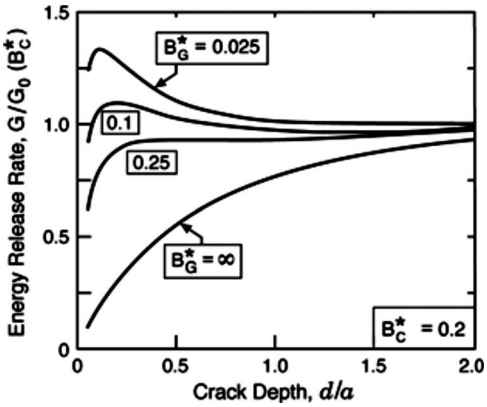


Fig. 7 Normalized energy release rate computed numerically for $B_C^* = h_C a/k = 0.2$ and various $B_G^* = h_G a/k$ with $G_0(B_C^*)$ given by Eq. (12). The mode mix, ψ, is plotted in Fig. 5.

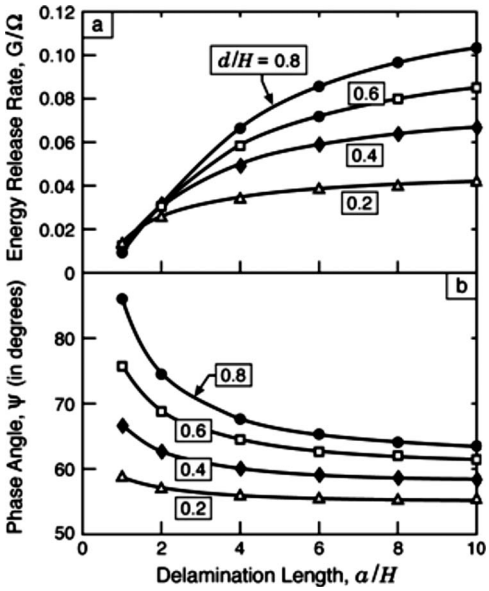


Fig. 8 Energy release rate and mode mix in Problem III computed numerically for cracks at various depths for $B_C=0$ and $B_G=1$. $H_S/H=3$, $k_{substrate}/k=100$, $E_{substrate}/E=5$, $\nu_{substrate}=0.3$, and $\nu_{TBC}=0.2$.

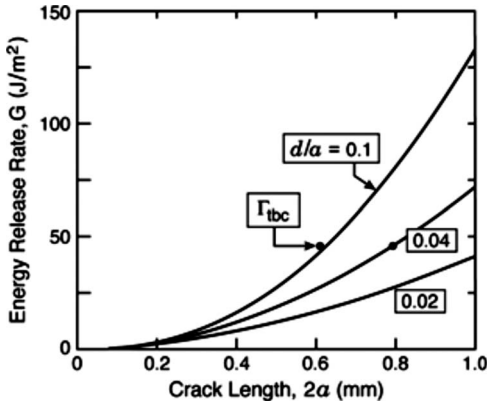


Fig. 9 Energy release rate computed numerically for shallow cracks at various depths for $B_C=0$ and $B_G=1$ for $q_0=0.4$ MW/m^2 and $E=200$ GPa ($H=1$ mm, $k=1$ W/m K, $\nu=0.2$, and $\alpha=11\times10^{-6}$/K). Critical flaw sizes are indicated based on a representative mode I toughness, Γ_{TBC}.

strate, $T_{substrate}$, to be nearly uniform. Interaction of a relatively short, deep crack ($a/H\ll1$) with the substrate significantly reduces the energy release rate. Detailed trends will not be presented, rather, one result will be quoted in Sec. 5.1.

5.1 Short Crack ($a/H\ll1$) at the Interface With the Substrate. Consider a plane strain crack of length $2a$ on the interface between two semifinite half spaces. The half space above the interface has moduli, coefficient of thermal expansion, and thermal conductivity taken to be the same as those in Problem I. The half space below the interface has identical moduli to that above the interface with an infinite thermal conductivity (recall that the thermal conductivity of the metal substrate is typically ten times that of the coating). The remote heat flow is q_0 and the heat transfer across the crack is h_C, as before.

It has not been possible to find the solution to this problem in literature, although a general solution for the interface crack with zero conductivity is available [11]. It is relatively straightforward to show that this problem with nonzero h_C is pure mode II, and that the stress intensity factor, K_{II}, is exactly half that for the isolated crack of length $2a$ in the uniform material (Problem I), except that the corresponding heat transfer coefficient across the crack must be taken as $2h_C$. (Solution details are omitted.) Thus, the short interface crack has a greatly reduced energy release rate given precisely by

$$G=\frac{1}{4}G_0(2\ \mathrm{B}_C^*) \tag{17}$$

where G_0 is give by Eq. (12). The result holds in the limit when there is no heat conducted across the crack, $h_C=0$, and agrees with the result in Ref. [11]. It does not depend on the coefficient of thermal expansion of the lower half space.

5.2 Long Crack ($a/H\gg1$). Numerical results for Problem III are presented in Fig. 8 for the specific case of a substrate with $H_S/H=3$, $k_{substrate}/k=100$, $E_{substrate}/E=5$, and $\nu_{substrate}=0.3$. The result depends on Poisson's ratio of the coating, taken as $\nu=0.2$. The two relevant Biot numbers are defined as $B_C=h_CH/k$ and $B_G=h_GH/k$. The energy release rate can be expressed in the form

$$G=\Omega F(a/H,d/H,B_G,B_C) \tag{18}$$

where F is a dimensionless function of the arguments shown. The trend with the normalized depth for a nonconducting crack ($B_C=0$) reveals that the surface enhancement of G only occurs for the shallow cracks considered in Sec. 3. For long cracks ($a/H\gg1$), G increases with depth below the surface. In all cases, G increases and ψ decreases with crack extension, although the asymptotic limit for very long cracks is nearly attained for $a/H=10$. The implication is that any crack that attains modest length, $a/H\approx1$, once critical, will propagate unstably without arrest.

6 Implications

The application of the foregoing results is illustrated for a YSZ-TBC on a superalloy substrate. The coating has thickness, $H=1$ mm, with thermal conductivity, $k=1$ W/m K, thermal expansion coefficient, $\alpha=11$ ppm/°C, and in-plane modulus, $E=20$ GPa [5]. The initial examples regard the temperatures drops across the boundary layer and coating as, respectively, $T_G-T_{0surface}=400$°C and $T_{0surface}-T_{0interface}=400$°C. Other values are invoked as the arguments emerge. The ensuing thermal parameters are $h_G=1$ kW/m^2 K, $B_G=1$, and $q_0=0.4$ MW/m^2. This set of parameters results in an energy release rate coefficient, $\Omega\approx580$ J m^{-2}. While this is quite large relative to the mode I toughness of the TBC ($\Gamma\approx50$ J m^{-2}) [12,13], note that the actual energy release rate is much smaller because of qualifying terms substantially less than unity, as elaborated below. A preamble before proceeding is that the phase angle, in all cases, is in the range $50\leq\psi\leq90$ deg. Consequently for a medium with isotropic fracture resistance, the crack would extend diagonally down through the coating (not parallel to the substrate). Fracture anisotropy, if sufficient, could result in parallel delaminations, but this seems unlikely given the degree of mode mixity unless the anisotropy is quite large. The exception is delaminations in the coating just above the substrate. These could oscillate in the coating as they extend (on average) parallel to the surface, in accordance with the appropriate mixed mode toughness. Given that this situation is the most realistic, it is considered first. The preliminary estimates assume an insulating crack ($h_c=0$) to obtain the maximum possible energy release rates.

The most basic result is the trend in $G_0(0)$ as a function of crack length, ascertained for 50 μm$\leq2a\leq1$ mm, for various levels of heat flux within the range $0.4\leq q_0\leq2$ MW/m^2 (Fig. 9). To interpret these plots, the fracture toughness must be superposed. For this purpose, the mode I toughness of YSZ ($\Gamma\approx50$ J m^{-2}) [12,13] has been included in the figure. This choice

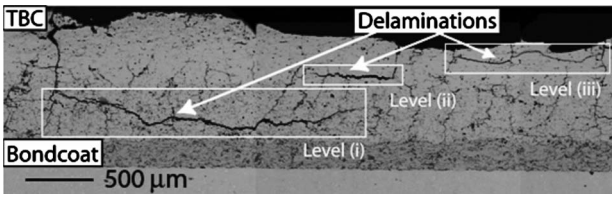


Fig. 10 Delaminations in the TBC on an engine shroud [5]. The delaminations just below the surface occur within the CMAS infiltrated regions, which give rise to a significantly increased Young's modulus.

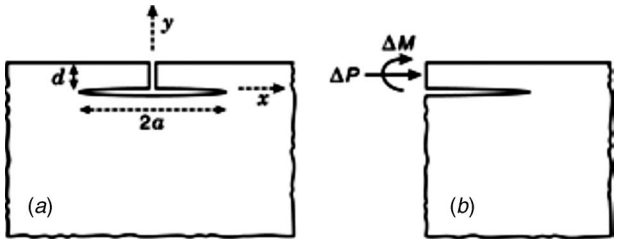


Fig. 11 (*a*) Cut above the crack creating a simply connected region. (*b*) Resultant force/length and moment/length required to eliminate displacement discontinuity across the cut in (*a*).

again represents a worst case from a delamination susceptibility perspective, since the delaminations are mixed mode with appreciably higher toughness. Recall that, for short cracks, $G = G_0(0)/4$, it is apparent that, even at the highest heat flux, the energy release rate only becomes sufficiently large to attain the toughness when the delaminations exceed about 1 mm in length ($a/H > 1$). The same conclusion is reached by referring to Fig. 7, recalling that $\Omega \approx 580$ J m^{-2}. Such long cracks do not pre-exist in these systems, but could form due to other thermomechanical phenomena [1,5,14]. Once delaminations of this length have been created just above the substrate, the heat flux induced energy release rate would lead to catastrophic extension, with associated spalling.

The situation for short cracks just below the surface can be judged by combining the information contained in Figs. 3, 4, and 7 with that in Fig. 8. An example is presented for short cracks, $0.1 \le a/H \le 0.5$, just beneath the surface, $0.01 \le d/a \le 0.1$. The result for a representative flux, $q_0 = 0.4$ MW/m^2, and a conventional choice of the modulus ($E = 20$ GPa) indicates that G always remains below the mode I toughness. The corresponding result for a case wherein the top of the TBC has either sintered or been filled with CMAS ($E = 200$ GPa) is plotted in Fig. 9. For this case, G can exceed Γ for shallow cracks in the length and depth ranges $2a \approx 0.6$ mm and $d \approx 30$ μm. Namely, moderately long shallow cracks are susceptible to delamination. Moreover, recalling that increasing the heat flux by a factor 1.5 would increase G by a factor 2.25 (because of the scaling, $\Omega \sim q_0^2$) infers that situations capable of generating an extreme heat flux would allow short shallow cracks to delaminate. Indeed, delaminations of this type reported in CMAS infiltrated airfoils (type (iii) in Fig. 10) [5] had previously defied explanation. In summary, high heat flux appears to be capable of extending subsurface delaminations, but only when the modulus has been elevated by either CMAS penetration or sintering. Otherwise, the energy release rates remain well below the toughness, unless large delaminations have already formed by other mechanisms.

Acknowledgment

Assistance in identifying relevant thermal crack solutions existing in the literature from J.R. Barber and S. Dag is gratefully acknowledged.

Appendix: Asymptotic Analysis for Shallow Nonconducting Cracks—Problem II

Problem II in Fig. 1(*b*) in the limit $d/a \ll 1$ is considered for plane strain cracks with $h_C = 0$. The first step in the analysis is to obtain the temperature distributions above and below the crack. Because the layer above the crack is thin and because no heat is conducted across the crack, the y-dependence of the temperature in this layer is negligible. Conservation of heat under the steady-state conditions of interest requires

$$\frac{d^2T}{d^2x} - \frac{h_G}{kd}T = -\frac{h_G}{kd}T_G, \quad |x| < a, \quad 0 \le y \le d \tag{A1}$$

Zones whose size is of order d exist at the ends of the crack, in which the temperature transitions from that in the strip to the surface temperature $T_{0\text{surface}}$ are given by Eq. (5). This zone shrinks to zero as d/a becomes small. These zones are ignored and the boundary conditions for Eq. (A1) are taken as $T(\pm a) = T_{0\text{surface}}$. Thus, the temperature distribution above the crack is

$$T = T_G - \frac{q_0}{h_G}\frac{\cosh(\sqrt{\mathrm{B}_G^*\, a/d}\, x/a)}{\cosh\sqrt{\mathrm{B}_G^*\, a/d}}, \quad |x| < a, \quad 0 \le y \le d \tag{A2}$$

An approximation to the temperature distribution on the bottom surface of the crack also exploits the facts that the crack is thermally insulating and $d/a \ll 1$. Consider the thermal problem for the half space below the crack line along $y = 0$. For $|x| < a$, $\partial T/\partial y = 0$. For $|x| > a$, $T = T_0(0)$, where $T_0(0)$ is given by Eq. (5), with transition zones of order d between these two conditions at the crack ends. If these zones are ignored, the conditions along $y = 0$ are identical to those for the classical problem of an isolated nonconducting crack subject to remote heat flux, q_0, in an infinite plane. The temperature distribution just below the crack is

$$T = T_0(0) - \frac{q_0 a}{k}\sqrt{1 - \left(\frac{x}{a}\right)^2}, \quad |x| < a, \quad y = 0^- \tag{A3}$$

The change in temperature due to presence of the crack, $\Delta T = T - T_0(y)$, determines the stress intensity factors

$$\Delta T = \frac{q_0}{h_G}\left[1 - \frac{y}{a}\mathrm{B}_0^* - \frac{\cosh(\sqrt{\mathrm{B}_G^*\, a/d}\, x/a)}{\cosh\sqrt{\mathrm{B}_G^*\, a/d}}\right], \quad |x| < a, \quad 0 \le y \le d \tag{A4}$$

$$\Delta T = -\frac{q_0 a}{k}\sqrt{1 - \left(\frac{x}{a}\right)^2}, \quad |x| < a, \quad y = 0^- \tag{A5}$$

Because the temperature change ΔT satisfies Laplace's equation, the associated strains generated under plane strain

$$\varepsilon_{xx} = \varepsilon_{yy} = \alpha(1+\nu)\Delta T, \quad \varepsilon_{xy} = 0 \tag{A6}$$

are compatible, producing no stress within the simply connected region shown in Fig. 11, a created by a cut along the y-axis above the crack. However, the displacements derived from these strains are discontinuous across the cut—it is the enforcement of their continuity that generates the stresses and stress intensity factors that arise from ΔT. The second step in the analysis is to compute the displacement discontinuity and then to determine the force/length, ΔP, and moment/length, ΔM, in Fig. 11(*b*). These are directly linked to the stress intensity factors by [15]

$$K_{\mathrm{I}} = \frac{1}{\sqrt{2d}}[\Delta P \cos\omega + 2\sqrt{3}\Delta M d^{-1}\sin\omega]$$

$$K_{II} = \frac{1}{\sqrt{2d}}[\Delta P \sin \omega - 2\sqrt{3}\Delta M d^{-1} \cos \omega] \tag{A7}$$

where $\omega = 52.1$ deg.

Given the symmetry of ΔT with respect to x and the fact that the associated shear strain, ε_{xy}, vanishes, one can readily show that the displacements derived from the strains in Eq. (A6) are such that u_y is continuous across the cut, while the discontinuity of u_x varies linearly across the cut. Denote the discontinuity across the cut by $[u_x] = 2\delta + 2\theta(0^+)y$, where $\theta(0^+)$ is the rotation of the cut surface on the right, taken as positive in the clockwise sense. The difference in average strain ε_{xx} from Eq. (A6) on the top and bottom surfaces of the crack obtained using Eqs. (A4) and (A5) gives

$$\frac{\delta}{a} = -\frac{(1+\nu)\alpha q_0}{h_G}\left[B_G^*\left(\frac{\pi}{4} + \frac{d}{a}\right) + 1 - \frac{\tanh\sqrt{B_G^* \, a/d}}{\sqrt{B_G^* \, a/d}}\right] \tag{A8}$$

The rotation discontinuity is given by $\theta(0^+) = \theta(a) + \int_0^a \kappa dx$, where $\theta(a)$ is the rotation at the right hand crack tip and κ is the curvature of the upper crack surface. Because ΔT varies linearly with y in the layer above the crack in Eq. (A4), the latter is immediately obtained using Eq. (A6) with

$$\kappa = \frac{(1+\nu)\alpha q_0}{k} \tag{A9}$$

Next, $\theta(a)$ can be obtained using the fact that $\partial\theta/\partial x = u_{x,xy} = (1+\nu)\alpha \partial \Delta T/\partial x$ along the lower surface of the crack because $\varepsilon_{xy} = 0$. Then, because symmetry dictates that the rotation vanishes at $x = 0$ below the crack, one obtains from Eq. (A5)

$$\theta(a) = -\frac{(1+\nu)\alpha q_0 a}{k} \tag{A10}$$

Together, Eqs. (A9) and (A10) give

$$\theta(0^+) = 0 \tag{A11}$$

The final step in the analysis is to enforce continuity of displacements across the cut by imposing ΔP and ΔM in Fig. 11(*b*). For slender layers ($d/a \ll 1$), the layer can be modeled as a plate clamped at its right end. The choices

$$\Delta P = -\frac{Ed}{(1-\nu^2)}\frac{\delta}{a} = \frac{E\alpha q_0 d}{(1-\nu)h_G}\left[B_G^*\left(\frac{\pi}{4} + \frac{d}{a}\right) + 1 - \frac{\tanh\sqrt{B_G^* \, a/d}}{\sqrt{B_G^* \, a/d}}\right], \quad \Delta M = 0 \tag{A12}$$

provide continuity by canceling δ in Eq. (A8) and $\theta(0^+)$ in Eq. (A11). Then, by Eq. (A7)

$$(K_I, K_{II}) = \frac{E\sqrt{d}}{\sqrt{2}(1-\nu)}\frac{\alpha q_0 a}{k}\left[\frac{\pi}{4} + \frac{d}{a} + \frac{1}{B_G^*}\left(1 - \frac{\tanh\sqrt{B_G^* \, a/d}}{\sqrt{B_G^* \, a/d}}\right)\right] \times (\cos\omega, \sin\omega) \tag{A13}$$

The term d/a in the square brackets above is negligible over the range of validity of Eq. (A13) and is on the order of terms already neglected in the analysis, thus it can be neglected. The results for G in Eq. (14) and ψ in Eq. (15) follow directly.

References

[1] Choi, S. R., Hutchinson, J. W., and Evans, A. G., 1999, "Delamination of Multilayer Thermal Barrier Coatings," Mech. Mater., **31**, pp. 431–447.

[2] Rabiei, A., and Evans, A.G., 2000, "Failure Mechanisms Associated With the Thermally Grown Oxide in Plasma-Sprayed Thermal Barrier Coatings," Acta Mater., **48**, pp. 3963–3976.

[3] Hutchinson, J. W., and Evans, A. G., 2002, "On the Delamination of Thermal Barrier Coatings in a Thermal Gradient," Surf. Coat. Technol., **149**, pp. 179–184.

[4] Evans, A. G., and Hutchinson, J. W., 2007, "The Mechanics of Coating Delamination in Thermal Gradients," Surf. Coat. Technol., **201**, pp. 7905–7916.

[5] Kraemer, S., Faulhaber, S., Chambers, M., Clarke, D. R., Levi, C. G., Hutchinson, J. W., and Evans, A. G., 2008, "Mechanisms of Cracking and Delamination Within Thick Thermal Barrier Systems in Aero-Engines Subject to Calcium-Magnesium-Alumino-Silicate (CMAS) Penetration," Mater. Sci. Eng., A, **490**, pp. 26–35.

[6] Johnson, C. A., 2007, GE Global Research Center, private communication.

[7] Florence, A. L., and Goodier, J. N., 1960, "Thermal Stress Due to Disturbance of Uniform Heat Flow," ASME J. Appl. Mech., **27**, pp. 635–639.

[8] Sih, G. C., 1963, "On the Singular Character of Thermal Stresses Near a Crack Tip," ASME J. Appl. Mech., **28**, pp. 587–589.

[9] Kuo, A. Y., 1990, "Effects of Crack Surface Heat Conductance on Stress Intensity Factors," ASME J. Appl. Mech., **57**(2), pp. 354–358.

[10] HKS, 2005, *ABAQUS/Standard User's Manual*, 6.5 ed., HKS, Providence, RI.

[11] Lee, K. Y., and Park, S.-J., 1995, "Thermal Stress Intensity Factors for Partially Insulated Interface Crack Under Uniform Heat Flow," Eng. Fract. Mech., **50**, pp. 475–482.

[12] Mercer, C., Williams, J. R., Clarke, D. R., and Evans, A. G., 2007, "On a Ferroelastic Mechanism Governing the Toughness of Metastable Tetragonal-Prime (t') Yttria-Stabilized Zirconia ," Proc. R. Soc. London, Ser. A, **463**, pp. 1393–1408.

[13] Lampenscherf, S., 2006, Siemens AG, private communication.

[14] Johnson, M., Ruud, C. A., Bruce, J. A., and Wortman, D., 1998, "Relationships between Residual Stress, Microstructure and Mechanical Properties of Electron Beam Physical Vapor Deposition Thermal Barrier Coatings," Surf. Coat. Technol., **108–109**, pp. 80–85.

[15] Hutchinson, J. W., and Suo, Z., 1992, "Mixed Mode Cracking in Layered Materials," Adv. Appl. Mech., **29**, pp. 63–191.

An Experimental Study on the Influence of Vortex Generators on the Shock-Induced Boundary Layer Separation at M=1.4

A. Zare Shahneh
e-mail: a.shahneh@qmul.ac.uk

F. Motallebi
e-mail: f.motallebi@qmul.ac.uk

School of Engineering and Material Sciences,
Queen Mary University of London,
London E1 4NS, U.K.

The results of an investigation into the effects that sub-boundary layer vortex generators (SBVGs) have on reducing normal shock-induced turbulent boundary layer separation are presented. The freestream Mach number and Reynolds number were $M=1.45$ and $R=15.9\times10^6/m$, respectively. Total pressure profiles, static pressure distributions, surface total pressure (Preston pressure) distributions, oil flow visualization, and Schlieren photographs were used in the result analysis. The effects of SBVG height and the location upstream of the shock were investigated. A novel tetrahedron shape SBVG with different lengths (30 mm and 60 mm) was used for these experiments. The effect of streamwise location of the longer SBVG on the interaction was also investigated. The location of the shock wave was controlled by an adjustable choke mechanism located downstream of the working section. The results show that an increase in the distance for the longer SBVG from $17.4\delta_R$ to $25.5\delta_R$ did not remove the separation entirely, but the shorter SBVG provided higher total pressure distribution within the boundary layer in the recovery region. This also provided a healthier boundary layer profile downstream of the interaction region with lower displacement thickness and shape factor.
[DOI: 10.1115/1.3086591]

1 Introduction

It is well known that the interaction of the shock wave with boundary layers could result in flow separation. The interaction between a turbulent boundary layer and a shock wave is a feature frequently encountered in modern aerodynamics. It occurs, for example, in the flow round an airfoil at transonic speeds and in the intakes of jet engines. It can be also responsible for a large loss in intake delivery pressure.

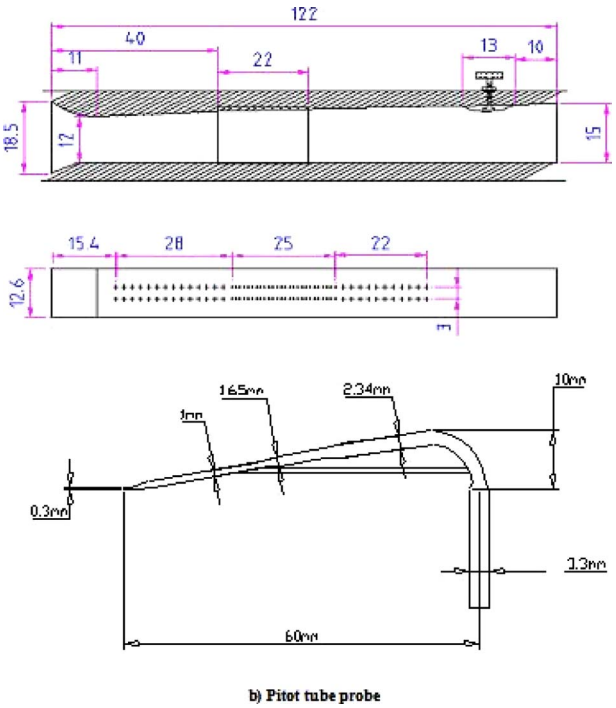
Numerous investigations have been reported by several authors to verify the validity of channel, cavity, bump, and vortex generator devices for the control of boundary layer separation. Wheeler invented three-sided submerged channels for the control of the flow separation [1]. Bur and Corbel investigated a bump system and compared it with a suction system in Onera [2]. A vortex generator achieves boundary layer control only at the penalty of considerable drag. A sub-boundary vortex generator produces vortices, which travel downstream along the surface, causing turbulence mixing between the boundary layer and the freestream, resulting in an unwanted drag increase but less than a vortex generator with the thickness of about boundary layer. Investigations on vane-type of vortex generators were carried out at NACA in 1950 [3], and many attempts have been made to state the effective role of vortex generator to mitigate buffeting, loss of control, and reduction of stability [4–7]. Later investigation on the effectiveness of the submerged vortex generators showed their effectiveness in controlling the separation of compressible turbulent boundary layers [8]. Wheeler [9] also introduced a vortex generator with the height of 60% of the local boundary layer thickness, which were positioned at about 22% of the chord. The result confirmed that it was able to reduce the drag of an LA2573 Liebeck low Reynolds number airfoil. Vane-type vortex generators are perhaps the most common type of SBVGs.

An extensive experimental investigation into the performance of the vane-type vortex generators was carried out at DERA Bedford boundary layer tunnel using 3D laser Doppler anemometry [10] in their wind tunnel. The nominal air speed in the tunnel was ranged between 10 m/s and 40 m/s. In these experiments, forward wedge, counter-rotating joined V-shape, and spaced counter-rotating vanes were investigated. They showed that SBVG devices with the heights of about one-quarter of the boundary layer can significantly reduce the size of the separated region. A recent survey has been done by Holden and Babinsky in a blowdown supersonic wind tunnel at the University of Cambridge [11] to investigate the effectiveness of the wedge-shaped and arrays of counter-rotating vanes in alleviating the boundary layer separation. The freestream Mach numbers were 1.5 and 1.3 at a unit Reynolds number of $28\times10^6/m$. Three shock positions have been set up in different experiments. The experiments showed that for the Mach number of 1.5, vane-type vortex generators can generate stronger vortices closer to the surface and therefore are more effective than wedge-type at eliminating the shock-induced boundary layer separation. With vortex generators, there are extra shock waves generated at the leading and trailing edges of the vortex generators and a strong re-expansion over the devices. This makes a higher total pressure loss leading to a higher wave drag.

Kuethe [12,13] and then Rao [14] investigated semicircular cross-section vortex generators, immersed within 65% or less of the boundary layer thickness, and a wire vortex generator was introduced by Ashil and Riddle [15]. The idea of applying pop-up models (circular and triangular in shape) has also been investigated [16,17]. For instance, in NASA, a set of cylinders was mounted horizontally to investigate the effect of these devices on flow separation [18]. None of wire or pop-up devices is satisfactory for the specific purpose of this research.

A review of different separation control devices [19] showed that low profile vortex generator and passive cavity are both capable to reduce separation; however, the authors suggested that

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 22, 2008; final manuscript received December 29, 2008; published online April 23, 2009. Review conducted by Nesreen Ghaddar.



b) Pitot tube probe

Fig. 1 Test section and Pitot tube

vortex generators cause more pressure recovery and therefore are more applicable for a supersonic diffuser, but passive causes can reduce wave drag and therefore is more applicable for an airfoil.

A comparison between Wheeler-type (small submerged "owe" shape) with vane-type vortex generators showed that the Vane-type counter-rotating vortex generator with $h/\delta=0.8$ provides the largest pressure recovery but also imposes the largest device drag penalty, but the Wheeler-type reduced the reattachment distance by up to 66% with higher pressure recovery. The 0.1δ height Wheeler-type performed as well as larger (0.4δ height) one.

This work is the latest streams of investigations that have been done by the authors to show the performance of tetrahedral vortex generator followed by the same purpose using flat vortex generator [20].

2 The Experiment Setup

The experiments were conducted in the M=1.4 supersonic induction wind tunnel at Queen Mary University of London. Although this Mach number is somewhat high for wing applications, the results obtained from the experiments are likely to be applicable to shock wave boundary layer interactions above M=1.3, in addition to shock-induced separations in supersonic intakes. The wind tunnel test section was of rectangular cross-section shape with glass windows on opposing walls to facilitate Schlieren and oil flow visualization studies. The tunnel width and height were 127 mm and 135 mm, respectively. Mean-flow pressure measurements were facilitated by the instrumented lower internal wall of the test section that was designed and manufactured for the purpose of these experiments. The instrumentation included static pressure tappings and Pitot-probe access ports distributed along the length of the liner, as shown in Fig. 1(*a*).

Two parallel lines of 50 static pressure tappings were located 15 mm either side of the liner centerline, the streamwise spacing between them being 10 mm. The static pressures could be recorded independently to record spanwise pressure variations, by blocking one tapping, or as an average from both tubes to a single transducer. A separate Pitot access port dummy-plug was made with a built-in static pressure tapping to facilitate the measurement of static pressure at three spanwise locations. The streamwise spacing of the measurement interval of the Pitot-probe was 30 mm. The Pitot-probe access ports were located along the centerline of the tunnel. From 15δ up to 40δ downstream of the interaction, the surface total pressure measurements were taken at larger intervals of 40 mm. The normal shock was held at the required location in the tunnel by adjusting a second adjustable throat; the tunnel injector pressure was kept constant at 760 KPa.

A roughness strip (1 mm in height and 2.5 mm in width) was used to trip wind tunnel bottom wall boundary layer to a fully developed turbulent boundary layer.

Total pressure profiles were measured using a flat ended Pitot-probe with tip dimensions of 0.3×1 mm (Fig. 1(*b*)).

A source of interference effects appears at supersonic flow, where a Pitot tube does not indicate the local total pressure. This is due to a detached shock wave stands ahead of the tube [21]. Here, the interference effect due to the detached shock wave is corrected.

The Pitot-probe was designed to enable detailed measurements to be made very close to the wall. This was made possible by flattening the tip of the probe to the smallest structurally allowable size, i.e., up to the limiting condition beyond which the tip sides were not sufficiently rigid to maintain a rectangular orifice shape. This ensured that the error induced by the normal velocity gradient present in a boundary layer would be kept at a minimum. The smallest possible external tip dimensions achieved were internal height of 0.1 mm, external height of 0.3 mm, and external width of 1.37 mm. The probe-tip design was based on the information provided by Ower and Pankhurst [22], Bryer and Pankhurst [23], and Motallebi [24]. The probe-tip was closely examined to ensure that no burrs from the manufacturing process obscured the rectangular orifice. No tip-displacement correction was applied to the vertical location of the Pitot tube tip [25].

In addition, a series of surface total pressure probes of 0.5 mm tip diameter was used to take measurements at smaller intervals than the Pitot ports allowed over the length of interest. All pressures were recorded using a Scanivalve digital pressure transducer array that had a measurement accuracy of 0.5%. The standard deviation of ten sets of pressure records was always less than 0.06%.

In order to obtain an appreciation of the changes to the flow structure with the different SBVG configurations, two forms of flow visualization were utilized. These were conventional black and white Schlieren imaging and surface oil flow visualization using a mixture of Shell Tellus oil, titanium dioxide, and oleic acid.

A computer controlled Digiplan stepper motor has been used to traverse the boundary layer probe. The accuracy of the movement of the stepper motor was better than ±0.075 mm, which is less than of 1% of undisturbed boundary layer physical thickness upstream of interaction zone. The standard deviation of ten sets of pressure records by Scanivalve for a certain point is in order of 0.06%.

The SBVG that has been used in these experiments has a tetrahedron geometry (Fig. 2). The model is a modified version of the DERA's proposed flat vane model [10]. The smaller SBVG (SS) had a length of 30 mm with a base dimension of 3×3 mm^2. The longer SBVG had a length of 60 mm with a base dimension of 5×3 mm^2. The trailing gap between both SBVGs was 5 mm. The larger SBVG was used at two streamwise locations upstream of the location of the shock wave in the freestream flow (i.e., 17.4δ and 25.5δ), while the shorter SBVG (SS) was placed at one location corresponding to 17.4δ upstream of the shock position. Table 1 shows the basic dimensions of the SBVGs used in these experiments.

3 Results and Discussion

In these experiments, the larger SBVG was placed at two different streamwise locations (130 mm or LS configuration and 190 mm or LL configuration). For the smaller SBVG (SS configura-

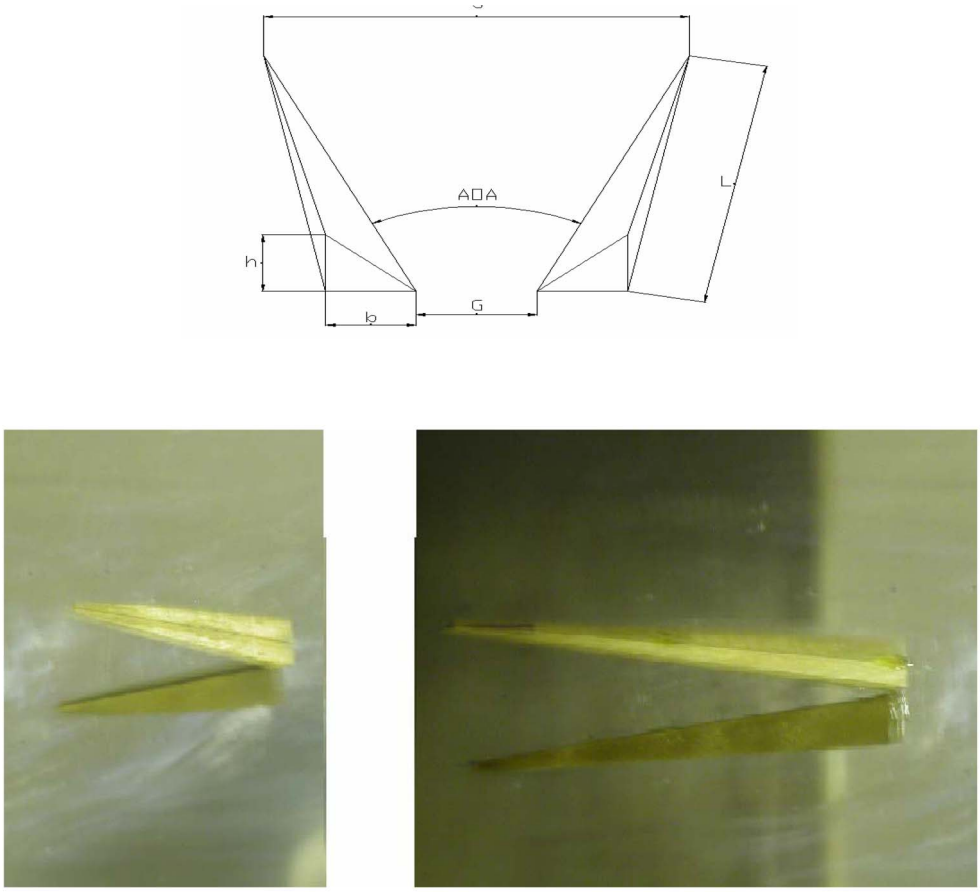


Fig. 2 SBVGs with 30 and 60 mm lengths. Flow direction is from left to right.

tion), only one streamwise location corresponding to 130 mm upstream of the shock wave was used. The distances 130 mm and 190 mm correspond to 17.4δ and 25.5δ, where δ is the physical thickness of the baseline boundary well upstream of the shock-boundary layer interaction zone.

Streamwise surface static pressure distribution for LS, LL, and SS configurations are compared with baseline (flow with no SBVG), and inviscid pressure jump across the normal shock wave is shown in Fig. 3. The point of "inflection" for the baseline produced roughly at the shock location ($x/\delta_R=0$). The static pressure distributions show quite similar behavior for all cases. All configurations start with an almost zero pressure gradient, a relatively sharp increase followed by a gradual increase further downstream of the shock wave, approaching to ideal pressure recovery. However, the LS configuration shows slightly better pressure recovery in the region about 10δ downstream of the shock wave.

The streamwise surface impact pressure (Preston pressure) distributions for all configurations are compared with the baseline distribution and is shown in Fig. 4. For all configurations, it shows the expected rapid decrease in the Preston pressure as the flow approaches the shock wave. However, the Preston pressure distribution for the LS configuration shows a higher level as compared with other configurations. As the flow passes the shock location, the streamwise Preston pressure starts to recover. Compared with the baseline case, all configurations (in particular, SS and LS) show a complete elimination of the separation region. It is also interesting to note the SS case provides the fastest recovery of the Preston pressure. This indicates that the trailing vortices produced by SS must have stayed close to the surface and therefore producing a fuller total pressure profile close to the surface. Both LS and LL configurations show inferior performance as compared with SS. Boundary layer velocity profiles for all configurations are shown in Figs. 5–12.

Figure 5 shows the boundary layer velocity profiles for LS, LL, SS, and baseline at $x/\delta_R=-5.36$. It clearly shows that the LL and SS configurations produce the larger deficit in mean velocity close to the surface, while the LS configuration produces less disturbances in the boundary layer velocity profile close to the surface.

Table 1 Basic Dimensions of the SBVGs and their relative locations upstream of the shock wave

SBVG	L (mm)	h (mm)	b (mm)	G (mm)	AOA (deg)	D^* (δ_R)	D^* (mm)
SS	30	3	3	3	29	17.4	130
LS	60	3	5	3	29	17.4	130
LL	60	3	5	3	29	25.5	190

SS=small and short, LS=large and short, LL=large, and long. δ_R is the reference boundary layer thickness equal to 7.1 mm; D is the distance between vortex generator and shock wave location.

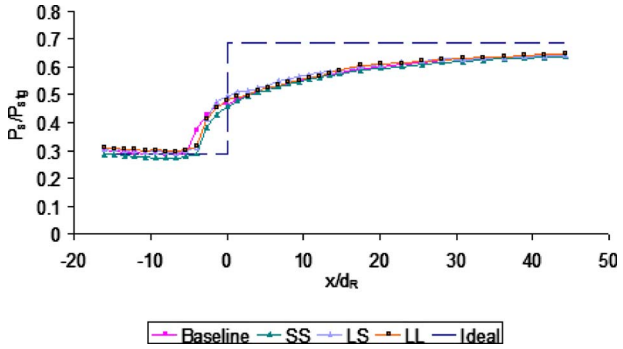


Fig. 3 Variation in static pressure ratio along the tunnel for different configurations

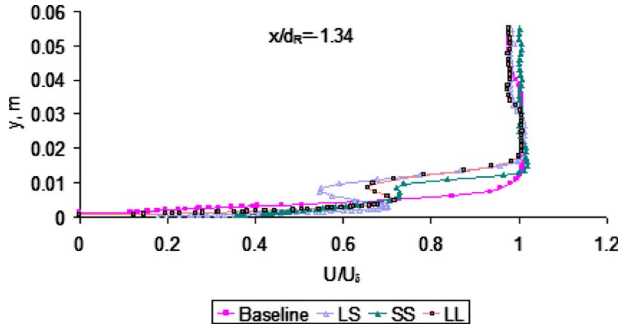


Fig. 6 Streamwise velocity profile for different configurations at $x/\delta_R=-1.34$

This indicates that in the case of LL, the boundary layer must have a larger displacement thickness and hence shallower velocity profile upstream of the shock wave. Further downstream at $x/\delta_R=-1.34$ (Fig. 6), LL clearly shows regions of separated flow, which is the result of an unhealthier boundary layer, a confirmation of the previous statement.

The boundary layer velocity profiles downstream of the shock wave at $x/\delta_R=2.68$ (Fig. 7) clearly show that, while in the baseline case, the flow is separated and the SS configuration starts to recover faster close to the surface. This trend continues further downstream, as shown in Figs. 8–11, corresponding to $x/\delta_R=6.7$, 10.72, 14.75, and 36.19 downstream of the shock location. This also agrees with the results obtained for Preston pressure distributions (Fig. 3). The recovery of the velocity profile in the last station (i.e., $x/\delta_R=36.19$, Fig. 11) shows that, compared with other configurations (LS and LL), SS produces a fuller profile.

Figure 12 shows the streamwise distributions of the boundary layer physical thickness for all configurations. It is clear from this figure that the LL configuration induces thicker and hence less stable boundary layer especially closer to the shock location, which confirms the variations in velocity profiles shown in Figs. 5–11.

The streamwise distribution of the boundary layer displacement thickness is shown in Fig. 13. This figure as suggested before indicates that generally LL configuration produces larger flow displacement as compared with SS, which is another indication of less stable flow within the boundary layer. The variation in the boundary layer shape factor for all cases is shown in Fig. 14. It again shows that SS configuration produces lowest shape factor throughout the interaction as compared with LL, which is another indication of a fuller velocity profile. All configurations show much lower H distribution compared with the baseline.

4 Flow Field and Surface Oil Flow Visualization

In order to gain an appreciation of the characteristics of the flow near the surface, a well-known method for surface flow vi-

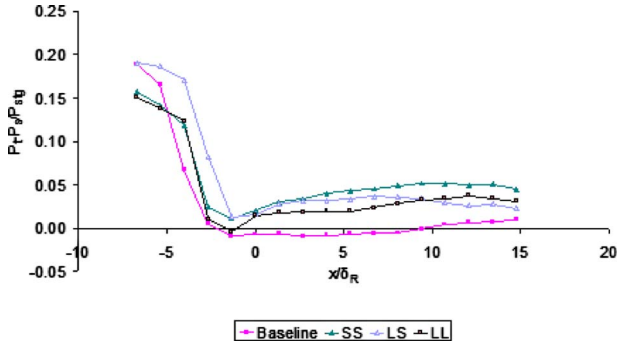


Fig. 4 Variation in dynamic pressure ratio along the tunnel for different configurations

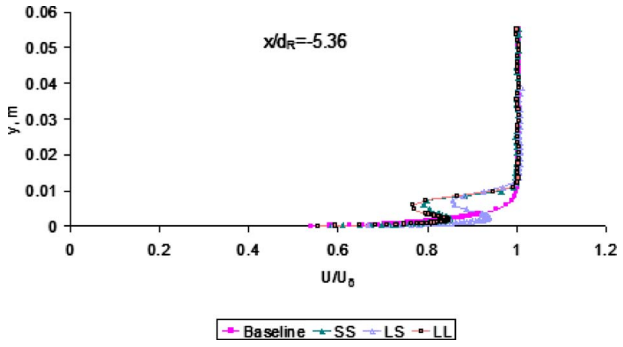


Fig. 5 Streamwise velocity profile for different configurations at $x/\delta_R=-5.36$

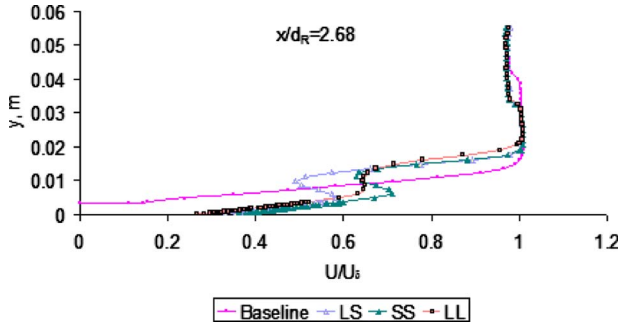


Fig. 7 Streamwise velocity profile for different configurations at $x/\delta_R=2.68$

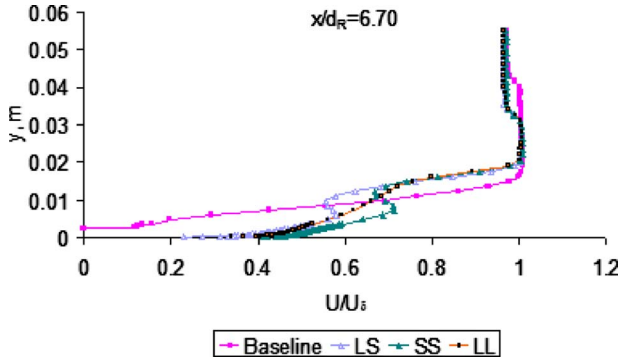


Fig. 8 Streamwise velocity profile for different configurations at $x/\delta_R=6.70$

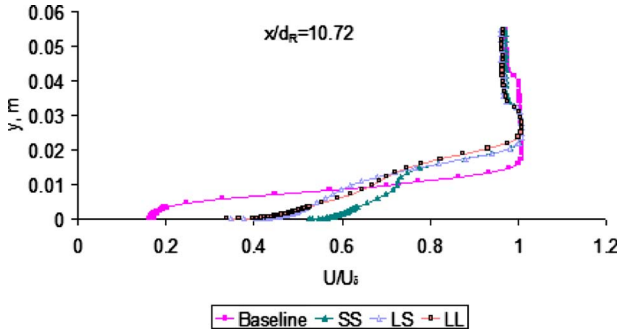


Fig. 9 Streamwise velocity profile for different configurations at x/δ_R=10.72

sualization was employed. A mixture of Shell Tellus oil, titanium dioxide, and oleic acid (an anticoagulant) was spread thinly and evenly on the surface of the liner in advance of the dedicated test run. A high intensity projector lamp was shone through the window of the working section that was covered with a sheet of ordinary tracing paper, so that the light would be more evenly spread onto the tunnel liner. A digital camera was supported on a tripod on the opposing side of the tunnel that was focused on the location of interest. As the flow developed, the oil was displaced by the forces imparted by the near-wall flow. Once the flow development had ceased, a still photograph was taken of the developed flow pattern to be used for later comparisons between the different SBVG configurations. The exact measures of the ingredients of the mixture were as follows: 40 ml Tellus oil, ten flat teaspoons of titanium dioxide, and 20 drops Oleic acid. The ingredients were mixed vigorously until a smooth solution was obtained and was then left for 1 h before mixing once more prior to use.

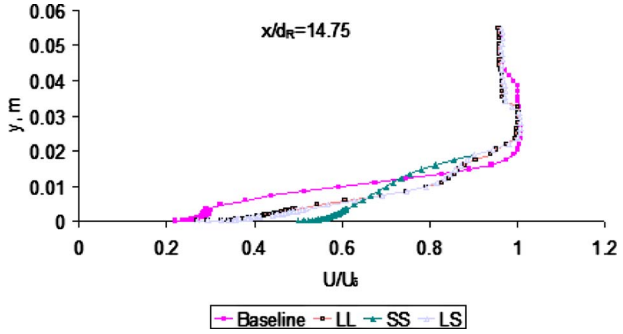


Fig. 10 Streamwise velocity profile for different configurations at x/δ_R=14.75

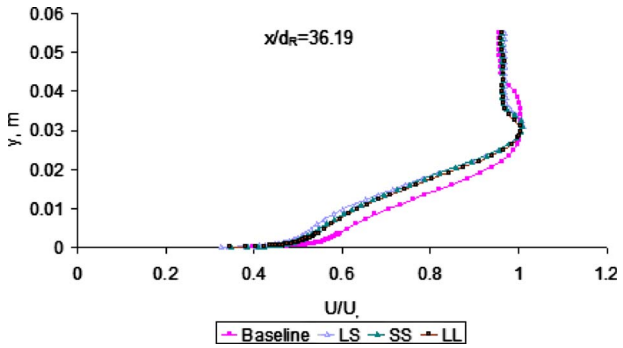


Fig. 11 Streamwise velocity profile for different configurations at x/δ_R=36.19

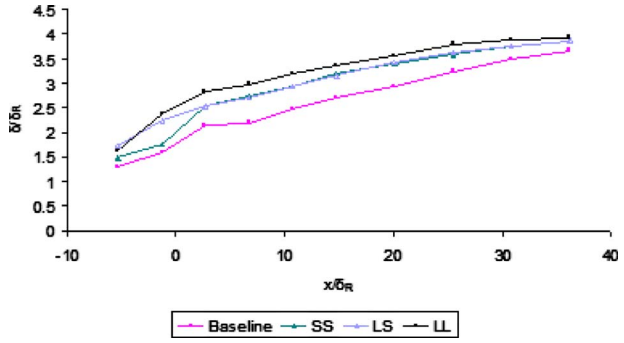


Fig. 12 Streamwise boundary layer thickness for different configurations

Figures 15(*a*)–15(*c*) show the surface oil flow visualization for the baseline, SS, and LS cases. The center of each vortical structure labeled "spiral node" formed at 25 mm from the sidewalls of the wind tunnel. The vortical structure of separation has been developed symmetrically where at the mid-span region, there is a region of quasi-two-dimensional flow with a width of about 20 mm. In the case of baseline, the flow was completely reversed in the centerline. The separation line is around 2δ upstream of the main shock wave, which is well within the location of front leg of Lambda structure, which is also about the point of rapid rise in static pressure. The reattachment point in the mid-span is located at about of separation is about 11.8δ downstream of the location of the main shock.

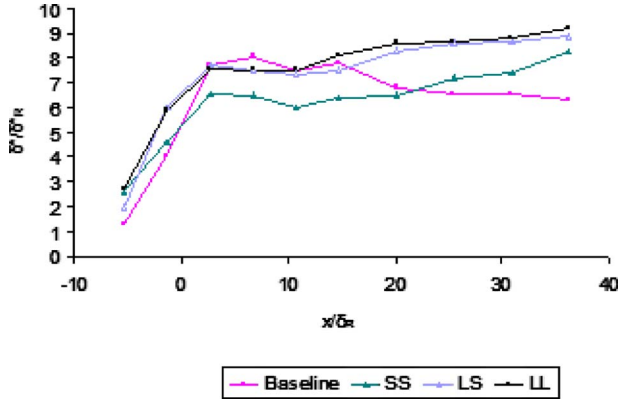


Fig. 13 Streamwise displacement thickness for different configurations

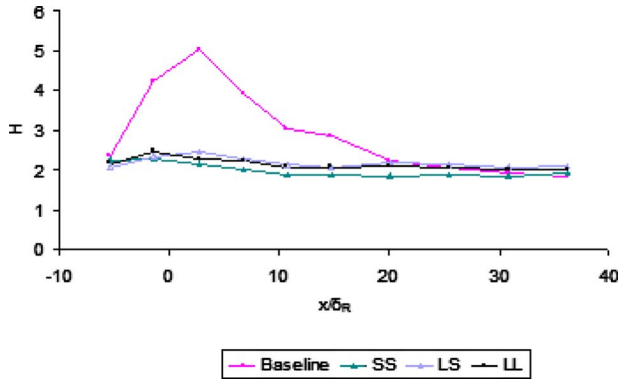
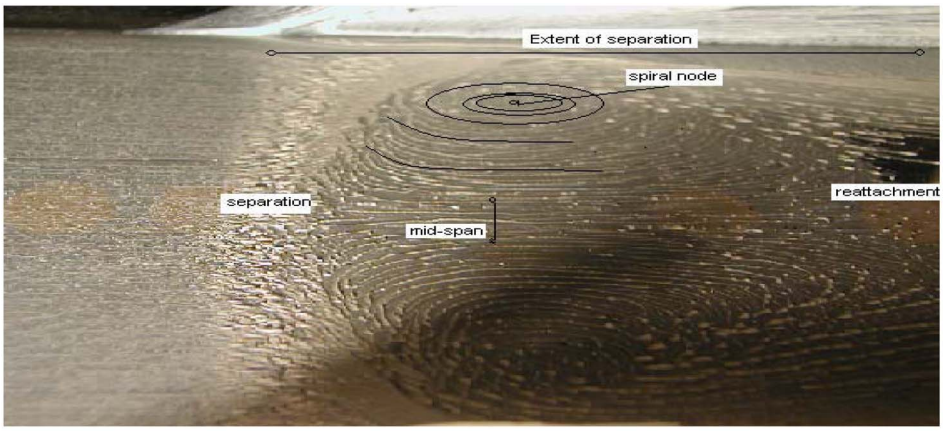
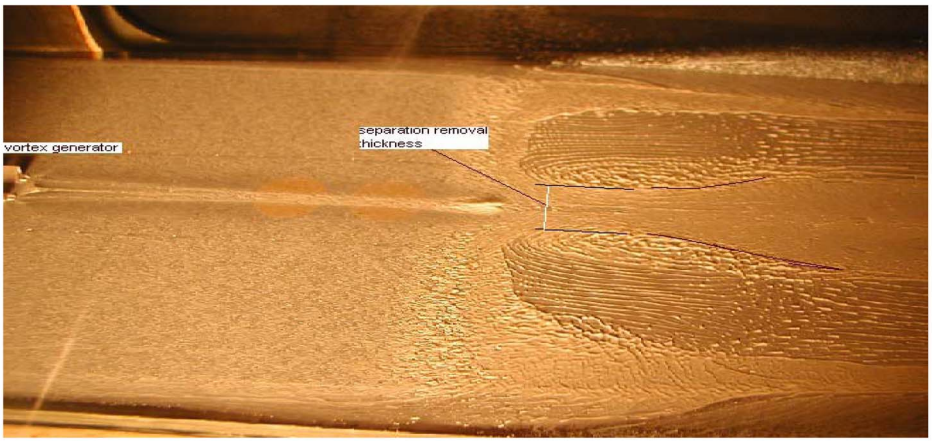


Fig. 14 Streamwise shape factor for different configurations



(*a*)



(*b*)

(*c*)

Fig. 15 Visualization of separation: (*a*) baseline, (*b*) SS, and (*c*) LS. Location of shock wave is above the separation line.

For SS, the flow around the midspan with a width of about 18 mm has been completely reattached. The separation removal thickness of the flow at the area of vanished separation is in the range 15–18 mm, where the leading gap of the pair of vortex generator is also 18 mm. Similar observation can be seen in the case of LS configuration.

For flow field flow visualization, the Z-Type Schlieren system was used to visualize flow field around the interaction region. Figure 16 shows a noticeable reduction in the extent of Lambda shock system for both SS and LS devices. The figure demonstrates the bifurcation height, which is almost equal for SS and LS and shorter than the baseline case: bifurcation point in baseline is

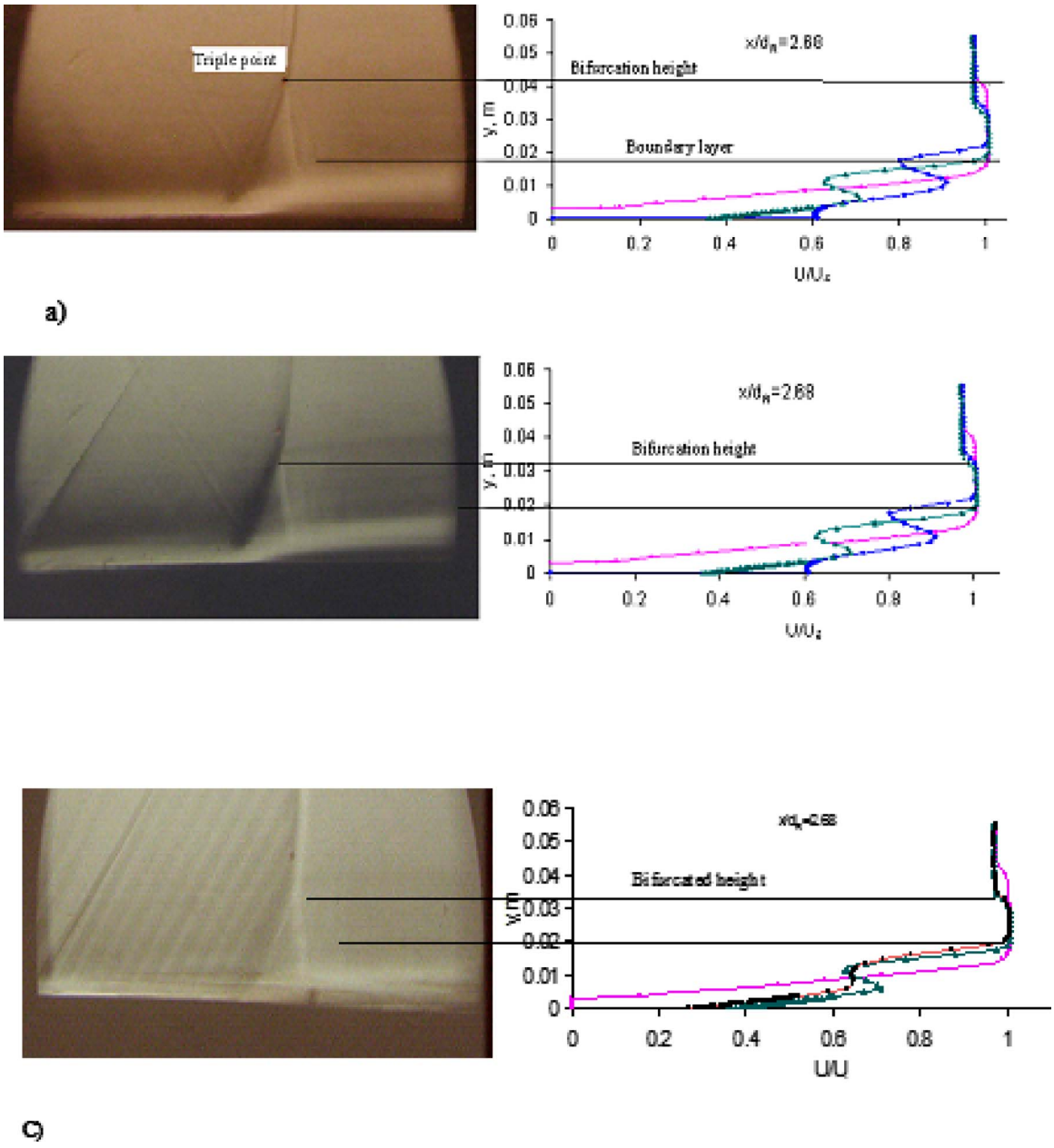


Fig. 16 Schlieren visualization of the affected area: (*a*) baseline, (*b*) SS, and (*c*) LS

0.042 m and reduced to 0.033 m and 0.033 m for SS and LS. Boundary layer thickness in baseline is 0.017 m and in SS and LS are 0.019 m and 0.02 m, correspondingly.

Both methods of flow visualization show agreement with the flow measurements in terms of the effects of SBVGs on the extent of the boundary layer separation.

5 Conclusion

The performance of tetrahedral vortex generators with different lengths and different locations were examined by studying the development of boundary layer profile and the streamwise variation in the Preston pressure. The results can be listed as follows.

- Sub-boundary vortex generator are capable to produce vortices without considerable parasitic drag.
- Tetrahedral vortex generator here called SS model could totally vanish the shock wave boundary layer separation.
- The long model (LS) located at $17.4\delta_R$ is capable to eliminate the separation. This model is mounted in a larger distance ($25.5\delta_R$) called LL, to verify if a larger distance can be effective enough when the device is longer. The result indicated that an increase in the distance for the long device from $17.4\delta_R$ to $25.5\delta_R$ did not remove the separation entirely. A comparison between SS and LS in Preston pressure shows that SS in the area of separation and further points has a higher total pressure, where its displacement thickness and shape factor are lower. Therefore, SS has a better performance.

Acknowledgment

The work presented here were taken from a research project carried out at Queen Mary University of London. I highly appreciate the laboratory technicians and workshop staff for their help to facilitate this research project.

Nomenclature

H = shape factor
h = vortex generator height
M = Mach number
P = pressure
P_s = wall static pressure
P_{stg} = settling chamber stagnation pressure
P_t = total pressure
Re = Reynolds number
U_δ = velocity at edge of boundary layer
X = streamwise coordinate direction
Y = vertical coordinate (normal to the wall)
δ = boundary layer thickness
δ_R = boundary layer thickness at reference point
δ^* = displacement thickness
θ = momentum thickness

References

[1] Wheeler, G. O., 1984, "Means for Maintaining Attached Flow of a Flowing Medium," U.S. Patent No. 4,455,045.
[2] Bur, R., and Corbel, B., 2000, "Experimental Study of Transonic Interaction With Shock and Boundary-Layer Control," Office National d'Etudes et de Recherches Aerospatiales, AIAA, Fluid 2000, Denver, CO.
[3] Donaldson C., 1950, "Investigation of a Simple Device for Preventing Separation Due to Shock and Boundary-Layer Interaction," NACA Report No. RM L5OB02a.
[4] Smith, A. N., Babinsky, H., Fulker, J. L., and Ashill, P. A., 2001, "Control of Normal Shock Wave/Turbulent Boundary-Layer Interaction Using Streamwise Slots," 39th Aerospace Science Meeting and Exhibit, Reno, NV, AIAA Paper No. 2001-0739.
[5] Westkaemper, J. C., and Whitten, J. W., 1970, "Drag of Vane-Type Vortex Generators in Compressible Flow," J. Spacecr. Rockets, **7**, pp. 1269–1271.
[6] Lin J.C., "Application of Micro-Vortex Generator for Turbulent Flow Separation Control," NASA Langley Research Centre, Hampton, Virginia USA, Report No. 23681-2199.
[7] Lin, J. C., Robinson, S. K., and McGhee, R. J., 1994, "Separation Control on High-Lift Airfoils Via Micro-Vortex Generators," J. Aircr., **31** (6), pp. 1317–1323.
[8] Lin, J. C., Howard, F. G., and Selby, G. V., 1990, "Small Submerged Vortex Generators for Turbulent Flow Separation Control," Journal of Spacecraft, **27** (5), pp. 503–507.
[9] Wheeler, G., 1991, "Low Drag Vortex Generators," U.S. Patent No. 5,058,837.
[10] Ashill, P. R., Fulker, J. L., and Hackett, K. C., 1999, "Research at DERA on Sub Boundary Layer/Vortex Generators," Aerodynamic Department, Defence Evaluation and Research Agency (DERA), Bedford, Report No. MK41 6AE UK.
[11] Holden, H. A., and Babinsky, H., 2004, "Vortex Generators Near Shock/Boundary Layer Interactions," AIAA Paper No. 2004–1242.
[12] Kuethe, A. M., 1971, "Boundary Layer Control of Flow Separation and Heat Exchanger," U.S. Patent No. 3,516,264.
[13] Kuethe, A. M., 1973, "Boundary Layer Control of Flow Separation and Heat Exchanger," U.S. Patent No. 3,741,285.
[14] Rao, D. M., and Kariya, T. T., 1988, "Boundary Layer Submerged Vortex Generators for Separation Control—An Exploratory Study," AIAA Paper No. 88-3546-CP.
[15] Ashil, P. R., and Riddle, G. L., 1995, "Sub Boundary Layer Vortex Generators," International Patent Classification No. B64C 21/10, 23/06.
[16] Motallebi, F., 2003, "Flow Control," *Lecture Note*, QMW U. of London, Engineering Department of Queen Mary University of London.
[17] Manor D., Dima C., Schoch P., and Polo J., 1993, "Using Pop-Up Vortex Generators on the Wing Surface to Greatly Increase the Lift and Stall Angle of Attack,"AIAA Paper No. 93-1016.
[18] Dolling, D. S., Fourier, E., and Shau, Y. R., 1992, "Effect of Vortex Generators on the Growth of a Compressible Shear Layer," J. Propul. Power, **8**(5), pp. 1049–1056.
[19] Lin J. C., 2002, "Review of Research on Low-Profile Vortex Generators to Control Boundary Layer Separation," Prog. Aerosp. Sci., **38**, pp. 389–420.
[20] Cohen, S. G., and Motallebi, F., 2006, "Sub-Boundary Layer Vortex Generators for the Control of Shock-Induced Separation," Aeronaut. J., April, pp. 215–226.
[21] Anderson, J. D., 1991, *Fundamentals of Aerodynamics*, McGraw-Hill, New York.
[22] Ower, E., and Pankhurst, R. C., 1966, *The Measurement of Airflow*, 4th ed., Pergamon, New York.
[23] Bryer, D. W., Pankhurst, R. C., 1971, "Pressure-Probe Methods for Determining Wind Speed and Flow Direction," NPL, HMSO.
[24] Motallebi, F., 1994, "Mean Flow Study of Two-Dimensional Subsonic Turbulent Boundary Layers," AIAA J., **32**(11), pp. 2153–2161.
[25] Allen, J. M., 1972, "Pitot Probe Displacement in a Supersonic Turbulent Boundary Layer," NASA Report No. TN-D 6759.

Computation of Thermal Stress Intensity Factors for Bimaterial Interface Cracks Using Domain Integral Method

Ratnesh Khandelwal

J. M. Chandra Kishen[1]
e-mail: chandrak@civil.iisc.ernet.in

Department of Civil Engineering,
Indian Institute of Science,
Bangalore 560 012, India

The concept of domain integral used extensively for J integral has been applied in this work for the formulation of J_2 integral for linear elastic bimaterial body containing a crack at the interface and subjected to thermal loading. It is shown that, in the presence of thermal stresses, the J_k domain integral over a closed path, which does not enclose singularities, is a function of temperature and body force. A method is proposed to compute the stress intensity factors for bimaterial interface crack subjected to thermal loading by combining this domain integral with the J_k integral. The proposed method is validated by solving standard problems with known solutions. [DOI: 10.1115/1.3086588]

Keywords: bimaterial interface, J_2 integral, stress intensity factor, thermal load

1 Introduction

Interfacial delamination and fracture are commonly observed problems in many of the advanced composite materials that are used heavily in aerospace, nuclear industry, and microelectronics field. The important parameters for evaluating the delamination fracture strength are Mode 1 and Mode 2 stress intensity factors (SIFs) or the energy release rate of a crack between the two dissimilar materials. For practical applications, the SIFs of bimaterial interface cracks are generally obtained by the finite element method (FEM). Energy approaches such as the crack closure integral method [1], the *J*-integral method [2], and the virtual crack extension method [3] are reliable methods for calculating the energy release rate using FEM.

Fracture at a bimaterial interface is essentially mixed mode, even when the geometry is symmetric with respect to a crack, and loading is pure Mode 1. This is due to the differences in the elastic properties across an interface, which would disrupt the symmetry [4]. Consequently, both tensile and shear stresses act on the interface ahead of the crack, and opening and sliding displacements of the crack flanks occur behind the crack tip. The linear elastic solutions of the crack-tip stress and displacement fields show that the stresses and displacements ahead of the crack front behave in an oscillatory manner. Due to this oscillatory behavior, the definition of the stress intensity factors needs special consideration, and in addition crack face contact may occur for a small distance near the crack tip. Mode 1 and Mode 2 stress intensity factors cannot be decoupled to represent tension and shear stress fields, as seen in the case of homogeneous materials. The Appendix gives the linear elastic stress and displacement fields for a crack lying between the interfaces of two different materials. In order to obtain the SIFs, combinations of different methods have been proposed by different researchers. A combination of the *J*-integral and the superposition method with FEM known as the *M*-integral method was applied by Yau and Wang [5]. Matos et al. [6] applied the virtual crack extension method to the calculation of mixed-mode SIFs for interface cracks in conjunction with the superposition method and FEM. Banks-Sills et al. [7] used the *M*-integral in conjunction with FEM and the concept of traction weight function to determine the stress intensity factors arising from residual thermal stresses in a bimaterial finite body. The *M*-integral method with virtual crack extension method and the boundary element method (BEM) have been used by Miyazaki et al. [8,9] for the determination of fracture parameters. Ryoji and Sang-Bong [10] applied the displacement extrapolation method along with the BEM.

The problem of determining the SIFs of bimaterial interface cracks when subjected to thermal loads becomes more complicated. A great amount of residual stresses are developed near an interface between dissimilar materials because of the differences in the coefficient of linear thermal expansion between the two jointed materials. Hence, thermal stresses play an important role in an interface crack than for a crack in a homogeneous material [11]. However, none of the techniques mentioned above can analyze the SIF of an interface crack in the presence of thermal loads. Studies on this coupled thermomechanical stress fracture problem due to the local intensification of temperature gradient and stress field in recent years have led to significant progress in analytical research. Among the available methods [10–15] for calculating fracture parameters in the literature, the conservative *M*-integral method used in conjunction with the domain integral method given recently by Banks-Sills and Dolev [12] are well suited for bimaterial thermal interface problems. However, this method requires an auxiliary stress and displacement field for a given crack configuration, thus making the computations tedious.

The domain integral method has emerged as being well suited for bimaterial interface crack problems. In the domain integral method, a crack-tip contour integral is expressed as a volume integral over a finite domain surrounding the crack tip. The process of recasting the contour line integral into an area integral for 2D problems and surface integral into a volume integral for 3D problems is advantageous for numerical purposes because accurate fracture parameters can be obtained without having to precisely capture the details of the singular fields in the vicinity of the crack tip.

The subject of path-independent integrals (J_k) of fracture mechanics has received considerable attention from researchers [15–18] with regard to numerical determination of the mixed-mode stress intensity factors. Khandelwal and Chandra Kishen [19] proposed an analytical expression for J_k integral for bimaterial interface problem subjected to mechanical loading and

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 4, 2008; final manuscript received August 10, 2008; published online April 23, 2009. Review conducted by Martin Ostoja-Starzewski.

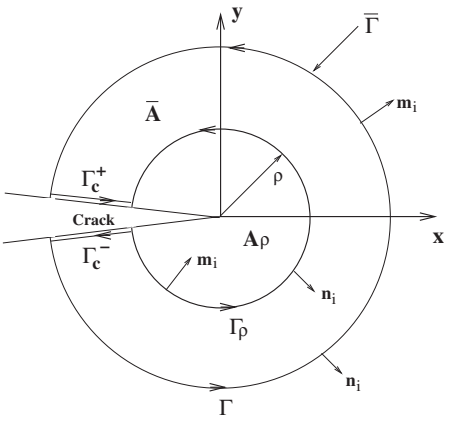


Fig. 1 Closed integration path $\bar{\Gamma}$ in a homogenous body ($\bar{\Gamma} = \Gamma_c^- + \Gamma + \Gamma_c^+ - \Gamma_\rho$)

showed that this integral is path independent in a modified sense and is useful in the determination of stress intensity factors of bimaterial interface cracks. J_k integrals are line integrals evaluated along a counterclockwise contour enclosing and shrinking onto the crack tip. Knowles and Sternberg [20] defined J_k as a complex quantity given by

$$J = J_1 - iJ_2 \tag{1}$$

The quantity J corresponds to the energy release rate for the movement of crack edge in any direction. For homogeneous media with traction-free crack surface, both components of J, J_1 and J_2, are path independent. Although J_1 can start anywhere from the lower crack surface and end anywhere along the upper crack surface, the same does not hold true for J_2 wherein two additional line integrals along the crack surfaces are added up, which causes the inclusion of a singular region in the integration. Therefore, direct calculation of J_2 from numerical solution becomes a tedious job. To overcome the difficulty in calculating J_2, an approximate analytical expression for J_2, in conjunction with the finite element method, has been proposed by Khandelwal and Chandra Kishen [19] for bimaterial interface subjected to mechanical loading.

In this work, the concept of J_2 domain integral is proposed for computing the stress intensity factors of bimaterial interface cracks subjected to thermal loads. It is shown that, in the presence of thermal stress, the J_k domain integral over a closed path, which does not enclose singularities, is a function of temperature and body force. A method is proposed to compute the stress intensity factors for bimaterial interface crack subjected to thermal loading by combining this domain integral with the J_k integral. The proposed method is validated by solving standard problems with known solutions.

2 Formulation

The path-independent integral J_k for a 2D elastic isothermal homogeneous material in the absence of body forces and thermal loading can be expressed as

$$J_k = \lim_{\Gamma_\rho \to 0} \oint_{\Gamma_\rho} (W^{\text{el}} \delta_{ki} - \sigma_{ij} u_{j,k}) n_i d\Gamma \tag{2}$$

where W^{el} is the elastic strain energy density and in the present isothermal case is equal to total strain energy density, W^T, as both elastic and total strains are equal in isothermal case. n_i is the unit outward normal to the closed path Γ_ρ taken in anticlockwise sense, as shown in Fig. 1, u_i is the displacement vector referred in the Cartesian coordinate system, σ_{ij} is the stress tensor, and $d\Gamma$ is the arc length measured along the contour Γ_ρ.

Here

$$W^{\text{el}} = \int_0^{\epsilon^{\text{el}}} \sigma_{ij} d\epsilon_{ij}^{\text{el}} \tag{3}$$

$$W^T = W^{\text{el}} = \tfrac{1}{2} \sigma_{ij} \epsilon_{ij} \tag{4}$$

where ϵ_{ij} is the total strain and in present case is equal to elastic strain $\epsilon_{ij}^{\text{el}}$. σ_{ij} may be represented as

$$\sigma_{ij} = \lambda \epsilon_{kk} \delta_{ij} + 2\mu \epsilon_{ij} \tag{5}$$

where δ_{ij} is the Kronecker delta function defined as follows:

$$\delta_{ij} = \begin{cases} 1 & \text{when } i = j \\ 0 & \text{when } i \neq j \end{cases} \tag{6}$$

μ is the shear modulus and λ is the Lame's constant, which is given by

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)} \quad \text{for plane strain}$$
$$\lambda = \frac{\nu E}{(1-\nu^2)} \text{for plane stress} \tag{7}$$

Here E is the modulus of elasticity and ν the Poisson ratio. In the presence of thermal loads, the stress field is modified and can be written in the Duhamel–Neumann constitutive equation as

$$\sigma_{ij} = \lambda \epsilon_{kk} \delta_{ij} + 2\mu \epsilon_{ij} - \beta \theta \delta_{ij}, \quad i,j,k = 1,2$$
$$\sigma_{33} = \lambda \epsilon_{kk} - \beta \theta \quad \text{for plane strain} \tag{8}$$
$$\epsilon_{33} = \frac{1+\nu}{1-\nu} \alpha \theta - \frac{\nu}{1-\nu} \epsilon_{kk} \quad \text{for plane stress}$$

Here $\epsilon_{kk} = \epsilon_{11} + \epsilon_{22}$. If α be the coefficient of thermal expansion, then β is defined as

$$\beta = \frac{E\alpha}{(1-2\nu)'} \quad \text{for plane strain}$$
$$\beta = \frac{E\alpha}{(1-\nu)} \quad \text{for plane stress} \tag{9}$$

It may be mentioned here that, in the presence of body forces and thermal loads, the total strain energy density W^T is still related to the stress and strain tensors, as given in Eq. (4). However, in this case, the elastic strain energy density W^{el} of the system is different from total strain energy density W^T. W^{el} and W^T are related by

$$W^{\text{el}} = \tfrac{1}{2} \sigma_{ij} (\epsilon_{ij} - \alpha \theta \delta_{ij}) = W^T - \tfrac{1}{2} \alpha \theta \sigma_{ii} \tag{10}$$

Let us define the term W^F, called thermo-elastic strain energy density, where

$$W^F = W^T - \tfrac{\beta}{2} \theta \epsilon_{ii} \tag{11}$$

In the presence of body force, mechanical, and thermal loads, using the chain rule of differentiation, the derivative of different form of the strain energy density with respect to coordinate axis x_k can be written as

$$\frac{\partial W^{\text{el}}}{\partial x_k} = \frac{\partial W^T}{\partial x_k} - \frac{\alpha}{2} [\theta \sigma_{ii,k} + \theta_{,k} \sigma_{ii}]$$
$$\frac{\partial W^T}{\partial x_k} = \frac{\partial}{\partial x_j} \left(\sigma_{ij} \frac{\partial u_i}{\partial x_k} \right) + f_i \frac{\partial u_i}{\partial x_k} + \beta \left[\frac{1}{2} (\theta \epsilon_{ii})_{,k} - \theta_{,k} \epsilon_{ii} \right] \tag{12}$$
$$\frac{\partial W^F}{\partial x_k} = \frac{\partial}{\partial x_j} \left(\sigma_{ij} \frac{\partial u_i}{\partial x_k} \right) + f_i \frac{\partial u_i}{\partial x_k} - \beta \theta_{,k} \epsilon_{ii}$$

2.1 Domain Integral. It has been shown in Ref. [21] that conservative area integrals produce more stable results than the

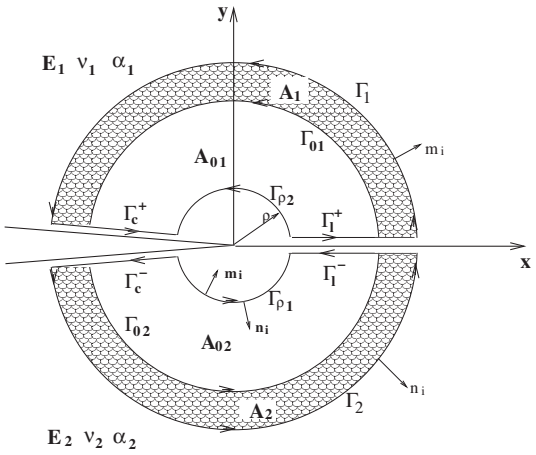


Fig. 2 Simple connected regions, A^1 and A^2, enclosed by integration paths $\Gamma^1=\Gamma_l^+ + \Gamma_1 + \Gamma_c^+ - \Gamma_{\rho 1}$ and $\Gamma^2=\Gamma_c^- + \Gamma_2 + \Gamma_l^- - \Gamma_{\rho 2}$, respectively, in a bimaterial body

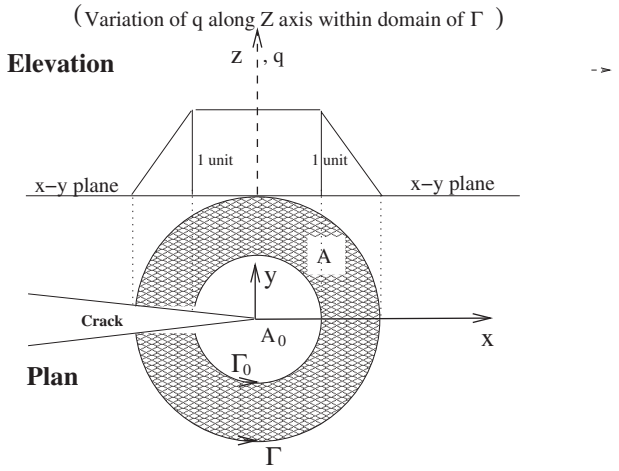


Fig. 3 Plot showing the distribution of function q within the domain of Γ

line integrals. Hence, in this work, the J_k line integral as defined above is first applied to the bimaterial crack configuration system and then converted into its domain integral form.

Redefining the J_k line integral of Eq. (2) in terms of an outer unit vector m_i, normal to the area $\bar{A}$, enclosed by the loop $\bar{\Gamma}$, and excluding the area enclosed within Γ_ρ, where $m_i=-n_i$ on Γ_ρ, as shown in Fig. 1, we obtain

$$J_k = \lim_{\Gamma_\rho \to 0} \oint_{\Gamma_\rho} (\sigma_{ij} u_{j,k} - W^{\mathrm{el}} \delta_{ki}) m_i d\Gamma \tag{13}$$

Shown in Fig. 2 are two elastic half planes of different materials bounded together with material: Occupying $y>0$ has Young's modulus E_1, Poisson's ratio ν_1, and thermal expansion coefficient α_1, whereas that occupying $y<0$ has the respective values as E_2, ν_2, and α_2. Also, the plane defined by $x<0$ and $y=0$ is unbounded and represents a crack. Γ^1 and Γ^2 are closed contours taken arbitrarily in the anticlockwise direction with the contact points on crack surface and interface at the same distance from the crack tip. The inner paths, $\Gamma_{\rho 1}$ and $\Gamma_{\rho 2}$ are semicircular anticlockwise contours having the same radius ρ about the crack tip. Now, the summation of the above two closed contour defined as $\bar{\Gamma}$ is written as $\bar{\Gamma}=\Gamma^1+\Gamma^2$, and $\Gamma^1=\Gamma_l^+ + \Gamma_1 + \Gamma_c^+ - \Gamma_{\rho 1}$ and $\Gamma^2=\Gamma_c^- + \Gamma_2 + \Gamma_l^- - \Gamma_{\rho 2}$. Also, A^1 and A^2 are the simply connected domain area enclosed by closed contours Γ^1 and Γ^2, respectively, and the whole area can be written as $\bar{A}=A^1+A^2$.

Furthermore in Fig. 2, the areas A^1 and A^2 are shown to be divided into shaded and unshaded parts. The shaded part of Fig. 2 represents area A that is further divided into two parts, A_1 and A_2, in the upper and lower materials, respectively. The inner core area (unshaded) A_0 is composed of A_{01} and A_{02} as shown. Hence,

$$\bar{A} = A^1 + A^2 = A_1 + A_{01} + A_2 + A_{02} = A_0 + A \tag{14}$$

For the sake of clarity and conciseness, let us take

$$L = (\sigma_{ij} u_{j,k} - W^F \delta_{ki}) \tag{15}$$

where W^F is as defined in Eq. (11). Multiply L with some arbitrary function q, which is smooth and continuously differentiable within area $\bar{A}$ whose value equals to unity on Γ_ρ and zero on outer contours, Γ_1 and Γ_2, respectively [22,21]. The distribution of function q within the domain of $\bar{A}$ is as shown in Fig. 3, which shows that q is further assumed to be unity within the unshaded domain A_0. The function q defined within an eight-noded isoparametric element is [21]

$$q = \sum_{m=1}^{8} N_m(\zeta, \eta) q_m \tag{16}$$

where N_m are the finite element shape functions of an eight-noded isoparametric element, ζ and η are the coordinates in the parent element, and q_m is the value of q at the nodal points.

The calculation of the J-integral is carried out within a ring of elements at a distance from the crack tip. This is seen to produce better results as compared with performing calculations within the crack-tip elements. The elements within the ring move as a rigid body. For each of these elements, q is unity, so that the derivative of q with respect to x_j is zero. For all elements outside the ring, q is zero, so that again the derivative of q is zero. For elements belonging to the ring, the vector q_m in Eq. (16) is chosen so that the virtual crack extension does not disturb the relative nodal point positions in their new locations; for example, a regular element with nodes at the midsides contains only midside nodes after distortion [21].

Taking separately the integral of function L given in Eq. (15) along the upper and lower closed loops Γ^1 and Γ^2, respectively, we obtain for $n=1,2$,

$$I_n = \oint_{\Gamma^n} (\sigma_{ij} u_{j,k} - W^F \delta_{ki}) q m_i d\Gamma \tag{17}$$

Equation (15) can be expanded as

$$\oint_{\Gamma^1} (L) q d\Gamma = \int_{\Gamma_c^+} (L) q d\Gamma + \oint_{\Gamma_1} (L) q d\Gamma + \int_{\Gamma_l^+} (L) q d\Gamma - \oint_{\Gamma_{\rho 1}} (L) q d\Gamma \tag{18}$$

$$\oint_{\Gamma^2} (L) q d\Gamma = \int_{\Gamma_c^-} (L) q d\Gamma + \oint_{\Gamma_2} (L) q d\Gamma + \int_{\Gamma_l^-} (L) q d\Gamma - \oint_{\Gamma_{\rho 2}} (L) q d\Gamma \tag{19}$$

Since the function q vanishes along the outer contour, the integrands along Γ_1 and Γ_2 will be zero. Furthermore, the integrals along $\Gamma_{\rho 1}$ and $\Gamma_{\rho 2}$ can be combined and written as

$$\oint_{\Gamma^1+\Gamma^2}(L)qd\Gamma=\int_{\Gamma_c^+ +\Gamma_c^-}(L)qd\Gamma+\int_{\Gamma_l^+ +\Gamma_l^-}(L)qd\Gamma+\oint_{\Gamma_\rho}(L)d\Gamma \tag{20}$$

Since $m_1=0$ and $m_2=\pm 1$ on Γ_c^-, Γ_c^+, Γ_l^-, and Γ_l^+, the above equation for traction-free crack surface is simplified to

$$\oint_{\Gamma_\rho}(\sigma_{ij}u_{j,k}-W^F\delta_{ki})m_i d\Gamma=\oint_{\Gamma^1+\Gamma^2}Lqd\Gamma+\int_{\Gamma_c^-+\Gamma_c^+}W^F qm_2\delta_{k2}d\Gamma-\int_{\Gamma_l^-+\Gamma_l^+}(\sigma_{j2}u_{j,k}-W^F\delta_{k2})qm_2 d\Gamma \tag{21}$$

J_k-integral is defined, with its second part in terms of elastic strain energy density W^{el}, as shown in Eq. (13). Converting W^F of the left hand side of the Eq. (21) to W^{el} using Eq. (11) and taking the limit as $\rho\to 0$,

$$\lim_{\Gamma_\rho\to 0}\oint_{\Gamma_\rho}(\sigma_{ij}u_{j,k}-W^{\mathrm{el}}\delta_{ki})m_i d\Gamma-\lim_{\rho\to 0}\sum_{m=1}^{2}\oint_{\Gamma_{\rho_m}}\left[\frac{\alpha_m}{2}\theta\sigma_{ii}-\frac{\beta_m}{2}\theta\epsilon_{ii}\right]m_k d\Gamma=\oint_{\Gamma^1+\Gamma^2}Lqd\Gamma+\int_{\Gamma_c^-+\Gamma_c^+}W^F qm_2\delta_{k2}d\Gamma-\int_{\Gamma_l^-+\Gamma_l^+}(\sigma_{j2}u_{j,k}-W^F\delta_{k2})qm_2 d\Gamma \tag{22}$$

Now for two dimensional bodies, the second integral in the left hand side of above equation is reduced to 0. Using Eq. (13),

$$J_k=\oint_{\Gamma^1+\Gamma^2}(\sigma_{ij}u_{j,k}-W^F\delta_{ki})m_i qd\Gamma+\int_{\Gamma_c^-+\Gamma_c^+}W^F qm_2\delta_{k2}d\Gamma-\int_{\Gamma_l^-+\Gamma_l^+}(\sigma_{j2}u_{j,k}-W^F\delta_{k2})qm_2 d\Gamma \tag{23}$$

Now applying the divergence theorem to the first line integral of Eq. (23), we obtain

$$J_k=\int_{A^1+A^2}\frac{\partial}{\partial x_i}[(\sigma_{ij}u_{j,k}-W^F\delta_{ki})q]dA+\int_{\Gamma_c^-+\Gamma_c^+}W^F qm_2\delta_{k2}d\Gamma-\int_{\Gamma_l^-+\Gamma_l^+}(\sigma_{j2}u_{j,k}-W^F\delta_{k2})qm_2 d\Gamma \tag{24}$$

Here, areas A^1 and A^2 are as discussed earlier. Differentiating the first integrand of Eq. (24) and simplifying, we obtain using Eq. (14)

$$J_k=\int_{A_0+A}[(\sigma_{ij}u_{j,k}-W^F\delta_{ki})]\frac{\partial q}{\partial x_i}dA+\int_{A_0+A}\left[\frac{\partial}{\partial x_i}\left(\sigma_{ij}\frac{\partial u_j}{\partial x_k}\right)-\frac{\partial W^F}{\partial x_k}\right]qdA+\int_{\Gamma_c^-+\Gamma_c^+}W^F qm_2\delta_{k2}d\Gamma-\int_{\Gamma_l^-+\Gamma_l^+}(\sigma_{j2}u_{j,k}-W^F\delta_{k2})qm_2 d\Gamma \tag{25}$$

The conditions along straight crack surface and interface, which are parallel to x-axis requires the following conditions to be satisfied.

1. $m_1 ds=n_1 ds=dy=0$.
2. $m_2=n_2=+1$ for Γ_c^- and Γ_l^-.
3. $m_2=n_2=-1$ for Γ_c^+ and Γ_l^+.
4. For traction-free crack surfaces, $\sigma_{ij}m_i=0$.
5. Continuity conditions of displacements and tractions across the interface requires $[|u_j|]=0$ and $[|\sigma_{j2}|]=0$.

Furthermore, $q=1$ and $\partial q/\partial x_i=0$ in domain A_0. Using Eq. (12) and the above mentioned conditions, Eq. (25) reduces to

$$J_k=\int_A(\sigma_{ij}u_{j,k}-W^F\delta_{ki})\frac{\partial q}{\partial x_i}dA+\sum_{n=1}^{2}\beta_n\int_{A_n}\theta_{,k}\epsilon_{ii}qdA+\sum_{n=1}^{2}\beta_n\int_{A_{0n}}\theta_{,k}\epsilon_{ii}dA-\int_{A^1+A^2}f_i\frac{\partial u_i}{\partial x_k}qdA+\int_{\Gamma_c^-+\Gamma_c^+}W^F qm_2\delta_{k2}d\Gamma-\int_{\Gamma_l^-+\Gamma_l^+}(\sigma_{j2}u_{j,k}-W^F\delta_{k2})qm_2 d\Gamma \tag{26}$$

The above equation for $k=1,2$ in the absence of body forces would reduce to

$$J_1=\int_A(\sigma_{ij}u_{j,1}-W^F\delta_{1i})\frac{\partial q}{\partial x_i}dA+\sum_{n=1}^{2}\beta_n\int_{A_n}\theta_{,1}\epsilon_{ii}qdA+\sum_{n=1}^{2}\beta_n\int_{A_{0n}}\theta_{,1}\epsilon_{ii}dA \tag{27}$$

$$J_{2\rho}=\int_A(\sigma_{ij}u_{j,2}-W^F\delta_{2i})\frac{\partial q}{\partial x_i}dA+\sum_{n=1}^{2}\beta_n\int_{A_n}\theta_{,2}\epsilon_{ii}qdA+\sum_{n=1}^{2}\beta_n\int_{A_{0n}}\theta_{,2}\epsilon_{ii}dA+\int_{\Gamma_c^-+\Gamma_c^+}W^F qm_2 d\Gamma-\int_{\Gamma_l^-+\Gamma_l^+}(\sigma_{j2}u_{j,2}-W^F)qm_2 d\Gamma \tag{28}$$

where $A_0(=A_{01}+A_{02})$ indicates the area A_0, which excludes the inner core area of radius ρ.

If we allow the outer path to be arbitrary, but make the inner path circular with radius ρ about the crack tip, as shown in Fig. 1, then the integral in Eq. (27) for the limit when $\Gamma_\rho\to 0$ will be reduced to the well known J-integral (J_1 in our case). Furthermore, Eq. (28) does not converge to a certain fixed value; it is defined for a very small value of ρ lying within the singularity dominated zone and is defined as the J_2-integral. As explained in the previous work of the authors [19], for the limit as $\Gamma_\rho\to 0$, Eqs. (27) and (28) can be reduced to closed form expressions in terms of stress intensity factors as

$$J_1=\left[\frac{(1+\kappa_1)}{16\mu_1}+\frac{(1+\kappa_2)}{16\mu_2}\right](K_1^2+K_2^2) \tag{29}$$

$$J_{2\rho}=-\frac{1}{32\pi\epsilon}\left[\frac{(1+\kappa_1)}{\mu_1}(1-e^{-2\pi\epsilon})+\frac{(1+\kappa_2)}{\mu_2}(e^{2\pi\epsilon}-1)\right][(K_1^2-K_2^2)\sin 2\epsilon\log(\rho)+2K_1K_2\cos 2\epsilon\log(\rho)] \tag{30}$$

where $A_{1\rho}$ and $A_{2\rho}$ are the areas within the upper and lower closed loops Γ^1 and Γ^2, respectively. μ_j is the shear modulus of material j, $\kappa_j=(3-\nu_j)/(1+\nu_j)$ for plane stress, $\kappa_j=3-4\nu_j$ for plane strain, and ν_j is Poisson's ratio of material j. ϵ is the bimaterial constant as defined in the Appendix.

2.2 Computation of Bimaterial Stress Intensity Factors. The main objective of this work is to compute the stress intensity factors for bimaterial interface cracks when subjected to thermal loads, and this is done using the above definitions of J_1 and $J_{2\rho}$.

Equations (29) and (30) can be written in the form

$$(K_1^2+K_2^2)=B_1J_1 \tag{31}$$

$$(K_1^2 - K_2^2)C_1 + 2K_1K_2C_2 = J_{2\rho} \tag{32}$$

where

$$\frac{1}{B_1} = \left[\frac{(1+\kappa_1)}{16\mu_1} + \frac{(1+\kappa_2)}{16\mu_2}\right]$$

$$C_1 = -\frac{1}{32\pi\epsilon}\left[\frac{(1+\kappa_1)}{\mu_1}(1-e^{-2\pi\epsilon}) + \frac{(1+\kappa_2)}{\mu_2}(e^{2\pi\epsilon}-1)\right]\sin(2\epsilon\log\rho)$$

$$C_2 = -\frac{1}{32\pi\epsilon}\left[\frac{(1+\kappa_1)}{\mu_1}(1-e^{-2\pi\epsilon}) + \frac{(1+\kappa_2)}{\mu_2}(e^{2\pi\epsilon}-1)\right]\cos(2\epsilon\log\rho)$$

Elimination of K_2 from Eqs. (31) and (32) results in

$$K_1^4(4C_1^2+4C_2^2) - 4K_1^2[J_1B_1(C_1^2+C_2^2) + J_{2\rho}C_1] + (J_1B_1C_1 + J_{2\rho})^2 = 0 \tag{33}$$

Hence, K_1 can be obtained by solving the above equation. Substituting for K_1 in Eq. (31), we can compute K_2.

The sign of K_1 and K_2 are determined by monitoring the magnitude of the crack opening displacement near the crack tip. The crack opening displacement may be defined as

$$\Delta u = u^+ - u^- \tag{34}$$

$$\Delta v = v^+ - v^- \tag{35}$$

where the + and − sign refers to the upper and lower crack faces, respectively. The signs of K_1 and K_2 correspond to the signs of Δu and Δv.

3 Validation

In order to validate the domain integral formulation, three benchmark problems of plane strain, one under uniform temperature and the other two under thermal flow, are considered. The domain integral calculation of both J_1 and $J_{2\rho}$ have been computed using a newly developed finite element and fracture analysis code using MATLAB with eight-noded isoparametric elements for both mechanical and thermal analysis. The preprocessing input for this code is obtained from ANSYS [23].

To determine the optimized mesh size and the sensitivity of mesh refinement on the J_k domain integral, convergence studies has been performed. Based on this study, an optimized mesh size is used for computation of the J_k domain integrals. Mesh refinement is done in the region surrounding the crack tip for the body subjected to thermal loads. In the region away from the crack tip, a uniform mesh is used.

3.1 Dissimilar Semi-Infinite Plate With Double Edge Crack Subjected to Uniform Temperature Change. A jointed dissimilar semi-infinite plate with double edge crack subjected to uniform temperature change of 100°C [11] is analyzed under plane strain conditions. The geometry of the plate is shown in Fig. 4. Due to the symmetry in geometry and loading about the y-axis, only the left half is analyzed. The dimensions of the plate are taken as 200 units × 400 units with the jointed interface of 1 unit length. The material properties used in the analysis are shown in Table 1 and are assumed to be independent of temperature variation. Figures 5 and 6 show the displacement boundary condition and finite element mesh along with the contours used, respectively. The same finite element mesh is used for thermal analysis too.

The results of the convergence study on this problem are shown in Fig. 7 for a value of $\rho = 8.845 \times 10^{-3}$. It is seen that with increasing mesh density, the computed value of $J_{2\rho}$ converges to the analytically obtained value. It is observed that for a finite element mesh with 15,000–17,000 number of elements, there is very little variation in calculated $J_{2\rho}$. The variation in computed J_1 is found to be within 1% of the analytical value for all the mesh sizes considered in this study. As a trade-off between solution accuracy and computation time, a finite element mesh of 15,869 eight-noded elements with 48,320 nodes have been used for subsequent analysis. The solution in terms of the stress intensity factors for this problem was given by Erdogan [24] as

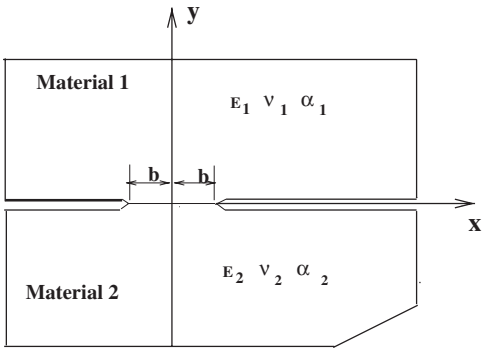


Fig. 4 Semi-infinite bimaterial plate subjected to constant uniform temperature load

Table 1 Material properties used in the analysis

Properties	Material 1	Material 2
Young modulus (Pa)	1000×10^9	100×10^9
Poisson ratio	0.3	0.3
Temperature (°C)	100	100
Coefficient of thermal expansion (/°C)	1.0×10^{-6}	1.0×10^{-7}
Coefficient of heat conduction (W/m °C)	100	100

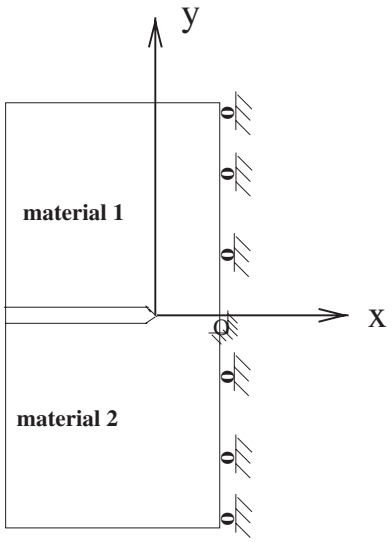


Fig. 5 Displacement boundary condition for jointed dissimilar semi-infinite plate with double edge crack under uniform temperature load

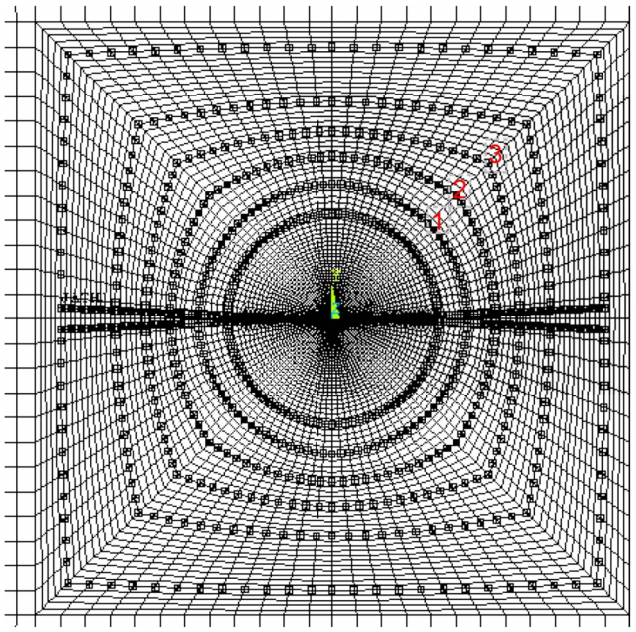


Fig. 6 A typical FE mesh and shape of the contour path used for all the problems

$$\hat{K}_1 = -2\epsilon\sigma_0(\alpha_2\eta_2 - \alpha_1\eta_1)\Delta T\sqrt{\pi b}(2b)^{-i\epsilon}$$
$$\hat{K}_2 = -\sigma_0(\alpha_2\eta_2 - \alpha_1\eta_1)\Delta T\sqrt{\pi b}(2b)^{-i\epsilon} \tag{36}$$

where $\eta_i = 1$ for plane stress and $\eta_i = 1 + \nu_i$ for plane strain case; σ_0 is defined by

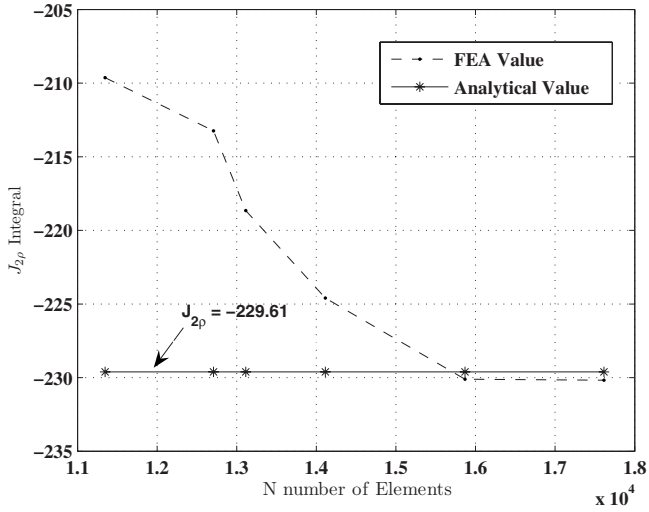


Fig. 7 Influence of mesh refinement on $J_{2\rho}$ for jointed dissimilar semi-infinite plate with double edge crack ($\rho = 8.845 \times 10^{-3}$)

$$\sigma_0 = \frac{4\mu_1\mu_2\cosh(\varepsilon\pi)}{(\mu_1 + \mu_2\kappa_1 + \mu_1\kappa_2 + \mu_2)} \tag{37}$$

where α_1 and α_2 are the coefficients of linear thermal expansion for materials 1 and 2, respectively. ΔT is the temperature excursion.

Let $K_0 = \sigma\sqrt{\pi b}$, where $\sigma = \sigma_0(\alpha_2\eta_2 - \alpha_1\eta_1)\Delta T$. The normalized complex stress intensity factor is defined as

$$\tilde{K} = \frac{\hat{K}L^{i\epsilon}}{\sigma\sqrt{\pi a}} \tag{38}$$

where L and a are the length parameters and σ is stress. For the present case, normalized stress intensity factors with $L = 2a$ and $a = b$ can be written as

$$\tilde{K}_1 = -2\varepsilon$$
$$\tilde{K}_2 = 1 \tag{39}$$

Since the body is subjected to an uniform temperature distribution, the area integral part involving temperature in Eq. (27) becomes zero and the equation reduces to the normal domain integral. Similarly, in Eq. (28), the area integral involving temperature vanishes.

The J_1 and $J_{2\rho}$ domain integrals have been computed using a new finite element and fracture analysis code developed using MATLAB. The J_1 domain integral is obtained using only the outer contour, whereas the $J_{2\rho}$ domain integral is computed together with the line integral along the outer contour incorporating the portion of material interface. Three different contour paths are considered in order to demonstrate the path independence of the J_1 integral.

Table 2 shows the results of J_1 domain integral obtained for the three different paths together with the analytical result. It is seen that the error in the average of all the three paths of the J_1 integral is almost negligible. Furthermore, there is no variation in the computed J_1 values for the three different paths, thus demonstrating the path independence.

Table 3 shows the results of $J_{2\rho}$ computed for four different values of the inner circle of radius ρ and also for three different integration paths. It is seen that the computed results are in close agreement with the analytical ones. In addition, the results do not vary much for the three different paths indicating the path independence of $J_{2\rho}$ integral.

Table 4 shows the normalized Mode 1 and Mode 2 stress intensity factors for the four different values of ρ together with the analytical values, as given by Erdogan [24]. It is seen that there is good agreement in the computed solution with an average error of about -0.03% in $\tilde{K}_1$ and -0.02% in $\tilde{K}_2$. It may be noted that the SIF $K_1 + iK_2$ defined in this paper is related to the SIF $\hat{K}_1 + i\hat{K}_2$ defined in the work by Erdogan [24] as

$$K_1 + iK_2 = \frac{(\hat{K}_1 + i\hat{K}_2)}{\cosh\pi\epsilon} \tag{40}$$

3.2 Jointed Dissimilar Semi-Infinite Plates With Double Edge Cracks Subjected to Uniform Thermal Flow. In this case study, a semi-infinite bimaterial plate with double edge cracks and subjected to constant heat flux of $q_f = 10^5$ W/m^2 in the negative

Table 2 Results of the J_1 integral for all the three case studies (plane strain)

J_1	Analytical	Path 1	Path 2	Path 3	Average	% Error
Case study 1	549.38	549.39	549.36	549.35	549.36	0.003
Case study 2	14050.12	14054.9	14054.5	14054.23	14054.5	−0.03
Case study 3	1.2477×10^8	1.2484×10^8	1.2483×10^8	1.2482×10^8	1.2483×10^8	−0.045

Table 3 Case study 1: results of $J_{2\rho}$ for different values of ρ (plane strain)

ρ	Analytical $J_{2\rho}$	Numerical $J_{2\rho}$ Path 1	Path 2	Path 3	Average $J_{2\rho}$	% error
6.591×10^{-3}	−247.22	−252.71	−249.9	−248.90	−250.33	−1.26
8.845×10^{-3}	−229.61	−232.16	−229.27	−228.9	−230.11	−0.22
1.004×10^{-2}	−221.88	−219.4	−218.93	−218.37	−218.9	1.34
1.259×10^{-2}	−207.87	−205.19	−205.08	−204.34	−204.87	1.44

y-direction is considered (Fig. 8). The material properties are the same as those considered in the previous case study, as depicted in Table 1. Since the geometry of this plate and the one considered in the previous case study are the same, the same finite element mesh is used here with plane strain conditions. Figure 9 shows the left half of the geometry with displacement and thermal boundary conditions. No thermal insulation is assumed along the cracks. This is achieved by making the coincident nodes on the crack surfaces to coincide in the thermal analysis. Only during mechanical load analysis the constraints on the nodes lying on the crack surfaces are removed.

Ikeda and Sun [11] solved this problem numerically using the virtual crack extension method along with superposition method, and Banks-Sills and Dolev [12] solved using the M-integral concept. Brown and Erdogan [25] presented the exact solution for the thermal stress intensity factors as

$$K_1 = \frac{q_f\sigma_0(\alpha_2\eta_2k_1 - \alpha_1\eta_1k_2)b\sqrt{\pi b}[1-4\epsilon^2]}{2k_1k_2}(2b)^{-i\epsilon}$$

$$K_2 = \frac{q_f\sigma_0(\alpha_2\eta_2k_1 - \alpha_1\eta_1k_2)b\sqrt{\pi b}(2\epsilon)}{k_1k_2}(2b)^{-i\epsilon} \tag{41}$$

where η_1 and η_2 are as defined in the previous case study, k_1 and k_2 are the coefficients of heat conduction, and q_f is the thermal flow.

Since the crack surfaces are not insulated, i.e., the presence of crack does not influence the temperature distribution, the effect of far field flux on the body in negative y-direction will cause linear variation in temperature as a function of y. Therefore, the area integral part involving the temperature in Eq. (27) vanishes and the whole expression on the right hand side reduces to the normal domain integral. On the other hand, in Eq. (28), the area integral involving the temperature is nonzero as such with temperature gradient in the y-direction attaining a constant value.

The J_1 and $J_{2\rho}$ integrals have been computed as in the previous case study. Table 2 again shows the results of the J_1 integral computed for the three different paths along with the analytical results. It is seen that the error in the average of all three paths of the J_1 integral is less than 0.05%. Furthermore, there is negligible variation in computed J_1 values for the three different paths, thus demonstrating the path independence.

Table 5 shows the results of $J_{2\rho}$ computed for the four different values of the inner circle of radius ρ and also for three different

Table 4 Case study 1: normalized Mode 1 and Mode 2 stress intensity factors for different values of ρ (plane strain)

ρ	$\tilde{K}_1$	$\tilde{K}_2$
6.591×10^{-3}	−0.09889	1.0069
8.845×10^{-3}	−0.09887	1.0067
1.004×10^{-2}	−0.09880	1.0068
1.259×10^{-2}	−0.09883	1.0070
Average	−0.09885	1.0068
Erdogan's solution	−0.09888	1.0066
% error	−0.03	−0.02

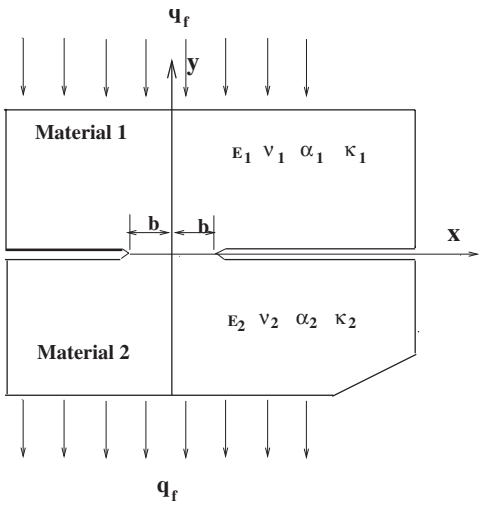


Fig. 8 Semi-infinite bimaterial plate with noninsulated crack subjected to uniform thermal flow

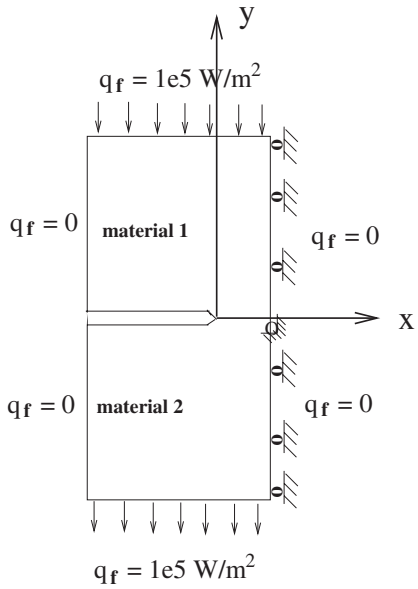


Fig. 9 Displacement and thermal boundary conditions for jointed dissimilar semi-infinite plate with noninsulated double edge crack under uniform thermal flow

Table 5 Case study 2: results of $J_{2\rho}$ for different values of ρ (plane strain)

ρ	Analytical $J_{2\rho}$	Numerical $J_{2\rho}$ Path 1	Path 2	Path 3	Average $J_{2\rho}$	% error
3.534×10^{-3}	4.146×10^{3}	4.312×10^{3}	4.251×10^{3}	4.060×10^{3}	4.209×10^{3}	−1.5
5.530×10^{-3}	3.388×10^{3}	3.485×10^{3}	3.434×10^{3}	3.326×10^{3}	3.415×10^{3}	−0.80
7.695×10^{-3}	2.818×10^{3}	2.861×10^{3}	2.82×10^{3}	2.701×10^{3}	2.794×10^{3}	0.83
8.845×10^{-3}	2.575×10^{3}	2.584×10^{3}	2.564×10^{3}	2.484×10^{3}	2.544×10^{3}	1.2

integration paths. It is seen that the computed results are in close agreement with the analytical ones. In addition, the results do not vary much for the three different paths indicating the path independence of $J_{2\rho}$-integral.

Table 6 shows the normalized Mode 1 and Mode 2 stress intensity factors for four different values of ρ together with the analytical values, as given by Erdogan [24]. It is seen that there is good agreement in the computed solution with average errors of -0.01% in $\tilde{K}_1$ and -0.03% in $\tilde{K}_2$.

3.3 An Infinite Body With Insulated Central Crack Subjected to Uniform Thermal Flow. In this case study, analysis is done on an infinite body with central insulated crack subjected to constant heat flux of $q_f=10^5$ W/m^2 in the negative y-direction under plane strain conditions, as shown in Fig. 10. Both the mechanical and thermal properties are considered to be the same as in the previous examples. The dimensions of semi-infinite plate are taken as 80 units$\times$160 units with a crack length of 4 units. Taking advantage of the symmetry in geometry and loading, only the right half of this bimaterial infinite plate is modeled using the finite element software ANSYS [23]. Figure 11 shows the right half of the geometry with displacement and thermal boundary conditions.

Table 6 Case study 2: normalized Mode 1 and Mode 2 stress intensity factors for different values of ρ (plane strain)

ρ	$\tilde{K}_1$	$\tilde{K}_2$
3.534×10^{-3}	0.4959	0.1258
5.530×10^{-3}	0.4958	0.1258
7.695×10^{-3}	0.4959	0.1258
8.845×10^{-3}	0.4958	0.1257
Average	0.4959	0.1258
Erdogan's solution	0.4958	0.1257
% error	−0.01	−0.03

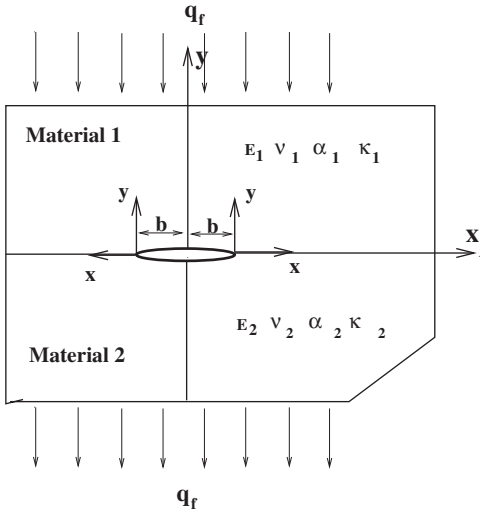


Fig. 10 Semi-infinite bimaterial plate with insulated central crack subjected to uniform thermal flow

The results of the convergence study on this problem are shown in Fig. 12 for a value of $\rho=7.806\times10^{-3}$. It is seen that with increasing mesh density, the computed $J_{2\rho}$ converges to the analytically obtained value, and for FE mesh using 17,000–19,000 number of elements there are very little variations. Therefore, a finite element mesh consisting of 17,882 elements with 54,195 nodes have been used for subsequent analysis. The finite element mesh and contour path used for this problem are of the same type, as shown in Fig. 6. In this example, all integrals in Eqs. (27) and (28) are nonzero and contribute to J_1- and J_2-integrals, thereby increasing the complexity of computations.

As in the previous two cases, Table 2 shows the results of J_1-integral obtained for three different contour paths together with the analytical results. The analytical values shown in this table are taken from Ref. [12]. It is seen that the error in the average of all the three paths of the J_1-integral is less than 0.05%. Furthermore, the variation in the computed J_1 values for the three different paths does not vary much, thus demonstrating the path independence.

Table 7 shows the results of $J_{2\rho}$ computed for four different values of the inner circle of radius ρ and also for three different integration paths. It is seen that the computed results are in close agreement with the analytical ones. In addition, the results do not vary much for the three different paths indicating the path independence of $J_{2\rho}$ integral.

Table 8 shows the normalized Mode 1 and Mode 2 stress intensity factors for the four different values of ρ together with the

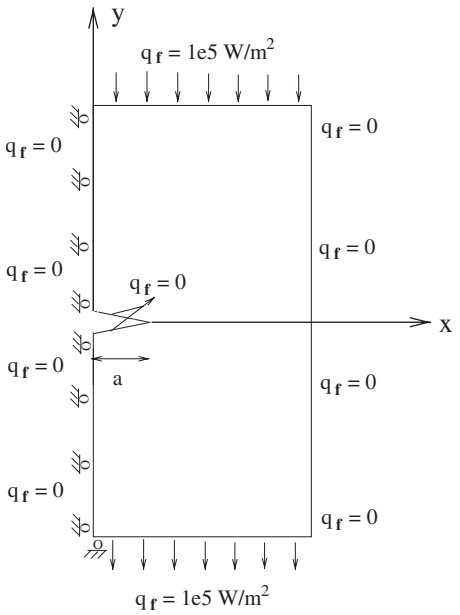


Fig. 11 Displacement and thermal boundary conditions for semi-infinite bimaterial plate with insulated central crack subjected to uniform thermal flow

Table 7 Case study 3: results of $J_{2\rho}$ for different values of ρ (plane strain)

ρ	Analytical $J_{2\rho}$	Numerical $J_{2\rho}$ Path 1	Path 2	Path 3	Average $J_{2\rho}$	% error
5.607×10^{-3}	8.767×10^{7}	8.912×10^{7}	8.863×10^{7}	8.850×10^{7}	8.875×10^{7}	−1.22
7.806×10^{-3}	8.469×10^{7}	8.685×10^{7}	8.457×10^{7}	8.298×10^{7}	8.480×10^{7}	−0.13
8.976×10^{-3}	8.336×10^{7}	8.391×10^{7}	8.281×10^{7}	8.198×10^{7}	8.290×10^{7}	0.55
1.146×10^{-3}	8.096×10^{7}	8.213×10^{7}	7.979×10^{7}	7.909×10^{7}	8.033×10^{7}	0.77

Table 8 Case study 3: normalized Mode 1 and Mode 2 stress intensity factors for different values of ρ (plane strain)

ρ	$\tilde{K}_1$	$\tilde{K}_2$
5.607×10^{-3}	−4.1175	0.4232
7.806×10^{-3}	−4.1174	0.4233
1.99×10^{-3}	−4.1174	0.4234
1.146×10^{-2}	−4.1174	0.4235
Average	−4.1175	
Erdogan'ssolution	−4.1163	0.4245
% error	−0.026	0.27

analytical values obtained from Ref. [12]. It is seen that there is good agreement in the computed solution with average errors of −0.03% in K_1 and 0.27% in K_2.

4 Conclusions

For a homogeneous crack problem, closed form expression for J_2 is available and it is physically identified as the energy release rate due to the crack tip advancing normal to its original orientation but the same cannot be said for bimaterial interface crack problems since it is shown to be nonexistent. In this work, the concept of J_2 domain integral is proposed for computing the stress intensity factors of bimaterial interface cracks subjected to thermal loads. It is shown that, in the presence of thermal stress, the J_k domain integral over a closed path, which does not enclose singularities, is not equal to zero but is the function of the temperature and body force. A method is proposed to compute the stress intensity factors for bimaterial interface crack subjected to thermal loading by combining this domain integral with the J_k integral. The proposed method is validated by solving standard problems with known solutions.

In order to properly characterize the near-tip behavior, it is required that the cutoff radius ρ be taken small enough so the integration contour for $J_{2\rho}$ be inside the region dominated by singularity zone. It is seen that the results of K_1 and K_2 are insensitive to different selections of ρ.

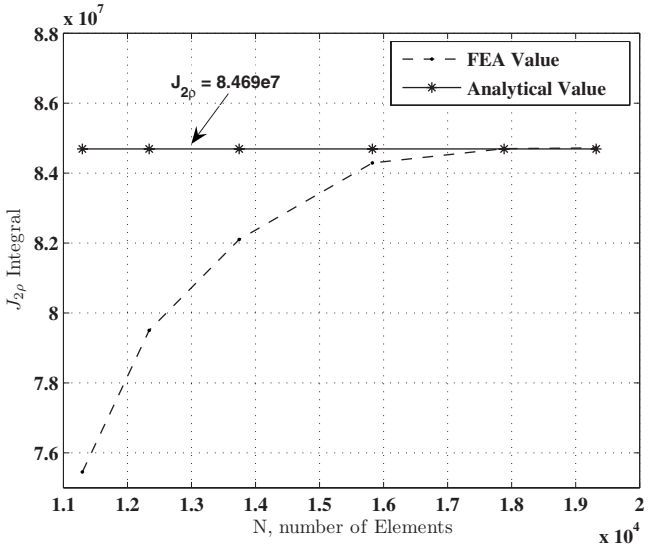


Fig. 12 Influence of mesh refinement on $J_{2\rho}$ for semi-infinite bimaterial plate under uniform thermal flow ($\rho=7.806\times10^{-3}$)

Appendix: Elastic Fields for Interface Cracks

Williams [26] performed an asymptotic analysis of the elastic fields at the tip of an open interface crack and found that the stresses and displacements behaved in an oscillatory manner as

$$\sigma \sim r^{-1/2}(\sin,\cos)(\varepsilon \log r) \tag{A1}$$

$$u \sim r^{1/2}(\sin,\cos)(\varepsilon \log r) \tag{A2}$$

$$\varepsilon = \frac{1}{2\pi}\log\left(\frac{\kappa_1\mu_2+\mu_1}{\kappa_2\mu_1+\mu_2}\right) \tag{A3}$$

where μ_j is the shear modulus of material j, $\kappa_j=(3-\nu_j)/(1+\nu_j)$ for plane stress and $\kappa_j=3-4\nu_j$ for plane strain, and ν_j is Poisson's ratio of material j.

Considering the bimaterial interface crack in Fig. 13, for traction-free plane problems the near-tip normal and shear stresses σ_{yy} and τ_{xy} may conveniently be expressed in terms of complex stress intensity factors [27] as

$$\sigma_{yy}+i\tau_{xy} = \frac{(K_1+iK_2)r^{i\varepsilon}}{\sqrt{(2\pi r)}} \tag{A4}$$

where $i=\sqrt{-1}$, K_1 and K_2 are the components of the complex stress intensity factor $K=K_1+iK_2$. From Eq. (A4), it is seen that the usual definition of the stress intensity factors as $\lim_{r\to0}\{(2\pi r)^{1/2}\sigma\}$ will not work and it produces logarithmically infinite factors. Furthermore, any attempt to define the stress intensity factors without reference to a characteristic length, such as the crack length, will produce dimensionally meaningless results [28].

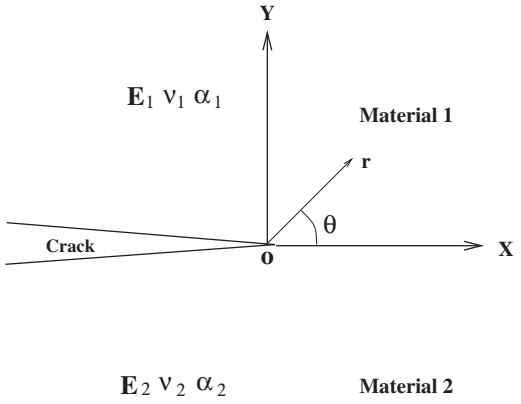


Fig. 13 A bimaterial interface crack

References

[1] Irwin, G. R., 1957, "Analysis of Stresses and Strains Near the End of a Crack Traversing a Plate," ASME J. Appl. Mech., **24**(3), pp. 361–364.

[2] Rice, J. R., 1968, "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks," ASME J. Appl. Mech., **35**(2), pp. 379–386.

[3] Parks, D. M., 1974, "A Stiffness Derivative Finite Element Technique for Determination of Crack Tip Stress Intensity Factors," Int. J. Fract., **10**, pp. 487–502

[4] Carlsson, L. A., and Prasad, S., 1993, "Interface Fracture of Sandwich Beams," Eng. Fract. Mech., **44**(4), pp. 581–590.

[5] Yau, J. F., and Wang, S. S., 1984, "An Analysis of Interface Crack Between Dissimilar Isotropic Materials Using Conservation Integrals in Elasticity," Eng. Fract. Mech., **20**, pp. 423–432.

[6] Matos, P. P. L., McMeeking, R. M., Charalambides, P. G., and Drory, M. D., 1989, "A Method of Calculating Stress Intensities in Bimaterial Fracture," Int. J. Fract., **40**, pp. 235–254.

[7] Bank-Sills, L., Travitzky, N., Ashkenazi, D., and Eliasi, R., 1999, "A Methodology for Measuring Interface Fracture Toughness of Composites Materials," Int. J. Fract., **99**, pp. 143–161.

[8] Miyazaki, N., Ikeda, T., Soda, T., and Munakata, T., 1993, "Stress Intensity Factor Analysis of Interface Crack Using Boundary Element Method—Application of Contour-Integral Method," Eng. Fract. Mech., **45**(5), pp. 599–610.

[9] Miyazaki, N., Ikeda, T., Soda, T., and Munakata, T., 1993, "Stress Intensity Factor Analysis of Interface Crack Using Boundary Element Method—Application of Virtual Crack Extension Method," JSME Int. J., Ser. A, **36**, pp. 36–42.

[10] Ryoji, Y., and Sang-Bong, C., 1989, "Efficient Boundary Element Analysis of Stress Intensity Factors for Interface Cracks in Dissimilar Materials," Eng. Fract. Mech., **34**, pp. 179–188.

[11] Ikeda, T., and Sun, C. T., 2001, "Stress Intensity Factor Analysis for an Interface Crack Between Dissimilar Isotropic Materials Under Thermal Stress," Int. J. Fract., **111**, pp. 229–249.

[12] Banks-Sills, L., and Dolev, O., 2004, "The Conservative *M* Integral for Thermal-Elastic Problems," Int. J. Fract., **125**, pp. 149–170.

[13] Hong, C. C., and Stern, M., 1978, "The Computation of Stress Intensity Factors in Dissimilar Materials," J. Elast., **8**(1), pp. 21–36.

[14] Sun, C. T., and Jih, C. J., 1987, "On Strain Energy Release Rates for Interfacial Cracks in Bi-Material Media," Eng. Fract. Mech., **28**, pp. 13–20.

[15] Wilson, W. K., and Yu, I. W., 1979, "The Use of the *J* Integral in Thermal Stress Crack Problem," Int. J. Fract., **15**, pp. 377–387.

[16] Chang, J. H., and Pu, L. P., 1997, "A Finite Element Approach for J_2 Calculation in Anisotropic Materials," Compos. Struct., **62**(4), pp. 635–647.

[17] Eischen, J. W., 1987, "An Improved Method for Calculating J_2 Integral," Eng. Fract. Mech., **26**(5), pp. 691–700.

[18] Shih, C. F., Moran, B., and Nakamura, T., 1986, "Energy Release Rate Along Three Dimension Crack Front in a Thermally Stress Body," Int. J. Fract., **30**, pp. 79–102.

[19] Khandelwal, R., and Chandra Kishen, J. M., 2006, "Complex Variable Method of Computing J_k for Bi-Material Interface Cracks," Eng. Fract. Mech., **73**, pp. 1568–1580.

[20] Knowles, J. K., and Sternberg, E., 1972, "On a Class of Conservation Laws in Linearilized and Finite Elastostatics," Arch. Ration. Mech. Anal., **44**, pp. 187–211.

[21] Banks-Sills, L., and Sherman, D., 1992, "On the Computation of Stress Intensity Factors for Three-Dimensional Geometries by Means of the Stiffness Derivative and J-Integral Methods," Int. J. Fract., **53**, pp. 1–20.

[22] Li, F. Z., Shih, C. F., and Needleman, A., 1985, "A Comparison of Methods for Calculating Energy Release Rates," Eng. Fract. Mech., **21**, pp. 405–421.

[23] 2005, ANSYS, University Advanced, v. 10.0, ANSYS, Inc., Canonsburg, PA.

[24] Erdogan, F., 1965, "Stress Distribution in Bounded Dissimilar Materials With Cracks," Trans. ASME, J. Appl. Mech., **32**, pp. 403–410.

[25] Brown, E. J., and Erdogan, F., 1968, "Thermal Stress in Bounded Materials Containing Cuts on the Interface," Int. J. Fract., **6**, pp. 517–529.

[26] Williams, M. L., 1959, "The Stresses Around a Fault or Crack in Dissimilar Media," Bull. Seismol. Soc. Am., **49**(2), pp. 199–204.

[27] Rice, J. R., Suo, Z., and Wang, J. S., 1990, *Mechanics and Thermodynamics of Brittle Interfacial Failure in Bi-Material Systems*, G. C. Ruhle, A. G. Evans, M. F. Ashby, and J. P. Hirth, eds., Pergamon, New York.

[28] Comninou, M., 1990, "An Overview of Interface Cracks," Eng. Fract. Mech., **37**, pp. 197–208.

Vibration and Snap-Through of Bent Elastica Strips Subjected to End Rotations

R. H. Plaut
Fellow ASME
Department of Civil and Environmental Engineering,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061-0105

L. N. Virgin[1]
Mem. ASME
Department of Civil and Environmental Engineering,
Duke University,
Durham, NC 27708-0287
e-mail: l.virgin@duke.edu.

A flexible strip is rotated at its ends until it forms a deep circular arc above its ends. Then the ends are kept immovable and are rotated downward until the arch suddenly snaps into an inverted configuration. The strip is analyzed as an inextensible elastica. Two-dimensional equilibrium shapes, vibration modes and frequencies, and critical rotations for snap-through are determined using a shooting method. Experiments are also conducted and results are compared with those from the analysis. The agreement is good. In addition, a microelectromechanical systems (MEMS) example is analyzed, in which an electrostatic force below a buckled strip causes the strip to snap downward, and the critical force is obtained as a function of the vertical gap. [DOI: 10.1115/1.3086783]

Keywords: bent elastica, snap-through, vibration

1 Introduction

Snap-through of arches or buckled beams is an interesting phenomenon. Most previous studies have considered downward loads that can cause the structure (in whole or in part) to suddenly jump from a configuration above the horizontal line connecting its end supports into an inverted configuration below the horizontal. Such behavior can be useful in switches, for example, and there has been recent interest in applications for MEMS [1,2]. Snap-through can also be induced by rotation of the ends of the structure. Analyses of shallow arches under end rotations have been presented in Refs. [3–5], and this type of "loading" (displacement control) is investigated here for nonshallow (deep) bent strips.

A related problem was analyzed in Ref. [6]. A flexible strip was attached at its ends to a substrate comprised of two rigid plates connected with a hinge. The plates were initially flat, and then one of the plates was rotated upward. In contrast to that problem, here the ends of the strip remain a fixed distance apart rather than coming together. Also, experiments are included in the present study.

A slender elastic strip is considered. It is unstrained when straight, and the ends are rotated and moved so that, when self-weight is neglected, the strip has an initial circular shape with constant bending moment and no internal axial or shear forces. Then the ends are rotated downward quasistatically with equal magnitudes, so that the resulting equilibrium shapes remain symmetric about the midpoint of the strip. A portion of the strip near each end displaces below the horizontal. At a critical end rotation, the central portion of the strip suddenly snaps downward into an inverted position. The equilibrium shapes before snap-through are investigated, along with small vibrations about equilibrium.

In the analysis, the strip is modeled as an inextensible elastica with fixed ends. (Elastica arches were treated in Refs. [7,8].) A shooting method is utilized to obtain numerical solutions for equilibrium shapes, and for vibration modes and frequencies. The critical end rotation is obtained by determining when the lowest frequency reduces to zero. In conjunction with the analysis, experiments are conducted with polycarbonate strips. Displacements, vibration frequencies, and critical values of the end rotations are recorded. The experimental results are compared with those from the analytical work, and the correlation is good.

The analytical formulation is presented in Sec. 2, and the experiments are described in Sec. 3. Section 4 contains numerical and experimental results. A related problem is examined in Sec. 5, directly related to MEMS, in which a strip with pinned ends is buckled upward and then a downward electrostatic force causes snap-through. Concluding remarks are given in Sec. 6.

2 Analytical Formulation and Numerical Solution Procedure

Consider a thin, flexible, inextensible, unshearable, uniform, elastic strip with length S_0, bending stiffness EI, cross-sectional area A, mass per unit length μ, weight per unit length W, and mass moment of inertia $\mu I/A$. The ends are clamped a distance L apart, as shown in Fig. 1. The rectangular cross section has width B and depth D, so that $A=BD$ and $I=BD^3/12$. The initial height of the bent strip is H_0.

The effect of the weight of the strip on the equilibrium configuration and vibration frequencies was considered in the initial analysis, and then was neglected after being shown to be negligible for the polycarbonate strips used in the experiments. Damping was neglected in the analysis, and only in-plane motion was examined, since the moment of inertia with respect to out-of-plane displacements is much greater than that for in-plane displacements. Transverse and axial inertia forces are included. Rotary inertia was also included in the formulation [9], but then was found to have a negligible effect for the polycarbonate strips and was ignored.

From the left end, the arc length is S, the horizontal and vertical coordinates are $X(S,T)$ and $Y(S,T)$, respectively, and the angle in radians is $\theta(S,T)$, where T denotes time. The total arc length is S_0, and the angle with the horizontal at $S=0$ is denoted α. In the initial circular equilibrium configuration (assuming negligible self-weight), the angle is $\theta_e(S)$, with $\theta_e(0)=\alpha_0$, and the height is H_0. The downward rotation of the ends from the initial configuration is ϕ.

The bending moment $M(S,T)$ is positive if it is counterclockwise on a positive face, the horizontal force $P(S,T)$ is positive in compression, and the vertical force $Q(S,T)$ is positive if downward on a positive face. Dimensional vibration frequencies are denoted Ω.

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 18, 2008; final manuscript received December 29, 2008; published online April 23, 2009. Review conducted by William Scherzinger.

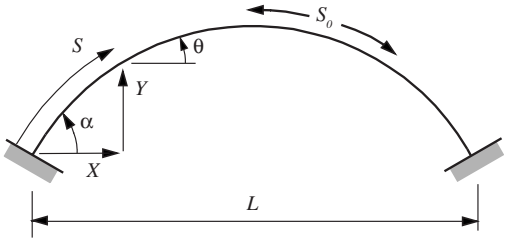


Fig. 1 Schematic of bent strip

Based on the geometry, the moment-curvature relationship for an elastica, and dynamic equilibrium, the governing equations are [10]

$$\partial X/\partial S = \cos\,\theta, \quad \partial Y/\partial S = \sin\,\theta$$

$$\partial\theta/\partial S = M/EI, \quad \partial M/\partial S = (\mu I/A)\partial^2\theta/\partial T^2 + Q\cos\,\theta - P\sin\,\theta \quad (1)$$

$$\partial P/\partial S = -\mu\partial^2 X/\partial T^2, \quad \partial Q/\partial S = -\mu\partial^2 Y/\partial T^2 - W$$

The following nondimensional quantities are defined:

$$x = X/S_0, \quad y = Y/S_0, \quad s = S/S_0, \quad m = MS_0/EI, \quad p = PS_0^2/EI$$

$$q = QS_0^2/EI, \quad t = (T/S_0^2)\sqrt{EI/\mu}, \quad \omega = \Omega S_0^2\sqrt{\mu/EI}, \quad l = L/S_0 \quad (2)$$

$$h_0 = H_0/S_0, \quad w = WS_0^3/EI, \quad \gamma = D^2/(12S_0^2), \quad \phi = \alpha_0 - \alpha$$

The governing equations in nondimensional form are then given by

$$\partial x/\partial s = \cos\,\theta, \quad \partial y/\partial s = \sin\,\theta$$

$$\partial\theta/\partial s = m, \quad \partial m/\partial s = \gamma\partial^2\theta/\partial t^2 + q\cos\,\theta - p\sin\,\theta \quad (3)$$

$$\partial p/\partial s = -\partial^2 x/\partial t^2, \quad \partial q/\partial s = -\partial^2 y/\partial t^2 - w$$

where $0<s<1$ and $0<x<l$, with $l<1$.

The subscript e will denote quantities associated with an equilibrium configuration, and the subscript d will be used for dynamic quantities corresponding to small vibrations about equilibrium. The variables in Eq. (3) are written in the following form:

$$x(s,t) = x_e(s) + x_d(s)\sin\,\omega t, \quad y(s,t) = y_e(s) + y_d(s)\sin\,\omega t$$

$$\theta(s,t) = \theta_e(s) + \theta_d(s)\sin\,\omega t, \quad m(s,t) = m_e(s) + m_d(s)\sin\,\omega t \quad (4)$$

$$p(s,t) = p_e + p_d(s)\sin\,\omega t, \quad q(s,t) = q_e(s) + q_d(s)\sin\,\omega t$$

From Eqs. (3) and (4), the governing equations for equilibrium are found to be

$$x_e' = \cos\,\theta_e, \quad y_e' = \sin\,\theta_e$$

$$\theta_e' = m_e, \quad m_e' = q_e\cos\,\theta_e - p_e\sin\,\theta_e \quad (5)$$

where p_e is constant and

$$q_e(s) = w(1/2 - s) \quad (6)$$

so that $q_e=0$ if the self-weight w is neglected. In that case, the initial shape is circular, the corresponding bending moment $m(s)$ and curvature $\theta_0'(s)$ are constant, $p_e=0$, and the nondimensional span and initial height are related to the end angles by

$$l = \frac{\sin\,\alpha_0}{\alpha_0}, \quad h_0 = \frac{1-\cos\,\alpha_0}{2\alpha_0} \quad (7)$$

where α_0 is in radians.

The governing equations for linear vibrations about the equilibrium state are given by

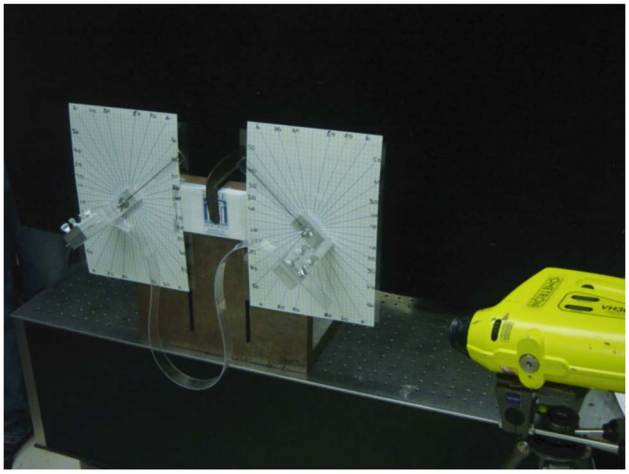


Fig. 2 Photo of experimental system

$$x_d' = -\theta_d\sin\,\theta_e, \quad y_d' = \theta_d\cos\,\theta_e$$

$$\theta_d' = m_d, \; m_d' = -\gamma\omega^2\theta_d + (q_d - p_e\theta_d)\cos\,\theta_e - (p_d + q_e\theta_d)\sin\,\theta_e \quad (8)$$

$$p_d' = \omega^2 x_d, \quad q_d' = \omega^2 y_d$$

If $\alpha_0>142.5$ deg, corresponding to $l=0.1224$, the two sides of the strip touch each other when a certain end rotation is attained. Similar self-contact at a point has been described in Refs. [6,10] and references cited therein.

Numerical solutions are obtained using a shooting method with the subroutines NDSOLVE and FINDROOT in MATHEMATICA [11]. For the case of equal end rotations and symmetric equilibrium shapes, the left half of the strip is considered. The end rotation ϕ and nondimensional half-span $l/2$ are specified, and q_e is given by Eq. (6). The values of p_e and $m_e(0)$ are varied until the solution of Eq. (5) satisfies $x_e(0.5)=l/2$ and $\theta_e(0.5)=0$ with sufficient accuracy.

For symmetric equilibrium configurations involving self-contact at a point, the left half of the strip is considered, and the portion from the left end to the contact point, whose length is denoted a, is analyzed. For the upper loop above the contact point, it is known that $q_e=0$, $m_e(a)=3.027634/(0.5-a)$, $p_e=[m_e(a)/0.96163]^2$, and $m_e(0.5)=1.71018\sqrt{p_e}$ [12]. As in Ref. [10], coordinate systems are set up at the contact point, and the shooting method is applied using geometrical conditions at the left end, contact point, and midpoint of the strip. The length a is specified, and the angle at the left end is obtained from the resulting shape.

To compute vibrations using Eq. (8), a shooting method is again applied. Resulting parameters from the equilibrium solution are utilized. The quantity $m_d(1)$ is given a value, since the amplitude of the vibration mode is arbitrary, and the quantities $p_d(0)$, $q_d(0)$, and ω^2 are varied until the conditions $x_d(1)=0$, $y_d(1)=0$, and $\theta_d(1)=0$ are satisfied. Frequencies for the initial bent strip are the same as for a clamped circular arch, as described in the Appendix.

3 Experiments

A thin polycarbonate strip was used to verify some of the theoretical calculations. A number of different thicknesses were available, but most of the tests were conducted on strips 1.524 mm thick. The width was fixed at 25.4 mm, and the length varied such that the initial end angle resulted in a circular arc. The specific weight was 11.2 kN/m^3 (although weight was typically neglected in the theoretical analysis) and Young's modulus was 2.4 GPa.

The experimental apparatus is shown in Fig. 2. Aluminum blocks clamped the strip at either end, and then each end could be

Table 1 Effect of strip thickness and self-weight on lowest frequency

	$\alpha_0=30$ deg		$\alpha_0=60$ deg		$\alpha_0=90$ deg	
Thickness	$\phi=0$	$\phi=\alpha_0$	$\phi=0$	$\phi=\alpha_0$	$\phi=0$	$\phi=\alpha_0$
0.508 mm	58.2	45.2	52.4	45.8	43.5	43.1
1.016 mm	57.0	45.0	55.7	43.8	46.6	44.2
1.524 mm	58.0	44.5	52.8	45.3	44.3	42.7
Th. ($w=0$)	58.93	44.18	51.97	43.34	43.27	40.73
Th. (up)	58.67	43.89	51.80	43.15	43.05	40.51
Th. (down)	59.19	44.48	52.14	43.54	43.49	40.94

rotated by a prescribed angle ϕ. The central point deflection (in the vertical direction) d was then recorded as a function of end angle.

For each value of ϕ, with $\alpha_0=30$ deg, 60 deg, 90 deg, and 120 deg, the lowest frequencies of small-amplitude vibration were measured using a laser velocity vibrometer. Small oscillations were initiated by applying a small perturbation, and the frequency content was extracted using the B&K PULSE software (Naerum, Denmark).

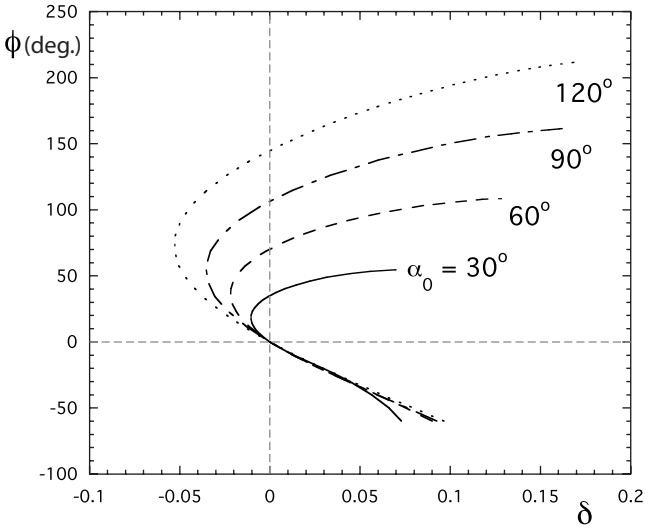


Fig. 3 End rotation as function of midpoint vertical deflection for $\alpha_0=30$ deg, 60 deg, 90 deg, and 120 deg

4 Results

The results are described in terms of nondimensional quantities. The effect of self-weight on the vibration frequencies is considered first. For the polycarbonate strip with a thickness of 1.524 mm, the weights per unit length w are 0.784, 1.21, and 2.65, respectively, for initial end angles $\alpha_0=30$ deg, 60 deg, and 90 deg. The lowest vibration frequency was computed for the case of self-weight with the bent strip in an upward initial position and also in a downward initial position (initial end angles of -30 deg, -60 deg, and -90 deg), as well as for $w=0$. Results are presented in the last three rows of Table 1 for the initial shape ($\phi=0$) and for end rotations $\phi=\alpha_0$ (so that the ends of the strip are horizontal). The effect of self-weight is negligible in Table 1, based on the parameters for the strips used in the experiments, and it is assumed that $w=0$ in subsequent results (it is possible that self-weight would have a greater influence for longer strips). Table 1 also includes experimental values for the lowest frequency for the standard strip thickness and for two thinner strips. There is finite precision in both the accuracy of setting the initial end angle and measuring the central vertical deflection. In some cases the clarity of a specific frequency may have been influenced by the proximity of an out-of-plane mode. No further data will be presented for the thinner strips.

Next, the effect of rotary inertia is examined. The same cases as in Table 1 are analyzed. The frequencies in Table 1 do not include rotary inertia, and when it is added in the equations of motion, the first four significant digits for the frequencies do not change. Rotary inertia is, therefore, neglected in subsequent results, i.e., $\gamma=0$ in Eq. (3).

Most of the results correspond to initial end angles $\alpha_0=30$ deg, 60 deg, 90 deg, and 120 deg. For these cases, the values of the base length l and initial height h_0 are listed in Table 2. Figure 3 shows the theoretical equilibrium paths for these four cases. The abscissa is the downward equilibrium deflection of the midpoint of the bent strip from its initial shape, denoted δ, i.e., $\delta=h_0-y_e(0.5)$. The end rotation ϕ is the ordinate. As the ends are rotated downward, the central portion of the strip initially moves upward ($\delta<0$). Then, with further end rotation, δ becomes positive.

Table 2 Base length, height, and critical end rotation

α_0	30 deg	60 deg	90 deg	120 deg
l	0.955	0.827	0.637	0.413
h_0	0.128	0.239	0.318	0.358
ϕ_{cr}	54.4 deg	108 deg	162 deg	212 deg
ϕ_{cr}/α_0	1.82	1.81	1.80	1.77

If the ends are rotated upward, $\phi<0$. The upward deflection of the center of the arch is shown in Fig. 3 for the range -60 deg $<\phi<0$. The curves for $\alpha_0=60$ deg, 90 deg, and 120 deg are close to each other for the whole range.

Equilibrium shapes for the case $\alpha_0=90$ deg at rotational increments of 20 deg are sketched in Fig. 4, along with the theoretical equilibrium path (solid curve) and experimental data (dots). The last shape is close to that for the critical end rotation $\phi_{cr}=162$ deg for snap-through. The analytical and experimental results are very close to each other. Figure 5 shows the results for $\alpha_0=30$ deg, 60 deg, and 120 deg. The scales on the axes are not identical.

The upper end of each curve in Figs. 3–5 corresponds to the point at which snap-through occurs. The corresponding values of ϕ_{cr} are listed in Table 2. The results demonstrate that snap-through is experienced when the magnitude of the angle of the ends below the horizontal is about 80% of its initial value above the horizontal (i.e., $\phi_{cr}\approx 1.8\alpha_0$).

Snap-through is determined by observing when the lowest vibration frequency ω_1 reduces to zero, and the corresponding vibration mode is the buckling mode. Figure 6 shows how the lowest frequency changes as ϕ is increased for $\alpha_0=30$ deg, 60 deg, 90 deg, and 120 deg. The corresponding mode is antisymmetric about the center of the strip ($s=0.5$), and hence the instability is of the bifurcation type. The upper ends of the equilibrium paths in Fig. 3 and the frequency curves in Fig. 6 have slopes that are almost zero; a zero slope would imply a limit point. Bifurcation occurs for all initial end angles α_0, even for very small angles. This is in contrast to results based on the usual shallow-arch theory, in which no instability occurs if α_0 is between zero and some value α_1, a limit point occurs if α_0 is in some range $\alpha_1<\alpha_0<\alpha_2$, and snap-through is associated with bifurcation if $\alpha_0>\alpha_2$. The shallow-arch theory assumes that the structure is compressible, whereas here the strip is assumed to be inextensible.

The effect of the end rotation ϕ on the lowest two or three frequencies is seen in Figs. 7(*a*)–7(*d*) for $\alpha_0=30$ deg, 60 deg, 90 deg, and 120 deg, respectively. Results from the analysis are shown as open circles, and results from the experiments are

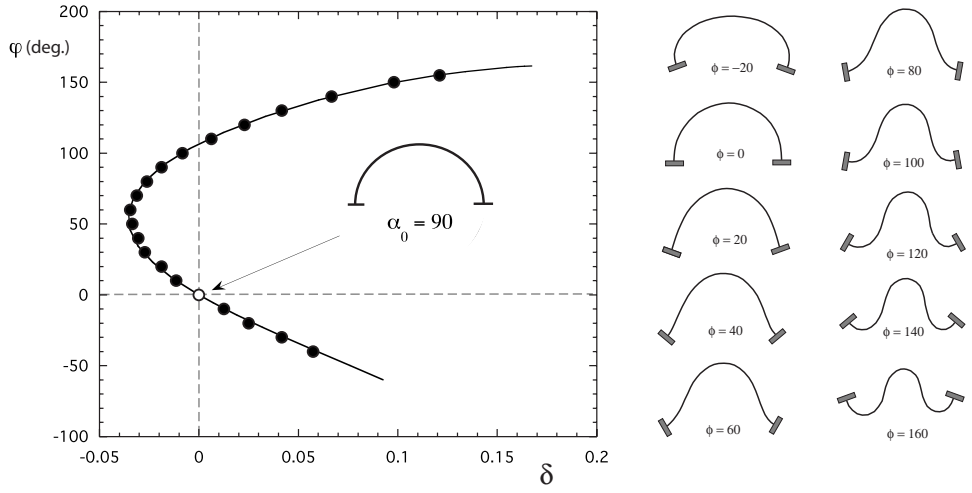


Fig. 4 End rotation as function of midpoint vertical deflection for α_0=90 deg; dots denote experimental data

shown as closed circles. For all α_0 values, the theoretical and experimental frequencies are very close to each other.

The first and third curves emanating from ϕ=0 are associated with modes having antisymmetric normal displacements, and the second curve corresponds to a symmetric mode. The first three mode shapes are depicted in Fig. 8 for the case α_0=90 deg and ϕ=149 deg. The amplitude of each mode is arbitrary, since the vibration analysis is based on linear equations. The second mode has no nodes, which is possible for the inextensible strip due to the changes in curvature in equilibrium (e.g., see Refs. [6,10]). The dashed curve is the equilibrium configuration for this case, and the solid curve shows the shape of the strip during vibration. The first and third modes involve a rocking type of motion, whereas the second mode is essentially a vertical motion.

For α_0=120 deg, the two curves in Fig. 7(*d*) for antisymmetric modes exhibit a veering phenomenon (i.e., they approach each other but do not intersect), and the curve for symmetric modes intersects the "third" curve twice. Figure 9 presents a close-up view of Fig. 7(*d*) in the approximate range 140 deg$<\phi<$200 deg. The veering and the upper crossing point are seen in this figure for both the theoretical and experimental data.

Frequencies associated with upward end rotations ($\phi<0$) are included in Fig. 7. The axial force in the strip tends to become more tensile as ϕ decreases, yet the second and third frequencies tend to decrease. In Fig. 7(*a*) for α_0=30 deg, the first two frequency curves intersect when ϕ reaches a certain negative value.

If the bent strip exhibits snap-through and then the ends are rotated further (i.e., ϕ is increased past ϕ_{cr}), the post-snap-through equilibrium states can be determined from Fig. 7 due to symmetry of the system about the horizontal axis. A frequency for ϕ is also a frequency for end rotation $2\alpha_0-\phi$. Therefore, for $\phi_{cr}<\phi<2\alpha_0$, the corresponding frequencies for small positive values of ϕ based on this relationship are applicable, and for $\phi>2\alpha_0$, corresponding values for $\phi<0$ can be used.

For the case α_0=150 deg, self-contact occurs when ϕ=171 deg. The corresponding equilibrium configuration of the strip is depicted in Fig. 10. If ϕ is increased further, the analysis predicts that self-contact at a point continues until ϕ reaches 306 deg, at which time the normal internal compressive force between the two sides at the contact point reduces to zero. The plot of the midpoint deflection δ for the range $0<\phi<306$ deg is shown in Fig. 11. In tests for this case of α_0=150 deg, however, out-of-plane displacement and snap-through occur before ϕ reaches 306 deg.

5 Electrostatic Force on MEMS Device

The problem treated in this section is related to the above work and to the motivation for this study. In Ref. [13], a beam was buckled upward by a compressive force, the ends were constrained to be stationary, and then a downward electrostatic force was applied. The slopes of the beam were relatively small, and shallow-arch theory was utilized in the analysis. Some experiments were also carried out with a microbeam. A similar study was conducted in Ref. [14] for a shallow, initially curved beam with clamped ends. These MEMS devices exhibit snap-through instability that can be utilized as a switch.

Here, a strip is bent upward into a circular shape and the ends are pinned, i.e., there is no bending moment at the ends, unlike the work in Secs. 2–4. In dimensional terms, a horizontal electrode is situated below the bent strip at a distance (gap) G_0 below the pinned supports. A voltage difference V_0 is applied, and the membrane deflects downward toward the plate due to the electrostatic Coulomb force. When the voltage reaches a critical value, snap-through occurs and the strip jumps into an inverted position. (In Refs. [13,14], this is called pull-in instability if the strip contacts the electrode when it jumps downward.)

Using the standard model [13,14], the downward force per unit of horizontal length is $\epsilon_0 B V_0^2/[2(G_0+Y)^2]$, where ϵ_0 is the dielectric constant of air. Hence, the electrostatic force is inversely proportional to the square of the distance between the electrode and the strip. With W=0, the last of Eq. (1) is replaced with

$$\partial Q/\partial S = -\mu \partial^2 Y/\partial T^2 - \frac{\epsilon_0 B V_0^2 \cos\theta}{2(G_0+Y)^2} \tag{9}$$

The nondimensional terms in Eq. (2) are used, along with

$$g_0 = \frac{G_0}{S_0}, \quad f = \frac{6 S_0 \epsilon_0 V_0^2}{D^3 E} \tag{10}$$

Equilibrium is governed by Eq. (5) plus

$$q_e' = \frac{-f\cos\theta_e}{(g_0+y_e)^2} \tag{11}$$

and vibrations are governed by Eq. (8), except that the last of those equations is replaced with

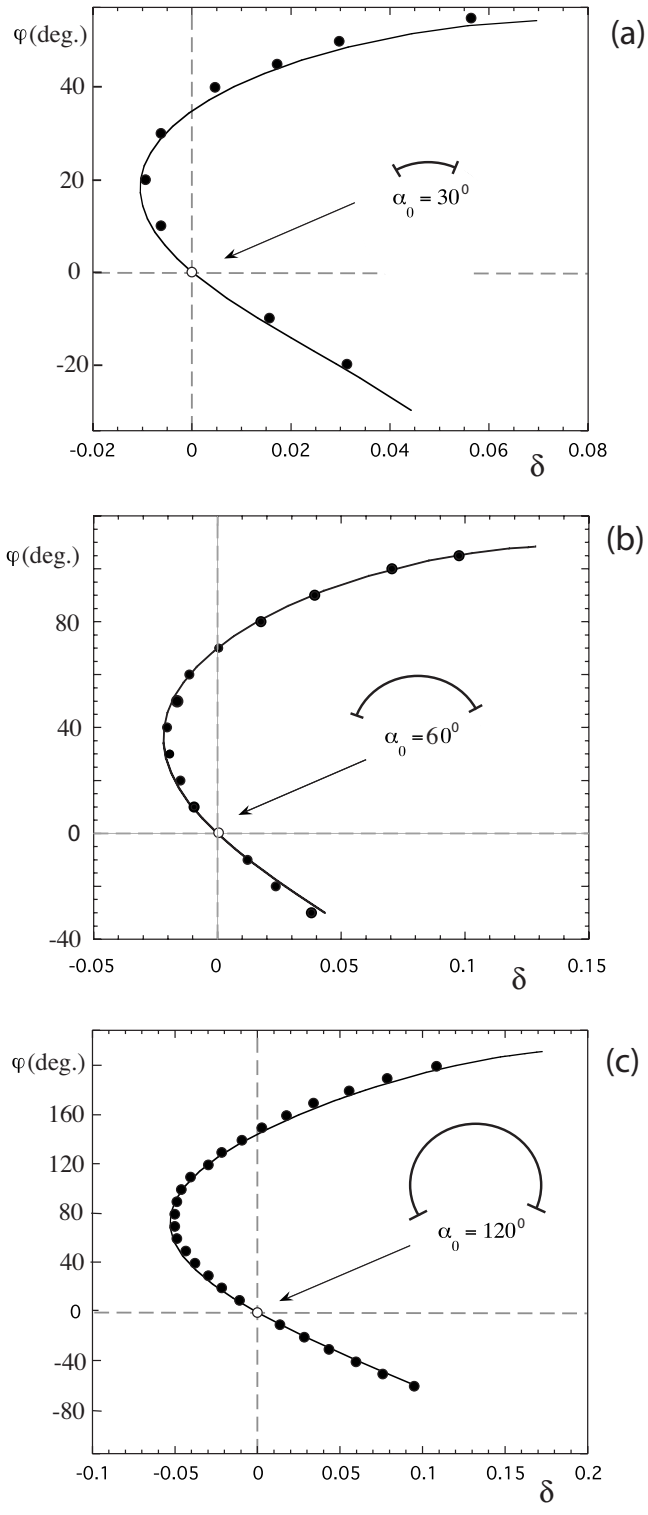


Fig. 5 End rotation as function of midpoint vertical deflection for (*a*) α_0=30 deg, (*b*) 60 deg, and (*c*) 120 deg; dots denote experimental data

$$q_d' = \omega^2 y_d + \frac{2f y_d \cos\theta_e}{(g_0 + y_e)^3} + \frac{f\theta_d \sin\theta_e}{(g_0 + y_e)^2} \tag{12}$$

The parameters w and γ are zero. At $s=0$ and 1, m_e and m_d are zero, and the end angles are not specified during loading. The stable equilibrium states are again symmetric with respect to $s=0.5$.

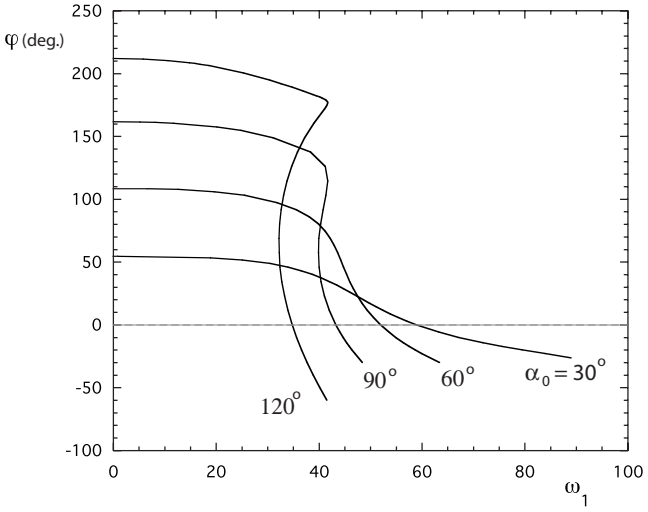


Fig. 6 End rotation as function of lowest frequency for α_0 =30 deg, 60 deg, 90 deg, and 120 deg

The appropriate nondimensional span l for a circular initial shape is determined by setting α_0, varying p until $\theta(0.5)=0$, and using $x(0.5)=l/2$ to get l. For $\alpha_0=15$ deg, 30 deg, 45 deg, and 60 deg, respectively, this yields $l=0.982926$, 0.932432, 0.850656, and 0.741020. In the examples, which use these initial angles, the bent strip remains above the x-axis until snap-through occurs, and then the whole strip suddenly inverts downward.

In the shooting method for equilibrium, f, g_0, and l are specified. The quantities p_e, $q_e(0)$, and $\theta_e(0)$ are varied until $x_e(0.5)=l/2$, $\theta_e(0.5)=0$, and $q_e(0.5)=0$. For vibration, the quantity $\theta_d(0)$ is given a value, and $p_d(0)$, $q_d(0)$, and ω^2 are varied until $x_d(1)=0$, $y_d(1)=0$, and $m_d(1)=0$. The force magnitude f is increased until the lowest vibration frequency decreases to zero, giving the critical value f_{cr} for snap-through. Again the buckling mode is antisymmetric with a single node at the midpoint of the strip, such as in Fig. 8(*a*).

Results for the critical force magnitude f_{cr} are computed for the range $0.01 \le g_0 \le 0.05$. The resulting relationship is almost linear, with $f_{cr} \approx c_0 + c_1 g_0$ where $(c_0, c_1) = (0.003, 3)$ for $\alpha_0=15$ deg, (0.06,11) for $\alpha_0=30$ deg, (0.26,25) for $\alpha_0=45$ deg, and (0.75,46) for $\alpha_0=60$ deg. In other words, for the numerical examples investigated, for a given bent strip, the square of the critical voltage is approximately a linear function of the vertical gap between the horizontal electrode and the ends of the strip.

6 Concluding Remarks

In-plane snap-through of a thin elastic strip has been examined. The ends of the strip are rotated upward such that the strip is initially circular and is "deep," and then the ends are rotated downward symmetrically (with constant horizontal span) until the central portion of the strip suddenly inverts. The onset of the snap-through is associated with a bifurcation point.

In the analysis, the strip is modeled as an inextensible elastica, and two-dimensional equilibrium shapes, vibration frequencies, and vibration modes are determined numerically using a shooting method. Initial angles of 30 deg, 60 deg, 90 deg, and 120 deg are considered in detail, and some results are also obtained for an initial angle of 150 deg (in which case the two sides of the strip contact each other above the center of the span as the ends are rotated). When the ends of the strip are angled downward, the first and third vibration modes involve primarily rocking motions, whereas the second mode is essentially a vertical vibration with no internal nodes.

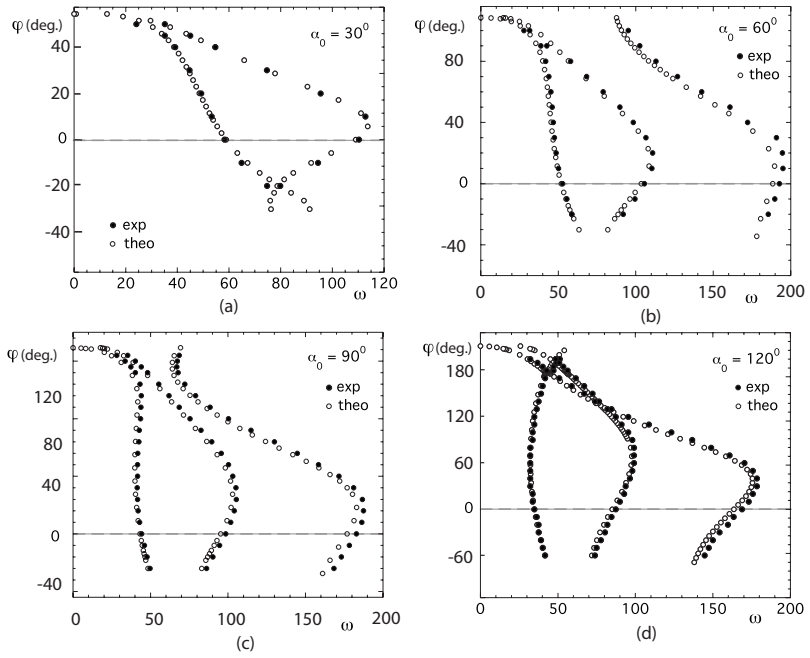


Fig. 7 End rotation as function of lowest frequencies for α_0 equal to (*a*) 30 deg, (*b*) 60 deg, (*c*) 90 deg, and (d) 120 deg

The critical value of the end rotation is obtained by observing when the lowest frequency reduces to zero. Snap-through occurs when the magnitude of the downward end angle is approximately 80% of the initial upward angle.

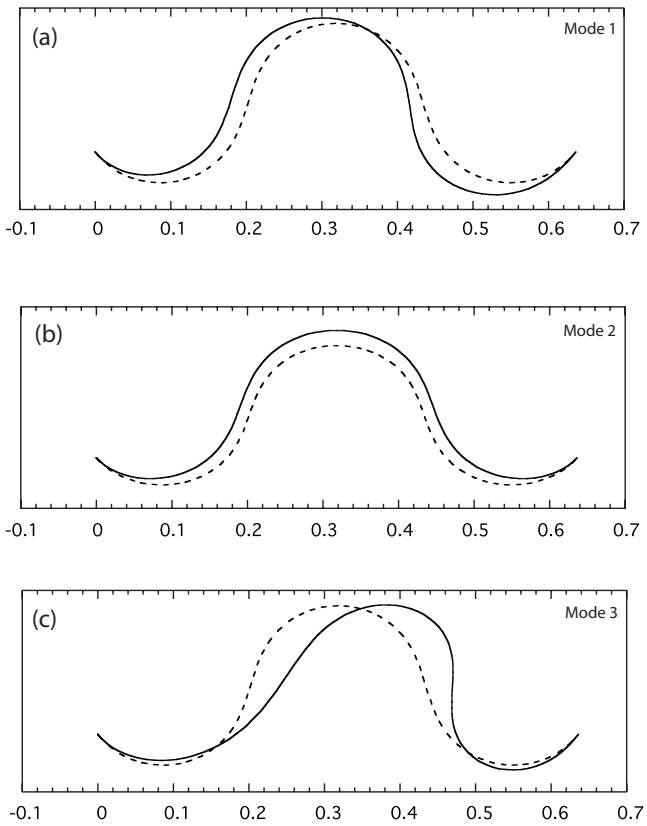


Fig. 8 Mode shapes for α_0=90 deg and ϕ=149 deg

Experiments were conducted with polycarbonate strips. Equilibrium shapes and vibration frequencies were measured. The results agreed well with the analytical results.

In the standard model for shallow curved beams or arches, snap-through does not occur if the height of the structure is below a threshold value [3,4]. In the present problem, it occurs for small heights as well as large heights. The reason is that the strip considered here is assumed to be inextensible. The stable equilibrium states are symmetric, but the strip exhibits an asymmetric shape as it suddenly moves downward into a fully inverted configuration.

Extensibility could be included in the elastica analysis [15]. On the right-hand sides of the first, second, and fourth of Eq. (1), the following factor would be included: $1-[(P\cos\theta+Q\sin\theta)/(EA)]$. In Eq. (3), the factor would be $1-[(p\cos\theta+q\sin\theta)/(\lambda^2)]$ where λ is the slenderness ratio S_0/r with r being the radius of gyration. For the polycarbonate strips, λ is large (e.g., λ=726 when α_0=30 deg) and the effect of extensibility on the results is negligible.

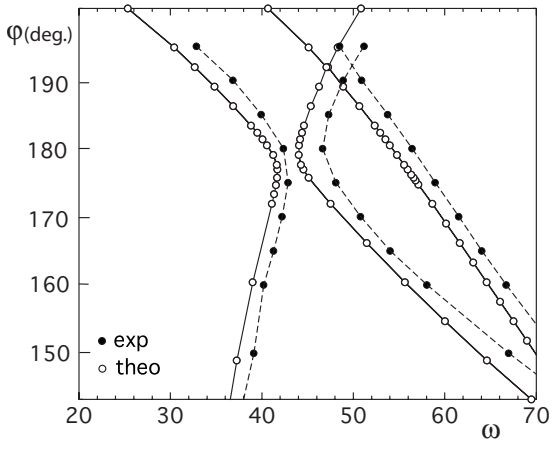


Fig. 9 Expansion of the end rotation as function of lowest frequencies for α_0=120 deg

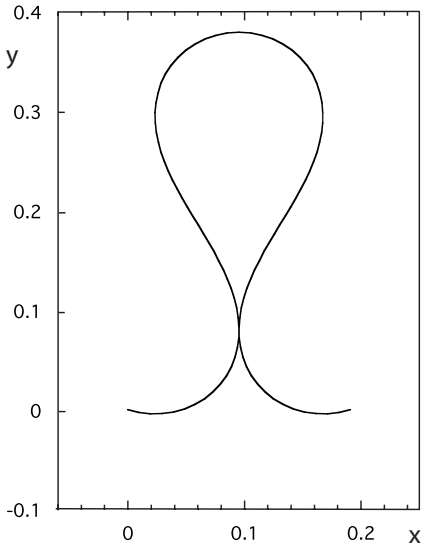


Fig. 10 Self-contact equilibrium shape for $\alpha_0=150$ deg

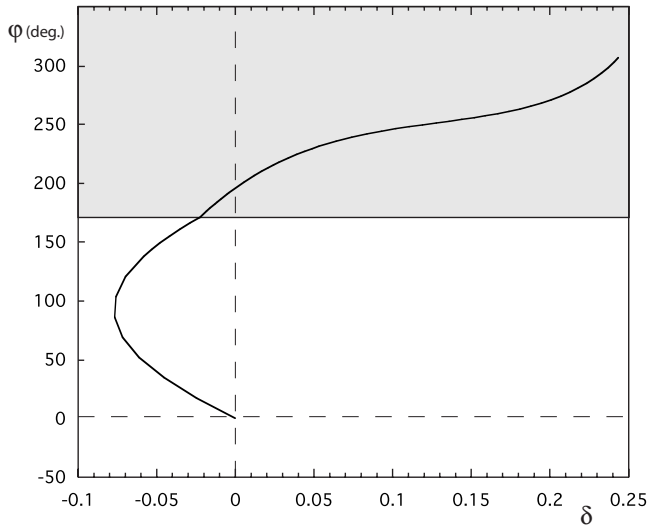


Fig. 11 End rotation as function of midpoint vertical deflection for $\alpha_0=150$ deg. The shaded area for $\phi>171$ deg indicates self-contact.

A related problem with applications in MEMS was also treated analytically. The ends of the strip were pinned and were compressed until the strip buckled upward. Then a downward electrostatic force was applied by a horizontal electrode below the strip. Initial end angles of 15 deg, 30 deg, 45 deg, and 60 deg were considered, and the critical value of the force was found to be almost a linear function of the vertical distance between the electrode and the supports of the strip.

Acknowledgment

The authors are grateful to the reviewers for their helpful suggestions.

Appendix

For the initially circular, inextensible, bent strip (neglecting self-weight and rotary inertia), the internal forces p_e and q_e are zero, and the constant bending moment is $m_e(s)=-2\alpha_0$. The vibration frequencies are the same as those for the corresponding circular, inextensible arch with clamped ends, even though the arch is unstrained and hence has no bending moment. The arch frequencies can be obtained in the following manner [16]. The equation

$$n^6 - 2n^4 + (1-\zeta)n^2 - \zeta = 0 \tag{A1}$$

is solved for the roots n_1^2, n_2^2, and n_3^2 using the analytical solution for a cubic equation. Then the equation

$$n_1(n_2^2-n_3^2)\tan(n_1\alpha_0) + n_2(n_3^2-n_1^2)\tan(n_2\alpha_0) + n_3(n_1^2-n_2^2)\tan(n_3\alpha_0) = 0 \tag{A2}$$

is solved numerically for ζ (e.g., using FINDROOT in MATHEMATICA). The value of ω used here is related to ζ by $\omega=4\alpha_0^2\sqrt{\zeta}$. For $\alpha_0=30$ deg, 60 deg, 90 deg, 120 deg, and 150 deg, respectively, Eqs. (A1) and (A2), as well as the elastica analysis of this paper, yield the lowest frequency to be 58.9327, 51.9693, 43.2726, 34.7846, and 27.5826.

References

[1] Miao, M., Wu, G., Jin, Y., Zhang, J., and Hao, Y., 2005, "A Bulk Micromachined MEMS Digital Distributed Phase Shifter Based on Metal-Air-Metal Capacitors Formed by Intrinsic Stress," Int. J. Nonlinear Sci. Numer. Simul., **6**, pp. 425–430.
[2] Maurini, C., Pouget, J., and Vidoli, S., 2007, "Distributed Piezoelectric Actuation of a Bistable Buckled Beam," Eur. J. Mech. A/Solids, **26**, pp. 837–853.
[3] Plaut, R. H., 1983, "Optimal Arch Form for Stability Under End Moments," *Developments in Mechanics*, *Proceedings of the 18th Midwestern Mechanics Conference*, Iowa City, IA, May 16–18, J. S. Arora, R. L. Benedict, Y. K. Liu, and V. C. Patel, eds., Vol. 12, pp. 87–89.
[4] Chen, J.-S., and Lin, J.-S., 2005, "Exact Critical Loads for a Pinned Half-Sine Arch Under End Couples," ASME J. Appl. Mech., **72**, pp. 147–148.
[5] Plaut, R. H., 2009, "Snap-Through of Shallow Elastic Arches Under End Moments," ASME J. Appl. Mech., **76**, p. 014504.
[6] Plaut, R. H., Taylor, R. P., and Dillard, D. A., 2004, "Postbuckling and Vibration of a Flexible Strip Clamped at Its Ends to a Hinged Substrate," Int. J. Solids Struct., **41**, pp. 859–870.
[7] Perkins, N. C., 1990, "Planar Vibration of an Elastica Arch: Theory and Experiment," ASME J. Vibr. Acoust., **112**, pp. 374–379.
[8] Patricio, P., Adda-Bedia, M., and Ben Amar, M., 1998, "An Elastica Problem: Instabilities of an Elastic Arch," Physica D, **124**, pp. 285–295.
[9] de Silva, C. W., ed., 2005, Vibration and Shock Handbook, Taylor & Francis, Boca Raton, FL.
[10] Santillan, S. T., Virgin, L. N., and Plaut, R. H., 2006, "Post-Buckling and Vibration of Heavy Beam on Horizontal or Inclined Rigid Foundation," ASME J. Appl. Mech., **73**, pp. 664–671.
[11] Wolfram, S., 1996, *The Mathematica Book*, 3rd ed., Cambridge University Press, Cambridge.
[12] Wang, C.-Y., 1981, "Folding of Elastica—Similarity Solutions," ASME J. Appl. Mech., **48**, pp. 199–200.
[13] Zhang, Y., Wang, Y., Li, Z., Huang, Y., and Li, D., 2007, "Snap-Through and Pull-In Instabilities of an Arch-Shaped Beam Under an Electrostatic Loading," J. Microelectromech. Syst., **16**, pp. 684–693.
[14] Krylov, S., Ilic, B. R., Schreiber, D., Seretensky, S., and Craighead, H., 2008, "The Pull-In Behavior of Electrostatically Actuated Bistable Microstructures," J. Micromech. Microeng., **18**, p. 055026.
[15] Atanackovic, T. M., 1997, *Stability Theory of Elastic Rods*, World Scientific, Singapore.
[16] Wasserman, Y., 1977, "The Influence of the Behaviour of the Load on the Frequencies and Critical Loads of Arches With Flexibly Supported Ends," J. Sound Vib., **54**, pp. 515–526.

The Energy Absorption Characteristics of Square Mild Steel Tubes With Multiple Induced Circular Hole Discontinuities—Part I: Experiments

S. B. Bodlani

S. Chung Kim Yuen[1]
e-mail: steeve.chungkimyuen@uct.ac.za

G. N. Nurick

Blast Impact and Survivability Research Unit (BISRU),
Department of Mechanical Engineering,
University of Cape Town,
Private Bag,
Rondebosch, 7700, South Africa

This two-part article reports the results of experimental and numerical works conducted on the energy absorption characteristics of thin-walled square tubes with multiple circular hole discontinuities. Part I presents the experimental tests in which dynamic and quasistatic axial crushings are performed. The mild steel tubes are 350 mm in length, 50 mm wide, and 1.5 mm thick. Circular hole discontinuities, 17 mm in diameter, are laterally drilled on two or all four opposing walls of the tube to form opposing hole pairs. The total number of holes varies from 2 to 10. The results indicate that the introduction of holes decreases the initial peak force but an increase in the number of holes beyond 2 holes per side does not further significantly decrease the initial peak force. The findings show that strategic positioning of holes triggers progressive collapse hence improving energy absorption. The results also indicate that the presence of holes may at times disrupt the formation of lobes thus compromising the energy absorption capacity of the tube. In Part II, the finite element package ABAQUS/EXPLICIT version 6.4–6 is used to model the dynamic axial crushing of the tubes and to investigate the action of the holes during dynamic loading at an impact velocity of 8 m/s. [DOI: 10.1115/1.3114971]

Keywords: square tubes, imperfections, multiple hole cut-outs, energy absorption

1 Introduction

High impact events such as transport-related accidents can result in life-changing injuries or loss of life to vehicle occupants. Attempts to minimize this damage primarily require the absorption of the kinetic energy of the crash in an effort to attain what is termed good *vehicle crashworthiness* [1–3]. Improving vehicle crashworthiness is of specific importance in the light of an ever-increasing number of vehicles on the road and the subsequent increase in vehicular accident deaths. According to the 2004 World Health Organization report [4], on average, 3000 people are killed daily in road accidents. Such statistical figures have ensured that the subject of vehicle crashworthiness continues to draw the attention of academic and industrial researchers worldwide [5–11].

In order to achieve good vehicle crashworthiness the common agreement among researchers [1,2,5–15] is to have an energy absorption device that achieves one or both of the following: decrease the initial peak force and/or absorb as much energy as possible. In fulfilling these criteria the device will enable a reduction in the acceleration perceived by vehicle occupants and the amount of energy transferred from the vehicle to its occupants in a crash. Crush force efficiency (CFE) and stroke energy (SE), given in Eqs. (1) and (2), are typical parameters used to assess the performance of these energy absorbing devices in achieving these criteria.

$$\mathrm{CFE} = \frac{P_m}{P_i} \tag{1}$$

$$\mathrm{SE} = \frac{\delta}{L} \tag{2}$$

The ideal value for CFE is unity [5]. The CFE parameter is useful in determining the extent to which a constant deceleration is achieved, thus a value of unity would indicate that the device has reduced the initial peak force (P_i), also referred to as ultimate peak force, sufficiently to enable a smooth deceleration during impact. The ideal value that would indicate maximum energy absorption for SE is 100%. Due to the compaction of the tubes however, this is not physically possible. Abramowicz and Jones [6] found that the maximum attainable value of SE for square tubes is 73%.

Various studies [2,3,5–11] have been conducted on the axial compression of thin-walled tubes and have enabled the characterization of the manner in which they fail, namely, Euler and progressive bucklings. Euler buckling is characterized by a small deformation after loading begins followed by a very large global bending such that the tube can no longer sustain further axial loading [5,7]. Progressive buckling is associated with an initial peak force followed by a lower periodic force with relatively constant amplitude for the remainder of the load application event [5,7]. Progressive buckling proceeds through the development of folds or lobes resulting in extensive plastic strain. This strain is the mechanism through which progressive buckling absorbs significantly more energy than Euler buckling [2,5,8]. Furthermore, the incidence of either Euler or progressive mode is governed by the tube geometry of the tube. Abramowicz and Jones [9] found

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANISMS. Manuscript received October 2, 2007; final manuscript received November 2, 2007; published online April 27, 2009. Review conducted by Robert M. McMeeking.

that the length to width ratio (L/C) at which progressive collapse transitioned to Euler increased with an increase in the width to thickness ratio (C/t).

This ability to effectively dissipate energy makes thin-walled tubes an ideal manner of improving vehicle crashworthiness [1,2,5–9]. Consequently they are widely used in the construction of the mainframes of cars [2,3]. These thin-walled tubes are joined together to form a hollow frame commonly known as the *space-frame*. Various studies [1–3,5–18] have investigated the manner through which thin-walled tubes dissipate energy and the optimization of that energy absorption. This allows thin-walled tubes to be applied in a manner that fully exploits their energy absorbing capabilities [5].

One of the ways to improve energy absorption is to insert imperfections, also known as triggers or initiators or discontinuities, into the tube wall. These discontinuities can be geometrical in form such as dents and through-hole circular, elliptical, or slotted cut-outs or material changes. Chung Kim Yuen and Nurick [11] provided an extensive review of the application of induced geometric and material modifications to the improvement of energy absorption in thin-walled structures. It was noted that these efforts to improve energy absorption must ensure that tubes are able to deform in a predictable manner when their energy absorbing qualities are required. More particularly, these induced imperfections can decrease the initial peak force and control the particular mode of buckling induced in the tube. Studies [5,7,13,14] investigating the effect of a single discontinuity pair inserted at mid-height have concluded that the introduction of discontinuities into the tube wall does decrease the initial peak force. Furthermore, these discontinuities "trigger" the collapse of the tube. This is as discussed by Abramowicz [2] who stressed the need for the relevant absorber to be "triggered" in such a way that the formation of natural folds is promoted. The formation of natural folds encourages progressive buckling and energy absorption is maximized.

The prevailing conclusion among Refs. [5,7,13,14] is that for circular discontinuities an increase in hole diameter results in an increase in the reduction in the initial peak force. There is however a limit beyond which increasing the discontinuity severity will have an adverse effect on energy absorption. For example, Arnold and Altenhof [7] and Marshall and Nurick [13] found that a discontinuity of hole diameter larger than 32 mm in 50 mm wide square tubes caused splitting.

Gupta and Gupta [14] investigated the effect of multiple circular holes on the energy absorption characteristics of circular tubes. The specimens used contained two or four opposing circular hole discontinuities laterally drilled on one, two, or three planes spaced symmetrically along the length of the tube. The results indicate that the introduction of more than two opposing holes did not decrease the initial peak force significantly further than that achieved by just two opposing holes. Furthermore, Gupta and Gupta [14] reported that buckling was triggered at the location of the holes in one of the planes and the collapse mode varied depending on the number of planes.

Montanini et al. [10] noted that both the size and the positioning of the trigger need to be taken into account for effective crush initiation. Lee et al. [15] furthermore found that if induced discontinuities were not located at the natural hinges of lobes then they would cause irrevocable disturbances to the collapse mode that developed and lead to global bending even if the geometry of the tube was such that it would have otherwise buckled in progressive mode. Consequently, a strategic placement of holes in the zones where the lobe hinges are expected to fall when the tube undergoes progressive buckling is required. Montanini et al. [10] suggested the application of theoretical findings to calculate the positioning of the triggers. Abramowicz and Jones [6,16] proposed analytical predictions for the wavelength of a single lobe in the progressive buckling of square tubes, in that the original (before lobe formation) height of the lobe is given as

$$l/t = 0.99(C/t)^{2/3} \tag{3}$$

The work described in this paper extends the knowledge described above by investigating the effect of multiple circular discontinuities as crush initiators in the collapse mode progression. Since the aim is that the discontinuities act as points of crush initiation at which lobes are induced, the spacing of successive hole pairs relative to one another is considered carefully. Furthermore, the effect of placing holes on both or all four tube walls is also reported. For the quasistatic tests the manner by which collapse is triggered and progresses during axial loading is closely examined. The nature of the dynamic tests precludes similar observations and consequently a validated numerical model is presented in Part II [17] to observe the role of the discontinuities in triggering collapse during dynamic axial loading.

2 Experimental Program

2.1 Test Specimens. Mild steel industrial seam weld tubes with nominal width (C) 50 mm, thickness (t) 1.5 mm, and length (L) 350 mm are used. These dimensions give geometrical ratios of L/C and C/t corresponding to progressive buckling ($L/C<10$ for $5.5<C/t<38$) as described by Abromowicz and Jones [6,16]. Using Eq. (3), the wavelength ($2l$) for the tubes in this paper should be 32 mm. However, on the measurement of the actual lobe wavelength of as-received tube (plain with no discontinuities) of the above dimensions, the theoretical prediction appears to underestimate the actual lobe wavelength, which is approximately 50 mm. The objective is to place discontinuities at regular intervals of 50 mm in order to attempt to trigger lobe formation where the lobe would naturally form in the absence of holes. Opposing hole pairs of diameter 17 mm are laterally drilled midwidth of the walls of the tube. The specimen configurations are illustrated in Fig. 1. Specimen group A has holes on two tube walls and specimen groups B_H50 and B_HEq have holes on all four tube walls. In group A, illustrated in Fig. 1(a), the hole are spaced 50 mm apart where there are multiple holes per sides. In group B_H50, illustrated in Fig. 1(b), the holes are also 50 mm apart on each tube wall and adjacent hole pairs are placed on alternating walls. Group B_HEq, illustrated in Fig. 1(c), contains hole pairs spaced symmetrically along the length of the tube with adjacent hole pairs being on alternating tube walls. A number is appended after the letters "A" or "B" in each group name to indicate the number of holes present in the particular specimen.

2.2 Quasistatic Axial Crush Testing. Quasistatic axial crushing is performed in a Zwick hydraulic load cell with a maximum capacity of 200 kN. The bottom 50 mm of the tube is fixed by a clamp during testing. All tubes are crushed at a cross head speed of 30 mm/min (strain rate 1.7×10^{-3} s^{-1}) for a distance of 220 mm to correspond to the maximum tube compaction of 73% [6]. Data acquisition software, connected to the load cell and a computer, records the force-displacement history of the crushing event. The energy absorbed (E_a^s) and mean crushing force (P_m) are calculated using Eqs. (4) and (5). Equation (4) applies the trapezium rule for integrating the area under the force-displacement graph and represents the energy absorbed during the entire quasistatic axial crushing event.

$$E_a^s = \sum_{j=2}^{N} 0.5(P_i + P_{i-1})(\delta_i - \delta_{i-1}) \tag{4}$$

$$P_m = \frac{E_a^s}{\delta} \tag{5}$$

2.3 Dynamic Axial Crush Testing. The dynamic crush tests are performed in a drop hammer rig. The specimen is clamped for a length of 50 mm at its base and placed on the anvil of the rig. The mass is raised to the desired height and released to impact the

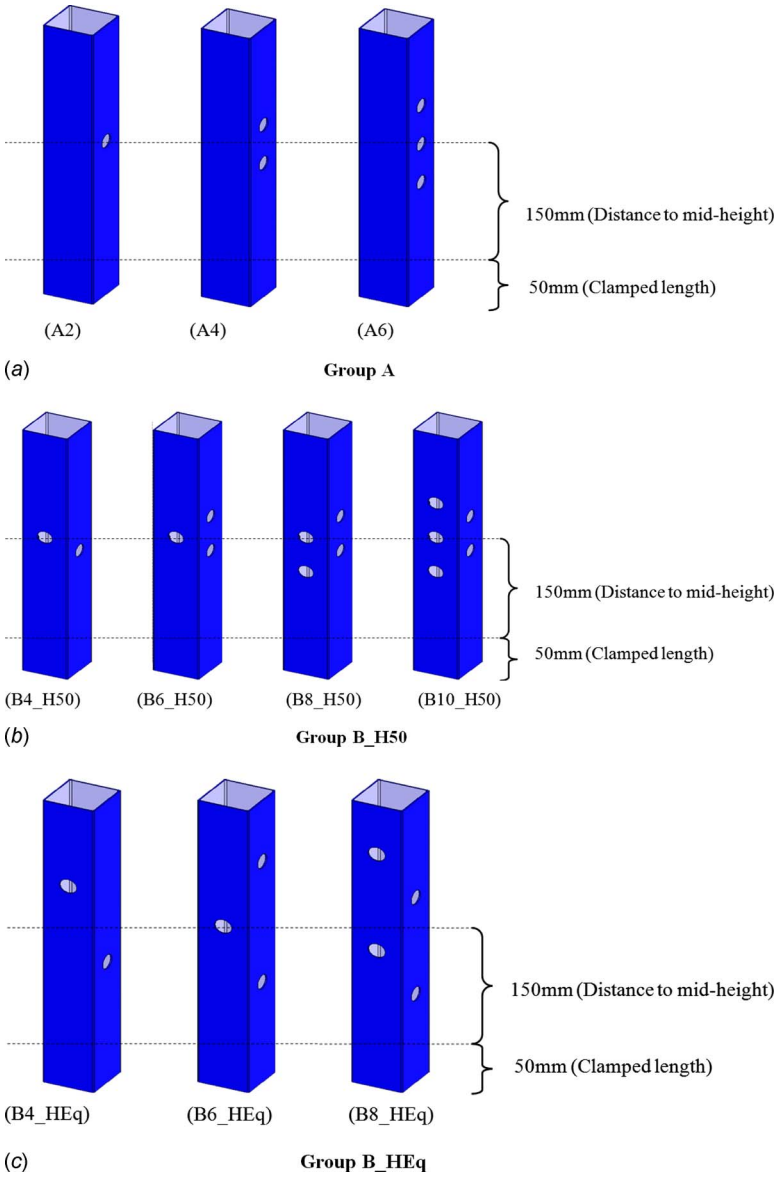


Fig. 1 Test specimen configurations

tube in an axial direction. The energy absorbed (E_a^d) during the dynamic loading event is calculated using Eq. (6), assuming no losses to friction, in which the crushed distance (δ) is measured for each specimen.

$$E_a^d = Mg(h + \delta) \tag{6}$$

Replacing energy absorbed in the quasistatic tests (E_a^s) in Eq. (5) with E_a^d yields the mean force for the dynamic tests. The drop hammer rig used is not equipped with data acquisition equipment and hence force-displacement histories are not recorded. A finite element model is validated and used to analyze the force-displacement histories of the dynamic tests and is reported in Ref. [17].

3 Experimental Results and Observations

3.1 Quasistatic Tests. Table 1 presents the results of the quasistatic axial crush tests. The letter "S" is prefixed to the specimen name to indicate the quasistatic test and a number is suffixed to distinguish between different specimens of the same configuration. Furthermore, specimens without holes are referred to as "plain_tube" and where the diameter of discontinuities present in the specimen differs from the standard 17 mm, the diameter is included in the specimen name preceded by the letter "d." Two failure modes are observed in the quasistatic tests, namely, progressive collapse (PS) and mixed mode (MM), and are illustrated in Fig. 2. Mixed mode refers to a collapse mode that is a combination of progressive collapse and Euler buckling. The test is stopped if MM collapse mode develops.

3.1.1 Specimen Group A

3.1.1.1 Configuration A2. As expected, the introduction of an opposing hole discontinuity pair at midheight causes a reduction in initial peak force. The initial peak force for specimen SA2_1 is 10% lower than that for the specimens without holes (splain_tube_1 and splain_tube_2). Furthermore a specimen of diameter 17 mm (SA2_1) results in a lower initial peak force than one of diameter 12.5 mm (SA2_d12.5_1). This result is concurrent with literature [5,14] in which the understanding is that the larger the hole diameter, the lower the initial peak force. This reduction in force is due to the localization of deformation at the hole discontinuity. The discontinuities act as stress raisers within the area in

Table 1 Results for quasistatic axial crushing of square tubes with and without hole discontinuities

Specimen name	P_i (kN)	P_m (kN)	E_a (kJ)	CFE	SE (%)	Failure mode
splain_tube_1	101.7	32.0	7.0	0.32	73.3	PS
splain_tube_2	101.6	31.4	6.9	0.31	73.3	PS
SA2_d12.5_1	94.8	32.7	7.1	0.34	73.3	PS
SA2_1	91.3	33.4	7.3	0.37	73.3	PS
SA4_1	86.9	31.9	3.4	0.37	35.5	MM
SA4_2	90.2	34.6	3.7	0.38	35.8	MM
SA6_1	89.4	33.4	7.3	0.37	73.3	PS
SA6_2	91.0	35.1	3.8	0.39	35.6	MM
SB4_H50_1	84.0	32.2	7.0	0.38	73.33	PS
SB6_H50_1	81.6	33.3	7.3	0.41	73.33	PS
SB8_H50_1	80.5	32.1	7.0	0.4	73.33	PS
SB10_H50_1	79.6	33.6	3.5	0.42	34.6	MM
SB10_H50_2	79.9	32.6	7.2	0.41	73.3	PS
SB4_HEq_1	86.6	29.9	6.6	0.35	73.3	PS
SB6_HEq_1	89.3	31.62	7.0	0.35	73.3	PS
SB8_HEq_1	86.9	34.02	7.5	0.39	73.3	PS

which they are placed allowing the area to fail at a comparatively lower applied force than it would in the absence of discontinuities [5]. Photographs taken during the crushing event are given in Fig. 3(*a*) for specimen SA2_d12.5_1. The letters below each photograph correspond to those in Fig. 3(*b*), which illustrates the corresponding force-displacement curve together with that of a plain tube (splain_tube_2). Specimen SA2_1 displays similar crushing behavior to SA2_d12.5_1 but with a larger reduction in initial peak force. The crushing event for specimen SA2_1 is illustrated in Fig. 4.

Although specimen SA2_1 collapsed in the PS mode, a skew initial lobe is observed in point (b) of Fig. 4. This is characterized by an upturning of the outer hinge of the lobe such that it contacts the tube wall above it and interferes with the subsequent lobe formation adjacent to it. This is visible at points (c), (d), and (e) of Fig. 4. The incidence of a skew lobe for the initial lobe appears to be a random occurrence since it is observed for some initial lobes but not all.

3.1.1.2 Configuration A4 and A6. Increasing the number of holes from two to six holes affords a slight decrease in the initial peak force. The average initial peak force for the specimens of four and six holes is 89.34 kN. Increasing the number of holes does however produce a change in the collapse mode from progressive collapse in specimens of configuration A2 to mixed mode in A4 and A6 specimens. The crushing event for SA4_1 is illustrated in Fig. 5. It is illustrated that one pair of opposing holes in SA4_1 act as crush initiators. After the initial lobe is formed, a similar incidence of skew lobe formation as observed in specimen SA2_1 (Fig. 4) is observed at point (c) of Fig. 5. This skew lobe results in a destabilization of the subsequent lobes forming above it and the collapse mode of the tube is compromised as illustrated by points (d)–(f) in Fig. 5. The centerline of the tube shifts sideways resulting in an offset in the axis through which the load is being applied as seen at points (e) and (f) at which point the tube can no longer sustain further axial loading. This collapse mode is defined as MM since progressive collapse is first observed followed by the global bending of the tube. MM collapse is also observed in specimens SA4_2 and SA6_2. Specimens SA4_1 and SA4_2 are illustrated in Fig. 6.

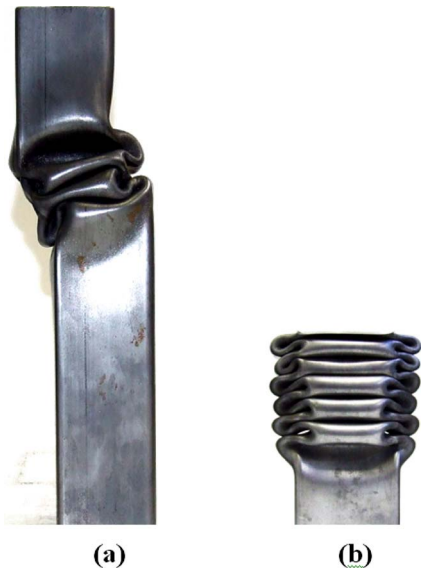


Fig. 2 Photographs of quasistatically crushed tubes, which collapsed in (*a*) MM and (*b*) PS modes

3.1.2 Specimen Group B_H50 and B_HEq

3.1.2.1 Group B_H50. The average initial peak force for all group B_H50 specimens is 81 kN. This represents a 20% reduction in the initial peak force of a plain tube. Furthermore, in Table 1 it is observed that there is a slight decrease in initial peak force as the number of holes is increased from four to ten. The crushing event for specimen SB6_H50_1 is illustrated in Fig. 7. The progressive collapse in specimens of group B_H50 is triggered by two pairs of opposing holes simultaneously. The arrows at point (b) of Fig. 7 indicate the triggers that concurrently initiate the first lobe. Subsequent collapse continues progressively first below and then above the initial lobe. Most specimens in group B_H50 are subject to a skew initial lobe but the lobes forming subsequently are able to straighten the initial lobe and progressive collapse continues undisturbed, except in specimen SB10_1 whose crushing event is illustrated in Fig. 8.

To illustrate the lobe initiation role that hole discontinuities play in the progressive collapse of tubes, the crushing process for specimen SB10_H50_1, which failed in MM collapse, is given by the photographs in Fig. 8. As noted for group B_H50 specimens, two sets of opposing holes simultaneously trigger the initiation of the first lobe visible at point (b) of Fig. 8. The skew lobe formation that occurs in specimen SB10_H50_1 is more severe than that observed in the other group B_H50 specimens. The initial lobe rotates to contact the tube wall as indicated at point (d) of Fig. 8. The hole discontinuity marked with an arrow displaces the action

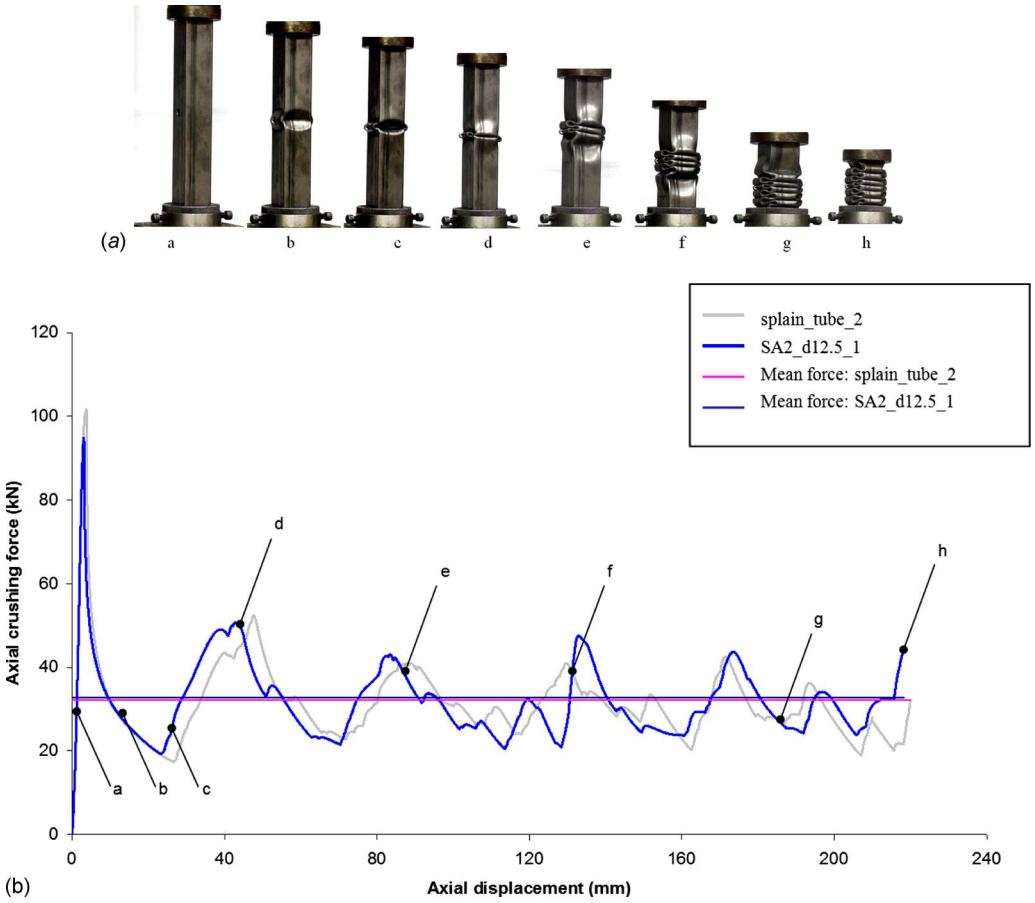


Fig. 3 (*a*) **Photographs of the quasistatic crushing of specimen SA2_d12.5_1;** (*b*) **force-displacement curve for the quasistatic crushing of specimens SA2_d12.5_1 and splain_tube-2**

of the skew lobe and enforces the initiation of another lobe, which subsequently prevents the specimen from bending further to the right as illustrated by points (d)–(f). It appears that the impaction of the skew lobe on the tube wall is severe enough to cause the tube to bend in the opposite direction as shown in point (g) resulting in MM collapse, even after the hole discontinuity straightens the tube in point (e). Figures 7 and 8 illustrate that the hole discontinuities fall at the hinges of the lobes for group B_H50.

3.1.2.2 Group B_HEq. The initial peak force for all group B_HEq specimens averages 88 kN, a 13% reduction from that of a plain tube, with no significant change with increasing number of holes. All specimens collapsed in the PS mode. In specimen SB4_HEq_1, crushing commences at the top pair of opposing holes, progressing in a steady manner above the initial lobe. Crushing continues below the initial lobe down the length of the tube in a steady manner as for a tube without hole discontinuities. The same crush behavior was observed for specimen SB6_d17_HEq_1. Some hole discontinuities fall at the lobe hinges and some do not for both specimens. Photographs of the final crushed specimens are given in Fig. 9 for specimens SB4_HEq_1 and SB6_HEq_1. The force-displacement curve and the crushing process for specimen SB8_HEq_1 is given in Figs. 10 and 11, respectively. The buckling of specimen SB8_HEq_1 initiates at one of the hole discontinuity pairs as shown by point (b) of Fig. 11. Progressive buckling occurs in an upwards direction. Once crushing finishes above the initial lobe, another lobe is triggered at a hole discontinuity located further down the tube. This leads to the formation of a relatively large lobe as shown in

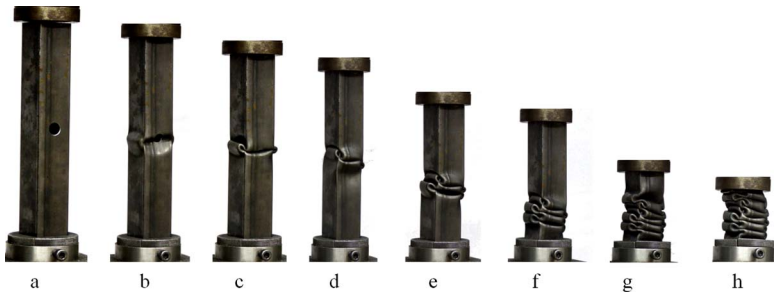


Fig. 4 Photographs of the transient quasistatic crushing of specimen SA2_1

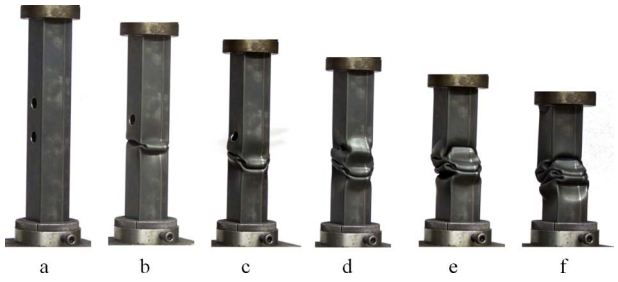


Fig. 5 Photographs of the transient quasistatic crushing of specimen SA4_1

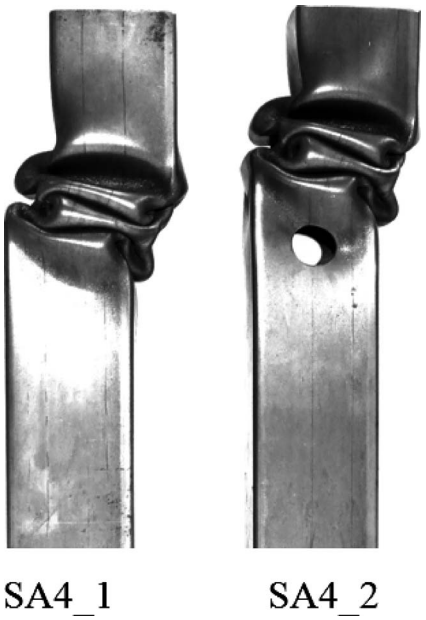


Fig. 6 Photographs of specimens SA4_1 and SA4–2 at the end of quasistatic crushing

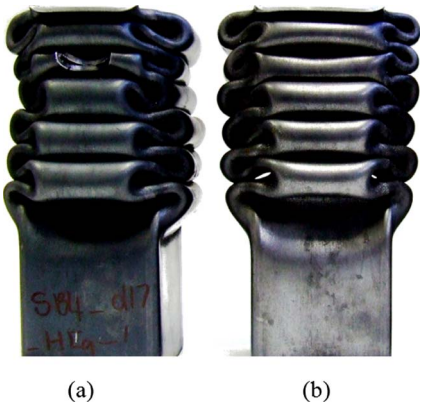


Fig. 9 Photographs of specimens (*a*) SB4_HEq_1 and (*b*) SB6_HEq_1 at the end of quasistatic crushing

points (g)–(i). The positioning of the hole pairs at different distances relative to one another thus imposes a certain lobe wavelength on the tube.

3.2 Dynamic Tests. The results of all specimens tested dynamically are presented in Table 2. The letter "D" is prefixed to the specimen name to indicate the dynamic test and, as before, a number is suffixed to distinguish between different specimens. All specimens failed in PS mode except the three marked with an asterisk (*) in Table 2, which failed in MM. Velocity is calculated by Eq. (7), which assumes zero friction losses and applies conservation of energy.

$$V = \sqrt{2g(h + \delta)} \tag{7}$$

3.2.1 Specimen group A

3.2.1.1 Configuration A4. The specimens of configuration A4 tested at different heights are shown in Fig. 12. Expectedly, an increase in height and hence impact velocity results in an increase in stroke efficiency, which is visible by the increase in crushed distance in Fig. 12. The collapse mode appears to develop in a

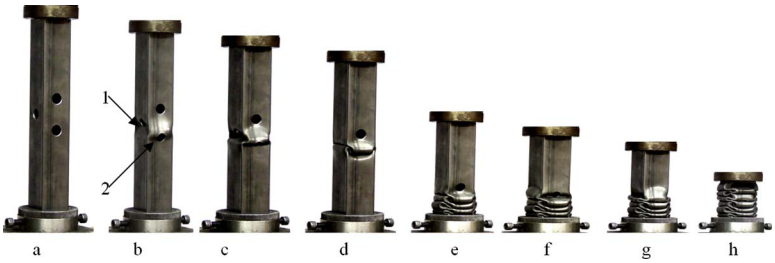


Fig. 7 Photographs of the transient quasistatic crushing of specimen SB6_H50_1

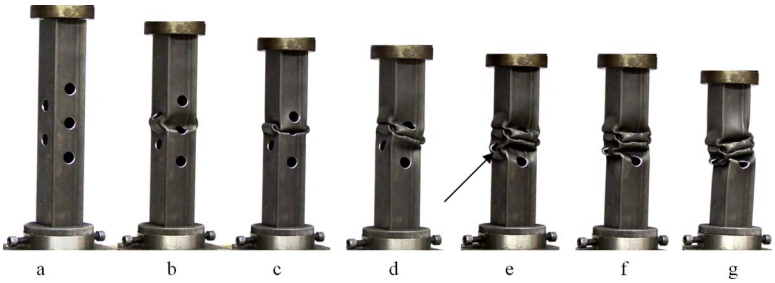


Fig. 8 Photographs of the transient quasistatic crushing of specimen SB10_H50_1

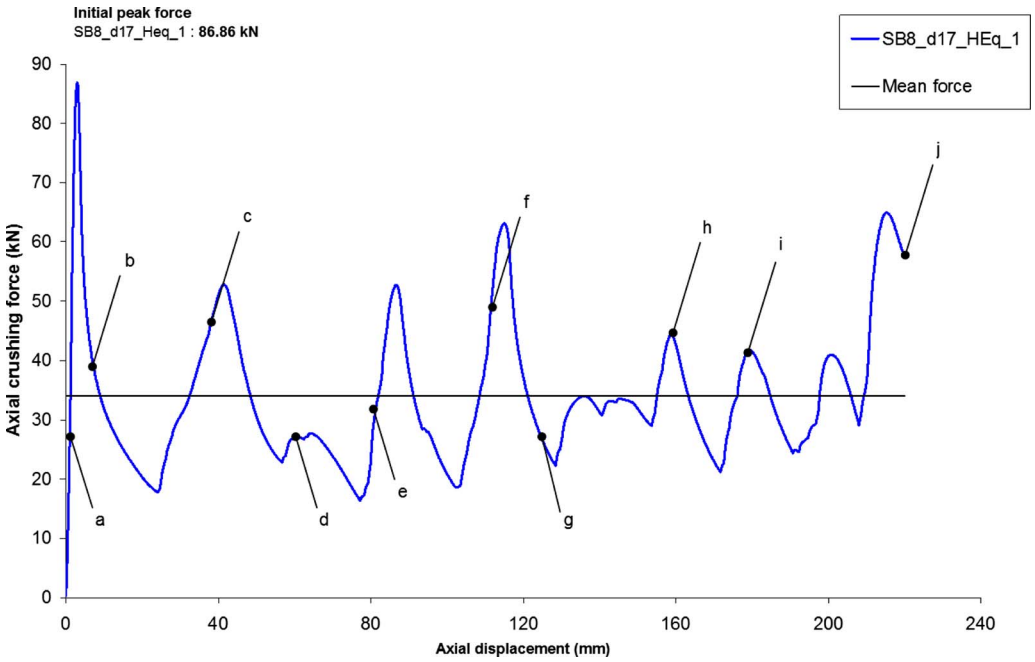


Fig. 10 Force-displacement curve for the quasistatic crushing of specimen SB8_HEq_1

steadier manner with increasing height as illustrated by the lack of visible disturbances in the formation of lobes for specimen DA4_3 in photograph (c) of Fig. 12 as opposed to the other two specimens (photographs (a) and (b)).

3.2.1.2 A6. Specimens DA6_1, DA6_2, and DA6_3 are illustrated in Fig. 13. The mean force is slightly higher for specimens with six holes (Fig. 13) than for those with four holes (Fig. 12), while those with four holes have a slightly better stroke efficiency resulting in an overall equivalent amount of energy absorption. For a height of 3.3 m, specimen DA4_1 has a stroke efficiency of 64.4%, while DA6_1 has a stroke efficiency of 61.6%; the energy absorbed is 11.2 kJ and 11.1 kJ, respectively.

3.2.2 Specimen Group B_H50 and B_HEq

3.2.2.1 Group B_H50. All group B_H50 specimens tested dynamically buckled in the PS mode. Specimens tested at heights of 2.5 m and 3.3 m, respectively, are given in Fig. 14. The figure illustrates that specimens of group B_H50 are characterized by a steady and repeatable mode of collapse. Figure 14 also illustrates that there is a random cluster of lobes formed at either the impact end, nonimpact end, or the middle of the tube. No significant changes in mean force, stroke efficiency, or energy absorption are notable for an increase in the number of holes within the specimens tested at the same height. The average values for mean force, stroke efficiency, and energy absorbed for specimens tested at 4m are 57.4 kN, 79.8%, and 13.7 kJ, respectively. These values are similar to those for the plain tube, d_plain_tube_2, tested at the same height; a mean force of 57.5 kN, a stroke efficiency of 79.6%, and an energy absorption of 13.7 kJ. This indicates that the addition of holes has no quantitative effect on the energy absorption parameters of a plain tube. The absence of collapse in Euler bending or MM in this group indicates that the addition of holes in the specified configuration ensures the development of progressive collapse. A definite occurrence of progressive collapse is a desirable feature in energy absorbing devices in which progressive collapse is required to provide the necessary energy absorbing attributes. Figure 15 further illustrates the repeatability achieved by specimen group B_H50 where configuration B6_H50 specimens both tested at 3.3 m produce similar collapse profiles.

Fig. 11 Photographs of the transient quasistatic crushing of specimen SB8_HEq_1

3.2.2.2 Group B_HEq. Specimen group B_HEq is characterized by the formation of lobes of different sizes within the same specimen. Though the overall collapse mode remains progressive, the change in the lobe size compromises the overall stability of the collapse mode in some specimens. Figures 16–18 illustrate photographs of the specimens of group B_HEq after dynamic axial crushing at the indicated heights.

The irregular lobes formed in group B_HEq specimens are imposed by the distances between successive discontinuities and have wavelengths larger than that would develop in a plain tube. The result is an uneven distribution of lobes such that these irregular lobes destabilize the collapse causing the tube to tilt to one side as viewed in photograph (b) of Figs. 16 and 17, respectively. The collapse mode is compromised because these irregular lobes result in the tube leaning over precariously as if about to collapse in mixed mode. It is noted earlier for the quasistatic tests that specimen configuration B6_HEq is characterized by regular sized lobes similar to those present in a plain tube. For the dynamic tests, the same observation is made for specimen DB6_HEq_3,

Table 2 Test results for square tubes subjected to dynamic axial loading

Specimen name	h (m)	V (m/s)	δ (mm)	E_a (kJ)	P_m (kN)	SE (%)
d_plain_tube_1	3.3	8.0	210.0	11.2	53.5	70.0
d_plain_tube_2	4.0	8.9	238.7	13.7	57.5	79.6
DA2_1	3.3	8.0	172.0	11.1	64.6	57.3
DA2_2	3.3	8.0	176.1	11.1	63.1	58.7
DA2_3	3.3	8.0	209.3	11.2	53.7	69.8
DA4_1	3.3	8.0	193.3	11.2	57.9	64.4
DA4_2*	3.3	8.0	-	-	-	-
DA4_3	4.0	8.9	241.8	13.8	56.9	80.6
DA4_4	2.5	7.0	155.8	8.6	55.2	52.0
DA6_1	3.3	8.0	184.6	11.1	60.3	61.6
DA6_2	4.0	8.9	235.5	13.7	58.3	78.5
DA6_3*	2.5	7.0	-	-	-	-
DB4_H50_1	3.3	8.0	193.6	11.2	57.6	64.6
DB4_H50_2	3.3	8.0	199.4	11.2	56.2	66.5
DB4_H50_3	4.0	8.9	240.8	13.8	57.1	80.3
DB4_50_4	2.5	7.0	154.5	8.6	55.6	51.5
DB6_H50_1	3.3	8.0	190.5	11.2	58.5	63.5
DB6_H50_2	3.3	8.0	200.1	11.2	56.1	66.7
DB6_H50_3	4.0	8.9	237.4	13.7	57.8	79.1
DB6_H50_4	2.5	7.0	167.5	8.6	51.5	55.8
DB8_H50_1	3.3	8.0	201.4	11.2	55.6	67.1
DB8_H50_2	4.0	8.9	239.8	13.7	57.3	79.9
DB8_H50_3	2.5	7.0	167.2	8.7	51.7	55.7
DB10_H50_1	3.3	8.0	192.1	11.2	58.2	64.0
DB10_H50_2	4.0	8.9	239.1	13.7	57.4	79.7
DB10_H50_3	2.5	7.0	158.6	8.6	54.4	52.9
DB4_Heq_1	3.3	8.0	184.7	11.2	60.4	61.6
DB4_Heq_2	2.5	7.0	180.6	8.7	48.0	60.2
DB4_Heq_3*	4.0	8.9	-	-	-	-
DB6_Heq_1	3.3	8.0	179.5	11.1	62.1	59.8
DB6_Heq_2	2.5	7.0	181.9	8.7	47.7	60.6
DB6_Heq_3	4.0	8.9	238.4	13.7	57.5	79.5
DB8_Heq_1	3.3	8.0	216.0	11.3	52.1	72.0
DB8_Heq_2	2.5	7.0	177.9	8.7	48.7	59.3
DB8_Heq_3	3.3	8.0	203.4	11.2	55.2	67.8

illustrated in Fig. 19. The mean force, energy absorbed and stroke efficiency for specimen DB6_HEq_3 are 57.5 kN, 13.7 kJ, and 79.5%, respectively. These values are similar to those of d_plain-_tube_2 also tested at 3.3 m, which are 57.5 kN, 13.7 kJ, and 79.6%, respectively.

4 Discussion of Experimental Findings

4.1 Initial Peak Force. The reduction in initial peak force relative to a plain tube is illustrated in Fig. 20 for each specimen configuration. Figure 20 illustrates that group B_H50 results in the highest reduction in initial peak force. The reduction in initial peak force for group B_HEq is generally similar to that of group A. This implies that not only is having holes on all four sides of the tube better for initial peak force reduction but also fixing hole spacing appropriately as group B_H50 is better than the symmetrical spacing of group B_HEq. This confirms that having the discontinuities at the estimated tube wavelength improves crashworthiness as suggested by Montanini et al. [10] and Lee et al. [15].

4.2 Collapse Mode. Group B_H50 has the most stable collapse mode, which was not prone to mixed mode collapse. Specimen group A is the most unstable. Of the 11 group A specimens tested dynamically and quasistatically, 5 collapsed in mixed mode,

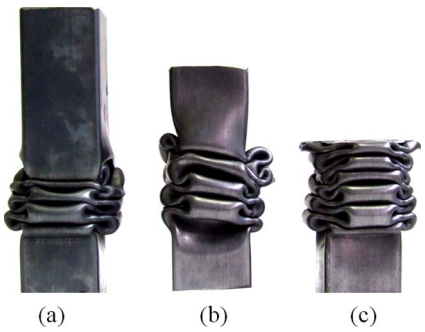


Fig. 12 Photographs of specimens (*a*) DA4_4, (*b*) DA4–1, and (*c*) DA4_3 after loading at heights of 2.5 m, 3.3 m, and 4 m, respectively

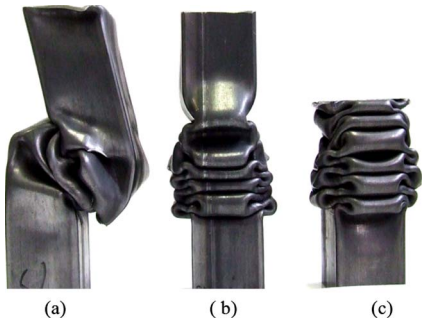


Fig. 13 Photographs of specimens (*a*) DA6_3, (*b*) DA6_1, and (*c*) DA6_2 after load at heights of 2.5 m, 3.3 m, and 4 m, respectively

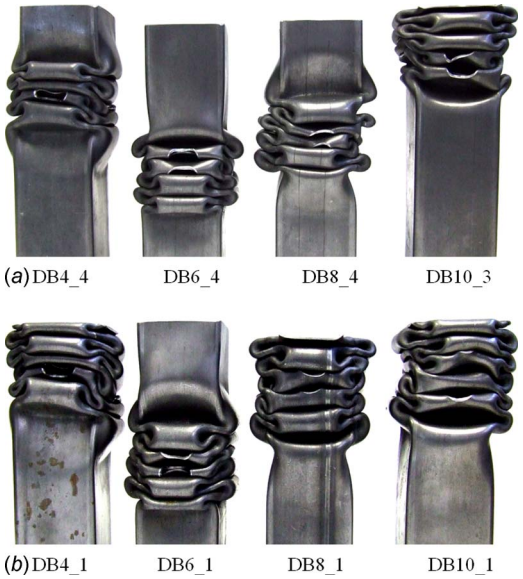


Fig. 14 Photographs of specimens of group B-H50 after loading at heights of (*a*) 2.5 m and (*b*) 3.3 m

3 of which were tested dynamically. Both groups B_H50 and B_HEq exhibit mostly progressive collapse mode but the lobes were of significantly different shapes. Group B_H50 produces lobes of regular sizes much like those of a plain tube. In group B_HEq however, the lobes induced are of a nonuniform size.

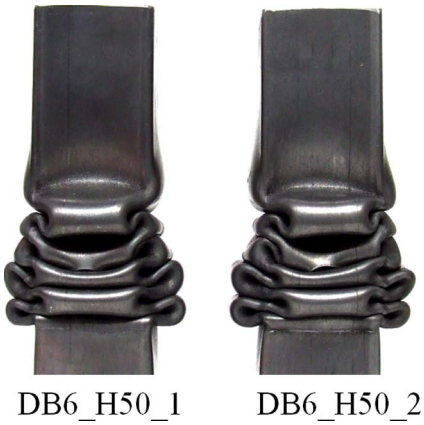


Fig. 15 Photographs of specimens of configuration B6_H50 after loading at a height of 3.3 m

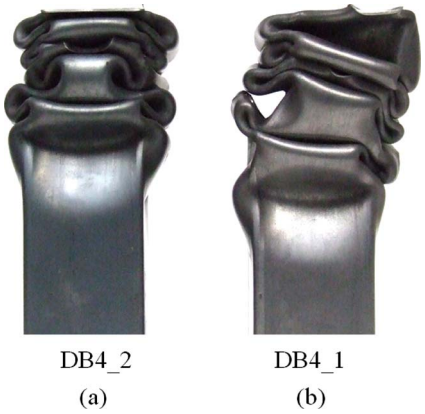


Fig. 16 Photographs of specimens of configuration B4_HEq after loading at heights of (*a*) 2.5 m and (*b*) 3.3 m

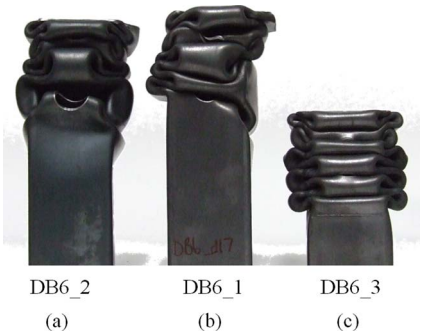


Fig. 17 Photographs of specimens of configuration B6_HEq after loading at heights of (*a*) 2.5 m, (*b*) 3.3 m, and (*c*) 4 m

4.3 Crush Force Efficiency. The crush force efficiency for groups B_H50 and B_HEq is illustrated in Fig. 21. Within each specimen group, CFE fluctuates about some average for an increase in the number of holes. The average CFE for group A, B_H50, and B_HEq is 0.38, 0.40, and 0.36, respectively. The

Fig. 18 Photographs of specimens of configuration B8_HEq after loading at a height of 3.3 m

Fig. 19 Photograph of specimen DB6_HEq_3 after loading at a height of 4 m

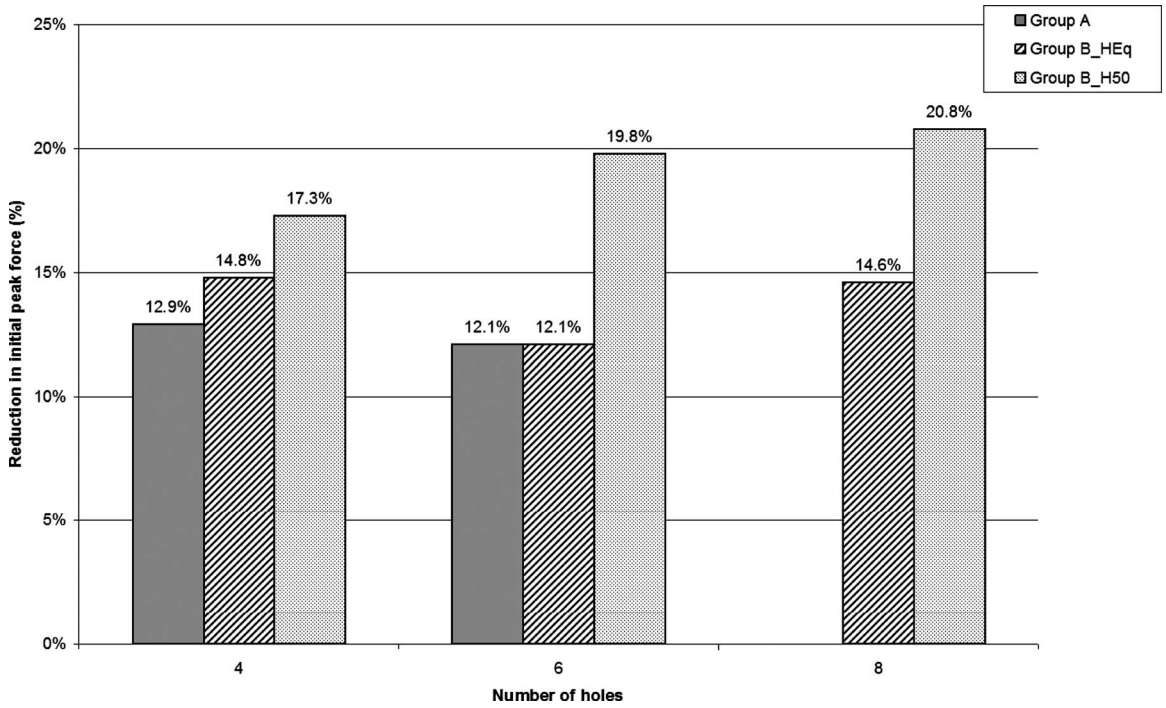


Fig. 20 Reduction in initial peak force for specimens of groups A, B_ HEq, and B_H50 crushed quasistatically

average CFE for group B_H50 is 27% higher than that of the plain tubes, which averages 0.31 for plain tubes tested quasistatically.

4.4 Stroke Efficiency. The stroke efficiencies for all specimens of group A, group B_H50 and B_HEq crushed dynamically at height of 3.3m averaged 65%. This is less than the stroke efficiency of a plain tube crushed at the same height which is 70%. This better stroke efficiency does not necessarily imply that the introduction of holes is futile since it has been shown that groups with holes (B_H50) improve other energy absorption parameters (CFE).

5 Conclusions

Dynamic and quasistatic axial crush tests are performed on mild steel tubes of various hole configurations. The presence of

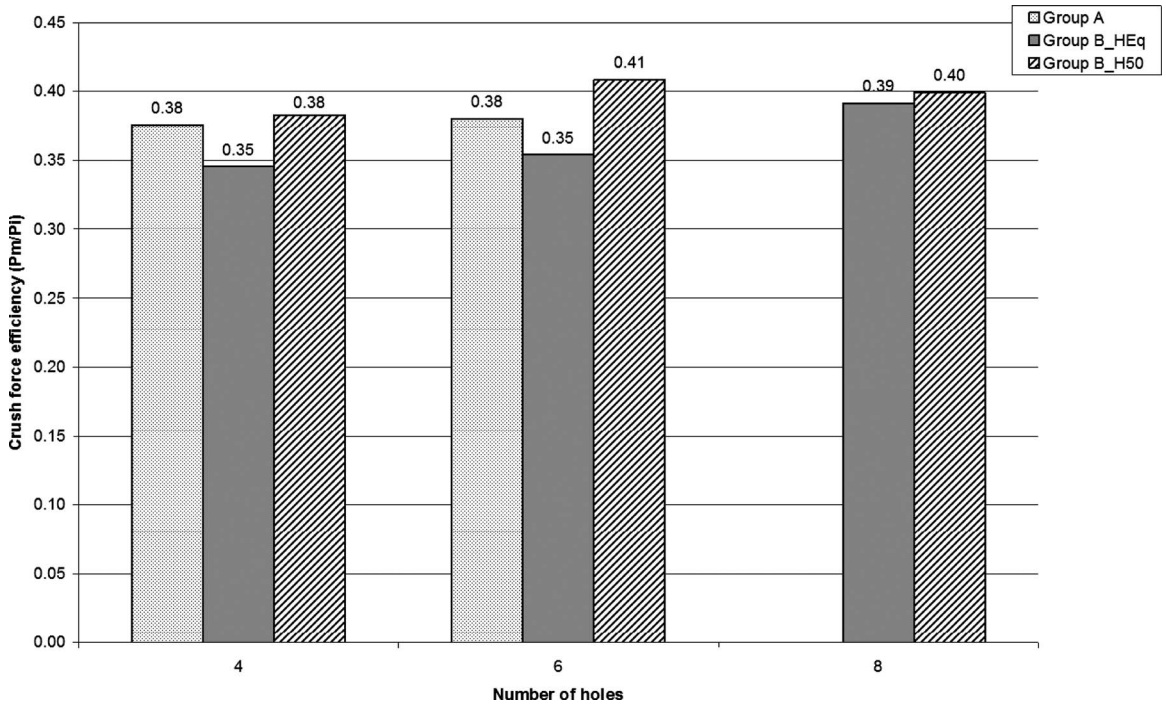


Fig. 21 Crush force efficiency for groups B_H50 and B_HEq for specimens crushed quasistatically

holes decreases the initial peak force at all times; this is however not improved by an increase in the number of holes. Specimen group B_H50 significantly improves the crush force efficiency of a plain tube. The mode of crush initiation that develops in group B_H50 is associated with lower initial peak forces. The location of holes at points corresponding to the natural hinges of lobes, as seen in group B_H50, appears to greatly improve energy absorption by ensuring that progressive collapse mode is repeatedly triggered. This repeatability is important for energy absorption devices as it means that this group can be expected to crush in the preferred progressive mode at most times. Specimen group A displays both progressive and mixed mode collapse in specimens of the same hole configuration. This means that hole discontinuities may have a negative effect on the energy absorption characteristics if unfavorably positioned. Group A specimens, however, show improved performance under dynamic loading.

Acknowledgment

The authors wish to thank Glen Newins and Dylan Jacobs at the Mechanical Engineering Workshop, UCT for manufacturing the test specimens.

Nomenclature

C = width of side of square tube
d = hole diameter
E_a^d = energy absorbed (dynamic)
E_a^s = energy absorbed (quasistatic)
H = initial height of mass
l = half the wavelength of a single lobe
L = initial length of tube
m = mass
P_i = initial peak force
P_m = mean crushing force
t = wall thickness
V = impact velocity
δ = crushed distance (stroke)

References

[1] Lu, G., and Yu, T., 2003, *Energy Absorption of Structures and Materials*, Woodhead, Cambridge.
[2] Abramowicz, W., 2003, "Thin-Walled Structures as Impact Energy Absorbers," Thin-Walled Struct., **41**, pp. 91–107.
[3] Langseth, M., and Hopperstad, O. S., 1996, "Static and Dynamic Axial Crushing of Square Thin-Walled Aluminium Extrusions," Int. J. Impact Eng., **18**(7–8), pp. 949–968.
[4] World Health Organization and World Bank, 2004, "World Report on Road Traffic Injury Prevention," *Prevention*, World Health Organization, Geneva.
[5] Cheng, Q., Altenhof, W., and Li, L., 2006, "Experimental Investigations on the Crush Behaviour of AA6061-T6 Aluminium Square Tubes With Different Types of Through-Hole Discontinuities," Thin-Walled Struct., **44**, pp. 441–454.
[6] Abramowicz, W., and Jones, N., 1984, "Dynamic Axial Crushing of Square Tubes," Int. J. Impact Eng., **2**, pp. 179–208.
[7] Arnold, B., and Altenhof, W., 2004, "Experimental Observations on the Crush Characteristics of AA6061 T4 and T6 Structural Square Tubes With and Without Circular Discontinuities," Int. J. Crashworthiness, **9**(1), pp. 73–87.
[8] Tarigopula, V., Langseth, M., Hopperstad, O. S., and Clausen, A. H., 2006, "An Experimental and Numerical Study of Energy Absorption in Thin-Walled High-Strength Steel Sections," Int. J. Impact Eng., **32**(5), pp. 847–882.
[9] Abramowicz, W., and Jones, N., 1997, "Transition From Initial Global Bending to Progressive Buckling of Tubes Loaded Statically and Dynamically," Int. J. Impact Eng., **19**(5–6), pp. 415–437.
[10] Montanini, R., Belingardi, G., and Vadori, R., 1997, "Dynamic Axial Crushing of Triggered Aluminium Thin-Walled Columns," 30th International Symposium on Automotive Technology & Automation, Florence, Italy, Jun. 16–19, pp. 437–444.
[11] Chung Kim Yuen, S., and Nurick, G. N., 2008, "The Energy Absorbing Characteristics of Tubular Structures With Geometric and Material Modifications: An Overview," Appl. Mech. Rev., **61**(2), p. 020802.
[12] Langseth, M., and Hopperstad, O. S., 1997, "Local Buckling of Square Thin-Walled Aluminium Extrusions," Thin-Walled Struct., **27**(1), pp. 117–126.
[13] Marshall, N. S., and Nurick, G. N., 1998, "The Effect of Induced Deformations on the Formation of the First Lobe of Symmetric Progressive Buckling of Thin Walled Square Tubes," Structures Under Shock and Impact (SUSI 98), Thessaloniki, Greece, Jun. 24–26, N. Jones, D. G. Talaslidis, C. A. Brebbia, and G. D. Manolis, eds., pp. 155–168.
[14] Gupta, N. K., and Gupta, S. K., 1993, "Effect of Annealing, Size and Cut-Outs on Axial Collapse Behaviour of Circular Tubes," Int. J. Mech. Sci., **35**(7), pp. 597–613.
[15] Lee, S., Hahn, C., Rhee, M., and Oh, J., 1999, "Effect of Triggering on the Energy Absorption Capacity of Axially Compressed Aluminum Tubes," Mater. Des., **20**(1), pp. 31–40.
[16] Abramowicz, W., and Jones, N., 1986, "Dynamic Progressive Buckling of Circular and Square Tubes," Int. J. Impact Eng., **4**(4), pp. 243–270.
[17] Bodlani, S. B., Chung Kim Yuen, S., and Nurick, G. N., 2009, "The Energy Absorption Characteristics of Square Mild Steel Tubes With Multiple Induced Circular Hole Discontinuities—Part II: Numerical Simulations," ASME J. Appl. Mech., **76**, p. 041013.
[18] Otubushin, A., 1998, "Detailed Validation of a Non-Linear Finite Element Code Using Dynamic Axial Crushing of a Square Tube," Int. J. Impact Eng., **21**(5), pp. 349–368.

The Energy Absorption Characteristics of Square Mild Steel Tubes With Multiple Induced Circular Hole Discontinuities—Part II: Numerical Simulations

S. B. Bodlani

S. Chung Kim Yuen[1]
e-mail: steeve.chungkimyuen@uct.ac.za

G. N. Nurick

Department of Mechanical Engineering,
Blast Impact and Survivability Research Unit
(BISRU),
University of Cape Town,
Private Bag,
Rondebosch 7700, South Africa

This paper is Part II of a two-part article and presents the results of numerical simulations conducted to investigate the energy absorption characteristics of square tubes subjected to dynamic axial loading. Part I reports the experimental results of both quasistatic and dynamic tests. The validated model is used to study the crushing characteristics of tubes with multiple induced circular hole discontinuities using the finite element package ABAQUS/EXPLICIT version 6.4-6. Holes of diameter 17 mm are used as crush initiators, which are laterally drilled into the tube wall to form opposing hole pairs. Holes of diameters 12.5 mm and 25 mm are also used to assess the effects of hole diameter on energy absorption. Two hole spacing configurations are investigated, one in which the hole pairs are placed at regular intervals of 50 mm along the tube wall and another in which the hole pairs are spaced symmetrically along the tube length. Holes are also drilled on either two or all four opposing tube walls. The number of holes is varied from 2 to 10. The results indicate that the introduction of the holes decreases the initial peak force. However, an increase in the number of holes, beyond two holes, does not further significantly decrease the initial peak force. A study of the crushing history of the tubes reveals that crushing is initiated at the location of the holes. The results also indicate that the type of hole spacing determines how crushing is initiated at the hole locations. The model satisfactorily predicts the resultant collapse shapes but overpredicts the crushed distance. [DOI: 10.1115/1.3114967]

Keywords: square tubes, imperfections, multiple circular cut-outs, energy absorption

1 Introduction

Part I [1] in this two-part paper reports on the experimental investigation aimed to determine the effect of inserting multiple circular hole discontinuities on the energy absorption characteristics of axially loaded square tubes. Data from the quasistatic tests provided complete force-displacement history of the crushing event. The nature of the quasistatic test also allows observation of the collapse history during the entire crushing event and hence the effects of the holes on the initiation and development of the lobe formation. The rapidity of the dynamic test, however, precludes the same level of analysis by removing the ability to visually observe the crushing process without the use of expensive equipment, such as high speed camera. With the unavailability of the relevant equipment, the force-displacement history of the dynamic loading is not recorded. These constraints limit the understanding to be garnered from the dynamic test; in particular, the action of the discontinuities in inducing crushing cannot be monitored and reported on. In a discussion about the design of energy absorbing devices, Abramowicz [2] stressed the need for the relevant absorber to be "triggered" in such a way that the formation of natural folds is promoted. This allows the peak force to be reduced in such a way that any instability is arrested in the region of the discontinuity. The use of finite element analysis (FEA), once validated, can provide better insight into the initiation and progression of crushing in tubes loaded dynamically and enables parameters inaccessible experimentally to be investigated [3].

Simulation by the finite element method is generally accepted as a good indication of material and structural behavior under various loading conditions [2]. Chung Kim Yuen and Nurick [3] provided an overview of the use of imperfections to improve energy absorption in thin-walled tubes and discussed the various researchers [2,4–8], who successfully used commercial finite element packages to model the response of thin-walled structures under axial loading. Popular finite element packages noted in the review [3] are LS-DYNA, used by Tarigopula et al. [8], Langseth et al. [7], and Ostubishin [5]; ABAQUS, used by Lee et al. [6]; and PAM-CRASH, also used by Lee et al. [6]. Common among the abovementioned researchers is the use of four-noded shell elements to model the tubular sections and, in particular, to characterize the initial peak force, lobe wavelength, collapse mode, and energy absorption in general.

When more than one discontinuity is inserted into the tube wall, the location of the discontinuities becomes important. Montanini et al. [7] used DYNA3D to assess the effect of various shapes of triggers on the collapse mode and energy absorption characteristics of thin-walled square tubes. It was suggested that the use of theoretical predictions to calculate the length of the lobe wavelength may aid in obtaining appropriate positioning for triggers

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received October 2, 2007; final manuscript received November 2, 2007; published online April 27, 2009. Review conducted by Robert M. McMeeking.

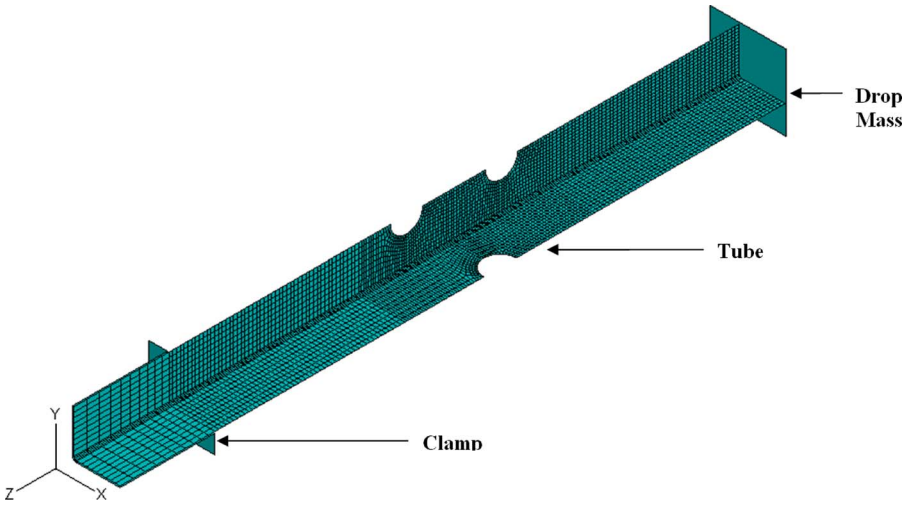


Fig. 1 Finite element analysis model assembly for configuration B6_H50

[7]. Lee et al. [6] investigated the effect of multiple triggers on the energy absorption characteristics of square tubes. Both ABAQUS and PAM-CRASH were used to pre-estimate the folding sites (hinges) at which the triggers should be inserted. It was shown that the insertion of triggers at the simulation predicted sites improved energy absorption [6]. Conversely, it was found that if induced discontinuities were not located at the lobe hinges, then they would cause irrevocable disturbances to the collapse mode and lead to global bending even if the geometry of the tube was such that it would have otherwise buckled in progressive mode [6].

In Part I [1], the actual lobe wavelength of a plain tube (no discontinuities), which is 350 mm in length (L), 50 mm in width (C), and 1.5 mm in thickness (t), is measured and found to be approximately 50 mm. Examining the crushing history as viewed through the simulation generated images allows the role of the hole discontinuities in the initiation and progression of collapse to be studied closely for the dynamic tests as for quasistatic tests. Furthermore, FEA allows many more tests to be "conducted" than is physically or economically possible experimentally [2]. The numerical package used for the simulation is ABAQUS/EXPLICIT version 6.4-6. The simulation models a drop test of mass of 330 kg and an impact velocity of 8 m/s for various hole discontinuity configurations.

2 Finite Element Analysis Formulation

In order to minimize processing time, a quarter model with two symmetry planes is used to model the tube, mass, and clamp assembly. The tube geometry is modeled as a deformable shell extrusion of thickness 1.6 mm and depth 350 mm using a structured mesh of S4R elements. The shell elements allow the assumption that the tube experiences no through-thickness stresses. The clamp and the mass are modeled as rigid bodies. No stresses or strains are monitored for the clamp and the mass thus allowing them to be meshed with as few elements as possible while simultaneously preserving their geometric shape. Two interaction properties are defined for the model assembly. The first one assigns frictionless tangential behavior and hard-contact normal behavior, which allows separation after contact and is applied to define the contact between the clamp and the tube. This property is also applied to the tube, mass, and clamp in their interaction with themselves and thus also defines the contact behavior between successive lobes in the tube. The second interaction property assigns a tangential coefficient of friction of 0.3 and is applied to define the contact between the mass and the tube. This value of the coefficient of friction was used successfully by Tarigopula et al. [8] to prevent lateral motion between the tube and the mass. The assembly of the model is illustrated in Fig. 1, in which loading occurs in the positive direction of the indicated z-axis.

The mass is assigned an initial velocity of 8 m/s, which, in the experimental tests, corresponds to a drop height of 3.3 m under the assumption of zero friction losses. The following boundary conditions are further defined; the mass is constrained in all rotational and translational degrees of freedom except for translation parallel to the z-axis under the action of gravity, and all translational and rotational degrees of freedom are constrained for the clamp. Two symmetry planes, parallel to the z-axis, are created on the tube to give the quarter model across which reflection enables the entire tube geometry to be discerned. Translational motion in the z-axis direction is constrained for the bottom-most edge of the tube. The interaction property defined for the clamp enables the boundary conditions applied to the clamp to also constrain the motion of the clamped end of the tube in all directions except translation along the z-axis.

Tensile tests conducted on the tube material give a yield stress (σ_y) of 328 MPa. These data are converted into true stress (σ_t) and logarithmic plastic (ε_t) strain and is assigned to the tube. Strain rate effects are incorporated using the Cowper–Symonds relationship [9] with constants $D=844\ \text{s}^{-1}$ and $q=2.204$, obtained for common South African mild steel by Marais et al. [10]. Temperature effects are also incorporated using Young's modulus (E) and yield stress (σ_y) dependence on temperature as proposed by Masui et al. [11]. Other general material properties assigned to the tube are density (ρ) 7850 kg/m^3, Young's modulus (E) of 210 GPa, inelastic heat fraction of 0.9, Poisson's ratio of 0.3, and specific heat of 450 kJ/kg K. The initial temperature is set to 293 K.

3 Validation of Finite Element Analysis

The finite element model generally overestimates the crushed distances for dynamic tests. The model can also not be used to validate the crush forces due to the unavailability of the relevant instrumentation in the drop test rig. Nevertheless, the finite element results show good correlation for crush initiation and final crushed shape and are therefore used as a comparative study.

3.1 Crush Initiation Validation. Observations made in the crush initiation for the quasistatic tests are compared with the FEA results. This comparison is made on the assumption that quasistatic tests and dynamic tests performed at velocities under 10 m/s should, according to Jones [9], yield similar collapse characteristics. A comparison of the quasistatic crush initiation for

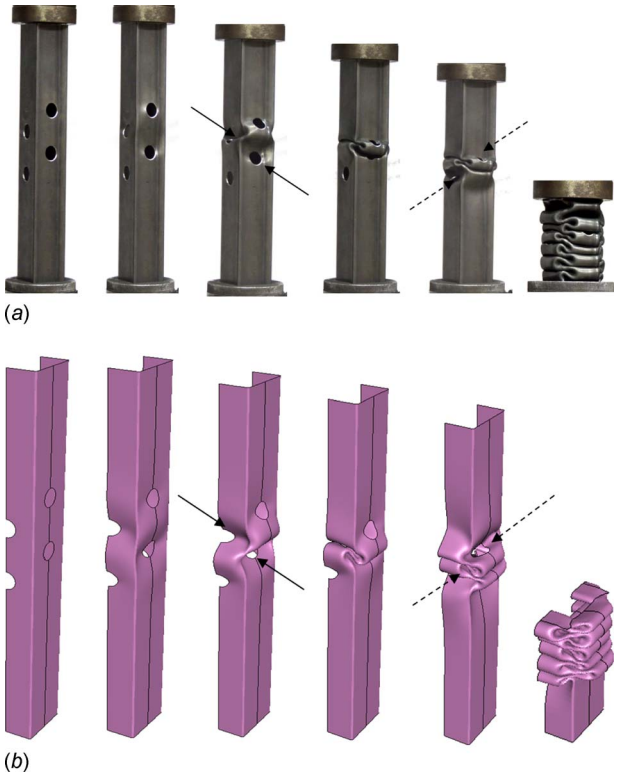


Fig. 2 Comparison of (*a*) experimental and (*b*) simulation crush initiations for configuration B8_H50

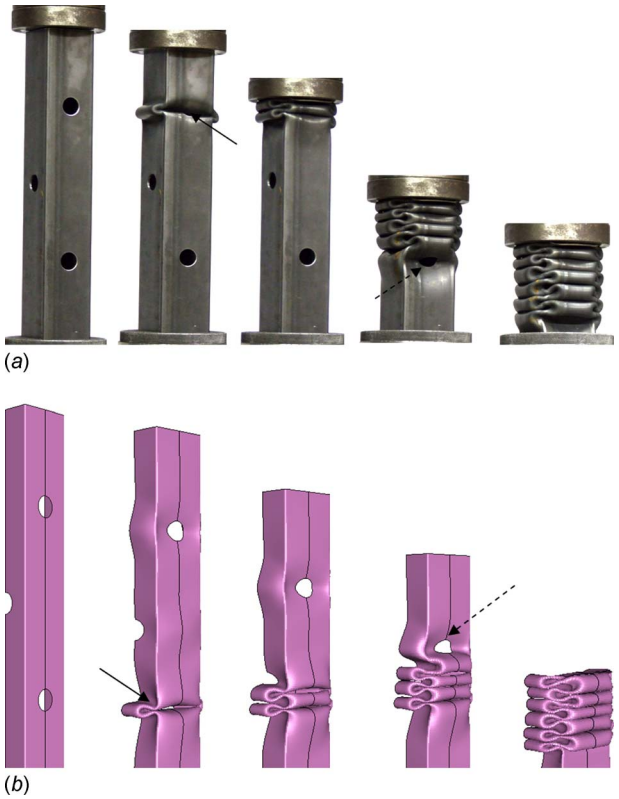


Fig. 3 Comparison of (*a*) experimental and (*b*) simulation crush initiations for configuration B6_HEq

configurations B8_H50 and B6_HEq against the simulation results is given in Figs. 2 and 3. For B8_H50, it is clear in both experimental and FEA images, Fig. 2(*a*) and 2(*b*), respectively, that crush initiation occurs by the same mechanism in which two hole pairs act together to initiate the first lobe. The hole pairs that initiate crushing are indicated by the solid arrows. The remaining two hole pairs fall on the hinges of lobes forming subsequent to the initial lobe and are indicated by the dashed arrows in Figs. 2(*a*) and 2(*b*). For B6_HEq, the initiation of crushing occurs at a single hole pair for both simulation and experimental results, as illustrated by the solid arrows in Fig. 3(*a*) and 3(*b*), respectively. Subsequent crushing occurs progressively from the initial lobe even though other hole pairs are available at remote locations away from the initial lobe. That is, the tube crushes as if the hole discontinuities were absent but it is observed that these other hole pairs fall on the hinges of the lobes, as illustrated by the dashed arrows in Fig. 3.

3.2 Crushed Shape Validation. All the simulation models crush in the progressive mode. This was not the case for the experiments in which some specimens collapse in the mixed mode (MM) where progressive buckling commences but is disturbed and global bending results. It is, however, observed in the experiments [1] that the occurrence of mixed mode appears to be random, occurring in some specimens of the same configuration. Also holes fall on the inner hinges of lobes for all simulations while in the experimental results holes fall on either the inner or outer hinge for specimens of the same configuration.

3.2.1 Group A. The presence of a hole pair appears to decrease the size of a lobe compared with those adjacent to it. This observation was also made by Marshall and Nurick [12] and is noted presently. The final crushed profile for A4_d17 is illustrated in Fig. 4 for both experimental and simulation tests. The arrows in Fig. 4 indicate lobes of decreased size compared with their adjacent counterparts. The numbered arrows illustrate that the simulation predicts the same number of lobes as the experimental tests.

3.2.2 Group B_HEq. A comparison of the experimental and FEA final profiles for configurations B4_HEq, B6_HEq, and B8_HEq are given in Figs. 5–7. It is noted that the simulation results overestimate the crushed distance of the experimental results and this is evident in Figs. 5–7. The evolution of the lobes and their shape is still closely predicted by the simulation, however. Perceptibly large lobes, indicated by the arrows in Fig. 5, are induced by the prevailing hole spacing in the tubes. Figure 6 illustrates that the final shape B6_HEq is accurately predicted by the model and is characterized by five uniform lobes. The collapse shape of B8_HEq is characterized by curling of some, as illustrated by the solid arrows in Figs. 7(*a*) and 7(*b*). Figure 7 also shows, for both simulation and experimental results, an uncrushed length at the impact end (indicated by the braces) and the formation of a large lobe, which is indicated by the dashed arrows. The experimental specimens of group B_HEq in general are prone to

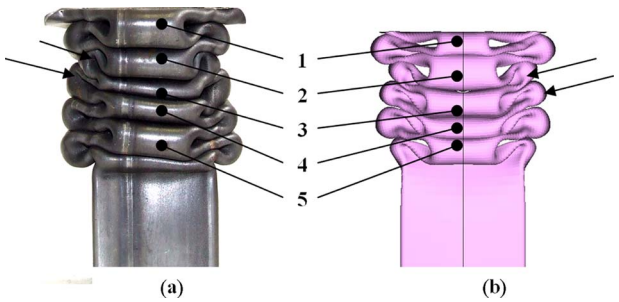


Fig. 4 Comparison of (*a*) experimental and (*b*) simulation crush modes for configuration A4_d17

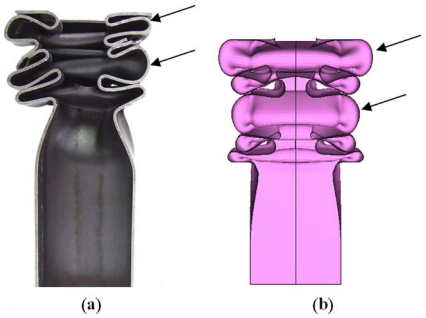


Fig. 5 Comparison of experimental (*a*) and simulation (*b*) crush modes for configuration B4_HEq

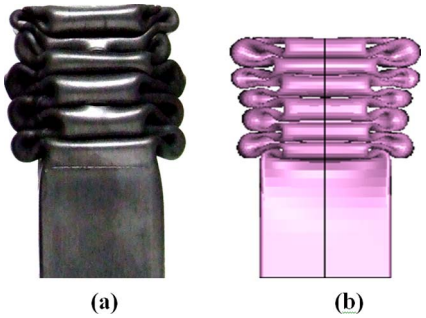


Fig. 6 Comparison of (*a*) experimental and (*b*) simulation crush modes for configuration B6_HEq

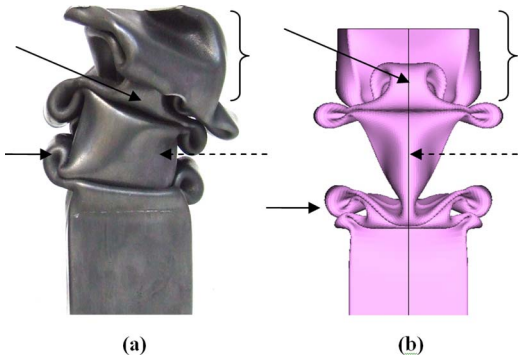


Fig. 7 Comparison of experimental and simulation crush modes for configuration B8_HEq

instability and tend to lean to one side. These comparisons show that the simulation predicts the collapse shape of the specimens well.

4 Analysis Results and Observations

4.1 Group A

4.1.1 A2. The crush behavior for specimens of two opposing holes at midheight is similar for the different hole diameters, and the force-displacement curves are given with that of the as-received (plain—no initiators) tube in Fig. 8. A marked reduction in initial peak force is evident between the tubes with holes and the plain tube. The extent to which the initial peak force (P_i) is reduced depends on the size of the discontinuity such that a hole diameter of 12.5 mm (A2_d12.5) produces the smallest reduction in P_i and a diameter of 25 mm (A2_d25) produces the largest reduction. The circled region indicates a high peak force at the end of crushing, which is caused by the compaction of the tube.

Figure 9 illustrates the initial stages of collapse for (a) a plain tube and (b) A2_d17. The progressive buckling for the plain tube commences at the impact end, as shown in Fig. 9(*a*). Progressive buckling proceeds in a steady manner all the way down the tube. Small deflections near the clamped end of the tube are observed in Fig. 9(*a*). Langseth and Hopperstad [13] observed similar hardly discernable buckles distributed over the entire length of quasistatically and dynamically loaded aluminum square tubes before the onset of localized buckling. The last lobe formed in what is termed extensional mode, which is characterized by a seamless fold running all the way around the tube. Langseth and Hopperstad [13] also observed extensional mode on the peripheral lobes in square tubes, which were expected to buckle in symmetric mode tested in the vicinity of 8 m/s. In Fig. 9(*b*), deflections of the tube wall are observed alongside the hole discontinuity pair in A2_d17. The hole discontinuity pair gradually takes on an elliptical shape as the deflections progress. The deflections continue to expand outward and eventually result in the formation of a lobe and the flattening of the hole. Progressive buckling proceeds from the first lobe on either the top or the bottom half. Once that half is completely crushed, buckling continues on the other half, progressing from the location of the first lobe. The mean force for specimens of two opposing holes remains relatively constant for an increase in diameter and the average of the three specimens of diameters 12.5 mm, 17 mm, and 25mm is 42.44 kN and that of a plain tube is 46.40 kN. This slight decrease in mean force may be

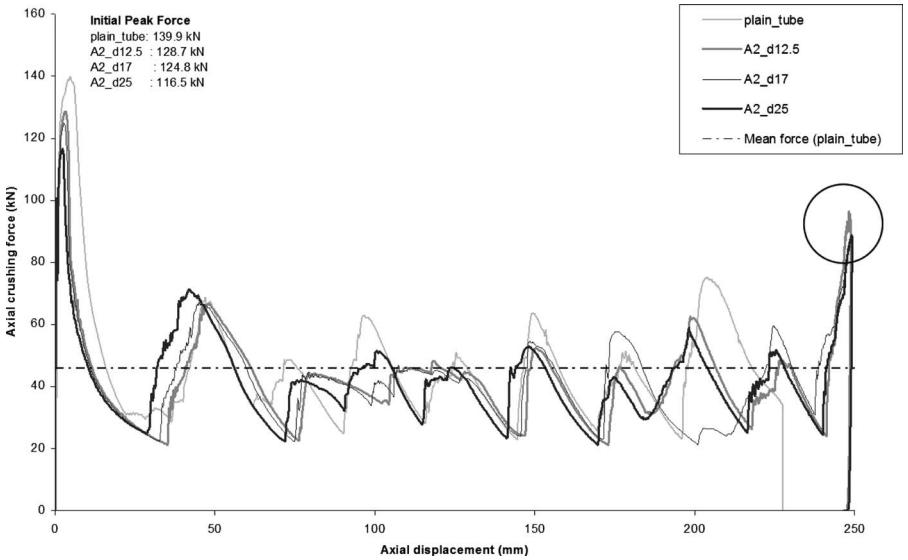


Fig. 8 Force-displacement curves for A2_d12.5, A2_d17, A2_d25, and a plain tube

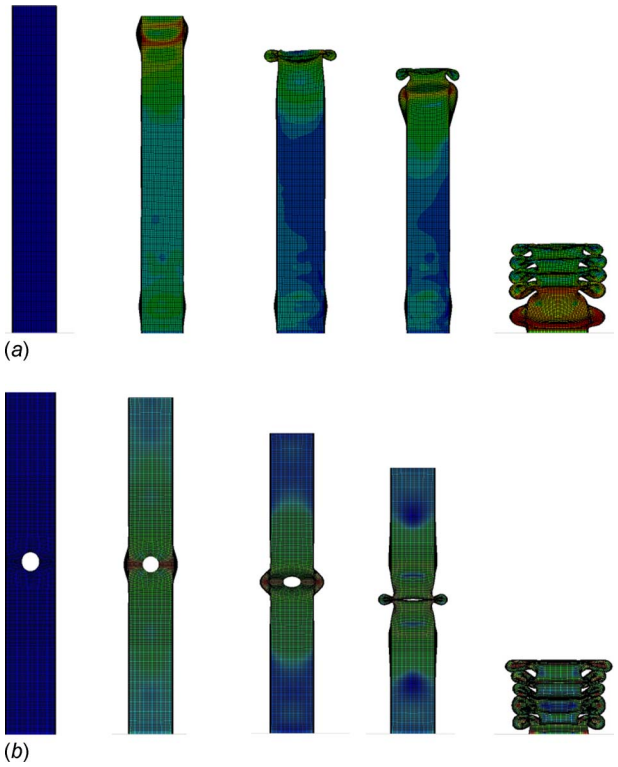


Fig. 9 Initiation of buckling for a (*a*) plain tube and (*b*) A2_d17

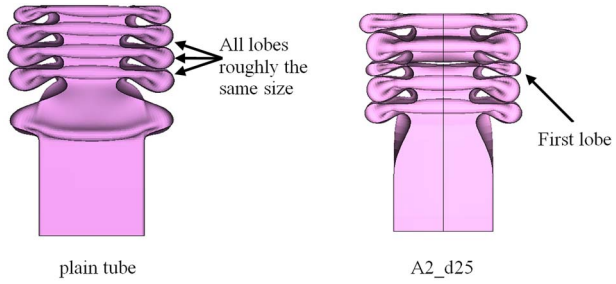


Fig. 10 Difference in size of first lobe between a plain tube and specimen A2_d25

attributed mainly to the decrease in initial peak force effected by the presence of the hole discontinuities.

Though configuration A2 specimens failed in the same progressive mode as the plain tube, some changes in the deformation shape are observed. The insertion of increasing hole diameters in the plain tube results in a reduction in the size of the first lobe but with no change in the collapse mode. Figure 10 illustrates this observation for the specimen of hole diameter 25 mm. The lobes of the plain tube are approximately the same size. The reduction in the size of the first lobe is most prominent for specimen A2_d25 and is less apparent for specimens A2_d12.5 and A2_d17, which have comparatively smaller hole diameters.

4.1.2 A4, A6, and A8. The effects of increasing the number of holes are similar across specimens of varying diameters; consequently, only the specimens with hole diameter 25 mm are taken up for further discussion. Figure 11 presents the force-displacement curves for group A specimens of diameter 25. The mean force for A4, A6, and A8 remains relatively constant across specimens with the same hole diameter. For example, the mean forces for specimens A4_d25, A6_d25, and A8_d25 are 42.80 kN, 43.15 kN, and 43.68 kN, respectively; the average of which is 43.20 kN. This average is similar to the average mean force for specimens of configuration A2 (A2_d12.5, A2_17, and A2_d25), which is 42.44 kN. This implies that the introduction of further hole discontinuities after the initial pair affords no further change in mean force.

The lobe formation processes of specimens A4_d25 and A8_d25 are alike, particularly at the initial stages of buckling. This can be attributed to the geometric similarity of these specimens at the location of the first lobe, which results in the same manner of crush initiation. The initial stages of crushing are illustrated in Fig. 12 for A4_d25 and A8_d25 and in Fig. 13 for A6_d25. Hardly discernable lobes are visible on the tube walls alongside the hole discontinuities, as noted for specimens of configuration A2. On continued crushing, however, buckling does not localize at any single hole discontinuity pair but rather between two hole pairs, as illustrated in Figs. 12(*a*) and 12(*b*), respectively. This is in contrast to specimens of A2 in which the first lobe forms by an outward deflection of the tube walls directly adjacent to the hole as a result of the hole assuming an elliptical shape and falling into the lobe. On continued crushing, subsequent lobes form such that the hole discontinuities are at the hinges of the lobes. In A6_d25, Fig. 13, the first lobe is induced at one of the hole pairs located at midheight. The first lobe forms in a skew manner char-

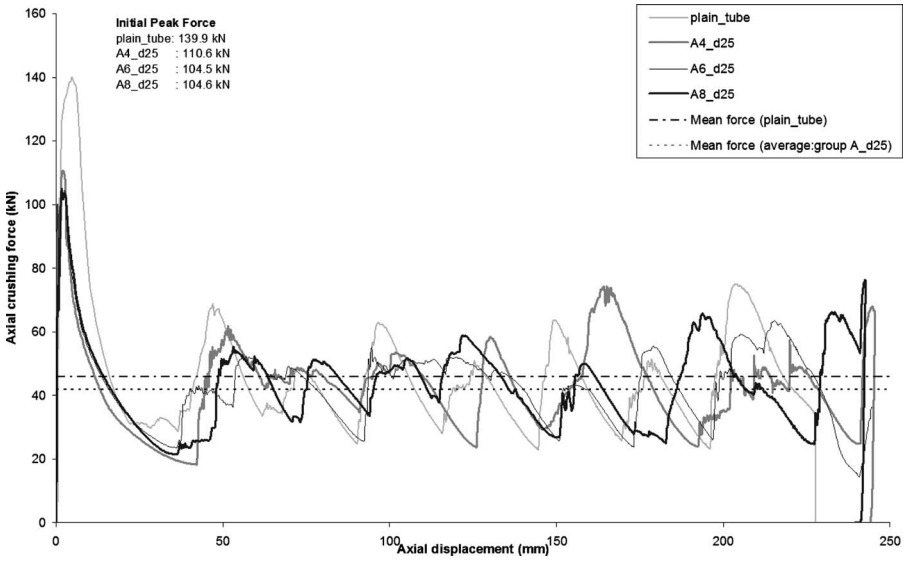
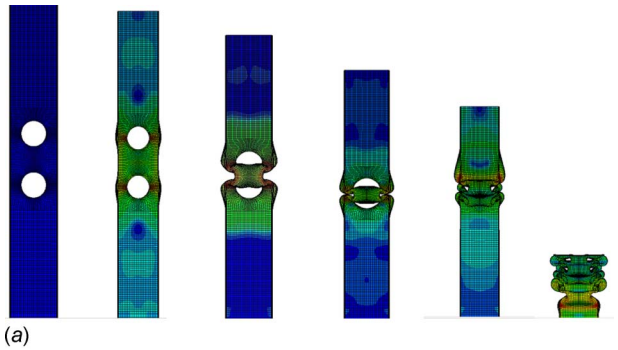


Fig. 11 Force-displacement curves for group A and a plain tube



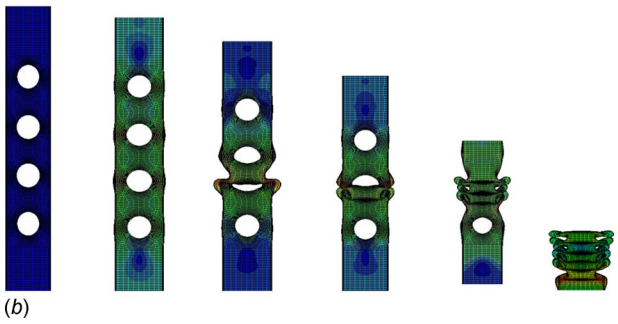


Fig. 12 Initial transient response for (*a*) A4_d25 and (*b*) A8_d25

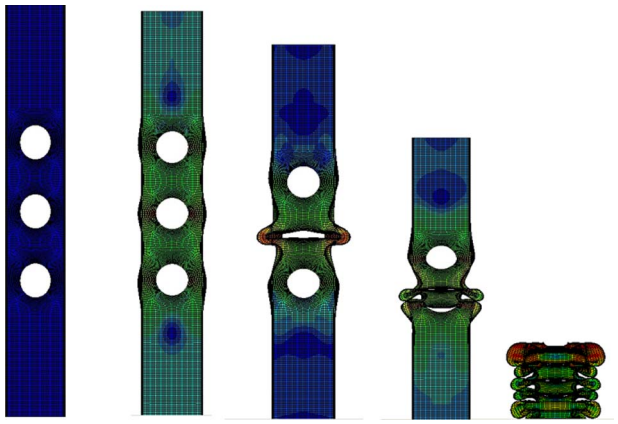


Fig. 13 Initial transient response for specimen A6_d25

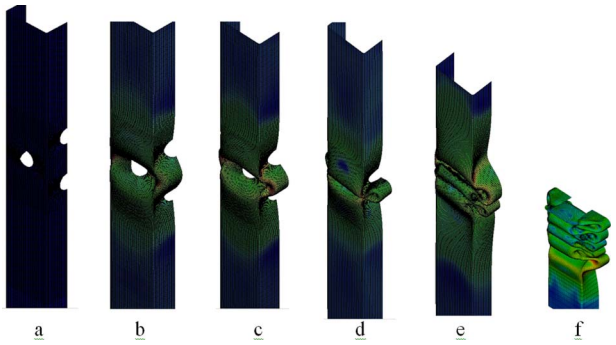


Fig. 15 Transient response for B6_H50

acterized by a bending of the lobe in the downward direction. Further loading results in progressive buckling first above the initial lobe and then below it.

4.2 Group B_H50 and B_HEq. Group B contains all specimens with hole discontinuities on all four tube walls. Group B is further subdivided into groups B_H50 and B_HEq. In both groups, there is a slight decrease in initial peak force with an increase in number of holes. As expected, this decrease increases with an increase in hole diameter. The crush modes observed for groups B_H50 and B_HEq are similar for specimens of different diameters and thus only the behavior of the specimens with holes of diameter 17 mm are discussed in detail.

4.2.1 B_H50. The collapse behaviors of all specimens of group B_H50 are very similar in that the initial force is greatly reduced by the action of more than two or three hole pairs in inducing crushing. The force-displacement curves for B4_H50, B6_H50, B8_H50, and B10_H50 are given Fig. 14 together with that of a plain tube. The circled region indicates compaction of the tube. The crushing process for B8_H50 is described in Fig. 2 and is similar to that of B4_H40. In both specimens, two hole pairs simultaneously induce the initial lobe. The crushing process for B6_H50, which crushed in the same manner as B10_H50, is illustrated in Fig. 15. Figure 15(*c*) illustrates that two or three hole pairs simultaneously trigger the initiation of the first lobe such that the hole pairs lie in the hinges of the initial lobe. Further loading produces progressive buckling first above and then below the initial lobe.

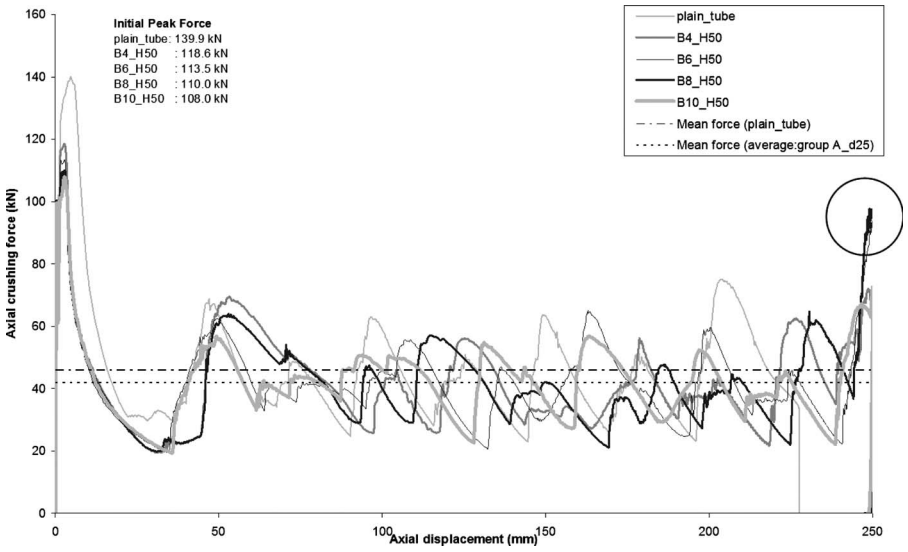


Fig. 14 Force-displacement curves for group B_H50 and a plain tube

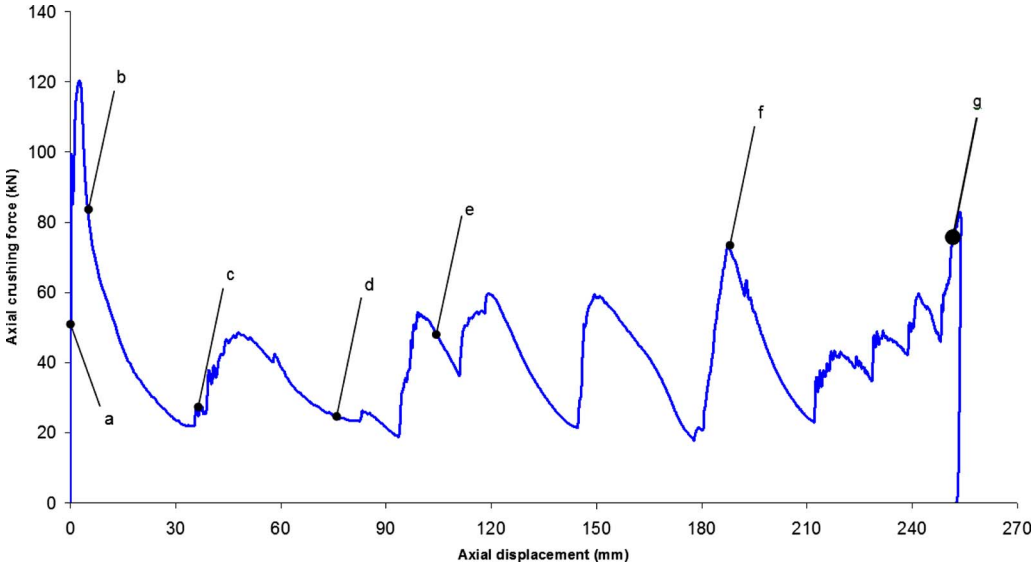


Fig. 16 Force-displacement curve for B4_HEq

4.2.2 B_HEq. This group is characterized by progressive collapse with the formation of nonuniformly sized lobes. The sizes of the lobes forming are dependent on the spacing between successive hole pairs. Each specimen displays a unique progressive crushed shape and is discussed separately below.

4.2.2.1 B4_HEq. The force-displacement curve and the transient response for B4_HEq are provided in Figs. 16 and 17, respectively. Figure 17(*c*) illustrates that for specimen B4_HEq crushing initiates at one of the two hole discontinuity pairs. Subsequent crushing occurs above the initial lobe without being influenced by the other hole discontinuities. The next lobe is not formed at the second hole discontinuity pair but rather between the initial lobe and the second hole. The second hole discontinuity pair does, however, fall into the hinge of the lobe formed above the initial lobe. Crushing then resumes below the initial lobe. The hole spacing between successive hole discontinuity pairs is 100 mm; this results in the formation of lobes larger than those encountered in a plain tube.

4.2.2.2 B6_d17_HEq. The hole spacing for successive hole pairs along the tube length is 75 mm for specimen B6_HEq. The force-displacement curve of specimen B6_HEq is similar to that of a plain tube in that the force-amplitude and length of successive periods is similar. This is due to the uniformly sized lobes that develop in the specimen. Some areas of decreased force amplitude are, however, present and correspond to the formation of lobes whose hinges contain hole discontinuities. Figure 18 shows specimen B6_HEq after crushing. As can be observed in Fig. 18, the lobes in specimen B6_HEq are of a uniform size. The actual crushing process for the specimen initiates at the bottom-most hole and continues in the same way as in specimen B4_HEq in which subsequent progressive crushing continues above the initial lobe. The 75 mm spacing between successive holes allows one lobe to form over a 50 mm length and the subsequent one is split between two 25 mm lengths (recall that the wavelength of single lobe is 50 mm). This allows the hole discontinuities to fall on the inner hinges of lobes even though crushing does not initiate remotely at those holes.

4.2.2.3 B8_Heq. The force-displacement curve and transient response for B8_HEq are illustrated in Figs. 19 and 20, respectively. The initial lobe in specimen B8_HEq forms at a single hole discontinuity pair, as shown in Fig. 20(*c*). The second lobe forms remotely at the top pair of holes. This is different from the other

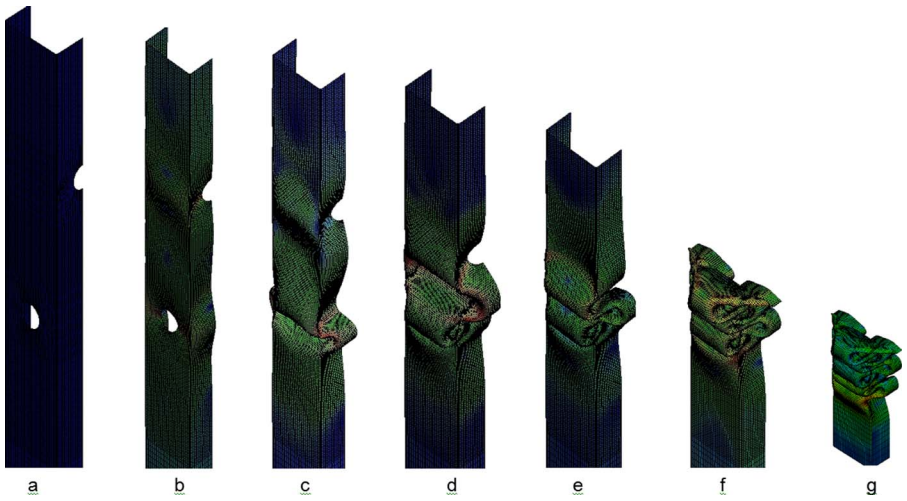


Fig. 17 Transient response for specimen B4_HEq

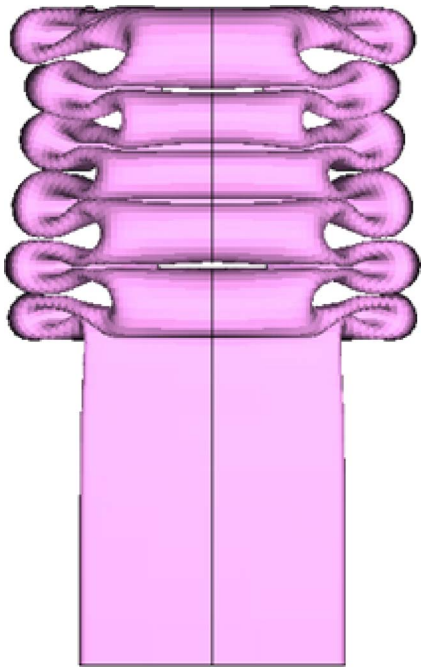


Fig. 18 B6_HEq after crushing

specimens, such as B6_HEq, in which lobes forming subsequent to the initial lobe develop in the space between successive hole pairs as opposed to remotely at another pair. Further crushing results in the formation of a very large lobe, as shown in Fig. 20(*h*).

5 Discussion of Findings

5.1 Initial Peak Force. Since the intention is to improve the energy absorption of a plain tube, it is useful to assess the extent to which the initial peak force is reduced when hole discontinuities are introduced. There is a slight decrease in initial peak force when the number of holes is increased from 4 to 8 in groups A, B_H50, and B_HEq. The percent reduction in initial peak force is illustrated in Fig. 21 for specimens of diameter 17 mm. For a given number of holes, group B_H50 yields the largest reduction in initial peak force. The difference in initial peak force between specimens of group A and those of group B_HEq is almost negligible for the same hole diameter. This is attributed to the mode of initiation prevalent in groups A and B_HEq, in which crush initiation occurs at one of the hole discontinuity pairs. This is in contrast to group B_H50 in which two hole discontinuity pairs

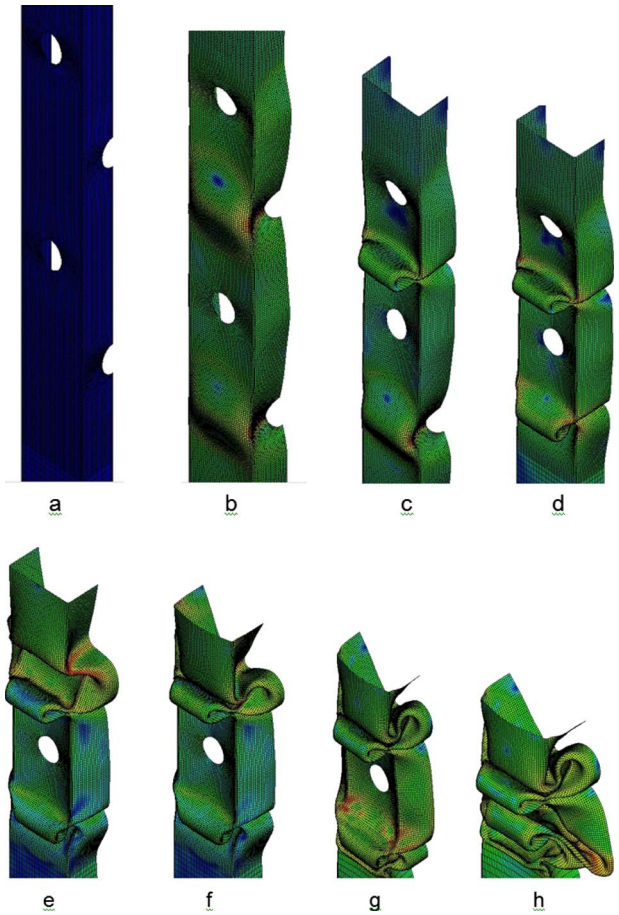


Fig. 20 Transient response for B8_HEq

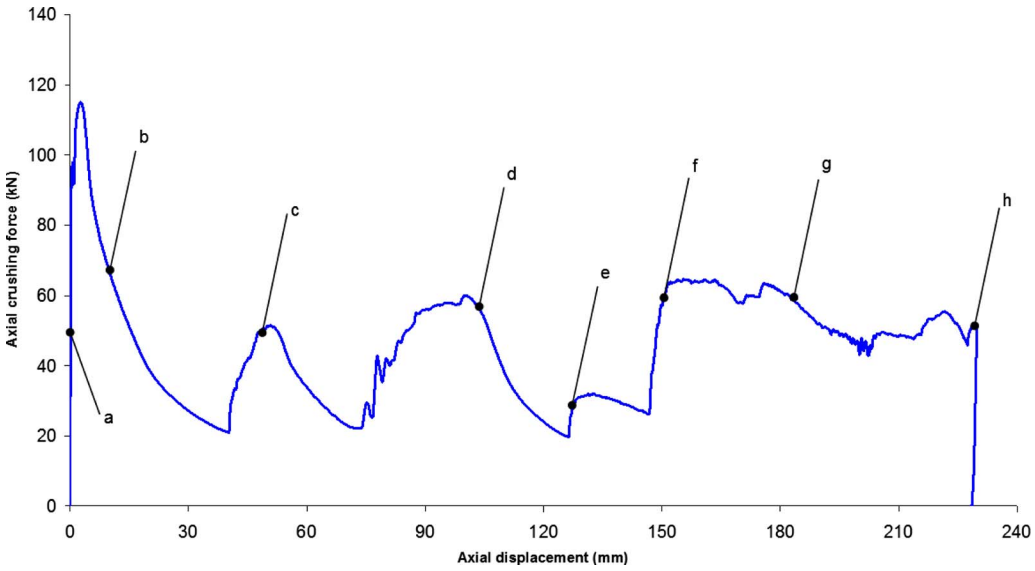


Fig. 19 Force-displacement curve for specimen B8_HEq as obtained by FE Analysis

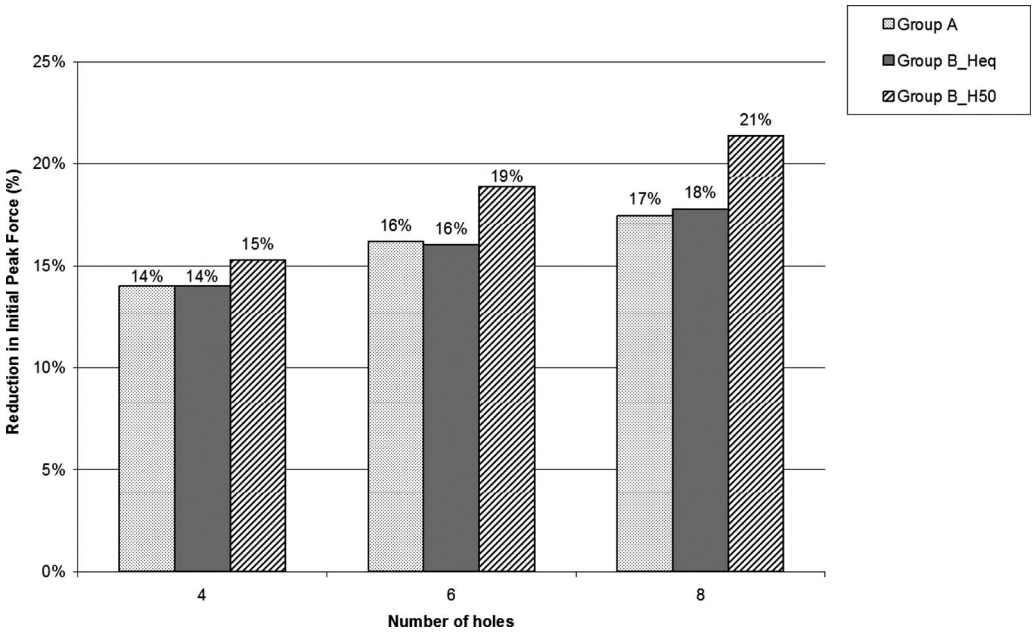


Fig. 21 Reduction in the initial peak force of a plain tube for specimens of diameter 17 mm

simultaneously induce the initial lobe. This dual action results in the lower initial peak forces encountered in group B_H50. The extent to which the initial peak force can be reduced via the introduction of holes is, however, finite. In other words, there is a point beyond which increasing the number of holes will have a negligible effect on the initial peak force. It is expected that an increase in the number of holes for specimens of hole diameter 25 mm will result in lower initial peak forces than the same increase in number of holes for specimens of diameter 12.5 mm or 17 mm. This trend is due to the ability of the hole discontinuity diameter (d) to scale the initial peak force (P_i) by reducing P_i with an increase in d.

5.2 Collapse Mode. All specimens collapsed in the progressive mode. This is expected since the geometry of the tubes is such that Euler buckling is not expected. A major difference in the collapse profile encountered in the different groups is that the fixed spacing groups (groups A and B_H50) produced lobes of uniform size very much like those found on a plain tube, while group B_HEq produced irregular-sized lobes. This implies that the hole discontinuities fall on the hinges of the lobes formed thus inducing a specific lobe wavelength that is different from the tube's natural wavelength.

5.3 CFE. The crush force efficiency (CFE) parameter indicates the extent to which the initial peak force is reduced and the mean force preserved in order to maximize energy absorption. The values of CFE are given in Fig. 22 for the different groups for specimens of diameter 17 mm. The difference in crush force efficiency between the different groups is negligible. What is evident,

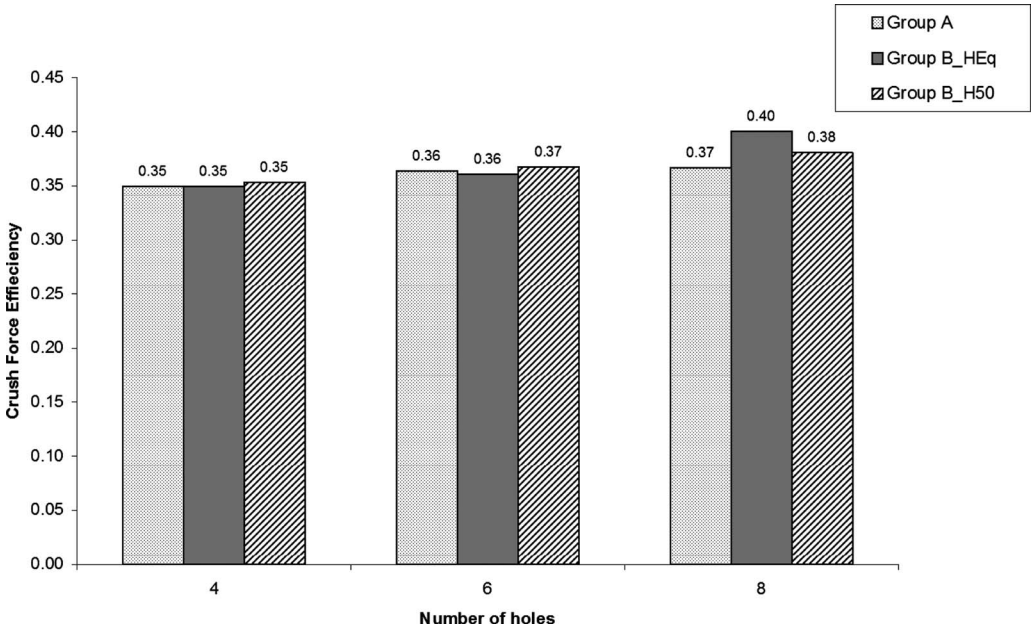


Fig. 22 Crush force efficiency for specimens of diameter 17 mm

however, is a slight increase in CFE with an increase in number of holes. This is expected given the decrease in initial peak force noted for an increase in number of holes. The mean force remained relatively constant at an average of 42 kN for specimens of diameter 17 mm. This is a slight reduction in the mean force of a plain tube, which is 46 kN. This reduction may simply be due to the reduction in initial peak force effected by the introduction of hole discontinuities.

5.4 SE. Stroke efficiency (SE) is synonymous with the amount of energy absorbed given that an increase in stroke automatically indicates an increase in energy absorption. The stroke efficiency did not change significantly between the different groups or with an increase in number of holes. The average stroke efficiency of specimens of groups A, B_HEq, and B_H50 for a diameter of 17 mm is 82% while that for a plain tube is 76%.

6 Conclusions

Dynamic and quasistatic axial crush tests are performed on mild steel tubes of various hole configurations. The test rig used for the experimental dynamic tests is unable to provide data concerning the crushing dynamics of the tubes. Finite element analysis is hence performed to model the dynamic tests. The finite element results show that the insertion of hole discontinuities decreases the initial peak force. Group B_H50, in which two hole discontinuity pairs are close enough to simultaneously induce the initial lobe, produced the largest reduction in initial peak force from that of a plain tube. This indicates that the mode of crush initiation has a large effect on the initial peak force experienced. The crush force efficiency, stroke efficiency, and amount of energy absorbed remained relatively unchanged for an increase in the number of holes and across different groups. The lobes formed in group B_HEq specimens are irregular and dependent on the prevailing hole spacing in the specimen. The finite element simulation satisfactorily predicts the final collapse shape of the tubes but overestimates the crushed distances.

Nomenclature

C = width of side of a square tube
D = Cowper–Symonds coefficient, 844 s^{-1}
E_a = energy absorbed
h = initial height of mass
l = half the wavelength of a single lobe
L = initial length of tube
m = mass
q = strain-rate hardening exponent, 2.207
P_i = initial peak force
P_m = mean crushing force
t = wall thickness
V = impact velocity
δ = crushed distance (stroke)
ε_t = true plastic strain
σ_t = true stress
σ_y = yield stress
ν = Poisson's ratio

References

[1] Bodlani, S. B., Yuen, S. Chung Kim, and Nurick, G. N., 2009, "The Energy Absorption Characteristics of Square Mild Steel Tubes With Multiple Induced Circular Hole Discontinuities—Part I: Experiments." ASME J. Appl. Mech., **76**, p. 041012.

[2] Abramowicz, W., 2003, "Thin-Walled Structures as Impact Energy Absorbers," Thin-Walled Struct., **41**, pp. 91–107.

[3] Yuen, S. Chung Kim, and Nurick, G. N., 2008, "The Energy Absorbing Characteristics of Tubular Structures With Geometric and Material Modifications: An Overview," Appl. Mech. Rev., **61**(2), p. 020802.

[4] Montanini, R., Belingardi, G., and Vadori, R., 1997, "Dynamic Axial Crushing of Triggered Aluminium Thin-Walled Columns," 30th International Symposium on Automotive Technology & Automation, Florence, Italy, Jun. 16–19, pp. 437–444.

[5] Otubushin, A., 1998, "Detailed Validation of a Non-Linear Finite Element Code Using Dynamic Axial Crushing of a Square Tube," Int. J. Impact Eng., **21**(5), pp. 349–368.

[6] Lee, S., Hahn, C., Rhee, M., and Oh, J., 1999, "Effect of Triggering on the Energy Absorption Capacity of Axially Compressed Aluminum Tubes," Mater. Des., **20**(1), pp. 31–40.

[7] Langseth, M., Hopperstad, O. S., and Berstad, T., 1999, "Crashworthiness of Aluminium Extrusions: Validation of Numerical Simulation, Effect of Mass Ratio and Impact Velocity," Int. J. Impact Eng., **22**, pp. 829–854.

[8] Tarigopula, V., Langseth, M., Hopperstad, O. S., and Clausen, A. H., 2005, "An Experimental and Numerical Study of Energy Absorption in Thin-Walled High-Strength Steel Sections," Int. J. Impact Eng., **32**(5), pp. 847–882.

[9] Jones, N., 1989, *Structural Impact*, Cambridge University Press, Cambridge.

[10] Marais, S. T., Tait, R. B., Cloete, T. J., and Nurick, G. N., 2004, "Material Testing at High Strain Rate Using the Split Hopkinson Pressure Bar," Latin American Journal of Solids and Structures, **1**, pp. 319–338.

[11] Masui, T., Nunokawa, T., and Hiramatsa, T., 1987, "Shape Correction of Hot Rolled Steel Using an On Line Leveler," J. Jpn. Soc. Technol. Plast., **28**(312), pp. 81–87.

[12] Marshall, N. S., and Nurick, G. N., 1998, "The Effect of Induced Deformations on the Formation of the First Lobe of Symmetric Progressive Buckling of Thin Walled Square Tubes," *Structures Under Shock and Impact* (SUSI 98), Thessaloniki, Greece, Jun. 24–26, N. Jones, D. G. Talaslidis, C. A. Brebbia, and G. D. Manolis, eds., pp. 155–168.

[13] Langseth, M., and Hopperstad, O. S., 1996, "Static and Dynamic Axial Crushing of Square Thin-Walled Aluminium Extrusions," Int. J. Impact Eng., **18**(7–8), pp. 949–968.

The Effect of Fiber Diameter Distribution on the Elasticity of a Fiber Mass

Mårten Alkhagen[1]

Staffan Toll

Department of Applied Mechanics,
Chalmers University of Technology,
SE-41296 Gothenburg, Sweden;
School of Textiles,
University College of Borås,
SE-50190 Borås, Sweden

A random mass of loose fibers interacting by fiber-fiber contact is considered. As proposed in a previous paper, the elastic response is modeled based on the statistical mechanics of bending and torsion of fiber segments between fiber-fiber contact points. Presently we show how the statistical approach can be used to account for a distribution of fiber diameters rather than just a single diameter. The resulting expression has the same form and the same set of parameters as its single-diameter counterpart, except for two dimensionless reduction factors, which depend on the fiber diameter distribution only and reduce to unity for monodisperse fibers. Uniaxial compressibility experiments are performed on several materials with different bimodal fiber diameter distributions and are compared to model predictions. Even though no additional parameters were introduced to model the effect of mixed fiber diameters, the behavior is accurately predicted. Notably, the effect of the nonuniform fiber diameter is strong: A mixture of two fiber diameters differing by a factor of 2 can reduce the response by an order of magnitude, compared to the case of uniform diameter. [DOI: 10.1115/1.2966178]

Introduction

Fiber masses occur in a variety of technologies. A mechanically demanding application is press fabrics, which are used in the press dewatering section of paper machines. The top layer of a press fabric consists of a carded web of nylon fibers, whose task is to resist the high pressure in the press nip while retaining a sufficient porosity to let the water escape at a high velocity.

The effort to theoretically estimate the compressibility of loose fiber masses was pioneered by van Wyk [1], who derived a simple power-law relation between applied pressure and fiber volume fraction. In his derivations he made several simplifications, e.g., fiber bending is the only deformation mechanism, the fibers are randomly orientated throughout the deformation, and all fibers have the same size, altogether leading to an overstiff model. After the work of van Wyk there have been many attempts to improve the accuracy of the micromechanical approach; Corte and Kallmes [2], Komori and Makishima [3], Pan [4], and Komori and Itoh [5,6] mainly focused on deriving a more compliant model compared to van Wyk's by incorporating orientational effects, curved fiber segments, and steric hindrance. More recently, Beil and Roberts [7,8] have performed detailed fiber-level simulations of uniaxial compression. In a previous paper [9] we developed a constitutive theory allowing for more general deformation paths, by letting contacts form and break continually. The theory was based on the independent rotation of approximately straight fiber segments and the neglect of slip at contacts. The latter restricts the theory to predominantly compressive deformations. Because the theory allows for multiple deformation mechanisms, such as fiber torsion and fiber bending, it captures a wider range of fiber volume fractions compared to previous models.

All the above theories have, however, been limited to uniform fiber diameter distributions, i.e., all fibers were assumed to have the same diameter [4,5,9–15]. Yet modern press fabrics are often constructed with a tailored mixture of fiber diameters to achieve maximum control of compressibility and fluid permeability.

The present work generalizes our previous theory [9] to incorporate the effects of a nonuniform fiber diameter. To this end we will need to reconsider the analysis of the contact compliances, the structure tensors, and the geometrical properties with regard to the fiber diameter distribution. This is done without introducing any new adjustable parameters. Finally, we propose a quantification of the influence of mixed fiber diameters on the compressibility of the fiber mass in terms of diameter distribution reduction factors. The theoretical outcome of these reduction factors is compared to compressive experiments performed on several fiber masses with different fiber diameter distributions.

Constitutive Theory

In Ref. [9] we derived a constitutive relationship for the stress response to mainly compressive deformations. The most important and fundamental steps of the derivation will be recapitulated here. We start from the macroscopic Cauchy stress

$$\boldsymbol{\Sigma} = \frac{1}{V}\mathrm{sym}\sum \mathbf{pr} \tag{1}$$

where $\mathbf{p}$ is the contact force, $\mathbf{r}$ is the position of a contact point, and the summation is carried out over all contact points on all particles in a representative volume V. The stress tensor must be symmetric on account of the angular momentum balance. Each contact gives rise to two contact points, one on each particle, with the associated contact force and normal $(\mathbf{p},\mathbf{n})$ and $(-\mathbf{p},-\mathbf{n})$, respectively. The contact vector $\mathbf{r}$ is referred to a local origin at the centroid of a fiber segment whose length corresponds to the fiber crimp spacing $2b$, as shown in Fig. 1.

If we consider a contact between two particles defined by the basis vectors, as illustrated in Fig. 2, it is understood that the dynamics of the contact will depend on the geometric configuration of several other contacts. Since it is practically impossible to know the complete configuration we have to describe the contact by a limited number of contact variables. In our previous article [9], we used the axis orientations $\mathbf{e}$, $\mathbf{e}'$ and the position vectors $\mathbf{r}$, $\mathbf{r}'$ of the two contacting fibers, respectively, as contact variables. Hence the contact force $\mathbf{p}$ is considered as a stochastic function of those variables. Next, we rewrote Eq. (1) as $\boldsymbol{\Sigma}=N\,\mathrm{sym}\langle\bar{\mathbf{p}}\mathbf{r}\rangle_c$, where $\bar{\mathbf{p}}$ is the conditional expectation of $\mathbf{p}$ given the contact variables $(\mathbf{e},\mathbf{e}',\mathbf{r},\mathbf{r}')$, N is the number of contacts per unit volume

[1]Present address: SCA Hygiene Products AB, R&D Personal Care, SE-40503 Gothenburg, Sweden.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 10, 2007; final manuscript received May 30, 2008; published online April 27, 2009. Review conducted by Anthony Waas.

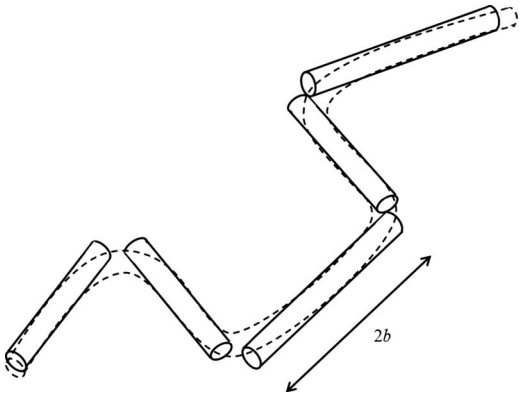


Fig. 1 Segmentation of a fiber, where 2*b* is the fiber crimp spacing

(defined such that each contact point is counted twice, once for each particle involved), and the angle brackets $\langle\cdot\rangle_c$ denote an average among all contacts. Then time differentiation of the stress yields

$$\dot{\boldsymbol{\Sigma}} + (\boldsymbol{\delta}:\mathbf{L})\boldsymbol{\Sigma} = N\,\mathrm{sym}(\langle\bar{\mathbf{p}}\mathbf{r}\rangle_c \cdot \mathbf{L}^t + \langle\bar{\mathbf{p}}\dot{r}\rangle_c) \tag{2}$$

where $\mathbf{L}=(\nabla\mathbf{V})^t$ is the macroscopic velocity gradient and $\mathbf{V}$ is the macroscopic velocity field. To construct a model of a fiber mass we assumed the expectation of the convected force rate to be

$$\bar{\mathbf{p}} = (\overline{\dot{p}_n}\mathbf{n} + \overline{\dot{p}_e}\mathbf{e} + \overline{\dot{p}_\theta}\boldsymbol{\theta}) + \mathbf{L}\cdot\bar{\mathbf{p}} \tag{3}$$

where $\overline{\dot{p}_n}=\overline{\dot{r}_n}/\overline{\langle s_n\rangle_c}$, $\overline{\dot{p}_e}=\overline{\dot{r}_e}/\overline{\langle s_e\rangle_c}$, $\overline{\dot{p}_\theta}=\overline{\dot{r}_\theta}/\overline{\langle s_\theta\rangle_c}$, and the compliances are defined as $s_n=\partial r_n/\partial p_n$, $s_e=\partial r_e/\partial p_e$, and $s_\theta=\partial r_\theta/\partial p_\theta$.

Finally, we obtained the fiber mass model

$$\overset{\Delta}{\boldsymbol{\Sigma}} = N\left[\frac{\langle r_n^2\rangle_c}{\overline{\langle s_n\rangle_c}}\langle\mathbf{nnnn}\rangle_c + \frac{\langle r_e^2\rangle_c}{\overline{\langle s_e\rangle_c}}\langle\mathbf{eeee}\rangle_c + \frac{\langle r_\theta^2\rangle_c}{\overline{\langle s_\theta\rangle_c}}\langle\boldsymbol{\theta\theta\theta\theta}\rangle_c\right]:\mathbf{L} \tag{4}$$

Here N is the number of contacts per unit volume, $\langle\mathbf{nnnn}\rangle_c$, $\langle\mathbf{eeee}\rangle_c$, and $\langle\boldsymbol{\theta\theta\theta\theta}\rangle_c$ are structure tensors, sensitive to the change in orientation of the fiber mass during deformation, and

$$\overset{\Delta}{\boldsymbol{\Sigma}} \equiv \dot{\boldsymbol{\Sigma}} + (\boldsymbol{\delta}:\mathbf{L})\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\cdot\mathbf{L}^t - \mathbf{L}\cdot\boldsymbol{\Sigma} \tag{5}$$

is known as the Truesdell stress rate. Thus far, no assumptions about the distribution of fiber diameters have been made. The previous theory proceeded by deriving the various averages in Eq. (4) under the assumption that the fiber diameter was constant. In the following we shall rederive these quantities for a more general fiber diameter distribution.

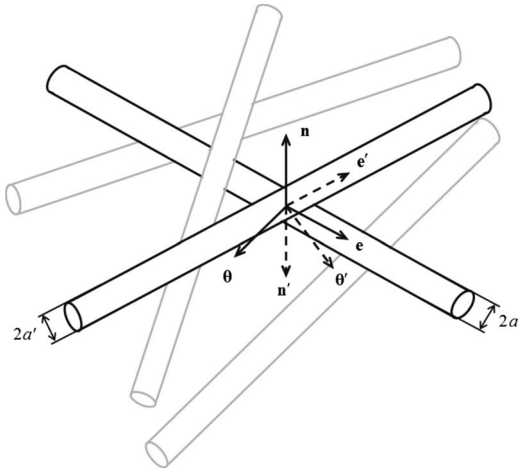


Fig. 2 Schematic of a contact point with the fiber diameters and the local basis vectors indicated. The primes refer to the contacting fiber.

Contact Point Distribution and Structure Tensors. In our previous article [9] we considered a mass of fibers possessing uniform properties (constant radius and Young's modulus). Here we will generalize to an arbitrary fiber radius distribution. To do this, we will need to take appropriate averages of the different parameters in Eq. (4). The fiber content will be characterized by a length fraction of fibers L, which is the total fiber length per unit volume. Furthermore, we consider an orientation distribution $\Psi_l(\mathbf{e})$ and radius distribution $\Psi_l(a)$ by length and assume that these distributions are mutually independent. In analyzing the distribution of fiber-to-fiber contacts, we make the assumption that the distribution of contacts between fibers is equivalent to the distribution of intersections of interpenetrable "phantom" fibers with the same distribution of orientations and radii. We introduce a phantom fiber with orientation $\mathbf{e}$ and radius a, referred to as a test fiber. Then a segment of an arbitrary phantom fiber with orientation $\mathbf{e}'$, radius a', and length Δl intersects a given segment of the test fiber only if its centerline intersects a surface $2(a+a')\Delta l$. Equivalently, the center of the phantom fiber segment must lie within a volume of the size $2(a+a')|\mathbf{e}\times\mathbf{e}'|\Delta l\Delta l'$. Thus, the number of phantom fibers intersecting the test fiber in the orientation-radius interval $d\mathbf{e}'da'$, per unit length of the test fiber, is given by

$$dn_l(\mathbf{e},a,\mathbf{e}',a') = 2L(a+a')|\mathbf{e}\times\mathbf{e}'|\Psi_l(\mathbf{e}')\Psi_l(a')d\mathbf{e}'da' \tag{6}$$

Integration over all possible fiber orientations and radii of the intersecting fibers, and using the assumption that orientations and radii are uncorrelated, gives the expected number of *contacts* per unit length of the test fiber, characterized by $\mathbf{e}$ and a as follows:

$$\begin{aligned} n_l(\mathbf{e},a) &= 2L\iint (a+a')|\mathbf{e}\times\mathbf{e}'|\Psi_l(\mathbf{e}')\Psi_l(a')d\mathbf{e}'da' \\ &= 2L\langle(a+a')|\mathbf{e}\times\mathbf{e}'|\rangle_l \\ &= 2L(a+\langle a\rangle_l)\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l \end{aligned} \tag{7}$$

where the average by length is defined as

$$\langle\cdot\rangle_l = \iint(\cdot)\Psi_l(\mathbf{e})\Psi_l(a)d\mathbf{e}da \tag{8}$$

The average number of contact points per unit volume (defined such that each contact is counted twice, once for each particle involved) is given by

$$N = L\langle n_l\rangle_l = 4L^2\langle a\rangle_l\langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l\rangle_l \tag{9}$$

where the double average by length is defined as

$$\langle\langle\cdot\rangle_l\rangle_l = \iint\iint(\cdot)\Psi_l(\mathbf{e})\Psi_l(\mathbf{e}')\Psi_l(a)\Psi_l(a')d\mathbf{e}d\mathbf{e}'dada' \tag{10}$$

Let us now consider the average of any given quantity $(\cdot)$ taken over all contacts. This average is derived by summation over all contacts and division by the total number of contacts as follows:

$$\langle\cdot\rangle_c = \frac{L}{N}\langle(\cdot)n_l\rangle_l$$

or by inserting Eqs. (7) and (8) as follows:

$$\begin{aligned} \langle\cdot\rangle_c = \frac{2L^2}{N}\iint\iint(\cdot)(a+a')|\mathbf{e}\times\mathbf{e}'| \\ \times\Psi_l(\mathbf{e})\Psi_l(\mathbf{e}')\Psi_l(a)\Psi_l(a')d\mathbf{e}d\mathbf{e}'dada' \end{aligned} \tag{11}$$

The average among contacts can be rewritten in terms of averages by length, by introducing Eqs. (9) and (10) into Eq. (11), as follows:

$$\langle\cdot\rangle_c = \frac{\langle\langle(\cdot)(a+a')|\mathbf{e}\times\mathbf{e}'|\rangle_l\rangle_l}{2\langle a\rangle_l\langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l\rangle_l} \tag{12}$$

In particular, if a function, $q(\mathbf{e},\mathbf{e}')$ depends on the orientations $\mathbf{e}$ and $\mathbf{e}'$ only, then the average among contacts is

$$\langle q(\mathbf{e},\mathbf{e}')\rangle_c = \frac{\langle\langle q(\mathbf{e},\mathbf{e}')|\mathbf{e}\times\mathbf{e}'|\rangle_l\rangle_l}{\langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l\rangle_l} \tag{13}$$

and, equivalently, for a function $p(a,a')$,

$$\langle p(a,a')\rangle_c = \frac{\langle\langle p(a,a')(a+a')\rangle_l\rangle_l}{2\langle a\rangle_l} \tag{14}$$

Since the distributions of a and $\mathbf{e}$ are independent, the fourth moment of the fiber axis vector $\mathbf{e}$ is, by combining Eqs. (10) and (13),

$$\langle\mathbf{eeee}\rangle_c = \frac{1}{f}\iint \mathbf{eeee}|\mathbf{e}\times\mathbf{e}'|\Psi_l(\mathbf{e})\Psi_l(\mathbf{e}')d\mathbf{e}d\mathbf{e}' \tag{15}$$

where

$$f = \langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l\rangle_l = \iint |\mathbf{e}\times\mathbf{e}'|\Psi_l(\mathbf{e})\Psi_l(\mathbf{e}')d\mathbf{e}d\mathbf{e}' \tag{16}$$

Similarly the fourth order moment of the contact normal, defined as $\mathbf{n}=\pm(\mathbf{e}\times\mathbf{e}')/|\mathbf{e}\times\mathbf{e}'|$, is

$$\langle\mathbf{nnnn}\rangle_c = \frac{1}{f}\iint \frac{(\mathbf{e}\times\mathbf{e}')^4}{|\mathbf{e}\times\mathbf{e}'|^3}\Psi_l(\mathbf{e})\Psi_l(\mathbf{e}')d\mathbf{e}d\mathbf{e}' \tag{17}$$

The structure tensors, Eqs. (15) and (17), are identical to the structure tensors found in our previous article [9]. This was expected since all fiber fractions (defined by a) are assumed to have the same orientation distribution. The function f relates to the overall probability of fibers crossing one another; it may be used as a scalar measure of the degree of alignment: $f=0$ for parallel fibers, $f=\pi/4$ for 3D random fibers, and $f=2/\pi$ for planar random fibers.

The orientation distribution will be modeled in the same way as in our previous article [9] by assuming affine rotation of the fiber axes as follows:

$$\mathbf{e} = \frac{\mathbf{F}\cdot\mathbf{e}_r}{\sqrt{\mathbf{e}_r\cdot\mathbf{F}^t\cdot\mathbf{F}\cdot\mathbf{e}_r}} \tag{18}$$

where $\mathbf{e}_r$ is the orientation of a fiber segment in a reference configuration (for example 3D random orientation) and $\mathbf{F}=\partial\mathbf{X}/\partial\mathbf{X}_r$ is the overall deformation gradient with respect to the reference configuration.

Compliances. In our previous paper [9], we found that in uniaxial compression the θ-component in Eq. (4) will be negligible compared to the $\mathbf{e}$-component. Thus, from hereon, the θ-component will be omitted. Another, and seemingly necessary, assumption is that the parameters $\bar{s}_n$, $\bar{s}_e$, r_n, and r_e are all approximately independent of $\mathbf{n}$ and $\mathbf{e}$ (this is certainly true for random, planar random, and unidirectional orientations). To estimate the compliance parameters we identify two sources of contact force: fiber bending and fiber torsion. The contacts are not allowed to slip, i.e., frictional dissipation is not included. As the fiber segments are assumed straight, the position of a contact point relative to the fiber segment centroid in the normal direction is simply

$$r_n \approx a \tag{19}$$

In the axial direction the contact position will be distributed in some way. Most probably, the contacts along a segment are randomly distributed and thus,

$$r_e \approx xb \tag{20}$$

where b is half of the crimp spacing and x is a stochastic variable randomly distributed in the interval $0\leqslant x\leqslant 1$. The crimp spacing in these materials scales approximately with the fiber radius, cf. Ref. [9],

$$b \approx \beta a \tag{21}$$

Notice that due to Eqs. (19)–(21), the fiber radii a, a' and the stochastic variables x, x' have now replaced $\mathbf{r}$, $\mathbf{r}'$ as contact variables.

Normal Compliance. The normal compliance is governed by the bending of the test fiber, which is described as a generic Euler beam

$$s_n = \frac{\partial r_n}{\partial p_n} = \frac{\lambda^3}{6\pi k_b E a^4} \tag{22}$$

where E is the fiber Young's modulus, a is the fiber radius, λ is the contact spacing (the distance between two consecutive contact points along the fiber), and k_b is a geometric factor, which is unity if the fiber beam is loaded at its midsection and fixed at its end-sections. Because a is a contact variable and k_b is a constant, only λ is allowed to vary, for given values of the contact variables $(\mathbf{e},\mathbf{e}',\mathbf{r},\mathbf{r}')$, and the average among contacts of the expectation of the normal compliance (Eq. (22)) is

$$\langle\bar{s}_n\rangle_c = \left\langle\frac{\overline{\lambda^3}}{6\pi k_b E a^4}\right\rangle_c \tag{23}$$

where the expectation of the contact spacing cubed, $\overline{\lambda^3}$, needs to be estimated. Again, using the assumption that the contacts along a fiber are randomly distributed, the distribution of the contact spacing λ is exponential given by

$$f(\lambda) = \frac{1}{\bar{\lambda}}\mathbf{e}^{-\lambda/\bar{\lambda}} \tag{24}$$

Thus, the expectation of the contact spacing cubed is obtained as

$$\overline{\lambda^3} = \bar{\lambda}^{-1}\int_0^\infty \lambda^3\mathbf{e}^{-\lambda/\bar{\lambda}}d\lambda = 6\bar{\lambda}^3 \tag{25}$$

Inserting this into Eq. (23) and considering that the expectation of the contact spacing is the reciprocal of contacts per unit length of the fibers, $\bar{\lambda}=n_l^{-1}$, yields

$$\langle\bar{s}_n\rangle_c = \left\langle\frac{1}{\pi k_b E a^4 n_l^3}\right\rangle_c = \frac{1}{8\pi k_b E L^3}\left\langle\frac{1}{a^4(a+\langle a\rangle_l)^3\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l^3}\right\rangle_c \tag{26}$$

where we have used Eq. (7) to replace n_l. Now, using Eqs. (13) and (14) to express the above equation in terms of averages by length yields

$$\langle\bar{s}_n\rangle_c = \frac{1}{16\pi k_b E L^3\langle a\rangle_l}\left\langle\frac{1}{a^4(a+\langle a\rangle_l)^2}\right\rangle_l\langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l^{-3}\rangle_l \tag{27}$$

where $\langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l^{-3}\rangle_l \approx f^{-3}$ (exact for random, planar random, and unidirectional orientations). Using Eqs. (14) and (19), the average among contacts of r_n^2 is

$$\langle r_n^2\rangle_c = \frac{\langle a^3 + a^2\langle a\rangle_l\rangle_l}{2\langle a\rangle_l} = \frac{\langle a^3\rangle_l + \langle a\rangle_l\langle a^2\rangle_l}{2\langle a\rangle_l} \tag{28}$$

Finally, using Eqs. (9), (27), and (28) to express the first right-hand side term in Eq. (4) yields

$$N\frac{\langle r_n^2\rangle_c}{\langle\bar{s}_n\rangle_c} = 32\pi k_b E L^5 f^4\frac{\langle a\rangle_l\langle a^3\rangle_l + \langle a\rangle_l^2\langle a^2\rangle_l}{\langle a^{-4}(a+\langle a\rangle_l)^{-2}\rangle_l} \tag{29}$$

In practice it is more direct to calculate the volumetric averages of a instead of the averages among contacts or units of fiber length.

Therefore we rewrite the above equation in terms of volumetric averages, and to do that we need to express the total fiber length per unit volume in terms of volumetric averages as follows:

$$L = \frac{1}{V}\sum_i l^i = \frac{\Phi}{V^F}\sum_i \frac{v^i}{\pi a^{i2}} = \frac{\Phi}{\pi}\langle a^{-2}\rangle_v \tag{30}$$

where V^F is the total fiber volume, l^i is the length and v^i the volume of a certain fiber fraction i, $\Phi = V^F/V$ is the total fiber volume fraction, and the angle bracket with subscript v denotes the volumetric average. Now the average of an arbitrary function of fiber radius $q(a)$ among units of fiber length is transformed into its volumetric counterpart as follows:

$$\langle q(a)\rangle_l = \frac{1}{VL}\sum_i l^i q(a) = \frac{\pi}{\langle a^{-2}\rangle_v}\frac{1}{V^F}\sum_i \frac{v^i}{\pi a^{i2}} q(a) = \frac{\langle a^{-2}q(a)\rangle_v}{\langle a^{-2}\rangle_v} \tag{31}$$

Finally, using Eqs. (30) and (31) to rewrite Eq. (29) yields

$$N\frac{\langle r_n^2\rangle_c}{\langle \bar{s}_n\rangle_c} = \frac{256k_bEf^4\Phi^5}{\pi^4}\Gamma_b \tag{32}$$

where Γ_b is the *reduction factor* describing the radius dependence

$$\Gamma_b = \frac{\langle a\rangle_v\langle a^{-1}\rangle_v\langle a^{-2}\rangle_v^4 + \langle a^{-1}\rangle_v^2\langle a^{-2}\rangle_v^3}{8\langle a^{-6}(a + \langle a^{-1}\rangle_v\langle a^{-2}\rangle_v^{-1})^{-2}\rangle_v} \tag{33}$$

When the radius distribution is uniform the reduction factor is unity.

Axial Compliance. The axial compliance is governed by the torsion and bending of the *contacting* fiber

$$s_e = \frac{\partial r_e}{\partial p_e} = s_e^t + s_e^b = \frac{16\lambda'}{3\pi k_t E a'^2} + \frac{\lambda'^3}{6\pi k_b E a'^4} \tag{34}$$

where the primed quantities refer to the contacting fiber [9] and k_t is unity for a simple torsion bar that is loaded by a couple ap_θ at its midsection and fixed at its end-sections. The average of the bending part of the axial compliance is equal to the averaged normal compliance, cf. Eqs. (22) and (27),

$$\langle \bar{s}_e^b\rangle_c = \left\langle \overline{\frac{\lambda'^3}{6\pi k_b E a'^4}} \right\rangle_c \equiv \left\langle \overline{\frac{\lambda^3}{6\pi k_b E a^4}} \right\rangle_c = \frac{1}{16\pi k_b E L^3 f^3 \langle a\rangle_l}\left\langle \frac{1}{a^4(a + \langle a\rangle_l)^2} \right\rangle_l \tag{35}$$

The torsional part of the axial compliance is

$$\langle \bar{s}_e^t\rangle_c = \left\langle \frac{16\bar{\lambda}'}{3\pi k_t E a'^2} \right\rangle_c \equiv \left\langle \frac{16\bar{\lambda}}{3\pi k_t E a^2} \right\rangle_c \tag{36}$$

which is rewritten using $\bar{\lambda} = n_l^{-1}$ together with Eq. (7) as follows:

$$\langle \bar{s}_e^t\rangle_c = \left\langle \frac{16}{3\pi k_t E a^2 n_l} \right\rangle_c = \frac{8}{3\pi k_t E L}\left\langle \frac{1}{a^2(a + \langle a\rangle_l)\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l} \right\rangle_c \tag{37}$$

To express the above equation in terms of averages by length, Eqs. (13) and (14) are used,

$$\langle \bar{s}_e^t\rangle_c = \frac{4}{3\pi k_t E L\langle a\rangle_l}\langle a^{-2}\rangle_l\langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l^{-1}\rangle_l \tag{38}$$

where $\langle\langle|\mathbf{e}\times\mathbf{e}'|\rangle_l^{-1}\rangle_l \approx f^{-1}$. Using Eqs. (20) and (21), the average of r_e^2 among contacts is

$$\langle r_e^2\rangle_c \approx \beta^2\langle x^2\rangle_c\langle a^2\rangle_c \tag{39}$$

where $\langle x^2\rangle_c = \frac{1}{3}$ and thus, by Eqs. (14) and (39),

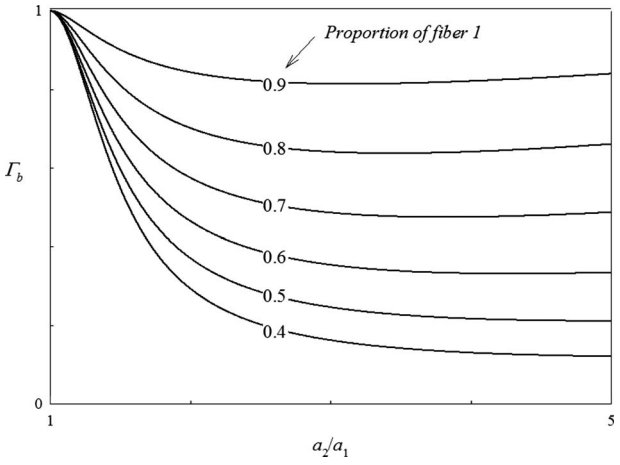


Fig. 3 The stiffness reduction versus the fiber radius ratio for various mixing proportions, according to Eq. (33)

$$\langle r_e^2\rangle_c \approx \frac{\beta^2\langle a^3 + a^2\langle a\rangle_l\rangle_l}{6\langle a\rangle_l} = \beta^2\frac{\langle a^3\rangle_l + \langle a\rangle_l\langle a^2\rangle_l}{6\langle a\rangle_l} \tag{40}$$

Now, using Eqs. (9), (38), and (40) to express the torsional part of the second right-hand side term in Eq. (4) yields

$$N\frac{\langle r_e^2\rangle_c}{\langle \bar{s}_e^t\rangle_c} = \frac{\pi k_t E L^3 f^2}{2}\frac{\beta^2\langle a^3\rangle_l + \langle a\rangle_l\langle a^2\rangle_l}{\langle a\rangle_l\langle a^{-2}\rangle_l} \tag{41}$$

and finally by virtue of Eqs. (30) and (31) this can be rewritten in terms of volumetric averages as follows:

$$N\frac{\langle r_e^2\rangle_c}{\langle \bar{s}_e^t\rangle_c} = \frac{k_t\beta^2}{\pi^2}Ef^2\Phi^3\Gamma_t \tag{42}$$

where Γ_t is the reduction factor describing the radius dependence of the torsional part of the axial stiffness

$$\Gamma_t = \frac{\langle a\rangle_v\langle a^{-1}\rangle_v\langle a^{-2}\rangle_v^2 + \langle a^{-1}\rangle_v^2\langle a^{-2}\rangle_v}{2\langle a^{-4}\rangle_v} \tag{43}$$

which of course becomes unity at uniform conditions. The bending part of the axial stiffness is, except for the factor $\beta^2/3$, identical to the normal stiffness, cf. Eq. (32),

$$N\frac{\langle r_e^2\rangle_c}{\langle \bar{s}_e^b\rangle_c} = \frac{256k_b\beta^2Ef^4\Phi^5}{3\pi^4}\Gamma_b \tag{44}$$

where the reduction factor Γ_b is given by Eq. (33).

Constitutive Equation. Introducing Eqs. (32), (42), and (44) into Eq. (4) and ignoring the θ-component yields the result

$$\overset{\Delta}{\boldsymbol{\Sigma}} = \left[\frac{256k_bE}{\pi^4}f^4\Phi^5\Gamma_b\langle\mathbf{nnnn}\rangle_c + \frac{E\beta^2}{\left[\frac{1}{\pi^2}k_tf^2\Phi^3\Gamma_t\right]^{-1} + \left[\frac{256}{3\pi^4}k_bf^4\Phi^5\Gamma_b\right]^{-1}}\langle\mathbf{eeee}\rangle_c\right]:\mathbf{L} \tag{45}$$

Figures 3 and 4 plot the reduction factors versus the fiber radius ratio, a_2/a_1, for a fiber mass consisting of discrete fractions of two fiber radii, a_1 and a_2. If the fiber radius ratio is unity then the reduction factors become unity, and as a_2/a_1 increases, the values of the reduction factors are reduced. Upon further increasing a_2/a_1, one would expect the reduction factors to decrease mono-

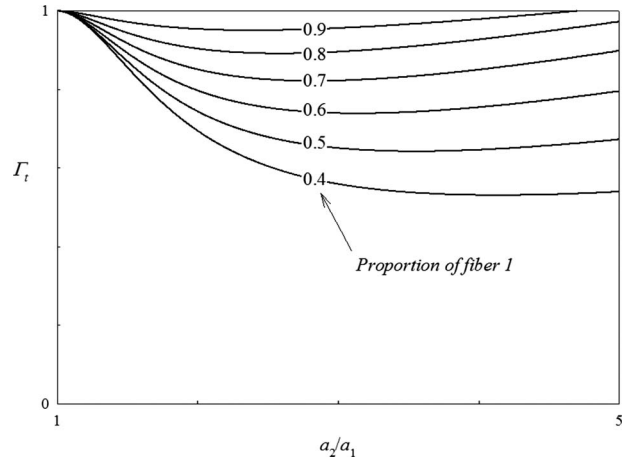


Fig. 4 The stiffness reduction versus the fiber radius ratio for various mixing proportions, according to Eq. (43)

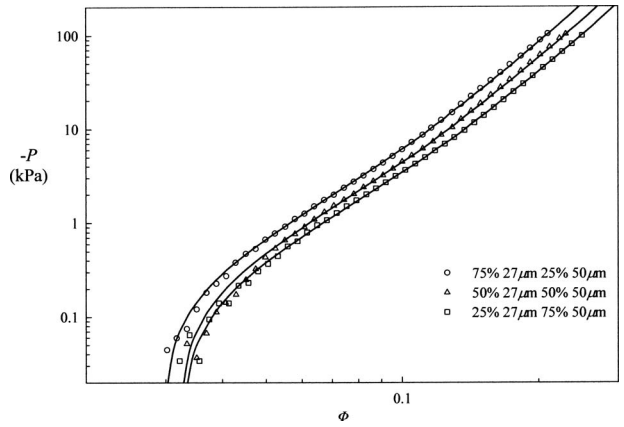


Fig. 5 Experimental data for fiber masses where $d_1=27$ μm and $d_2=50$ μm and the 1-fraction Φ_1 are 0.25, 0.50, and 0.75, respectively. The solid lines are model fits to determine the experimental reduction factors Γ_b^{expt} and Γ_t^{expt}.

tonically and level out as $a_2/a_1 \to \infty$. The model, however, predicts a minimum after which the factors increase toward infinity as $a_2/a_1 \to \infty$. This is of course an anomaly and hence the model will clearly not be applicable to other than moderate radius ratios.

Experiment

The compression measurements were performed using a parallel plate instrument [16] for soft compressible solids, where the stress is measured by means of a local stress transducer to remove any free edge effects. The load cell sensitivity is about 0.1 kPa in compression.

The tested materials were carded webs of fibers of a type used in the industrial press felt: crimped polyamide-6 fibers. Fibers of four different diameters, 27 μm, 35 μm, 50 μm, and 67 μm, were mixed in different proportions. The mixtures were then carded in a laboratory carding machine (CORMATEX CC/400) into a uniform web. The web was wound up on a reel and the resulting fiber mass was cut into appropriate lengths that were rotated at 90 deg and then carded again. This procedure was repeated three times to ensure good mixing. To verify an even dispersion of the fibers, differently colored fibers were carded and then examined by optical microscopy. Sixteen different compositions were prepared, as shown in Table 1. The samples were cut out of the web by means of a punching machine. Each sample was weighed and measured in order to determine the surface weight.

Table 1 Materials

a_1	a_2	a_1/a_2	Φ_1[a]
27	67	2.48	0.25
35	67	1.91	0.25
27	50	1.85	0.25
50	67	1.34	0.25
27	67	2.48	0.50
35	67	1.91	0.50
27	50	1.85	0.50
50	67	1.34	0.50
27	67	2.48	0.75
35	67	1.91	0.75
27	50	1.85	0.75
50	67	1.34	0.75
27	—	—	1
35	—	—	1
50	—	—	1
67	—	—	1

[a]Φ_1=volumetric proportion of fibers with radius a_1.

To erase the stress history from carding and storage, the samples were compressed to a stress of $P=-100$ kPa from which the samples were allowed to relax to a stress of $P=-80$ kPa. After this treatment, the samples were used in the experiments. Compression data were sampled at a constant displacement rate of 50 mm/min, a constant temperature of 21.0°C, and a relative humidity of 31%.

Results

Figure 5 presents some typical experimental data for the samples consisting of two fibers ($d_1=27$ μm and $d_2=50$ μm) mixed in various proportions. The scatter at low compressions is due to the limited resolution of the load cell. The repeatability and reproducibility of the measurements were consistent for all samples. The experimental values of the reduction factors, Γ_b^{expt} and Γ_t^{expt}, were determined by fitting the model to the data; the resulting curves are shown as solid lines in Fig. 5. The values of the dimensionless constants β, k_b, and k_t, and the ratio Φ_0/Φ_r were determined for materials of uniform fiber diameter; in Ref. [9], $k_b=2.2$ and $k_t=0.05$. The initial condition of the numerical computation was $P=0$ at $\Phi=\Phi_0$, defined as the volume fraction where a nonzero compressive stress was first registered.

As expected, and clearly seen in Fig. 5, the pressure response at a given volume fraction decreases as the fraction of the finer fibers increases. At higher volume fractions $\langle e_3 e_3 e_3 e_3 \rangle$ vanishes rapidly, so that the main contribution to the total stress comes from the first term of Eq. (45) and thus it was straightforward to resolve Γ_b^{expt} from the experimental data. Then the experimental magnitudes of the torsional reduction factor Γ_t^{expt} could be determined from the experimental data at intermediate volume fractions. Here, the second term of Eq. (45) dominates, and the Γ_t and Γ_b terms in the denominator are of similar magnitude. The results are plotted in Figs. 6 and 7, and compared to the theoretical estimates of Γ_b and Γ_t, Eqs. (33) and (43). As seen from the plot, the experimental values are in close agreement with the theoretical estimates.

Concluding Remarks

In a previous article [9] we derived a constitutive equation for fiber masses of uniform fiber diameter. To estimate the effect of a *nonuniform* fiber diameter distribution we have presently rederived the constitutive equation and, apart from allowing a nonuniform fiber diameter, the fundamental assumptions are identical to those of the previous model. The resulting constitutive equation differs from the previous one by the appearance of two reduction factors, which depend on the diameter distribution only. In the

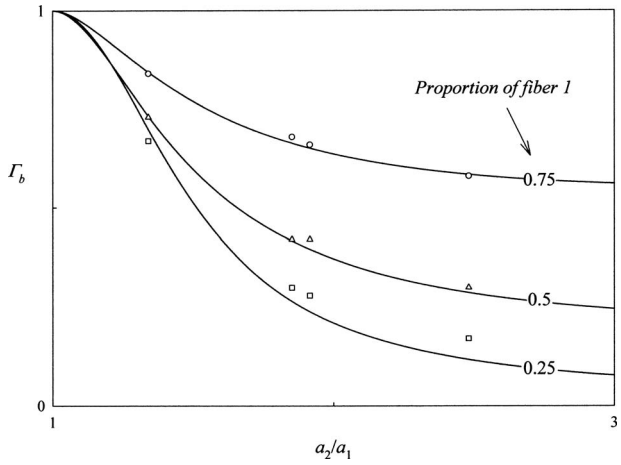


Fig. 6 Γ_b according to Eq. (33) (solid lines) compared to experimental data (the squares correspond to $\Phi_1=0.25$, the triangles to $\Phi_1=0.5$, and the circles to $\Phi_1=0.75$)

special case of a monodisperse fiber diameter, the reduction factors reduce to unity and the previous model is recovered. The reduction factors, Eqs. (33) and (43), imply that the response of a fiber mass is independent of the absolute scale of the fibers. Notice that this is not only true for monodisperse fibers but also for polydisperse ones as long as all fibers scale in the same way.

The experimental results are in excellent agreement with the theoretical ones, cf. Figs. 6 and 7. The reduction factors (especially Γ_b) for our most extreme compositions seem to be slightly underestimated by the model, as shown in Fig. 6, indicating that we are close to the limit of what the model can handle. The explanation for the deviation is the fundamental choice of relation between the expected force rate and the expected contact displacement rate by the *total average* of the expected compliance rather than by the local compliance of a particular contact. A result of this assumption is that when the proportion of the finer fibers *decreases*, their contribution to the total average will be exaggerated, leading to a too compliant deformation. Nevertheless, without this assumption, the stiffer contacts (e.g., coarser fibers) would contribute far more excessively to the total average leading to a severely overstiff model. Apparently, the aforementioned assumption is valid when the differences in the mixed fiber properties are moderate.

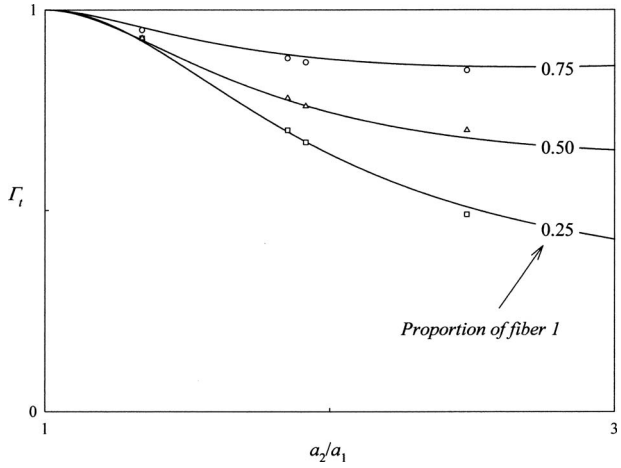


Fig. 7 Γ_t according to Eq. (43) (solid line) compared to experimental data (the squares correspond to $\Phi_1=0.25$, the triangles to $\Phi_1=0.5$, and the circles to $\Phi_1=0.75$)

However, at moderate diameter ratios, it seems as though the two effects discussed in the two previous paragraphs are either negligible or cancel each other out, and the main conclusion is that the present approach successfully predicts the reduction in the pressure response for a fiber mass consisting of two (or more) different fiber diameters compared to a fiber mass of uniform fiber diameter. These predictions are remarkably accurate, considering that no new adjustable parameters have been introduced.

It is finally noted that the error of neglecting the difference in fiber diameter is substantial. Judging from the bottom data in Fig. 6, it may lead to predictions that are off by an order of magnitude.

Acknowledgment

Albany International (AB) is acknowledged for supplying the fibers and important information on the materials used.

References

[1] van Wyk, C. M., 1946, "Note on the Compressibility of Wool," J. Text. Inst., **37**, pp. T285–T292.
[2] Corte, H., and Kallmes, O., 1962, *Formation and Structure of Paper*, 1st ed., Vol. 1, Technical Section of the British Papers and Board Makers Association, London.
[3] Komori, T., and Makishima, K., 1977, "Numbers of Fiber-to-Fiber Contacts in General Fiber Assemblies," Text. Res. J., **47**, pp. 13–17.
[4] Pan, N., 1993, "A Modified Analysis of the Microstructural Characteristics of General Fiber Assemblies," Text. Res. J., **63**, pp. 336–345.
[5] Komori, T., and Itoh, M., 1997, "Analyzing the Compressibility of a Random Fiber Mass Based on the Modified Theory of Fiber Contact," Text. Res. J., **67**, pp. 204–210.
[6] Komori, T., and Itoh, M., 1994, "A Modified Theory of Fiber Contact in General Fiber Assemblies," Text. Res. J., **64**, pp. 519–528.
[7] Beil, N. B., and Roberts, W. W. , Jr., 2002, "Modeling and Computer Simulation of the Compressional Behavior of Fiber Assemblies, Part I: Comparison to van Wyk's Theory," Text. Res. J., **72**, pp. 341–351.
[8] Beil, N. B., and Roberts, W. W. Jr., 2002, "Modeling and Computer Simulation of the Compressional Behavior of Fiber Assemblies, Part II: Hysteresis, Crimp, and Orientational Effects," Text. Res. J., **72**, pp. 375–382.
[9] Alkhagen, M., and Toll, S., 2007, "Micromechanics of a Compressed Fiber Mass," Trans. ASME, J. Appl. Mech., **74**, pp. 723–731.
[10] Komori, T., and Itoh, M., 1991, "A New Approach to the Theory of the Compression of Fiber Assemblies," Text. Res. J., **61**, pp. 420–428.
[11] Komori, T., and Itoh, M., 1991, "Theory of the General Deformation of Fiber Assemblies," Text. Res. J., **61**, pp. 588–594.
[12] Lee, D. H., and Lee, J. K., 1985, "Initial Compressional Behaviour of Fibre Assembly," *Objective Measurements: Application to Product Design and Process Control*, S. Kawabata, R. Postle, and M. Niwa, eds., The Textile Machinery Society of Japan, Osaka, pp. 613–622.
[13] Neckár, B., 1997, "Compression and Packing Density of Fibrous Assemblies," Text. Res. J., **67**, pp. 123–130.
[14] Toll, S., 1998, "Packing Mechanics of Fiber Reinforcements," Polym. Eng. Sci., **38**, pp. 1337–1350.
[15] Toll, S., and Manson, J.-A. E., 1995, "Elastic Compression of a Fiber Network," Trans. ASME, J. Appl. Mech., **62**, pp. 223–226.
[16] Alkhagen, M., and Toll, S., 2002, "A Triaxial Rheometer for Soft Compressible Solids," J. Rheol., **46**, pp. 31–47.

Exact Solutions for the Generalized Euler's Problem

Q. S. Li
Department of Building and Construction,
City University of Hong Kong,
Kowloon, Hong Kong, P.R.C.

This paper is concerned with buckling analysis of a nonuniform column with classical/nonclassical boundary conditions and subjected to a concentrated axial force and distributed variable axial loading, namely, the generalized Euler's problem. Exact solutions are derived for the buckling problem of nonuniform columns with variable flexural stiffness and under distributed variable axial loading expressed in terms of polynomial functions. Then, more complicated buckling problems are considered such as that the distribution of flexural stiffness of a nonuniform column is an arbitrary function, and the distribution of axial loading acting on the column is expressed as a functional relation with the distribution of flexural stiffness and vice versa. The governing equation for such problems is reduced to Bessel equations and other solvable equations for seven cases by means of functional transformations. A class of exact solutions for the generalized Euler's problem involved a nonuniform column subjected to an axial concentrated force and axially distributed variable loading is obtained herein for the first time in literature.
[DOI: 10.1115/1.2937151]

Keywords: buckling, structural stability, exact solution, nonuniform column

1 Introduction

Structural components with varying cross sections are commonly used in buildings, bridges, mechanical and aeronautical engineering in an effort to achieve a better distribution of strength and weight of structures, and sometimes to satisfy architectural and functional requirements. The stability of such structural members subjected to compressive forces is a subject of considerable scientific and practical interest that has been studied extensively, and is still receiving attention in literature. Since previous studies on this topic are quite extensive, only selected research works on finding exact solutions or closed-form solutions for the buckling of nonuniform columns are cited further.

Euler [1] pioneered the study of buckling of a column subjected to a concentrated force or under its own weight. Since then, exact buckling load solutions for simple cases, such as a nonuniform column with specific nonuniformity under a concentrated axial load at its top or end, columns with several types of variable cross sections subjected to uniformly distributed axial loading or a cuneiform column under its own weight were obtained by several researchers including Greenhill [2], Morley [3], Dinnik [4,5], Karman and Biot [6], Timosenko and Gere [7], and others. Gere and Carter [8] derived exact buckling solutions for several special types of tapered columns with simple boundary conditions in terms of Bessel functions. Ermopoulos [9] studied the buckling of tapered bars, axially compressed by concentrated loads applied at various locations along their axes. The law of stiffness variation chosen in his study was a power function with second order. Williams and Aston [10] provided several curves that enable the buckling loads to be found for tapered columns with a uniform axial force applied over all their length. Arbabi and Li [11] presented a semianalytical procedure for the buckling of elastic columns with step varying thickness. Siginer [12] investigated the stability of a column whose flexural rigidity follows a continuous linear variation along the column. The exact solutions for buckling of a nonuniform column under a concentrated force and variably distributed axial loading (as shown in Fig. 1) were obtained by Li et al. [13,14] for the following two cases:

$$\text{Case 1:}\quad K(x)=K(0)\left(1+\beta\frac{x}{L}\right)^{\gamma},\quad N(x)=N_0\left(1+\beta\frac{x}{L}\right)^{c}$$

$$\text{Case 2:}\quad K(x)=K(0)e^{\beta(x/L)},\quad N(x)=N(0)e^{c(x/L)}$$

where $K(x)$, $N(x)$, and L are the flexural stiffness, axial force, and length of the column, respectively, and β, γ, and c are the parameters relating to the variations of the stiffness and axial loading.

Recently, Elishakoff and co-worker [15–19] obtained several closed-form solutions for buckling of inhomogeneous columns, which considered a column with variable stiffness under axially distributed loading. In these studies, the buckling mode of a column was given first; then, the variation of the flexural stiffness along the axis of the column was determined so that the column had the prescribed buckling mode. The variations of stiffness, axial loading, and buckling mode were all selected to be polynomial functions.

In the first part of this paper, exact solutions are obtained for the buckling problem of a nonuniform column with variable flexural stiffness and distributed variable axial loading all described in terms of polynomial functions. In the present buckling analysis, the concerned nonuniform column is also subjected to a concentrated axial force at its top and various kinds of boundary conditions including classical and nonclassical ones are taken into account. Furthermore, more complicated problems are considered in this study, such as that the distribution of flexural stiffness of a nonuniform column is an arbitrary function, and the distribution of axial loading acting on the column is expressed as a functional relation with the distribution of flexural stiffness and vice versa. The governing equation for such buckling problems is reduced to Bessel equations and other solvable equations for several cases by means of functional transformations. Exact solutions for the buckling problem of a nonuniform column subjected to an axial concentrated force at its top and axially distributed variable loading and with various classical/nonclassical boundary conditions are thus presented in this paper.

Euler's problem, namely, buckling of a uniform column subjected to a concentrated force or under its own weight, is generalized to buckling of a nonuniform column with classical/nonclassical boundary conditions and subjected to a concentrated axial force and axially distributed variable loading. It is referred as the generalized Euler's problem in this paper. Literature review

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 15, 2007; final manuscript received March 18, 2008; published online April 27, 2009. Review conducted by Edmundo Corona.

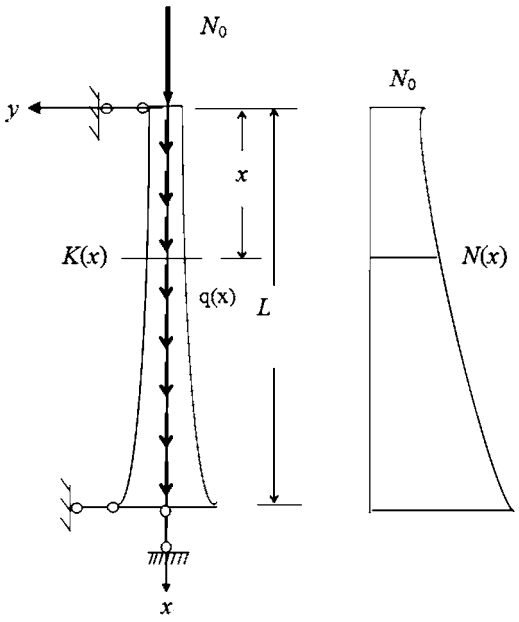


Fig. 1 A H-H column subjected to a concentrated axial force at its top and a distributed axial force

reveals that, generally, the previous studies on the buckling analysis of a nonuniform column under variably distributed loading have directed their investigations to special functions for describing the distributions of flexural stiffness and axial distributed forces to derive analytical solutions. A successful attempt is made herein to present the exact solutions for the generalized Euler's problem, which includes more complicated problems such as arbitrary distribution of flexural stiffness or axial forces. It is noted that some of the exact solutions presented in this paper are obtained for the first time in literature. Although in the absence of the exact solutions, the generalized Euler's problem can be solved using approximated methods or numerical techniques. However, the present exact solutions could provide adequate insight into the physics of the problem and can be easily implemented for structural stability analysis. Furthermore, the availability of the exact solutions can help in examining the accuracy of the approximated or numerical solutions. Therefore, it is always desirable to obtain the exact solutions to such an important problem.

2 Polynomial Solutions for the Generalized Euler's Problem

The governing differential equation for buckling of a nonuniform column subjected to a concentrated axial force N_0 at its top and a variably distributed axial loading $q(x)$ applied over all its length (as shown in Fig. 1) is expressed by

$$\frac{d^2}{d^2x}\left[K(x)\frac{d^2y(x)}{dx^2}\right]+\frac{d}{dx}\left[N(x)\frac{dy(x)}{dx}\right]=0 \tag{1}$$

where $K(x)=EI(x)$ is the flexural stiffness, $E=$Young's modulus, it may be a constant or a function of x, $I=$the moment of inertia, $N(x)=N_0+\int_0^x q(x)dx=$ axial force, $y(x)=$displacement, x is the axial coordinate, and the origin of the coordinate system is set at the top of the column, as shown in Fig. 1.

It is assumed that the variations of the stiffness, axial force, and displacement are described by the following polynomial functions:

$$K(\xi)=\sum_{i=0}^{m} b_i\xi^i \tag{2}$$

$$N(\xi)=N_0\sum_{0}^{n}\beta_i\xi^i \tag{3}$$

$$y(\xi)=\sum_{i=0}^{p} a_i\xi^i \tag{4}$$

where $\xi=x/L$ is a nondimensional axial coordinate and L is the length of the column. All β_i are known, and a_i may be determined from the boundary conditions of the column and it is now required to find the critical value of N_0.

Substituting Eqs. (2)–(4) into Eq. (1), one obtains

$$\frac{d^2}{d\xi^2}\left[\sum_{i=0}^{m}(b_i\xi^i)\frac{d^2}{d\xi^2}\sum_{i=0}^{p}(a_i\xi^i)\right]+N_0L^2\frac{d}{d\xi}\left[\sum_{i=0}^{n}(\beta_i\xi^i)\frac{d}{d\xi}\sum_{i=0}^{p}(a_i\xi^i)\right]=0 \tag{5}$$

This equation is valid for the variations of $K(x)$ and $N(x)$ expressed in terms of polynomial functions. It can be seen from the above equation that m, n, and p must satisfy the following relationship:

$$m+(p-2)-2=n+(p-1)-1$$

or

$$m=n+2 \tag{6}$$

In order to ensure the four boundary conditions at the two ends of a column be satisfied, the order of the displacement polynomial function must be greater or equal to 4.

To simplify the analysis, p is taken as 4 in this study, then Eq. (5) becomes

$$\sum_{i=0}^{n+2}[2i(i-1)a_2b_i\xi^{i-2}+6i(i+1)a_3b_i\xi^{i-1}+12(i+1)(i+2)a_4b_i\xi^i]+N_0L^2\sum_{0}^{n}[ia_1\beta_i\xi^{i-1}+2(i+1)a_2\beta_i\xi^i+3(i+2)a_3\beta_i\xi^{i+1}+4(i+3)a_4\beta_i\xi^{i+2}]=0 \tag{7}$$

To ensure Eq. (7) to be valid, all coefficients in front of ξ^i must vanish; hence, a set of equations is obtained as follows:

$$\text{For } \xi^0,\quad 4a_2b_2+12a_3b_1+24a_4b_0+N_0L^2(a_1\beta_1+2a_2\beta_0)=0 \tag{8}$$

$$\text{For } \xi^1,\quad 6a_2b_3+18a_3b_2+36a_4b_1+N_0L^2(a_1\beta_2+2a_2\beta_1+3a_3\beta_0)=0 \tag{9}$$

$$\text{For } \xi^i(2\le i\le n),\quad 2(i+1)a_2b_{i+2}+6(i+2)a_3b_{i+1}+12(i+2)a_4b_i+N_0L^2(a_1\beta_{n+1}+2a_2\beta_i+3a_3\beta_{i-1}+4a_4\beta_{i-2})=0 \tag{10}$$

$$\text{For } \xi^{n+1},\quad 6(n+3)a_3b_{n+2}+12(n+3)a_4b_{n+1}+N_0L^2(3a_3\beta_n+4a_4\beta_{n-1})=0 \tag{11}$$

$$\text{For } \xi^{n+2},\quad 3(n+4)b_{n+2}+N_0L^2\beta_n=0 \tag{12}$$

It is evident that there is a number of $(n+3)$ equations involved and a number of $(n+4)$ unknowns, which are solvable if an arbitrary unknown is given a value. If the arbitrary unknown coefficient is selected to be b_{n+2}, the critical value of the concentrated axial force is thus determined from Eq. (12) as

$$N_0 = -\frac{3(n+4)}{\beta_n L^2} b_{n+2} \tag{13}$$

It is observed from the above equation that the coefficient b_{n+2} must be a negative constant. In general, all β_i and L are given, only b_i $(i=0,1,2,\ldots,n+1)$ are needed to be determined. First of all, b_{n+1} is determined from Eq. (11) as

$$b_{n+1} = -\frac{a_3}{2a_4} b_{n+2} - \frac{N_0 L^2 (3a_3\beta_n + 4a_4\beta_{n-1})}{12(n+3)a_4} \tag{14}$$

After b_{n+1} is found, $b_i (2 \leqslant i \leqslant n)$ can be determined from Eq. (10) as

$$b_i = -\frac{a_3}{2a_4} b_{i+1} - \frac{a_2}{6a_4} b_{i+2} - \frac{N_0 L^2 (a_1\beta_{i+1} + 2a_2\beta_i + 3a_3\beta_{i-1} + 4a_4\beta_{i-2})}{12(i+2)a_4} \tag{15}$$

Using Eqs. (9) and (8), one obtains

$$b_1 = -\frac{a_2}{6a_4} b_3 - \frac{a_3}{2a_4} b_2 - \frac{N_0 L^2 (a_1\beta_2 + 2a_2\beta_1 + 3a_3\beta_0)}{36a_4} \tag{16}$$

$$b_0 = -\frac{a_2}{6a_4} b_2 - \frac{a_3}{2a_4} b_1 - \frac{N_0 L^2 (a_1\beta_1 + 2a_2\beta_0)}{24a_4} \tag{17}$$

Equations (13)–(17) are called the general solutions of the governing differential equation. The particular solutions for various supporting conditions can be obtained according to the formulas presented above and the specific boundary conditions as follows.

1. *Hinged-hinged (H-H) column*. The boundary conditions for this case (Fig. 1) can be written as

$$y(0) = 0, \quad y''(0) = 0 \tag{18a}$$

$$y(1) = 0, \quad y''(1) = 0 \tag{18b}$$

Applying Eqs. (18*a*) and (18*b*) results in

$$a_0 = 0, \quad a_1 = 1, \quad a_2 = 0, \quad a_3 = -2, \quad a_4 = 1 \tag{19}$$

i.e., $y(\xi)$ is given by

$$y(\xi) = \xi - 2\xi^3 + \xi^4 \tag{20}$$

$y(\xi)$ is the buckling mode shown in Fig. 2

Substituting Eq. (19) into Eqs. (14)–(17), one obtains

$$b_{n+1} = b_{n+2} - \frac{N_0 L^2 (-6\beta_n + 4\beta_{n-1})}{12(n+3)} \tag{21a}$$

$$b_i = b_{i+1} - \frac{N_0 L^2 (\beta_{i+1} - 6\beta_{i-1} + 4\beta_{i-2})}{12(i+2)} \tag{21b}$$

$$b_1 = b_2 - \frac{N_0 L^2 (\beta_2 - 6\beta_0)}{36} \tag{21c}$$

$$b_0 = b_1 - \frac{N_0 \beta_1 L^2}{24} \tag{21d}$$

2. *Clamped-free (C-F) column*. The boundary conditions for this case (shown in Fig. 3) can be written as

$$y''(0) = 0, \quad y'''(0) = 0 \tag{22a}$$

$$y(1) = 0, \quad y'(1) = 0 \tag{22b}$$

Similarly, we have

$$a_0 = 3, \quad a_1 = -4, \quad a_2 = 0, \quad a_3 = 0, \quad a_4 = 1 \tag{23}$$

i.e., $y(\xi)$ is given by

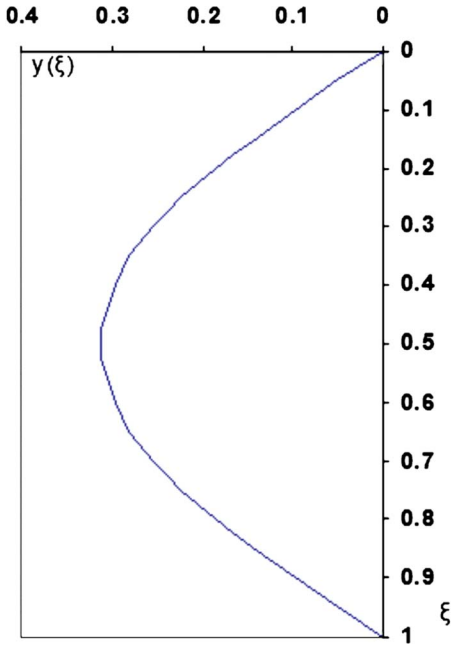


Fig. 2 The buckling mode of a H-H column

$$y(\xi) = 3 - 4\xi + \xi^4 \tag{24}$$

$y(\xi)$ is the buckling mode shown in Fig. 4.

Substituting Eq. (23) into Eqs. (14)–(17) results in

$$b_{n+1} = -\frac{N_0 L^2 \beta_{n-1}}{3(n+3)} \quad (n > 2) \tag{25a}$$

$$b_i = -\frac{N_0 L^2 (-\beta_{i+1} + \beta_{i-2})}{3(i+2)} \quad (2 \leqslant i \leqslant n) \tag{25b}$$

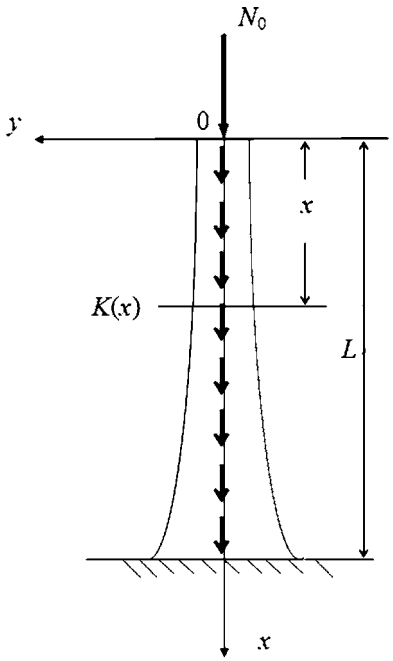


Fig. 3 A C-F column subjected to a concentrated axial force at its top and a distributed axial force

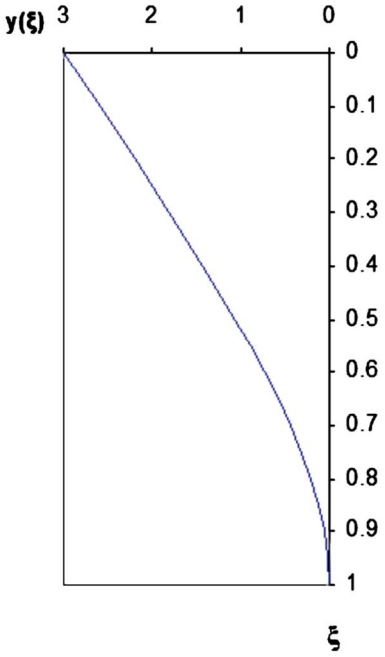


Fig. 4 The buckling mode of a C-F column

$$b_1 = \frac{N_0 L^2 \beta_2}{9} \tag{25c}$$

$$b_0 = \frac{N_0 L^2 \beta_1}{6} \tag{25d}$$

3. *Clamped-clamped (C-C) column*. The boundary conditions for this case are

$$y(0) = 0, \quad y'(0) = 0 \tag{26a}$$

$$y(1) = 0, \quad y'(1) = 0 \tag{26b}$$

Hence, the following results are obtained:

$$a_0 = 0, \quad a_1 = 0, \quad a_2 = 1, \quad a_3 = -2, \quad a_4 = 1 \tag{27}$$

i.e., $y(\xi)$ is given by

$$y(\xi) = \xi^2 - 2\xi^3 + \xi^4 \tag{28}$$

Substituting Eq. (27) into Eqs. (14)–(17) yields

$$b_{n+1} = b_{n+2} - \frac{N_0 L^2(-3\beta_n + 2\beta_{n-1})}{6(n+3)} \tag{29a}$$

$$b_i = b_{i+1} - \frac{b_{i+2}}{6} - \frac{N_0 L^2(\beta_i - 3\beta_{i-1} + 2\beta_{i-2})}{6(i+2)} \tag{29b}$$

$$b_1 = -\frac{b_3}{6} + b_2 - \frac{N_0 L^2(\beta_2 - 3\beta_0)}{18} \tag{29c}$$

$$b_0 = -\frac{b_2}{6} + b_1 - \frac{N_0 L^2 \beta_0}{12} \tag{29d}$$

4. *Camped-hinged (C-H) column*. Considering a column with a hinged end at $\xi=0$ and a clamped support at $\xi=1$, the boundary conditions for this case can be written as

$$y(0) = 0, \quad y''(0) = 0 \tag{30a}$$

$$y(1) = 0, \quad y'(1) = 0 \tag{30b}$$

Applying Eqs. (30*a*) and (30*b*) to Eq. (4) yields

$$a_0 = 0, \quad a_1 = 1, \quad a_2 = 0, \quad a_3 = -3, \quad a_4 = 2 \tag{31}$$

$$y(\xi) = \xi - 3\xi^3 + 2\xi^4 \tag{32}$$

Substituting Eq. (31) to Eqs. (14)–(17) results in

$$b_{n+1} = \frac{3}{4} b_{n+2} - \frac{N_0 L^2(-9\beta_n + 8\beta_{n-1})}{24(n+3)} \quad (n > 2) \tag{33a}$$

$$b_i = -\frac{3}{4} b_{i+1} - \frac{N_0 L^2(\beta_{i+1} - 9\beta_{i-1} + 8\beta_{i-2})}{24(i+2)} \quad (2 \leqslant i \leqslant n) \tag{33b}$$

$$b_1 = \frac{3}{4} b_2 - \frac{N_0 L^2(\beta_2 - 9\beta_0)}{72} \tag{33c}$$

$$b_0 = \frac{3}{4} b_1 - \frac{N_0 L^2 \beta_1}{48} \tag{33d}$$

5. *Hinged-guided (H-G) column*. A column with a hinged end at $\xi=0$ and a guided end at $\xi=1$ is considered here. The corresponding boundary conditions for this case are

$$y(0) = 0, \quad y''(0) = 0 \tag{34a}$$

$$y'(1) = 0, \quad y'''(1) = 0 \tag{34b}$$

Using Eqs. (4), (34*a*), and (34*b*), one obtains

$$a_0 = 0, \quad a_1 = 8, \quad a_2 = 0, \quad a_3 = -4, \quad a_4 = 1 \tag{35}$$

The buckling mode is thus described by

$$y(\xi) = 8\xi - 4\xi^3 + \xi^4 \tag{36}$$

Substituting Eq. (35) into Eqs. (14)–(17) yields

$$b_{n+1} = 2b_{n+2} - \frac{N_0 L^2(-3\beta_n + \beta_{n-1})}{3(n+3)} \tag{37a}$$

$$b_i = 2b_{i+1} - \frac{N_0 L^2(2\beta_{i+1} - 3\beta_{i-1} + \beta_{i-2})}{3(i+2)} \tag{37b}$$

$$b_1 = 2b_2 - \frac{N_0 L^2(2\beta_2 - 3\beta_0)}{9} \tag{37c}$$

$$b_0 = 2b_1 - \frac{N_0 L^2 \beta_1}{3} \tag{37d}$$

6. *Clamped-guided (C-G) column*. Next, considering a C-G column with a clamped end at $\xi=0$ and a guided end at $\xi=1$, the boundary conditions for this case are

$$y(0) = 0, \quad y'(0) = 0 \tag{38a}$$

$$y'(1) = 0, \quad y'''(1) = 0 \tag{38b}$$

Using Eqs. (4), (38*a*), and (38*b*) leads to

$$a_0 = 0, \quad a_1 = 0, \quad a_2 = 4, \quad a_3 = -4, \quad a_4 = 1 \tag{39}$$

The buckling mode is expressed by

$$y(\xi) = 4\xi^2 - 4\xi^3 + \xi^4 \tag{40}$$

Substituting Eq. (40) into Eqs. (14)–(17) results in

$$b_{n+1} = 2b_{n+2} - \frac{N_0 L^2(-3\beta_n + \beta_{n-1})}{3(n+3)} \tag{41a}$$

$$b_i = 2b_{i+1} - \frac{2}{3} b_{i+2} - \frac{N_0 L^2(2\beta_i - 3\beta_{i-1} + \beta_{i-2})}{3(i+2)} \tag{41b}$$

$$b_1 = -\frac{2}{3} b_3 + 2b_2 - \frac{N_0 L^2(2\beta_2 - 3\beta_0)}{9} \tag{41c}$$

$$b_0 = -\frac{2}{3}b_2 + 2b_1 - \frac{N_0 L^2 \beta_0}{3} \tag{41d}$$

7. *Spring-spring (S-S) column*. The boundary conditions for this case are

$$M(0) = -K_{\varphi 0} y'(0), \quad Q(0) = K_{u0} y(0) + N_0 y'(0) \tag{42a}$$

$$y''(0) = \mu_{\varphi 0} y'(0), \quad y'''(0) = -\mu_{u0} y(0) - (\mu_0 \mu_{\varphi 0} + \mu_{n0}) y'(0) \tag{42b}$$

$$M(1) = K_{\varphi 1} y'(1), \quad Q(1) = -K_{u1} y(1) + N_1 y'(1) \tag{42c}$$

$$y''(1) = -\mu_{\varphi 1} y'(1), \quad y'''(1) = \mu_{u1} y(1) - (\mu_1 \mu_{\varphi 1} - \mu_{n1}) y'(1) \tag{42d}$$

The definitions of the parameters involved in the above equations are given below. Applying these boundary conditions to Eq. (4) results in

$$a_0 = \frac{1}{|D|}\begin{vmatrix} 0 & -\mu_{\varphi 0} & 2 & 0 \\ 0 & a_{22} & 0 & 6 \\ c_3 & \mu_{\varphi 1} & a_{33} & a_{43} \\ c_4 & a_{24} & a_{34} & a_{44} \end{vmatrix}, \quad a_1 = \frac{1}{|D|}\begin{vmatrix} 0 & 0 & 2 & 0 \\ \mu_{u0} & 0 & 0 & 6 \\ 0 & c_3 & a_{33} & a_{43} \\ -\mu_{u1} & c_4 & a_{34} & a_{44} \end{vmatrix}$$

$$a_2 = \frac{1}{|D|}\begin{vmatrix} 0 & -\mu_{\varphi 0} & 0 & 0 \\ \mu_{u0} & a_{22} & 0 & 6 \\ 0 & \mu_{\varphi 1} & c_3 & a_{43} \\ -\mu_{u1} & c_{24} & c_4 & a_{44} \end{vmatrix}, \tag{43}$$

$$a_3 = \frac{1}{|D|}\begin{vmatrix} 0 & -\mu_{\varphi 0} & 2 & 0 \\ \mu_{u0} & a_{22} & 0 & 6 \\ 0 & \mu_{\varphi 1} & a_{33} & c_3 \\ -\mu_{u1} & a_{24} & a_{34} & c_4 \end{vmatrix}$$

where

$$\mu_0 = \frac{K'(0)}{K(0)}, \quad \mu_{\varphi 0} = \frac{K_{\varphi 0}}{K(0)}, \quad \mu_{n0} = \frac{N_0}{K(0)}, \quad \mu_1 = \frac{K'(1)}{K(1)}$$

$$\mu_{u0} = \frac{K_{u0}}{K(0)}, \quad \mu_{u1} = \frac{K_{u1}}{K(1)}, \quad \mu_{\varphi 1} = \frac{K_{\varphi 1}}{K(1)}, \quad \mu_{n1} = \frac{N_1}{K(1)}$$

$$a_{22} = \mu_0 \mu_{\varphi 0} + \mu_{n0}, \quad c_3 = -12 - 4\mu_{\varphi 1},$$

$$c_4 = \mu_{u1} - 4\mu_1 \mu_{\varphi 1} + 4\mu_{n1} - 24$$

$$a_{24} = \mu_1 \mu_{\varphi 1} - \mu_{u1} - \mu_{n1}, \quad a_{33} = 2 + 2\mu_{\varphi 1}, \quad a_{43} = 6 + 3\mu_{\varphi 1}$$

$$a_{34} = 2\mu_1 \mu_{\varphi 1} + 2\mu_{n1} - \mu_{u1}, \quad a_{44} = 3\mu_1 \mu_{\varphi 1} - 3\mu_{n1} - \mu_{u1} + 6$$

$$|D| = \begin{vmatrix} 0 & -\mu_{\varphi 0} & 2 & 0 \\ \mu_{u0} & a_{22} & 0 & 6 \\ 0 & \mu_{\varphi 1} & a_{33} & a_{43} \\ -\mu_{u1} & a_{24} & a_{34} & a_{44} \end{vmatrix}$$

in which K_{u0}, $K_{\varphi 0}$, K_{u1}, and $K_{\varphi 1}$ are the constants of the translational spring and rotational spring attached to the column at $x=0$ and L, respectively; a_4 is an arbitrary constant.

Substituting Eq. (43) into Eq. (4) and Eqs. (13)–(17), the buckling mode $y(\xi)$, the critical buckling force N_0, and the corresponding flexural stiffness $K(\xi)$ can be determined.

3 Functional Solutions for the Generalized Euler's Problem

In order to obtain exact solutions for more complicated cases of the generalized Euler's problem, the governing differential equation for buckling of a nonuniform column under a concentrated axial force at its top end and axially distributed variable loading is rewritten in the following form [14]:

$$\frac{d^2 M(x)}{dx^2} - \frac{1}{N(x)}\frac{dN(x)}{dx}\frac{dM(x)}{dx} + \frac{N(x)}{K(x)}M(x) = \frac{C_0}{N(x)}\frac{dN(x)}{dx} \tag{44}$$

where

$$\frac{dN(x)}{dx} = q(x)$$

$q(x)$ is the intensity of axial distributed force at x, $M(x)$, $N(x)$, and $K(x)$ are the bending moment, axial force, and flexural stiffness, respectively, and C_0 is an integral constant, which can be expressed as [14]

$$C_0 = N(x)\theta(x) - Q(x) \tag{45}$$

where $\theta(x)$ and $Q(x)$ are the slope and shear force at section x of the column, respectively. Obviously, for a cantilever column $C_0 = 0$, as for this case, the following equation is valid:

$$Q(x) = N(x)\theta(x) \tag{46}$$

As is well known, the general solution of Eq. (44) can be written as

$$M(x) = C_1 S_1(x) + C_2 S_2(x) + C_0 S_N(x) \tag{47}$$

where C_1 and C_2 are integral constants; $S_1(x)$ and $S_2(x)$ are the homogeneous solutions of Eq. (44); $S_N(x)$ is a special solution of Eq. (44), which is given by

$$S_N(x) = -S_1(x)\int \frac{N'(x)S_2(x)dx}{N(x)[S_1(x)S_2'(x) - S_2(x)S_1'(x)]} + S_2(x)\int \frac{N'(x)S_1(x)dx}{N(x)[S_1(x)S_2'(x) - S_2(x)S_1'(x)]} \tag{48}$$

where the prime indicates differentiation with respect to x.

It is evident that the key step for solving Eq. (44) is to determine the homogeneous solutions. In order to achieve this objective, the following functional transformations are adopted:

$$\varsigma = \int N(x)dx, \quad P(\zeta) = N^{-1}(x)K^{-1}(x) \tag{49}$$

$$K(x) = \text{an arbitrary function of } x, \quad N(x) = K^{-1}(x)P^{-1}\left(\int N(x)dx\right) \tag{50a}$$

or

$$N(x) = \text{an arbitrary function of } x, \quad K(x) = N^{-1}(x)P^{-1}\left(\int N(x)dx\right) \tag{50b}$$

Substituting Eq. (50*a*) or Eq. (50*b*) into the homogeneous equation of Eq. (44) yields

$$\frac{d^2 m(\varsigma)}{d\varsigma^2} + P(\varsigma)m(\varsigma) = 0, \quad m(\varsigma) = M(x) \tag{51}$$

It can be seen that the expression of Eq. (51) is much simpler than that of Eq. (44). Therefore, the solution of Eq. (51) is easier to be found. Especially ς is a function of x, $m(\varsigma)$ is actually a functional expression, and the solution of Eq. (51) is thus a functional solution; i.e., one solution of Eq. (51) actually represents a class of

solutions of the homogenous equation of Eq. (44). If the expressions of $N(x)$ and $K(x)$ are given, $P(\zeta)$ can be determined from Eq. (49), and the homogeneous solutions can be obtained by solving Eq. (51).

Obviously, the solutions of Eq. (51) are dependent on the expression of $P(\varsigma)$; such solutions for several cases that are important in engineering practice are given in the Appendix.

After $S_i(x)(i=1,2)$ and $S_N(x)$ are found, differentiating Eq. (47) with respect to x results in

$$Q(x) = C_1 S_1'(x) + C_2 S_2'(x) + C_0 S_N'(x) \tag{52}$$

Substituting Eq. (52) into Eq. (45) leads to

$$\theta(x) = C_1 \frac{S_1'(x)}{N(x)} + C_2 \frac{S_2'(x)}{N(x)} + C_0 \frac{1 + S_N'(x)}{N(x)} \tag{53}$$

Integrating Eq. (53) with respect to x yields the displacement as

$$y(x) = C_1 \int \frac{S_1'(x)}{N(x)} dx + C_2 \int \frac{S_2'(x)}{N(x)} dx + C_0 \int \frac{1 + S_N'(x)}{N(x)} dx + C_3 \tag{54}$$

The eigenvalue equation for the buckling of the generalized Euler's problem can be established based on the specific boundary conditions as follows.

1. *H-H column*. In order to easily derive the eigenvalue equation, the boundary conditions for this case (shown in Fig. 1) are written as

$$M(0) = 0, \quad M(L) = 0, \quad Q(L) = Q(0) \tag{55}$$

Applying these conditions to Eqs. (47) and (52) leads to the eigenvalue equation as

$$\begin{vmatrix} S_1(0) & S_2(0) & S_N(0) \\ S_1(L) & S_2(L) & S_N(L) \\ S_1'(L) - S_1'(0) & S_2'(L) - S_2'(0) & S_N'(L) - S_N'(0) \end{vmatrix} = 0 \tag{56}$$

2. *C-F column*. As mentioned before, it was proved that $C_0 = 0$ for this case. Using the boundary conditions (Fig. 3)

$$M(0) = 0, \quad Q(L) = 0 \tag{57}$$

results in the eigenvalue equation as

$$S_1(0) S_2'(L) - S_2(0) S_1'(L) = 0 \tag{58}$$

3. *C-H column*. Using the boundary conditions

$$M(0) = 0, \quad \theta(L) = 0, \quad Q(L) = Q(0) \tag{59}$$

the eigenvalue equation is derived as

$$\begin{vmatrix} S_1(0) & S_2(0) & S_N(0) \\ S_1'(L) & S_2'(L) & 1 + S_N'(L) \\ S_1'(L) - S_1'(0) & S_2'(L) - S_2'(0) & S_N'(L) - S_N'(0) \end{vmatrix} = 0 \tag{60}$$

4. *S-S column*. The boundary conditions for this case are

$$\begin{gathered} M(0) = -K_{\varphi 0} \theta(0), \quad M(L) = K_{\varphi L} \theta(L) \\ Q(0) = K_{u0} y(0) + N(0) \theta(0), \quad Q(L) = -K_{uL} y(L) + N(L) \theta(L) \end{gathered} \tag{61}$$

where K_{u0}, $K_{\varphi 0}$, K_{uL}, and $K_{\varphi L}$ are the constants of the translational spring and rotational spring attached to the column at $x=0$ and L, respectively.

Applying these boundary conditions to Eqs. (47) and (52)–(54) yields the eigenvalue equation as

$$\begin{vmatrix} S_1(0) + \dfrac{K_{\varphi 0} S_1'(0)}{N_0} & S_2(0) + \dfrac{K_{\varphi 0} S_2'(0)}{N_0} & S_N(0) + \dfrac{K_{\varphi 0}(1 + S_N'(0))}{N_0} & 0 \\ S_1(L) + \dfrac{K_{\varphi L} S_1'(L)}{N(L)} & S_2(L) + \dfrac{K_{\varphi L} S_2'(L)}{N(L)} & S_N(L) - \dfrac{K_{\varphi L}(1 + S_N'(L))}{N(L)} & 0 \\ 0 & 0 & -1 & -K_{u0} \\ K_{uL} \displaystyle\int_0^L \frac{S_1'(x)}{N(x)} dx & K_{uL} \displaystyle\int_0^L \frac{S_2'(x)}{N(x)} dx & -1 + K_{uL} \displaystyle\int_0^L \frac{1 + S_N'(x)}{N(x)} dx & K_{uL} \end{vmatrix} = 0 \tag{62}$$

4 Examples and Discussions

4.1 Several Special Cases. Based on the derived solutions given in Eqs. (13), (20), and (21*a*)–(21*d*) for H-H columns, several special cases of the solutions are discussed herein.

(1) *Special Case 1*.

$$n = 0, \quad \beta_0 = 1, \quad N(L) = N_0 = \text{const} \tag{63}$$

This special case represents a nonuniform column with simply supported at its both ends and subjected to a concentrated axial force N_0 at its top.

Substituting Eq. (63) into Eqs. (13) and (21*a*)–(21*d*) results in

$$N_0 = -\frac{12 b_2}{L^2} \tag{64}$$

$$b_0 = b_1 = -b_2 \tag{65}$$

If setting $b_2 = -K(0)$, then the flexural stiffness variation is expressed by

$$K(\xi) = K(0)(1 + \xi - \xi^2) \tag{66}$$

The resulted $y(\xi)$, N_0, and $K(\xi)$ are the same as those obtained by Elishakoff [16], thus verifying the reliability of the derived solutions.

(2) *Special Case 2*.

$$n = 1, \quad \beta_0 = 1, \quad N(\xi) = N_0(1 + \beta_1 \xi) \tag{67}$$

This case represents a nonuniform column with simply supported as its both ends and subjected to a concentrated force N_0 at its top and a uniform distributed force applied over all its length with intensity $\beta_1 N_0$.

Substituting Eq. (67) into Eq. (13) and Eqs. (14)–(17) leads to

$$N_0 = -\frac{15 b_3}{\beta_1 L^2} \tag{68a}$$

$$b_2 = \left(\frac{10}{\beta_1} - 7 \right) \frac{b_3}{8} \tag{68b}$$

$$b_1 = -\left(\frac{10}{\beta_1} + 7 \right) \frac{b_3}{8} \tag{68c}$$

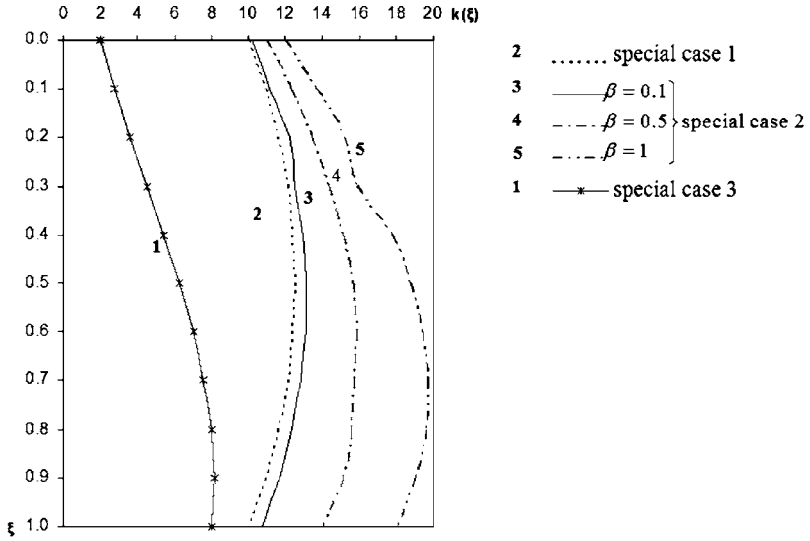


Fig. 5 Comparison among the three special cases

$$b_0 = -\left(\frac{10}{\beta_1} + 2\right)\frac{b_3}{8} \tag{68d}$$

If setting $b_2 = -8K(0)$, then we have

$$K(\xi) = K(0)\left[\left(\frac{10}{\beta_1} + 2\right) + \left(\frac{10}{\beta_1} + 7\right)\xi - \left(\frac{10}{\beta_1} - 7\right)\xi^2 - 8\xi^3\right] \tag{69}$$

(3) *Special Case 3*.

$$n = 0, \quad \beta_0 = 0, \quad \beta_1 = 1, \quad N(\xi) = N_0\xi \tag{70}$$

This case represents a nonuniform column subjected to a uniformly distributed force with intensity N_0. Obviously, it is equivalent to the Special Case 2 when $\beta_0 = 0$.

Using Eqs. (13)–(17) and (70) yields

$$b_1 = b_2 = -\tfrac{7}{8}b_3 \tag{71a}$$

$$b_0 = -\tfrac{2}{8}b_3 \tag{71b}$$

N_0 is the same as that in Eq. (68*a*) if setting $\beta_1 = 1$.

Setting $b_3 = -8K(0)$ results in

$$K(\xi) = K(0)(2 + 7\xi + 7\xi^2 - 8\xi^3) \tag{72}$$

The above results obtained from the present solutions are the same as those obtained by Elishakoff [15].

The above discussions and comparisons for the three special cases not only verify the reliability of the present solutions but also illustrate that the existing analytical solutions in literature for the buckling problems of nonuniform columns and bars actually result as special cases of the exact solutions obtained in this paper. This is due to the fact that this study considered more general buckling problems than previous researchers did, which is referred to as the generalized Euler's problem, namely, buckling of a nonuniform column with classical/nonclassical boundary conditions and subjected to a concentrated axial force and axially distributed variable loading. In particular, more complicated buckling problems of nonuniform columns such as arbitrary distribution of flexural stiffness or axial forces were considered in the present study.

A comparison among the results of the three special cases is shown in Fig. 5, where all $K(\xi)$ represent the buckling mode described by Eq. (20) and the critical buckling force is expressed as follows:

$$N_0 = \frac{120K(0)}{L^2} \tag{73}$$

It can be seen from Fig. 5 that the stiffness $K(\xi)$ for Special Case 3 is less than those for Special Cases 1 and 2, because Special Case 3 represents the column subjected to a uniformly distributed force with intensity N_0 only, and the internal axial force is $N_0\xi_1$ at $\xi = \xi_1$, which is less than N_0 when $\xi < 1$, Its maximal value is $N(0)$ at the bottom end; while Special Case 1 is corresponding to the column subjected to a concentrated force at its top, the internal axial force from the top to the bottom end is a constant N_0; such a case thus requires greater stiffness $K(\xi)$ than that for Special Case 3. It is also found that the required stiffness for Special Case 2 is greater than that for Special Case 1, because Special Case 2 represents the column subjected to a concentrated force N_0 at its top and a uniformly distributed force; hence, the internal axial force $N(\xi)$ for Special Case 2 is greater than that for Special Case 1. It is shown in this figure that $K(\xi)$ for Special Case 2 increases as β_1 increases, because $N(\xi)$ increases as β_1 becomes larger.

4.2 Stability Analysis of a High-Rise Structure. A high-rise structure with 205 m height (as shown in Fig. 6) is considered herein as a numerical example to illustrate the application and effectiveness of the present solutions and analysis procedure. The values of the flexural stiffness and axial force at several sections of this structure are listed in Table 1. It is required to determine the critical buckling force and the safety coefficient for buckling of the structure.

The stability analysis procedure based on the obtained exact solutions is as follows.

(1) *Selection of the functions for describing the distributions of flexural stiffness and axial force* According to the values of the flexural stiffness at several sections of this structure listed in Table 1, the following functions are selected for describing the variations of the flexural stiffness and axial force.

$$K(x) = K(0)\left(1 + \beta\frac{x}{L}\right), \quad N(x) = N_0\left(1 + \beta\frac{x}{L}\right) \tag{74}$$

(2) *Determination of* $P(\zeta)$. Using Eqs. (50*b*) and (74), we have

$$P(\zeta) = (d\zeta)^c \tag{75}$$

where

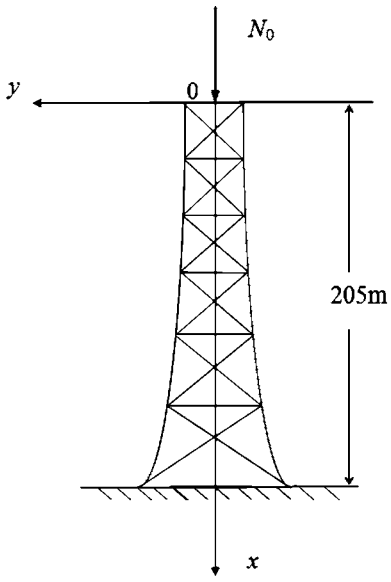


Fig. 6 Sketch of a high-rise structure

$$\varsigma = \int N(x)dx = \frac{NL}{2\beta}\left(1+\beta\frac{x}{L}\right)^2, \quad c = -\frac{3}{2}, \quad d = (\alpha N)^{2/3}\frac{2\beta}{NL}$$

(3) *Determination of the solutions for buckling of the column*. The expression of $P(\varsigma)$ determined above belongs to a special case of Case 2 in the Appendix; the solutions are found to be

$$S_1(\varsigma) = (d\varsigma)^{1/2}J_2[\alpha(d\varsigma)^{1/4}] \tag{76a}$$

$$S_2(\varsigma) = (d\varsigma)^{1/2}Y_2[\alpha(d\varsigma)^{1/4}] \tag{76b}$$

$$\alpha = \frac{4}{b} \tag{76c}$$

where $J_2(\cdot)$ and $Y_2(\cdot)$ are Bessel functions of the first and second kinds, of order 2, respectively.

(4) *Determination of the critical buckling force and the safety coefficient for buckling*. The eigenvalue equation for buckling of this structure is Eq. (58). Substituting Eqs. (76*a*)–(76*c*) into Eq. (58) and solving it result in the critical buckling force as $N_{cr}(0) = 2.357\times10^7$ N The safety coefficient, k, for buckling of the structure is thus determined as

$$k = \frac{N_{cr}(0)}{N_0} = \frac{23.57\times10^6}{3.207\times10^6} = 7.350$$

5 Conclusions

In this study, the exact solutions were derived for the buckling problems of nonuniform columns with variable flexural stiffness and distributed variable axial loading expressed in terms of polynomial functions. Although several closed-form solutions for buckling of a nonuniform column subjected to a concentrated axial force at its top or a axially distributed force were proposed by previous researchers, more complicated cases, such as that a nonuniform column with arbitrary variation of flexural stiffness is expressed as a functional relation with the distribution of the flexural stiffness and vice versa, were considered in this paper. The governing equation for such problems was reduced to Bessel equations and other solvable equations for several cases by means of functional transformations. The exact solutions for the buckling problem of nonuniform columns subjected to an axial concentrated force and axially distributed variable loading were thus obtained. It was shown through the examples that several existing closed-form solutions for the buckling problems of nonuniform columns actually were special cases of the exact solutions obtained in this paper.

In this paper, the governing equation for buckling of nonuniform columns subjected to a concentrated axial compressive force and axially distributed variable loading was given by the second order differential equation for bending moment $M(x)$. Then, differentiating and integrating $M(x)$ with respect to x can determine shear force, slope, and displacement. The present procedure is actually simpler than conventional approaches, which need solving the fourth order differential equation for displacement.

Acknowledgment

The work described in this paper was fully supported by a grant from the Research Grants Council of Hong Kong Special Administrative Region, China (Project No. CityU 116906)

Appendix: Exact Solutions of Equation (51) for Several Cases

Case 1.

$$P(\varsigma) = ae^{b\varsigma} - c \tag{A1}$$

$$S_1(\varsigma) = J_v(\lambda e^{b\varsigma/2}), \quad S_2(\varsigma) = Y_v(\lambda e^{b\varsigma/2})$$
$$\lambda^2 = \frac{4a}{b^2}, \quad v^2 = \frac{4c}{b^2} \tag{A2}$$

where $J_v(\cdot)$ and $Y_v(\cdot)$ are Bessel functions of the first and second kinds, of order v, respectively.

Case 2.

$$P(\varsigma) = (a+b\varsigma)^c, \quad c \neq -2 \tag{A3}$$

$$S_1(\varsigma) = (a+b\varsigma)^{1/2}J_v[\alpha(a+b\varsigma)^{1/2v}]$$
$$S_2(\varsigma) = (a+b\varsigma)^{1/2}Y_v[\alpha(a+b\varsigma)^{1/2v}] \tag{A4}$$
$$\alpha^2 = \frac{4v^2}{b^2}, \quad v^2 = \frac{1}{c+2}$$

When $c=-2$, Eq. (51) becomes an Euler equation.

Case 3.

$$P(\varsigma) = (a\varsigma^2 + b\varsigma + c)^{-2} \tag{A5}$$

$$\left.\begin{aligned} S_1(\varsigma) &= (a\varsigma^2+b\varsigma+c)^{1/2}\sin\sqrt{A}\xi \\ S_2(\varsigma) &= (a\varsigma^2+b\varsigma+c)^{1/2}\cos\sqrt{A}\xi \end{aligned}\right\} \quad \text{for } A>0 \tag{A6}$$

Table 1 Values of flexural stiffness and axial force at several sections of the structure. Note that the data in parentheses are calculated by Eq. (74).

x (m)	0	25	50	70	90	110	135	170	205
$K(x)$	1.241	1.780	2.423	3.002	3.641	4.347	5.319	7.830	8.552
	(1.241)	(1.780)	(2.423)	(3.002)	(3.641)	(4.347)	(5.319)	(7.830)	(8.552)
$N(x)$	3.207	3.841	4.486	4.987	5.494	6.008	6.636	7.528	4.419
	(3.207)	(3.841)	(4.486)	(4.987)	(5.494)	(6.008)	(6.636)	(7.528)	(4.419)

$$A = 1 + ac - \frac{1}{4}b^2, \quad \xi = \int \frac{d\varsigma}{a\varsigma^2 + b\varsigma + c}$$

$$\left.\begin{aligned} S_1(\varsigma) &= (a\varsigma^2 + b\varsigma + c)^{1/2} e^{\sqrt{A}\xi} \\ S_2(\varsigma) &= (a\varsigma^2 + b\varsigma + c)^{1/2} e^{-\sqrt{A}\xi} \end{aligned}\right\} \quad \text{for } A < 0 \tag{A7}$$

$$\left.\begin{aligned} S_1(\varsigma) &= (a\varsigma^2 + b\varsigma + c)^{1/2} \\ S_2(\varsigma) &= (a\varsigma^2 + b\varsigma + c)^{1/2} \xi \end{aligned}\right\} \quad \text{for } A = 0 \tag{A8}$$

Case 4.

$$P(\varsigma) = a(\varsigma^2 + b)^{-2}, \quad a > 0, \quad b > 0 \tag{A9}$$

$$S_1(\varsigma) = (\varsigma^2 + b)^{1/2} \sin \xi$$

$$S_2(\varsigma) = (\varsigma^2 + b)^{1/2} \cos \xi \tag{A10}$$

$$\xi = \left(\frac{a+b}{b}\right)^{1/2} \arctan \frac{\varsigma}{b^{1/2}}$$

Case 5.

$$P(\varsigma) = a(\varsigma^2 - b)^{-2}, \quad a > 0, \quad b > 0 \tag{A11}$$

$$S_1(\varsigma) = (\varsigma^2 - b)^{1/2} \sin \xi$$

$$S_2(\varsigma) = (\varsigma^2 - b)^{1/2} \cos \xi \tag{A12}$$

$$\xi = \frac{1}{2}\left(\frac{a-b}{b}\right)^{1/2} \ln \frac{b^{1/2} + \varsigma}{b^{1/2} - \varsigma}, \quad |\varsigma| < b^{1/2}$$

Case 6.

$$P(\varsigma) = ac^{b\zeta} \tag{A13}$$

This is a special case of Eq. (A1). The solutions are the same as those of Eq. (A2) when setting $\upsilon = 0$.

Case 7.

$$P(\varsigma) = a\varsigma^{c-2} - b^2\varsigma^{2c-2} \tag{A14}$$

$$S_1(\varsigma) = \varsigma^{1/2(1-c)} \Phi\left(\frac{a}{2bc}, -\frac{1}{2c}; \frac{2b}{c}\varsigma^c\right) \tag{A15}$$

$$S_2(\varsigma) = \left(\frac{2b}{c}\varsigma^c\right)^{1-(1/2)c} \Phi\left(\frac{a}{2bc} - \frac{1}{2c} + 1, 2 - \frac{1}{2c}; \frac{2b}{c}\varsigma^c\right)$$

where $\Phi(x,x,x)$ represents the Φ function [20].

References

[1] Euler, L., 1757, "Sur La Forces Des Colunnes," Memoires de l'Academie Royale des Sciences et Belles Letters, Berlin, **13**, pp. 252–282.

[2] Greenhill, A. G., 1881, "Determination of the Greatest Height Consistent With Stability That a Vertical Pole or Mast Can Be Made, and of the Greatest Height to Which a Tree of Given Proportions Can Grow," Proc. Cambridge Philos. Soc., **4**, pp. 765–774.

[3] Morley, A., 1917, "Critical Loads for Long Tapering Struts," Engineering (London), **104**, p. 295.

[4] Dinnik, A. N., 1912, "Buckling Under Own Weight," Proceedings of the Don Polytechnic Institute, **1**, p. 19.

[5] Dinnik, A. N., 1955, "Selected Papers, Application of Bessel Function to the Theory of Elasticity Problems," Ukrainian Academy of Science, **2**, pp. 73–87.

[6] Karman, T. V., and Biot, M. A., 1940, *Mathematical Methods in Engineering*, McGraw-Hill, New York.

[7] Timoshenk, S. P., and Gere, G. M., 1961, *Theory of Elastic Stability*, McGraw-Hill, New York.

[8] Gere, J. M., and Carter, W. O., 1962, "Critical Buckling Loads for Tapered Columns," ASCE J. Struct. Div., **88**(1), pp. 1346–1354.

[9] Ermopoulos, J. C., 1986, "Buckling of Tapered Bars Under Stepped Axial Loads," J. Struct. Eng., **112**(6), pp. 1346–1354.

[10] Williams, F. W., and Aston, G., 1989, "Exact or Lower Bound Tapered Column Buckling Loads," J. Struct. Eng., **115**(5), pp. 1088–1100.

[11] Arbabi, F., and Li, F., 1991, "Buckling of Variable Cross-Section Columns—Integral Equation Approach," J. Struct. Eng., **117**(8), pp. 2426–2441.

[12] Siginer, A., 1992, "Buckling of Columns of Variable Flexural Rigidity," J. Eng. Mech., **118**(3), pp. 640–643.

[13] Li, Q. S., Cao, H., and Li, G., 1994, "Stability Analysis of Bars With Multi-Segments of Varying Cross-Section," Comput. Struct., **53**(5), pp. 1085–1089.

[14] Li, Q. S., Cao, H., and Li, G., 1995, "Stability Analysis of Bars With Varying Cross-Section," Int. J. Solids Struct., **32**(21), pp. 3217–3228.

[15] Elishakoff, I., 2000, "A Closed-Form Solution for the Generalized Euler Problem," Proc. R. Soc. London, Ser. A, **456**, pp. 2409–2417.

[16] Elishakoff, I., 2001, "Inverse Buckling Problem for Inhomogeneous Columns," Int. J. Solids Struct., **38**, pp. 457–464

[17] Elishakoff, I., 2001, "Euler's Problem Revisited: 222 Years Later," Meccanica, **36**, pp. 265–272.

[18] Elishakoff, I., and Guédé, Z., 2001, "Novel Closed-Form Solutions in Buckling of Inhomogeneous Columns Under Distributed Variable Loading," Chaos, Solitons Fractals, **12**, pp. 1075–1089.

[19] Elishakoff, I., 2005, *Eigenvalues of Inhomogeneous Structures*, CRC, Boca Raton, FL.

[20] Kamke, E., 1972, *Differentialgleichungen, Lö sungsmethoden und Lö sungen*, Chelsea, New York.

Nonlinear Field Theory of Fracture Mechanics for Paramagnetic and Ferromagnetic Materials

Xiaohong Chen
Sikorsky Aircraft Corporation,
6900 Main Street,
P.O. Box 9729,
Stratford, CT 06615-9129

A nonlinear field theory of fracture mechanics is developed for crack propagation in paramagnetic and ferromagnetic materials from the global energy balance equation and the non-negative global dissipation requirement. The crack-front generalized $\tilde{J}$-integral is equivalent to the generalized energy release rate serving as the thermodynamic driving force for crack propagation and also related to the generalized energy-momentum tensor in a way similar to the material force method. On the basis of the developed theory, the generalized energy release rate method, the generalized $\tilde{J}$-integral method, and the extended essential work of fracture method are proposed for quasistatic and dynamic fracture characterization of magnetosensitive materials in the presence of magnetothermomechanical coupling and dissipative effects. The present work overcomes the drawbacks and limitations of conventional fracture mechanics and resolves the controversial issues on magnetoelastic fracture criterion. Especially, the crack-front generalized $\tilde{J}$-integral has an odd dependence on the magnetic induction intensity factor for a Griffith-type crack in a magnetoelastic solid. [DOI: 10.1115/1.3086784]

1 Introduction

With the wide applications of materials consisting of piezoelectric and piezomagnetic components, particularly in smart structures and devices, growing interests have been developed on magneto-electro-elastic fracture mechanics [1–33]. Fracture criteria for these intelligent materials are of vital importance for reliability assessment of structural integrity. Shindo [1,2] first studied the distribution of mechanical and magnetic fields in an infinite body with a planar or penny-shaped crack following a linear theory for soft ferromagnetic elastic materials. Path-independent integrals have been employed to study defected bodies since the pioneering work of Eshelby [34–36]. Based on the rotationally invariant quasimagnetostatic theory of elastic paramagnets and soft ferromagnets, Sabir and Maugin [3] constructed the path-independent integrals with the magnetic enthalpy and yielded the same results as the canonical field-theoretic approach using the notions of Eshelby stress and material forces [5]. Their studies indicate that the magnetic field has a negative contribution to the energy release rate so that its presence is beneficial from the point of view of fracture toughness. Such a conclusion may be doubtful since a similar prediction based on the path-independent integral constructed with the electric enthalpy [7–13] has been shown contradictory to experimental observation on crack growth in piezoelectric materials [14–19].

Many attempts, e.g., Refs. [20–25], have been made to resolve this controversy. However, as pointed out by McMeeking [24], the major difficulty with this subject is that the theoretical treatments are all incomplete, and some of the theoretical treatments also introduce features that are hard to justify. On the other hand, thermodynamics/thermomechanics has been used to study thermoelasticity, magneto-electro-elasticity, viscoelasticity, plasticity, and fracture [37–43]. Most recently, Chen [44] proposed a new formation of the energy flux integral and the energy-momentum tensor for studying the crack driving force in electroelastodynamic fracture and showed that the dynamic energy release rate thus obtained has an odd dependence on the electric displacement intensity factor, which is in agreement with experimental evidence. Chen [45] further developed a nonlinear field theory of fracture mechanics for piezoelectric and ferroelectric materials based on fundamental principles of thermodynamics.

The objective of the present work is to develop a consistent thermodynamic formulation of a nonlinear field theory of fracture mechanics for paramagnetic and ferromagnetic materials from the global energy balance equation and the non-negative global dissipation requirement. In Sec. 2, crack propagation in the presence of magnetothermomechanical coupling and dissipative effects is determined in terms of the generalized energy release rate as the thermodynamic driving force conjugate to the crack variable. In Sec. 3, on the basis of the developed theory, the generalized energy release rate method, the generalized $\tilde{J}$-integral method, and the extended essential work of fracture (EWF) method are proposed for quasistatic and dynamic fracture characterization of magnetosensitive materials, and the inter-relation of these methods and their correlation with conventional methods are explored. In Sec. 4, a Griffith-type magnetoelastic crack problem is investigated as a case study. In Sec. 5, some important features of this formulation for an in-depth understanding of the magnetoelastic fracture behavior are discussed. In the end, the concluding remarks are given.

2 Thermodynamic Formulation

Continuum mechanics of electromagnetic solids has been presented in the monographs by Eringen [39] and Maugin [40] with Maxwell equations, mass balance equation, linear momentum balance equation, angular momentum balance equation, and energy balance equation in global and local forms. The following theoretical framework is developed for a cracked body under combined magnetic, thermal, and mechanical loadings involving dissipative mechanisms.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 28, 2008; final manuscript received December 22, 2008; published online April 28, 2009. Review conducted by Zhigang Suo.

2.1 Maxwell Equations. The magnetic field $\mathbf{H}$ is introduced by

$$\mathbf{H} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{M} \tag{1}$$

where μ_0 is the vacuum permeability, $\mathbf{B}$ is the magnetic induction, and $\mathbf{M}$ is the magnetization.

The Maxwell equations in the absence of electric effect are simplified as

$$\nabla \cdot \mathbf{B} = 0 \tag{2}$$

$$\nabla \times \mathbf{H} = 0 \tag{3}$$

$$\frac{\partial \mathbf{B}}{\partial t} = 0 \tag{4}$$

2.2 Mass Balance. Following the conservation law of mass, the global mass balance equation over a cracked body V_t bounded by the material surface ∂V_t is given by

$$\frac{d}{dt}\int_{V_t} \rho dV = 0 \tag{5}$$

where ρ is the mass density.

Hence, the local mass balance equation is obtained as

$$\frac{d\rho}{dt} = -\rho \nabla \cdot \mathbf{v} \tag{6}$$

where $\mathbf{v}=\dot{\mathbf{u}}$ is the velocity vector and $\mathbf{u}$ is the displacement vector.

2.3 Linear and Angular Momentum Balance. Following the conservation law of linear momentum, the global linear momentum balance equation over the cracked body V_t is given by

$$\frac{d}{dt}\int_{V_t} \rho \mathbf{v} dV = \int_{\partial V_t} \mathbf{t}_{(n)} dS + \int_{V_t} (\rho \hat{\mathbf{f}} + {}_m\mathbf{f}) dV \tag{7}$$

where $\mathbf{t}_{(n)}=\mathbf{n}\cdot\boldsymbol{\sigma}$ is the mechanical traction acting on the surface ∂V_t, $\boldsymbol{\sigma}$ is the Cauchy stress tensor, $\hat{\mathbf{f}}$ is the mechanical body force per unit mass, and ${}_m\mathbf{f}=(\nabla\mathbf{B})\cdot\mathbf{M}$ is the magnetic body force per unit volume.

Hence, the local linear momentum balance equation is obtained as

$$\rho\frac{d\mathbf{v}}{dt} = \nabla \cdot \boldsymbol{\sigma} + \rho\hat{\mathbf{f}} + {}_m\mathbf{f} \tag{8}$$

Following the conservation law of angular momentum, the global angular momentum balance equation over the cracked body V_t is given by

$$\frac{d}{dt}\int_{V_t} \mathbf{r}\times\rho\mathbf{v} dV = \int_{\partial V_t} (\mathbf{r}\times\mathbf{t}_{(n)}) dS + \int_{V_t} (\mathbf{r}\times\rho\hat{\mathbf{f}} + \mathbf{r}\times{}_m\hat{\mathbf{f}} + {}_m\mathbf{c}) dV \tag{9}$$

where $\mathbf{r}$ is the position vector from the fixed point O, and ${}_m\mathbf{c}=\mathbf{M}\times\mathbf{B}$.

Hence, the local angular momentum balance leads to

$$\varepsilon_{kij}\sigma_{ij} + {}_mc_k = 0 \tag{10}$$

With the use of the second-order tensor ${}_m\boldsymbol{\sigma}$, Eqs. (8) and (10) can be rewritten as

$$\rho\frac{d\mathbf{v}}{dt} = \nabla\cdot(\boldsymbol{\sigma} + {}_m\boldsymbol{\sigma}) + \rho\hat{\mathbf{f}} \tag{11}$$

$$\varepsilon_{kij}(\sigma_{ij} + {}_m\sigma_{ij}) = 0 \tag{12}$$

where ${}_m\boldsymbol{\sigma} = -\mathbf{B}\otimes\mathbf{M} + (1/\mu_0)\mathbf{B}\otimes\mathbf{B} + (\mathbf{M}\cdot\mathbf{B} - {}_m\mu^f)\mathbf{I}$ is the magnetic stress and ${}_m\mu^f = (1/2\mu_0)\mathbf{B}\cdot\mathbf{B}$ is the energy density of the free magnetic field.

2.4 Energy Balance. Following the conservation law of energy, the global energy balance equation over the cracked body V_t is given by

$$\frac{dE}{dt} + \frac{dK}{dt} = \dot{W} + \dot{Q} \tag{13}$$

where E is the internal energy, K is the kinetic energy, $\dot{W}$ is the power by external forces, and $\dot{Q}$ is the heat exchange rate.

The internal energy E and the kinetic energy K over the cracked body V_t are defined as

$$E \equiv \int_{V_t} \rho\hat{e} dV \tag{14}$$

$$K \equiv \int_{V_t} \rho\hat{k} dV \tag{15}$$

where $\hat{e}$ is the internal energy per unit mass and $\hat{k}$ is the kinetic energy per unit mass.

The power by external forces is given by

$$\dot{W} = \int_{\partial V_t} \mathbf{t}_{(n)}\cdot\mathbf{v} dS + \int_{V_t} \rho\hat{\mathbf{f}}\cdot\mathbf{v} dV + \int_{V_t} {}_m w dV \tag{16}$$

where ${}_m w = {}_m\mathbf{f}\cdot\mathbf{v} - \mathbf{M}\cdot\dot{\mathbf{B}}$ is the magnetic power per unit volume.

The heat flux $\mathbf{j}_q$ is introduced to describe the heat exchange rate through the boundary ∂V_t as

$$\dot{Q} = -\int_{\partial V_t} \mathbf{n}\cdot\mathbf{j}_q dS \tag{17}$$

Hence, the global energy balance equation over the cracked body V_t becomes

$$\frac{d}{dt}\int_{V_t} \rho(\hat{k}+\hat{e}) dV = -\int_{V_t} \nabla\cdot\mathbf{j}_q dV + \int_{V_t} (\nabla\cdot\boldsymbol{\sigma} + \rho\hat{\mathbf{f}} + {}_m\hat{\mathbf{f}})\cdot\mathbf{v} dV + \int_{V_t} \boldsymbol{\sigma}:\mathbf{v}\nabla dV - \int_{V_t} \mathbf{M}\cdot\dot{\mathbf{B}} dV \tag{18}$$

2.5 Helmholtz Free Energy. The Helmholtz free energy per unit mass is defined as

$$\hat{h} \equiv \hat{e} - T\hat{s} \tag{19}$$

where T is the absolute temperature and $\hat{s}$ is the entropy per unit mass.

Substituting Eq. (19) into Eq. (18) yields

$$\int_{V_t} \rho T\frac{d\hat{s}}{dt} dV = -\int_{V_t} \nabla\cdot\mathbf{j}_q dV + \int_{V_t} (\nabla\cdot\boldsymbol{\sigma} + \rho\hat{\mathbf{f}} + {}_m\hat{\mathbf{f}})\cdot\mathbf{v} dV + \int_{V_t} (\boldsymbol{\sigma}:\mathbf{v}\nabla - \mathbf{M}\cdot\dot{\mathbf{B}} - \rho\hat{s}\dot{T}) dV - \frac{d}{dt}\int_{V_t} \rho(\hat{k}+\hat{h}) dV \tag{20}$$

2.6 Non-Negative Global Dissipation Requirement. The entropy production is defined as

$$\frac{d_i\hat{s}}{dt} \equiv \frac{d\hat{s}}{dt} + \frac{1}{\rho}\nabla\cdot\mathbf{j}_s \tag{21}$$

where $\mathbf{j}_s$ is the entropy flux.

The non-negative global dissipation requirement is given by

$$\int_{V_t} \rho T \frac{d_i \hat{s}}{dt} dV \geq 0 \tag{22}$$

2.7 State Equations. Substituting Eqs. (20) and (21) into Eq. (22) with the use of Eq. (8) yields

$$\int_{V_t} \rho T \frac{d_i \hat{s}}{dt} dV = -\int_{\partial V_t} \mathbf{n} \cdot (\mathbf{j}_q - T\mathbf{j}_s) dS - \int_{V_t} \mathbf{j}_s \cdot \nabla T dV + \int_{V_t} \rho \dot{\mathbf{v}} \cdot \mathbf{v} dV + \frac{1}{2}\int_{V_t} J^{-1}{}_t\boldsymbol{\Sigma} : \dot{\mathbf{C}} dV + \int_{V_t} J^{-1}\hat{\mathbf{H}} \cdot \hat{\dot{\mathbf{B}}} dV - \int_{V_t} \rho \hat{s}\dot{T} dV - \frac{d}{dt}\int_{V_t} \rho\left(\hat{k} + \hat{h} + \frac{{}_m u^f}{\rho}\right) dV \geq 0 \tag{23}$$

where ${}_t\boldsymbol{\Sigma} = J\mathbf{F}^{-1}\,{}_t\boldsymbol{\sigma}\mathbf{F}^{-T}$ is the second Piola–Kirchhoff total stress tensor, ${}_t\boldsymbol{\sigma} = \boldsymbol{\sigma} + {}_m\boldsymbol{\sigma}$ is symmetric, $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ is the right Cauchy–Green deformation tensor, $\hat{\mathbf{H}} = \mathbf{H}\cdot\mathbf{F}$, $\hat{\mathbf{B}} = J\mathbf{F}^{-1}\cdot\mathbf{B}$, $J = \det(\mathbf{F})$, $\mathbf{F} = \partial\mathbf{x}/\partial\mathbf{X}$ is the deformation gradient, $\mathbf{X}$ is the position in the reference configuration, $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$ is the position in the current configuration, and $\mathbf{u} = \mathbf{x} - \mathbf{X}$.

In the reference configuration V_R with respect to which the deformation gradient $\mathbf{F}$ is measured, the Helmholtz free energy including the contribution of the energy of the free magnetic field, $\tilde{h} = \hat{h} + {}_m u^f/\rho$, is taken to be a functional of the histories of deformation, temperature, temperature gradient, and magnetic induction, with the crack parameter, A, as an internal state variable:

$$\tilde{h} = \tilde{h}(\mathbf{C}(t-\tau), T(t-\tau), \nabla_R T(t-\tau), \hat{\mathbf{B}}(t-\tau); A, \mathbf{X}) \tag{24}$$

The *Hamiltonian* density is given by

$$\mathcal{H}(\mathbf{v}, \mathbf{C}(t-\tau), T(t-\tau), \nabla_R T(t-\tau), \hat{\mathbf{B}}(t-\tau); A, \mathbf{X}) = \hat{k}(\mathbf{v}; A, \mathbf{X}) + \tilde{h}(\mathbf{C}(t-\tau), T(t-\tau), \nabla_R T(t-\tau), \hat{\mathbf{B}}(t-\tau); A, \mathbf{X}) \tag{25}$$

In order for the non-negative global dissipation requirement inequality (23) to be always valid, it is necessary and sufficient that state equations fulfill the following conditions:

$$\frac{\delta\tilde{h}}{\delta T_{,K}} = 0 \tag{26}$$

$${}_t\Sigma_{KL} = 2\rho_0 \frac{\delta\tilde{h}}{\delta C_{KL}} \tag{27}$$

$$\hat{H}_K = \rho_0 \frac{\delta\tilde{h}}{\delta\hat{B}_K} \tag{28}$$

$$\hat{s} = -\frac{\delta\tilde{h}}{\delta T} \tag{29}$$

$$\mathbf{j}_q = T\mathbf{j}_s \tag{30}$$

$$\int_{V_t} \rho T \frac{d_i \hat{s}}{dt} dV = \int_{V_t} T\mathbf{j}_q \cdot \nabla\frac{1}{T} dV + \int_{V_t} \rho\hat{\Lambda} dV + \tilde{G}\dot{A} \geq 0 \tag{31}$$

where the thermodynamic force, $\tilde{G}$, the conjugate to the crack variable, A, and the time-dependent dissipation rate, $\hat{\Lambda}$, are given by

$$\tilde{G} = -\frac{\delta}{\delta A}\int_{V_t} \rho\mathcal{H} dV \tag{32}$$

$$\hat{\Lambda} = -\frac{\partial\tilde{h}}{\partial t} \tag{33}$$

From Eq. (26), the Helmholtz free energy functional does not depend on the temperature gradient. From Eq. (31), the total dissipation is associated with heat conduction, relaxation/retardation, and crack propagation. Since inequality (31) should always be satisfied, kinetic laws for specific irreversible processes can be determined accordingly.

2.8 Coupled Heat Transfer Equation. Thermodynamic flux for heat conduction is postulated to depend linearly on corresponding thermodynamic force, that is,

$$\mathbf{j}_q = \mathbf{L}^{qq} \cdot \nabla\frac{1}{T} \tag{34}$$

where $\mathbf{L}^{qq}$ is positive definite.

Substituting Eqs. (21), (29), (30), and (34) into Eq. (31) yields the coupled heat transfer equation for the cracked body,

$$-\int_{V_t} \rho T \frac{\delta\tilde{h}}{\delta T} dV = -\int_{V_t} \nabla\cdot\left(\mathbf{L}^{qq}\cdot\nabla\frac{1}{T}\right) dV + \int_{V_t} \rho\hat{\Lambda} dV + \tilde{G}\dot{A} \tag{35}$$

2.9 Time-Dependent Dissipation Rate Inequality. The time-dependent dissipation rate satisfies the following inequality:

$$\hat{\Lambda} \geq 0 \tag{36}$$

2.10 Thermodynamic Consistent Time-Dependent Fracture Criterion. In terms of the generalized energy release rate, $\tilde{G}$, as the thermodynamic driving force conjugate to the crack variable, A, crack propagation in coupled magnetic, thermal, and mechanical fields is determined by

$$\tilde{G} - R = 0 \tag{37}$$

where R is the crack resistance.

Substituting Eqs. (14), (17), (19), (21), (30), and (31) into the global energy balance equation (13) yields

$$\dot{W} = \frac{d}{dt}\int_{V_t} \rho\hat{h} dV + \frac{d}{dt}\int_{V_t} \rho\hat{k} dV + \int_{V_t} \rho\hat{s}\dot{T} dV + \int_{V_t} \rho\hat{\Lambda} dV + \tilde{G}\dot{A} \tag{38}$$

For stable crack growth, substituting Eq. (37) into Eq. (38) yields

$$\dot{W} = \frac{dH}{dt} + \frac{dK}{dt} + \int_{V_t} \rho\hat{s}\dot{T} dV + \int_{V_t} \rho\hat{\Lambda} dV + R\dot{A} \tag{39}$$

where $H \equiv \int_{V_t} \rho\hat{h} dV$. Integrating Eq. (39) over the time domain gives the following expression for the total work:

$$\Delta W = \Delta H + \Delta K + \int_{t_0}^{t}\int_{V_t} \rho\hat{s}\dot{T} dV dt + \int_{t_0}^{t}\int_{V_t} \rho\hat{\Lambda} dV dt + \int_{t_0}^{t} R\dot{A} dt \tag{40}$$

With the use of the Poynting vector $\boldsymbol{S} = (\mathbf{v}\times\mathbf{B})\times\mathbf{H}$, Eq. (16) leads to

$$\dot{W} = \int_{\partial V_t} \mathbf{n}\cdot(\boldsymbol{\sigma} + {}_m\boldsymbol{\sigma})\cdot\mathbf{v}dS + \int_{V_t}\rho\hat{\mathbf{f}}\cdot\mathbf{v}dV - \int_{\partial V_t}\mathbf{n}\cdot\boldsymbol{S}dS - \frac{d}{dt}\int_{V_t}{}_m u^f dV \tag{41}$$

Substituting Eq. (41) into Eq. (38) yields

$$\int_{\partial V_t}\mathbf{n}\cdot(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})\cdot\mathbf{v}dS - \int_{\partial V_t}\mathbf{n}\cdot\boldsymbol{S}dS = \frac{d}{dt}\int_{V_t}\rho(\tilde{h}+\hat{k})dV - \int_{V_t}\rho\hat{\mathbf{f}}\cdot\mathbf{v}dV + \int_{V_t}\rho\hat{s}\dot{T}dV + \int_{V_t}\rho\hat{\Lambda}dV + \tilde{G}\dot{A} \tag{42}$$

Consider a three-dimensional cracked body $\tilde{V}_t$ bounded by a surface $\tilde{\Gamma}$ in a reference frame ($\tilde{\mathbf{X}}=\mathbf{X}-\mathbf{c}t$) traveling with the crack front at a speed $\mathbf{c}$. Equation (42) can be rewritten as

$$F(\tilde{\Gamma}) \equiv \int_{\tilde{\Gamma}}[\mathbf{n}\cdot(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})\cdot\mathbf{v} - \mathbf{n}\cdot\boldsymbol{S} + (\tilde{\rho}\tilde{h}+\tilde{\rho}\hat{k})\mathbf{n}\cdot\mathbf{c}]d\tilde{\Gamma} = \int_{\tilde{V}_t}\frac{\tilde{\partial}}{\tilde{\partial}t}(\tilde{\rho}\tilde{h}+\tilde{\rho}\hat{k})d\tilde{V} - \int_{\tilde{V}_t}\tilde{\rho}\hat{\mathbf{f}}\cdot\mathbf{v}d\tilde{V} + \int_{\tilde{V}_t}\tilde{\rho}\hat{s}\dot{T}d\tilde{V} + \int_{\tilde{V}_t}\tilde{\rho}\hat{\Lambda}d\tilde{V} + \tilde{G}\dot{A} \tag{43}$$

The above expression represents the total energy flux through the surface $\tilde{\Gamma}$. It is noted that the global energy balance for the existence of a moving discontinuous line/surface in a material surface/volume [39,40] may lead to a similar expression, but it has not been used to study crack problems.

As the surface $\tilde{\Gamma}\to 0$, the energy flux integral is related to the generalized energy release rate by

$$\lim_{\tilde{\Gamma}\to 0}\left\{\frac{F(\tilde{\Gamma})}{Bc}\right\} = \lim_{\tilde{\Gamma}\to 0}\left\{\frac{1}{Bc}\int_{\tilde{\Gamma}}[\mathbf{n}\cdot(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})\cdot\mathbf{v} - \mathbf{n}\cdot\boldsymbol{S} + (\tilde{\rho}\tilde{h}+\tilde{\rho}\hat{k})\mathbf{n}\cdot\mathbf{c}]d\tilde{\Gamma}\right\} = \tilde{G} \tag{44}$$

where $c=|\mathbf{c}|$ is the crack velocity and B is the thickness along the crack front.

3 Application to Fracture Characterization

3.1 Generalized Energy Release Rate Method. The generalized energy release rate criterion given by Eq. (37) is a generalization of the Griffith–Irwin strain energy release rate criterion [46]. The time-dependent fracture criterion given by Eq. (39) is a generalization of the rate-dependent criterion for viscoelastic crack growth [47]. The generalized energy release rate, $\tilde{G}$, and the time-dependent dissipation rate, $\hat{\Lambda}$, can be calculated from Eqs. (32) and (33).

3.2 Generalized $\tilde{J}$-Integral Method. The generalized $\tilde{J}$-integral is defined as

$$\tilde{J}_{\tilde{\Gamma}} \equiv \frac{F(\tilde{\Gamma})}{Bc} \tag{45}$$

From Eq. (44), the crack-front generalized $\tilde{J}$-integral is the energy flux toward the crack-front per unit crack advance, which is equivalent to the generalized energy release rate, that is,

$$\tilde{J}_0 = \lim_{\tilde{\Gamma}\to 0}\left\{\frac{F(\tilde{\Gamma})}{Bc}\right\} = \tilde{G} \tag{46}$$

The relation between the global and local generalized $\tilde{J}$-integrals is obtained from Eq. (43) as

$$\tilde{J}_g = \tilde{J}_l + \frac{1}{Bc}\int_{\tilde{V}_g-\tilde{V}_l}\frac{\tilde{\partial}}{\tilde{\partial}t}(\tilde{\rho}\tilde{h}+\tilde{\rho}\hat{k})d\tilde{V} - \frac{1}{Bc}\int_{\tilde{V}_g-\tilde{V}_l}\tilde{\rho}\hat{\mathbf{f}}\cdot\mathbf{v}d\tilde{V} + \frac{1}{Bc}\int_{\tilde{V}_g-\tilde{V}_l}\tilde{\rho}\hat{s}\dot{T}d\tilde{V} + \frac{1}{Bc}\int_{\tilde{V}_g-\tilde{V}_l}\tilde{\rho}\hat{\Lambda}d\tilde{V} \tag{47}$$

where $\tilde{V}_g$ and $\tilde{V}_l$ are the volumes bounded by the closed surfaces $\tilde{\Gamma}_g$ and $\tilde{\Gamma}_l$ including crack faces.

Hence, the generalized $\tilde{J}$-integral is a generalization of the conventional J-integral and dynamic contour integral [46,48] and Schapery's [49] crack-tip model for viscoelastic fracture. For steady-state crack propagation in the absence of mechanical body force ($\hat{\mathbf{f}}=0$), thermal effect ($\dot{T}=0$), and time-dependent dissipation ($\hat{\Lambda}=0$), the generalized $\tilde{J}$-integral becomes path independent as the closed surface is chosen in the present work.

For a flat, straight through-crack, if a field quantity is invariant in a reference frame traveling with the crack front in the $\tilde{\mathbf{e}}_1$-direction at a speed $\mathbf{c}=c\tilde{\mathbf{e}}_1$, the field quantity depends on t only through the combination $\tilde{\mathbf{X}}=\mathbf{X}-\mathbf{c}t$. The energy flux integral (44) takes the special form

$$F(\tilde{\Gamma}) = \int_{\tilde{\Gamma}}\mathbf{n}\cdot[-(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})\cdot\mathbf{u}\tilde{\nabla} + \mathbf{B}\otimes\mathbf{H}\cdot\mathbf{u}\tilde{\nabla} - (\mathbf{H}\cdot\mathbf{B})\mathbf{u}\tilde{\nabla} + (\tilde{\rho}\tilde{h}+\tilde{\rho}\hat{k})\mathbf{I}]\cdot\mathbf{c}d\tilde{\Gamma} \tag{48}$$

With the introduction of the generalized energy-momentum tensor,

$$\tilde{\mathbf{b}} \equiv -(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})\cdot\mathbf{u}\tilde{\nabla} + \mathbf{B}\otimes\mathbf{H}\cdot\mathbf{u}\tilde{\nabla} - (\mathbf{H}\cdot\mathbf{B})\mathbf{u}\tilde{\nabla} + (\tilde{\rho}\tilde{h}+\tilde{\rho}\hat{k})\mathbf{I} \tag{49}$$

the generalized $\tilde{J}$-integral for steady-state crack propagation can be expressed as

$$\tilde{J}_{\tilde{\Gamma}} = \frac{1}{B}\int_{\tilde{\Gamma}}\mathbf{n}\cdot\tilde{\mathbf{b}}d\tilde{\Gamma}\cdot\tilde{\mathbf{e}}_1 = \tilde{\mathbf{J}}_{\tilde{\Gamma}}\cdot\tilde{\mathbf{e}}_1 \tag{50}$$

where $\tilde{\mathbf{J}}_{\tilde{\Gamma}}=(1/B)\int_{\tilde{\Gamma}}\mathbf{n}\cdot\tilde{\mathbf{b}}d\tilde{\Gamma}$ is the configuration force on singularity.

Hence, the generalized $\tilde{J}$-integral is related to the generalized energy-momentum tensor in a way similar to that given in Refs. [3–6]. Nevertheless, the generalized $\tilde{J}$-integral and the generalized energy-momentum tensor constructed with the use of the Helmholtz free energy in this formulation are different from those yielded with the use of the magnetic enthalpy [3,4].

3.3 Extended Essential Work of Fracture Method. From Eq. (40), the total work, W_f, from the start of loading until the final fracture can be partitioned into the essential work of fracture, W_e, and the nonessential work of fracture, W_{ne}, as

$$W_f = W_e + W_{\text{ne}} \tag{51}$$

$$W_e \equiv \int_{t_0}^{t_f} w_e\dot{A}dt = \int_{t_0}^{t_f} R\dot{A}dt \tag{52}$$

$$W_{\text{ne}} = \Delta K + \int_{t_0}^{t_f}\int_{V_t}\rho\hat{s}\dot{T}dVdt + \int_{t_0}^{t_f}\int_{V_t}\rho\hat{\Lambda}dVdt \tag{53}$$

where w_e is the specific essential work of fracture.

The specific essential work of fracture is geometry independent due to its equivalence to crack resistance, and the nonessential

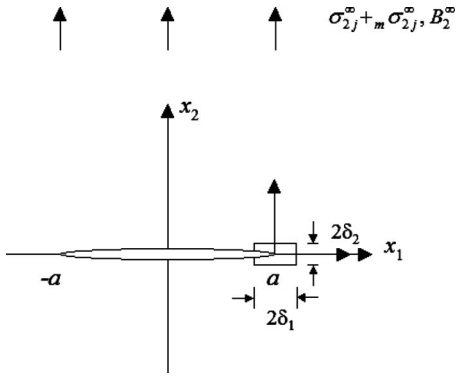


Fig. 1 A Griffith-type crack in a magnetoelastic solid

work of fracture is geometric dependent due to its association with kinetic energy change, thermal effect, and time-dependent dissipation. Hence, this formulation provides a fundamental basis for extending the EWF method, developed by Cotterell, Reddel, Mai, and co-workers [50–52] from the unified theory of fracture by Broberg [53,54], to fracture characterization of magnetosensitive materials involving dynamic, thermal, and hysteresis effects.

It turns out that the generalized energy release rate method, the generalized $\tilde{J}$-integral method, and the extended essential work of fracture method should give consistent results, irrespective of material systems, loading, and environmental conditions. The critical generalized energy release rate, $\tilde{G}_c$, the critical crack-front generalized $\tilde{J}$-integral, $\tilde{J}_c$, and the specific essential work of fracture, w_e, are equivalent as a measure of crack resistance.

4 Griffith-Type Magnetoelastic Crack Problem

4.1 Basic Field Equations. For a quasistatic crack problem in the absence of mechanical body force ($\hat{\mathbf{f}}=0$), Eq. (11) becomes

$$\nabla\cdot(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})=0 \tag{54}$$

With the use of Eqs. (2)–(4) and (54), general solutions exist for the magnetic induction $\mathbf{B}$ and the total stress ${}_t\boldsymbol{\sigma}=\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma}$ with inverse square root singularity at the crack tip as the total stress tensor ${}_t\boldsymbol{\sigma}$ is used instead of the Cauchy stress tensor $\boldsymbol{\sigma}$ for crack problems in magnetoelastic solids [28–32].

4.2 Path-Independent Integral. With the use of $(\mathbf{n}\times\mathbf{H})\times\mathbf{B}=(\mathbf{n}\cdot\mathbf{B})\mathbf{H}-(\mathbf{H}\cdot\mathbf{B})\mathbf{n}$, expression (50) can be rewritten as

$$\tilde{J}_{\tilde{\Gamma}}=-\frac{1}{B}\int_{\tilde{\Gamma}}\mathbf{n}\cdot(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})\cdot\mathbf{u}\tilde{\nabla}\cdot\tilde{\mathbf{e}}_1 d\tilde{\Gamma}+\frac{1}{B}\int_{\tilde{\Gamma}}[(\mathbf{n}\times\mathbf{H})\times\mathbf{B}]\cdot\mathbf{u}\tilde{\nabla}\cdot\tilde{\mathbf{e}}_1 d\tilde{\Gamma}+\frac{1}{B}\int_{\tilde{\Gamma}}(\tilde{\rho}\tilde{h}+\tilde{\rho}\hat{k})\mathbf{n}\cdot\tilde{\mathbf{e}}_1 d\tilde{\Gamma} \tag{55}$$

From Eq. (47), due to the path independence of the generalized $\tilde{J}$-integral for steady-state crack propagation in the absence of mechanical body force, thermal effect, and time-dependent dissipation, the global and local generalized $\tilde{J}$-integrals have the same value as the closed contours are chosen, that is,

$$\tilde{J}_g=\tilde{J}_l=\tilde{J}_0=\tilde{G} \tag{56}$$

4.3 Griffith-Type Crack. Consider a Griffith-type crack in a magnetoelastic solid with the reference coordinate system affixed to the crack tip, as shown in Fig. 1. A generalized plane crack problem deals with a mathematical slit crack or an elliptical cylinder cavitylike crack under coupled magnetomechanical boundary conditions. For simplicity, the remote loads considered here are $\sigma_{2j}^{\infty}+{}_m\sigma_{2j}^{\infty}$ and B_2^{∞}.

For an elliptical inclusion problem, the jump conditions across the interface are given by

$$\mathbf{n}\cdot[[\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma}]]=0,\quad [[\mathbf{u}]]=0 \tag{57}$$

$$\mathbf{n}\cdot[[\mathbf{B}]]=0,\quad \mathbf{n}\times[[\mathbf{H}]]=0 \tag{58}$$

where $[[\cdots]]$ represents the jump of the field quantity inside the double square brackets across the interface.

For a slit crack problem, the crack surface boundary conditions may generally be written as

$$\mathbf{n}^+\cdot(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})^+=-\mathbf{n}^-\cdot(\boldsymbol{\sigma}+{}_m\boldsymbol{\sigma})^-=-\{\sigma_{21}^0+{}_m\sigma_{21}^0,\sigma_{22}^0+{}_m\sigma_{22}^0,\sigma_{23}^0+{}_m\sigma_{23}^0\}^T \tag{59}$$

$$\mathbf{n}^+\cdot\mathbf{B}^+=-\mathbf{n}^-\cdot\mathbf{B}^-=-B_2^0 \tag{60}$$

The total traction should be considered in the coupled magnetomechanical boundary conditions along the crack faces and the remainder surface of the solid. For an elliptical inclusion problem, Eqs. (57) and (58) are exact, but the corresponding boundary value problem needs to be solved in both the solid region and the cavity region [28,29]. For a slit crack problem, B_2^0 in Eq. (60) is either prescribed under magnetically impermeable crack face condition or determined through solving the boundary value problem with magnetically permeable or semipermeable crack face condition [30–33].

Choose the contour as shown in Fig. 1. This is a convenient choice because $n_1=0$ along the segments parallel to the $\tilde{X}_1$-axis. The contour is shrunk onto the crack tip by first letting $\delta_2\to 0$ and then $\delta_1\to 0$. In analogy to the purely elastodynamic case [48], there is no contribution to $\tilde{J}_0$ from the segments parallel to the $\tilde{X}_2$-axis and the segments along crack faces. Moreover, the second term on the right side of Eq. (55) along the segments parallel to the $\tilde{X}_1$-axis vanishes. Consequently, $\tilde{J}_0$ can be computed by evaluating only the first term on the right side of Eq. (55) along the segments parallel to the $\tilde{X}_1$-axis, that is,

$$\tilde{J}_0=2\lim_{\delta_1\to 0}\int_{-\delta_1}^{\delta_1}[\sigma_{2j}(\tilde{X}_1,0)+{}_m\sigma_{2j}(\tilde{X}_1,0)]\frac{\partial u_j(\tilde{X}_1,0^+)}{\partial\tilde{X}_1}d\tilde{X}_1 \tag{61}$$

Hence, the crack-front generalized $\tilde{J}$-integral is equal to the following crack closure integral,

$$\tilde{J}_0=\lim_{\delta a\to 0}\frac{1}{2\delta a}\int_0^{\delta a}[\sigma_{2j}(\tilde{X}_1,0)+{}_m\sigma_{2j}(\tilde{X}_1,0)]\Delta u_j(\tilde{X}_1-\delta a,0^{\pm})d\tilde{X}_1 \tag{62}$$

where $\Delta u_j(\tilde{X}_1-\delta a,0^{\pm})=u_j(\tilde{X}_1-\delta a,0^+)-u_j(\tilde{X}_1-\delta a,0^-)$.

From the near-tip field solutions for magnetoelastic crack problems [28–32] with the replacement of the Cauchy stress tensor by the total stress tensor, the crack-front generalized $\tilde{J}$-integral is calculated as

$$\tilde{J}_0=\tfrac{1}{4}(K_{\mathrm{II}},K_{\mathrm{I}},K_{\mathrm{III}},0)[H](K_{\mathrm{II}},K_{\mathrm{I}},K_{\mathrm{III}},K_B)^T \tag{63}$$

where $K_{\mathrm{I}}=\sqrt{\pi a}[(\sigma_{22}^{\infty}+{}_m\sigma_{22}^{\infty})-(\sigma_{22}^0+{}_m\sigma_{22}^0)]$, $K_{\mathrm{II}}=\sqrt{\pi a}[(\sigma_{21}^{\infty}+{}_m\sigma_{21}^{\infty})-(\sigma_{21}^0+{}_m\sigma_{21}^0)]$, and $K_{\mathrm{III}}=\sqrt{\pi a}[(\sigma_{23}^{\infty}+{}_m\sigma_{23}^{\infty})-(\sigma_{23}^0+{}_m\sigma_{23}^0)]$ are Modes I, II, and III total stress intensity factors, $K_B=\sqrt{\pi a}(B_2^{\infty}-B_2^0)$ is the magnetic induction intensity factor, and $[H]$ is the Irwin matrix.

For the Mode III elliptical cavitylike crack problem by reducing the cavity to a magnetically impermeable, permeable, or semipermeable crack [29], the crack-front generalized $\tilde{J}$-integral becomes

$$\tilde{J}_0 = \tfrac{1}{4}(0,0,K_{\mathrm{III}},0)[H](0,0,K_{\mathrm{III}},K_B)^T \tag{64}$$

For the Mode I slit crack problem with magnetically impermeable, permeable, or semipermeable crack face condition [32], the crack-front generalized $\tilde{J}$-integral becomes

$$\tilde{J}_0 = \tfrac{1}{4}(0,K_{\mathrm{I}},0,0)[H](0,K_{\mathrm{I}},0,K_B)^T \tag{65}$$

The effects of purely applied magnetic field on a magnetically permeable crack in a soft ferromagnetic solid have recently been studied by Gao et al. [33] with inclusion of the Maxwell stresses in the boundary conditions not only on the crack faces but also at infinity. It is found that all the field variables in the solid are uniform, which means that there is no crack-tip field singularity when a mathematical slit crack is dealt with in this case.

5 Discussion

The first feature of this formulation is that Eq. (56) shows that the generalized $\tilde{J}$-integral is path independent for steady-state crack propagation in the absence of mechanical body force, thermal effect, and time-dependent dissipation so that fracture criteria based on the global generalized $\tilde{J}$-integral and the local generalized $\tilde{J}$-integral are identical for such crack problems as the closed contours are chosen.

The second feature of this formulation is that Eq. (62) shows that the crack-front generalized $\tilde{J}$-integral is equal to the mechanical part of the crack closure integral with the replacement of the Cauchy stress tensor by the total stress tensor, which is analogous to the semi-empirical fracture criterion proposed by Park and Sun [18,19] for piezoelectric materials. Nevertheless, the difference lies in the replacement of the Cauchy stress tensor by the total stress tensor and the equivalence of the crack-front generalized $\tilde{J}$-integral to the generalized energy release rate instead of the mechanical strain energy release rate.

The third feature of this formulation is that Eq. (63) shows that the crack-front generalized $\tilde{J}$-integral is an odd function of the magnetic induction intensity factor, which is consistent with experimental observation on crack growth in piezoelectric materials [14–19].

The fourth feature of this formulation is that Eq. (63) indicates that the coupled magnetomechanical boundary conditions affect crack propagation through changing the field intensity factors, which involves the opening and bridging of the crack.

The fifth feature of this formulation is that Eq. (63) indicates that the contribution of pure magnetic loading to the crack driving force through the total stress and magnetic induction intensity factors depends on the adopted crack face condition. As remarked by Gao et al. [33], a conclusion must await further experimental data.

The sixth feature of this formulation is that Eq. (63) implies that the magnetic induction intensity factor has no contribution to the crack driving force in the absence of piezomagnetic effect, but the magnetic stress may still play a role in crack propagation through the total stress intensity factors.

6 Concluding Remarks

Based on fundamental principles of thermodynamics, a nonlinear field theory of fracture mechanics is developed for crack propagation in the presence of magnetothermomechanical coupling and dissipative effects. Through rigorous derivation, this formulation successfully captures the crack-tip singularity of coupled fields, offers the right expression for the crack driving force, and yields the generally applicable fracture criterion for both conservative and dissipative magnetosensitive materials. The developed theory lays a foundation for the application of the generalized energy release rate method, the generalized $\tilde{J}$-integral method, and the extended essential work of fracture method to quasistatic and dynamic fracture characterization of paramagnetic and ferromagnetic materials. The present work can also be extended to include gradient effects in thin films and microelectromechanical systems (MEMS).

Acknowledgment

The author is grateful to the anonymous reviewers for valuable comments and suggestions. In particular, the comments have introduced the author to the work by Gao et al. [33].

References

[1] Shindo, Y., 1977, "The Linear Magnetoelastic Problem for a Soft Ferromagnetic Elastic Solid With a Finite Crack," ASME Trans. J. Appl. Mech., **44**, pp. 47–50.

[2] Shindo, Y., 1978, "Magnetoelastic Interaction of a Soft Ferromagnetic Elastic Solid With a Penny-Shaped Crack in a Constant Axial Magnetic Field," Trans. ASME, J. Appl. Mech., **45**, pp. 291–296.

[3] Sabir, M., and Maugin, G. A., 1996, "On the Fracture of Paramagnets and Soft Ferromagnets," Int. J. Non-linear Mech., **31**, pp. 425–440.

[4] Fomethe, A., and Maugin, G. A., 1998, "On the Crack Mechanics of Hard Ferromagnets," Int. J. Non-Linear Mech., **33**, pp. 85–95.

[5] Maugin, G. A., Epstein, M., and Trimarco, C., 1992, "Pseudomomentum and Material Forces in Inhomogeneous Materials (Application to the Fracture of Electromagnetic Materials in Electromagnetoelastic Fields)," Int. J. Solids Struct., **29**, pp. 1889–1900.

[6] Trimarco, C., 2007, "Material Electromagnetic Fields and Material Forces," Arch. Appl. Mech., **77**, pp. 177–184.

[7] Pak, Y. E., and Hermann, G., 1986, "Conservation Laws and the Material Momentum Tensor for the Elastic Dielectric," Int. J. Eng. Sci., **24**, pp. 1365–1372.

[8] Pak, Y. E., 1990, "Crack Extension Force in a Piezoelectric Material," Trans. ASME J. Appl. Mech., **57**, pp. 647–653.

[9] Maugin, G. A., and Epstein, M., 1991, "The Electroelastic Energy-Momentum Tensor," Proc. R. Soc. London, Ser. A, **433**, pp. 299–312.

[10] Suo, Z., Kuo, C. M., Barnett, D. M., and Willis, J. R., 1992, "Fracture Mechanics for Piezoelectric Ceramics," J. Mech. Phys. Solids, **40**, pp. 739–765.

[11] Dascalu, C., and Maugin, G. A., 1994, "Energy-Release Rates and Path-Independent Integrals in Electroelastic Crack Propagation," Int. J. Eng. Sci., **32**, pp. 755–765.

[12] Maugin, G. A., 1994, "On the J-Integral and Energy-Release Rates in Dynamic Fracture," Acta Mech., **105**, pp. 33–47.

[13] Dascalu, C., and Maugin, G. A., 1995, "On the Dynamic Fracture of Piezoelectric Materials," Q. J. Mech. Appl. Math., **48**, pp. 237–254.

[14] Pak, Y. E., and Tobin, A., 1993, "On Electric Field Effects in Fracture of Piezoelectric Materials," Mechanics of Electromagnetic Materials and Structures, AMD (Am. Soc. Mech. Eng.) **161**, pp. 51–62.

[15] Tobin, A., and Pak, Y. E., 1993, "Effects of Electric Fields on Fracture Behavior of PZT Ceramics," *Smart Materials*, Vol. 1916, V. K. Varadan, ed., SPIE, Bellingham, WA, pp. 78–86.

[16] Cao, H. C., and Evans, A. G., 1994, "Electric-Field-Induced Fatigue Crack Growth in Piezoelectric Ceramics," J. Am. Ceram. Soc., **77**, pp. 1783–1786.

[17] Lynch, C. S., Yang, W., Collier, L., Suo, Z., and McMeeking, R. M., 1995, "Electric Field Induced Cracking in Ferroelectric Ceramics," Ferroelectrics, **166**, pp. 11–30.

[18] Park, S. B., and Sun, C. T., 1993, "Effect of Electric Field on Fracture of Piezoelectric Ceramics," Int. J. Fract., **70**, pp. 203–216.

[19] Park, S. B., and Sun, C. T., 1995, "Fracture Criteria for Piezoelectric Ceramics," J. Am. Ceram. Soc., **78**, pp. 1475–1480.

[20] Gao, H., Zhang, T.-Y., and Tong, P., 1997, "Local and Global Energy Release Rate for an Electrically Yielded Crack in a Piezoelectric Ceramic," J. Mech. Phys. Solids, **45**, pp. 491–510.

[21] Fulton, C. C., and Gao, H., 2001, "Effect of Local Polarization Switching on Piezoelectric Fracture," J. Mech. Phys. Solids, **49**, pp. 927–952.

[22] Li, S., 2003, "On Global Energy Release Rate of a Permeable Crack in a Piezoelectric Ceramic," ASME J. Appl. Mech., **70**, pp. 246–252.

[23] Zhang, T. Y., Zhao, M. H., and Gao, C. F., 2005, "The Strip Dielectric Breakdown Model," Int. J. Fract., **132**, pp. 311–327.

[24] McMeeking, R. M., 2001, "Towards a Fracture Mechanics for Brittle Piezoelectric and Dielectric Materials," Int. J. Fract., **108**, pp. 25–41.

[25] McMeeking, R. M., 2004, "The Energy Release Rate for a Griffith Crack in a Piezoelectric Material," Eng. Fract. Mech., **71**, pp. 1149–1163.

[26] Lin, C. B., and Yeh, C. S., 2002, "The Magnetoelastic Problem of a Crack in a Soft Ferromagnetic Solid," Int. J. Solids Struct., **39**, pp. 1–17.

[27] Liang, W., Fang, D., Shen, Y., and Soh, A. K., 2002, "Nonlinear Magnetoelastic Coupling Effects in a Soft Ferromagnetic Material With a Crack," Int. J. Solids Struct., **39**, pp. 3997–4011.

[28] Gao, C.-F., Kessler, H., and Balke, H., 2003, "Crack Problems in Magnetoelectroelastic Solids. Part I: Exact Solution of a Crack," Int. J. Eng. Sci., **41**, pp. 969–981.

[29] Gao, C.-F., Tong, P., and Zhang, T.-Y., 2004, "Fracture Mechanics for a Mode III Crack in a Magnetoelectroelastic Solid," Int. J. Solids Struct., **41**, pp. 6613–6629.
[30] Wang, B. L., and Mai, Y.-W., 2003, "Cracking of Electromagnetic Elastic Solids," Key Eng. Mater., **251–252**, pp. 303–312.
[31] Wang, B. L., and Mai, Y.-W., 2004, "Fracture of Piezoelectromagnetic Materials," Mech. Res. Commun., **31**, pp. 65–73.
[32] Wang, B. L., and Mai, Y.-W., 2007, "Applicability of the Crack-Face Electromagnetic Boundary Conditions for Fracture of Magnetoelectroelastic Materials," Int. J. Solids Struct., **44**, pp. 387–398.
[33] Gao, C.-F., Mai, Y.-W., and Wang, B.-L., 2008, "Effects of Magnetic Fields on Cracks in a Soft Ferromagnetic Material," Eng. Fract. Mech., **75**, pp. 4863–4875.
[34] Eshelby, J. D., 1951, "The Force on an Elastic Singularity," Philos. Trans. R. Soc. London, Ser. A, **244**, pp. 87–112.
[35] Eshelby, J. D., 1970, "Energy Relations and the Energy-Momentum Tensor in Continuum Mechanics," *Inelastic Behavior of Solids*, M. F. Kanninen, ed., McGraw-Hill, New York, pp. 77–114.
[36] Eshelby, J. D., 1975, "The Elastic Energy-Momentum Tensor," J. Elast., **5**, pp. 321–335.
[37] Trueddell, C., and Noll, W., 2004, *The Non-Linear Field Theories of Mechanics*, 3rd ed., S. S. Antman, ed., Springer-Verlag, Berlin.
[38] Fung, Y. C., and Tong, P., 2001, *Classical and Computational Solid Mechanics*, World Scientific, Singapore.
[39] Eringen, A. C., 1980, *Mechanics of Continua*, 2nd ed., Robert E. Krieger Publishing Company, Inc., Malabar, FL.
[40] Maugin, G. A., 1988, *Continuum Mechanics of Electromagnetic Solids*, North-Holland, Amsterdam.
[41] Maugin, G. A., 1992, *The Thermomechanics of Plasticity and Fracture*, Cambridge University Press, Cambridge, UK.
[42] Dorfmann, A., and Ogden, R. W., 2004, "Nonlinear Magnetoelastic Deformations," Q. J. Mech. Appl. Math., **57**, pp. 599–622.
[43] Chen, X., 2007, "Coupled Hygro-Thermo-Viscoelastic Fracture Theory," Int. J. Fract., **148**, pp. 47–55.
[44] Chen, X., 2009, "Crack Driving Force and Energy-Momentum Tensor in Electroelastodynamic Fracture," J. Mech. Phys. Solids, **57**, pp. 1–9.
[45] Chen, X., 2008, "Nonlinear Field Theory of Fracture Mechanics for Piezoelectric and Ferroelectric Materials, unpublished.
[46] Kanninen, M. F., and Popelar, C. H., 1985, *Advanced Fracture Mechanics*, Oxford University Press, New York.
[47] Christensen, R. M., 1982, *Theory of Viscoelasticity—An Introduction*, 2nd ed., Academic, New York.
[48] Freund, L. B., 1998, *Dynamic Fracture Mechanics*, Cambridge University Press, Cambridge, UK.
[49] Schapery, R. A., 1975, "A Theory of Crack-Initiation and Growth in Viscoelastic Media," Int. J. Fract., **11**, pp. 141–159.
[50] Cotterell, B., and Reddel, J. K., 1977, "The Essential Work of Plane Stress Ductile Fracture," Int. J. Fract., **12**, pp. 267–277.
[51] Mai, Y.-W., and Cotterell, B., 1986, "On the Essential Work of Ductile Fracture in Polymers," Int. J. Fract., **32**, pp. 105–125.
[52] Mai, Y.-W., Wong, S.-C., and Chen, X.-H., 2000, "Application of Fracture Mechanics for Characterization of Toughness of Polymer Blends," *Polymer Blends*, D. R. Paul and C. B. Bucknall, eds., Wiley, New York, pp. 17–58.
[53] Broberg, K. B., 1971, "Crack Growth Criteria and Non-Linear Fracture Mechanics," J. Mech. Phys. Solids, **19**, pp. 407–418.
[54] Broberg, K. B., 1975, "On Stable Crack Growth," J. Mech. Phys. Solids, **23**, pp. 215–237.

Acoustic Radiation From a Laminated Composite Plate Excited by Longitudinal and Transverse Mechanical Drives

X. W. Yin[1]
Wuxi Branch,
Jiangsu Institute of Special Equipment Safety Supervision and Inspection,
Wuxi, Jiangsu 214174, China
e-mail: x.w.yin@163.com

H. F. Cui
Wuxi Institute of Technology,
Wuxi, Jiangsu 214073, China

*Analytical expressions are derived for the vibration and far-field acoustic radiation from a fluid-loaded laminated composite plate, which is excited by three types of mechanical drives: point, line, and uniformly distributed forces in longitudinal and transverse directions. Through numerical results, it is shown that the lamination schemes of the laminated composite plate play an important role in its far-field acoustic radiation, especially when the plies of the plate are asymmetric with its middle plane. This paper would help better understand the main mechanism of acoustic radiation from laminated composite plates, which has not been adequately addressed in its companion paper (Yin, et al., 2007, "Acoustic Radiation From a Laminated Composite Plate Reinforced by Doubly Periodic Parallel Stiffeners," J. Sound Vibrat., **306**, pp. 877–889).* [DOI: 10.1115/1.3086429]

1 Introduction

It is well known that a vibrating plate may exhibit three types of modes (or waves) [1–8], namely, longitudinal, shear, and bending. Of these three modes, bending is referred to as out-of-plane motion, and longitudinal and shear, when grouped together, as in-plane motions. In general, the bending motions of the plate can be sufficiently described by its transverse displacement, while the longitudinal (or extensional) and shear motions can be usually described by two in-plane displacement components.

In the past few decades, in contrast to the in-plane motions of plates, more attention was paid to the bending vibration [3–7], not only because the bending is prone to excitation from external sources but also because such motion can interact strongly with the surrounding acoustic medium. In plate bending theories, the classical Kirchhoff–Love assumptions are often employed, from which the plate suffers no extension in its middle plane as well as in the transverse direction. This has the consequence that the structural response to a driving force normal to the plate is solely determined by the bending wave equation and there is no coupling to the in-plane vibrations. In most of the publications, acoustic radiation from a uniform plate is mainly attributed to its bending motions, and hence the corresponding transverse displacement is considered to be the unique displacement component that makes contribution to the induced acoustic radiation.

However, in engineering applications, plates are usually joined together through various types of connections. In this case, the above mentioned three types of motions may be converted into each other [8]. So the prediction of the in-plane vibrations of the plates is also of practical importance in the sense of motion (or wave) conversion.

Moreover, with the knowledge of plate theory, there is one acknowledged concept governing the coupling between the in-plane and out-of-plane vibrations. For thin plates [4], whose rotatory inertia and shear deformations are assumed to be ignored, their bending motions are decoupled from extensional and shear motions, while for moderately thick or even thicker plates, their displacements are fairly coupled, e.g., through Poisson's ratio and the corrected shear modulus. In this case, only thick plate theories [8] or even more accurate three dimensional elastic theories can work well.

With the increasingly extensive use of laminated composite materials, the dynamics of plates made of such materials is more complex than that of uniform plates [9]. In the classical laminated plate theory (CLPT) [10–13], the laminates are also assumed to be transversely inextensible; nevertheless, their transverse and in-plane displacements are coupled through the extension-bending coupling stiffnesses B_{ij} (see Appendix A in Ref. [14]) in a mathematical sense. For instance, for an asymmetric laminate whose ply stacking sequence, material, and geometry (i.e., ply thickness) are not symmetric with its middle plane, its extension and bending may be coupled not only through coupling stiffnesses but also through Poisson's effects. In particular, for a symmetric laminate, its extension-bending coupling stiffnesses become zero, that is, the extension and bending vibrations are decoupled from each other (if Poisson's effects are of insignifance, e.g., the plates are thin enough). Many articles [9,14] are concentrated in the problem of acoustic radiation from a composite laminated plate mainly due to transverse vibration, and little attention is paid to the effects of in-plane motions on the overall vibration and the acoustic radiation.

Particularly, in our previous work [14], when the composite laminated plates are driven by a harmonic transverse point force, there is almost no difference between the sound pressure levels from the symmetric and antisymmetric laminates, the reason for which was not addressed. In this paper, the work in Ref. [14] is extended to develop solutions to the vibration and acoustic radiation from a laminated composite plate, which is driven by point, line, and uniformly distributed forces in longitudinal and transverse directions. Through theoretical expressions and numerical results, we try to give a further discussion on the underlying mechanism of sound radiation from composite laminated plates driven by different types of mechanical loadings in different directions.

2 Theoretical Sketch

Consider an infinite laminated composite plate composed of N orthotropic layers, the mth one of which has principal material coordinates $(x^{(m)}, y^{(m)}, z^{(m)})$ oriented along a vector with an angle $\theta^{(m)}$ to the global laminate coordinate, x, as shown in Fig. 1. An acoustic fluid occupies the half-space $z>0$, which interacts with the plate at the interface between them. Since the thickness of the plate is very small compared with its other two dimensions, the fluid-plate interface is considered to be lying on the plane $z=0$ without loss of accuracy. Three types of forces are applied at the plate in its middle plane, namely, point forces F_k, line forces N_k, and uniformly distributed forces p_{vk}, where the lower index k denotes the x-, y-, and z-component of the quantities corresponding to $k=1,2,3$, respectively. For problems of harmonic vibration, all of the loadings have a time dependence $e^{i\omega t}$, where ω is the circular frequency, and $e^{i\omega t}$ will henceforth be suppressed throughout.

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received August 21, 2007; final manuscript received December 9, 2008; published online April 21, 2009. Review conducted by Professor Sridhar Krishnaswamy.

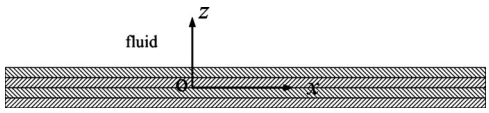


Fig. 1 An infinite laminated composite plate with Cartesian coordinate system

The classical laminated plate theory is employed to describe the motions of a fluid-loaded laminated composite plate in a concise matrix form [13,14]:

$$\begin{bmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{Bmatrix} u_0 \\ v_0 \\ w \end{Bmatrix} = \begin{Bmatrix} -F_1 - N_1 - p_{v1}(x,y) \\ -F_2 - N_2 - p_{v2}(x,y) \\ [F_3 + N_3 + p_{v3}(x,y) - p_a(x,y,0)] \end{Bmatrix} \tag{1}$$

where L_{kl} $(k=1,2,3;\ l=1,2,3)$ are differential operators for a thin laminated composite plate, whose detailed expressions were given in Appendix A of Ref. [14]. u_0 and v_0 are the in-plane displacements in the plate's middle plane, and w is the transverse displacement of the plate. It is generally noted that $p_a(x,y,z)$ is the acoustic pressure in the fluid; in particular, $p_a(x,y,0)$ is the reactive pressure acting on the plate.

In the governing equation (1), the external point forces acting at (x_k, y_k) are defined as follows [4]:

$$F_k = Q_k \delta(x - x_k)\,\delta(y - y_k) \quad (k = 1,2,3) \tag{2a}$$

and the line forces at $x = x_k$ are

$$N_k = T_k \delta(x - x_k) \quad (k = 1,2,3) \tag{2b}$$

where $\delta(\cdot)$ is the Dirac delta function.

Similarly, uniformly distributed forces, which act on a rectangle with dimension of $2L_x \times 2L_y$ along the x-, y- and z-directions, respectively, are expressed as follows:

$$p_{vk}(x,y) = Q_{vk} H(L_x - |x - x_{vk}|) H(L_y - |y - y_{vk}|) \quad (k = 1,2,3) \tag{2c}$$

where $H(\cdot)$ denotes the Heaviside function. The center of the rectangle is located at (x_{vk}, y_{vk}) $(k=1,2,3)$.

Taking the Fourier transform of Eq. (1) by the following transform pair

$$\tilde{w}(\alpha,\beta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(x,y) e^{i(\alpha x + \beta y)} dx dy \tag{3a}$$

$$w(x,y) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{w}(\alpha,\beta) e^{-i(\alpha x + \beta y)} d\alpha d\beta \tag{3b}$$

yields

$$\begin{bmatrix} \tilde{L}_{11} & \tilde{L}_{12} & \tilde{L}_{13} \\ \tilde{L}_{21} & \tilde{L}_{22} & \tilde{L}_{23} \\ \tilde{L}_{31} & \tilde{L}_{32} & \tilde{L}_{33} \end{bmatrix} \begin{Bmatrix} \tilde{u}_0 \\ \tilde{v}_0 \\ \tilde{w} \end{Bmatrix} = \begin{Bmatrix} -\tilde{F}_1 - \tilde{N}_1 - \tilde{p}_{v1} \\ -\tilde{F}_2 - \tilde{N}_2 - \tilde{p}_{v2} \\ \tilde{F}_3 + \tilde{N}_3 + \tilde{p}_{v3} - \tilde{p}_a \end{Bmatrix} \tag{4}$$

where $\tilde{L}_{kl}(k=1,2,3;\ l=1,2,3)$ are the transformed algebraic operators for a thin laminated composite plate, which were given in Appendix B of Ref. [14].

2.1 The Transformed External Forces. The transformed point forces in Eq. (4) are also readily written as follows:

$$\tilde{F}_k = Q_i e^{i(\alpha x_k + \beta y_k)} \quad (k = 1,2,3) \tag{5a}$$

Similarly, the transformed line forces in Eq. (4) are

$$\tilde{N}_k = T_k e^{i\alpha x_k} \quad (k = 1,2,3) \tag{5b}$$

and the transformed surface loadings in Eq. (4) are also given as follows:

$$\tilde{p}_{vk}(\alpha,\beta) = 4Q_{vk} \sin(\alpha L_x) \sin(\beta L_y)/(\alpha\beta) \quad (k = 1,2,3) \tag{5c}$$

2.2 The Transformed Fluid Loading. In Cartesian coordinates, the acoustic pressure in the half-infinite space above $z>0$ satisfies the Helmholtz equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) p_a + \left(\frac{\omega^2}{c_0^2}\right) p_a = 0 \tag{6}$$

where c_0 is the speed of sound in the fluid. The coupling between the plate and the fluid satisfies the momentum equation in z axis

$$\left.\frac{\partial p_a}{\partial z}\right|_{z=0} = \omega^2 \rho_0 w \tag{7}$$

where ρ_0 is the density of the fluid. Taking the Fourier transform of Eqs. (6) and (7) yields

$$\tilde{p}_a(\alpha,\beta,0) = -\omega^2 \rho_0 \tilde{w}(\alpha,\beta)/\gamma(\alpha,\beta) \tag{8}$$

where

$$\gamma^2 = \alpha^2 + \beta^2 - \omega^2/c_0^2 \tag{9}$$

and γ is to be evaluated such that $\mathrm{Re}(\gamma) \geq 0$, $\mathrm{Im}(\gamma) \geq 0$ if $\mathrm{Re}(\gamma) = 0$, in order that the radiation conditions for outgoing waves are met.

3 Solution of the Transformed Equations

Substituting Eq. (8) into Eq. (4) and gathering up the terms involving the transformed transverse displacement $\tilde{w}$ lead to

$$\begin{bmatrix} \tilde{L}_{11} & \tilde{L}_{12} & \tilde{L}_{13} \\ \tilde{L}_{21} & \tilde{L}_{22} & \tilde{L}_{23} \\ \tilde{L}_{31} & \tilde{L}_{32} & \tilde{L}_{33} - \omega^2\rho_0/\gamma \end{bmatrix} \begin{Bmatrix} \tilde{u}_0 \\ \tilde{v}_0 \\ \tilde{w} \end{Bmatrix} = \begin{Bmatrix} -\tilde{F}_1 - \tilde{N}_1 - \tilde{p}_{v1} \\ -\tilde{F}_2 - \tilde{N}_2 - \tilde{p}_{v2} \\ \tilde{F}_3 + \tilde{N}_3 + \tilde{p}_{v3} \end{Bmatrix} \tag{10}$$

By solving Eq. (10), the transformed transverse displacement $\tilde{w}(\alpha,\beta)$ is

$$\tilde{w}(\alpha,\beta) = (-(\tilde{F}_1 + \tilde{N}_1 + \tilde{p}_{v1})A_{13} - (\tilde{F}_2 + \tilde{N}_2 + \tilde{p}_{v2})A_{23} + (\tilde{F}_3 + \tilde{N}_3 + \tilde{p}_{v3})A_{33})\Delta^{-1} \tag{11}$$

where the dispersion denominator Δ is the determinant of the 3×3 matrix on the left-hand side of Eq. (10). To save writing, we introduce three elements, i.e., A_{13}, A_{23}, and A_{33}, which were listed in Appendix C of Ref. [14]. The plate's lamination schemes, i.e., material, geometry, and ply stacking sequence are also found implicit in the four factors.

The far-field acoustic radiation in spherical coordinates (R, θ, ϕ) can be given with the application of one standard procedure, i.e., stationary phase approximation

$$P(R,\theta,\phi) = -\rho_0 \omega^2 \tilde{w}(\alpha_0,\beta_0) e^{-ik_0 R}/2\pi R \tag{12}$$

where the stationary point is defined as follows:

$$\alpha_0 = (\omega/c_0)\sin\theta\cos\phi, \quad \beta_0 = (\omega/c_0)\sin\theta\sin\phi \tag{13}$$

and R is the distance from the field point to the origin. It is clear from Eq. (12) that the far-field pressure at an observation point (R, θ, ϕ) only involves contribution from transformed displacement at a single wavenumber pair (α_0, β_0), which is specified by Eq. (13), and the contribution from all other wavenumbers may be filtered out due to the rapidly oscillatory integral in the expression for the far field. By the way, it will be noted that the origin of the spherical coordinates shall be coincident with the origin of the plate's coordinates. Besides, the θ and ϕ axes shall be orientated along the x and z axes for the plate's coordinates, respectively.

Table 1 Parameter values of the laminated composite plates

Ply No.	$E_x^{(k)}$ (Pa)	$E_y^{(k)}$ (Pa)	$\nu_{xy}^{(k)}$	$\nu_{yx}^{(k)}$	$\rho^{(k)}$ (kg/m^3)	$h^{(k)}$ (m)	Ply angle (deg)
1	3×10^{10}	3×10^{11}	0.03	0.3	7800	0.0015	75
2	2×10^{10}	2×10^{11}	0.03	0.3	7800	0.0015	60
3	1×10^{10}	1×10^{11}	0.03	0.3	7800	0.0015	45
4	1×10^{10}	1×10^{11}	0.03	0.3	7800	0.0015	±45
5	2×10^{10}	2×10^{11}	0.03	0.3	7800	0.0015	±60
6	3×10^{10}	3×10^{11}	0.03	0.3	7800	0.0015	±75

4 Numerical Examples and Discussion

In our previous paper [14], when symmetric and antisymmetric laminates are driven by a point force, the curves for their far-field sound radiation levels are almost the same; the problems such as on what cases their curves would display differences have not been addressed. In this section, numerical results based on Eq. (12) are presented for the far-field acoustic radiation from a laminated composite plate. As depicted in Table 1, two lamination schemes, i.e., symmetric (75/60/45/45/60/75) and antisymmetric (75/60/45/–45/–60/–75) plies are used for numerical calculation.

As a standard procedure, the sound pressure levels at the observation point are normalized with reference to 1 μPa pressure and corrected to the field point ($R_0=1.0$ m, $\theta_0=\phi_0=45$ deg).

When we examine the elements $\tilde{L}_{kl}$ in Eq. (4) through the laminated composite plate theory (see Appendixes A and B in Ref. [14]), the bending of a laminated composite plate includes the contribution of the in-plane motions due not only to Poisson's effect but also to the particular feature of the laminated plate, i.e., extension-bending coupling. In order to induce different types of motions, i.e., bending and in-plane vibrations, several load cases, namely, point, line, and uniformly distributed forces in different directions are applied to the plates. In the numerical calculation, the driving points for the point, line forces are located at (0.0, 0.0) and 0.0, respectively, and the driving rectangle for uniformly distributed forces is with dimension of 0.01×0.01 and centered on the point (0.0, 0.0).

As shown in Figs. 2, 5, and 8, the symmetric and antisymmetric laminates are excited by longitudinal forces; in this case, the in-plane displacement components would be excited and dominant compared with transverse displacement. For further discussion, the plates are also driven by transverse forces, as well as by a combination of in-plane and transverse forces.

In Fig. 2, the numerical results for the acoustic radiation are presented for the plates driven by a longitudinal point force $Q_1=1.0$ N. The far-field sound pressure levels for the antisymmetric laminates are far greater than those for the symmetric laminates. Nevertheless, when the plates are driven by a transverse point force $Q_3=1.0$ N, the curves in Fig. 3 show that the sound pressure levels for the antisymmetric laminates are almost identical to those for the symmetric laminates.

When the driving point forces are the combination of the longitudinal and transverse forces, there is one remarkable feature arresting our attention. Shown in Fig. 4, the curves have almost no differences with those in Fig. 3, which confirms in an explicit numerical way that the contribution from the in-plane motions has trivial impact on the total acoustic field, and this is why almost all of the publications are dedicated to the sound radiation induced by transverse motions.

What should be pointed out is that the superposition principle in linear acoustics still works here. Since the sound pressure level is a logarithmic quantity, and the sound pressures for the plates driven by in-plane forces is far smaller than those for the plates driven by transverse forces, as shown in Figs. 2 and 3, the differences of the sound pressure levels for the plates driven by in-plane forces, as shown in Fig. 2, are indistinguishable in Fig. 4.

Shown in Figs. 6 and 7, the plates are driven by a transverse line force $T_3=1.0$ N/m and by a combination of a longitudinal line force $T_1=1.0$ N/m and a transverse line force $T_3=1.0$ N/m, respectively. Shown in Figs. 9 and 10, the plates are driven by a transverse distributed force $Q_{v3}=1.0$ N/m^2 and by a combination of a longitudinal distributed force $Q_{v1}=1.0$ N/m^2 and a transverse distributed force $Q_{v3}=1.0$ N/m^2, respectively.

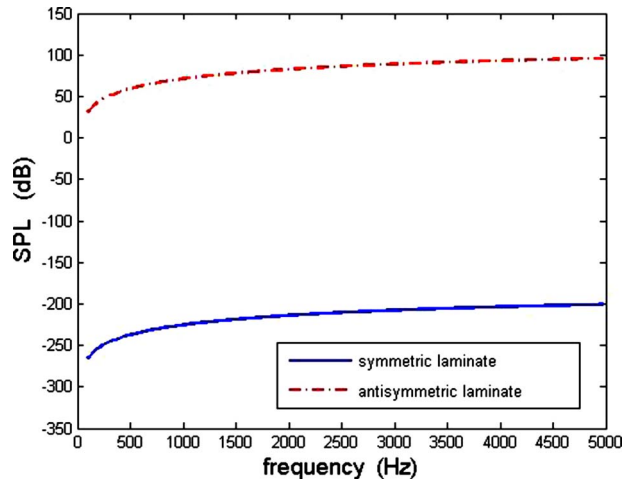


Fig. 2 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates ($Q_1=1.0$ N)

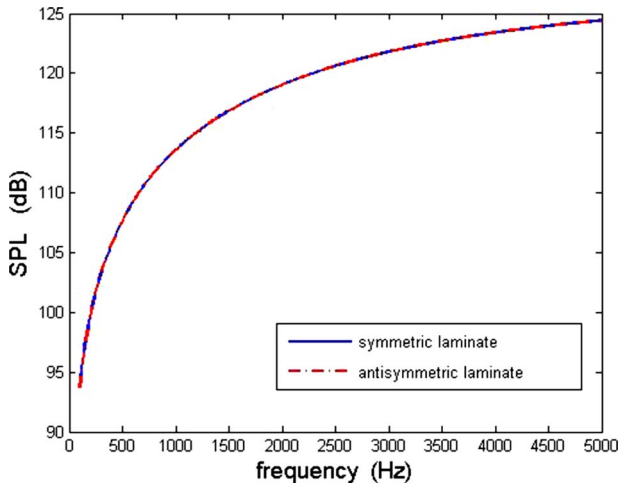


Fig. 3 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates ($Q_3=1.0$ N)

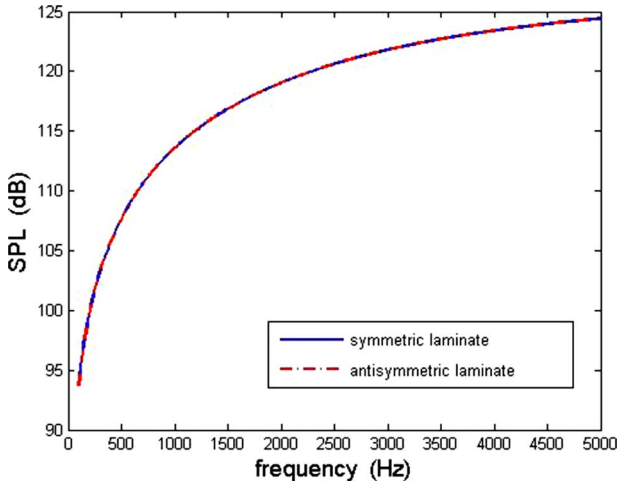


Fig. 4 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates (Q_1=1.0 N and Q_3=1.0 N)

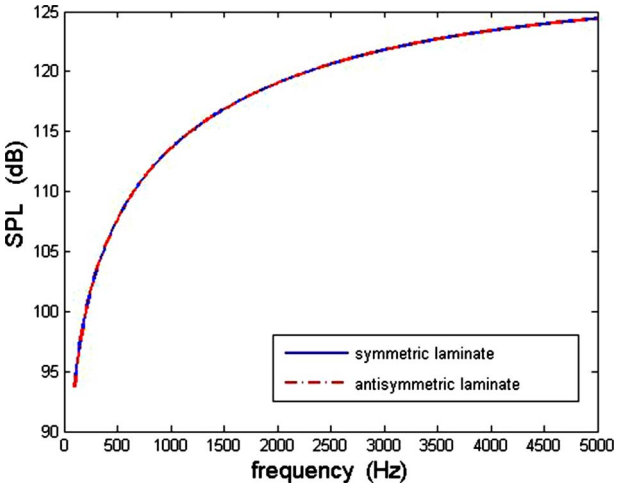


Fig. 6 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates (T_3=1.0 N/m)

The curves in Figs. 5–7 and in Figs. 8–10 show the similar aforementioned features, which correspond to the plates driven by line and distributed forces, respectively.

From Figs. 2–10, we make the following concluding remarks:

(1) As stated in Refs. [13,14], the extension-bending coupling stiffnesses B_{ij} are equal to zero for symmetric laminates. So for this type of symmetric laminates, the extensional and bending motions are coupled only through Poisson's ratios; in particular, if Poisson's ratios are small enough, e.g., the condition of the well known thin plate or laminate definition is satisfied, they can be fully decoupled from each other. This is why the sound radiation levels for the symmetric laminates driven by longitudinal forces are far more smaller than those for the antisymmeric laminates, as shown in Figs. 2, 5, and 8.

(2) While for asymmetric laminates, their extension-bending coupling stiffnesses B_{ij} are not equal to zero, their extensional and bending motions are coupled. However, this coupling effect is relatively small in some cases, as shown in Figs. 4, 7, and 10, since the curves in these figures are almost identical to those in Figs. 3, 6, and 9; that is, the additional point, line, and distributed forces have little impact in the sound pressure levels shown in Figs. 4, 7, and 10. Nevertheless, this coupling effect in the dynamics of asymmetric laminates shall be dealt with carefully once the longitudinal motions are dominant, as shown in Figs. 2, 5, and 8, which remind us that if the longitudinal driving forces in the asymmetric laminates are far greater than the transverse forces, the extensional-bending coupling must be taken into account. For example, if we make a measurement on the sound pressure from a ship hull, which is made of asymmetric laminate material and driven by a propeller, this effect shall be considered beforehand.

(3) Whatever type of composite laminated plates, symmetric or asymmetric, are driven by transverse forces, as well as by a combination of longitudinal and transverse forces, the bending of the plates would be so dominant in the acoustic field that the contribution from the in-plane motions is trivial. In these cases, the lamination schemes of the laminates are of little significance in its vibration and acoustic radiation.

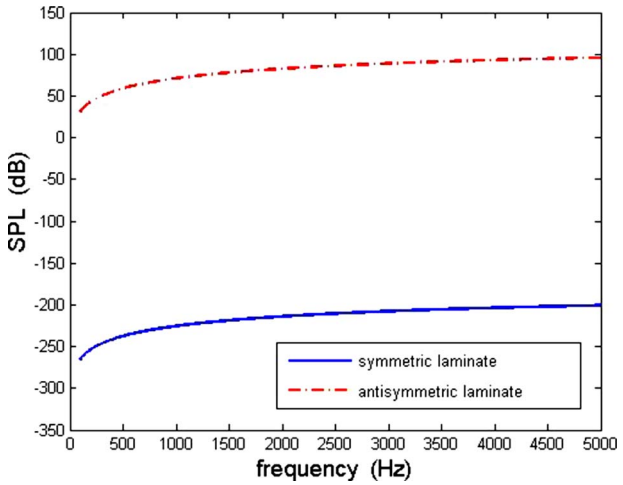


Fig. 5 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates (T_1=1.0 N/m)

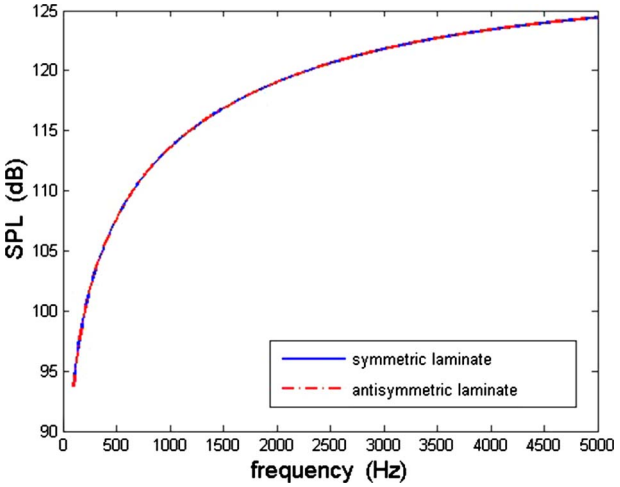


Fig. 7 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates (T_1=1.0 N/m and T_3=1.0 N/m)

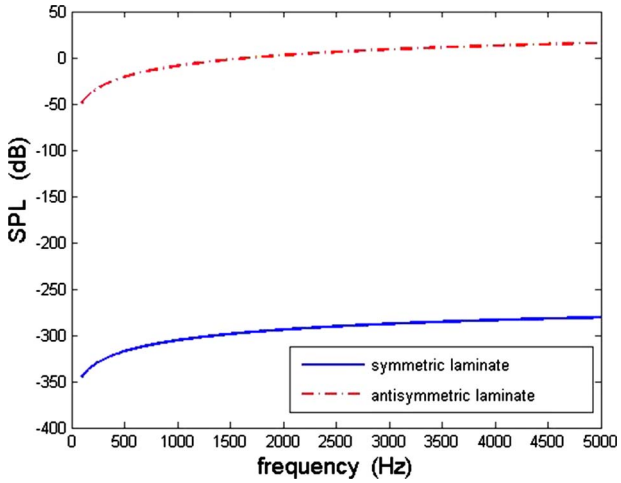


Fig. 8 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates (Q_{v1}=1.0 N/m^2)

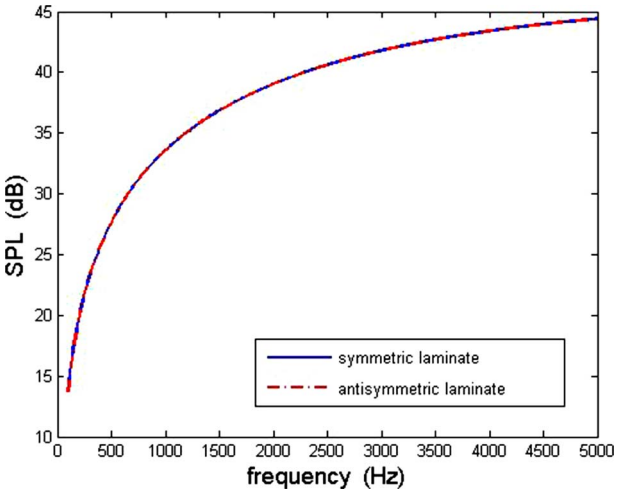


Fig. 10 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates (Q_{v1}=1.0 N/m^2 and Q_{v3}=1.0 N/m^2)

5 Conclusions

Analytical expressions are derived for the vibration and far-field acoustic radiation from a fluid-loaded laminated composite plate, which is excited by three types of mechanical drives: point, line, and uniformly distributed forces in the longitudinal and transverse directions. This work is an extension of that done by Yin et al. [14]. Through numerical results, it is shown that the lamination schemes of the laminated composite plate play an important role in its far-field acoustic radiation, especially when the plies of the plate are asymmetric with its middle plane.

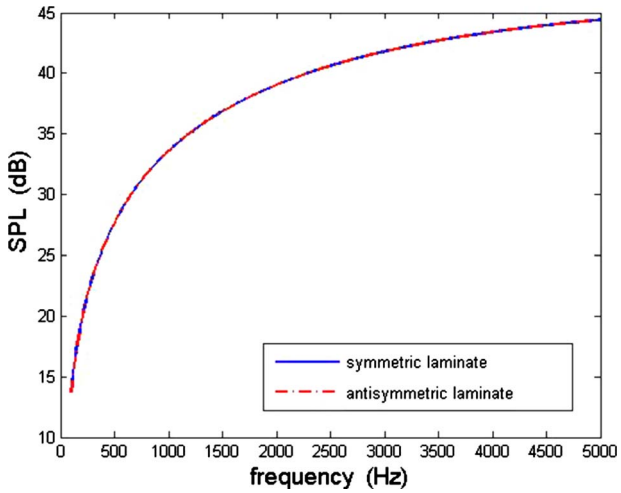


Fig. 9 Numerical results for the far-field acoustic radiation from symmetric and antisymmetric laminates (Q_{v3}=1.0 N/m^2)

References

[1] Gorman, D. J., 2005, "Free In-Plane Vibration Analysis of Rectangular Plates With Elastic Support Normal to the Boundaries," J. Sound Vib., **285**, pp. 941–966.

[2] Gorman, D. J., 2006, "Exact Solutions for the Free In-Plane Vibration of Rectangular Plates With Two Opposite Edges Simply Supported," J. Sound Vib., **294**, pp. 131–161.

[3] Smith, J. D., 2007, "Symmetric Wave Corrections to the Line Driven, Fluid Loaded, Thin Elastic Plate," J. Sound Vib., **305**, pp. 827–842.

[4] Mace, B. R., 1980, "Sound Radiation From a Plate Reinforced by Two Sets of Parallel Stiffeners," J. Sound Vib., **71**, pp. 435–441.

[5] Mead, D. J., and Yaman, Y., 1991, "The Harmonic Response of Rectangular Sandwich Plates With Multiple Stiffening: A Flexural Wave Analysis," J. Sound Vib., **145**, pp. 409–428.

[6] Rumerman, H. M. L., 1975, "Vibration and Wave Propagation in Ribbed Plates," J. Acoust. Soc. Am., **57**, pp. 370–373.

[7] Keltie, R. F., 1993, "Structural Acoustic Response of Finite Rib-Reinforced Plates," J. Acoust. Soc. Am., **94**, pp. 880–887.

[8] Cuschieri, J. M., and McCollum, M. D., 1996, "In-Plane and Out-of-Plane Waves' Power Transmission Through an L-Shape Junction Using the Mobility Power Flow Approach," J. Acoust. Soc. Am., **100**, pp. 857–870.

[9] Chakrabarti, A., and Sheikh, A. H., 2007, "Vibration of Composites and Sandwich Laminates Subjected to In-Plane Partial Edge Loading," Compos. Sci. Technol., **67**, pp. 1047–1057.

[10] Hwang, Y. F., Kim, M., and Zoccola, P. J., 2000, "Acoustic Radiation by Point- or Line-Excited Laminated Plates," ASME J. Vibr. Acoust., **122**, pp. 189–195.

[11] Reddy, J. N., 1984, "A Simple Higher-Order Theory for Laminated Composite Plates," ASME J. Appl. Mech., **51**, pp. 745–752.

[12] Reddy, J. N., and Liu, C. F., 1985, "A Higher-Order Shear Deformation Theory for Laminated Elastic Shells," Int. J. Eng. Sci., **23**, pp. 319–330.

[13] Reddy, J. N., 1997, *Mechanics of Laminated Composite Plates: Theory and Analysis*, CRC, Boca Raton, FL.

[14] Yin, X. W., Gu, X. J., Cui, H. F., and Shen, R. Y., 2007, "Acoustic Radiation From a Laminated Composite Plate Reinforced by Doubly Periodic Parallel Stiffeners," J. Sound Vib., **306**, pp. 877–889.

Digital Photothermoelastic and Numerical Analysis of Transient Thermal Stresses in Cracked Bimaterial Interfaces

B. Neethi Simon
e-mail: neethisimon@gmail.com

K. Ramesh
e-mail: kramesh@iitm.ac.in

Department of Applied Mechanics,
IIT Madras,
Chennai - 600 036, India

Transient thermal stresses of a bimaterial specimen with interface cracks, under uniform cooling by convection, are analyzed by photothermoelasticity and a coupled temperature-displacement, finite element scheme. The stress intensity factors of the interface crack are determined by a multiparameter overdeterministic system of equations in a least-squares sense using the experimental data and by J-integral, numerically. The study showed that a normal temperature variation can lead to significant stresses due to the mismatch of thermal expansion coefficients.
[DOI: 10.1115/1.3086782]

Keywords: photothermoelasticity, bimaterial interface, SIF evaluation, transient thermal loading

1 Introduction

Photoelasticity is a whole-field optical experimental technique, which provides the principal stress difference and the orientation of principal stress direction. With the advent of digital photoelasticity, the analysis of two-dimensional (2D) problems has become faster and more accurate. In photothermoelasticity, thermal, rather than mechanical, loads are used. Several investigators have successfully used photothermoelasticity to analyze transient thermal stresses for homogeneous bodies [1].

Bimaterial interfaces occur in several components of engineering importance such as electronic packages and bimetal thermostats, and bimaterial strips are simplified models used in the literature to analyze these. A number of theoretical studies on bimaterial strips subjected to thermal loading conditions have been done [2]. Wang and Hsu [2] investigated the steady state thermal stresses in bimaterial interfaces, photoelastically and numerically. During thermal cycling of bimaterial specimens, the induced thermal stresses are predominant near the edges of the interface, which may eventually cause fatigue failure by initiating interface cracks [3]. Kokini and Smith [1] studied the interfacial crack problem using photothermoelasticity and evaluated the transient stress intensity factors (SIFs) for an initially cool, edge-heated polycarbonate-brass interface, pasted using an epoxy adhesive. They observed a drop in the value of K_I, after an initial steady increase.

In this paper, a bimaterial specimen with symmetric interface edge cracks, subjected to whole-field cooling by convection, is studied. This is of key importance because almost all engineering components are subjected to cyclic variations in temperature between day and night, and get heated or cooled by convective heat transfer. The transient thermal stresses are recorded experimentally using digital photothermoelasticity. A transient coupled temperature-displacement 2D finite element (FE) analysis is also carried out using a standard FE package and is validated by plotting photoelastic fringe contours, using a postprocessing software developed in-house [4]. The transient thermal SIFs are evaluated using the *J*-integral and compared with those obtained experimentally from photothermoelasticity using the method of least squares [5].

2 Experimental Analysis

2.1 Preparation of the Bimaterial Specimen. Polycarbonate (PSM-1) and aluminum of the same thickness (5.5 mm), whose properties are given in Table 1, are initially cut from the sheets to the dimensions given in Fig. 1. The bonding surfaces are carefully roughened for perfect bonding. The two machined pieces are first kept at a slightly elevated temperature (36°C) for about 2 h, by keeping in the close vicinity of a hot air blower, until they reach steady state. It is to be noted that the material properties of polycarbonate remain invariant in the range −10°C to 55°C [6].

The adhesive is prepared by mixing a general-purpose polyester resin, accelerator, and catalyst. Polycarbonate being a polyester-based plastic, the choice of polyester adhesive ensures a true bimaterial interface. Teflon tapes are placed on the edges, with silicon grease applied on them, for easy removal after curing, to form the two symmetric edge cracks of length 18 mm each. The adhesive is applied on the surfaces to be bonded, light pressure is applied, and the specimen is allowed to set at the same elevated temperature conditions for about 6 h. After curing, the Teflon tape is easily removed, forming the two interface edge cracks, as indicated in Fig. 1.

2.2 Experimental Procedure. Figure 2 is a schematic representation of the experimental setup for circular polariscope dark-field optical arrangement. The bimaterial specimen, bonded at a relatively high temperature (36°C), is suddenly exposed to room temperature (28°C), and it loses heat to the surroundings purely by convection. The aluminum and polycarbonate portions of the specimen tend to contract, each according to their own thermal expansion coefficients, as their temperatures drop. However, at the interface, they are bonded and hence restrained to have the same level of contraction. This gives rise to thermal stresses, predominantly around the interface region. As the specimen temperature decreases further, higher stress levels exist. The presence of the interface crack further intensifies the stresses present at the crack tip. The dark-field isochromatics are recorded at time intervals of 15 s, successively, using a monochrome charge coupled device (CCD) camera (Sony XC-ST50), until steady state is reached (about 15 min). A software module is written using VC++ to carry out automated recording of images at specified time intervals.

Table 1 Material properties

S. No.	Property	Units	Aluminum	PSM-1 [6]
1	Density, ρ	g/cc	2.823	1.17
2	Expansion coefficient, α	/°C	22.7×10^{-6}	146×10^{-6}
3	Thermal conductivity, k	W/m K	142	0.365
4	Specific heat capacity, C_P	J/g °C	0.963	1.1052
5	Young's modulus, E	GPa	71	2.39
6	Poisson's ratio, ν	-	0.33	0.38
7	Material stress fringe value	N/mm fringe	-	7

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 16, 2008; final manuscript received January 16, 2009; published online April 28, 2009. Review conducted by Thomas W. Shield.

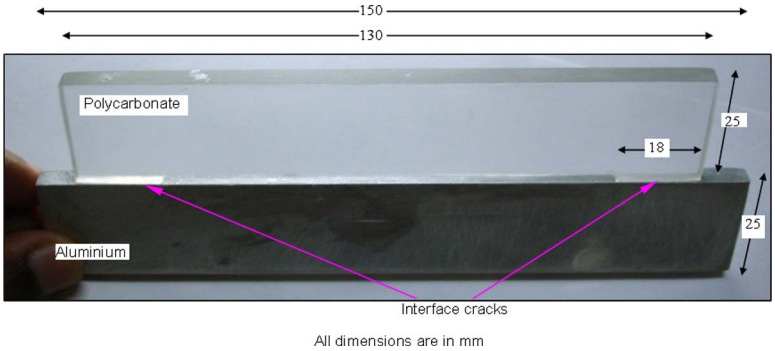


Fig. 1 Aluminum-polycarbonate bimaterial specimen with interface cracks and dimensions indicated

2.3 SIF Evaluation Using the Method of Least-Squares. Ravichandran and Ramesh [5] gave, for use by experimentalists, the Cartesian stress components of the stress field equation, for the top half plane of an interface crack, tangential to a bimaterial joint, loaded at the boundary as

$$\begin{Bmatrix}(\sigma_{xx})_1\\(\sigma_{yy})_1\\(\tau_{xy})_1\end{Bmatrix} = \sum_{n=0,2,4}^{\infty}\frac{K_{In}}{Q}r^{(n-1)/2}\begin{Bmatrix} S\left\{3\cos\left[\left(\frac{n-1}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]-(n-1)\sin\theta\sin\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]+2\varepsilon\sin\theta\cos\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]\right\}-\frac{1}{S}\cos\left[\left(\frac{n-1}{2}\right)\theta+\varepsilon\ln\frac{r}{L}\right]\\ S\left\{\cos\left[\left(\frac{n-1}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]+(n-1)\sin\theta\sin\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]-2\varepsilon\sin\theta\cos\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]\right\}+\frac{1}{S}\cos\left[\left(\frac{n-1}{2}\right)\theta+\varepsilon\ln\frac{r}{L}\right]\\ S\left\{-\sin\left[\left(\frac{n-1}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]-(n-1)\sin\theta\cos\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]-2\varepsilon\sin\theta\sin\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]\right\}+\frac{1}{S}\sin\left[\left(\frac{n-1}{2}\right)\theta+\varepsilon\ln\frac{r}{L}\right]\end{Bmatrix}$$
$$+\sum_{n=1,3,5}^{\infty}\frac{K_{In}}{R}r^{(n-1)/2}\begin{Bmatrix}4\cos\left(\frac{n-1}{2}\right)\theta-(n-1)\sin\theta\sin\left(\frac{n-3}{2}\right)\theta\\ (n-1)\sin\theta\sin\left(\frac{n-3}{2}\right)\theta\\ -2\sin\left(\frac{n-1}{2}\right)\theta-(n-1)\sin\theta\cos\left(\frac{n-3}{2}\right)\theta\end{Bmatrix}$$
$$+\sum_{n=0,2,4}^{\infty}\frac{K_{IIn}}{Q}r^{(n-1)/2}\begin{Bmatrix} S\left\{3\sin\left[\left(\frac{n-1}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]+(n-1)\sin\theta\cos\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]+2\varepsilon\sin\theta\sin\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]\right\}+\frac{1}{S}\sin\left[\left(\frac{n-1}{2}\right)\theta+\varepsilon\ln\frac{r}{L}\right]\\ S\left\{\sin\left[\left(\frac{n-1}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]-(n-1)\sin\theta\cos\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]-2\varepsilon\sin\theta\sin\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]\right\}-\frac{1}{S}\sin\left[\left(\frac{n-1}{2}\right)\theta+\varepsilon\ln\frac{r}{L}\right]\\ S\left\{\cos\left[\left(\frac{n-1}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]-(n-1)\sin\theta\sin\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]+2\varepsilon\sin\theta\cos\left[\left(\frac{n-3}{2}\right)\theta-\varepsilon\ln\frac{r}{L}\right]\right\}+\frac{1}{S}\cos\left[\left(\frac{n-1}{2}\right)\theta+\varepsilon\ln\frac{r}{L}\right]\end{Bmatrix}$$
$$+\sum_{n=1,3,5}^{\infty}\frac{K_{IIn}}{R}r^{(n-1)/2}\begin{Bmatrix}2\sin\left(\frac{n-1}{2}\right)\theta+(n-1)\sin\theta\cos\left(\frac{n-3}{2}\right)\theta\\ 2\sin\left(\frac{n-1}{2}\right)\theta-(n-1)\sin\theta\cos\left(\frac{n-3}{2}\right)\theta\\ -(n-1)\sin\theta\sin\left(\frac{n-3}{2}\right)\theta\end{Bmatrix} \tag{1}$$

where $Q=2\sqrt{2\pi}\cosh\pi\varepsilon$, $R=\sqrt{2\pi}(1+w)$, $S=e^{-\varepsilon(\pi-\theta)}$, $\varepsilon=(1/2\pi)\times\ln((G_2\kappa_1+G_1)/(G_1\kappa_2+G_2))$, $w=(1+\kappa_1)G_2/(1+\kappa_2)G_1$, and $\kappa_i=(3-\upsilon_i)/(1+\upsilon_i)$, $i=1,2$. K_{In} and K_{IIn} are the stress field parameters, ε is the bimaterial constant, G_i are the shear moduli, and ν_i are Poisson's ratios of the two materials, respectively $(i=1,2)$.

The method of least-squares described in Ref. [5] is used to evaluate the multiparameters governing the stress field iteratively. The fringe order and the corresponding positional coordinates at fringe locations are collected from the thinned dark-field image in such a way that, when plotted, they capture the basic geometric features of the fringe field. Since the number of parameters involved in characterizing the stress field are not known a priori, the iteration is started with two parameters each of K_I and K_{II} series. The iteration is stopped using the fringe order error minimization criteria. Using the parameters obtained, the program developed using VC++ [5] recalculates the fringe orders at every point in the data field, for reconstruction and visual comparison with the experiment. Using the solution of the parameters thus obtained as

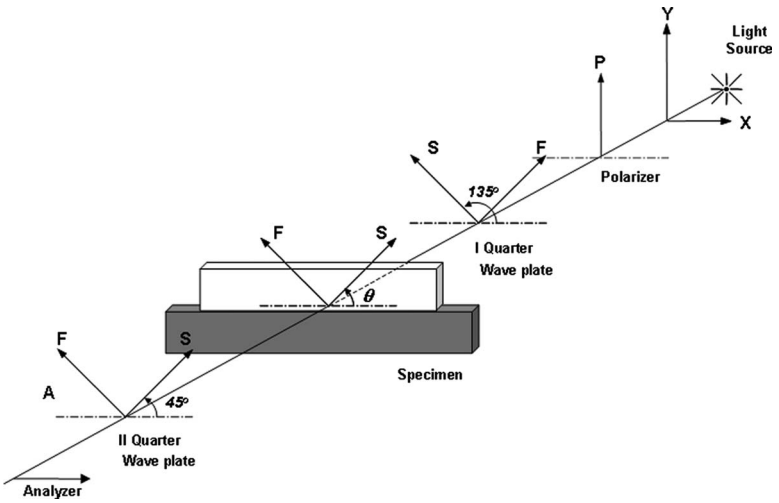


Fig. 2 Circular polariscope dark-field arrangement

starting values, the number of parameters in each series is iteratively increased until the convergence error obtained is of the order of 0.05 [5].

In order to appreciate the convergence visually, the reconstructed dark-field image for the steady state condition of the convection-cooled bimaterial specimen is shown for two-, three-, and four-parameter solutions in Figs. 3(*a*)–3(*c*) with the data points echoed; the corresponding experimental dark-field image is shown in Fig. 3(*d*). From the four-parameter solution, the values of K_I and K_{II} are 0.084 MPa$\sqrt{}$m and 0.176 MPa$\sqrt{}$m, respectively, and the convergence error obtained is 0.046.

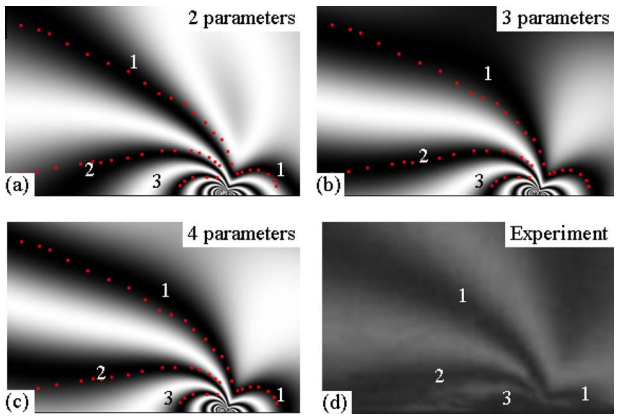


Fig. 3 SIF evaluation for steady state condition of whole-field convection-cooled bimaterial specimen

3 Finite Element Modeling

For the bimaterial specimen under study, a half symmetry 2D model is created using the standard FE package, ABAQUS v6.6, and the final mesh obtained after mesh-convergence study has 2742 plane stress coupled temperature-displacement, biquadratic elements (CPS8RT), as shown in Fig. 4(*a*). The interface crack is created by specifying duplicate nodes along the crack face, and quarter-point collapsed-quadrilateral elements surround the crack tip (Fig. 4(*b*)). Symmetry about the vertical axis (edge *AF* in Fig. 4(*a*)) is specified and point *F* in Fig. 4(*a*) is roller supported so that it is not overconstrained.

Simulating the experimental conditions with concern for even small temperature variations is of key importance in getting a good comparison with the experiment [7]. Convection heat loss occurring from all faces has been accounted for in the FE analysis. The empirical formulas for convection coefficient of natural convection in still air at atmospheric pressure [8] are used. Convection coefficient for all vertical faces (edges *BC* and *DE* in Fig. 4(*a*) and the front and back faces) is $1.42(\Delta T/L)^{0.25}$ as for a vertical plate and that for the horizontal faces (edges *AB* and *CD* in Fig. 4(*a*)) is $1.32(\Delta T/L)^{0.25}$, where, L is the characteristic length dimension computed by dividing the area of the face by its perimeter, and ΔT represents the difference in temperature from the ambient. Convection coefficient values, thus specified, are tem-

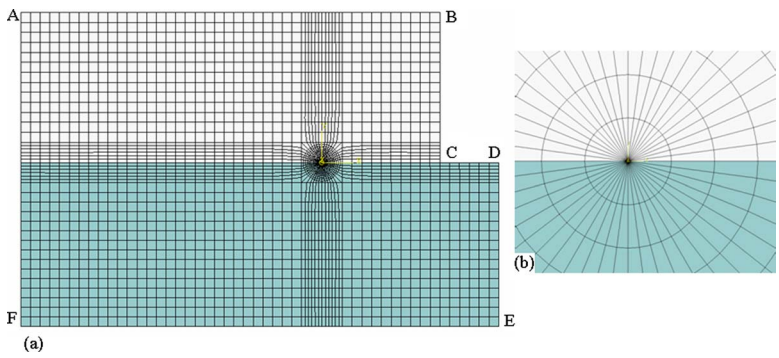


Fig. 4 (*a*) FE mesh and (*b*) zoomed FE mesh near the crack tip

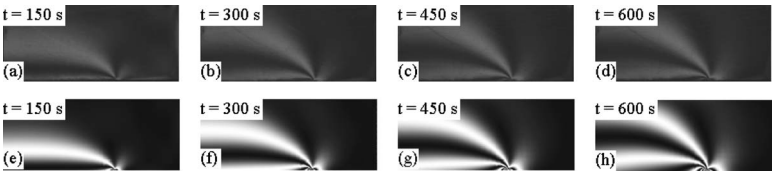


Fig. 5 Transient isochromatic fringe patterns with monochromatic light source for region surrounding the crack tip cooled by convection to 28 °C from a bonding temperature of 36 °C. (*a*)–(*d*) Experimental isochromatics. (*e*)–(*h*) Isochromatics plotted from FE results.

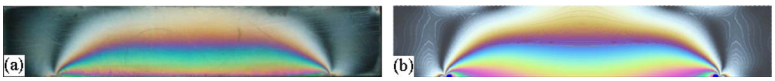


Fig. 6 Steady-state whole-field isochromatic fringe patterns with white light source for bimaterial specimen cooled to 28 °C from the bonding temperature of 36 °C. (*a*) Experimental isochromatics. (*b*) Isochromatics plotted from FE results.

perature dependent. The body heat flux, equivalent to the heat lost due to convection from front and back faces, is $-2h\Delta T/B$, where B is the thickness of the specimen. This is incorporated for the entire bimaterial specimen body by defining a user subroutine [9]. Convection heat loss from the crack faces to the ambient is negligible and they are modeled as completely insulating. The bottom face of aluminum (edge *EF* in Fig. 4(*a*)) is placed on a wooden block during the course of the experiment and hence is insulated.

4 Experimental and Numerical Results

The transient dark-field isochromatics for the polycarbonate portion of the bimaterial specimen, cooled by convection from 36 °C to 28 °C, is given in Figs. 5(*a*)–5(*d*), for a few time steps, in the vicinity of the crack tip. A transient fully coupled temperature-displacement analysis is performed and the postprocessing software of Karthick Babu and Ramesh [4] is used to plot the isochromatic fringe contours from the FE results (Figs. 5(*e*)–5(*h*)), which compare very well with the experimentally recorded images. In order to make a whole-field comparison, the steady state dark-field image is taken using a digital camera (Sony DSC-P150) with a white light source (Fig. 6(*a*)) and compared with the color isochromatic fringe patterns reconstructed from the FE results (Fig. 6(*b*)). It is seen that the comparison is very good and also that a fairly high fringe order is produced even for a small temperature variation of 8 °C.

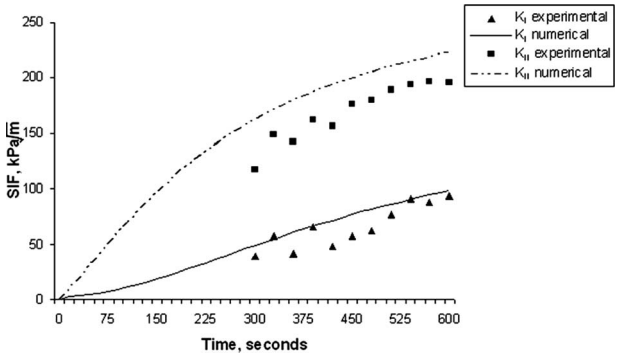


Fig. 7 Comparison of transient experimental and numerical SIFs for convection-cooled bimaterial specimen

SIFs have been evaluated experimentally only from the time step at 300 s up to steady state at 600 s, as sufficient data could be picked up only from that stage. The comparison between the SIFs evaluated experimentally and those evaluated numerically by *J*-integral is given in Fig. 7. It is seen that the trend of the graphs is the same, although the numerical estimate is consistently higher. A probable cause is that the assumption that the crack faces are insulating might have influenced the temperature distribution around the crack and hence the SIFs. It can also been seen from the graph that K_{II} dominates throughout, indicating that the shearing stresses are dominant, and both K_I and K_{II} increase monotonically to steady state.

As an attempt to increase the thermal load on the bimaterial specimen, the ambient temperature was brought down further, by 3 °C, to 25 °C. It was observed that the interface crack propagated quickly, which is seen in the record of isochromatics (Fig. 8).

5 Closure

Convection-cooled interface-cracked bimaterial specimen has been analyzed experimentally using digital photothermoelasticity and using FE method and they compare well. It has been found that for bimaterial specimens such as this, even a small temperature variation in the order of 8 °C produces high interfacial thermal stresses, which are predominantly shear. Although when the specimen is heated along the edge, the transient K_I peaks to a value higher than the steady state K_I [1], for the case of uniform cooling by convection, the steady state SIFs are maximum and there is no such peak observed. In design engineering practice, for

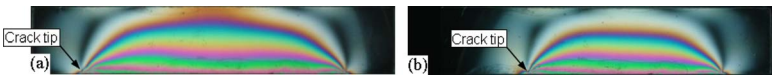


Fig. 8 Experimental whole-field isochromatic fringe patterns with white light source for bimaterial specimen cooled to 25 °C from bonding temperature of 36 °C. (*a*) Before the crack propagated. (*b*) After the crack propagated.

bimaterial specimens such as this, it is important to study the case of uniform cooling by convection. The current study has shown that only the steady state stresses are a maximum, and it is therefore both necessary and sufficient that a steady state analysis be carried out in design engineering practice.

Acknowledgment

The authors are thankful to Professor T. Sundararajan, for the valuable discussion they had on thermal numerical modeling. The authors gratefully acknowledge Mr. Raja Guru Prasath and Mr. Harihara Subramanian for their help in the experiments.

References

[1] Kokini, K., and Smith, C. C., 1989, "Interfacial Transient Thermal Fracture of Adhesively Bonded Dissimilar Materials," Exp. Mech., **29**, pp. 312–317.
[2] Wang, W. C., and Hsu, J. S., 2004, "Re-Examination on Thermal Stresses of Bonded Structures," J. Strain Anal. Eng. Des., **39**, pp. 261–270.
[3] Ileri, L., and Madenci, E., 1995, "Crack Initiation and Growth in Electronic Packages," *Proceedings of 45th Electronic Components and Technology Conference*, 1995, Las Vegas, NV, pp. 970–976.
[4] Karthick Babu, P. R. D., and Ramesh, K., 2006, "Development of Fringe Plotting Scheme From 3D FE Results," Commun. Numer. Methods Eng., **22**, pp. 809–821.
[5] Ravichandran, M., and Ramesh, K., 2005, "Evaluation of Stress Field Parameters for an Interface Crack in a Bimaterial by Digital Photoelasticity," J. Strain Anal. Eng. Des., **40**, pp. 327–343.
[6] Miskioglu, I., Gryzagorides, J., and Burger, C. P., 1981, "Material Properties in Thermal-Stress Analysis," Exp. Mech., **21**, pp. 295–301.
[7] Benolina Neethi, S., and Ramesh, K., 2007, "Numerical and Digital Photoelastic Assessment of Bimaterial Interfaces in Electronic Packages," ICTACEM, IIT Kharaghpur.
[8] Holman, J. P., 2002, *Heat Transfer*, 9th ed., McGraw-Hill, New York, Chap. 7.
[9] 2006, ABAQUS Version 6.6 Documentation, Dassault Systèmes Simulia Corp., Providence, RI.