ASME
SETTING THE STANDARD

Journal of Applied Mechanics

Published Bimonthly by ASME

VOLUME 73 • NUMBER 4 • JULY 2006

TECHNICAL PAPERS

525 An Effective Numerical Approach for Multiple Void-Crack Interaction
Xiangqiao Yan

536 Constitutive Modeling of the Finite Deformation Behavior of Membranes Possessing a Triangulated Network Microstructure
M. Arslan and M. C. Boyce

544 Transient Response of a Finite Bimaterial Plate Containing a Crack Perpendicular to and Terminating at the Interface
Xian-Fang Li and L. Roy Xu

555 Unified Probabilistic Approach for Model Updating and Damage Detection
Ka-Veng Yuen, James L. Beck, and Lambros S. Katafygiotis

565 Multidomain Topology Optimization for Structural and Material Designs
Zheng-Dong Ma, Noboru Kikuchi, Christophe Pierre, and Basavaraju Raju

574 Elastic Bending of Steel-Polymer-Steel (SPS) Laminates to a Constant Curvature
M. Weiss, B. F. Rolfe, M. Dingle, and J. L. Duncan

580 Stability and Andronov-Hopf Bifurcation of Steady-State Motion of Rotor System Partly Filled With Liquid: Continuous and Discrete Models
N. V. Derendyaev, A. V. Vostrukhov, and I. N. Soldatov

590 Reconfiguration of a Rolling Sphere: A Problem in Evolute-Involute Geometry
Tuhin Das and Ranjan Mukherjee

598 Ultrasonic Field Modeling in Multilayered Fluid Structures Using the Distributed Point Source Method Technique
Sourav Banerjee, Tribikram Kundu, and Dominique Placko

610 The Effects of Vibrations on Particle Motion Near a Wall in a Semi-Infinite Fluid Cell
Samer Hassan, Masahiro Kawaji, Tatyana P. Lyubimova, and Dmitry V. Lyubimov

622 Exact Solutions for the Functionally Graded Plates Integrated With a Layer of Piezoelectric Fiber-Reinforced Composite
M. C. Ray and H. M. Sachade

633 On the Crashworthiness of Shear-Rigid Sandwich Structures
Dirk Mohr and Tomasz Wierzbicki

642 Transient Growth and Stick-Slip in Sliding Friction
Norbert Hoffmann

648 Complex Flow Dynamics in Dense Granular Flows—Part I: Experimentation
Piroz Zamankhan and Mohammad Hadi Bordbar

658 Isotropic Clamped-Free Thin Annular Circular Plate Subjected to a Concentrated Load
Ajayi O. Adewale

664 Post-buckling and Vibration of Heavy Beam on Horizontal or Inclined Rigid Foundation
S. T. Santillan, L. N. Virgin, and R. H. Plaut

672 The Arithmetic Mean Theorem of Eshelby Tensor for Exterior Points Outside the Rotational Symmetrical Inclusion
Min-Zhong Wang and Bai-Xiang Xu

(Contents continued on inside back cover)

Editor
ROBERT M. McMEEKING
Assistant to the Editor
LIZ MONTANA

APPLIED MECHANICS DIVISION
Executive Committee
(Chair) **W.-K. LIU**
T. N. FARRIS
K. RAVI-CHANDAR
D. J. INMAN
Z. SUO

Associate Editors
Y. ABOUSLEIMAN (2008)
E. ARRUDA (2007)
J. CAO (2008)
E. CORONA (2008)
H. ESPINOSA (2007)
H. GAO (2006)
S. GOVINDJEE (2006)
Y. Y. HUANG (2008)
S. KRISHNASWAMY (2008)
K. M. LIECHTI (2006)
A. M. MANIATTY (2007)
I. MEZIC (2006)
M. P. MIGNOLET (2006)
S. MUKHERJEE (2006)
O. O'REILLY (2007)
K. RAVI-CHANDAR (2006)
T. W. SHIELD (2008)
N. SRI NAMACHCHIVAYA (2006)
Z. SUO (2006)
T. E. TEZDUYAR (2006)
N. TRIANTAFYLLIDIS (2006)
B. A. YOUNIS (2006)

PUBLICATIONS DIRECTORATE
Chair, **ARTHUR G. ERDMAN**

OFFICERS OF THE ASME
President, **TERRY E. SHOUP**
Executive Director, **V. R. CARTER**
Treasurer, **T. PESTORIUS**

PUBLISHING STAFF
Managing Director, Publishing
PHILIP DI VIETRO
Manager, Journals
COLIN MCATEER
Production Coordinator
JUDITH SIERANT
Production Assistant
MARISOL ANDINO

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).

COPYRIGHT © 2006 by The American Society of Mechanical Engineers. For authorization to photocopy material for internal or personal use under those circumstances not falling within the fair use provisions of the Copyright Act, contact the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923, tel: 978-750-8400, www.copyright.com. Request for special permission or bulk copying should be addressed to Reprints/Permission Department.

Canadian Goods & Services Tax Registration #126148048.

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 73, Number 4** **JULY 2006**

679 Null-Field Integral Equation Approach for Plate Problems With Circular Boundaries
Jeng-Tzong Chen, Chia-Chun Hsiao, and Shyue-Yuh Leu

TECHNICAL BRIEFS

694 A Brief Evaluation of Approximation Methods for Microcrack Shielding Problems
Xiangqiao Yan

697 Ice or Snow in the Tempel 1 Comet?
Nicola M. Pugno

698 Relation Between Stochastic Failure Location and Strength in Brittle Materials
Sefi Givli and Eli Altus

702 Continuum Damage Mechanics and the Life-Fraction Rule
U. Stigh

705 Circumferential Waves of Infinite Hollow Poroelastic Cylinders
M. Tajuddin and S. Ahmed Shah

709 Extending Koiter's Simplified Equations to Shells of Arbitrary Closed Cross Section
James G. Simmonds

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

An Effective Numerical Approach for Multiple Void-Crack Interaction

Xiangqiao Yan
Research Laboratory on Composite Materials,
Harbin Institute of Technology,
Harbin 150001, China
e-mail: yanxiangqiao@hotmail.com

This paper presents a numerical approach to modeling a general system containing multiple interacting cracks and voids in an infinite elastic plate under remote uniform stresses. By extending Bueckner's principle suited for a crack to a general system containing multiple interacting cracks and voids, the original problem is divided into a homogeneous problem (the one without cracks and voids) subjected to remote loads and a multiple void-crack problem in an unloaded body with applied tractions on the surfaces of cracks and voids. Thus the results in terms of the stress intensity factors (SIFs) can be obtained by considering the latter problem, which is analyzed easily by means of the displacement discontinuity method with crack-tip elements (a boundary element method) proposed recently by the author. Test examples are included to illustrate that the numerical approach is very simple and effective for analyzing multiple crack/void problems in an infinite elastic plate. Specifically, the numerical approach is used to study the microdefect-finite main crack linear elastic interaction. In addition, complex crack problems in infinite/finite plate are examined to test further the accuracy and robustness of the boundary element method. [DOI: 10.1115/1.2127955]

1 Introduction

Many brittle materials, such as rocks, ceramics and intermetallics, have a preexisting subscale flaw structure in the form of voids, cracks, and other inclusions. Accurate prediction of the explicit fracture responds of these materials would require an accurate accounting of the growth of the voids and cracks and their interaction.

In recent years, multiple void-crack-related micromechanical analysis has become an area of research [1–17]. Analysis of crack-crack interaction was carried out using the superposition and the pseudotraction concept [4–9]. Analysis of hole-hole interactions using the pseudo-traction concept was also included in the work by Horii and Nemat-Nasser [7]. In the work by Hu et al. [16], the holes were modeled as unknown pseudotractions on the hole surfaces and the cracks were modeled as an unknown distribution of dislocations.

This paper presents a numerical approach to modeling a general system containing multiple interacting cracks and voids in an infinite elastic plate under remote uniform stresses. By extending Bueckner's principle [18] suited for a crack to a general system containing multiple interacting cracks and voids, the original problem is divided into a homogeneous problem (the one without cracks and voids) subjected to remote loads and a multiple void-crack problem in an unloaded body with applied tractions on the surfaces of cracks and voids. Thus the results in terms of the stress intensity factors (SIFs) can be obtained by considering the latter problem, which is analyzed easily by means of the displacement discontinuity method with crack-tip elements (a boundary element method) proposed recently by the author [19]. Test examples are included to illustrate that the numerical approach is very simple and effective for analyzing multiple crack/void problems in an infinite elastic plate. Specifically, the numerical approach is used to study the microdefect-finite main crack linear elastic interaction. Two models are used. One is "A Finite Main Crack Interaction with a Collinear Elliptical Microdefect (Hole)." The other is "A Finite Main Crack Interaction with a Pair of Symmetric Collinear Elliptical Microdefects (Holes)." In the numerical computations, the ratio of the elliptical horizontal axis length to the main crack length varies from 0.05 to 0.4 and the elliptical aspect ratio is taken as 0.0, 0.5 and 1.0. Thus the present numerical results reveal the effect that the relative size of the elliptical microdefect (hole) to the main crack has on the microdefect-finite main crack interaction. In addition, complex crack problems in infinite/finite plate are examined to indicate further the accuracy and robustness of the boundary element method.

By the way, it is pointed out here that this paper concentrates on presenting an efficient numerical approach to modeling a general system containing multiple interacting cracks and voids in an infinite elastic plate under remote uniform stresses. The numerical approach involves a generation of Bueckner's principle and a displacement discontinuity method with crack-tip elements proposed recently by the author [19]. In addition, complex crack problems in infinite/finite plate are examined to indicate further the accuracy and robustness of the boundary element method. In Ref. [19], the displacement discontinuity method with crack-tip elements is described in detail and branch crack problems are analyzed to indicate that the boundary element method is an accurate and efficient approach for calculating the SIFs of the branch crack problems.

2 A Description of the Present Numerical Approach

The numerical approach for analyzing a void-crack interaction problem presented in this paper involves a generation of Bueckner's principle and a displacement discontinuity method with crack-tip elements (an indirect boundary element method) proposed recently by the author.

2.1 A Generalization of Bueckner's Principle. Now consider a two-dimensional system containing a finite number of nonintersecting cracks and elliptical voids. Specifically, consider an infinite elastic plate containing M arbitrarily oriented cracks and N elliptical voids under remote uniform stresses $\sigma_{yy}^{\infty},\sigma_{xx}^{\infty},\sigma_{xy}^{\infty}$. Let

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 20, 2005; final manuscript received September 20, 2005. Review conducted by Z. Suo. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

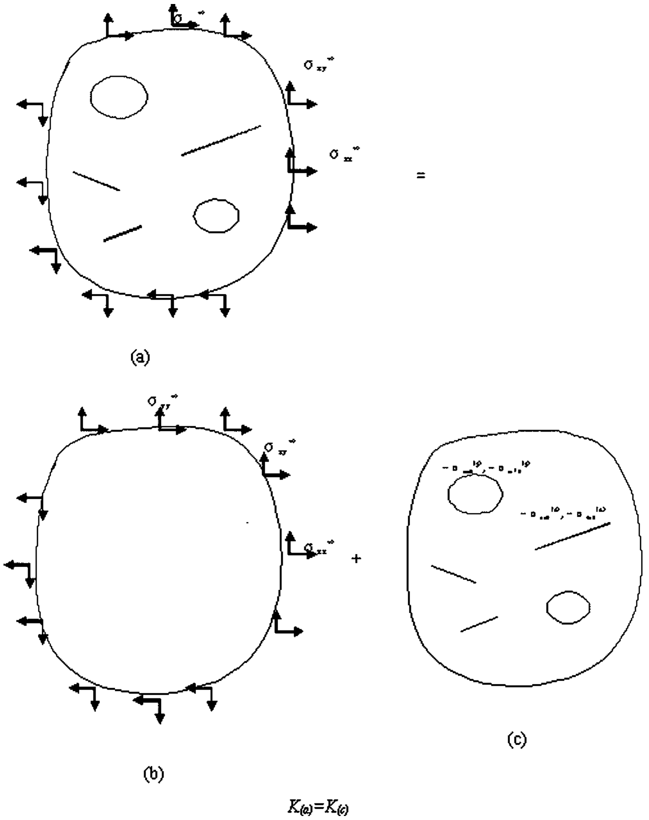


Fig. 1 A generalization of Bueckner's principle

(x,y) be the global Cartesian coordinates. Let $a^{(i)}$ and $\phi^{(i)}$ be the half-length of the ith crack and its orientation angle with respect to the x axis, and $(x_c^{(i)}, y_c^{(i)})$ be the coordinates of the ith crack center. Let $R_x(j)$ and $R_y(j)$ be the horizontal and longitudinal radii of the jth elliptical void, and $(x_v^{(j)}, y_v^{(j)})$ be the coordinates of its center. Denote the local polar coordinate system associated with the jth void by $r(j)$ and $\theta(j)$, and the local coordinates normal and tangential to the ith crack by $s^{(i)}$ and $t^{(i)}$. The boundary conditions for the surfaces of the cracks and holes are traction free, giving

$$\sigma_{ss}^{(i)} = 0, \quad \sigma_{ts}^{(i)} = 0, \quad i = 1,2,\ldots,M,$$
$$\sigma_{rr}^{(j)} = 0, \quad \sigma_{r\theta}^{(j)} = 0, \quad j = 1,2,\ldots,N. \tag{1}$$

We shall refer to the above-described boundary value problem as the original problem.

Bueckner [18] derived an important result, which is related to the principle of superposition. He demonstrated the equivalence of the SIFs resulting from external loading on a body and those resulting from internal tractions on the crack face. The SIFs for a crack in a loaded body may be determined by considering the crack to be in an unloaded body with applied tractions on the crack surface only. These surface tractions are equal in magnitude but opposite in sign to those evaluated along the line of the crack site in the uncracked configuration.

Here, we try to extend Bueckner's principle [18] suited for a crack to a general system containing multiple interacting cracks and voids. The original problem (see Fig. 1(a)) is divided into a homogeneous problem (see Fig. 1(b)) (the one without cracks and holes) subjected to remote loads and a multiple void-crack problem (see Fig. 1(c)) in an unloaded body with applied tractions on the surfaces of cracks and holes. The applied tractions on the ith crack surface are equal in magnitude but opposite in sign to those evaluated along the line of the ith crack site in the uncracked configuration, which are

$$\sigma_{sso}^{(i)} = \sigma_{xx}^{\infty} \sin\phi^{(i)} \sin\phi^{(i)} - 2\sigma_{xy}^{\infty} \sin\phi^{(i)} \cos\phi^{(i)} + \sigma_{yy}^{\infty} \cos\phi^{(i)} \cos\phi^{(i)},$$
$$\sigma_{tso}^{(i)} = (\sigma_{yy}^{\infty} - \sigma_{xx}^{\infty}) \sin\phi^{(i)} \cos\phi^{(i)} + \sigma_{xy}^{\infty}(\cos\phi^{(i)} \cos\phi^{(i)} - \sin\phi^{(i)} \sin\phi^{(i)}), \tag{2a}$$

which are called initial stresses here. Denoting the orientation angle of a tangent at any point $[r(j),\theta(j)]$ on surface of the jth hole with respect to the x axis by $\beta^{(j)}$, the initial stresses at this point are

$$\sigma_{rro}^{(j)} = \sigma_{xx}^{\infty} \sin\beta^{(j)} \sin\beta^{(j)} - 2\sigma_{xy}^{\infty} \sin\beta^{(j)} \cos\beta^{(j)} + \sigma_{yy}^{\infty} \cos\beta^{(j)} \cos\beta^{(j)},$$
$$\sigma_{r\theta o}^{(j)} = (\sigma_{yy}^{\infty} - \sigma_{xx}^{\infty}) \sin\beta^{(j)} \cos\beta^{(j)} + \sigma_{xy}^{\infty}(\cos\beta^{(j)} \cos\beta^{(j)} - \sin\beta^{(j)} \sin\beta^{(j)}). \tag{2b}$$

Of course, the initial stresses defined in (2b) varies with different point $[r(j),\theta(j)]$ on surface of the jth hole. Thus, the results in terms of the SIFs can be obtained by considering the latter problem (see Fig. 1(c)), which is analyzed easily by means of the displacement discontinuity method with crack-tip elements proposed recently by the author [19].

2.2 Description of the Displacement Discontinuity Method With Crack-Tip Elements.

Here, the displacement discontinuity method with crack-tip elements proposed recently by the author [19] is described. It consists of the (nonsingular) constant displacement discontinuity element presented by Crouch and Starfield [20] and the crack-tip displacement discontinuity elements due to the author.

Constant Displacement Discontinuity Element and Its Numerical Formulations. The displacement discontinuity D_i in $|x| < a$, $y=0$ in an infinite plate is defined as the difference in displacement between the two sides of the segment [20]:

$$D_x = u_x(x,0_-) - u_x(x,0_+),$$
$$D_y = u_y(x,0_-) - u_y(x,0_+). \tag{3}$$

The solution to the subject problem is given by Crouch and Starfield [20]. The displacements and stresses can be written as

$$u_x = D_x[2(1-\nu)F_3(x,y) - yF_5(x,y)] + D_y[-(1-2\nu)F_2(x,y) - yF_4(x,y)],$$
$$u_y = D_x[(1-2\nu)F_2(x,y) - yF_4(x,y)] + D_y[2(1-\nu)F_3(x,y) - yF_5(x,y)], \tag{4}$$

and

$$\sigma_{xx} = 2GD_x[2F_4(x,y) + yF_6(x,y)] + 2GD_y[-F_5(x,y) + yF_7(x,y)],$$
$$\sigma_{yy} = 2GD_x[-yF_6(x,y)] + 2GD_y[-F_5(x,y) - yF_7(x,y)],$$
$$\sigma_{xy} = 2GD_x[-F_5(x,y) + yF_7(x,y)] + 2GD_y[-yF_6(x,y)]. \tag{5}$$

G and v in these equations are shear modulus and Poisson's ratio, respectively. Functions F_2 through F_7 are described in Ref. [20].

Equations (4) and (5) are used by Crouch and Starfield [20] to set up a constant displacement discontinuity boundary element method. The numerical formulations of the boundary element method are outlined below.

The displacements and stresses denoted by Eqs. (4) and (5) can be rewritten as

$$u_x = C_{xx}(x,y)D_x + C_{xy}(x,y)D_y,$$
$$u_y = C_{yx}(x,y)D_x + C_{yy}(x,y)D_y, \tag{6}$$

and

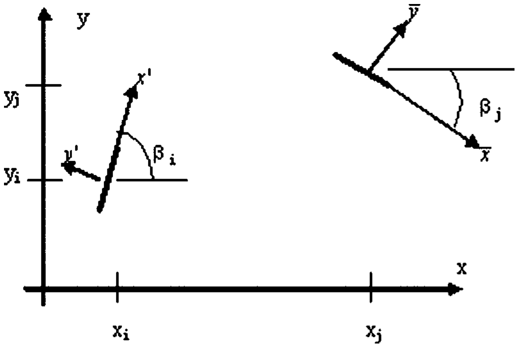


Fig. 2 Schematic of two different segments over which two different sets of displacement discontinuities are imposed

$$\sigma_{xx} = E_{xxx}(x,y)D_x + E_{xxy}(x,y)D_y,$$

$$\sigma_{yy} = E_{yyx}(x,y)D_x + E_{yyy}(x,y)D_y,$$

$$\sigma_{xy} = E_{xyx}(x,y)D_x + E_{xyy}(x,y)D_y. \tag{7}$$

Now let us consider two sets of displacement discontinuities applied to two different segments, as depicted in Fig. 2. According to Eqs. (6) and (7), it is possible to express the displacements and stresses in the center of the ith segment due to the displacement discontinuities applied to the jth segment (denoted by $D^j_{\bar{x}}$ and $D^j_{\bar{y}}$) with regard to the local reference system $\bar{x},\bar{y}$ of the same element; let these displacements and stresses be $u^i_{\bar{x}}, u^i_{\bar{y}}$ and $\sigma^i_{\overline{xx}}, \sigma^i_{\overline{yy}}, \sigma^i_{\overline{xy}}$. The rotation transformation of the $\bar{x},\bar{y}$ system by an angle $\gamma=\beta_i-\beta_j$ turns the same system into the x', y' one, so that by means of the rotation transformation formulas for stresses and displacements [see (2.8.1) and (2.8.5) in Ref. [20]], it is possible to express the displacements and stresses in the center of the ith segment with reference to the local x', y' system. Here, by setting and $D^j_{\bar{y}}=D^j_n$, at the same time, noting that $u^i_{x'}$ and $u^i_{y'}$ are equivalent to u^i_s and u^i_n, and that $\sigma^i_{y'y'}$ and $\sigma^i_{x'y'}$ are equivalent to σ^i_n and σ^i_s, we can obtain

$$u^i_s = p_{sn}(\bar{x}_i,\bar{y}_i)D^j_n + p_{ss}(\bar{x}_i,\bar{y}_i)D^j_s,$$

$$u^i_n = p_{nn}(\bar{x}_i,\bar{y}_i)D^j_n + p_{ns}(\bar{x}_i,\bar{y}_i)D^j_s, \tag{8}$$

$$\sigma^i_n = s_{y'y'n}(\bar{x}_i,\bar{y}_i)D^j_n + s_{y'y's}(\bar{x}_i,\bar{y}_i)D^j_s,$$

$$\sigma^i_s = s_{x'y'n}(\bar{x}_i,\bar{y}_i)D^j_n + s_{x'y's}(\bar{x}_i,\bar{y}_i)D^j_s, \tag{9}$$

where $\bar{x}_i,\bar{y}_i$ are the coordinates of the center of the ith segment with regard to the local reference system $\bar{x},\bar{y}$.

Evidently, $p_{sn}(\bar{x}_i,\bar{y}_i)$, etc. and $s_{y'y'n}(\bar{x}_i,\bar{y}_i)$, etc. in Eqs. (8) and (9) are the boundary influence coefficients for the displacements and for the stresses, respectively. Then Eqs. (8) and (9) can be rewritten in the form

$$u^i_n = L^{ij}_{nn}D^j_n + L^{ij}_{ns}D^j_s,$$

$$u^i_s = L^{ij}_{sn}D^j_n + L^{ij}_{ss}D^j_s, \tag{10}$$

$$\sigma^i_n = O^{ij}_{nn}D^j_n + O^{ij}_{ns}D^j_s,$$

$$\sigma^i_s = O^{ij}_{sn}D^j_n + O^{ij}_{ss}D^j_s. \tag{11}$$

L^{ij}_{nn}, etc. and O^{ij}_{nn}, etc. being the influence coefficients.

In the case where there are m segments over which displacement discontinuities are imposed, either linked or separated from each other, the global effect on the ith segment can be obtained as superimposition of the effects of any single segments, including the ith one,

$$u^i_n = \sum_{j=1}^{m} L^{ij}_{nn}D^j_n + \sum_{j=1}^{m} L^{ij}_{ns}D^j_s,$$

$$u^i_s = \sum_{j=1}^{m} L^{ij}_{sn}D^j_n + \sum_{j=1}^{m} L^{ij}_{ss}D^j_s, \tag{12}$$

$$\sigma^i_n = \sum_{j=1}^{m} O^{ij}_{nn}D^j_n + \sum_{j=1}^{m} O^{ij}_{ns}D^j_s,$$

$$\sigma^i_s = \sum_{j=1}^{m} O^{ij}_{sn}D^j_n + \sum_{j=1}^{m} O^{ij}_{ss}D^j_s. \tag{13}$$

Thus any structural problem can be described by congruence and equilibrium equations like (12) and (13) which constitute a system of linear algebraic equations since the index i can vary over the rang $[1,m]$. The terms on the left-hand side of (12) and (13) represent the resulting effect, displacement or stress, due to all displacement discontinuities applied, thus representing the boundary conditions given over each segment.

If only kinematic boundary conditions are given, then the elastic problem is represented by equations like (12); while static conditions lead to a solving system like (13). If the elastic problem provides mixed boundary conditions, the solving system can be obtained by choosing segment by segment the appropriate equations of type (12) or (13). The solution of a general problem will be provided by a system of linear algebraic equations like

$$\begin{aligned} q^i_n &= \sum_{j=1}^{m} K^{ij}_{nn}D^j_n + \sum_{j=1}^{m} K^{ij}_{ns}D^j_s, \\ q^i_s &= \sum_{j=1}^{m} K^{ij}_{sn}D^j_n + \sum_{j=1}^{m} K^{ij}_{ss}D^j_s, \end{aligned} \qquad (i=1,2,\ldots,m) \tag{14}$$

where K^{ij}_{nn}, etc. and q'_n etc. are, respectively, the influence coefficients L^{ij}_{nn}, etc. and u'_n, etc. in Eq. (12) for kinematic boundary conditions and the influence coefficients O^{ij}_{nn}, etc. and σ'_n, etc. in Eq. (13) for static conditions.

Once the displacement discontinuities $[D^j_n, D^j_s\ (j=1,2,\ldots,m)]$ are found by solving the system (14), stresses and displacements at any point of the continuum medium can be evaluated through (12) and (13) after recalculating the influence coefficients.

Crack-Tip Displacement Discontinuity Elements and Its Numerical Formulations. By using Eqs. (4) and (5), recently, the author [19] presented crack-tip displacement discontinuity elements, which can be classified as the left and the right crack-tip displacement discontinuity elements to deal with crack problems in general plane elasticity. The following gives basic formulas of the left crack-tip displacement discontinuity element.

For the left crack-tip displacement discontinuity element, its displacement discontinuity functions are

$$D_x = H_s\left(\frac{a+\xi}{a}\right)^{1/2}, \quad D_y = H_n\left(\frac{a+\xi}{a}\right)^{1/2}, \tag{15}$$

where H_s and H_n are the tangential and normal displacement discontinuity quantities at the center of the element, respectively. Here, it is noted that the element has the same unknowns as the two-dimensional constant displacement discontinuity element. But it can be seen that the displacement discontinuity functions defined in (15) can model the displacement fields around the crack tip. The stress field determined by the displacement discontinuity functions (15) possesses $r^{-1/2}$ singularity around the crack tip.

Based on Eqs. (4) and (5), the displacements and stresses at a point (x,y) due to the left crack-tip displacement discontinuity element can be obtained,

$$u_x = H_s[2(1-\nu)B_3(x,y) - yB_5(x,y)] + H_n[-(1-2\nu)B_2(x,y) - yB_4(x,y)],$$

$$u_y = H_s[(1-2\nu)B_2(x,y) - yB_4(x,y)] + H_n[2(1-\nu)B_3(x,y) - yB_5(x,y)], \quad (16)$$

and

$$\sigma_{xx} = 2GH_s[2B_4(x,y) + yB_6(x,y)] + 2GH_n[-B_5(x,y) + yB_7(x,y)],$$

$$\sigma_{yy} = 2GH_s[-yB_6(x,y)] + 2GH_n[-B_5(x,y) - yB_7(x,y)],$$

$$\sigma_{xy} = 2GH_s[-B_5(x,y) + yB_7(x,y)] + 2GH_n[-yB_6(x,y)], \quad (17)$$

where functions B_2 through B_7 are described in Ref. [19].

It can be seen by comparing Eqs. (16) and (17) with Eqs. (4) and (5) that the displacements and stresses due to the crack-tip displacement discontinuity possess the same forms as those due to a constant displacement discontinuity, with $F_i(x,y) \times (i=2,3,\ldots,7)$ in Eqs. (4) and (5) being replaced by $B_i(x,y) \times (i=2,3,\ldots,7)$, D_x and D_y by H_s and H_n, respectively.

For the right crack tip, formulas similar to (15)–(17) can be obtained and are not given here.

According to the similarity of Eqs. (16) and (17) to Eqs. (4) and (5), it is easy to perform numerical formulations like (12) and (13) for the crack-tip displacement discontinuity elements,

$$u^i_n = Q^{ij}{}_{nn}H^j{}_n + Q^{ij}{}_{ns}H^j{}_s,$$

$$u^i_s = Q^{ij}{}_{sn}H^j{}_n + Q^{ij}{}_{ss}H^j{}_s, \quad (18)$$

$$\sigma^i_n = R^{ij}{}_{nn}H^j{}_n + R^{ij}{}_{ns}H^j{}_s,$$

$$\sigma^i_s = R^{ij}{}_{sn}H^j{}_n + R^{ij}{}_{ss}H^j{}_s. \quad (19)$$

$Q^{ij}{}_{nn}$, etc. and $R^{ij}{}_{nn}$, etc. being the influence coefficients.

Numerical Formulations of the Boundary Element Method. As mentioned above, the displacement discontinuity method with crack-tip elements proposed recently by the author consists of the (nonsingular) constant displacement discontinuity element presented by Crouch and Starfield and the crack-tip displacement discontinuity elements due to the author. The numerical formulations of the boundary element method are obtained easily as follows:

$$\begin{aligned} q^i{}_n &= \sum_{j=1}^{m} W^{ij}{}_{nn}D^{j'}{}_n + \sum_{j=1}^{m} W^{ij}{}_{ns}D^{j'}{}_s, \\ q^i{}_s &= \sum_{j=1}^{m} W^{ij}{}_{sn}D^{j'}{}_n + \sum_{j=1}^{m} W^{ij}{}_{ss}D^{j'}{}_s. \end{aligned} \quad (i=1,2,\ldots,m) \quad (20)$$

The influence coefficients $W^{ij}{}_{nn}$, etc. in Eqs. (20) are, respectively, $L^{ij}{}_{nn}$, etc., in Eq. (12) and $O^{ij}{}_{nn}$, etc., in Eqs. (13) for kinematic and static boundary conditions of constant displacement discontinuity element, $Q^{ij}{}_{nn}$, etc. in Eqs. (18) and $R^{ij}{}_{nn}$, etc. in Eqs. (19) for kinematic and static boundary conditions of crack tip displacement discontinuity element. The displacement discontinuities $D^{j'}{}_n$ and $D^{j'}{}_s$ in Eqs. (20) are, respectively, $D^j{}_n$ and $D^j{}_s$ for constant displacement discontinuity element, and $H^j{}_n$ and H^j for crack tip displacement discontinuity element.

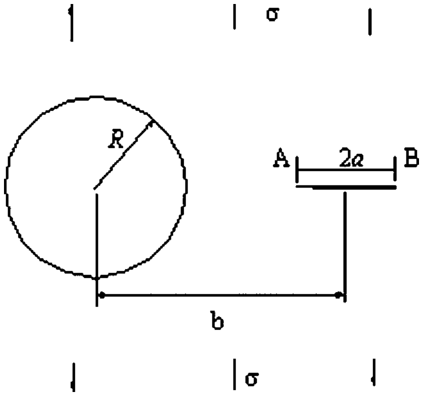


Fig. 3 One circular hole and one crack subjected to remote tension

Computational Formulas of Stress Intensity Factors and Some Examples. The objective of many analyses of linear elastic crack problems is to obtain the SIFs K_{I} and K_{II} at the crack tips. Based on the displacement field around the crack tip, the following formulas exist

$$K_I = -\frac{\sqrt{2\pi}GH_n}{4(1-\nu)\sqrt{a}}, \quad K_{II} = -\frac{\sqrt{2\pi}GH_s}{4(1-\nu)\sqrt{a}}. \quad (21)$$

To indicate further the accuracy and robustness of the numerical approach, some examples are given below.

Crack problems in infinite plate. This subsection is concerned with two multiple crack/void problems: One is the interaction problem of a circular hole with a crack in an infinite plane elasticity medium subjected to remote uniform stress σ (see Fig. 3). The other is the interaction problem of one crack in horizontal position and another in inclined position subjected to remote uniform stress σ (see Fig. 4). The two examples illustrate that the numerical approach is very effective for analyzing the multiple crack/void problems in infinite plate.

For the void-crack interaction problem shown in Fig. 3, the following cases are considered

$$R/a = 2$$

$$b/a = 3.2, 3.5, 4, 5, 8$$

The numbers of boundary elements discretized on the crack and the circular hole are 200 and 800, respectively. The present numerical results of the SIFs normalized by $\sigma\sqrt{\pi a}$ are given in Table 1. Table 1 also shows results obtained by Erdogan et al. [21] by using the singular integral equation method. In addition, we also check the effect of element discretizations on numerical results for the case $b/a=3.5$, $R/a=2$. Let the number of boundary elements discretized on the crack and the circular hole be denoted

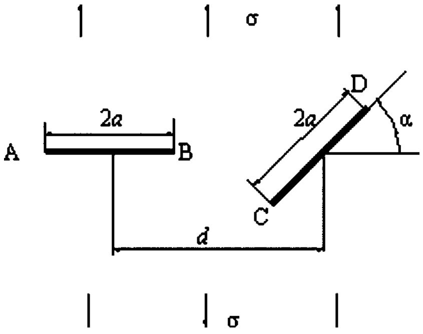


Fig. 4 One crack in horizontal position and another in inclined position subjected to remote uniform stress σ

Table 1 Normalized SIFs for the interaction between one circular hole and one crack

b/a	3.2		3.5		4		5		8	
	F_A	F_B	F_A	F_B	F_A	F_B	F_A	F_B	F_A	F_B
Present	2.2650	1.4149	1.7173	1.2887	1.3918	1.1871	1.1732	1.1007	1.0438	1.0326
Ref. [23]	2.274	1.417	1.722	1.290	1.394	1.188	1.174	1.102	1.045	1.033
Relative Error (%)	0.4	0.1	0.3	0.1	0.2	0.1	0.1	0.1	0.1	0.0

by N_{crack} and N_{circular}, respectively. Table 2 gives the variation of the SIFs normalized by $\sigma\sqrt{\pi a}$ as the number of discretized elements varies. From Tables 1 and 2, it can be seen that the numerical approach is very effective for analyzing the void-crack interaction problem.

For the multiple-crack interaction problem shown in Fig. 4, the following cases are considered:

$$2a/d = 0.1, 0.9$$

$$\alpha = 0°, 30°, 60°, 90°$$

The total number of boundary elements discretized on crack *AB* and crack *CD* is 60. The present numerical results of the SIFs normalized by $\sigma\sqrt{\pi a}$ at crack tips *A*, *B*, *C*, and *D* are given in Tables 3 and 4. In Table 3, the results obtained by Chen [24] by using the singular integral equation method for the multiple crack problem of plane elasticity are also given. The exact results of a single crack inclined to tensile axis in an infinite plane elasticity reported in Ref. [23] are also given in Table 4. These numerical results indicate further that the numerical approach is also very effective for analyzing the multiple-crack interaction problem.

Crack problems in finite plate. Pan [25] pointed out that "the displacement discontinuity method [20] is quite suitable for cracks in infinite domain where there is no no-crack boundary. However, it alone may not be efficient for finite domain problems, since the kernel functions in DDM involve singularities with order higher than those in the traditional displacement BEM." The displacement discontinuity method with crack-tip elements is used in this subsection to calculate the SIFs of complex plane cracks in a finite plate. These numerical results show that the numerical approach is also simple, yet very accurate for complex plane cracks in a finite plate.

Figure 5 shows the schematic of an inclined center crack in a rectangular plate under tension. For this crack problem, the following geometrical parameters are considered

$$\theta = 15°, 30°, 45°, 60°, 75°$$

$$a/W = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$$

Regarding discretization, here, the number of elements discretized on a crack is 20 and the other boundaries are discretized according to the limitation that all boundary elements have approximately equal length. The present numerical results of the SIFs normalized

Table 2 Variation of normalized SIFs for the interaction between one circular hole and one crack with the number of discretized elements

N_{crack} N_{circular}	25 100	50 200	75 300	150 600	
F_{A}	1.6954	1.7082	1.7111	1.7181	1.722 Ref. [23]
Relative Error (%)	1.5	0.8	0.6	0.2	
F_{B}	1.2826	1.2862	1.2864	1.2890	1.290 Ref. [23]
Relative Error (%)	0.6	0.3	0.3	0.1	

Table 3 Normalized SIFs for interaction between one crack in horizontal position and another in inclined position

				$(2a/d=0.9)$				
α	$F_{\text{I}A}$	$F_{\text{I}A}$ [24]	$F_{\text{II}A}$	$F_{\text{II}A}$ [24]	$F_{\text{I}B}$	$F_{\text{I}B}$ [24]	$F_{\text{II}B}$	$F_{\text{II}B}$ [24]
0°	1.1075	1.1174	0.0000	0.	1.4182	1.4539	0.0000	0.
30°	1.0844	1.0939	−0.0464	−0.0472	1.2729	1.2933	−0.0679	−0.0663
60°	1.0227	1.0310	−0.0296	−0.0300	1.0654	1.0757	−0.0393	−0.0394
90°	0.9961	1.0040	0.0000	0.	0.9990	1.0071	0.0000	0
α	$F_{\text{I}C}$	$F_{\text{I}C}$ [24]	$F_{\text{II}C}$	$F_{\text{II}C}$ [24]	$F_{\text{I}D}$	$F_{\text{I}D}$ [24]	$F_{\text{II}D}$	$F_{\text{II}D}$ [24]
0°	1.4182	1.4539	0.0000	0.	1.1075	1.1174	0.0000	0.
30°	1.0179	1.0252	0.5451	0.5580	0.8501	0.8566	0.4444	0.4478
60°	0.3112	0.3101	0.5101	0.5149	0.3067	0.3086	0.4489	0.4525
90°	0.0314	0.0305	0.0136	0.0133	0.0314	0.0305	−0.0136	−0.0133

Table 4 Normalized SIFs for interaction between one crack in horizontal position and another in inclined position

						$(2a/d=0.1)$						
α(°)	$F_{\text{I}A}$	$F_{\text{II}A}$	$F_{\text{I}B}$	$F_{\text{II}B}$	$F_{\text{I}C}$	$F_{\text{I}C}$ [23]	$F_{\text{II}C}$	$F_{\text{II}C}$ [23]	$F_{\text{I}D}$	$F_{\text{I}D}$ [23]	$F_{\text{II}D}$	$F_{\text{II}D}$ [23]
0	0.9933	0.0000	0.9934	0.0000	0.9934	1.0000	0.0000	0.0000	0.9933	1.0000	0.0000	0.0000
30	0.9930	−0.0005	0.9931	−0.0006	0.7463	0.7500	0.4301	0.4330	0.7462	0.7500	0.4302	0.4330
60	0.9924	−0.0005	0.9925	−0.0006	0.2496	0.2500	0.4302	0.4330	0.2495	0.2500	0.4301	0.4330
90	0.9921	0.0000	0.9921	0.0000	0.0012	0.0000	0.0001	0.0000	0.0012	0.0000	−0.0001	0.0000

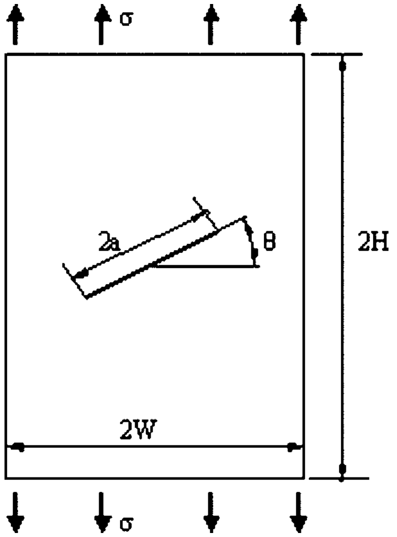


Fig. 5 An inclined center crack in rectangular plate under tension

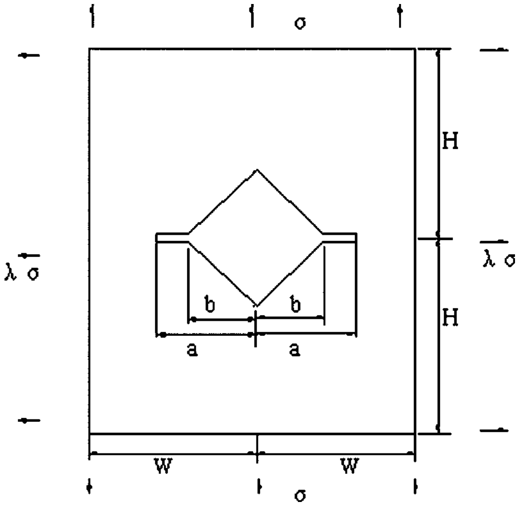


Fig. 6 Schematic of a fair of cracks emanating from a square hole in rectangular plate under biaxial loads

by $\sigma\sqrt{\pi a}$ are given in Table 5. For the comparison purpose, Table 5 also lists those reported in Ref. [23]. From Table 5, it is found that the present numerical results are in very good agreement with those reported in Ref. [23].

Figure 6 shows the schematic of a fair of cracks emanating from a square hole in a rectangular plate under biaxial loads. For this problem, the symmetric conditions about x-axis and y-axis can be used. For the purpose of illustrating the accuracy of the numerical results, the following case is considered first

$$\lambda = 0$$

$$b/W = 0.1, \quad a/b = 1.1$$

which can be regarded approximately as a fair of cracks emanating from a square hole in an *infinite* plate under tension Regarding discretization, here, the number of elements discretized on a quarter of square hole is 100 and the other boundaries are discretized according to the limitation that all boundary elements have approximately equal length. The present numerical result of the SIFs normalized by $\sigma\sqrt{\pi a}$ is given in Table 6. For the comparison purpose, Table 6 also lists the numerical result reported in Ref [23]. From Table 6, it is found that the present numerical result is in excellent agreement with that reported in Ref. [23].

Then the following cases are considered. Load parameter λ is chosen as

$$\lambda = 0, \quad 1, -1$$

and two types of geometric parameters are chosen as

$$H/W = 1$$

$$b/W = 0.25$$

$$a/W = 0.26, 0.27, 0.28, 0.29, 0.30, 0.35, 0.40, 0.50, 0.60, 0.70, 0.80, 0.85, 0.90$$

and

$$H/W = 1$$

$$b/W = 0.50$$

$$a/W = 0.51, 0.52, 0.525, 0.53, 0.54, 0.55, 0.60, 0.70, 0.78, 0.85, 0.90$$

Regarding discretization, the number of elements discretized on a quarter of square hole for the former and latter is, respectively, 100 and 200, and the other boundaries are discretized according to

Table 5 SIFs normalized by $\sigma\sqrt{\pi a}$ for an inclined center crack in a rectangular plate under tension

		θ (deg)									
	a/W	15 present	[23]	30 present	[23]	45 present	[23]	60 present	[23]	75 present	[23]
F_{I}	0.1	0.9360	0.9391	0.7532	0.7557	0.5023	0.5046	0.2519	0.2527	0.0678	0.0678
	0.2	0.9599	0.9577	0.7749	0.7730	0.5191	0.5181	0.2618	0.2605	0.0712	0.0701
	0.3	0.9962	0.9904	0.8074	0.8025	0.5437	0.5406	0.2758	0.2730	0.0758	0.0736
	0.4	1.0488	1.0402	0.8529	0.8456	0.5767	0.5719	0.2937	0.2896	0.0814	0.0783
	0.5	1.1240	1.1128	0.9145	0.9046	0.6188	0.6119	0.3155	0.3099	0.0880	0.0837
	0.6	1.2324	1.2183	0.9970	0.984	0.6708	0.6611	0.3407	0.3332	0.0954	0.0896
	0.7	1.3952	1.378	1.1090	1.091	0.7345	0.721	0.3693	0.359	0.1036	0.0957
	0.8	1.6607	1.653	1.2663	1.245	0.8122	0.795	0.4007	0.388	0.1119	0.102
F_{II}	0.1	0.2496	0.2502	0.4325	0.4339	0.4995	0.5018	0.4336	0.4352	0.2505	0.2516
	0.2	0.2514	0.2510	0.4373	0.4367	0.5072	0.5072	0.4422	0.4417	0.2563	0.2560
	0.3	0.2540	0.2527	0.4438	0.4417	0.5178	0.5162	0.4540	0.4521	0.2641	0.2631
	0.4	0.2578	0.2560	0.4529	0.4497	0.5318	0.5290	0.4688	0.4660	0.2736	0.2721
	0.5	0.2640	0.2619	0.4657	0.4517	0.5496	0.5458	0.4864	0.4827	0.2845	0.2825
	0.6	0.2747	0.2725	0.4843	0.480	0.5719	0.5674	0.5066	0.5022	0.2962	0.2939
	0.7	0.2933	0.290	0.5116	0.508	0.5998	0.595	0.5293	0.524	0.3085	0.3060
	0.8	0.3272	0.307	0.5527	0.550	0.6342	0.630	0.5544	0.549	0.3210	0.319

Table 6 Comparison of normalized SIFs for a fair of cracks emanating from a square hole

	$W/b=H/b=10$ (finite)	Infinite [23]
$a/b=1.1$	1.0864	1.07

the limitation that all boundary elements have approximately equal length. The present numerical results of the SIFs normalized by $\sigma\sqrt{\pi a}$ for the former and latter are listed, respectively, in Tables 7 and 8, from which the effect of load parameter λ on SIFs is found:

(1) For small cracks, the effect of load parameter λ on SIFs is very obvious. For example, for the cases ($b/W=0.25$, $a/W=0.26$) and ($b/W=0.5$, $a/W=0.51$), normalized SIF when $\lambda=-1$ is, respectively, 26.5% and 45.1% more than the one when $\lambda=0$.

(2) With increase of a/W, this effect decreases. For example, for the cases ($b/W=0.25$, $a/W=0.9$) and ($b/W=0.5$, $a/W=0.9$), normalized SIF when $\lambda=-1$ is, respectively, 1.8% and 13.4% more than the one when $\lambda=0$.

Table 7 Normalized SIFs for a fair of cracks emanating from a square hole in a square plate under biaxial loads

	($b/W=0.25$) λ		
a/W	0	1	−1
0.26	1.1974	0.8806	1.5142
0.27	1.2320	0.9228	1.5412
0.28	1.2503	0.9521	1.5485
0.29	1.2634	0.9765	1.5503
0.30	1.2743	0.9983	1.5503
0.35	1.3222	1.0911	1.5533
0.40	1.3743	1.1741	1.5745
0.50	1.5021	1.3401	1.6641
0.60	1.6622	1.5247	1.7997
0.70	1.8657	1.7509	1.9805
0.80	2.1681	2.0807	2.2555
0.85	2.4148	2.3443	2.4853
0.90	2.8337	2.7841	2.8833

Table 8 Normalized SIFs for a fair of cracks emanating from a square hole in a square plate under biaxial loads

	($b/W=0.5$) λ		
a/W	0	1	−1
0.51	1.8217	0.9996	2.6438
0.52	1.9020	1.0600	2.7440
0.525	1.9303	1.0841	2.7765
0.53	1.9551	1.1064	2.8038
0.54	1.9988	1.1481	2.8495
0.55	2.0382	1.1878	2.8886
0.60	2.2137	1.3780	3.0494
0.70	2.5575	1.7803	3.3347
0.78	2.8537	2.1564	3.5510
0.85	3.1676	2.5809	3.7543
0.90	3.5110	3.0414	3.9806

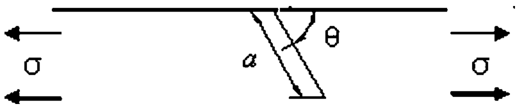


Fig. 7 An oblique edge crack in a half-infinite plane under uniform tension

Crack problems in half-infinite plane. Shown in Fig. 7 is an oblique edge crack in a half-infinite plane under uniform tension. In using the boundary element method in analyzing this kind of half-infinite plane problem, what region is chosen to analyze is first to be determined. For this, according to the geometry character of the problem, a rectangular region, $BB'C'C$, in which the length of CC' equals two times the one of BC, as shown in Fig. 8, is chosen to model the half-infinite plane. Let the ratio of the length of BC to the one of crack be gr, called a boundary geometry factor. Table 9 gives the SIFs normalized by $\sigma\sqrt{\pi a}$ as the boundary geometry factor varies for an edge crack normal to surface in a half-infinite plane under uniform tension. In this analysis, the number of boundary elements discretized on the crack is 15 and elements on other boundaries are discretized according to the limitation condition that all elements have approximately the same length. From the Table 9, it is found that the stress intensity factor is convergent with the boundary geometry factor and relative error is 2.58% when $gr=15$. Table 10 gives the present numerical results of the SIFs normalized by $\sigma\sqrt{\pi a}$ for an oblique edge crack in a half-infinite plane under uniform tension. For the comparison purpose, Table 10 lists also the results reported in Ref. [23]. From Table 10, it is found that the present numerical results are in very good agreement with those reported in Ref. [23].

3 Results and Discussions

In this section, specifically, two models are adopted to study the microdefect-finite main crack interaction. One is "A Finite Main Crack Interaction with a Collinear Elliptical Microdefect (hole)." The other is "A Finite Main Crack Interaction with a Pair of Symmetric Collinear Elliptical Microdefects (holes)."

3.1 A Finite Main Crack Interaction With a Collinear Elliptical Microdefect (Hole). Shown in Fig. 9 is a finite main crack and a collinear elliptical microdefect (hole) under uniform far-field tension σ normal to the crack faces, where the length of

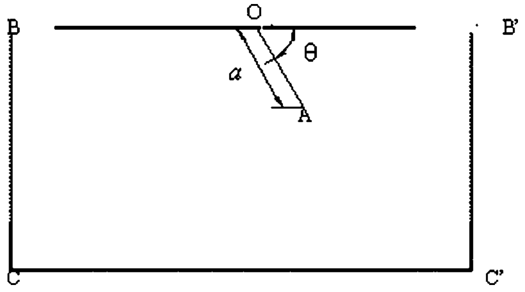


Fig. 8 Schematic of modeling a half-infinite plane

Table 9 Variation of normalized SIFs with a boundary geometry factor when $\theta=90°$

gr	7	10	15	20
F	1.2535	1.1878	1.1499	1.1363
F [23]	1.121			

Table 10 Normalized SIFs for an oblique edge crack in a half-infinite plane under uniform tension

θ (deg)	F_{I}	F_{I} [23]	F_{II}	F_{II} [23]
15	0.2222	0.232	0.2349	0.226
30	0.4615	0.463	0.3426	0.336
45	0.7438	0.705	0.3479	0.364
60	0.9424	0.920	0.3072	0.306
75	1.0948	1.069	0.1756	0.174
90	1.1499	1.121	0.0010	0.0

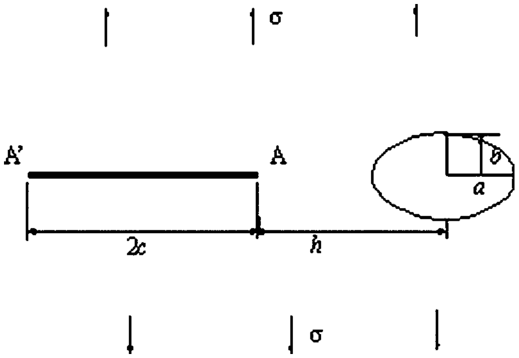


Fig. 9 A collinear elliptical microdefect (hole) in the vicinity of a finite main crack

the main crack $A'A$ is denoted by $2c$, the microdefect is characterized by an elliptical hole whose horizontal and vertical axis radius are denoted by a and b and h is utilized to specify the location of the center of the microdefect. In this analysis, let the length of the finite main crack $2c$ be constant. Thus, ratios a/c and a/h can be used to indicate, respectively, the relative magnitude of the microdefect size and the relative distance away from the crack tip A to the center of the microdefect. The following cases are considered

$$a/c = 0.05, 0.1, 0.2, 0.4$$

$$a/h = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$$

$$b/a = 0.5, 1.0$$

Regarding discretization, the number of elements discretized on the microdefect is kept constant 300, while the number (N) of elements discretized on the finite main crack $A'A$ varies with the ratio a/c and is given in Table 11. The SIFs at the main crack tip A normalized by $\sigma\sqrt{\pi c}$ are listed in Table 12. To compare with a finite main crack interaction with a collinear microcrack, Table 12 also gives the numerical results for a finite main crack interaction with a collinear microcrack, i.e., those corresponding to $b/a=0.0$. The present numerical results corresponding to $a/c=0.05$

Table 11 Variation of number of elements discretized on the finite main crack with the ratio *a/c*

a/c	0.05	0.1	0.2	0.4
N	750	375	188	94

can be used to reveal approximately the interaction of a semi-infinite main crack with a collinear microdefect whose analysis solution has been obtained by Gong and Meguid [22]:

$$K_I^{MA}/K_I = 1 + \frac{1}{4}(1+\varepsilon^2)(a/h)^2 + \frac{1}{128}(-49\varepsilon^4 + 12\varepsilon^3 + 46\varepsilon^2 + 23) \times (a/h)^4 + \cdots, \qquad (22)$$

where $\varepsilon = b/a$, K_I^{MA} and K_I are, respectively, the SIFs when there is the microdefect and there is not the microdefect. The analytical results obtained by using the first three terms in formula (22) are listed in Table 13. By comparing the present numerical results corresponding to $a/c=0.05$ in Table 12 with the analytical results shown in Table 13, it is found that to the extent that the ratio a/h reaches 0.7 the agreement is very good with relative error less than 1.5%, which indicates, on one hand, that the present numerical results are very accurate, and on the other hand, that the formula (22) is very effective when a/h is less than 0.7. But when the microdefect is close enough to the main crack tip A, for example, $a/h=0.9$, the analytical results obtained by using the first three terms in formula (22) have relative errors 11.9%, 19.2% and 15.0%, respectively, for the elliptical aspect ratios $b/a=0.0$, 0.5 and 1.0 compared to the present study results.

The present numerical results shown in Table 12 can be used to develop rules for a collinear microdefect interaction with a finite main crack. The following laws exist:

(1) When the microdefect is far enough away from the finite main crack tip, for example, $a/h<0.1$, the microdefect has no effect on the SIFs at the main crack tip A.

(2) The closer the microdefect is to the finite main crack tip, the more obvious is the effect of the elliptical hole aspect ratio on the SIFs at the main crack tip A. For example, for the case ($a/c=0.05$, $a/h=0.7$),

$$\frac{F(b/a=0.5, a/h=0.7, a/c=0.05) - F(b/a=0.0, a/h=0.7, a/c=0.05)}{F(b/a=0.0, a/h=0.7, a/c=0.05)} = 5.1\%,$$

Table 12 Normalized SIFs for finite main crack interaction with a collinear microdefect

		a/h								
a/c	b/a	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.05	0.0	1.0015	1.0098	1.0238	1.0451	1.0761	1.1214	1.1908	1.3087	1.5679
	0.5	0.9995	1.0089	1.0257	1.0522	1.0927	1.1535	1.252	1.428	1.8266
	1.0	1.0001	1.0126	1.0360	1.0730	1.1285	1.2095	1.3273	1.5121	1.8618
0.1	0.0	1.0018	1.0105	1.0252	1.0472	1.0789	1.1250	1.1954	1.3147	1.5760
	0.5	1.0014	1.0112	1.0282	1.0550	1.0954	1.1573	1.2574	1.4366	1.8403
	1.0	1.0018	1.0140	1.0369	1.0742	1.1322	1.2121	1.3345	1.5270	1.8858
0.2	0.0	1.0015	1.0112	1.0270	1.0502	1.0834	1.1311	1.2034	1.3253	1.5908
	0.5	1.0022	1.0127	1.0305	1.0580	1.0994	1.1627	1.2657	1.4483	1.8561
	1.0	1.0024	1.0146	1.0374	1.0747	1.1318	1.2167	1.3451	1.5447	1.9143
0.4	0.0	1.0005	1.0115	1.0291	1.0544	1.0901	1.1408	1.2168	1.3435	1.6167
	0.5	1.0023	1.0138	1.0332	1.0625	1.1061	1.1723	1.2788	1.4697	1.8988
	1.0	1.0025	1.0165	1.0435	1.0881	1.1556	1.2524	1.3916	1.6811	1.9826

Table 13 Normalized SIFs determined by using the first three terms in formula (22)

	a/h								
b/a	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0	1.0025	1.0104	1.0247	1.0470	1.0796	1.1254	1.1882	1.2720	1.3819
0.5	1.0032	1.0130	1.0309	1.0587	1.0994	1.1566	1.2348	1.3394	1.4764
1.0	1.0050	1.0204	1.0471	1.0866	1.1411	1.2134	1.3069	1.4256	1.5742

$$\frac{F(b/a=1.0,a/h=0.7,a/c=0.05)-F(b/a=0.0,a/h=0.7,a/c=0.05)}{F(b/a=0.0,a/h=0.7,a/c=0.05)}=11.5\%, \tag{23}$$

while for the case (a/c=0.05, a/h=0.9),

$$\frac{F(b/a=0.5,a/h=0.9,a/c=0.05)-F(b/a=0.0,a/h=0.9,a/c=0.05)}{F(b/a=0.0,a/h=0.9,a/c=0.05)}=16.5\%,$$

$$\frac{F(b/a=1.0,a/h=0.9,a/c=0.05)-F(b/a=0.0,a/h=0.9,a/c=0.05)}{F(b/a=0.0,a/h=0.9,a/c=0.05)}=18.7\%, \tag{24}$$

where F is the normalized SIFs listed in Table 12.

(3) The effect of the finite main crack size on SIFs at the main crack tip A can be characterized by means of a dimensionless quantity a/c. From Table 12, it is found that when the microdefect is close enough to the main crack tip A, the dimensionless quantity a/c has a certain influence on the SIFs at the main crack tip A. When a/h=0.9, for example, for b/a=0.0,

$$\frac{F(b/a=0,a/h=0.9,a/c=0.4)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}=3.1\%, \tag{25}$$

for b/a=0.5,

$$\frac{F(b/a=0.5,a/h=0.9,a/c=0.4)-F(b/a=0.5,a/h=0.9,a/c=0.05)}{F(b/a=0.5,a/h=0.9,a/c=0.05)}=4.0\%, \tag{26}$$

and for b/a=1.0,

$$\frac{F(b/a=1,a/h=0.9,a/c=0.4)-F(b/a=1,a/h=0.9,a/c=0.05)}{F(b/a=1,a/h=0.9,a/c=0.05)}=6.5\%. \tag{27}$$

(4) By comparing the present study results with those for a semi-infinite crack interaction with a microcrack [9–11] which can be regarded as ones corresponding to a/c=0.05, b/a=0.0 listed in Table 12, it is found that when the microdefect is close enough to the main crack tip, the elliptical hole aspect ratio b/a and the dimensionless quantity a/c have obvious influence on SIFs at the main crack tip A. When a/h=0.9, for example, for a/c=0.2,

$$\frac{F(b/a=0.5,a/h=0.9,a/c=0.2)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}=18.4\%,$$

$$\frac{F(b/a=1,a/h=0.9,a/c=0.2)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}=22.1\%, \tag{28}$$

and for a/c=0.4

$$\frac{F(b/a=0.5,a/h=0.9,a/c=0.4)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}=21.1\%,$$

$$\frac{F(b/a=1,a/h=0.9,a/c=0.4)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}=26.5\%. \tag{29}$$

3.2 A Finite Main Crack Interaction With a Pair of Symmetric Collinear Elliptical Microdefects (Holes).

Shown in Fig. 10 are a finite main crack and a pair of symmetric collinear elliptical microdefects (holes) under uniform far-field tension σ normal to the crack faces. The symmetric condition for this problem is available. Regarding the discretization of boundary elements, the number of elements discretized on one microdefect is kept constant of 300, while the number (N) of elements discretized on the half-finite main crack varies with the ratio a/c and is given in Table 14. The numerical results of the SIFs at the main crack tip normalized by $\sigma\sqrt{\pi c}$ are listed in Table 15. To allow comparison of a finite main crack interaction with a pair of symmetric collinear microcracks, Table 15 also gives the numerical results for b/a=0.0.

It can be seen from Tables 12 and 15 that for very small microdefect, for example, a/c=0.05, the SIFs at the main crack tip obtained by the model shown in Fig. 10 is almost equal to that by the model shown in Fig. 9 with relative error less than 1.4%, which illustrates that the present numerical results corresponding to a/c=0.05 given in Table 15 are very accurate because the

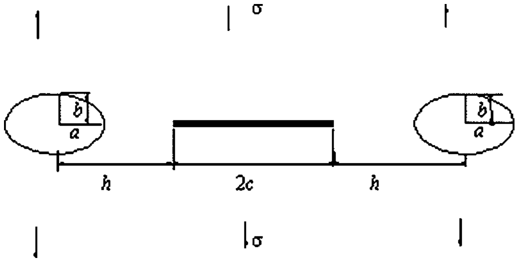


Fig. 10 A pair of symmetric collinear elliptical microdefects (holes) in the vicinity of a finite main crack

present numerical results corresponding to $a/c=0.05$ listed in Table 12, as described above, can be used to reveal very accurately the interaction of a semi-infinite main crack with a collinear microdefect.

By comparing the numerical results listed in Table 15 with those in Table 12, it is found that when a pair of collinear microdefects are close enough to the respective main crack tips neighboring there is some difference between the models shown in Figs. 9 and 10. For example, when $a/c=0.2$, $a/h=0.9$,

$$\frac{F^*(b/a=0.5,a/h=0.9,a/c=0.2)-F(b/a=0.5,a/h=0.9,a/c=0.2)}{F(b/a=0.5,a/h=0.9,a/c=0.2)}$$
$$=4.1\%,$$

$$\frac{F^*(b/a=1,a/h=0.9,a/c=0.2)-F(b/a=1,a/h=0.9,a/c=0.2)}{F(b/a=1,a/h=0.9,a/c=0.2)}$$
$$=5.0\%, \tag{30}$$

where F^* is the normalized SIFs listed in Table 15; when $a/c=0.4$, $a/h=0.9$,

$$\frac{F^*(b/a=0.5,a/h=0.9,a/c=0.4)-F(b/a=0.5,a/h=0.9,a/c=0.4)}{F(b/a=0.5,a/h=0.9,a/c=0.4)}$$
$$=7.3\%,$$

$$\frac{F^*(b/a=1,a/h=0.9,a/c=0.4)-F(b/a=1,a/h=0.9,a/c=0.4)}{F(b/a=1,a/h=0.9,a/c=0.4)}$$
$$=8.9\%. \tag{31}$$

Table 14 Variation of the number of elements discretized on the half-finite crack with the ratio *a/c*

a/c	0.05	0.1	0.2	0.4
N	375	188	94	49

By comparing the present numerical results with those for a semi-infinite crack interaction with a microcrack [9–11] which can be regarded as ones corresponding to $a/c=0.05$, $b/a=0.0$ listed in Table 12, it is found that when a pair of collinear microdefects are close enough to the respective main crack tips neighboring the elliptical hole aspect ratio b/a and dimensionless quantity a/c have obvious influence on the SIFs at the main crack tip. For example, when $a/h=0.9$, $a/c=0.2$,

$$\frac{F^*(b/a=0.5,a/h=0.9,a/c=0.2)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}$$
$$=23.2\%,$$

$$\frac{F^*(b/a=1,a/h=0.9,a/c=0.2)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}$$
$$=28.2\%, \tag{32}$$

and when $a/h=0.9$, $a/c=0.4$,

$$\frac{F^*(b/a=0.5,a/h=0.9,a/c=0.4)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}$$
$$=30.0\%,$$

$$\frac{F^*(b/a=1,a/h=0.9,a/c=0.4)-F(b/a=0,a/h=0.9,a/c=0.05)}{F(b/a=0,a/h=0.9,a/c=0.05)}$$
$$=37.2\%. \tag{33}$$

4 Concluding Remarks

This paper presents a numerical approach to modeling a general system containing multiple interacting cracks and voids in an infinite elastic plate under remote uniform stresses. Test examples illustrate that the numerical approach is very simple and effective for analyzing multiple crack/void interaction problems in an infinite elastic plate. Especially, it is found that the numerical approach is very accurate for studying the microdefect-finite main crack linear elastic interaction. In addition, it is illustrated that the displacement discontinuity method with crack-tip elements presented by the author is very accurate for analyzing complex crack problems in infinite/finite elastic plate.

By the way, it is pointed out that the displacement discontinuity method with crack-tip elements differs from hybrid boundary element codes [26] which, when used to analyze the SIFs of a branched crack, require the plate to be modeled as a finite plate of huge dimensions by fictitious stress elements [20], while the crack could be modeled by displacement discontinuity elements. This

Table 15 Normalized SIFs for finite main crack interaction with a pair symmetric collinear microdefects

		a/h								
a/c	b/a	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	0.0	1.0023	1.0111	1.0259	1.0480	1.0798	1.1262	1.1969	1.3170	1.5806
0.05	0.5	1.0009	1.0111	1.0287	1.0562	1.0974	1.1602	1.2612	1.4430	1.8472
	1.0	1.0016	1.0152	1.0398	1.0801	1.1347	1.2190	1.3394	1.5313	1.8885
	0.0	1.0029	1.0130	1.0291	1.0526	1.0860	1.1343	1.2075	1.3311	1.6011
0.1	0.5	1.0031	1.0143	1.0330	1.0616	1.1043	1.1698	1.2733	1.4595	1.8797
	1.0	1.0036	1.0176	1.0428	1.0826	1.1420	1.2284	1.3568	1.5564	1.9368
	0.0	1.0034	1.0157	1.0343	1.0606	1.0972	1.1491	1.2270	1.3573	1.6399
0.2	0.5	1.0044	1.0176	1.0386	1.0697	1.1155	1.1846	1.2948	1.4915	1.9314
	1.0	1.0047	1.0206	1.0472	1.0893	1.1527	1.2464	1.3854	1.6034	2.0103
	0.0	1.0033	1.0191	1.0422	1.0736	1.1162	1.1753	1.2622	1.4054	1.7120
0.4	0.5	1.0051	1.0218	1.0471	1.0834	1.1355	1.2124	1.3341	1.5534	2.0375
	1.0	1.0054	1.0239	1.0549	1.1029	1.1749	1.2795	1.4359	1.6817	2.1509

brings about a higher computational effort. Using the displacement discontinuity method with crack-tip elements, the branched crack problem is analyzed easily with high accuracy [19]. Also, the boundary element method is extended easily to simulate the mixed mode fatigue crack propagation [19]. Because of an intrinsic feature of the boundary element method, a general crack growth problem can be solved in a single region formulation. In the numerical simulation, for each increment of crack extension, remeshing of existing boundaries is not necessary. Crack extension is modeled conveniently by adding new boundary elements on the incremental crack extension to the previous crack boundaries. As an example, the fatigue propagation process of cracks emanating from a circular hole in a plane elastic plate is simulated using the numerical simulation approach; see Ref. [19].

Finally, it is pointed out that over the few decades many numerical methods have been proposed to model crack problems. Finite element methods with non-singular and singular elements [27,28] enable the accurate computation of the SIFs. However, these methods require the finite element edges to coincide with the crack. This often complicates mesh generation. Some of the other prominent numerical methods are the boundary element method [29], the boundary collocation method [30], the body force method [31] and the integral equation method [32]. Recently, the extended finite element method (X-FEM) allows for the modeling of arbitrary geometric features independently of the finite element mesh. The crack modeling technique is presented in Ref. [32–35]. This method allows the modeling of crack growth without remeshing. In Ref. [36], the X-FEM is applied to multiple crack/void problems.

Acknowledgment

Special thanks are due to the National Natural Science Foundation of China (No. 10272037) for supporting the present work.

References

[1] Evans, A. G., and Faber, K. T., 1981, “Toughening of Ceramics by Circumferential Microcracking,” J. Am. Helicopter Soc., **64**(7), 394–398.

[2] Hutchinson, J. W., 1987, “Crack Tip Shielding by Micro-Cracking in Brittle Solids,” Acta Metall., **35**, 1605–1619.

[3] Charalambides, P., and McMeeking, R. M., 1987, “Finite Element Method Simulation of Crack Propagation in a Brittle Microcracking Solids,” Mech. Mater., **6**, 71–87.

[4] Chudnovsky, A., and Kachanov, M. 1983, “Interaction of a Crack With a Field of Microcracks,” Int. J. Eng. Sci., **21**, 1009–1018.

[5] Chudnovsky, A., Dolgopolsky, A., and Kachanov, M., 1987, “Elastic Interaction of a Crack With a Microcrack Array,” Int. J. Solids Struct., **23**, 1–21.

[6] Kachanov, M., and Montagut, E., 1986, “Interaction of a Crack Certain Microcrack Array,” Eng. Fract. Mech., **25**, 625–636.

[7] Horii, H., and Nemat-Nasser, S., 1985, “Elastic Fields of Interacting Inhomogeneities,” Int. J. Solids Struct., **21**, 731–745.

[8] Hori, M., and Nemat-Nasser, S., 1987, “Interacting Microcracksnear the Tip in the Process Zone of a Macrocrack,” J. Mech. Phys. Solids, **35**(5), 601–629.

[9] Gong, S. X., and Horii, H., 1989, “General Solution to the Problem of Microcracks Near the Tip of a Main Crack,” J. Mech. Phys. Solids, **37**, 27–46.

[10] Rose, L. R. F., 1986, “Microcrack Interaction With a Main Crack,” Int. J. Fract., **31**, 233–242.

[11] Rubinstein, A. 1985, “Macrocrack Interaction With Semi-Infinite Microcrack Array,” Int. J. Fract., **27**, 113–119.

[12] Rubinstein, A., 1986, “Macrocrack-Microdefect Interaction,” ASME J. Appl. Mech., **53**, 503–510.

[13] Shum, D. K. M., and Hutchinson, J. W., 1990, “On Toughening by Microcracks,” Mech. Mater., **9**, 83–91.

[14] Kachanov, M., J. W. Hutchinson and T. Wu, eds., 1993, Adv. Appl. Mech., **30**, 259–455.

[15] Wang, X. M., Gao, S., and Chen, Y. H., 1996, “Further Investigation for the Macro-Microcrack Interaction I—In the Infinite Isotropic Body,” Int. J. Solids Struct., **33**(27), 4051–4063.

[16] Hu, K. X., Chandra, A., and Huang, Y., 1993, “Multiple Void-Crack Interaction,” Int. J. Solids Struct., **30**(11), 1473–1489.

[17] Ducourthial, E., Bouchaud, E., and Chaboche, J. L., 2000, “Influence of Microcracks on a Propagation of Macrocracks,” Comput. Mater. Sci., **19**, 229–234.

[18] Buckner, H. F., 1958, “The Propagation of Cracks and the Energy of Elastic Deformation,” ASME J. Appl. Mech., **80**, 1225–1230.

[19] Yan, X., 2005, “An Efficient and Accurate Numerical Method of SIFs Calculation of a Branched Crack,” ASME J. Appl. Mech., **72**(3), 330–340.

[20] Crouch, S. L., and Starfield, A. M., 1983, *Boundary Element Method in Solid Mechanics, with Application in Rock Mechanics and Geological Mechanics*, London, Geore Allon & Unwin, Bonton, Sydney.

[21] Erdogan, F., Gupta, F. G. D., and Ratwani, M., 1974, “Interaction Between a Circular Inclusion and an Arbitrarily Oriented Crack,”ASME J. Appl. Mech., **31**, 1007–1013.

[22] Gong, S. X., and Meguid, S. A., 1992, “Microdefect Interacting With a Main Crack: A General Treatment,” Int. J. Mech. Sci., **34**, 933–945.

[23] Murakami, Y., 1987, *Stress Intensity Factors Handbook*, Pergamon Press, New York.

[24] Chen, Z., 1999, “Stress Intensity Factors for Curved and Kinked Cracks in Plane Extension,” Theor. Appl. Fract. Mech., **31**, 223–232.

[25] Pan, E., 1997, “A General Boundary Element Analysis of 2-D Linear Elastic Fracture Mechanics,” Int. J. Fract., **88**, 41–59.

[26] Scavia, C., 1992, “A Numerical Technique for the Analysis of Cracks Subjected to Normal Compressive Stresses,” Int. J. Numer. Methods Eng., **33**, 929–942.

[27] Barsoum, R., 1977, “Triangular Quarter-Point Elements as Elastic and Perfectly-Plastic Crack Tip Elements,” Int. J. Numer. Methods Eng., **11**, 85–98.

[28] Kwon, Y. W., and Akin, J. E., 1989, “Development of a Derivative Singular Element for Aplication to Crack Propagation Problem,” Comput. Struct., **31**(3), 467–471.

[29] Crust, T., 1988, *Boundary element Analysis in Computational Fracture Mechanics*, Kluwer, Dordrecht.

[30] Newman, J., 1971, “An Improved Method of Collocation for the Stress Analysis of Cracked Plates With Various Shaped Boundaries,” Technical Report TN D-6376, NASA.

[31] Nisitani, H., 1985, “Body Force Method for Determination of the Stress Intensity Factors,” Journal Aeronautical Society of India (Special Issue on Fracture Mechanics), **37**, 21–41.

[32] Sneddon, I., 1973, *Methods of Analysis and Solutions of Crack Problems*, Nordhoff International, Leyden.

[33] Belytschko, T., and Black, T., 1997, “Elastic Crack Growth in Finite Elements With Minimal Remeshing,” Int. J. Numer. Methods Eng., **40**(5), 601–620.

[34] Xu, Y. L., Moran, B., and Belytschko, T., 1997, “Self-Similar Crack Expansion Method for Three-Dimensional Crack Analysis,”ASME J. Appl. Mech., **64**(4), 729–737.

[35] Moes, N., Dolbow, J., and Belytschko, T., 1999, “A Finite Element Method With Crack Growth Without Remeshing,” Int. J. Numer. Methods Eng., **46**, 131–150.

[36] Daux, C., Moes, N., and Dolbow, J., 2000, “Arbitrary Branched and Intersecting Cracks With the Extended Finite Element Method,” Int. J. Numer. Methods Eng., **48**, 1741–1760.

Constitutive Modeling of the Finite Deformation Behavior of Membranes Possessing a Triangulated Network Microstructure

M. Arslan

M. C. Boyce
Fellow ASME

Department of Mechanical Engineering,
Institute for Soldier Nanotechnologies,
Massachusetts Institute of Technology,
Cambridge, MA

The mechanical behavior of the membrane of the red blood cell is governed by two primary microstructural features: the lipid bilayer and the underlying spectrin network. The lipid bilayer is analogous to a two-dimensional fluid in that it resists changes to its surface area, yet poses little resistance to shear. A skeletal network of spectrin molecules is cross-linked to the lipid bilayer and provides the shear stiffness of the membrane. Here, a general continuum level constitutive model of the large stretch behavior of the red blood cell membrane that directly incorporates the microstructure of the spectrin network is developed. The triangulated structure of the spectrin network is used to identify a representative volume element (RVE) for the model. A strain energy density function is constructed using the RVE together with various representations of the underlying molecular chain force-extension behaviors where the chain extensions are kinematically determined by the macroscopic deformation gradient. Expressions for the nonlinear finite deformation stress-strain behavior of the membrane are obtained by proper differentiation of the strain energy function. The stress-strain behaviors of the membrane when subjected to tensile and simple shear loading in different directions are obtained, demonstrating the capabilities of the proposed microstructurally detailed constitutive modeling approach in capturing the small to large strain nonlinear, anisotropic mechanical behavior. The sources of nonlinearity and evolving anisotropy are delineated by simultaneous monitoring of the evolution in microstructure including chain extensions, forces and orientations as a function of macroscopic stretch. The model captures the effect of pretension on the mechanical response where pretension is found to increase the initial modulus and decrease the limiting extensibility of the networked membrane.

[DOI: 10.1115/1.2130360]

1 Introduction

The mechanical behavior of the membrane of the red blood cell is governed by two primary microstructural features: the lipid bilayer and the underlying spectrin network. The lipid bilayer is analogous to a two-dimensional fluid in that it resists changes to its surface area, yet poses little resistance to shear as noted as early as 1948 by Ponder [1]. A skeletal network of spectrin molecules is cross-linked to the lipid bilayer and provides the shear stiffness of the membrane. Experiments have documented a membrane surface area modulus of approximately $5(10^2)$ dyn/cm and a membrane shear modulus of approximately $6(10^{-3})$ dyn/cm [2]. Evans [3], Skalak et al. [4], and Evans and Hochmuth [5] have noted that this membrane behavior is essentially a two-dimensional analogue to a rubbery solid which shears readily (a typical rubber shear modulus is of order MPa), yet is nearly incompressible (a typical rubber bulk modulus is of order GPa). These researchers have posed a hyperelastic constitutive model which models the membrane stress versus stretch behavior using a two-dimensional neo-Hookean-type representation of the strain energy density function, $\mathcal{U}^*$

$$\mathcal{U}^* = C(\lambda_1^2 + \lambda_2^2 - 2); \tag{1}$$

together with a constraint of constant surface area, $\lambda_1\lambda_2=1$, where λ_1, λ_2 are the principal stretches in the plane. Here, strain energy density is defined as the strain energy per unit initial surface area. More generally, this constitutive model is expressed in terms of the invariants of the planar left Cauchy–Green tensor, $\mathbf{B}_{2\mathrm{D}} = \mathbf{F}_{2\mathrm{D}}\mathbf{F}_{2\mathrm{D}}^T$, which are given by $I_{1_{2\mathrm{D}}} = trace(\mathbf{B}_{2\mathrm{D}})$ and $I_{2_{2\mathrm{D}}} = \det(\mathbf{B}_{2\mathrm{D}})$, where $\mathbf{F}_{2\mathrm{D}}$ is the two-dimensional (membrane) deformation gradient. Thus, the strain energy density is expressed as

$$\mathcal{U}^* = C(I_{1_{2\mathrm{D}}} - 2); \tag{2}$$

The preservation of surface area constraint gives

$$I_{2_{2\mathrm{D}}} = 1. \tag{3}$$

The Cauchy membrane stress tensor (here "membrane stress" is defined as the product of stress and initial membrane thickness) is then obtained by proper differentiation of $\mathcal{U}^*$ giving

$$\mathbf{T} = 2\frac{d\mathcal{U}^*}{dI_{1_{2\mathrm{D}}}}\mathbf{B}_{2\mathrm{D}} + h\mathbf{I}, \tag{4}$$

or,

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received May 17, 2005; final manuscript received September 30, 2005. Review conducted by R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

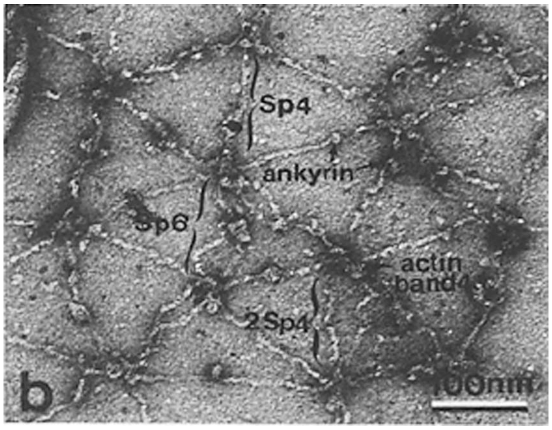


Fig. 1 Electron micrograph of a spread human erythrocyte cytoskeleton [6]

$$\mathbf{T} = 2C\mathbf{B}_{2\mathrm{D}} + h\mathbf{I}, \tag{5}$$

where h is the additional equibiaxial membrane stress required to satisfy the preservation of surface area constraint.

High resolution electron microscopy on spread red blood cell membranes reveals a two-dimensional triangulated network of cross-linked spectrin molecules (Fig. 1). In the spread state depicted in Fig. 1, the fully extended end-to-end distance or contour length of a spectrin molecule extended between two cross-link sites is approximately 150–210 nm [6]. In the unspread configuration, the initial end-to-end distance is 75 nm [7,8].

This triangulated network structure has been taken into account in molecular level modeling of the red blood cell in Discher et al. [9]; Boey et al. [10,11], and Wintz et al. [12] in simulations of deformation-induced structural instabilities ("phase transitions"), deformations, and cell aspiration, and, very recently, by Li et al. [13] in simulations of cell shape. In this paper, we develop a general continuum level hyperelastic constitutive model of the large stretch behavior of the red blood cell membrane that incorporates the detailed microstructure of the triangulated spectrin network and the force-extension behavior of spectrin molecules, building microstructural features into the pioneering modeling approach of Evans [3] and Skalak et al. [4]. The contributions of microstructural geometric nonlinearities as well as constituent molecular chain force-extension nonlinearities to the overall anisotropic nonlinear stress-strain behavior of the network are identified; the effects of network pretension on the membrane stress versus strain behavior are also investigated.

2 Constitutive Model

The constitutive model development for the general membrane stress-stretch behavior of the triangulated spectrin network follows the successful methodology of rubber elasticity (e.g., Treloar [14], Arruda and Boyce [15], Boyce and Arruda [16]) and can be broken down as follows:

- Idealization of the networked microstructure and identification of a representative volume element (RVE),
- Application of macroscopic deformation to the RVE,
- Assignment of mechanical behavior to constituent elements of the RVE,
- Determination of the RVE strain energy,
- Differentiation of the strain energy function to obtain the general multiaxial stress-stretch behavior.

2.1 Microstructure Idealization and Corresponding Representative Volume Element. Examination of spread red blood cell membranes by high resolution negative staining electron microscopy reveals the microstructure of the spectrin network (Fig. 1). A planar network is observed where typically 5–6 spectrin molecules emanate from each cross-link site forming a rather regular triangulated network. This structure has also been referred to as a hexagonal lattice structure [6], where we note six triangles form each hexagon. We idealize this microstructure as a perfectly triangulated network as shown in Fig. 2. Therefore, the representative unit cell is chosen to be an equilateral triangle (Fig. 2(*c*)). Noting that each chain in the unit cell triangle also contributes to its neighboring triangle, we use Voronoi tesselation to identify the area affiliated with the chosen unit cell RVE (Fig. 2(*a*)). Voronoi tesselation gives the area associated with the RVE constituents to be twice the area of the RVE triangle; this area will be needed for properly constructing the strain energy density of the network later.

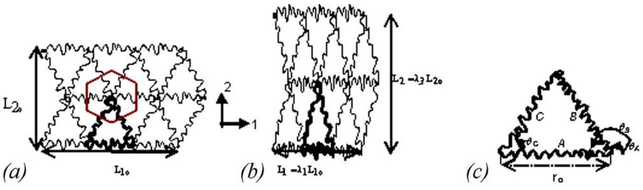


Fig. 2 Schematic of the triangulated network in (*a*) the undeformed state, also depicting Voronoi tessellation (the superposed hexagon) to identify the area of the RVE; (*b*) when stretched in the 2 direction (surface area is preserved); (*c*) the representative volume element.

2.2 Deformation of the Network RVE. An arbitrary deformation is applied to the unit cell equilateral triangle where the membrane deformation gradient $\mathbf{F}_{2\mathrm{D}}$ is defined in the 1–2 frame as: $\mathbf{F}_{2\mathrm{D}}=\partial\mathbf{x}/\partial\mathbf{X}=\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22}\end{bmatrix}$ where $\mathbf{x}$ is the deformed position of a material point and $\mathbf{X}$ is the reference position. The RVE is subjected to an arbitrary deformation gradient (Fig. 3), giving the stretch of constituent network chains A, B, C in terms of the macrosopic deformation gradient. The simplicity of the unit cell triangle RVE provides a unique, kinematically determined mapping of the macroscopic deformation gradient to the microscopic network deformation. Denoting the current end-to-end distance of each chain as r_i $(i=A,B,C)$, the axial stretch of each chain in the network is $\lambda_i=r_i/r_o$ $(i=A,B,C)$ and can be expressed in terms of an arbitrary deformation gradient:

$$\lambda_A = (F_{11}^2 + F_{21}^2)^{1/2}$$

$$\lambda_B = \frac{1}{2}[(F_{11} - F_{12}\sqrt{3})^2 + (F_{21} - F_{22}\sqrt{3})^2]^{1/2}$$

$$\lambda_C = \frac{1}{2}[(F_{11} + F_{12}\sqrt{3})^2 + (F_{21} + F_{22}\sqrt{3})^2]^{1/2} \tag{6}$$

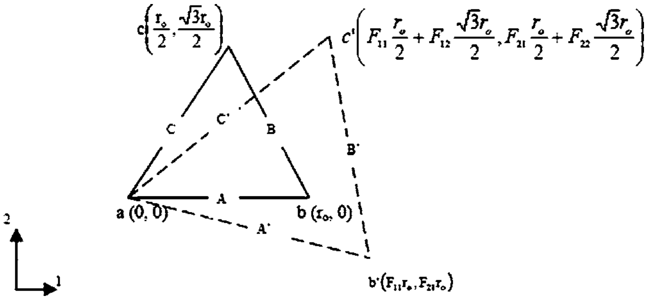

c$\left(\frac{r_o}{2},\frac{\sqrt{3}r_o}{2}\right)$
c'$\left(F_{11}\frac{r_o}{2}+F_{12}\frac{\sqrt{3}r_o}{2},F_{21}\frac{r_o}{2}+F_{22}\frac{\sqrt{3}r_o}{2}\right)$
C
C'
B
B'
A
A'
a (0, 0)
b (r_o, 0)
b'($F_{11}r_o$, $F_{21}r_o$)
2
1

Fig. 3 Schematic of the RVE in undeformed configuration (solid lines) and when subjected to an arbitrary deformation gradient (dashed lines)

2.3 Constitutive Stress-Stretch Behavior of Chains. The stress versus stretch relationship of the membrane will depend on the axial force-extension behavior of the constituent chains. This modeling approach can incorporate any chain force-extension behavior. Here, two representations of chain behavior are studied: linear axial extension behavior and non-Gaussian freely jointed molecular chain behavior. The linear force-extension behavior of a chain is given by:

$$P_{\text{ch}} = k_{\text{ch}} r_o (\lambda_{\text{ch}} - 1), \tag{7}$$

with corresponding strain energy:

$$\mathcal{U}_{\text{ch}_{\text{LC}}} = \frac{1}{2} k_{\text{ch}} r_o^2 (\lambda_{\text{ch}} - 1)^2, \tag{8}$$

where k_{ch} is the chain axial stiffness in units of force/length and r_o is the initial chain length. The force-extension behavior of a molecular chain is taken to follow that of a non-Gaussian freely jointed chain obtained from statistical mechanics treatments of long chain molecules (e.g., Treloar [14]) and is given by:

$$P_{\text{ch}_{\text{NGC}}} = \frac{N k_b \theta}{L_c} \beta; \tag{9}$$

with corresponding strain energy:

$$\mathcal{U}_{\text{ch}_{\text{NGC}}} = k_b \theta N \left[\frac{r}{L_c} \beta + \ln\left(\frac{\beta}{\sinh \beta} \right) \right]; \tag{10}$$

where N is the number of Kuhn segments along the chain, k_b is Boltzmann's constant, θ is the absolute temperature, $L_c = N\ell$ is the contour length of the chain, ℓ is the Kuhn segment length, r is the chain end-to-end distance and β is the inverse Langevin function. The Langevin function is defined as $\mathcal{L}(\beta) = \coth(\beta) - 1/\beta$, with inverse, $\beta = \mathcal{L}^{-1}(r/L_C)$.

2.4 Strain Energy Density of the RVE. The strain energy of the RVE, U is the summation of the strain energy in each chain:

$$\mathcal{U} = \mathcal{U}_A + \mathcal{U}_B + \mathcal{U}_C \tag{11}$$

Strain energy density is typically defined as the strain energy divided by the initial volume of the material. The membrane is a network of essentially single molecule thickness spanning a surface. Therefore the strain energy density, U^*, is taken here to be strain energy per unit initial surface area. As noted earlier, the initial area of the RVE is twice the area of the unit cell triangle following Voronoi tesselation:

$$\mathcal{U}^* = \frac{1}{2A_{\text{triangle}}} (\text{U}_A + \text{U}_B + \text{U}_C). \tag{12}$$

Using the strain energy expressions for each element found earlier together with the kinematics describing the deformation for each element, the following expression for the strain energy density function for the network of linear chains, $\mathcal{U}_{\text{LC}}^*$ is obtained:

$$\mathcal{U}_{\text{LC}}^* = \frac{1}{2A_{\text{triangle}}} \sum_{i=A,B,C} \frac{1}{2} k_{\text{ch}} r_o^2 (\lambda_i - 1)^2; \tag{13}$$

Noting that the chain areal density, υ, is three chains per $(2A)_{\text{triangle}}$, the strain energy density may be written as:

$$\mathcal{U}_{\text{LC}}^* = \frac{\upsilon}{3} k_{\text{ch}} r_o^2 \sum_{i=A,B,C} \frac{1}{2} (\lambda_i - 1)^2; \tag{14}$$

Similarly, for the non-Gaussian network, the strain energy density is given by:

$$\mathcal{U}_{\text{NGC}}^* = \frac{\upsilon}{3} \left\{ k_b \theta N \sum_{i=A,B,C} \left[\frac{\lambda_i r_o}{L_c} \beta_i + \ln\left(\frac{\beta_i}{\sinh \beta_i} \right) - \left(\frac{r_o}{L_c} \beta_o + \ln\left(\frac{\beta_o}{\sinh \beta_o} \right) \right) \right] \right\} \tag{15}$$

where r_o is the initial end-to-end distance (i.e., the initial chain length or distance between cross-links), λ_i $(i = A, B, C)$ are the chain stretches defined earlier as a function of the macroscopic deformation gradient, and $\beta_i = \mathcal{L}^{-1}(\lambda_i r_o / L_c)$. Note that the nonzero value of r_o implies a pretension in the molecular network.[1] This pretension is an important feature of the microstructure and strongly influences the initial modulus of the membrane and the entire nature of the multiaxial stress-stretch behavior as will be demonstrated in the results section. Network pretension is also influenced by swelling (as observed in elastomers and gels) and found to strongly affect mechanical behavior (see, for example, Boyce and Arruda [18]).

2.5 Stress-Stretch Relationships. The stress-stretch behavior is determined by proper differentiation of the strain energy density function. Given a strain energy density which is a function of the deformation gradient, the first Piola Kirchoff membrane stress is derived as:

$$\mathbf{T}_o = \frac{\partial(\mathcal{U}^*)}{\partial(\mathbf{F})}. \tag{16}$$

The Cauchy Stress tensor is then obtained from:

$$\mathbf{T} = \frac{1}{J} \mathbf{T}_o \mathbf{F}^T, \tag{17}$$

where $J = \det(\mathbf{F})$ is the surface area ratio (ratio of current area to original area). Here, we take $J = 1$ due to the preservation of area constraint imposed by the lipid bilayer which then necessitates an additional energy-indeterminate equibiaxial stress term, $h\mathbf{I}$, giving:

$$\mathbf{T} = \mathbf{T}_o \mathbf{F}^T + h\mathbf{I} \tag{18}$$

Noting that the $\mathcal{U}_i$ $(i = A, B, C)$ are functions of λ_i $(i = A, B, C)$, and λ_i $(i = A, B, C)$ are functions of the deformation gradient, gives:

$$\mathbf{T} = \left[\frac{\partial \mathcal{U}_A^*}{\partial \lambda_A} * \frac{\partial \lambda_A}{\partial \mathbf{F}} + \frac{\partial \mathcal{U}_B^*}{\partial \lambda_B} * \frac{\partial \lambda_B}{\partial \mathbf{F}} + \frac{\partial \mathcal{U}_C^*}{\partial \lambda_C} * \frac{\partial \lambda_C}{\partial \mathbf{F}} \right] \mathbf{F}^T + h\mathbf{I} \tag{19}$$

The $(\partial \lambda_i / \partial \mathbf{F})$ terms are independent of chain constitutive behavior and obtained by direct differentiation of the kinematically specified relationships of Eq. (6). For the linear chain model:

$$\frac{\partial \mathcal{U}_i^*}{\partial \lambda_i} = \frac{\upsilon}{3} k_{\text{ch}} (\lambda_i - 1), \quad i = A, B, C. \tag{20}$$

For the non-Gaussian chain model:

$$\frac{\partial \mathcal{U}_i^*}{\partial \lambda_i} = \frac{\upsilon}{3} k_b \theta N \frac{r_o}{L_c} \beta_{\text{i}}, \quad i = A, B, C. \tag{21}$$

The effect of a finite surface area modulus can be incorporated into the strain energy function by addition of the strain energy corresponding to surface area changes in a manner analogous to the treatment of finite compressibility in rubber elasticity (e.g., Boyce and Arruda [16]).

The presented formulation follows a formal continuum mechanics methodology for hyperelasticity: the macroscopic deformation gradient mapped to microscopic constituent elements; the strain energy of the system is formulated; and the strain energy density is then differentiated with respect to an appropriate and general large deformation measure (here, the deformation gradient) to obtain the stress. The approach is general and can account for any constituent chain behavior.[2]

---

[1] Pretension is also present in analogous three-dimensional formulations of classical statistical mechanics of rubber elasticity (Treloar [14], James and Guth [17]) where the initial network chain tension is balanced by internal pressure carried, for example, by intermolecular van der Waal interactions [17]. The spectrin network pretension will be balanced by a few possible sources including, for example, lipid bilayer stress, cytosol interactions, interactions with other protein molecules.

2.6 Determination of Material Properties. The shear modulus of the red blood cell membrane has been experimentally found to be between 0.006 and 0.010 dyn/cm (e.g. Mohandas and Evans [2]). The initial end-to-end distance of chains in the spectrin network is taken to be $r_o=75$ nm following Byers and Branton [7], Boal [8] which, from geometry, gives an areal chain density of $\upsilon=6.16(10^{14})/\mathrm{m}^2$. The fully extended contour length of the network chains has been found to be nominally 180 nm (Liu et al. [6]). In order to meet a target initial shear modulus value of approximately 0.010 dyn/cm, the linear chain stiffness was found to be $k_{\mathrm{ch}}=0.033$ dyn/cm. In order to meet this same shear modulus value for the non-Gaussian chain network while simultaneously meeting the initial end-to-end length and the contour length, the Kuhn segment length was found to be $\ell=10.5$ nm, giving $N=18$. Absolute temperature is taken to be, $\theta=300$ K, corresponding to the reported literature data which were obtained at 300 K.

3 Results

3.1 Uniaxial Tensile Behavior. The axial stress (T_{11}) versus axial stretch (λ) relationships for uniaxial tension in the 1 direction for the linear chain (LC) network and the non-Gaussian chain network (NGC) are obtained directly from Eq. (19) and given by:

$$T_{11_{\mathrm{LC}}}=\frac{\upsilon}{3}k_{\mathrm{ch}}r_o^2\left\{\lambda\left[\lambda-1+\lambda\left(\frac{1}{2}-(\lambda^2+3\lambda^{-2})^{-1/2}\right)\right]-3\lambda^{-2}\left(\frac{1}{2}-\left(\lambda^2+3\lambda^{-2}\right)^{-1/2}\right)\right\} \tag{22}$$

and

$$T_{11_{\mathrm{NGC}}}=\frac{\upsilon}{3}k_B\theta N\frac{r_o}{L_c}\{\lambda[\beta_A+\beta_B\lambda(\lambda^2+3\lambda^{-2})^{-1/2}]-3\beta_B\lambda^{-2}(\lambda^2+3\lambda^{-2})^{-1/2}\}. \tag{23}$$

The axial stress (T_{22}) versus axial stretch (λ) relationships for uniaxial tension in the 2 direction for the case of a network with constituent linear chains and the case of a network with constituent non-Gaussian chains, respectively, are given by:

$$T_{22_{\mathrm{LC}}}=\frac{\upsilon}{3}k_{\mathrm{ch}}r_o^2\left\{3\lambda^2\left[\frac{1}{2}-(3\lambda^2+\lambda^{-2})^{-1/2}\right]-\lambda^{-1}\left[-1+\lambda^{-1}+\lambda^{-1}\left(\frac{1}{2}-(3\lambda^2+\lambda^{-2})\right)^{-1/2}\right]\right\} \tag{24}$$

and

$$T_{22_{\mathrm{NGC}}}=\frac{\upsilon}{3}k_B\theta N\frac{r_o}{L_c}\{3\lambda^2\beta_B(3\lambda^2+\lambda^{-2})^{-1/2}-\lambda^{-1}[\beta_A+\beta_B\lambda^{-1}(3\lambda^2+\lambda^{-2})^{-1/2}]\}. \tag{25}$$

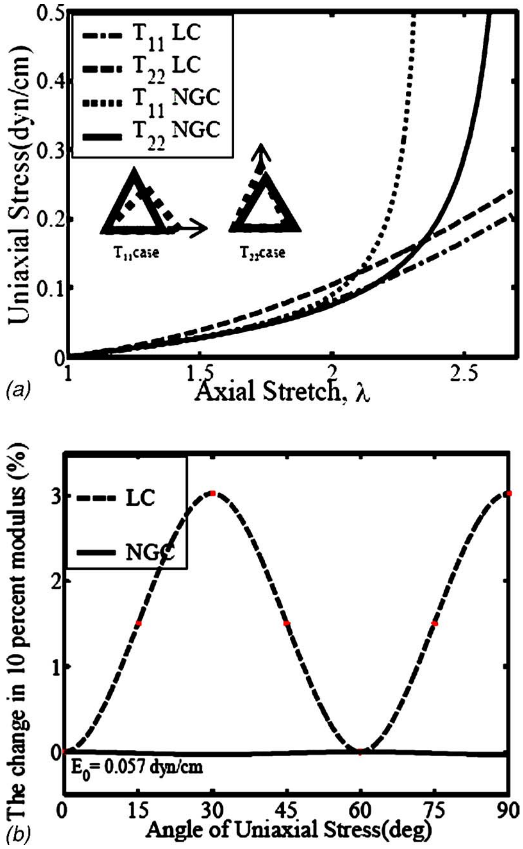


Fig. 4 (*a*) Uniaxial tensile stress vs stretch behavior in the 1 direction and the 2 direction for constituent linear and constituent non-Gaussian chain behaviors. (*b*) Percent change in the 10% strain secant modulus vs the angle of the uniaxial tension applied for various angles with respect to axis 1 for linear chain and non Gaussian chain behavior.

Figure 4(*a*) shows the uniaxial Cauchy membrane stress versus stretch behavior[3] for loading in the 1 and 2 directions, using LC and NGC representations. For small axial stretch ($\lambda<1.10$), both networks exhibit an essentially linear behavior. Nonlinear axial stress-stretch behavior becomes clearly apparent after a stretch of ~1.50. There are three different sources of nonlinearity in the stress-stretch response of the network:

i. Material nonlinearity due to evolution in the microstructure geometry (changes in length and angles of chains),
ii. Material nonlinearity arising from the nonlinear constituent chain behavior (for the non-Gaussian chain case),
iii. Geometric nonlinearity due to macroscopic shape change with deformation.

The sources of nonlinearity in the linear chain network arise from (i) and (iii), whereas all three sources of nonlinearity are inherent in the non-Gaussian chain network behavior.

The uniaxial stress-stretch curves are direction dependent for both the LC and NGC cases as shown by the different stress-stretch behaviors obtained in the 1 and 2 directions. The sixfold symmetry of the initial (undeformed) microstructure results in isotropy of the very initial modulus; expressions for initial uniaxial and shear modulus are provided in the appendix. However, even at an axial strain as small as 10%, anisotropy begins to manifest itself as shown in Fig. 4(*b*) which shows the 10% strain secant modulus as a function of loading direction. The anisotropy is more apparent and important in the large strain behavior, where the stress-stretch curves of Fig. 4(*a*) show significant differences (especially for NGC network) at large stretches.

To understand the predicted stress-stretch behavior, we examine the evolution in the triangulated network microstructure. Figure 5(*a*) shows the evolutions of chain stretch and chain orientation during uniaxial tension in the 1 direction, which are independent

---

[2] An alternative, equally general formulation has been provided in the supplemental materials of the recent parallel work of Li et al. [13], where the Worm-like chain model is used for the constituents and the network behavior is obtained using the virial stress theorem which is commonly used in atomistic and molecular level simulations to obtain stress (see, for example, Bergstrom and Boyce [19]).

[3] These uniaxial tension stress-stretch results were also reported in Arslan and Boyce [20].

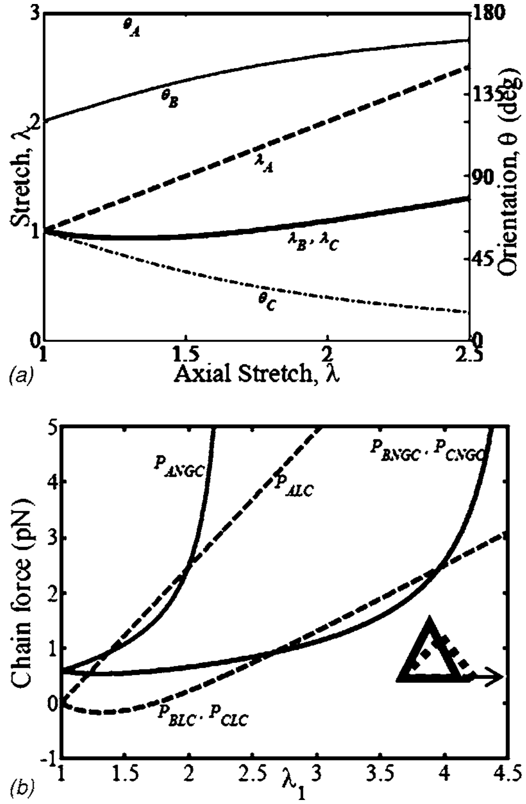


Fig. 5 (*a*) **Evolution of chain orientation and chain stretch with respect to axis 1 for chains A, B and C for uniaxial tension in the 1 direction. (*b*) Evolution of force in chains A, B and C with axial stretch in the 1 direction.**

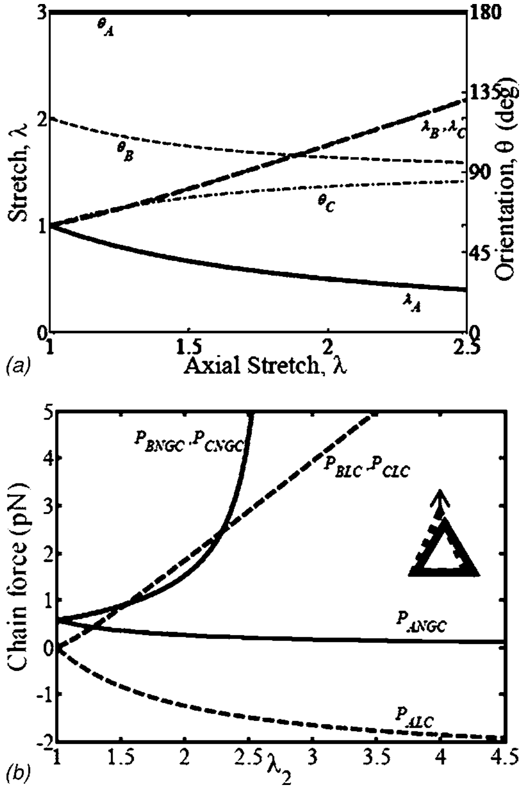


Fig. 6 (*a*) **Evolution of chain orientation and chain stretch with respect to axis 1 for chains A, B and C for uniaxial stress in the 2 direction. (*b*) Evolution of force in chains A, B and C with axial stretch in the 2 direction.**

of force-extension behavior since they are kinematically determined. Chains B and C respond identically to the applied loading because of the symmetry. As the network is deformed, chains stretch and orient to accommodate the imposed deformation. As shown in Fig. 5(*a*), since chain A is aligned with the loading direction, it does not reorient (θ_A is constant) and simply stretches by the imposed macroscopic stretch. Chains B and C are seen to monotonically orient towards the 1 direction; one observation is that, although these chains rearrange themselves as the applied macroscopic stretch increases, they exhibit a non-monotonic change in chain stretch with imposed macroscopic stretch: the chains first undergo a small amount of contraction from their initial end-to-end distance and then undergo extension.

Figure 5(*b*) shows chain force versus macroscopic stretch for uniaxial tension in the 1 direction. Here, for LC behavior, chain A is aligned with the loading direction and therefore shows a linear chain force versus macroscopic stretch behavior. However, chains B and C both stretch and rotate towards the loading direction, therefore their chain force values exhibit a nonlinear history with macroscopic stretch for the LC network. For the NGC, the force-extension curve for chain A shows the typical non-Gaussian behavior since it is directly axially stretched; chains B and C rotate and stretch with deformation and thus show a different nonlinear behavior and a later upturn in the chain force.

Chain force versus macroscopic stretch curves also provide insight regarding the direction dependence of the evolution of macroscopic stress with macroscopic stretch. In Fig. 4(*a*), the uniaxial stress behaviors in the 1 and the 2 directions are identical for NGC behavior until $\lambda > 1.50$. The curves depart at larger stretches and uniaxial membrane stress in the 1 direction curve gives an earlier upturn than uniaxial membrane stress in the 2 direction. Uniaxial tension in the 1 direction is dominated by chain A (Figs. 5(*a*) and 5(*b*)), which is directly axially stretched resulting in the earlier upturn in the macroscopic stress-stretch curve. Uniaxial tension in the 2 direction is dominated by chains B and C (Figs. 6(*a*) and 6(*b*)), which accommodate macroscopic stretch by both rotating and stretching, thus giving the upturn in macroscopic stress at a larger macroscopic stretch.

3.2 Simple Shear Behavior.

The shear stress (T_{12}) versus the shear strain ($\tan\gamma$) relationships for simple shear of the linear chain network and the non-Gaussian chain network are given by:

$$T_{12_{\mathrm{LC}}} = \frac{v}{2\sqrt{3}} k_{\mathrm{ch}} r_o^2 \left\{ \left(\frac{1}{2}(g_1^2+3)^{1/2} - 1\right) g_1 (g_1^2+3)^{-1/2} - \left(\frac{1}{2}(g_2^2+3)^{1/2} - 1\right) g_2 (g_2^2+3)^{-1/2} \right\}, \tag{26}$$

$$T_{12_{\mathrm{NGC}}} = \frac{v}{2\sqrt{3}} k_B \theta N \frac{r_o}{L_c} \{\beta_c g_1 (g_1^2+3)^{-1/2} - \beta_B g_2 (g_2^2+3)^{-1/2}\}, \tag{27}$$

where, $g_1 = (1+\sqrt{3}\tan\gamma)$, $g_2 = (1-\sqrt{3}\tan\gamma)$, and $\beta_i = \mathcal{L}^{-1}(\lambda_i r_o / L_c)$, $i = A, B, C$.

The shear stress (T_{21}) versus the shear strain ($\tan\gamma$) relationships for simple shear for the case of a network with constituent linear chains and the case of a network with constituent non-Gaussian chains, respectively, are given by:

$$T_{21_{\mathrm{LC}}} = \frac{v}{6} k_{\mathrm{ch}} r_o^2 \left\{ 2(f_1-1) f_1^{-1} \tan\gamma + \left(\frac{1}{2}(1+f_2^2)^{1/2} - 1\right) f_2 (1+f_2)^{-1/2} + \left(\frac{1}{2}(1+f_3^2)^{1/2} - 1\right) f_3 (1+f_3)^{-1/2} \right\}, \tag{28}$$

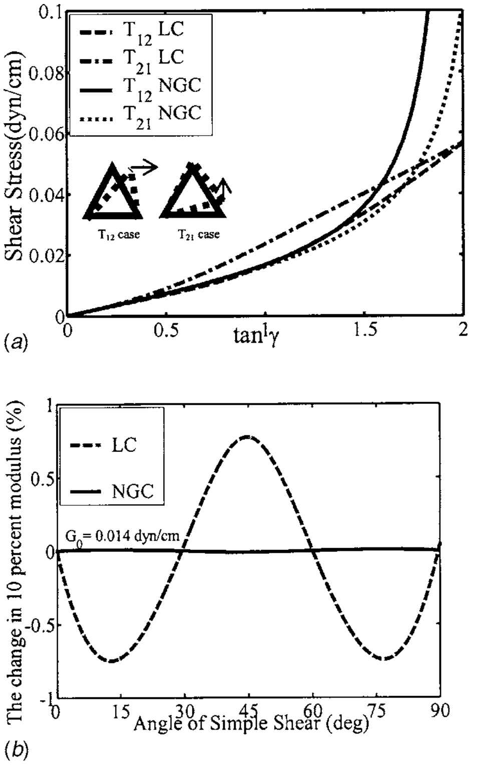


Fig. 7 (***a***) **Simple shear stress vs shear strain behavior in the 12 direction and in the 21 direction for constituent linear and constituent non-Gaussian chain behaviors. (*b*) Percent change in the 10% strain secant modulus vs the angle of the simple shear applied for various angles with respect to axis 1 for linear chain and non-Gaussian chain behavior.**

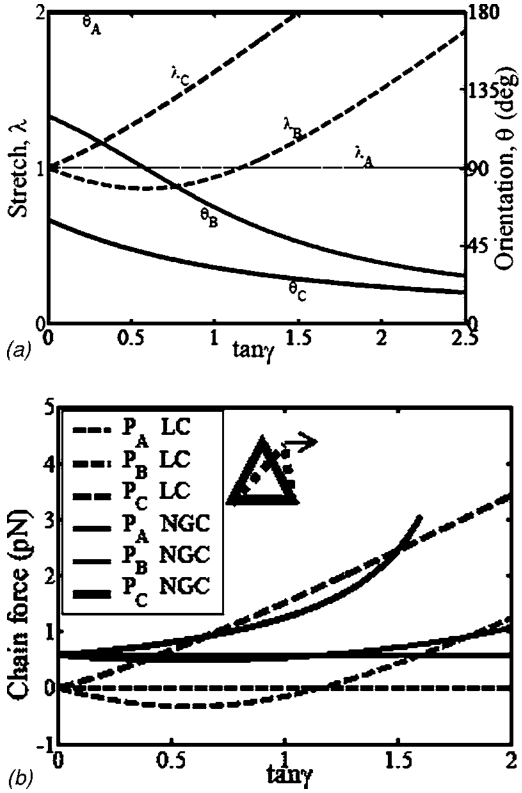


Fig. 8 (***a***) **Evolution of chain orientation and chain stretch with respect to axis 1 for chains A, B and C for simple shear in the 12 direction. (*b*) Evolution of force in chains A, B and C with simple shear in the 12 direction.**

$$T_{21_{\mathrm{NGC}}} = \frac{v}{6}k_B\theta N\frac{r_o}{L_c}\{2\beta_A f_1^{-1}\tan\gamma + \beta_c f_2(1+f_2^2)^{-1/2} - \beta_B f_3(1+f_3^2)^{-1/2}\}, \qquad (29)$$

where $f_1=(1+\tan\gamma^2)^{1/2}$; $f_2=(\tan\gamma-\sqrt{3})$; and $f_3=(\tan\gamma+\sqrt{3})$.

Figure 7(*a*) shows the behavior of the network for simple shear in the 12 and in the 21 directions, using linear chain and non-Gaussian chain models up to a macroscopic strain of $\tan\gamma=2$. The initial slope of these four curves is identical, giving the shear modulus to be 0.014 dyn/cm. The larger deformation behavior is nonlinear and anisotropic.

Figures 8(*a*) and 8(*b*) show the evolutions in microscopic chain stretches/orientations and chain forces, respectively, for simple shear in the 12 direction. Chain A is neither stretched nor rotated in simple shear in the 12 direction. Chains B and C rotate and stretch. Both B and C monotonically rotate towards alignment with the shear flow; chain C monotonically extends, whereas Chain B first contracts then extends. Note that, due to the pretension present in the NGC model, Chain B does not experience a compressive force during its contraction. Chain C extension governs the upturn in the macroscopic stress-strain behavior of the NGC model.

Figures 9(*a*) and 9(*b*) show the chain responses for simple shear in the 21 direction. All chains rotate towards alignment with the shear flow; Chains A and C monotonically extend, whereas Chain B first contracts then extends. The pretension prevents Chain B from experiencing compressive force during its contraction in the case of the NGC model. Chain A extension governs the upturn in the macroscopic stress-strain behavior.

3.3 Pretension Effects. The initial end-to-end distance; $r_o=75$ nm implies a pretension in the network (Qi et al. [21]). The NGC behavior naturally captures this effect of pretension. Referring back to Eq. (9), we note that the chain force is zero when the end-to-end distance $r_o=0$ and nonzero when $r_o\neq 0$; thus r_o sets the pretension of the network. To examine the influence of pretension on the membrane stress-stretch behavior, the value of r_o is varied. Changing r_o also corresponds to changing the areal density v of the membrane. As r_o is increased, v decreases. Figure 10 shows the uniaxial tensile behavior of the membrane for different r_o. Uniaxial tension in the 1 direction and in the 2 direction are both simulated for an r_o of 50, 75, 125, and 150 nm. As seen in Fig. 10(*a*), as pretension is increased over the range of 50 nm $<r_o<150$ nm, the initial modulus of the membrane increases (even though the chain density decreases) and the extensibility of the network decreases. Figure 10(*b*) depicts the dependence of the initial axial modulus on the pretension in a plot of E_o versus (r_o/L_c). For $(r_o/L_c)<0.3$, the modulus is weakly dependent on pretension; however, for $(r_o/L_c)>0.3$ the initial modulus becomes dramatically dependent on pretension. This suggests that the physiological conditions of $r_o\cong 75$ nm and $L_c\cong 180$ nm, giving $r_o/L_c=0.42$, are perhaps microstructurally optimized to provide a dramatic stiffening behavior at particular stretch magnitudes.

4 Summary

The mechanical behavior of the red blood cell membrane is governed by the lipid bilayer in the membrane and the spectrin network cross-linked to the bilayer. The lipid bilayer resists changes in surface area while posing little resistance to shear. The spectrin network provides shear stiffness to the membrane. In this paper a continuum level hyperelastic constitutive model is proposed for the red blood cell membrane stress-stretch behavior for arbitrary membrane deformations. The cases of networks constituted of chains with linear force-extension behavior and for chains

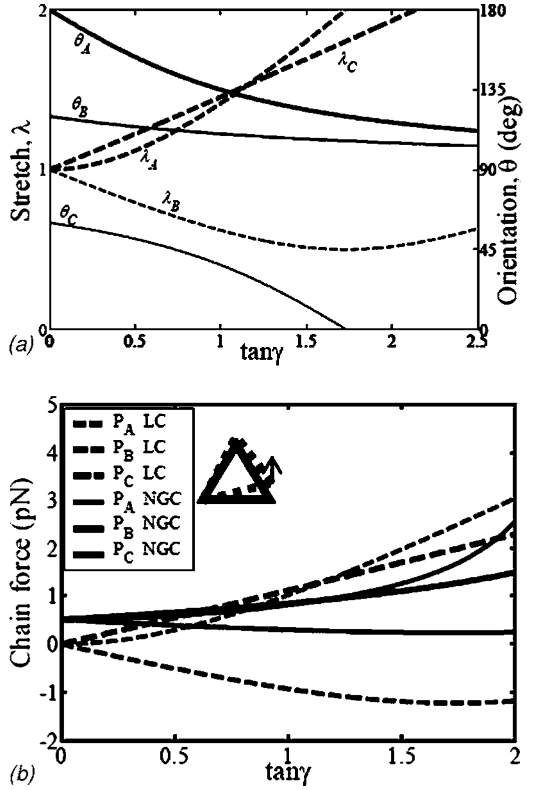


Fig. 9 (*a*) Evolution of chain orientation and chain stretch with respect to axis 1 for chains A, B and C for simple shear in the 21 direction. (*b*) Evolution of force in chains A, B and C with simple shear in the 21 direction.

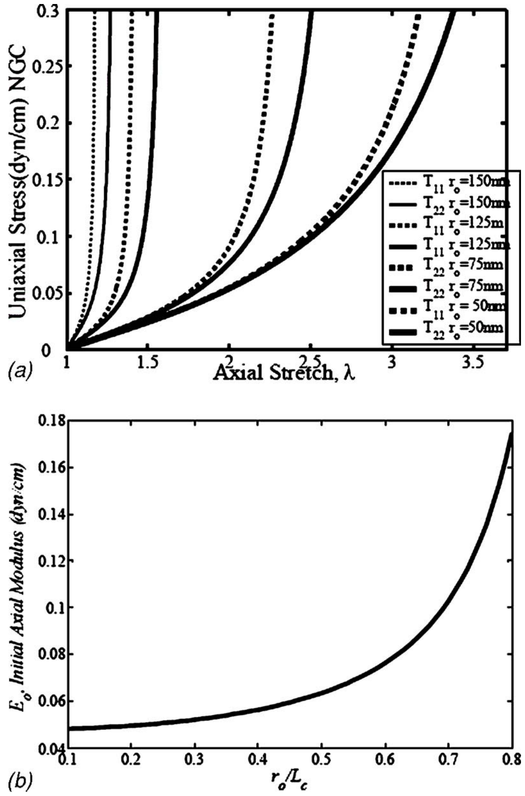


Fig. 10 Effect of pretension on (*a*) the uniaxial stress-stretch behavior and on (*b*) the initial axial modulus

with non-Gaussian force-extension behavior are studied. The skeletal network of spectrin is approximated as a perfectly triangulated network. Uniaxial tension and simple shear behaviors of the membrane are simulated for the two different constituent chain behaviors in different directions. The macroscopic nonlinear stress-stretch behavior is found to be anisotropic at large deformations. The evolution in microstructure (chain stretches, forces, and orientations) is naturally monitored during macroscopic deformation thus capturing the effects of structural evolution on macroscopic mechanical loading behaviors and vice versa. The effect of pretension on the initial modulus and the overall finite deformation stress-stretch behavior of the membrane is also studied; the initial modulus of the membrane is found to increase with chain pretension and the extensibility of the network is found to decrease with pretension. The proposed constitutive model can be used together with finite element modeling of the red blood cell to study complex deformation behaviors of the red blood cell; such cell deformation studies have been conducted by Dao et al. [22] utilizing the Evans, Skalak, Hochmuth neo-Hookean membrane model and, more recently, by Mills et al. [23] using a phenomenological higher-order I_1 (first stretch invariant) based constitutive model.[4] The proposed constitutive model also provides a framework to explore additional complexities on the mechanics of biomacromolecular network deformation including the effects of cross-linking proteins providing additional strain energy contributions during deformation, non-affine deformation in irregular triangulated networks (where mapping of macroscopic to microscopic deformation is not simply kinematically determined, but mechanical equilibrium must also be invoked), three-dimensional networked structures, and contributions of mechanically induced unfolding. Unfolding of modular biomacromolecules such as spectrin has been experimentally observed to have dramatic consequences on the single molecule force-extension behavior [e.g., [25,26]]; the unfolding event has been incorporated into recent eight-chain networks of protein solids, (Qi et al., [21,27]) and "four-chain" network representations of membranes (Qi, et al. [21]) whereby all chains experience the same stretching and thus unfold simultaneously. The triangulated network framework presented here provides a more physically representative microstructure for the spectrin network and will lead to unfolding in some sequential manner for the different chains (Arslan [28]) as evidenced by the different evolutions in chain forces with macroscopic stretching presented in this paper.

Acknowledgment

This research was supported by the U.S. Army through the Institute for Soldier Nanotechnologies, under Contract No. DAAD-19-02-D0002 with the U.S. Army research office. The content does not necessarily reflect the position of the Government and no official endorsement should be inferred.

Appendix

Expressions for the initial uniaxial and shear moduli: Note that the uniaxial modulus assumes constant surface area deformation of the membrane.

Initial axial modulus, E_o, *and shear modulus*, G_o, *for the LC case*:

$$E_o = \frac{v}{2}k_{\mathrm{ch}}r_o^2; \quad G_o = \frac{v}{8}k_{\mathrm{ch}}r_o^2.$$

Initial axial modulus, E_o, *and shear modulus*, G_o, *for the NGC case*:

---

[4] We note that Arruda and Boyce [15,16] and Boyce [24] have shown that higher order I_1-based models are phenomenological equivalents to non-Gaussian statistical models, in particular, to the eight-chain model of Arruda and Boyce [15].

$$E_o = \frac{\upsilon}{2} k_B \theta N \left(\frac{r_o}{L_c}\right)\left[3\beta + \left(\frac{r_o}{L}\right)\frac{\beta^2(\sinh\beta)^2}{(\sinh\beta)^2 - \beta^2}\right];$$

$$G_o = \frac{\upsilon}{8} k_B \theta N \left(\frac{r_o}{L_c}\right)\left[3\beta + \left(\frac{r_o}{L}\right)\frac{\beta^2(\sinh\beta)^2}{(\sinh\beta)^2 - \beta^2}\right].$$

References

[1] Ponder, E., 1948, *Hemolysis and Related Phenomena*, Grune and Stretton, New York.

[2] Mohandas, N., and Evans, E., 1994, "Mechanical Properties of the Red Cell Membrane in Relation to Molecular Structure and Genetic Effects," Annu. Rev. Biophys. Biomol. Struct., **23**, pp. 787–818.

[3] Evans, E. A., 1973, "A New Material Concept for the Red Cell Membrane," Biophys. J., **13**, pp. 926–940.

[4] Skalak, R., Tozeren, A., Zarda, R. P., and Chien, S., 1973, "Strain Energy Function of Red Blood Cell Membranes," Biophys. J., **13**, pp. 245–264.

[5] Evans, E., and Hochmuth, R. M., 1977, "Solid-Liquid Composite Model of Red Cell Membrane," J. Membr. Biol., **30**, pp. 351–362.

[6] Liu, S., Derick, L. H., and Palek, J., 1987, "Visualization of the Hexagonal Lattice in the Erythrocyte Membrane Skeleton," J. Cell Biol., **104**, pp. 527–536, by copyright permission of the Rockefeller University Press.

[7] Byers, T. J., and Branton, D., 1985, "Visualization of the Protein Associations in the Erythrocyte Membrane Skeleton," Proc. Natl. Acad. Sci. U.S.A., **82**, 6153–6157.

[8] Boal, D. H., 1994, "Computer Simulation of a Model Network for the Erythrocyte Cytoskeleton," Biophys. J., **67**, pp. 521–529.

[9] Discher, D. E., Boey, S. K, and Boal, D. H., 1997, "Phase Transitions and Anisotropic Responses of Planar Triangular Nets Under Large Deformation," Phys. Rev. E, **55**, pp. 4762–4772.

[10] Boey, S. K., Boal, D. H., and Discher, D. E., 1998, "Simulations of the Erythrocyte Cytoskeleton At Large Deformation, I. Microscopic Models," Biophys. J., **75**, pp. 1573–1583.

[11] Discher, D. E., Boal, D. H., and Boey, S. K., 1998, "Simulations of the Erythrocyte Cytoskeleton at Large Deformation, II. Micropipette Aspiration," Biophys. J., **75**, pp. 1584–1597.

[12] Wintz, W., Everaers, R., and Seifert, U., 1997, "Mesh Collapse in Two Dimensional Elastic Networks under Compression," J. Phys. I, **7**, pp. 1097–1111.

[13] Li, J., Dao, M., Lim, C. T., and Suresh, S., 2005, "Spectrin-Level Modeling of the Cytoskeleton and Optical Tweezers Stretching of the Erythrocyte," Biophys. J., **88**(5), pp. 3707–3719.

[14] Treloar, L. R. G., 1958, *The Physics of Rubber Elasticity*, Oxford, Clarendon Press.

[15] Arruda, E. M., and Boyce, M. C., 1993, "A Three-dimensional Constitutive Model for the Large Stretch Behavior of Elastomers," J. Mech. Phys. Solids, **41**, pp. 389–412.

[16] Boyce, M. C., and Arruda, E. M., 2000, "Constitutive Models of Rubber Elasticity: A Review," Rubber Chem. Technol., **73**, pp. 504–523.

[17] James, H. M., and Guth, E., 1943, "Theory of the Elastic Properties of Rubber," J. Chem. Phys., **11**, pp. 455–481.

[18] Boyce, M. C., and Arruda, E. M., 2001, "Swelling and Mechanical Stretching of Elastomeric Materials," Math. Mech. Solids, **6**, pp. 641–659.

[19] Bergstrom, J. S., and Boyce, M. C., 2001, "Deformation of Elastomeric Networks: Relation between Molecular Level Deformation and Classical Statistical Mechanics Models of Rubber Elasticity," Macromolecules, **34**, pp. 614–626.

[20] Arslan, M., and Boyce, M. C., 2005, "Constitutive Modeling of the Stress Stretch Behavior of Membranes Possessing a Triangulated Network Microstructure," Mater. Res. Soc. Symp. Proc., **874**, L.6.8.1–L.6.8.6.

[21] Qi, H. J., Ortiz, C., and Boyce, M. C., 2005, "Constitutive Model for the Stress-Strain Behavior of Biomacromolecular Networks Containing Folded Domains," (submitted).

[22] Dao, M., Lim, C. T., and Suresh, S., 2003, "Mechanics of the Human Red Blood Cell Deformed by Optical Tweezers," J. Mech. Phys. Solids, **51**, pp. 2259–2280.

[23] Mills, J. P., Qie, L., Dao, M., Lim, C. T., and Suresh, S., 2004, "Nonlinear Elastic and Viscoelastic Deformation of the Human Red Blood Cell with Optical Tweezers," Mech. Chem. Biosyst., **1**, pp. 169–180.

[24] Boyce, M. C., 1996, "Direct Comparison of the Gent and the Arruda-Boyce Constitutive Models of Rubber Elasticity," Rubber Chem. Technol., **69**, pp. 781–785.

[25] Rief, M., Pascual, J., Saraste, M., and Gaub, H. E., 1999, "Single Molecule Force Spectroscopy of Spectrin Repeats: Low Unfolding Forces in Helix Bundles," J. Mol. Biol., **286**, pp. 553–561.

[26] Law, R., Carl, P., Harper, S., Dalhaimer, P., Speicher, D., and Discher, D. E., 2003, "Cooperativity in Forced Unfolding of Tandem Spectrin Repeats," Biophys. J., **84**, pp. 533–544.

[27] Qi, H. J., Ortiz, C., and Boyce, M. C., 2005, "Protein Forced Unfolding and Its Effect on the Finite Deformation Stress-Strain Behavior of Biomacromolecular Solids," Mater. Res. Soc. Symp. Proc., **874**, pp. L.4.3.1–L.4.3.6.

[28] Arslan, M., 2005, "Constitutive Modeling of the Finite Deformation Behavior of Membranes Possessing a Triangulated Networked Microstructure," S.M. thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Xian-Fang Li
Institute of Mechanics and Sensor Technology,
School of Civil Engineering and Architecture,
Central South University,
Changsha 410083, China

L. Roy Xu[1]
Department of Civil and Environmental
Engineering,
Vanderbilt University,
Nashville, TN 37235
e-mail: l.roy.xu@vanderbilt.edu

Transient Response of a Finite Bimaterial Plate Containing a Crack Perpendicular to and Terminating at the Interface

The transient response of a finite bimaterial plate with a crack perpendicular to and terminating at the interface is analyzed for two types of boundaries (free-free and clamped-clamped). The crack surface is loaded by arbitrary time-dependent antiplane shear impact. The mixed initial-boundary value problem is reduced to a singular integral equation of a generalized Cauchy kernel for the crack tearing displacement density or screw dislocation density. The Gauss-Jacobi quadrature technique is employed to numerically solve the singular integral equation, and then the dynamic stress intensity factors are determined by implementing a numerical inversion of the Laplace transform. As an example, numerical calculations are carried out for a cracked bimaterial plate composed of aluminum (material I) and epoxy or steel (material II). The effects of material properties, geometry, and boundary types on the variations of dynamic stress intensity factors are discussed in detail. Results indicate that an overshoot of the normalized stress intensity factor of the crack tip at the interface decreases for a cracked bimaterial plate, and the occurrence of which is delayed for a cracked aluminum/epoxy plate compared to a pure aluminum plate with the same crack.
[DOI: 10.1115/1.2130734]

1 Introduction

Maintaining aging structures has become a very costly problem worldwide because tremendous efforts and considerable amount of time are required to ensure their safety and durability. A quick and effective technique is the use of a thin layer or film made of special materials to bond or cover the damaged or cracked areas of the aging structure. For example, repairing civil structures using composites in seismic environments has been very successful in limiting the earthquake damage. Very recently, another application has been in the use of advanced materials such as elastomer coatings for protecting structural failure from blast loading due to terrorist action. A common feature of these practical applications is the transient response of a finite bimaterial plate containing a crack perpendicular to and terminating at an interface [1]. Obviously, a fundamental investigation on the dynamic response of a thin layer bonded with a cracked material is very useful in providing guidance for practical applications. Since analytical results are difficult to obtain for a cracked bimaterial plate with a mode I or mode II crack under dynamic load [2], we will start from a simple mode III crack to study the effect of material property mismatch, etc. [3].

Dynamic behaviors of a cracked bimaterial plate under impact loading have been analyzed by a number of researchers including Stone et al. [4], Kuo [5,6], Ang [7], among others. In particular, for an antiplane shear crack embedded in a bimaterial, Karim and Kundu [8] treated transient surface response of layered isotropic and anisotropic half-spaces with interface cracks subjected to antiplane dynamic loading. Bostrom [9] considered elastic shear wave scattering for an antiplane interface crack, and Liu and Achenbach [10] studied wave scattering induced by cracks in a laminated plate. The elastodynamic problem of a semi-infinite interface crack and an interface crack of a finite crack length subjected to antiplane point impact loading has been dealt with by Kuo and Cheng [11] and Pramanik et al. [12], respectively. For a propagating antiplane shear interface crack, a fundamental solution has been derived by Ing and Ma [13]. An antiplane shear crack propagating in a self-similar manner along a bimaterial interface has been studied by Chung and Robinson [14]. Zhang [15] analyzed the dynamic stress intensity factors for a periodic array of collinear antiplane interface cracks. Furthermore, Wang and Gross [16] have proposed the scattering of shear waves for antiplane shear multicracks parallel to or lying along the interface of a layered medium. As we know, besides a class of cracks parallel to the interface or lying along the interface, another class of cracks inclined to or perpendicular to the interface is also often observed in laminated composite materials [17,18]. For such a class of cracks, Erdogan and Cook [19] studied the singularity of the static stress intensity factor for an antiplane shear crack normal to and terminating at the interface, and the stress intensity factor of an inclined crack terminating at a bimaterial interface has been analyzed by Bassani and Erdogan [20], and Ma and Hour [21]. The scattering of shear waves from an inclined crack terminating at a bimaterial interface has been formulated by Lu and Hanyga [22]. It is however noted that in the above-mentioned work, two dissimilar elastic bodies are assumed to be infinite, or at least one is infinite in extent. By using complex-variable integral equation along with a dislocation model, Wu and Chiu [23] gave a lower bound for the stress intensity factor having a fixed length-to-height ratio for a bimaterial with a rectangular cross section containing an antiplane shear crack terminating at the interface. However, for such a crack terminating at the interface of a finite bimaterial, the corresponding dynamic treatment is very limited. For a homogeneous material with a rectangular cross section, the transient response of multicracks has been determined by using Laplace finite element method by Chen et al. [24]. For a dynamic

---

[1] To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of ASME for publication in the APPLIED MECHANICS. Manuscript received June 12, 2005; final manuscript received September 26, 2005. Review conducted by S. Mukherjee. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California – Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

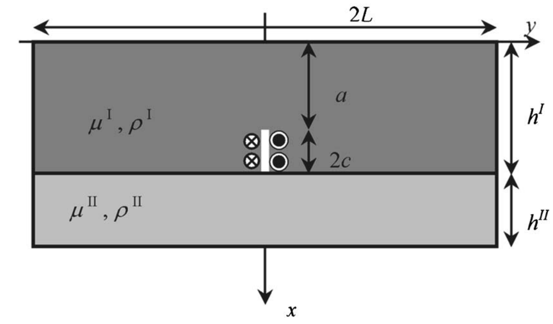

2L
y
a
$\mu^{\mathrm{I}}, \rho^{\mathrm{I}}$
2c
h^I
$\mu^{\mathrm{II}}, \rho^{\mathrm{II}}$
h^{II}
x

Fig. 1 A finite bimaterial plate with a through crack perpendicular to and terminating at the interface

elastic problem involving a cracked bimaterial with finite dimensions, the determination of the dynamic stress intensity factors becomes more complicated because of the interaction of the boundaries and excited waves.

Therefore, this paper is concerned with the dynamic analysis of a cracked finite bimaterial with a rectangular cross section. In Sec. 2, the governing equations and the corresponding initial-boundary conditions are stated. Section 3 is devoted to the derivation of a singular integral equation with a generalized Cauchy kernel, and in Sec. 4, the determination procedure of a numerical solution is provided. For two typical cracked aluminum/steel and aluminum/epoxy plates, the numerical dynamic stress intensity factors are determined by means of an inversion of the Laplace transform in Sec. 5. Finally, some numerical results are discussed in detail, and conclusions are extracted.

2 Formulation of the Problem

As shown in Fig. 1, the coordinate system is chosen such that the x axis is along the crack, and the y axis is along the upper surface of the cracked laminated plate. The z direction is perpendicular to the x-y plane. The thicknesses of the two perfectly bonded layers are denoted as h^{I} and $h^{\mathrm{II}}(h^{\mathrm{I}}+h^{\mathrm{II}}=h)$, respectively. Let the crack occupy the region $a<x<h^{\mathrm{I}}$, $y=0$, $-\infty<z<\infty$. The width of the plate is taken to be $2L$.

2.1 Initial-Boundary Conditions. From the viewpoint of fracture mechanics, the dynamic stress behavior around the crack tip is of significant interest. In other words, we study the transient response of a laminated plate when the crack surfaces are suddenly loaded by antiplane shear impact, i.e.

$$\tau_{zy}(x,0,t) = -\tau_0 f(x) H(t), \quad a < x < h^{\mathrm{I}},\ t > 0 \tag{1}$$

where τ_0 is a constant, $f(x)$ is a prescribed function depending on the position, and $H(t)$ is the Heaviside unit step function with respect to time. It should be noted that other impact functions could be used in place of $H(t)$ and be similarly analyzed without difficulty. In addition, the upper and lower surfaces of the laminated plate are assumed to be free of tractions, and can be stated as

$$\tau_{zx}(0,y,t) = 0 \quad \tau_{zx}(h,y,t) = 0 \quad -L < y < L,\ t > 0 \tag{2}$$

On the other hand, for other boundaries of the laminated plate, we consider two possible cases, free-free boundaries, and clamped-clamped boundaries. For these two cases, one can write the corresponding boundary conditions as

$$\tau_{zy}(x,\pm L,t) = 0 \quad 0 \leqslant x \leqslant h,\ t > 0 \quad \text{for free-free boundaries} \tag{3}$$

$$w(x,\pm L,t) = 0 \quad 0 \leqslant x \leqslant h,\ t > 0 \quad \text{for clamped-clamped boundaries} \tag{4}$$

2.2 Basic Equations. For a laminated plate subjected to antiplane shear load, the displacement components u and v along the x and y axes vanish, and there is only one out-of-plane displacement w, which gives two elastic stress components τ_{zx} and τ_{zy} based on the constitutive equations

$$\tau_{zx} = \mu \frac{\partial w}{\partial x} \quad \tau_{zy} = \mu \frac{\partial w}{\partial y} \tag{5}$$

where μ denotes the shear modulus of the elastic medium. With the aid of equations of motion, the out-of-plane displacement w must satisfy the wave equation [25]

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \frac{1}{c_s^2}\frac{\partial^2 w}{\partial t^2} \tag{6}$$

where c_s is the shear wave speed of the elastic medium, which is defined as

$$c_s = \sqrt{\frac{\mu}{\rho}} \tag{7}$$

In the present study, the laminated plate is composed of two dissimilar bonded materials. For convenience, the physical quantities in the elastic medium I (with a crack inside) occupying $0 \leqslant x \leqslant h^{\mathrm{I}}$, $|y| \leqslant L$, are denoted with superscript I, and similarly, those in the elastic medium II occupying $h^{\mathrm{I}} \leqslant x \leqslant h$, $|y| \leqslant L$, are denoted with superscript II.

3 Solution of the Problem

It is convenient to utilize the Laplace transform together with the integral equation technique [26] to solve the above-stated problem. In order to apply the Laplace transform, it is necessary to impose the condition that the laminated plate is initially at rest, i.e.

$$w = 0 \quad \frac{\partial w}{\partial t} = 0 \quad t < 0 \tag{8}$$

Under such circumstances, application of the Laplace transform, defined by

$$q^*(p) = \int_0^\infty q(t)\exp(-pt)\,dt \tag{9}$$

where the asterisk denotes the Laplace transform of a function with respect to t, p being the Laplace transform parameter, to the initial-boundary conditions (1)–(4) yield the following boundary conditions in the Laplace transform domain, respectively,

$$\tau_{zy}^*(x,0,p) = -\frac{\tau_0 f(x)}{p} \quad a < x < h^{\mathrm{I}} \tag{10}$$

$$\tau_{zx}^*(0,y,p) = 0 \quad \tau_{zx}^*(h,y,p) = 0 \quad -L < y < L \tag{11}$$

and

$$\tau_{zy}^*(x,\pm L,p) = 0 \quad 0 \leqslant x \leqslant h \quad \text{for free-free boundaries} \tag{12}$$

$$w^*(x,\pm L,p) = 0 \quad 0 \leqslant x \leqslant h \quad \text{for clamped-clamped boundaries} \tag{13}$$

3.1 Free-free Boundaries. First, we deal with the case of free-free boundaries. In this case, the symmetry of the problem allows us to consider only one half of the laminated plate, i.e., $y \geqslant 0$, hence, application of the Laplace transform to Eq. (6) yields a Poisson's equation in the Laplace transform space over a finite region. Employing the method of separation of variables along with the boundary condition (12), one may obtain eigenvalues of the resulting Poisson's equation and the corresponding eigenfunctions. Therefore, for $y \geqslant 0$ an appropriate solution of the problem in the Laplace transform domain may be written in the following form

$$w^{\mathrm{I}*}(x,y,p) = \int_0^\infty A(\xi)\cosh \alpha^{\mathrm{I}}(L-y)\cos \xi x d\xi + \sum_{n=0}^{\infty} B_n \cosh \gamma_n^{\mathrm{I}} x \sin \beta_n y \quad 0 \leqslant x \leqslant h^{\mathrm{I}} \tag{14}$$

$$w^{\mathrm{II}*}(x,y,p) = \sum_{n=0}^{\infty} C_n \cosh \gamma_n^{\mathrm{II}}(h-x)\sin \beta_n y \quad h^{\mathrm{I}} \leqslant x \leqslant h \tag{15}$$

with

$$\alpha^{\mathrm{I}} = \sqrt{\xi^2 + \left(\frac{p}{c_s^{\mathrm{I}}}\right)^2} \tag{16}$$

$$\beta_n = \frac{2n+1}{2L}\pi \tag{17}$$

$$\gamma_n^J = \sqrt{\beta_n^2 + \left(\frac{p}{c_s^J}\right)^2} \quad J = \mathrm{I}, \mathrm{II} \tag{18}$$

where $A(\xi)$, B_n, and $C_n (n=0,1,2,\ldots)$ are unknown to be determined from given boundary conditions. In fact, a direct check via substituting Eqs. (14) and (15) into the Laplace transformed equation (6) indicates that it is automatically satisfied. Moreover, the boundary conditions (11) and (12) are fulfilled identically by utilizing the constitutive relations (5). What remains is to determine $A(\xi)$, B_n, and $C_n (n=0,1,2,\ldots)$ through the application of boundary conditions at the crack surfaces in connection with the interface continuity conditions.

However, in what follows $A(\xi)$, B_n, and C_n are not solved directly. Instead, we eliminate these unknown quantities from the given boundary conditions to derive an integral equation, which can then be solved. To this end, from the expressions of displacements, with the aid of constitutive equations one can write the stress field in the Laplace transform domain as

$$\tau_{zx}^{\mathrm{I}*}(x,y,p) = -\mu^{\mathrm{I}} \int_0^\infty \xi A(\xi)\cosh \alpha^{\mathrm{I}}(L-y)\sin \xi x d\xi + \mu^{\mathrm{I}} \sum_{n=0}^{\infty} \gamma_n^{\mathrm{I}} B_n \sinh \gamma_n^{\mathrm{I}} x \sin \beta_n y \tag{19}$$

$$\tau_{zy}^{\mathrm{I}*}(x,y,p) = -\mu^{\mathrm{I}} \int_0^\infty \alpha^{\mathrm{I}} A(\xi)\sinh \alpha^{\mathrm{I}}(L-y)\cos \xi x d\xi + \mu^{\mathrm{I}} \sum_{n=0}^{\infty} \beta_n B_n \cosh(\gamma_n^{\mathrm{I}} x)\cos \beta_n y \tag{20}$$

for $0 \leqslant x \leqslant h^{\mathrm{I}}$, and

$$\tau_{zx}^{\mathrm{II}*}(x,y,p) = -\mu^{\mathrm{II}} \sum_{n=0}^{\infty} \gamma_n^{\mathrm{II}} C_n \sinh \gamma_n^{\mathrm{II}}(h-x)\sin \beta_n y \tag{21}$$

$$\tau_{zy}^{\mathrm{II}*}(x,y,p) = \mu^{\mathrm{II}} \sum_{n=0}^{\infty} \beta_n C_n \cosh \gamma_n^{\mathrm{II}}(h-x)\cos \beta_n y \tag{22}$$

for $h^{\mathrm{I}} \leqslant x \leqslant h$.

Now, by applying the continuity condition of displacement and stress at the interface $x=h^{\mathrm{I}}$ to Eqs. (14) and (15), (19) and (21) will yield two equations, then, making use of the orthogonal properties of sine functions, one can solve for B_n and C_n in terms of $A(\xi)$. Omitting the detailed procedure, the final results are

$$B_n = \frac{2\beta_n}{L \cosh \gamma_n^{\mathrm{I}} h^{\mathrm{I}}} \int_0^\infty \frac{A(\xi)\,\eta_B \cosh \alpha^{\mathrm{I}} L}{\beta_n^2 + (\alpha^{\mathrm{I}})^2} d\xi \tag{23}$$

$$C_n = \frac{2\beta_n}{L \cosh \gamma_n^{\mathrm{II}} h^{\mathrm{II}}} \int_0^\infty \frac{A(\xi)\,\eta_C \cosh \alpha^{\mathrm{I}} L}{\beta_n^2 + (\alpha^{\mathrm{I}})^2} d\xi \tag{24}$$

with

$$\eta_B = \frac{\xi \mu^{\mathrm{I}} \sin \xi h^{\mathrm{I}} - \mu^{\mathrm{II}} \gamma_n^{\mathrm{II}} \tanh \gamma_n^{\mathrm{II}} h^{\mathrm{II}} \cos \xi h^{\mathrm{I}}}{\mu^{\mathrm{I}} \gamma_n^{\mathrm{I}} \tanh \gamma_n^{\mathrm{I}} h^{\mathrm{I}} + \mu^{\mathrm{II}} \gamma_n^{\mathrm{II}} \tanh \gamma_n^{\mathrm{II}} h^{\mathrm{II}}} \tag{25}$$

$$\eta_C = \frac{\mu^{\mathrm{I}} \gamma_n^{\mathrm{I}} \tanh \gamma_n^{\mathrm{I}} h^{\mathrm{I}} \cos \xi h^{\mathrm{I}} + \xi \mu^{\mathrm{I}} \sin \xi h^{\mathrm{I}}}{\mu^{\mathrm{I}} \gamma_n^{\mathrm{I}} \tanh \gamma_n^{\mathrm{I}} h^{\mathrm{I}} + \mu^{\mathrm{II}} \gamma_n^{\mathrm{II}} \tanh \gamma_n^{\mathrm{II}} h^{\mathrm{II}}} \tag{26}$$

where in deriving the above results, the known equality [27]

$$\int_0^L \cosh \alpha^{\mathrm{I}}(L-y)\sin \beta_n y dy = \frac{\beta_n}{\beta_n^2 + (\alpha^{\mathrm{I}})^2} \cosh \alpha^{\mathrm{I}} L \tag{27}$$

has been used. Furthermore, substituting Eq. (23) into Eq. (14), we get the expression for the out-of-plane displacement given by the Fourier cosine integral as

$$w^{\mathrm{I}*}(x,0,p) = \int_0^\infty A(\xi)\cosh \alpha^{\mathrm{I}} L \cos \xi x d\xi \tag{28}$$

Consequently, similar to Erdogan [28], if introducing the crack tearing displacement density $g(x,t)$, defined by

$$g(x,t) = \frac{\partial w(x,0,t)}{\partial x} \tag{29}$$

the single-value condition of the displacement requires that $g(x,t)$ must obey the constraint

$$\int_a^{h^{\mathrm{I}}} g(x,t)dx = 0 \tag{30}$$

Moreover, by using the Fourier transform one can represent $A(\xi)$ by the following integral

$$A(\xi) = -\frac{2}{\pi \xi \cosh \alpha^{\mathrm{I}} L} \int_a^{h^{\mathrm{I}}} g^*(s,p)\sin \xi s ds \tag{31}$$

On the other hand, applying the boundary condition (10) at the crack surfaces to expression (20) for the antiplane stress gives

$$-\mu^{\mathrm{I}} \int_0^\infty \alpha^{\mathrm{I}} A(\xi)\sinh \alpha^{\mathrm{I}} L \cos \xi x d\xi + \mu^{\mathrm{I}} \sum_{n=0}^{\infty} \beta_n B_n \cosh \gamma_n^{\mathrm{I}} x = -\frac{\tau_0 f(x)}{p} \tag{32}$$

Furthermore, to eliminate auxiliary unknown functions $A(\xi)$ and B_n, we substitute Eq. (23) in connection with Eq. (31) into Eq. (32). By employing some known results, listed in the Appendix, and after a lengthy derivation, one can reach a singular integral equation with a generalized Cauchy kernel of the first kind

$$\frac{1}{\pi} \int_a^{h^{\mathrm{I}}} \frac{g^*(s,p)}{s-x} ds + \frac{1}{\pi} \frac{\mu^{\mathrm{I}} - \mu^{\mathrm{II}}}{\mu^{\mathrm{I}} + \mu^{\mathrm{II}}} \int_a^{h^{\mathrm{I}}} \frac{g^*(s,p)}{s+x-2h^{\mathrm{I}}} ds + \frac{1}{\pi} \int_a^{h^{\mathrm{I}}} g^*(s,p) R_{FF}(s,x,p) ds = -\frac{\tau_0 f(x)}{\mu^{\mathrm{I}} p} \tag{33}$$

where the Fredholm kernel $R_{FF}(s,x,p)$ is given in the Appendix. This resulting singular integral equation is similar to that derived by Erdogan and Cook [19], who analyzed the static elasticity problem of an antiplane shear crack perpendicular to and terminating at the interface of two half-planes, rather than a finite bimaterial plate as carried out in this investigation. Clearly, the effects of the boundaries can be described through the nonsingular term of the left-hand side of Eq. (33).

3.2 Clamped-clamped Boundaries.

Now we turn our attention to the case of clamped-clamped boundaries. The procedure of deriving a governing singular integral equation is similar to the case of free-free boundaries in the previous subsection. To fulfill all boundary conditions for the case of clamped-clamped boundaries, instead of expressions (14) and (15), the displacements in the Laplace transform domain in this case should be chosen as

$$w^{\mathrm{I}*}(x,y,p)=\int_0^\infty A(\xi)\sinh\alpha^{\mathrm{I}}(L-y)\cos\xi x d\xi + \sum_{n=1}^{\infty} B_n\cosh\gamma_n^{\mathrm{I}}x\sin\beta_n y \quad 0\leqslant x\leqslant h^{\mathrm{I}} \tag{34}$$

$$w^{\mathrm{II}*}(x,y,p)=\sum_{n=1}^{\infty} C_n\cosh\gamma_n^{\mathrm{II}}(h-x)\sin\beta_n y \quad h^{\mathrm{I}}\leqslant x\leqslant h \tag{35}$$

where α^{I} is defined as before

$$\beta_n=\frac{n}{L}\pi \tag{36}$$

$$\gamma_n^J=\sqrt{\beta_n^2+\left(\frac{p}{c_s^J}\right)^2} \quad J=\mathrm{I},\mathrm{II} \tag{37}$$

$A(\xi)$, B_n, and $C_n(n=1,2,\ldots)$ are unknown and are to be determined from the given boundary conditions. It is noted that the dependence of γ_n^J on β_n is the same for both free-free boundary and clamped-clamped boundary cases, while the dependence of β_n on n is not the same for the two cases. Furthermore, by using the constitutive equations one can write the stress field in the Laplace transform domain as follows

$$\tau_{zx}^{\mathrm{I}*}(x,y,p)=-\mu^{\mathrm{I}}\int_0^\infty \xi A(\xi)\sinh\alpha^{\mathrm{I}}(L-y)\sin\xi x d\xi + \mu^{\mathrm{I}}\sum_{n=1}^{\infty}\gamma_n^{\mathrm{I}}B_n\sinh\gamma_n^{\mathrm{I}}x\sin\beta_n y \tag{38}$$

$$\tau_{zy}^{\mathrm{I}*}(x,y,p)=-\mu^{\mathrm{I}}\int_0^\infty \alpha^{\mathrm{I}}A(\xi)\cosh\alpha^{\mathrm{I}}(L-y)\cos\xi x d\xi + \mu^{\mathrm{I}}\sum_{n=1}^{\infty}\beta_n B_n\cosh(\gamma_n^{\mathrm{I}}x)\cos\beta_n y \tag{39}$$

for $0\leqslant x\leqslant h^{\mathrm{I}}$, and

$$\tau_{zx}^{\mathrm{II}*}(x,y,p)=-\mu^{\mathrm{II}}\sum_{n=1}^{\infty}\gamma_n^{\mathrm{II}}C_n\sinh\gamma_n^{\mathrm{II}}(h-x)\sin\beta_n y \tag{40}$$

$$\tau_{zy}^{\mathrm{II}*}(x,y,p)=\mu^{\mathrm{II}}\sum_{n=1}^{\infty}\beta_n C_n\cosh\gamma_n^{\mathrm{II}}(h-x)\cos\beta_n y \tag{41}$$

for $h^{\mathrm{I}}\leqslant x\leqslant h$. Now, application of the continuity condition for both displacement and stress at the interface $x=h^{\mathrm{I}}$ to Eqs. (34) and (35), (38) and (40) yields two equations, the orthogonal properties of the sine function allows us to solve B_n and C_n in terms of $A(\xi)$, i.e.

$$B_n=\frac{2\beta_n}{L\cosh\gamma_n^{\mathrm{I}}h^{\mathrm{I}}}\int_0^\infty\frac{A(\xi)\eta_B\sinh\alpha^{\mathrm{I}}L}{\beta_n^2+(\alpha^{\mathrm{I}})^2}d\xi \tag{42}$$

$$C_n=\frac{2\beta_n}{L\cosh\gamma_n^{\mathrm{II}}h^{\mathrm{II}}}\int_0^\infty\frac{A(\xi)\eta_C\sinh\alpha^{\mathrm{I}}L}{\beta_n^2+(\alpha^{\mathrm{I}})^2}d\xi \tag{43}$$

with η_B and η_C given by Eqs. (25) and (26). Here in deriving the above results, we have used the following result [27]

$$\int_0^L \sinh\alpha^{\mathrm{I}}(L-y)\sin\beta_n y dy=\frac{\beta_n}{\beta_n^2+(\alpha^{\mathrm{I}})^2}\sinh\alpha^{\mathrm{I}}L \tag{44}$$

In a similar treatment as the previous subsection, substituting the obtained results into the remaining boundary conditions and eliminating the introduced auxiliary function $A(\xi)$, we can derive the singular integral equation. Omitting the details, the governing singular integral equation for this case is

$$\frac{1}{\pi}\int_a^{h^{\mathrm{I}}}\frac{g^*(s,p)}{s-x}ds+\frac{1}{\pi}\frac{\mu^{\mathrm{I}}-\mu^{\mathrm{II}}}{\mu^{\mathrm{I}}+\mu^{\mathrm{II}}}\int_a^{h^{\mathrm{I}}}\frac{g^*(s,p)}{s+x-2h^{\mathrm{I}}}ds + \frac{1}{\pi}\int_a^{h^{\mathrm{I}}}g^*(s,p)R_{CC}(s,x,p)ds=-\frac{\tau_0 f(x)}{\mu^{\mathrm{I}}p} \tag{45}$$

where the Fredholm kernel $R_{CC}(s,x,p)$ is given in the Appendix. The unknown density is also subjected to the following condition

$$\int_a^{h^{\mathrm{I}}}g^*(s,p)ds=0 \tag{46}$$

4 Numerical Treatment

In the previous section, the governing singular integral equation for the crack tearing displacement density has been derived for both cases of free-free boundaries and clamped-clamped boundaries. Because of the complicated form of the kernel, we appeal to a numerical approach to solve it. For numerical computation, we introduce the following normalized variables $\bar{s}$ and $\bar{x}$ such that

$$s=\frac{h^{\mathrm{I}}-a}{2}\bar{s}+\frac{h^{\mathrm{I}}+a}{2} \quad x=\frac{h^{\mathrm{I}}-a}{2}\bar{x}+\frac{h^{\mathrm{I}}+a}{2} \tag{47}$$

The above resulting singular integral equation for each case, Eq. (33) or Eq. (45) together with the constraint condition (46) can be rewritten as a normalized form

$$\frac{1}{\pi}\int_{-1}^{1}\frac{G^*(\bar{s},p)}{\bar{s}-\bar{x}}d\bar{s}+\frac{\lambda}{\pi}\int_{-1}^{1}\frac{G^*(\bar{s},p)}{\bar{s}+\bar{x}-2}ds+\frac{1}{\pi}\int_{-1}^{1}G^*(\bar{s},p)\bar{R}(\bar{s},\bar{x},p)d\bar{s} = -\frac{\bar{f}(\bar{x})}{p} \tag{48}$$

subjected to

$$\int_{-1}^{1}G^*(\bar{s},p)d\bar{s}=0 \tag{49}$$

where

$$\lambda=\frac{\mu^{\mathrm{I}}-\mu^{\mathrm{II}}}{\mu^{\mathrm{I}}+\mu^{\mathrm{II}}} \tag{50}$$

$$G^*(\bar{s},p)=\frac{1}{\tau_0}\mu^{\mathrm{I}}g^*(s,p) \tag{51}$$

$$\bar{R}(\bar{s},\bar{x},p)=\frac{h^{\mathrm{I}}-a}{2}R(s,x,p) \tag{52}$$

$$\bar{f}(\bar{x}) = f(x) \tag{53}$$

In the above expressions, $R(s,x,p)$ stands either for $R_{FF}(s,x,p)$ or $R_{CC}(s,x,p)$. Taking into account that $\bar{R}(\bar{s},\bar{x},p)$ is a continuous bounded function in $[-1,1]$, and that $1/(\bar{s}+\bar{x}-2)$ possesses a singularity of order 1 as $\bar{s}$ and $\bar{x}$ tend to the interface crack tip, simultaneously. Consequently, this is a singular integral equation with a generalized Cauchy kernel. To determine the singularity order of the stress field near the crack tip, following the method described in Muskhelishvili [29], the unknown function $G^*(\bar{x},p)$ in Eq. (48) can be assumed to be of the form

$$G^*(\bar{x},p) = \frac{\Omega(\bar{x},p)}{(1+\bar{x})^{\kappa_1}(1-\bar{x})^{\kappa_2}} \quad 0 < \kappa_1, \kappa_2 < 1 \tag{54}$$

where $\Omega(\bar{x},p)$ is assumed to be a Hölder continuous function over $[-1,1]$. Substituting Eq. (54) into Eq. (48), after some manipulations it follows that the constants κ_1 and κ_2 are determined from

$$\cot \pi\kappa_1 = 0 \quad \cos \pi\kappa_2 = -\lambda \tag{55}$$

It means that the function $G^*(\bar{x},p)$ has a usual square-root singularity near $\bar{x}=-1$ corresponding to the crack tip away from the interface (crack tip inside material I), and a $r^{-\kappa_2}$ $(0<\kappa_2<1)$ singularity of the crack tip at the interface, coinciding with the static results for an antiplane shear crack normal to and terminating at the bimaterial interface [19]. Solving the first equation in (55) one can write $G^*(\bar{x},p)$ below

$$G^*(\bar{x},p) = \frac{\Omega(\bar{x},p)}{(1-\bar{x})^{\kappa}(1+\bar{x})^{1/2}} \tag{56}$$

with

$$\kappa = 1 - \frac{1}{\pi}\cos^{-1}\left(\frac{\mu^{\mathrm{I}} - \mu^{\mathrm{II}}}{\mu^{\mathrm{I}} + \mu^{\mathrm{II}}}\right) \tag{57}$$

where the index 2 in κ_2 is omitted, i.e., κ_2 is abbreviated as κ. It is observed that the dynamic stress singularity is the same as the corresponding static one. In particular, the singularity order of the crack tip at the interface is related to the material properties. Moreover, it is readily concluded that if the uncracked layer is stiffer than the cracked layer, i.e., $\mu^{\mathrm{II}} > \mu^{\mathrm{I}}$. The singularity order of the crack tip at the interface is less than 1/2, while if the uncracked layer is more compliant than the cracked layer, meaning $\mu^{\mathrm{II}} < \mu^{\mathrm{I}}$, the singularity order of the crack tip is greater than 1/2. These results are different from an interfacial crack of a bimaterial since its material order (material I or II) is exchangeable. However, in this investigation, material I with a crack inside is fixed.

Now, in view of the expression for $G^*(\bar{x},p)$ given by Eq. (56), Eq. (48) subjected to the constraint (49) can be solved by approximating $\Omega(\bar{x},p)$ by a Lagrange interpolation polynomial at the zeros of Jacobi polynomial $P_n^{(-\kappa,-1/2)}(\bar{x})$ [30]. That is, suppose that $\bar{s}_j$ $(j=1,2,\ldots,n)$ are the zeros of Jacobi polynomial $P_n^{(-\kappa,-1/2)}(\bar{s})$. Making use of the Lagrange interpolation formula, we can write

$$\Omega(\bar{s},p) \approx \sum_{j=1}^{n} \Omega_j L_j(\bar{s}) \tag{58}$$

with

$$\Omega_j = \Omega(\bar{s}_j,p) \tag{59}$$

$$L_j(\bar{s}) = \prod_{1\leqslant m\leqslant n, m\neq j} \frac{\bar{s}-\bar{s}_m}{\bar{s}_j-\bar{s}_m} \tag{60}$$

Recalling the well-known Gauss-Jacobin quadrature formula [31]

$$\int_{-1}^{1} \frac{\Omega(u)}{(1-u)^{\kappa}(1+u)^{\iota}} du \approx \sum_{j=1}^{n} A_j\Omega_j \tag{61}$$

where A_j is the corresponding weight or the Christoffel numbers associated with the zeros of Jacobi polynomial $P_n^{(-\kappa,-1/2)}(u)$. It has been shown that the above approximation is exact for all $\Omega(u)$, if $\Omega(u)$ is a polynomial of the degree less than $2n-1$. Hence the singular integral equation with a generalized Cauchy kernel (48) with the condition (49) can be approximated by the following algebraic equation

$$\frac{1}{\pi}\sum_{j=1}^{n}\left\{\left[\frac{1}{\bar{s}_j-\bar{x}_k} + \frac{\lambda}{\bar{s}_j+\bar{x}_k-2} + \bar{R}(\bar{s}_j,\bar{x}_k,p)\right]A_j + B_{kj}\right\}\Omega_j = -\frac{\bar{f}(\bar{x}_k)}{p} \quad 1 \leqslant k \leqslant n-1 \tag{62}$$

$$\sum_{j=1}^{n} A_j\Omega_j = 0 \tag{63}$$

with

$$B_{kj} = L_j(\bar{x}_k)\left(q_0(\bar{x}_k) - \sum_{m=1}^{n}\frac{A_m}{\bar{s}_m-\bar{x}_k}\right) \tag{64}$$

$$q_0(\bar{x}_k) = -\frac{\pi\cot\pi\kappa}{(1-\bar{x}_k)^{\kappa}(1+\bar{x}_k)^{1/2}} - \sqrt{\frac{\pi}{2}}\frac{\Gamma(-\kappa)}{2^{\kappa}\Gamma(0.5-\kappa)} \times F\left(1,\kappa+0.5;\kappa+1;\frac{1-\bar{x}_k}{2}\right) \tag{65}$$

where $\Gamma(*)$ denotes the gamma function, and F denotes the hypergeometric function, respectively. In our calculations, according to the suggestions of Govorukha and Loboda [32], we take the collocation points $\bar{x}_k$ as the middle points of two adjacent zeros of the Jacobi polynomial. Thus Eqs. (62) and (63) form a system of n algebraic equations in n unknown Ω_j, and its solution can be determined by straightforward computation.

On the other hand, to determine the stress intensity factors at the crack tips, it is necessary to obtain the asymptotic stress fields ahead of the crack tips. Bearing in mind the known result [33]

$$\int_{-1}^{1}\frac{(1-u)^{-\kappa}(1+u)^{-\iota}}{(u-z)}du = -\frac{\pi\,\mathrm{cosec}\,\pi\kappa}{(z-1)^{\kappa}(z+1)^{\iota}} - \frac{\Gamma(-\kappa)\Gamma(1-\iota)}{2^{\kappa+\iota}\Gamma(1-\kappa-\iota)} \times F\left(1,\kappa+\iota;1+\kappa;\frac{1-z}{2}\right), \tag{66}$$

for $z \notin [-1,1]$, where z is a complex variable with $z=x+iy$, for the present analysis $\iota=1/2$, after some calculations, the stress field τ_{zy}^{I} in the Laplace transform domain in material I can be expressed as

$$\tau_{zy}^{\mathrm{I}*}(x,0,p) = \left(\frac{h^{\mathrm{I}}-a}{2}\right)^{\kappa+1/2}\frac{\tau_0\Omega(-1,p)}{(h^{\mathrm{I}}-x)^{\kappa}(a-x)^{1/2}} + O(1) \quad x \approx a-0 \tag{67}$$

where $O(1)$ denotes nonsingular terms. As regards the stress field τ_{zy}^{II} in material II, which is similar to Eq. (22), one can derive the leading term associated with the singular stress field to be given by the following singular integral

$$\tau_{zy}^{\mathrm{II}*}(x,0,p) = \frac{1}{\pi}\frac{2\mu^{\mathrm{II}}}{\mu^{\mathrm{I}}+\mu^{\mathrm{II}}}\int_{a}^{h^{\mathrm{I}}}\frac{g^*(s,p)}{s+x-2h^{\mathrm{I}}}ds + O(1), \quad x > h^{\mathrm{I}} \tag{68}$$

Letting x be located ahead of the crack at the interface, it follows from Eq. (68) that the asymptotic stress field takes the form

Table 1 Material constants of chosen elastic media

Material	μ (GPa)	ρ (kg/m^3)	κ
Al	26.1	2700	0.5
Steel	80.1	7850	0.33
Epoxy	1.49	1120	0.851

$$\tau_{zy}^{\mathrm{II}*}(x,0,p) = -\sqrt{\frac{\mu^{\mathrm{II}}}{\mu^{\mathrm{I}}}}\left(\frac{h^{\mathrm{I}}-a}{2}\right)^{\kappa+1/2}\frac{\tau_0\Omega(1,p)}{(x-h^{\mathrm{I}})^{\kappa}(x-a)^{1/2}} + O(1),\ x \approx h^{\mathrm{I}}+0 \tag{69}$$

In the above derivation we have applied the result

$$\sin \kappa\pi = \frac{2\sqrt{\mu^{\mathrm{I}}\mu^{\mathrm{II}}}}{\mu^{\mathrm{I}}+\mu^{\mathrm{II}}} \tag{70}$$

Based on the singularity of the crack tips, we use the following definition of the stress intensity factors

$$K_{\mathrm{hom}} = \lim_{x\to a-0}[2\pi(a-x)]^{1/2}\tau_{zy}^{\mathrm{I}}(x,0,t)$$
$$K_{\mathrm{int}} = \lim_{x\to h^{\mathrm{I}}+0}[2\pi(x-h^{\mathrm{I}})]^{\kappa}\tau_{zy}^{\mathrm{II}}(x,0,t) \tag{71}$$

where the stress intensity factors with the subscript "hom" and "int" represent those parameters corresponding to the crack tips inside the homogeneous material (material I) and at the interface, respectively. Substituting the asymptotic stress expressions (67) and (69) into the above definitions yields the desired stress intensity factors in the Laplace transform domain as follows

$$K_{\mathrm{hom}}^{*} = \tau_0(\pi c)^{0.5}2^{0.5-\kappa}\Omega(-1,p)$$
$$K_{\mathrm{int}}^{*} = -\tau_0\sqrt{\mu^{\mathrm{II}}/\mu^{\mathrm{I}}}(\pi c)^{\kappa}2^{\kappa-0.5}\Omega(1,p) \tag{72}$$

where κ is given by Eq. (57) and we have denoted c as the length of the half-crack, i.e.

$$c = \frac{h^{\mathrm{I}}-a}{2} \tag{73}$$

Clearly, if setting $\mu^{\mathrm{I}}=\mu^{\mathrm{II}}$, meaning $\kappa=1/2$, the above two stress intensity factors reduce to

$$K^{*} = \pm\tau_0(\pi c)^{1/2}\Omega(\mp 1,p) \tag{74}$$

and is consistent with the well-known results given by Chen and Sih [34].

5 Results and Discussions

To obtain the dynamic stress intensity factors in time domain, a numerical inversion of the Laplace transform by means of the Fourier series expansion formulated by Crump [35] is adopted in the present study. As compared to other methods for numerically inverting the Laplace transform based only on some values along the real-axis, this method has an advantage, i.e., it produces more accurate results since the Fourier series expansion at some complex-plane points contains more abundant information [36].

We choose aluminum (Al) as material I, and either epoxy or steel as material II, which is denoted as Al/Epoxy or Al/Steel, respectively. The relevant material constants are listed in Table 1 [37], where the last column is the singularity order κ of the crack tip at the interface when material II is perfectly bonded to aluminum, and is calculated using Eq. (57). The singularity order κ of Al/epoxy and Al/steel material combination are 0.851 and 0.33. The prescribed function $f(x)$ is taken to be unity.

Prior to the presentation of dynamic stress intensity factors, it is necessary to give relevant static results for validation purpose. For a perfectly bonded infinite bimaterial with an antiplane shear crack normal to and terminating at the interface, an explicit expression for the crack tearing displacement density or the screw dislocation density has been obtained by Chou [38], from which we readily derive the corresponding static stress intensity factors at two crack tips, given by

$$K_{\mathrm{hom}}^{\infty} = \tau_0(\pi c)^{0.5}\frac{\sqrt{2}\kappa}{\sin\frac{\pi\kappa}{2}} \quad K_{\mathrm{int}}^{\infty} = \tau_0(\pi c)^{\kappa}\frac{2^{3\kappa-1}(1-\kappa)}{\sin^2\frac{\pi\kappa}{2}}\cos\frac{\pi\kappa}{2} \tag{75}$$

The above results reduce to the well-known ones for a homogeneous isotropic material, meaning that $\kappa=0.5$, and $K_{\mathrm{hom}}^{\infty}=K_{\mathrm{int}}^{\infty}=\tau_0(\pi c)^{0.5}$. However, for a bimaterial consisting of two dissimilar bonded materials [3,39], due to the mismatched material constants, the stress intensity factor of the crack tip located at the interface, is no longer proportional to $c^{1/2}$, c being the half crack length. So the dimension of the interfacial stress intensity factor is no longer [stress][length]$^{1/2}$ or [F][m]$^{-3/2}$, but [stress][length]$^{\kappa}$ or [F][m]$^{\kappa-2}$. Moreover, the static stress intensity factors at two crack tips for an infinite cracked bimaterial are dependent on the material properties of the two mismatched materials, and vary with the singularity order κ or with the ratio of $\mu^{\mathrm{I}}/\mu^{\mathrm{II}}$, as shown in Fig. 2.

5.1 Dynamic Stress Intensity Factors of Different Bimaterial Combinations. In this section dynamic stress intensity factors at two crack tips are examined. For convenience, the normalized stress intensity factors are introduced and compared.

$$k_{\mathrm{hom}}(t) = \frac{K_{\mathrm{hom}}(t)}{K_{\mathrm{hom}}^{\infty}} \quad k_{\mathrm{int}}(t) = \frac{K_{\mathrm{int}}(t)}{K_{\mathrm{int}}^{\infty}} \tag{76}$$

First, to examine the efficiency and accuracy of our method, the response of a central crack is considered, i.e., $a=h^{\mathrm{II}}$. Figure 3 shows the variations of the normalized stress intensity factor $k_{\mathrm{hom}}(t)$ at the crack tip away from the interface versus the normalized time $c_s^{\mathrm{I}}t/c$ for Al/epoxy and Al/steel plates with $a:2c:L=5:2:10$ under the free-free type boundaries. For comparison, the response of the dynamic stress intensity factor for a centrally cracked aluminum plate with the same geometry is plotted in Fig. 3. It is easily found that at the early stage of applied impact loading, the response curves rise abruptly. The curve corresponding to the Al/steel laminate reaches a dynamic overshoot $k_{\mathrm{hom}}(t)=1.313$ when $c_s^{\mathrm{I}}t/c=2$, and its variation is similar to that of a cracked aluminum plate which has its overshoot $k_{\mathrm{hom}}(t)=1.259$ at $c_s^{\mathrm{I}}t/c=2$. However, for the Al/epoxy plate with $\kappa=0.851>0.5$, $k_{\mathrm{hom}}(t)=1.072$ at $c_s^{\mathrm{I}}t/c=2$, and its overshoot with $k_{\mathrm{hom}}(t)=1.22$ is postponed to occur at $c_s^{\mathrm{I}}t/c=4$. Since before the time at which either the stress wave from the other crack tip or those reflected from the boundaries arrive, the stress fields at the crack tips are precisely the same as those derived for a semi-infinite crack under the same loading. For comparison, we use Achenbach's analytical solution on a semi- infinite mode III crack under sudden uniform loading [40] to calculate the dynamic overshoots of the crack inside material I. Under the same time $c_s^{\mathrm{I}}t/c=2$, the analytical solution yields 1.273 for Al, 1.352 for Al/steel, and 1.029 for Al/epoxy, which are very close to our numerical values. As time further passes all curves approach certain constant values, corresponding to the static stress intensity factors for cracked plates with the same geometry. Since $k_{\mathrm{hom}}(t)=k_{\mathrm{int}}(t)=1$ for an infinite cracked bimaterial when $t\to\infty$, our calculated results are very close to 1, and a slight difference results from the effects of the boundaries. So our results are very satisfactory, and the method presented is efficient. On the other hand, the response history of $k_{\mathrm{int}}(t)$ (the crack tip at the interface) is presented in Fig. 4. Some similar transient trends can be seen. However, three overshoots of the dynamic stress intensity factors for Al/steel, Al, Al/epoxy are 1.186, 1.26, taking place at $c_s^{\mathrm{I}}t/c=2$, and 1.231 at $c_s^{\mathrm{I}}t/c=3.3$, respectively.

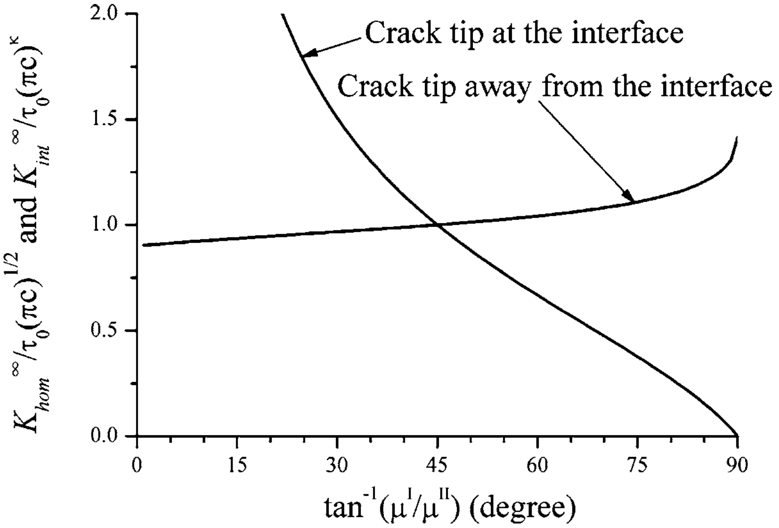


Fig. 2 Stress intensity factors $K^{\infty}_{\mathrm{hom}}/\tau_0(\pi c)^{1/2}$ and $K^{\infty}_{\mathrm{int}}/\tau_0(\pi c)^{1/2}$ as a function of $\tan^{-1}(\mu^{I}/\mu^{II})$

5.2 Effect of the Thickness of Material II. For a thin material II, the responses of the normalized dynamic stress intensity factors are displayed in Figs. 5 and 6 for the free-free boundaries with $a:c:h^{II}:L=2:1:0.1:10$. In this case, it is observed that as compared to a cracked pure Al plate, the peak of $k_{\mathrm{hom}}(t)$ increases a little bit for an Al/steel plate as $c_s^I t/c=2$, and the corresponding peak for an Al/epoxy plate is delayed to $c_s^I t/c=4$ with a basically unchanged value. However, the value of $k_{\mathrm{hom}}(t)$ at $c_s^I t/c=2$ for the Al/epoxy plate declines apparently, implying that in this case the interaction of the dynamic stress intensity factors is slow compared to the other two cases. For $k_{\mathrm{int}}(t)$ at the interface, the peak values of the dynamic stress intensity factors, irrespective of an Al/steel or Al/epoxy plate, become smaller than that of a cracked aluminum plate, in particular for a smaller value of h^{II}/a. These peaks are tabulated in Table 2. The table reveals that under impact loading, a thin layer/coating can lower the dynamic stress intensity factor peak of the crack tip at the interface.

Figures 7 and 8 show the effect of epoxy thickness on the dynamic stress intensity factor variations with $a:c:L=2:1:10$. The curves in Fig. 7 are completely overlapped during the period of $c_s^I t/c<2$ and then branch after $c_s^I t/c>2$, which is obviously due to the arrival of the shear wave generated from the crack tip at the interface. Furthermore, for a cracked Al/epoxy plate, it is seen that $k_{\mathrm{hom}}(t)$ reaches a peak at $c_s^I t/c=4$, and then drops as $c_s^I t/c$ rises, which is due to the reflected wave both from the boundary $x=0$ and from the interface $x=h^I$. As regards the crack tip at the interface, the overshoot of $k_{\mathrm{int}}(t)$ takes place before that of $k_{\mathrm{hom}}(t)$, which is partially attributed to the fact that it takes a shorter time for the reflected shear wave from the boundary $x=h^I+h^{II}$ to reach the crack tip at the interface $x=a+h^I$, than for the reflected shear wave from the boundary $x=0$ to reach the crack tip $x=a$, since $h^{II}\leqslant a$ is assumed in Fig. 8. In addition, when using steel as material II (instead of epoxy), the variations of the dynamic stress intensity factors $k_{\mathrm{hom}}(t)$ and $k_{\mathrm{int}}(t)$ at two crack tips are plotted in Figs. 9 and 10, and $k_{\mathrm{hom}}(t)$ and $k_{\mathrm{int}}(t)$ simultaneously reach their peaks at $c_s^I t/c=2$. Because of the influence of the reflected shear wave, the response curve corresponding to $h^{II}/c=0.1$ has the highest value among three curves, and that corresponding to $h^{II}/c=2$ has the lowest. These results indicate that for a thin layer bonded to a cracked media, dynamic stress intensity factors have more apparent transient behaviors.

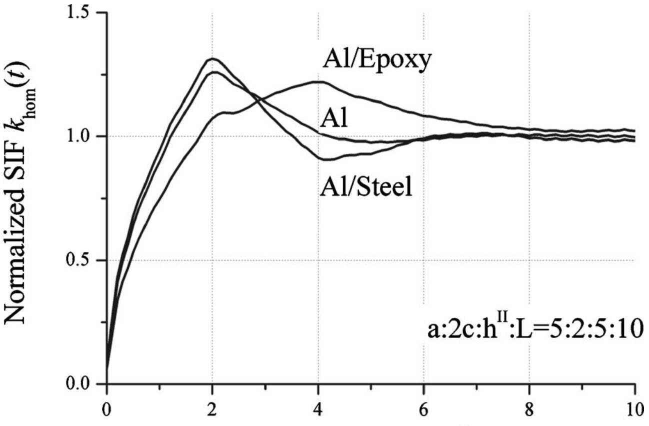


Fig. 3 Normalized stress intensity factors $k_{\mathrm{hom}}(t)$ versus normalized time $c_s^I t/c$ for a centrally cracked plate having free-free boundaries

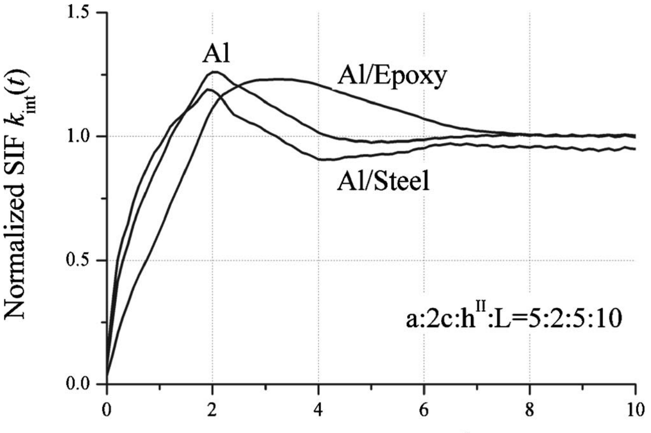


Fig. 4 Normalized stress intensity factors $k_{\mathrm{int}}(t)$ versus normalized time $c_s^I t/c$ for a centrally cracked plate having free-free boundaries

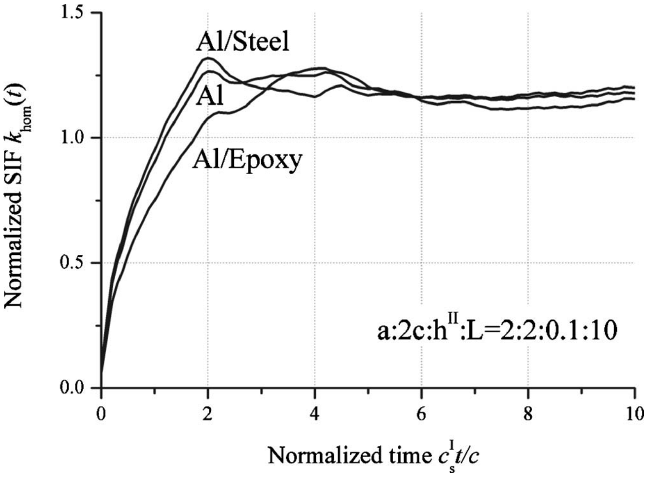


Fig. 5 Normalized stress intensity factors $k_{hom}(t)$ versus normalized time $c_s^I t/c$ for a cracked plate with $a:c:h^{II}:L=2:1:0.1:10$ having free-free boundaries

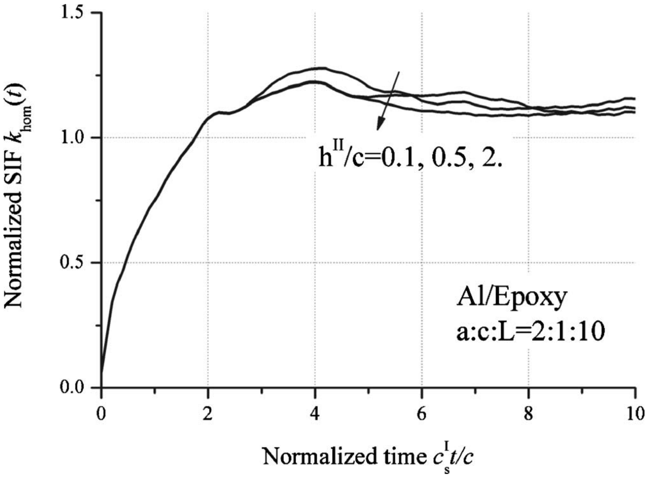


Fig. 7 The effects of the thickness of epoxy on $k_{hom}(t)$ for $a:c:L=2:1:10$ having free-free boundaries

5.3 Effect of the Boundary Conditions. Another interesting phenomenon is the influence of the boundary type on the dynamic stress intensity factor variation. Figures 11 and 12 illustrate the response of a cracked Al/epoxy plate under the free-free, and clamped-clamped boundaries with plate dimensions of $a:c:h^{II}=2:1:0.5$ and different L/c values. As expected, for $L/c=10$, the responses under the free-free and clamped-clamped boundaries are identical within $c_s^I t/c<10$. However, for $L/c=1$ and 2, there is significant discrepancy between the response curves corresponding to the free-free and clamped- clamped boundaries. Reflected waves can be generated when waves arrive at the free-free boundaries, which enhances the stress intensity factor, while transmitted wave occurs at the clamped-clamped boundaries, which weakens the stress intensity factor. Moreover, one can conclude that it takes $c_s^I t/c=1$ for excited shear waves to travel along the boundaries $y=\pm L$ for the case of $L/c=1$, and it takes the same time period to travel back to the crack tips. Therefore, in Figs. 11 and 12, the dynamic stress intensity factors are enhanced or weakened, occurring when $c_s^I t/c \geqslant 2$ for $L/c=1$. Similarly, the same conclusion can be inferred for $L/c=2$. Comparing the response curves for the free-free type boundaries to those for the clamped-clamped boundaries, the overshoot of the dynamic stress intensity factor for the free-free type boundaries is larger than that of the clamped-clamped boundaries, irrespective of $k_{hom}(t)$ or $k_{int}(t)$, suggesting that a crack in a bimaterial under the free-free boundaries is easier to extend than that under the clamped-clamped boundaries, in particular for smaller L.

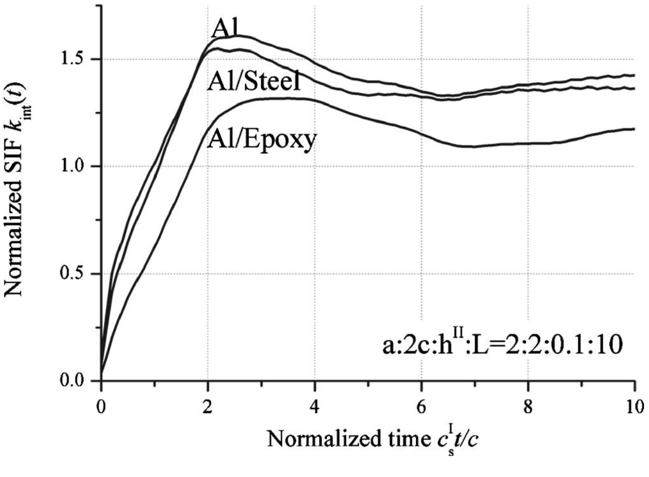


Fig. 6 Normalized stress intensity factors $k_{int}(t)$ versus normalized time $c_s^I t/c$ for a cracked plate with $a:c:h^{II}:L=2:1:0.1:10$ having free-free type boundaries

Table 2 Peak values of the normalized dynamic stress intensity factors with $a:c:L=2:1:10$

	k_{hom}			k_{int}		
h^{II}/c	2	0.5	0.1	2	0.5	0.1
Al	1.263	1.264	1.266	1.262	1.318	1.61
Al/steel	1.314	1.315	1.319	1.189	1.226	1.549
Al/epoxy	1.221	1.224	1.278	1.235	1.25	1.317

6 Conclusions

A rectangular bimaterial plate containing a through crack perpendicular to and terminating at the interface is considered. Making use of the Laplace transform, the associated initial-boundary value problem is converted to a singular integral equation with a generalized Cauchy kernel for the crack tearing displacement density. By using the Gauss-Jacobi quadrature method, the resulting singular integral equation is approximated to be a system of alge-

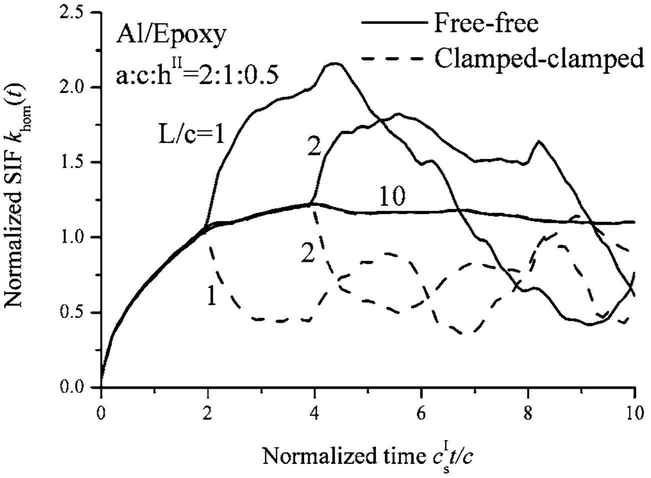


Fig. 11 The effects of the boundary types on $k_{hom}(t)$ for a cracked Al/epoxy plate with $a:c:h^{II}=2:1:0.5$ for various values of L/c

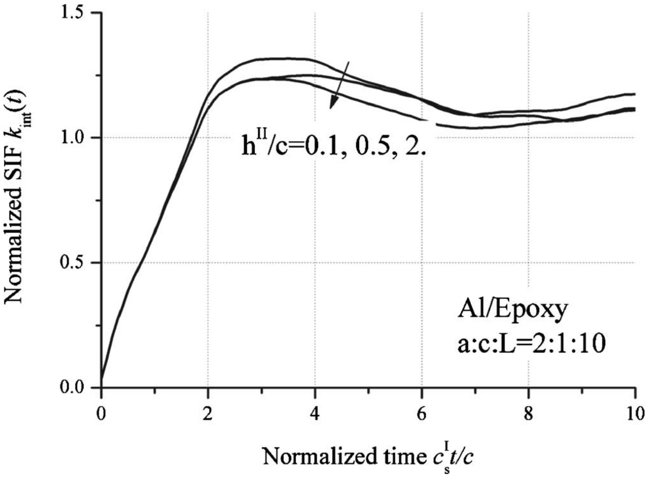


Fig. 8 The effects of the thickness of epoxy on $k_{int}(t)$ for $a:c:L=2:1:10$ having free-free boundaries

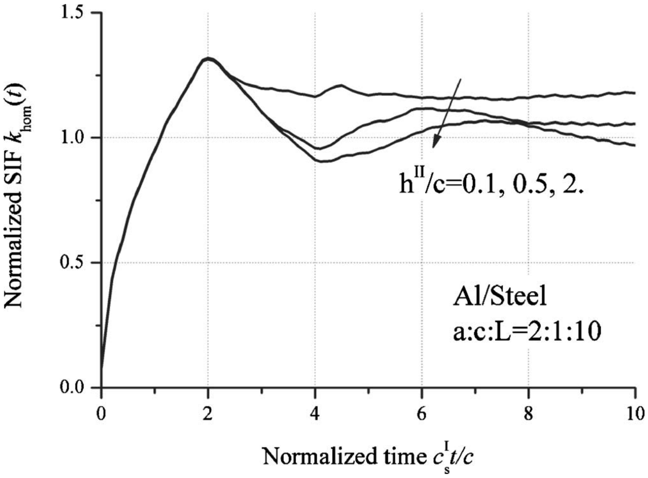


Fig. 9 The effects of the thickness of steel on $k_{hom}(t)$ for $a:c:L=2:1:10$ having free-free boundaries

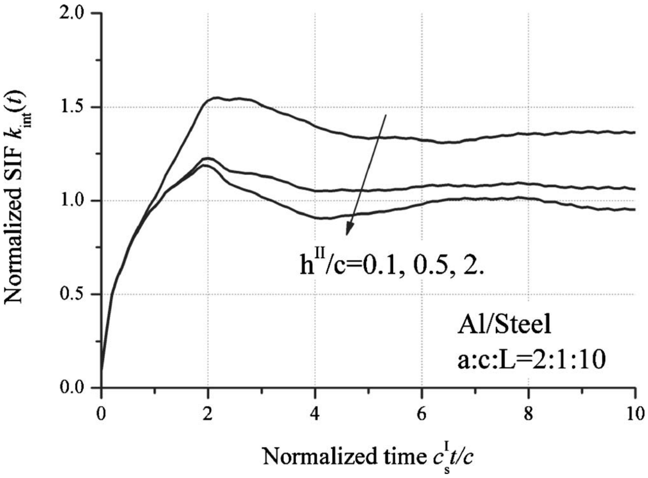


Fig. 10 The effects of the thickness of steel on $k_{int}(t)$ for $a:c:L=2:1:10$ having free-free boundaries

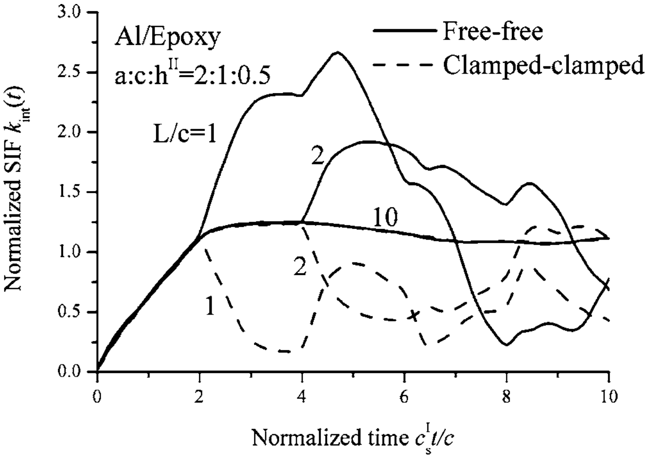


Fig. 12 The effects of the boundary types on $k_{int}(t)$ for a cracked Al/epoxy plate with $a:c:h^{II}=2:1:0.5$ for various values of L/c

braic equations. Numerical calculations are performed for typical bimaterial plates and the effects of material properties, geometry, and boundary types on the dynamic stress intensity factor variations are discussed. The following can be concluded:

- The dynamic elastic stress field for the crack tip at the interface exhibits a singularity of κ order, $\kappa=1-\cos^{-1}[(\mu^{I}-\mu^{II})/(\mu^{I}+\mu^{II})]$, while for the crack tip embedded in material I, still remains a square-root singularity. Also, the dynamic stress singularity order of the crack tip is the same as the static one.
- The overshoot of the normalized dynamic stress intensity factor of the crack tip at the interface for a cracked Al/epoxy or Al/steel plate declines as compared to that for a pure aluminum plate with the same crack.
- The occurrence of the overshoot of the dynamic stress intensity factor is delayed for a cracked Al/epoxy plate. Also, the transient response of the dynamic stress intensity factor is more evident for a thin film perfectly bonded to an edge-cracked layer.
- The free-free type boundaries cause increasing in the dynamic stress intensity factor, while the clamped-clamped type boundaries lead to decreasing in the dynamic stress intensity factor.

Acknowledgment

L.R.X. gratefully acknowledges the support from the Office of Naval Research Young Investigator Program (N00014-03-1-0505, Dr. Roshdy G. S. Barsoum, Program Officer), and the National Science Foundation, Structural Systems and Hazards Mitigation of Structures Program (CMS-0456807, Dr. Steven L. McCabe, Program Officer). Valuable suggestions from one anonymous reviewer are appreciated.

Appendix

In deriving singular integral equations (33) and (45), the following results have been used

$$\frac{2}{\pi}\int_0^\infty \sin \xi s \cos \xi x d\xi = \frac{1}{\pi}\left(\frac{1}{s-x}+\frac{1}{s+x}\right) \tag{A1}$$

$$\int_0^\infty \frac{a^2 \sin \xi b}{\xi(\xi^2+a^2)} d\xi = \frac{\pi}{2}(1-e^{-ab}) \quad a,b>0 \tag{A2}$$

$$\int_0^\infty \frac{a \cos \xi b}{\xi^2 + a^2} d\xi = \frac{\pi}{2} e^{-ab} \quad a,b > 0 \tag{A3}$$

$$\int_0^\infty \frac{a \sin \xi b \sin \xi c}{\xi^2 + a^2} d\xi = \begin{cases} (\pi/2) e^{-ab} \sinh ac, \ 0 < c < b \\ (\pi/2) e^{-ac} \sinh ab, \ 0 < b < c \end{cases} \tag{A4}$$

$$\int_0^\infty \frac{a^2 \sin \xi b \sin \xi c}{\xi^2 + a^2} d\xi = \begin{cases} (\pi/2) e^{-ab} \sinh ac, \ 0 < c < b \\ (\pi/2)(1 - e^{-ac} \cosh ab), \ 0 < b < c \end{cases} \tag{A5}$$

$$\sum_{n=0}^{\infty} \exp(-(2n+1)t) = \frac{1}{2 \sinh t} \tag{A6}$$

$$\sum_{n=1}^{\infty} \exp(-nt) = \frac{\exp(-t/2)}{2 \sinh t/2} \tag{A7}$$

In Eq. (33), the kernel $R_{FF}(s,x,p)$ takes the form

$$R_{FF}(s,x,p) = R_1(s,x,p) + R_2(s,x) + R_3(s,x,p) \tag{A8}$$

$$R_1(s,x,p) = \frac{1}{s+x} + 2\int_0^\infty \left(\frac{\alpha^{\mathrm{I}}}{\xi} \tanh \alpha^{\mathrm{I}} L - 1 \right) \sin \xi s \cos \xi x d\xi \tag{A9}$$

$$R_2(s,x) = -\frac{\pi}{L} \frac{\mu^{\mathrm{I}} - \mu^{\mathrm{II}}}{\mu^{\mathrm{I}} + \mu^{\mathrm{II}}} \left[\frac{1}{2 \sinh \beta(2h^{\mathrm{I}} - s - x)} - \frac{1}{2\beta(2h^{\mathrm{I}} - s - x)} \right] \tag{A10}$$

$$\begin{aligned} R_3(s,x,p) = -\frac{2\pi}{L} \sum_{n=0}^{\infty} \Bigg\{ & \frac{\mu^{\mathrm{I}} \gamma_n^{\mathrm{I}} - \mu^{\mathrm{II}} \gamma_n^{\mathrm{II}} \tanh \gamma_n^{\mathrm{II}} h^{\mathrm{II}}}{\mu^{\mathrm{I}} \gamma_n^{\mathrm{I}} \tanh \gamma_n^{\mathrm{I}} h^{\mathrm{I}} + \mu^{\mathrm{II}} \gamma_n^{\mathrm{II}} \tanh \gamma_n^{\mathrm{II}} h^{\mathrm{II}}} \\ & \times \left(\frac{\beta_n}{\gamma_n^{\mathrm{I}}} \right)^2 \frac{\cosh \gamma_n^{\mathrm{I}} x}{\cosh \gamma_n^{\mathrm{I}} h^{\mathrm{I}}} \exp(-\gamma_n^{\mathrm{I}} h^{\mathrm{I}}) \sinh(\gamma_n^{\mathrm{I}} s) \\ & - \frac{1}{2} \frac{\mu^{\mathrm{I}} - \mu^{\mathrm{II}}}{\mu^{\mathrm{I}} + \mu^{\mathrm{II}}} \exp[-2\beta_n(2h^{\mathrm{I}} - s - x)] \Bigg\} \end{aligned} \tag{A11}$$

where in the above expressions, $\beta = \pi/2L$, and $\beta_n = (2n+1)\pi/2L$.

In Eq. (45), the kernel $R_{CC}(s,x,p)$ takes the form

$$R_{CC}(s,x,p) = R_4(s,x,p) + R_5(s,x) + R_6(s,x,p) \tag{A12}$$

$$R_4(s,x,p) = \frac{1}{s+x} + 2\int_0^\infty \left(\frac{\alpha^{\mathrm{I}}}{\xi} \coth \alpha^{\mathrm{I}} L - 1 \right) \sin \xi s \cos \xi x d\xi \tag{A13}$$

$$R_5(s,x) = -\frac{\pi}{L} \frac{\mu^{\mathrm{I}} - \mu^{\mathrm{II}}}{\mu^{\mathrm{I}} + \mu^{\mathrm{II}}} \left[\frac{\exp \beta(s + x - 2h^{\mathrm{I}})}{2 \sinh \beta(2h^{\mathrm{I}} - s - x)} - \frac{1}{2\beta(2h^{\mathrm{I}} - s - x)} \right] \tag{A14}$$

$$\begin{aligned} R_6(s,x,p) = -\frac{2\pi}{L} \sum_{n=0}^{\infty} \Bigg\{ & \frac{\mu^{\mathrm{I}} \gamma_n^{\mathrm{I}} - \mu^{\mathrm{II}} \gamma_n^{\mathrm{II}} \tanh \gamma_n^{\mathrm{II}} h^{\mathrm{II}}}{\mu^{\mathrm{I}} \gamma_n^{\mathrm{I}} \tanh \gamma_n^{\mathrm{I}} h^{\mathrm{I}} + \mu^{\mathrm{II}} \gamma_n^{\mathrm{II}} \tanh \gamma_n^{\mathrm{II}} h^{\mathrm{II}}} \\ & \times \left(\frac{\beta_n}{\gamma_n^{\mathrm{I}}} \right)^2 \frac{\cosh \gamma_n^{\mathrm{I}} x}{\cosh \gamma_n^{\mathrm{I}} h^{\mathrm{I}}} \exp(-\gamma_n^{\mathrm{I}} h^{\mathrm{I}}) \sinh(\gamma_n^{\mathrm{I}} s) \\ & - \frac{1}{2} \frac{\mu^{\mathrm{I}} - \mu^{\mathrm{II}}}{\mu^{\mathrm{I}} + \mu^{\mathrm{II}}} \exp[-2\beta_n(2h^{\mathrm{I}} - s - x)] \Bigg\} \end{aligned} \tag{A15}$$

where $\beta = \pi/2L$, and $\beta_n = n\pi/L$.

References

[1] Xu, L. R., and Rosakis, A. J., 2002, "Impact Failure Characteristics in Sandwich Structures; Part II: Effects of Impact Speed and Interfacial Strength," Int. J. Solids Struct., **39**, pp. 4237–4248.

[2] Xu, L. R., Huang, Y. Y., and Rosakis, A. J., 2003, "Dynamic Crack Deflection and Penetration at Interfaces in Homogeneous Materials: Experimental Studies and Model Predictions," J. Mech. Phys. Solids, **51**, pp. 461–486.

[3] Hutchinson, J. W., and Suo, Z., 1992, "Mixed Mode Cracking in Layered Materials," Adv. Appl. Mech., **29**, pp. 63–191.

[4] Stone, S. F., Ghosh, M. L., and Mal, A. K., 1980, "Diffraction of Antiplane Shear-Waves by an Edge Crack," J. Appl. Mech., **47**, pp. 359–362.

[5] Kuo, A.-Y., 1984, "Transient Stress Intensity Factors of an Interfacial Crack between Two Dissimilar Anisotropic Half-Spaces. Part 1: Orthotropic Materials," J. Appl. Mech., **51**, pp. 71–76.

[6] Kuo, A.-Y., 1984, "Transient Stress Intensity Factors of an Interfacial Crack between Two Dissimilar Anisotropic Half-Spaces. Part 2: Fully Anisotropic Materials," J. Appl. Mech., **51**, pp. 780–786.

[7] Ang, W. T., 1988, "A Crack in an Anisotropic Layered Material Under the Action of Impact Loading," J. Appl. Mech., **55**, pp. 120–125.

[8] Karim, M. R., and Kundu, T., 1988, "Transient Surface Response of Layered Isotropic and Anisotropic Half-Spaces with Interface Cracks: SH Case," Int. J. Fract., **37**, pp. 245–262.

[9] Bostrom, A., 1987, "Elastic Wave Scattering from an Interface Crack–Antiplane Strain," J. Appl. Mech., **54**, pp. 503–508.

[10] Liu, G. R., and Achenbach, J. D., 1995, "Strip Element Method to Analyze Wave Scattering by Cracks in Anisotropic Laminated Plates," J. Appl. Mech., **62**, pp. 607–713.

[11] Kuo, M. K., and Cheng, S. H., 1991, "Elastodynamic Response due to Antiplane Point Impact Loadings on the Faces of an Interface Crack along Dissimilar Anisotropic Materials," Int. J. Solids Struct., **28**, pp. 751–768.

[12] Pramanik, R. K., Pal, S. C., and Ghosh, M. L., 1998, "Transient Response due to a Pair of Antiplane Point Impact Loading on the Faces of a Finite Griffith Crack at the Bimaterial Interface of Anisotropic Solids," Int. J. Eng. Sci., **36**, pp. 1197–1213.

[13] Ing, Y. S., and Ma, C. C., 1997, "Transient Analysis of a Subsonic Propagating Interface Crack Subjected to Antiplane Dynamic Loading in Dissimilar Isotropic Materials," J. Appl. Mech., **64**, pp. 546–556.

[14] Chung, Y. L., and Robinson, A. R., 1992. "The Transient Problem of a Mode-III Interface Crack," Eng. Fract. Mech., **41**, pp. 321–330.

[15] Zhang, Ch., 1991, "Dynamic Stress Intensity Factors for Periodically Spaced Collinear Antiplane Shear Cracks between Dissimilar Media," Theor. Appl. Fract. Mech., **15**, pp. 219–227.

[16] Wang, Y. S., and Gross, D., 2001, "Transfer Matrix Method of Wave Propagation in a Layered Medium with Multiple Interface Cracks: Antiplane Case," J. Appl. Mech., **68**, pp. 499–503.

[17] Leedy, K. D., and Stubbins, J. F., 2001, "Copper Alloy–Stainless Steel Bonded Laminates for Fusion Reactor Applications: Crack Growth and Fatigue," Mater. Sci. Eng., A, **297**, pp. 19–25.

[18] Schlichtinga, K. W., Padture, N. P., Jordanb, E. H., and Gell, M., 2003, "Failure Modes in Plasma-Sprayed Thermal Barrier Coatings," Mater. Sci. Eng., A, **342**, pp. 120–130.

[19] Erdogan, F., and Cook, T. S., 1974, "Antiplane Shear Crack Terminating at and Going Through a Bimaterial Interface," Int. J. Fract., **10**, pp. 227–240.

[20] Bassani, J. L., and Erdogan, F., 1979, "Stress Intensity Factors in Bonded Half Planes Containing Inclined Cracks and Subjected to Antiplane Shear Loading," Int. J. Fract., **15**, pp. 145–148.

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

[22] Lu, J.-F., and Hanyga, A., 2004, "Scattering of SH Wave by a Crack Terminating at the Interface," Comput. Mech., **34**, pp. 75–84.

[23] Wu, K. C., and Chiu, Y. T., 1991, "Antiplane Shear Interface Cracks in Anisotropic Bimaterials," J. Appl. Mech., **58**, pp. 399–403.

[24] Chen, W.-H., Chang, C.-L., and Tsai, C.-H., 2000, "Elastodynamic Fracture Analysis of Multiple Cracks by Laplace Finite Element Alternating Method," J. Appl. Mech., **67**, pp. 606–615.

[25] Freund, L. B., 1990, *Dynamic Fracture Mechanics*, Cambridge University Press, Cambridge.

[26] Sneddon, I. N., 1951, *Fourier Transforms*, McGraw-Hill, New York.

[27] Gradshteyn, I. S., and Ryzhik, I. M., 1980, *Table of Integrals, Series and Products*, Academic, New York.

[28] Erdogan, F., 1978, "Mixed Boundary-Value Problems in Mechanics," Mech. Today, **4**, pp. 1–86.

[29] Muskhelishvili, N. I., 1953, *Singular Integral Equations*, Noordhoff, Groningen, The Netherlands.

[30] Elliott, D., and Paget, D. F., 1979, "Gauss Type Quadrature Rules for Cauchy Principal Value Integrals," Math. Comput., **33**, pp. 301–309.

[31] Ghizzetti, A., and Ossicini, A., 1970, *Quadrature Formulae*, Academic, New York.

[32] Govorukha, V. B., and Loboda, V. V., 1996, "On a Numerical Method of the Solution of Singular Integral Equations," ZAMM, **76**, pp. 541–542.

[33] Erdelyi, A., Magnus, W., Oberhettinger, F., and Tricomi, F. G. eds., 1953, *Higher Transcendental Functions*, Vol. II, McGraw-Hill, New York.

[34] Chen, E. P., and Sih, G. C., 1977, *Elastodynamic Crack Problems*, Noordhoff International, Leyden, The Netherlands.

[35] Crump, K. S., 1976, "Numerical Inversion of Laplace Transforms using a Fourier Series Approximation," J. Assoc. Comput. Mach., **23**, pp. 89–96.

[36] Li, X.-F., and Duan, X. Y., 2003, "Comparison of Dynamic Response of a Piezoelectric Ceramic Containing Two Parallel Cracks via Two Methods of Laplace Inversion," Int. J. Fract., **122**, pp. L131–L136.

[37] Matthews, F. L., and Rawlings, R. D., 1994, *Composite Materials: Engineering and Science*, Chapman & Hall, London.

[38] Chou, T. W., 1970, "Dislocation Pileups and Elastic Cracks at a Bimaterial Interface," Metall. Trans., **1**, pp. 1245–1248.

[39] Barsoum, R. S., 1988, "Application of the Finite Element Iterative Method to the Eigenvalue Problem of a Crack between Dissimilar Media," Int. J. Numer. Methods Eng., **26**, pp. 541–554.

[40] Achenbach, J. D., 1970, "Brittle and Ductile Extension of a Finite Crack by a Horizontally Polarized Shear Wave," Int. J. Eng. Sci., **8**, pp. 947–966.

Ka-Veng Yuen
Department of Civil and Environmental Engineering,
University of Macau,
Macau, China
e-mail: kvyuen@umac.mo

James L. Beck
Division of Engineering and Applied Science,
California Institute of Technology,
Mail Code 104-44,
Pasadena, CA 91125
e-mail: jimbeck@caltech.edu

Lambros S. Katafygiotis
Department of Civil Engineering,
Hong Kong University of Science and Technology,
Clearwater Bay,
Kowloon, Hong Kong

Unified Probabilistic Approach for Model Updating and Damage Detection

A probabilistic approach for model updating and damage detection of structural systems is presented using noisy incomplete input and incomplete response measurements. The situation of incomplete input measurements may be encountered, for example, during low-level ambient vibrations when a structure is instrumented with accelerometers that measure the input ground motion and the structural response at a few instrumented locations but where other excitations, e.g., due to wind, are not measured. The method is an extension of a Bayesian system identification approach developed by the authors. A substructuring approach is used for the parameterization of the mass, damping and stiffness distributions. Damage in a substructure is defined as stiffness reduction established through the observation of a reduction in the values of the various substructure stiffness parameters compared with their initial values corresponding to the undamaged structure. By using the proposed probabilistic methodology, the probability of various damage levels in each substructure can be calculated based on the available dynamic data. Examples using a single-degree-of-freedom oscillator and a 15-story building are considered to demonstrate the proposed approach. [DOI: 10.1115/1.2150235]

1 Introduction

The problem of identification of the model parameters of a linear structural model using dynamic data has received much attention over the years because of its importance in model updating, response prediction, structural control and health monitoring. Many methodologies have been formulated, both in the time and frequency domain, for the cases of known and unknown input.

Structural health monitoring has been attracting much attention in the past two decades, including several workshops, e.g., Natke and Yao [1]; Agbabian and Masri [2]; Chang [3]; and special issues of journals, e.g., Journal of Engineering Mechanics (July 2000 and January 2004) and Computer-Aided Civil and Infrastructure Engineering (January 2001). Many methods have been developed. One such example is the class of direct methods using pattern recognition techniques (Mazurek and DeWolf [4]; Hearn and Testa [5]; Doebling et al. [6]; Lam et al. [7]; Smyth et al. [8]). Another example is the class of structural model-based inverse methods (Farhat and Hemez [9]; Pandey and Biswas [10]; Kim et al. [11]; Topole and Stubbs [12]; Hemez and Farhat [13]; Katafygiotis et al. [14]; Doebling et al. [15]; Vanik et al. [16]; Beck et al. [17]; Sohn and Farrar [18]; Ko et al. [19]; Ching and Beck [20]).

Recent interest has been shown in using Bayesian probabilistic approaches for model updating and damage detection as they allow for an explicit treatment of all the uncertainties involved (Geyskens et al. [21]; Beck and Katafygiotis [22]; Katafygiotis et al. [14]; Vanik et al. [16]; Katafygiotis and Yuen [23]; Yuen [24]). An advantage of the Bayesian approach is that it follows directly from the probability axioms and so there are no ad-hoc assumptions that lead to loss of information. In Beck and Katafygiotis [22], a methodology for model updating based on a Bayesian probabilistic system identification framework was presented. Although the framework is general, their presentation is for the case where the prediction error due to measurement noise and modeling error is modeled as Gaussian white noise.

In the present paper, the prediction error is modeled as the sum of a filtered white noise process, representing the input error (measurement noise plus unmeasured excitation) filtered through the system, plus another independent white noise process, representing the response measurement noise and modeling error. A Bayesian time-domain approach for modal identification by Yuen and Katafygiotis [25] is extended to handle the case of model updating with incomplete input measurements and with measurement noises in both input and output measurements. The proposed approach allows for the direct calculation of the probability density function (PDF) of the model parameters based on the data which can be then approximated by an appropriately selected multi-variate Gaussian distribution. By using data from the initial undamaged state and a later possibly damaged state, the probability of damage of various levels in specified substructures may be calculated. The formulation is presented for linear multi-degree-of-freedom (MDOF) systems. Examples using noisy simulated data from a single-degree-of-freedom (SDOF) oscillator and a 15-story building are given for illustration.

2 Model Formulation

2.1 Class of Structural Models. Consider a class of possible models for a structural system with N_d degrees of freedom (DOFs) and equation of motion

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{T}\mathbf{g} \tag{1}$$

where **M**, **C**, and **K** are the mass, damping and stiffness matrix of the system, respectively, $\mathbf{g} \in \mathbb{R}^{N_g}$ is the actual forcing vector and $\mathbf{T} \in \mathbb{R}^{N_d \times N_g}$ is the forcing distribution matrix. The mass, damping and stiffness matrices, **M**, **C**, and **K**, are defined in terms of mass, damping and stiffness parameters $\boldsymbol{\theta}_m$, $\boldsymbol{\theta}_c$, and $\boldsymbol{\theta}_s$ by

$$\mathbf{M} = \sum_{j=1}^{N_{\text{sub}}} \mathbf{M}_j(\boldsymbol{\theta}_m), \quad \mathbf{C} = \sum_{j=1}^{N_{\text{sub}}} \mathbf{C}_j(\boldsymbol{\theta}_c), \quad \mathbf{K} = \sum_{j=1}^{N_{\text{sub}}} \mathbf{K}_j(\boldsymbol{\theta}_s) \tag{2}$$

where N_{sub} is the number of substructures and $\mathbf{M}_j$, $\mathbf{C}_j$, and $\mathbf{K}_j$ are the contributions to the mass, damping and stiffness matrix of the

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 11, 2004; final manuscript received September 29, 2005. Review conducted by I. Mezic. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

jth substructure, respectively. Note that it is not necessary to require classical normal modes.

2.2 Stochastic Input Model. Assume that discrete-time input data $\{\mathbf{f}[k]: k=1,\ldots,N\}$ are available for the excitation where the index k refers to time $(k-1)\Delta t$ with Δt being the sampling interval. Define the uncertain input error $\boldsymbol{\eta}_f$ by

$$\mathbf{g}[k] = \mathbf{f}[k] + \boldsymbol{\eta}_f[k] \tag{3}$$

The input error $\boldsymbol{\eta}_f$ is modeled as zero-mean Gaussian discrete white noise with covariance matrix $\Sigma_{\eta f}(\boldsymbol{\theta}_\eta)$, where $\boldsymbol{\theta}_\eta$ is the parameter vector defining the covariance matrices of the input and output errors; this PDF maximizes the information entropy of the input error for a specified mean and covariance matrix. The components of $\mathbf{f}[k]$ that correspond to the unobserved excitation components of $\mathbf{g}[k]$ are set equal to zero. Thus, $\boldsymbol{\eta}_f$ models the input measurement noise for the observed components and it models the unobserved excitation for the unobserved components. The advantage of this formulation is that it can handle cases that include complete excitation measurements, incomplete excitation measurements and no excitation measurements (such as in ambient vibration tests).

2.3 Stochastic Output Model. Assume that discrete-time response data are available at N_o observed DOFs where the measured response $\mathbf{z}[k] \in \mathbb{R}^{N_o}$ is a linear combination of the model state vector $\mathbf{y}[k] = [\mathbf{x}[k]^T, \dot{\mathbf{x}}[k]^T]^T$ and the actual force $\mathbf{g}[k]$, plus an output error $\boldsymbol{\eta}_z[k] \in \mathbb{R}^{N_o}$ that accounts for modeling error and measurement noise in the response measurements. This output error is modeled as zero-mean Gaussian discrete white noise with covariance matrix $\Sigma_{\eta z}(\boldsymbol{\theta}_\eta)$. Thus, the measured response is given by

$$\begin{aligned}\mathbf{z}[k] &= \mathbf{L}_1\mathbf{y}[k] + \mathbf{L}_2\mathbf{g}[k] + \boldsymbol{\eta}_z[k] = \mathbf{L}_1\mathbf{y}[k] + \mathbf{L}_2\mathbf{f}[k] + \mathbf{L}_2\boldsymbol{\eta}_f[k] \\ &\quad + \boldsymbol{\eta}_z[k]\end{aligned} \tag{4}$$

where $\mathbf{L}_1 \in \mathbb{R}^{N_o \times 2N_d}$ and $\mathbf{L}_2 \in \mathbb{R}^{N_o \times N_g}$ are observation matrices that depend on the type of measurements (e.g., displacements or accelerations), and $\mathbf{y}[k]$ is given by Eq. (1). Furthermore, the errors $\boldsymbol{\eta}_z$ and $\boldsymbol{\eta}_f$ are modeled as stochastically independent.

2.4 Model and Damage Identification. The parameter vector $\boldsymbol{\theta}$ for identification from the excitation and response data is comprised of: (1) the mass, damping, and stiffness parameters $\boldsymbol{\theta}_m$, $\boldsymbol{\theta}_c$, and $\boldsymbol{\theta}_s$ that specify the mass matrix $\mathbf{M}$, damping matrix $\mathbf{C}$, and stiffness matrix $\mathbf{K}$; (2) the parameter vector $\boldsymbol{\theta}_\eta$ specifying the covariance matrices for the input and output errors $\boldsymbol{\eta}_f$ and $\boldsymbol{\eta}_z$, respectively; and (3) the $2N_d$ initial conditions for the structural state. In practice, the system may often be assumed to start from rest. In such a case, the initial conditions can be treated as known and equal to zero and can be excluded from the vector $\boldsymbol{\theta}$ of parameters for identification.

Let $\mathbf{Z}_{m,n}$ and $\mathbf{F}_{m,n}$ denote the response and the excitation measurement matrix from time $(m-1)\Delta t$ to $(n-1)\Delta t$, with $m \leq n$, respectively,

$$\mathbf{Z}_{m,n} = [\mathbf{z}[m], \ldots, \mathbf{z}[n]] \quad \text{and} \quad \mathbf{F}_{m,n} = [\mathbf{f}[m], \ldots, \mathbf{f}[n]], \quad m \leq n \tag{5}$$

The approach to damage detection is to first use the Bayesian framework presented in the next section to obtain the updated PDF (probability density function) $p(\boldsymbol{\theta}|\mathbf{Z}_{1,N}, \mathbf{F}_{1,N})$ of the parameter vector $\boldsymbol{\theta}$ given the measured input data $\mathbf{F}_{1,N}$ and output data $\mathbf{Z}_{1,N}$ where N denotes the total number of points in time where measurements are available. Then, this is used to compute the probability of damage of the jth substructure exceeding damage level d which is defined by

$$P_j^{\mathrm{dam}}(d) \equiv P\{\theta_j^{\mathrm{pd}} < (1-d)\,\theta_j^{\mathrm{ud}} | \mathbf{F}_{1,N}^{\mathrm{ud}}, \mathbf{Z}_{1,N}^{\mathrm{ud}}, \mathbf{F}_{1,N}^{\mathrm{pd}}, \mathbf{Z}_{1,N}^{\mathrm{pd}}\} \tag{6}$$

where subscripts "ud" and "pd" refer to undamaged and possibly damaged cases. Equation (6) gives the probability that the substructure stiffness parameter has decreased by a fractional amount of more than d. Based on the Gaussian approximation of the updated PDFs for θ_j^{ud} and θ_j^{pd}, one can easily calculate the probability damage as follows (Yuen et al. [26])

$$P_j^{\mathrm{dam}}(d) \approx \Phi\left(\frac{(1-d)\,\hat{\theta}_j^{\mathrm{ud}} - \hat{\theta}_j^{\mathrm{pd}}}{\sqrt{(1-d)^2(\hat{\sigma}_j^{\mathrm{ud}})^2 + (\hat{\sigma}_j^{\mathrm{pd}})^2}}\right) \tag{7}$$

where $\Phi(\cdot)$ is the standard Gaussian cumulative distribution function; $\hat{\theta}_j^{\mathrm{ud}}$ and $\hat{\theta}_j^{\mathrm{pd}}$ denote the most probable values of the stiffness parameters for the undamaged and (possibly) damaged structure, respectively; and $\hat{\sigma}_j^{\mathrm{ud}}$ and $\hat{\sigma}_j^{\mathrm{pd}}$ are the corresponding standard deviations of the stiffness parameters determined from the inverse of the Hessian matrix of the negative logarithm of the joint updated parameter PDF (Beck and Katafygiotis [22]).

3 Bayesian Model Updating

3.1 Exact Formulation of Updating. Using Bayes' theorem, the expression for the updated (posterior) PDF of the parameters $\boldsymbol{\theta}$ given the measured response $\mathbf{Z}_{1,N}$ and the measured input $\mathbf{F}_{1,N}$ is

$$p(\boldsymbol{\theta}|\mathbf{Z}_{1,N}, \mathbf{F}_{1,N}) = c_1 p(\boldsymbol{\theta}) p(\mathbf{Z}_{1,N}|\boldsymbol{\theta}, \mathbf{F}_{1,N}) \tag{8}$$

where c_1 is a normalizing constant such that the integral of the right hand side of Eq. (8) over the domain of $\boldsymbol{\theta}$ is equal to unity. We interpret $p(\boldsymbol{\theta}|\mathbf{Z}_{1,N}, \mathbf{F}_{1,N})$ as giving a measure of the plausibility of the values of $\boldsymbol{\theta}$ based on the data (Jaynes [27]). The factor $p(\boldsymbol{\theta})$ in Eq. (8) denotes the prior PDF of the parameters and it may be chosen based on engineering judgment. It may be treated as constant (noninformative prior) if all values of the parameters over some large but finite domain are felt to be equally plausible a priori. The likelihood $p(\mathbf{Z}_{1,N}|\boldsymbol{\theta}, \mathbf{F}_{1,N})$ is the dominant factor on the right hand side of Eq. (8). It reflects the contribution of the measured data $\mathbf{Z}_{1,N}$ and $\mathbf{F}_{1,N}$ in establishing the updated PDF of $\boldsymbol{\theta}$. Also, in order to establish the most probable (plausible) value of $\boldsymbol{\theta}$, denoted by $\hat{\boldsymbol{\theta}}$, one therefore needs to maximize $p(\boldsymbol{\theta})p(\mathbf{Z}_{1,N}|\boldsymbol{\theta}, \mathbf{F}_{1,N})$.

Since linear systems are considered and both the uncertain input and output measurement noise and unmeasured excitation are modeled as Gaussian, it follows that the likelihood $p(\mathbf{Z}_{1,N}|\boldsymbol{\theta}, \mathbf{F}_{1,N})$ in Eq. (8) is an N_oN-variate Gaussian distribution with appropriately calculated mean and covariance matrix. Direct calculation of this function for different values of $\boldsymbol{\theta}$ becomes computationally prohibitive for a large number N of data, as it requires repeated calculation of the determinant and inverse of the corresponding very high-dimensional $N_oN \times N_oN$ covariance matrices. Thus, although Eq. (8) offers a theoretically exact solution to the model updating problem, its computational implementation poses a challenge. In the next section, an approximation is presented which overcomes this difficulty and renders the Bayesian model updating problem computationally feasible.

3.2 Proposed Approximation for the Likelihood. The PDF $p(\mathbf{Z}_{1,N}|\boldsymbol{\theta}, \mathbf{F}_{1,N})$ in Eq. (8) can be written as a product of conditional PDFs

$$p(\mathbf{Z}_{1,N}|\boldsymbol{\theta}, \mathbf{F}_{1,N}) = p(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta}, \mathbf{F}_{1,N}) \prod_{k=N_p+1}^{N} p(\mathbf{z}[k]|\boldsymbol{\theta}, \mathbf{Z}_{1,k-1}, \mathbf{F}_{1,N}) \tag{9}$$

The following approximation is introduced [25]

$$p(\mathbf{Z}_{1,N}|\boldsymbol{\theta},\mathbf{F}_{1,N}) \approx p(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N}) \prod_{k=N_p+1}^{N} p(\mathbf{z}[k]|\boldsymbol{\theta},\mathbf{Z}_{k-N_p,k-1},\mathbf{F}_{1,N}) \tag{10}$$

Here, each conditional PDF factor depending on all the previous response measurements is approximated by a conditional PDF depending on only the most recent N_p response measurements. The sense of this approximation is that response measurements too far in the past do not provide significant information about the present observed response. Of course, one expects this to be true if N_p is so large that all the correlation functions have decayed to very small values. However, it is found that a significantly smaller value of N_p suffices for the approximation in Eq. (10) to be valid for practical purposes. In particular, it is found that a value for N_p of the order of $T_1/\Delta t$ is sufficient, where T_1 is the fundamental period of the system and Δt is the sampling interval. For example, assuming $\Delta t = \frac{1}{25}T_1$, it follows that a value of $N_p = 25$ is sufficient. The advantage of the approximation in Eq. (10) will become obvious in the subsequent sections where the expressions for computing the factors on the right hand side of Eq. (10) are given. In Sec. 3.2.1, the expression for the first factor $p(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})$ in Eq. (10) is given. In Sec. 3.2.2, a general expression for the conditional PDF $p(\mathbf{z}[k]|\boldsymbol{\theta},\mathbf{Z}_{k-N_p,k-1},\mathbf{F}_{1,N})$ in Eq. (10) is derived. Based on these results, $p(\mathbf{Z}_{1,N}|\boldsymbol{\theta},\mathbf{F}_{1,N})$ can be computed efficiently from Eq. (10).

The most probable parameters $\hat{\boldsymbol{\theta}}$ can then be obtained by minimizing $J(\boldsymbol{\theta}) = -\ln[p(\boldsymbol{\theta})p(\mathbf{Z}_{1,N}|\boldsymbol{\theta},\mathbf{F}_{1,N})]$. Also, the updated PDF of the parameters $\boldsymbol{\theta}$ can be well approximated by a Gaussian distribution $\mathcal{N}(\hat{\boldsymbol{\theta}},\mathbf{H}(\hat{\boldsymbol{\theta}})^{-1})$ with mean $\boldsymbol{\theta}$ and covariance matrix $\mathbf{H}(\hat{\boldsymbol{\theta}})^{-1}$, where $\mathbf{H}(\hat{\boldsymbol{\theta}})$ denotes the Hessian of $J(\boldsymbol{\theta})$ calculated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ (Beck and Katafygiotis [22]).

3.2.1 Expression for $p(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})$. Since the joint PDF $p(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})$ follows an N_oN_p-variate Gaussian distribution, it is specified by the mean and covariance matrix of $\mathbf{Z}_{1,N}$. Expressions for the mean and covariance are derived as a function of the identification parameter vector $\boldsymbol{\theta}$ as follows.

The equation of motion (1) can be rewritten in a state space form for the structural state vector $\mathbf{y} \equiv [\mathbf{x}^T,\dot{\mathbf{x}}^T]^T$

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{g} \tag{11}$$

where the system matrices $\mathbf{A} \in \mathbb{R}^{2N_d \times 2N_d}$ and $\mathbf{B} \in \mathbb{R}^{2N_d \times N_g}$ are given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{N_d} & \mathbf{I}_{N_d} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix} \tag{12}$$

and $\mathbf{B} = \begin{bmatrix} 0_{N_d} \\ \mathbf{M}^{-1} \end{bmatrix}\mathbf{T}$. Here, 0_{N_d} and $\mathbf{I}_{N_d}$ denote the $N_d \times N_d$ zero and identity matrix, respectively.

The continuous-time differential Eq. (11) is approximated by the following difference equation

$$\mathbf{y}[k+1] = \mathbf{A}_d\mathbf{y}[k] + \mathbf{B}_d\mathbf{g}[k] \tag{13}$$

where $\mathbf{y}[k]$ denotes the structural state vector at time $t_k = (k-1)\Delta t$, $\mathbf{A}_d \equiv e^{\mathbf{A}\Delta t}$ and $\mathbf{B}_d \equiv \mathbf{A}^{-1}(\mathbf{A}_d - \mathbf{I}_{2N_d})\mathbf{B}$. For notation convenience, denote the relationship between the state vector and the input of the above system using the function $\mathcal{L}$

$$\mathbf{y}[k] \equiv \mathcal{L}(k;\boldsymbol{\theta},\mathbf{G}_{1,N}), \quad k \leq N \tag{14}$$

where $\boldsymbol{\theta}$ is the vector comprised of the model parameters for identification described earlier and $\mathbf{G}_{1,N}$ denotes, in analogy to the definition of Eq. (5), the matrix comprised of the actual input force time history up to time $(N-1)\Delta t$, i.e., $\mathbf{G}_{1,N} = [\mathbf{g}[1],\mathbf{g}[2],\ldots,\mathbf{g}[N]]$.

It can be easily shown using Eq. (4) that the mean $\boldsymbol{\mu}[k] \equiv E[\mathbf{z}[k]|\boldsymbol{\theta},\mathbf{F}_{1,N}]$ is given by

$$\boldsymbol{\mu}[k] = \mathbf{L}_1\mathcal{L}(k;\boldsymbol{\theta},\mathbf{F}_{1,N}) + \mathbf{L}_2\mathbf{f}[k] \tag{15}$$

Thus, $\boldsymbol{\mu}[k]$ is equal to the model response calculated assuming that the only input is the measured excitation. The difference between $\mathbf{z}[k]$ and $\boldsymbol{\mu}[k]$ is the prediction error $\boldsymbol{v}[k]$

$$\boldsymbol{v}[k] = \mathbf{z}[k] - \boldsymbol{\mu}[k] \tag{16}$$

It is worth noting that $\boldsymbol{\mu}[k]$ in Eq. (15) can be simply calculated using the function "lsim" in MATLAB [28]. Collecting all the terms calculated by Eq. (15), $E(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})$ is given by

$$E(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N}) = [\boldsymbol{\mu}[1]^T,\ldots,\boldsymbol{\mu}[N_p]^T]^T \tag{17}$$

The covariance matrix $\Sigma_{Z,N_p} \equiv E[[\mathbf{Z}_{1,N_p} - E(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})] \times [\mathbf{Z}_{1,N_p} - E(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})]^T]$ is given by

$$\Sigma_{Z,N_p} = \begin{bmatrix} \Sigma_v[1,1] & \cdots & \Sigma_v[1,N_p] \\ & \ddots & \vdots \\ \text{sym} & & \Sigma_v[N_p,N_p] \end{bmatrix} \tag{18}$$

where $\Sigma_v[m,n]$, $m \leq n$, can be approximated by (see Appendix A)

$$\Sigma_v[m,n] \approx \mathbf{L}_1\mathbf{S}_\infty(\mathbf{A}_d^T)^{n-m}\mathbf{L}_1^T + \mathbf{L}_2\Sigma_{\eta f}\mathbf{B}_d^T(\mathbf{A}_d^T)^{n-m-1}\mathbf{L}_1^T(1-\delta_{m,n}) + (\mathbf{L}_2\Sigma_{\eta f}\mathbf{L}_2^T + \Sigma_{\eta z})\delta_{m,n} \tag{19}$$

where $\delta_{m,n}$ is the Kronecker delta and the matrix $\mathbf{S}_\infty$ can be obtained by solving the Lyapunov equation in discrete form (Lin [29])

$$\mathbf{S}_\infty = \mathbf{A}_d\mathbf{S}_\infty\mathbf{A}_d^T + \mathbf{B}_d\Sigma_{\eta f}\mathbf{B}_d^T \tag{20}$$

Furthermore, $\Sigma_v[n,m] = \Sigma_v[m,n]^T$, $m \leq n$, defines the elements in the lower triangle.

The joint PDF $p(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})$ is then the N_oN_p-variate Gaussian distribution

$$p(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N}) = \frac{1}{(2\pi)^{N_oN_p/2}|\Sigma_{Z,N_p}|^{1/2}} \times \exp\left\{-\frac{1}{2}[\mathbf{Z}_{1,N_p} - E(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})]^T \times \Sigma_{Z,N_p}^{-1}[\mathbf{Z}_{1,N_p} - E(\mathbf{Z}_{1,N_p}|\boldsymbol{\theta},\mathbf{F}_{1,N})]\right\} \tag{21}$$

3.2.2 Expression for $p(\mathbf{z}[k]|\boldsymbol{\theta},\mathbf{Z}_{k-N_p,k-1},\mathbf{F}_{1,N})$. Define the vector $\mathbf{W}[k]$, $k > N_p$, as follows: $\mathbf{W}[k] = [\mathbf{z}[k]^T,\mathbf{z}[k-1]^T,\ldots,\mathbf{z}[k-N_p]^T]^T$, which is comprised of all the response measurements appearing in $p(\mathbf{z}[k]|\boldsymbol{\theta},\mathbf{Z}_{k-N_p,k-1},\mathbf{F}_{1,N})$. Specifically, $\mathbf{W}[k]$ consists of $\mathbf{z}[k]$ followed by all vector elements of $\mathbf{Z}_{k-N_p,k-1}$ ordered in a descending time index order. Next, the expressions for the mean value and the covariance of the Gaussian joint PDF $p(\mathbf{W}[k]|\boldsymbol{\theta},\mathbf{F}_{1,N})$ are derived.

Clearly, the expected value of the vector $\mathbf{W}[k]$ given $\boldsymbol{\theta}$ and $\mathbf{F}_{1,N}$ is given by

$$E[\mathbf{W}[k]|\boldsymbol{\theta},\mathbf{F}_{1,N}] = [\boldsymbol{\mu}[k]^T,\boldsymbol{\mu}[k-1]^T,\ldots,\boldsymbol{\mu}[k-N_p]^T]^T \tag{22}$$

where $\boldsymbol{\mu}[k]$ is given by Eq. (15). The covariance matrix $\Sigma_W[k] = E\{[\mathbf{W}[k] - E(\mathbf{W}[k]|\boldsymbol{\theta},\mathbf{F}_{1,N})][\mathbf{W}[k] - E(\mathbf{W}[k]|\boldsymbol{\theta},\mathbf{F}_{1,N})]^T\}$ given $\mathbf{F}_{1,N}$ is easily shown to be

$$\Sigma_W[k] = \begin{bmatrix} \Sigma_v[k,k] & & \text{sym} \\ \vdots & \ddots & \\ \Sigma_v[k-N_p,k] & \cdots & \Sigma_v[k-N_p,k-N_p] \end{bmatrix} \tag{23}$$

where $\Sigma_v[m,n]$, $m \leq n$ is given by Eq. (19) and $\Sigma_v[n,m] = \Sigma_v[m,n]^T$, $m \leq n$. Therefore, the joint PDF $p(\mathbf{W}[k]|\boldsymbol{\theta},\mathbf{F}_{1,N})$, $N_p+1 \leq k \leq N$, is an $N_o(N_p+1)$-variate Gaussian distribution with mean given by Eq. (22) and covariance matrix $\Sigma_W[k]$ given by Eq. (23) which is independent of k when the approximation for

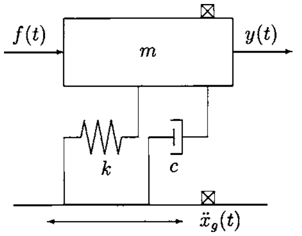


Fig. 1 Single-degree-of-freedom oscillator model (Example 1)

$\Sigma_v[m,n]$ in Eq. (19) is used. Then, the matrix Σ_W is partitioned as follows:

$$\Sigma_W = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix} \tag{24}$$

where Σ_{11}, Σ_{12}, and Σ_{22} have dimensions $N_o \times N_o$, $N_o \times N_o N_p$, and $N_o N_p \times N_o N_p$, respectively.

The mean and covariance matrix for the N_o-variate Gaussian PDF $p(\mathbf{z}[k] | \boldsymbol{\theta}, \mathbf{Z}_{k-N_p,k-1}, \mathbf{F}_{1,N})$ can be determined from the corresponding mean and covariance matrix for $\mathbf{W}[k]$ given $\boldsymbol{\theta}$ and $\mathbf{F}_{1,N}$. The mean $\mathbf{e}[k] \equiv E[\mathbf{z}[k] | \boldsymbol{\theta}, \mathbf{Z}_{k-N_p,k-1}, \mathbf{F}_{1,N}]$ is given by

$$\mathbf{e}[k] = \boldsymbol{\mu}[k] + \Sigma_{12}\Sigma_{22}^{-1}[\boldsymbol{v}[k-1]^T, \boldsymbol{v}[k-2]^T, \ldots, \boldsymbol{v}[k-N_p]^T]^T \tag{25}$$

where $\boldsymbol{\mu}[k]$ is given by Eq. (15) and the prediction errors $\boldsymbol{v}[m]$, $m = k-N_p, \ldots, k-1$, are given by Eq. (16). The covariance matrix $\Sigma_{\epsilon,N_p}[k]$ of the error $\boldsymbol{\epsilon}[k] = \mathbf{z}[k] - \mathbf{e}[k]$ is given by

$$\Sigma_{\epsilon,N_p}[k] \equiv E[\boldsymbol{\epsilon}[k]\boldsymbol{\epsilon}[k]^T] = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^T \tag{26}$$

which does not depend on k when the approximation for $\Sigma_v[m,n]$ in Eq. (19) is used. Thus, the conditional PDF $p(\mathbf{z}[k] | \boldsymbol{\theta}, \mathbf{Z}_{k-N_p,k-1}, \mathbf{F}_{1,N})$ is given by the following Gaussian distribution

$$p(\mathbf{z}[k] | \boldsymbol{\theta}, \mathbf{Z}_{k-N_p,k-1}, \mathbf{F}_{1,N}) \approx \frac{1}{(2\pi)^{N_o/2}|\Sigma_{\epsilon,N_p}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{z}[k] - \mathbf{e}[k])^T \Sigma_{\epsilon,N_p}^{-1}(\mathbf{z}[k] - \mathbf{e}[k]) \right\} \tag{27}$$

where $\mathbf{e}[k]$ is given by Eq. (25) and Σ_{ϵ,N_p} is given by Eq. (26). It is of interest to note that this probability distribution is equivalent to taking an auto-regressive model of order N_p for the prediction error $v[k]$ in Eq. (16).

The advantage of the approximation introduced in Eq. (10) is that all the conditional PDFs on the right hand side of Eq. (10) are conditioned on exactly N_p previous response measurement points and follow an N_o-variate Gaussian distribution with approximately the same covariance matrix Σ_{ϵ,N_p} which, therefore, needs to be calculated only once for a given parameter set $\boldsymbol{\theta}$. Thus, to compute $p(\mathbf{Z}_{1,N} | \boldsymbol{\theta}, \mathbf{F}_{1,N})$, one needs to calculate the inverse and the determinant of only the matrices Σ_{Z,N_p}, Σ_{22}, and Σ_{ϵ,N_p}, of dimension $N_o N_p \times N_o N_p$, $N_o N_p \times N_o N_p$, and $N_o \times N_o$, respectively. This effort is much smaller than that required in an exact formulation where one needs to calculate the inverse and the determinant of a matrix of dimension $N_o N \times N_o N$, where $N \gg N_p$ in general.

4 Illustrative Examples

4.1 Example 1: SDOF Oscillator. Consider a SDOF oscillator of mass $m = 1$ kg subjected to external force $f(t)$ and base acceleration $\ddot{x}_g(t)$, as shown in Fig. 1. Here, $f(t)$ is white noise with spectral intensity $S_{f0} = 0.02$ N^2 s and the base acceleration is taken to be the 1940 El-Centro earthquake record in the N-S direction. The parameters $\tilde{\boldsymbol{\theta}} = [\tilde{k}, \tilde{c}, \tilde{\sigma}_{\eta f}, \tilde{\sigma}_{\eta z}]^T$ used to generate the simulated data are: $\tilde{k} = 100.0$ N/m, $\tilde{c} = 0.04$ N s/m (corresponding to damping ratio of 2.0%), $\tilde{S}_{f0} = 0.02$ N^2 s, $\tilde{\sigma}_{\eta f} = 2.5005$ N and $\tilde{\sigma}_{\eta z} = 0.0036$ m. The chosen value of $\tilde{\sigma}_{\eta f}$ corresponds to the standard deviation combining the unmeasured input f and 10% measurement noise of the measured input $\ddot{x}_g$. Also, the chosen value of $\tilde{\sigma}_{\eta z}$ corresponds to a 10% rms output-error level, i.e., the noise is 10% of the rms of the noise-free response. The sampling time step is taken to be 0.02 s and the total time interval is $T = 50$ s, about 80 fundamental periods, so that the number of data points is $N = 2500$.

Table 1 Identification results of the stiffness parameters of the oscillator (Example 1)

Parameter	Actual $\tilde{\theta}$	Optimal $\hat{\theta}$	S.D. σ	COV α	β
k	100.00	100.32	1.3490	0.014	0.23
c	0.4000	0.5423	0.1391	0.348	1.02
$\sigma_{\eta f}$	2.5005	2.4476	0.0939	0.038	0.56
$\sigma_{\eta z}$	0.0036	0.0035	0.0001	0.016	1.24

Table 1 refers to the identification results using a single set of displacement response measurements $\hat{\mathbf{Z}}_{1,N}$ and base acceleration measurements $\hat{\mathbf{F}}_{1,N}$. It shows the exact values of the parameters, the most probable values $\hat{\boldsymbol{\theta}} = [\hat{k}, \hat{c}, \hat{\sigma}_{\eta f}, \hat{\sigma}_{\eta z}]^T$, the calculated standard deviations for the Gaussian approximation of the joint PDF of k, c, $\sigma_{\eta f}$, and $\sigma_{\eta z}$, the coefficient of variation for each parameter and the value of a "normalized error" β for each parameter. The parameter β represents the absolute value of the difference between the identified value and exact value, normalized with respect to the corresponding calculated standard deviation. Here, the value $N_p = 30$ (corresponding to one period of the oscillator) was used in Eq. (10). Note that the order of the square matrices that need to be inverted by the proposed approach is $N_p = 30$ which is much smaller than $N = 2500$ in an exact formulation. Repeating the identification with a value of $N_p = 60$ yielded identical results to the accuracy shown, verifying that using $N_p = T_1/\Delta t$ is sufficient.

Figure 2 shows contours in the (k,c) plane of the marginal updated PDF $p(k,c | \hat{\mathbf{Z}}_{1,N}, \hat{\mathbf{F}}_{1,N})$ calculated for the set of simulated data used for Table 1. Figure 3 shows a comparison between the conditional PDFs $p(k | \hat{\mathbf{Z}}_{1,N}, \hat{\mathbf{F}}_{1,N}, \hat{c}, \hat{\sigma}_{\eta f}, \hat{\sigma}_{\eta z})$ and $p(c | \hat{\mathbf{Z}}_{1,N}, \hat{\mathbf{F}}_{1,N}, \hat{k}, \hat{\sigma}_{\eta f}, \hat{\sigma}_{\eta z})$, respectively, obtained from: (i) Eqs. (8) and (10) (crosses) and (ii) the Gaussian approximation $\mathcal{N}(\hat{\boldsymbol{\theta}}, \mathbf{H}(\hat{\boldsymbol{\theta}})^{-1})$ described in Sec. 3.2 (solid). It is seen that the proposed Gaussian approximation is very accurate. Thus, the inverse Hessian matrix $\mathbf{H}(\hat{\boldsymbol{\theta}})^{-1}$ can be used to calculate the covariance matrix for the uncertainty in the value of the parameter $\boldsymbol{\theta}$, given the data $\hat{\mathbf{Z}}_{1,N}$ and $\hat{\mathbf{F}}_{1,N}$. In particular, this gives the variance $\sigma^2(\theta_j | \hat{\mathbf{Z}}_{1,N}, \hat{\mathbf{F}}_{1,N})$ in Table 1 for each parameter θ_j of $\boldsymbol{\theta}$.

Another set of data is simulated with the same parameters except that the stiffness is reduced by 5%, i.e., $k = 95$ N/m, to simulate damage. Identification results are shown in Table 2. By using the posterior PDFs for the undamaged and damaged oscillator, the probability of damage with respect to the fractional damage level d can be obtained. Figure 4 shows the probability of damage for different threshold levels d. It can be seen that it is almost with probability 1 that there is stiffness reduction in the damaged case. Furthermore, this damage is likely to be within the range from 0% to 10%, with median 5.6% and standard deviation 1.7%. The proposed approach is capable of indicating such a small level of damage with only a small amount of response data and unmeasured excitation that contributes about 63% of the rms response.

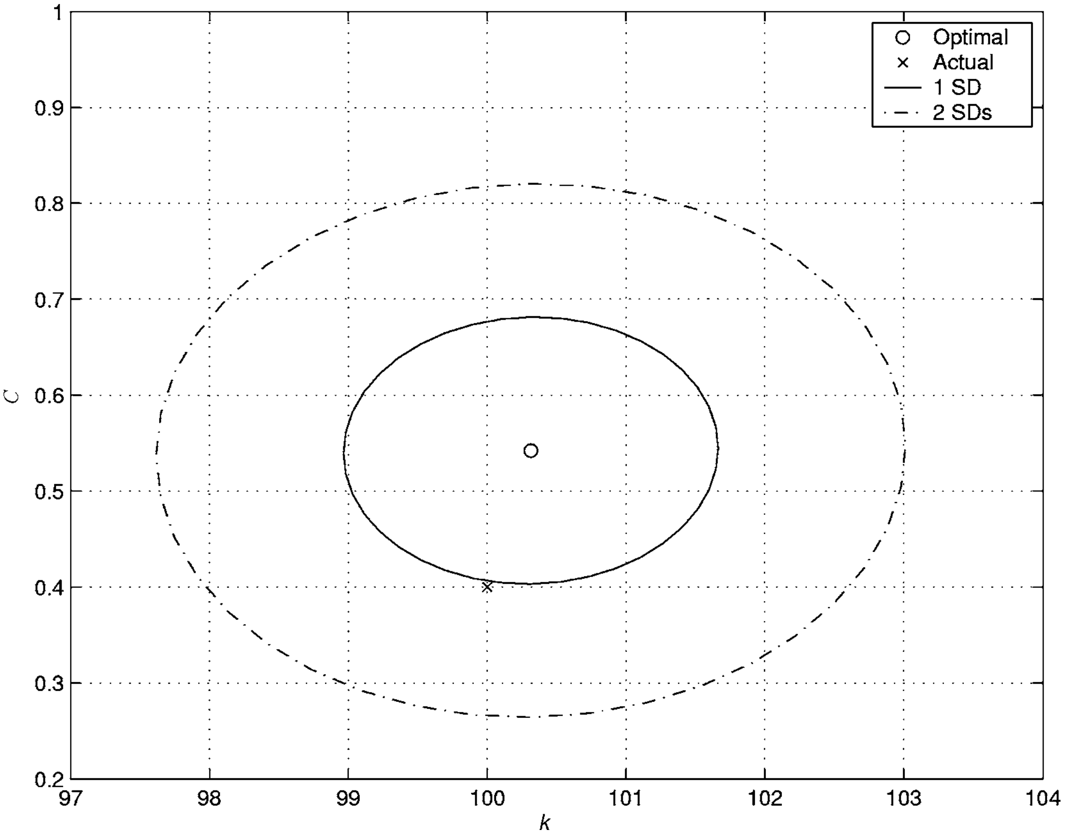


Fig. 2 Contours of the updated PDF projected onto the (k,c) plane of the undamaged oscillator (Example 1)

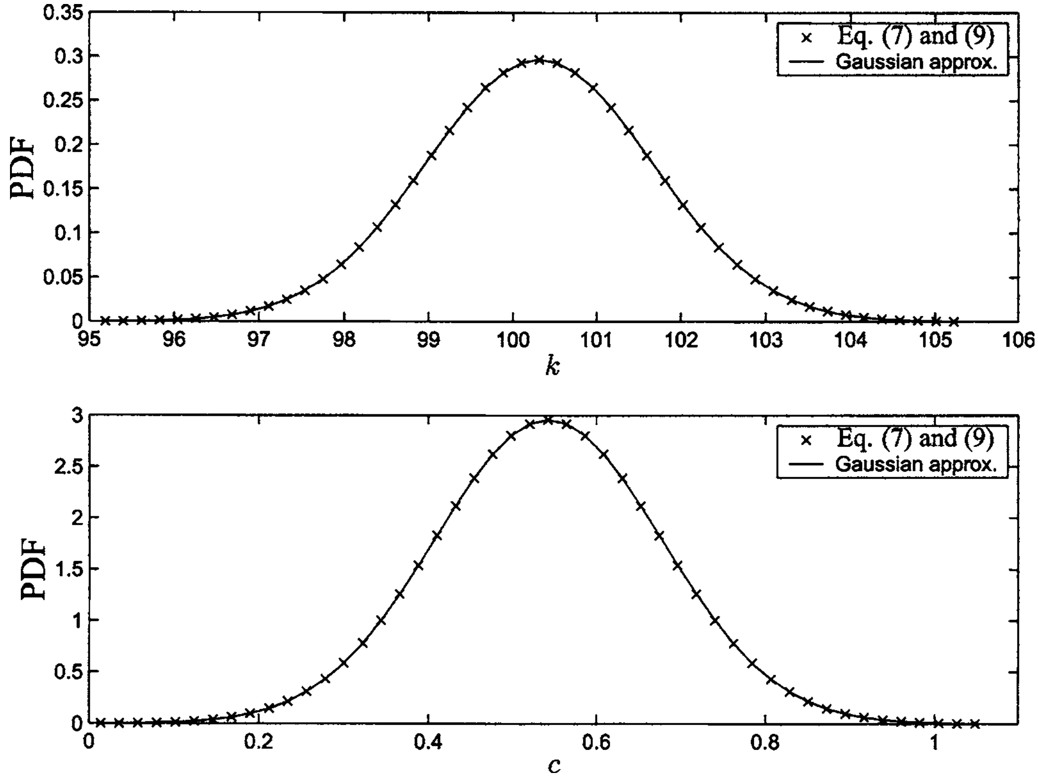


Fig. 3 Gaussian approximation for the conditional PDFs of the stiffness and damping coefficient of the undamaged oscillator (Example 1)

Table 2 Identification results of the stiffness parameters of the damaged oscillator (Example 1)

Parameter	Actual $\tilde{\theta}$	Optimal $\hat{\theta}$	S.D. σ	COV α	β
k	95.000	94.707	1.0805	0.011	0.27
c	0.4000	0.4099	0.1133	0.283	0.09
$\sigma_{\eta f}$	2.5005	2.5983	0.1059	0.042	0.93
$\sigma_{\eta z}$	0.0038	0.0039	0.0001	0.016	1.34

4.2 Example 2: Fifteen-Story Building Subjected to Earthquake and Wind Excitation. The second example uses simulated response data from the 15-story building shown in Fig. 5. The story height is 2.5 m. This building has uniformly distributed floor mass (100 ton each) and uniform story stiffness ($\tilde{k}_j=6.011\times 10^5$ kN/m, $j=1,2,\ldots,15$), so that the first four modal frequencies are 1.250, 3.737, 6.186, and 8.571 Hz, respectively. Rayleigh damping is chosen so the damping matrix is given by $\mathbf{C}=\alpha_M\mathbf{M}+\alpha_K\mathbf{K}$, where $\tilde{\alpha}_M=0.1177\ \mathrm{s}^{-1}$ and $\tilde{\alpha}_K=0.0006383$ s are used to simulate the data. As a result, the damping ratios of the first two modes are 1.0%.

For both undamaged and damaged cases, we assume that the measured response corresponds to the absolute acceleration at the 2nd, 5th, 8th, 11th and 14th DOF over a time interval $T=60$ s with a sampling interval $\Delta t=0.01$ s. Therefore, the total number of measured time points is $N=6000$ and corresponds to 48 fundamental periods.

The undamaged structure is subjected to stationary wind excitation (unmeasured) which has a uniform spectral intensity, $S_{f0}=5.0\ \mathrm{kN}^2$ s, at all DOFs and a correlation coefficient $\exp(-y/R)$, where y denotes the distance between two DOFs and R is a cor-

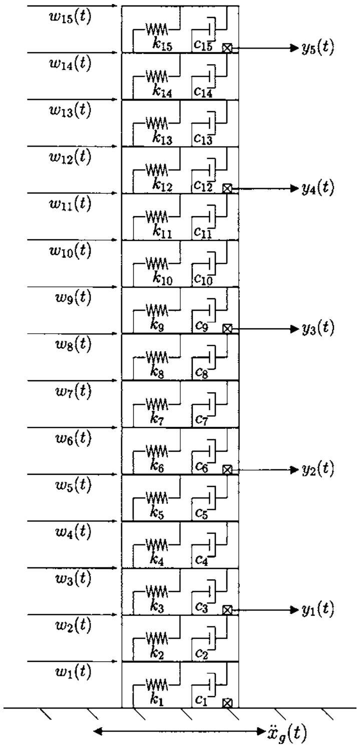


Fig. 5 Fifteen-story building model (Example 2)

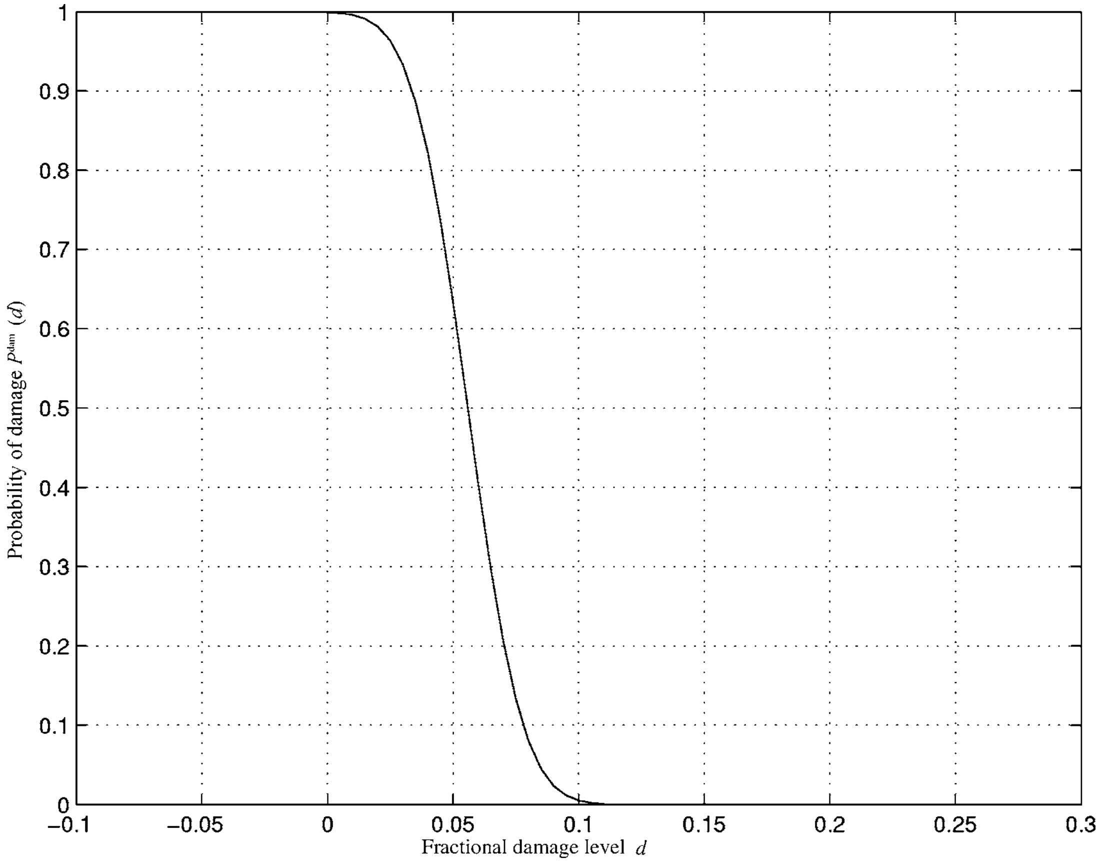


Fig. 4 Probability of damage for the stiffness (Example 1)

Table 3 Identification results of the undamaged structure (Example 2)

Parameter	Actual $\tilde{\theta}$	Optimal $\hat{\theta}$	S.D. σ	COV α	β
θ_1	1.0000	0.9978	0.0175	0.018	0.13
θ_2	1.0000	0.9909	0.0161	0.016	0.56
θ_3	1.0000	0.9939	0.0131	0.013	0.46
θ_4	1.0000	1.0071	0.0159	0.016	0.45
θ_5	1.0000	1.0224	0.0125	0.013	1.79
θ_6	1.0000	0.9849	0.0119	0.012	1.26
θ_7	1.0000	0.9848	0.0133	0.013	1.14
θ_8	1.0000	1.0228	0.0127	0.013	1.79
θ_9	1.0000	0.9779	0.0119	0.012	1.86
θ_{10}	1.0000	0.9987	0.0135	0.014	0.10
θ_{11}	1.0000	1.0076	0.0112	0.011	0.68
θ_{12}	1.0000	0.9812	0.0107	0.011	1.76
θ_{13}	1.0000	1.0083	0.0139	0.014	0.59
θ_{14}	1.0000	1.0187	0.0093	0.009	2.01
θ_{15}	1.0000	1.0047	0.0082	0.008	0.57
θ_{α_M}	1.0000	1.0062	0.4969	0.497	0.01
θ_{α_K}	1.0000	1.0165	0.0236	0.024	0.70
R	10.000	10.370	0.3753	0.038	0.99

relation distance, which is taken to be 10 m in the simulation of the data, but it is assumed unknown in the identification phase. The measurement noise for the response is taken to be 5%, i.e., the rms of the measurement noise for a particular channel of measurement is equal to 5% of the rms of the noise-free signal of the corresponding quantity. Identification using the proposed approach is carried out with a value of $N_p=100$, which corresponds to using previous data points of just over one fundamental period as the conditioning information at each time step in Eq. (10).

The stiffness and damping are based on the following nondimensional scaling parameters: stiffness parameters, $\theta_j, j=1,2,\ldots,15$ and damping parameters, θ_{α_M} and θ_{α_K}, that is, $k_j=\theta_j\tilde{k}_j$, $\alpha_M=\theta_{\alpha_M}\tilde{\alpha}_M$, and $\alpha_K=\theta_{\alpha_K}\tilde{\alpha}_K$. Table 3 shows the identification results for the undamaged structure. The second column in this table corresponds to the actual values of the parameters used for generation of the simulated measurement data; the third and fourth columns correspond to the most probable values and the calculated standard deviations, respectively; the fifth column lists the coefficient of variation for each parameter; and the last column shows the normalized error β described in Example 1. One observes that in all cases the actual parameters are at reasonable distances, measured in terms of the estimated standard deviations, from the most probable values, which confirms that the calculated uncertainties are consistent.

Figure 6 shows the contours in the (θ_1,θ_2) plane of the marginal updated PDF of θ_1 and θ_2. Figure 7 is a typical plot showing comparisons between the conditional PDFs of θ_1 and θ_2 (keeping all other parameters fixed at their most probable values) obtained from: (i) Eqs. (8) and (10) (crosses) and (ii) the Gaussian approximation $\mathcal{N}(\hat{\boldsymbol{\theta}},\mathbf{H}(\hat{\boldsymbol{\theta}})^{-1})$ described in Sec. 3.2 (solid). It is seen that the proposed Gaussian approximation is very accurate.

Next, damage is introduced by reducing the interstory stiffness of the first and third story by 15% and 10%, respectively. The damaged structure is subjected to wind excitation and earthquake ground motion. The wind excitation is assumed to have spectral intensity 2.5 kN^2 s with the same correlation model as before and the earthquake ground acceleration is taken to be equal to a 25% scaled version of the 1940 El-Centro earthquake N-S record.

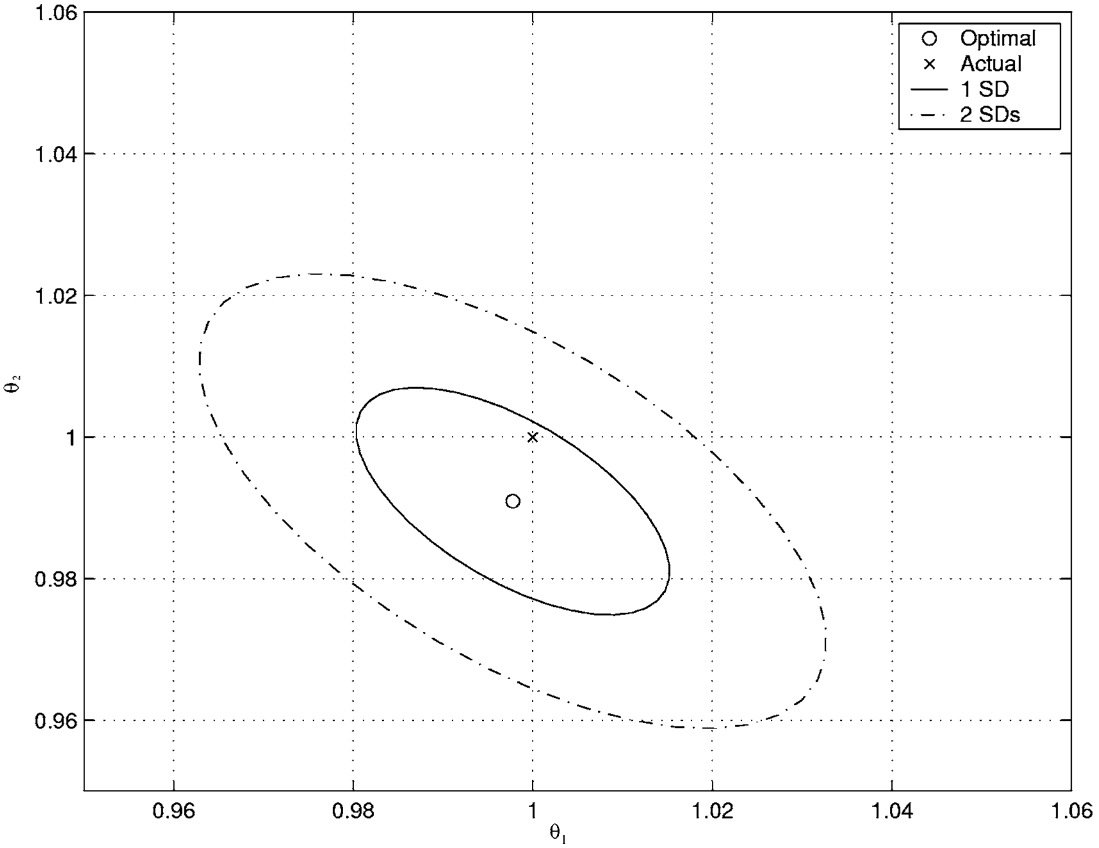


Fig. 6 Contours of the updated PDF projected onto the (θ_1,θ_2) plane of the undamaged structure (Example 2)

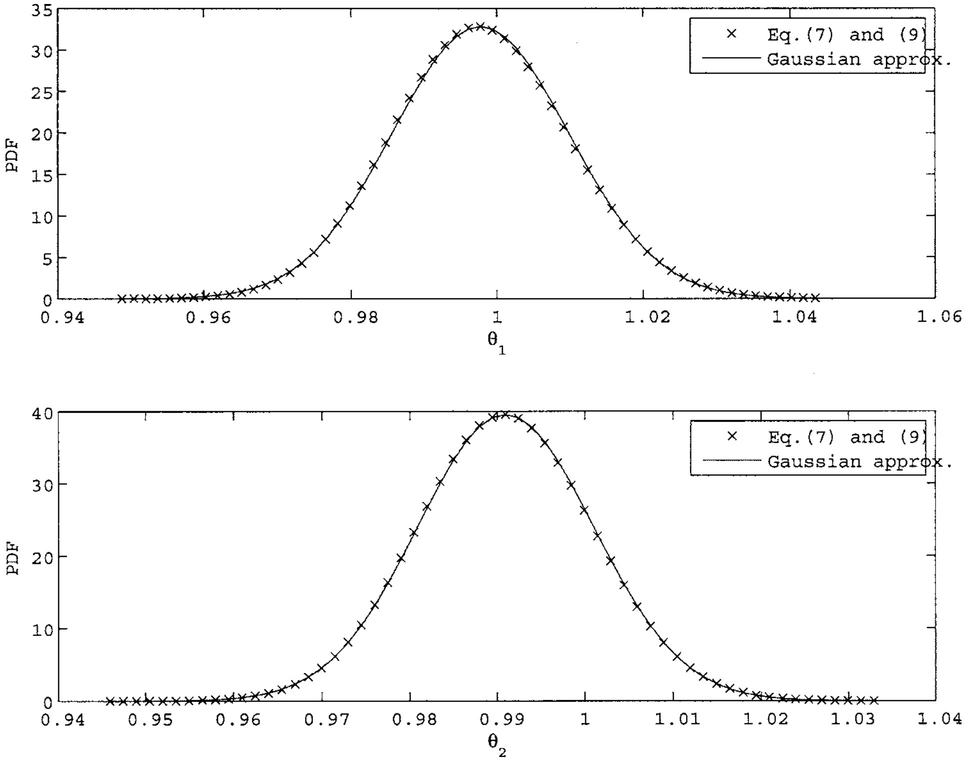


Fig. 7 Gaussian approximation for the conditional PDFs of the stiffness parameters θ_1 and θ_2 of the undamaged structure (Example 2)

Again, the wind excitation is assumed not to be measured but the earthquake ground motion is assumed to be measured with 5% measurement noise by a sensor at the base.

Figure 8 shows the displacement time histories at the first floor and the contribution from the earthquake only. Since the identification is based on acceleration, these data are assumed not to be available. It is shown here only for demonstration purposes. It can be seen that the earthquake ground motion dominates the response during the first 15 s but its contribution at later times is comparable with that from the wind excitation. If only the earthquake ground motion is considered, identification results will be poor, especially for the damping parameters, because the earthquake ground motion does not have much energy towards the end to explain the corresponding relatively strong response at these later times. Much smaller damping values, or even negative ones, will be identified in such case.

Identification results for the damaged structure are shown in Table 4. By using the posterior PDFs for the undamaged and damaged building, the probability of damage with respect to the fractional damage level can be obtained. Figure 9 shows the probability of damage with different threshold levels d. It can be seen that it is almost with probability 1 that there is stiffness reduction at the first and the third story. Furthermore, these damage levels have medians 14.8% and 10.5% and standard deviations 2.2% and 2.1%, respectively. The proposed approach is able to identify successfully both the location and severity of the damage. If a higher precision for the damage severity is desired, one solution is to obtain longer records of the structural excitation and response.

5 Concluding Remarks

A Bayesian approach to damage detection, location and assessment is presented using noisy incomplete excitation and response data. It is based on an approximate conditional probability density expansion of the updated PDF of the model parameters of a linear MDOF system using dynamic data. The updated posterior PDF can be accurately approximated by a multi-variate Gaussian distribution where the calculated mean and covariance matrix offer an estimate of the most probable values of the model parameters and their associated uncertainties. The updated PDFs from data in the undamaged state and in a possibly damaged state are used to calculate the probability of damage of different severity levels in each substructure. The approach was shown to successfully determine the location and probable level of damage from noisy incomplete data.

Acknowledgment

The first author would like to gratefully acknowledge the generous support by the Research Committee of University of Macau under Grant Nos. RG030/02-03S/YKV/FST and RG097/03-04S/YKV/FST.

Appendix A

Using Eqs. (4) and (13)–(15), the prediction error $\boldsymbol{v}[k]$ can be expressed as follows

$$\begin{aligned}\boldsymbol{v}[k] &= \mathbf{z}[k] - \boldsymbol{\mu}[k]\\ &= \mathbf{L}_1\mathcal{L}(k;\boldsymbol{\theta}^{\star},[\boldsymbol{\eta}_f[1],\ldots,\boldsymbol{\eta}_f[k-1]]) + \mathbf{L}_2\boldsymbol{\eta}_f[k] + \boldsymbol{\eta}_z[k]\\ &= \mathbf{L}_1\sum_{m=1}^{k-1}\mathbf{A}_d^{m-1}\mathbf{B}_d\boldsymbol{\eta}_f[k-m] + \mathbf{L}_2\boldsymbol{\eta}_f[k] + \boldsymbol{\eta}_z[k]\end{aligned} \tag{28}$$

where the parameter vector $\boldsymbol{\theta}^*$ has zero initial conditions and all other parameters are equal to the corresponding parameters in $\boldsymbol{\theta}$.

The covariance matrix $\Sigma_v[k,k+r] \equiv E[\boldsymbol{v}[k]\boldsymbol{v}^T[k+r]]$, $r \geqslant 0$, is given by

$$\begin{aligned}\Sigma_v[k,k+r] = \mathbf{L}_1\left[\sum_{m=1}^{k-1}\mathbf{A}_d^{m-1}\mathbf{B}_d\Sigma_{\eta f}\mathbf{B}_d^T(\mathbf{A}_d^T)^{m-1}\right](\mathbf{A}_d^T)^r\mathbf{L}_1^T\\ + \mathbf{L}_2\Sigma_{\eta f}\mathbf{B}_d^T(\mathbf{A}_d^T)^{r-1}\mathbf{L}_1^T(1-\delta_{r,0}) + (\mathbf{L}_2\Sigma_{\eta f}\mathbf{L}_2^T + \Sigma_{\eta z})\delta_{r,0}\end{aligned} \tag{29}$$

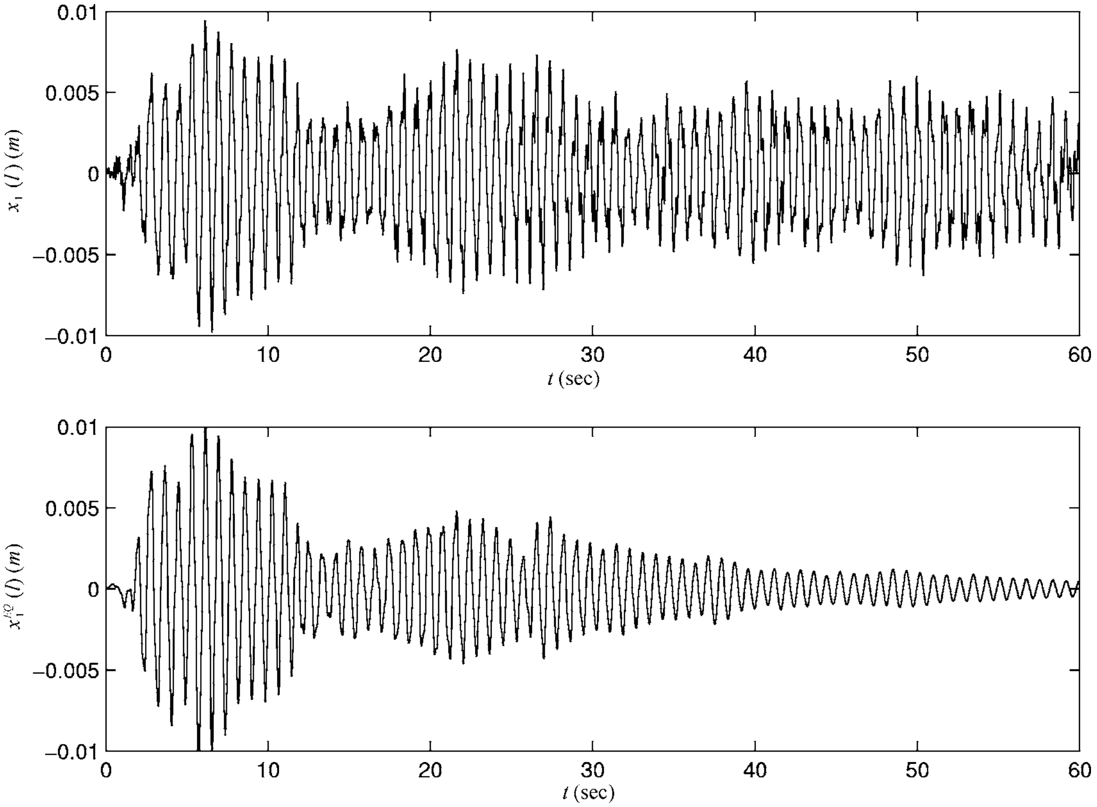


Fig. 8 Response time history (top) and its contribution from the earthquake only (bottom) at the first floor of the damaged structure (Example 2)

The sum $\mathbf{S}_k=\Sigma_{m=1}^{k-1}\mathbf{A}_d^{m-1}\mathbf{B}_d\Sigma_{\eta f}\mathbf{B}_d^T(\mathbf{A}_d^T)^{m-1}$ can be calculated by solving the following Lyapunov equation in discrete form [29]

$$\mathbf{S}_k=\mathbf{A}_d\mathbf{S}_k\mathbf{A}_d^T+[\mathbf{B}_d\Sigma_{\eta f}\mathbf{B}_d^T-\mathbf{A}_d^{k-1}\mathbf{B}_d\Sigma_{\eta f}\mathbf{B}_d^T(\mathbf{A}_d^T)^{k-1}] \tag{30}$$

For dissipative dynamical systems, the two-norm of the matrix $\mathbf{A}_d$ is less than unity, i.e., $\|\mathbf{A}_d\|_2<1$. As a result, the term $\mathbf{A}_d^{k-1}\mathbf{B}_d\Sigma_{\eta f}\mathbf{B}_d^T(\mathbf{A}_d^T)^{k-1}\rightarrow 0$ for large k. Therefore, Eq. (30) can be approximated by

$$\mathbf{S}_\infty=\mathbf{A}_d\mathbf{S}_\infty\mathbf{A}_d^T+\mathbf{B}_d\Sigma_{\eta f}\mathbf{B}_d^T \tag{31}$$

The advantage of this approximation is that the matrix $\mathbf{S}_\infty$ is no longer dependent on k, which improves the computational efficiency significantly.

Thus, the covariance matrix $\Sigma_v[k,k+r]$, $r\geqslant 0$, is readily obtained:

$$\Sigma_v[k,k+r]=\mathbf{L}_1\mathbf{S}_\infty(\mathbf{A}_d^T)^r\mathbf{L}_1^T+\mathbf{L}_2\Sigma_{\eta f}\mathbf{B}_d^T(\mathbf{A}_d^T)^{r-1}\mathbf{L}_1^T(1-\delta_{r,0}) \\ +(\mathbf{L}_2\Sigma_{\eta f}\mathbf{L}_2^T+\Sigma_{\eta z})\delta_{r,0} \tag{32}$$

Note that the right hand side of this expression does not depend on k.

Table 4 Identification results of the damaged structure (Example 2)

Parameter	Actual $\tilde{\theta}$	Optimal $\hat{\theta}$	S.D. σ	COV α	β
θ_1	0.8500	0.8495	0.0125	0.015	0.04
θ_2	1.0000	1.0103	0.0136	0.014	0.76
θ_3	0.9000	0.8887	0.0110	0.012	1.03
θ_4	1.0000	0.9938	0.0142	0.014	0.44
θ_5	1.0000	1.0074	0.0126	0.013	0.59
θ_6	1.0000	0.9758	0.0117	0.012	2.06
θ_7	1.0000	1.0214	0.0149	0.015	1.43
θ_8	1.0000	0.9979	0.0124	0.012	0.17
θ_9	1.0000	0.9759	0.0116	0.012	2.07
θ_{10}	1.0000	1.0293	0.0145	0.015	2.02
θ_{11}	1.0000	0.9971	0.0115	0.012	0.26
θ_{12}	1.0000	0.9981	0.0108	0.011	0.18
θ_{13}	1.0000	1.0129	0.0140	0.014	0.92
θ_{14}	1.0000	1.0062	0.0098	0.010	0.63
θ_{15}	1.0000	1.0024	0.0081	0.008	0.29
θ_{α_M}	1.0000	1.0149	0.2823	0.282	0.05
θ_{α_K}	1.0000	1.0248	0.0234	0.023	1.06
R	10.000	10.007	0.3649	0.037	0.02

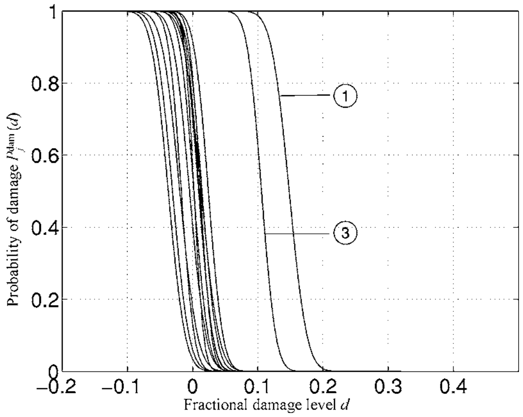


Fig. 9 Probability of damage for the stiffness parameters θ_j, $j=1,\ldots,15$ (Example 2)

References

[1] Natke, H. G., and Yao, J. T. P., eds., 1988, *Proceedings of the Workshop on Structural Safety Evaluation Based on System Identification Approaches*, Vieweg, Wiesbaden.

[2] Agbabian, M. S., and Masri, S. F., eds., 1988, *Proceedings of the International Workshop on Nondestructive Evaluation for Performance of Civil Structures*. Los Angeles, CA: Department of Civil Engineering, University of Southern California.

[3] Chang, F. K., ed., 2003, *Proceedings of the 4th International Workshop on Structural Health Monitoring*, Stanford University.

[4] Mazurek, D. F., and DeWolf, J. T., 1990, "Experimental Study of Bridge Monitoring Technique," J. Struct. Eng., **116**(9), pp. 2532–2549.

[5] Hearn, G., and Testa, R. B., 1991, "Modal Analysis for Damage Detection in Structures," J. Struct. Eng., **117**(10), pp. 3042–3063.

[6] Doebling, S. W., Farrar, C. R., Prime, M. B., and Shevitz, D. W., 1996, "Damage Identification and Health Monitoring of Structural and Mechanical Systems From Changes in Their Vibrations Characteristics: A Literature Review," Technical Report, LA-13070-MS, Los Alamos National Laboratory.

[7] Lam, H. F., Ko, J. M., and Wong, C. W., 1998, "Localization of Damaged Structural Connections Based on Experimental Modal and Sensitivity Analysis," J. Sound Vib., **210**(1), pp. 91–115.

[8] Smyth, A. W., Masri, S. F., Caughey, T. K., and Hunter, N. F., 2000, "Surveillance of Intricate Mechanical Systems on the Basis of Vibration Signature Analysis," J. Appl. Mech., **67**(3), pp. 540–551.

[9] Farhat, C., and Hemez, F. M., 1993, "Updating Finite Element Dynamics Models Using Element-by-Element Sensitivity Methodology," AIAA J., **31**(9), pp. 1702–1711.

[10] Pandey, A. K., and Biswas, M., 1994, "Damage Detection in Structures Using Changes in Flexibility," J. Sound Vib., **169**, pp. 3–17.

[11] Kim, H. M., Bartkowicz, T. J., Smith, S. W., and Zimmerman, D., 1995, "Health Monitoring of Large Structures," J. Sound Vib., **29**(4), pp. 18–21.

[12] Topole, K. G., and Stubbs, N., 1995, "Non-Destructive Damage Evaluation in Complex Structures From a Minimum of Modal Parameters," Int. J. Anal. Exp. Modal Anal., **10**(2), pp. 95–103.

[13] Hemez, F. M., and Farhat, C., 1995, "Structural Damage Detection via a Finite Element Model Updating Methodology," Int. J. Anal. Exp. Modal Anal., **10**(3), pp. 152–166.

[14] Katafygiotis, L., S., Papadimitriou, C., and Lam, H. F., 1998, "A Probabilistic Approach to Structural Model Updating," Soil Dyn. Earthquake Eng., **17**(7–8), pp. 495–507.

[15] Doebling, S. W., Farrar, C. R., and Prime, M. B., 1998, "A Review of Damage Identification Methods That Examine Changes in Dynamics Properties," Shock Vib. Dig., **30**(2), pp. 91–105.

[16] Vanik, M. W., Beck, J. L., and Au, S. K., 2000, "Bayesian Probabilistic Approach to Structural Health Monitoring," J. Eng. Mech., **126**(7), pp. 738–745.

[17] Beck, J. L., Au, S. K., and Vanik, M. W., 2001, "Monitoring Structural Health Using a Probabilistic Measure," Comput. Aided Civ. Infrastruct. Eng., **16**, pp. 1–11.

[18] Sohn, H., and Farrar, C. R., 2001, "Damage Diagnosis Using Time Series Analysis of Vibration Signals," Smart Mater. Struct., **10**, pp. 1–6.

[19] Ko, J. M., Sun, Z. G., and Ni, Y. Q., 2002, "Multi-Stage Identification Scheme for Detecting Damage in Cable-Stayed Kap Shui Mun Bridge," Eng. Struct., **24**(7), pp. 857–868.

[20] Ching, J., and Beck, J. L., 2004, "New Bayesian Model Updating Algorithm Applied to a Structural Health Monitoring Benchmark," Struct. Health Monit., **3**, pp. 313–332.

[21] Geyskens, P., Der Kiureghian, A., and Monteiro, P., 1998, "Bayesian Prediction of Elastic Modulus of Concrete," J. Struct. Eng., **124**(1), pp. 89–95.

[22] Beck, J. L., and Katafygiotis, L. S., 1998, "Updating Models and Their Uncertainties. I: Bayesian Statistical Framework," J. Eng. Mech., **124**(4), pp. 455–461.

[23] Katafygiotis, L. S., and Yuen, K.-V., 2001, "Bayesian Spectral Density Approach for Modal Updating Using Ambient Data," Earthquake Eng. Struct. Dyn., **30**(8), pp. 1103–1123.

[24] Yuen, K.-V., 2002, "Ph.D. Thesis: Model Selection, Identification and Robust Control for Dynamical Systems," Technical Report No. EERL 2002–03, California Institute of Technology, Pasadena.

[25] Yuen, K.-V., and Katafygiotis, L. S., 2002, "Bayesian Modal Updating Using Complete Input and Incomplete Response Noisy Measurements," J. Eng. Mech., **128**(3), pp. 340–350.

[26] Yuen, K.-V., Au, S. K., and Beck, J. L., 2004, "Two-Stage Structural Health Monitoring Methodology and Results for Phase I Benchmark Studies," J. Eng. Mech., **130**(1), pp. 16–33.

[27] Jaynes, E. T., 2003, *Probability Theory: The Logic of Science*, L. Bretthorst (ed.), Cambridge University Press.

[28] MATLAB, 1994, *Matlab User's Guide*, The MathWorks, Inc., Natick, MA.

[29] Lin, Y. K., 1976, *Probabilistic Theory of Structural Dynamics*, Krieger, Malabar, FL.

Zheng-Dong Ma
e-mail: mazd@umich.edu

Noboru Kikuchi
e-mail: kikuchi@umich.edu

Department of Mechanical Engineering,
University of Michigan,
2250 G. G. Brown Building,
Ann Arbor, MI 48109-2125

Christophe Pierre
e-mail: pierre@umich.edu

Basavaraju Raju

U.S. Army Tank-Automotive
and Armaments Command,
Warren, MI 48347-5000

Multidomain Topology Optimization for Structural and Material Designs

A multidomain topology optimization technique (MDTO) is developed, which extends the standard topology optimization method to the realm of more realistic engineering design problems. The new technique enables the effective design of a complex engineering structure by allowing the designer to control the material distribution among the subdomains during the optimal design process, to use multiple materials or composite materials in the various subdomains of the structure, and to follow a desired pattern or tendency for the material distribution. A new algorithm of Sequential Approximate Optimization (SAO) is proposed for the multidomain topology optimization, which is an enhancement and a generalization of previous SAO algorithms (including Optimality Criteria and Convex Linearization methods, etc.). An advanced substructuring method using quasi-static modes is also introduced to condense the nodal variables associated with the multidomain topology optimization problem, especially for the nondesign subdomains. The effectiveness of the new MDTO approach is demonstrated for various design problems, including one of "structure-fixture simultaneous design," one of "functionally graded material design," and one of "crush energy management." These case studies demonstrate the potential significance of the new capability developed for a wide range of engineering design problems. [DOI: 10.1115/1.2164511]

Introduction

Topology optimization has received extensive attention in the literature in recent years, and a number of review papers and books have been published on this topic [1–5]. Significant progress has been made through the use of a variety of approaches (also see [1–5]). Among various proposed approaches, a notable idea in modern topology optimization is to transform the topology optimization problem into an equivalent problem of "*optimum material distribution*" (OMD). For example, in the Homogenization Based Topology Optimization (HBTO) method [6] the material is distributed by considering specific "microstructures" in the design domain, and the structure is consequently optimized by varying the design variables associated with these "microstructures." The HBTO method has been extended to various applications areas, including structural design and material design [3–5]. It has also been applied to design problems for achieving static stiffness [6–8], mechanical compliance [9–11], desired eigenfrequencies [12–15], dynamic response characteristics [16,17], and other applications [3–5].

In the standard HBTO method, the structure is optimized within a single structural domain, subject to a given amount of the material for the entire structure. The optimization process determines the material distribution automatically, and there is no interaction with the designer. This process leaves little flexibility to the designer to control the material distribution in a way that may be desired. For example, based on previous experience, a designer may want to distribute more material in a certain region of the structure and less material in another region, etc. In general, the engineering design of a complex structure, such as an automotive vehicle structure, would benefit from having the designer hold more control over the material distribution, so that the final design can reflect valuable designer intuition and experience. A number of important applications can be pointed out that would result directly from having this capability. For instance, in a "*structure-fixture simultaneous design*" problem, the designer may want to assign different amounts of the material (or different materials) to the "structure" domain and the "fixture" domain. In a "*functionally gradient material*" design problem, the designer may want to distribute the material in such a way that the material density is gradually reduced (or increased) along a certain direction. In a "*crush energy management*" design problem, more material may need to be assigned to a region where buckling is not allowed to occur, while less material may be placed where buckling is desired in a crush event. Furthermore, the intuition of a designer can be used to improve manufacturability of a prototype design by visually controlling material distribution to the subdomains. In this paper, it is demonstrated how the standard topology optimization method can be extended into a multidomain topology optimization (MDTO) technique, by allowing for the assignment of different amounts of the material, as well as of different materials, to the various subdomains of a structure.

The multidomain topology optimization developed in this paper employs a new optimization algorithm, which is based on a Generalized Sequential Approximate Optimization (GSAO) method developed by Ma and Kikuchi [18]. Here the GSAO algorithm is enhanced for the special class of topology optimization problems. The new algorithm extends the capabilities of the standard topology optimization technique significantly, by using advanced updating rules, providing additional flexibility to switch between different updating rules, and offering more appropriate and effective control parameters for the optimization process algorithm. These enhancements result in improved convergence, higher computational efficiency, and a more stabilized (smoothed) iterative process for both free (eigenvalue) and forced dynamic response problems. In addition to the new optimization algorithm, the MDTO approach also employs an advanced Component Mode Synthesis (CMS) method [19], in which a complex structure is decomposed into a number of simpler substructures that can be solved for easily and efficiently. The CMS method can signifi-

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 2, 2004; final manuscript received October 5, 2005. Review conducted by S. Govindjee. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

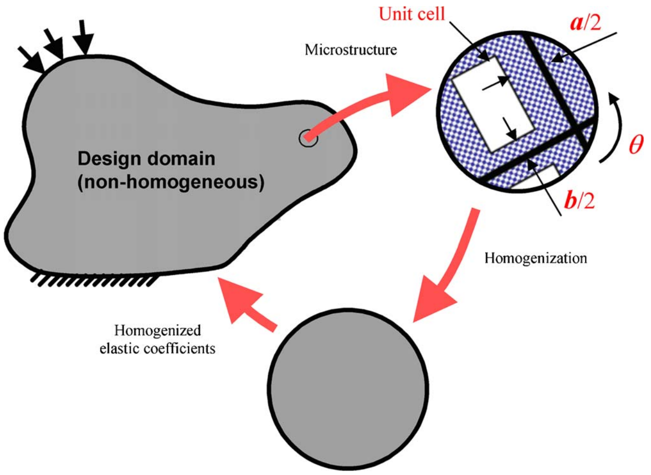


Fig. 1 Basic concept of the homogenization-based topology optimization method

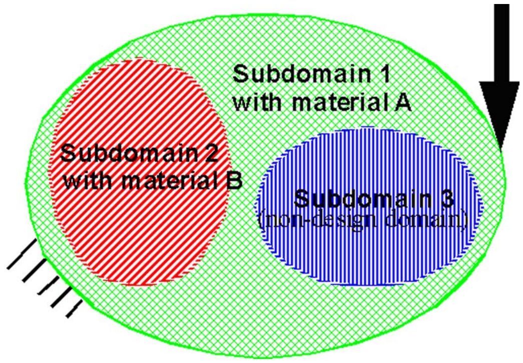


Fig. 2 A multidomain topology optimization problem

cantly extend the domain of solvability of topology optimization for practical engineering design problems, which is usually limited by the computer memory available as well as by the computational time and cost. The CMS approach makes it possible to deal with a structural model of very large size, since it yields a considerable reduction of the number of inactive degrees of freedom associated with the subdomains, especially for subdomains whose design is fixed at the current design or at the current iteration step. One further advantage of CMS is that structural reanalysis time is significantly reduced when design changes are limited to a few of the subdomains. The various examples presented in this paper illustrate how the new MDTO technique developed can be applied effectively to a wide range of topology optimization problems.

Topology Optimization

A technique for the topology optimization of structural systems was first developed by Bendsøe and Kikuchi in 1988 [6], and it has since become known as the homogenization-based topology optimization (HBTO) method. The basic idea in HBTO is to transform the optimal topology design problem into an equivalent Optimal Material Distribution (OMD) problem [6], as illustrated in Fig. 1. Here the structural domain is assumed to be filled with a nonhomogeneous composite material characterized by a variable microstructure. A typical microstructure is formed inside an empty rectangle in a unit cell with three design variables a, b, and θ, which are dimensions and orientation of the microstructure, respectively. Using homogenization, the effective material constants can be obtained as [17]:

$$\mathbf{D}^H = \mathbf{D}^H(a,b,\theta), \quad \rho^H = \rho^H(a,b) \tag{1}$$

where $\mathbf{D}^H$ denotes the effective elastic coefficient matrix at the material point, and ρ^H is the associated effective mass density. Both $\mathbf{D}^H$ and ρ^H are functions of the design variables a and b, and $\mathbf{D}^H$ is also a function of the orientation variable θ at the material point.

Using a standard finite element method, the state equation governing the dynamic response of the general structure shown in Fig. 1 can be obtained as:

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{f} \tag{2}$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the global mass, damping, and stiffness matrices of the structure, respectively. $\mathbf{M}$ and $\mathbf{K}$ can be obtained by assembling the elementary mass and stiffness matrices as

$$\mathbf{M} = \mathop{\mathrm{A}}_{e=1}^{n_{\mathrm{el}}} \mathbf{m}_e \text{ and } \mathbf{K} = \mathop{\mathrm{A}}_{e=1}^{n_{\mathrm{el}}} \mathbf{k}_e \tag{3}$$

where $\mathbf{m}_e=\mathbf{m}_e(a_e,b_e)$ and $\mathbf{k}_e=\mathbf{k}_e(a_e,b_e,\theta_e)$ are the mass and stiffness matrices of the finite element e, which are functions of a_e, b_e, and θ_e, where a_e, b_e, and θ_e are the discretized design variables of a, b, and θ, respectively, at the elementary level ($e=1,2,\ldots,n_{\mathrm{el}}$). In the special case of proportional damping, the viscous damping matrix $\mathbf{C}$ in Eq. (2) can be written as $\mathbf{C}=\beta\mathbf{K}$, where β is the damping coefficient.

Multidomain Topology Optimization

Multidomain topology optimization (MDTO) can be considered as a natural generalization of the standard, single-domain topology optimization discussed in the previous section. In contrast to single-domain optimization, in which a given amount of the material is assigned to the entire design domain, MDTO allows the designer to assign different amounts of the material, or even different materials, to the various subdomains of the structure. For example, Fig. 2 depicts a structural domain that is divided into three subdomains, where a certain amount of material A is assigned to subdomain 1, a different amount of material B is distributed into subdomain 2, and subdomain 3 is considered as a nondesign domain, for which the material distribution is not allowed to change at the current design stage. In general, a MDTO problem is defined as a topology optimization problem that has multiple (design and nondesign) domains; each design domain can be assigned a special amount of material (or material density), with option to distribute multiple materials in different design domains or in the same design domain. The objective function of the MDTO problem can be related to the performance/functionality of the whole structure or to that of a single domain or of multiple domains considered in the design problem.

The MDTO problem discussed in this paper is formulated as

$$\begin{gathered}\text{Minimize } f(\mathbf{X})\\ \text{Subject to: } h_j(\mathbf{X}) \leq 0 \; (j=1,2,\ldots,m)\\ \underline{x}_i \leq x_i \leq \bar{x}_i \; (i=1,2,\ldots,n)\\ \text{(and state equations)}\end{gathered} \tag{4}$$

where $f=f(\mathbf{X})$ denotes the objective function, which can be the compliance (or static stiffness) [6], the mean eigenvalue [13–15], the frequency response [16,17], the mutual energy [9–11], or a combination of the aforementioned objectives, or any others; $h_j=h_j(\mathbf{X})$ denotes the jth constraint function for the volume (or weight) of the jth substructure in the jth subdomain (where $j=1,2,\ldots,m$); m is the total number of the subdomains; $\mathbf{X}=\{x_1,x_2,\ldots,x_n\}^T$ denotes the vector of the design variables, which includes a_e, b_e, and θ_e ($e=1,2,\ldots,n_{\mathrm{el}}$) defined previously, where n is the total number of the design variables; and $\underline{x}_i$ and $\bar{x}_i$ are the lower and upper bounds of design variable x_i, respectively. Note that $f(\mathbf{X})$ in Eq. (4) also needs to satisfy the state equations for the structural problem at hand. These state equations may include, for example, the static equilibrium equation, the equation that defines the free vibration eigenvalue problem, or the equation for the dynamic forced response. Other additional constraint functions can also be considered in Eq. (4). Also note that since orientation variables θ_e $(e=1,2,\ldots,n_{\mathrm{el}})$ can be determined separately using either principal stress direction or an analytical updating equation in [17], in following discussions, we assume that design variable vector $\mathbf{X}$ in Eq. (4) contains only a_e's and b_e's.

Various optimization algorithms can be used to solve the optimization problem, Eq. (4). In this paper, we introduce a new algorithm based on the GSAO method developed by Ma and Kikuchi [18]. This algorithm can be considered as an extension of the algorithm used in the earlier stage of the topology optimization by Bendsøe and Kikuchi [6], and it can also be regarded as an extension of various popular optimization algorithms developed for structural optimization problems of large size. The new algorithm is described in the following section.

Optimization Algorithm

Topology optimization problems usually involve very large numbers of design variables, and thus require a highly efficient optimization algorithm. Because traditional mathematical programming methods are not practical for dealing with such large numbers of design variables, optimality criteria (OC) [20] and sequential approximate optimization (SAO) [18] methods are typically employed. Here, we consider the OC method (e.g., that used by Bendsøe and Kikuchi [6]) as a special case of the SAO methods (see [18] for an explanation). Other SAO methods, including Sequential Linear Programming (SLP), Convex Linearization (CONLIN) developed by Fleury and Braibant [21], Method of Moving Asymptotes (MMA) developed by Svanberg [22], and Diagonal Sequential Quadratic Programming (DSQP) developed by Fleury [23], were also introduced for large-scale structural optimization problems. Ma and Kikuchi [18] developed a generalized SAO (GSAO) method based on a general convex approximation, which can produce a number of different optimization algorithms. Since the GSAO algorithm reduces in special cases to the aforementioned SAO algorithms, including OC, CONLIN, MMA, and DSQP, it can be considered as a generalization of these algorithms. Note that while the work of Ma and Kikuchi [18] provided a general formula for the GSAO method, the current paper proposes a practical and executable algorithm that is based on the general method but is specifically tailored to the MDTO problems dealt with in this paper. Thus Ref. [18] focused on the foundation and general theory of the new method, and the current paper centers on the selection and practical use of a suitable algorithm.

The basic concept of the GSAO method is to approximate the original optimization problem, Eq. (4), as a sequence of simplified problems. At each iteration, an approximate optimization problem is solved by the dual method to obtain an approximation of the optimum for the original problem. This approximate optimum is then used as the starting point for the next iteration, and this iterative process is continued until a specified criterion is satisfied. The GSAO method approximates the optimization problem by linearizing the objective and constraint functions via a general intermediate variable, which can be written as

$$y_i = |x_i - c_i|^{\xi_i} \ (i=1,2,\ldots,n)$$

and

$$z_{ji} = |x_i - e_{ji}|^{\zeta_{ji}} \ (i=1,2,\ldots,n \text{ and } j=1,2,\ldots,m) \tag{5}$$

where y_i $(i=1,2,\ldots,n)$ are used to linearize the objective function ($f(\mathbf{X})$ in Eq. (4)), z_{ji} $(i=1,2,\ldots,n;\ j=1,2,\ldots,m)$ are used to linearize the constraint functions ($h_j(\mathbf{X})$ $(j=1,2,\ldots,m)$ in Eq. (4)), and ξ_i, ζ_{ji}, c_i, and e_{ji} are parameters that control the performance of the optimization algorithm. Different selections for these parameters will result in different updating rules for the GSAO algorithm. Existing popular optimization algorithms, such as OC, CONLIN, MMA, and DSQP, are special cases of the GSAO algorithm obtained by selecting special sets of the parameters in Eq. (5); this will be discussed later in the paper.

Linearizing the objective function and constraint functions in Eq. (4) via the intermediate variables defined in Eq. (5), one obtains a sequence of approximate optimization problems for the original optimization problem, as follows:

$$\text{Minimize } f_0^k + \sum_{i=1}^{n} a_i^k |x_i - c_i|^{\xi_i}$$

$$\text{Subject to } h_{0j}^k + \sum_{i=1}^{n} b_{ji}^k |x_i - e_{ji}|^{\zeta_{ji}} \leqslant 0 \ (j=1,2,\ldots,m)$$

$$\underline{x}_i \leqslant x_i \leqslant \bar{x}_i \ (i=1,2,\ldots,n) \tag{6}$$

where k $(k=1,2,\ldots)$ denotes the iteration number,

$$f_0^k = f(\mathbf{X}^k) - \sum_{i=1}^{n} a_i^k |x_i^k - c_i|^{\xi_i}$$

$$h_{0j}^k = h_j(\mathbf{X}^k) - \sum_{i=1}^{n} b_{ji}^k |x_i^k - e_{ji}|^{\zeta_{ji}}$$

$$a_i^k = \frac{1}{\xi_i}\,\mathrm{sign}\{x_i^k - c_i\}|x_i^k - c_i|^{1-\xi_i} f_{,x_i}^k$$

$$b_{ji}^k = \frac{1}{\zeta_{ji}}\,\mathrm{sign}\{x_i^k - e_{ji}\}|x_i^k - e_{ji}|^{1-\zeta_{ji}} h_{j,x_i}^k \tag{7}$$

where x_i^k $(i=1,2,\ldots,n;\ k=1,2,\ldots)$ are the design variables obtained at the previous, or $(k-1)$th iteration, (for $k=1$, x_i^1 $(i=1,2,\ldots,n)$ are the initial design variables), and

$$f_{,x_i}^k = \left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{X}=\mathbf{X}^k}$$

and

$$h_{j,x_i}^k = \left.\frac{\partial h_j}{\partial x_i}\right|_{\mathbf{X}=\mathbf{X}^k}.$$

By properly choosing the parameters in Eq. (5), the approximate optimization problem in Eq. (6) can be made to be always convex. It is then solved by using the dual method, and the dual problem can be written as

$$\underset{\boldsymbol{\lambda}}{\text{maximize}}\ L^k(\mathbf{X}^*(\boldsymbol{\lambda}),\boldsymbol{\lambda})$$

$$\text{subject to } \lambda_j > 0 \ (j=1,2,\ldots,m) \tag{8}$$

where, $\boldsymbol{\lambda}=\{\lambda_1,\lambda_2,\ldots,\lambda_m\}^T$ denotes the vector of the dual variables (which are also known as the Lagrange multipliers). Note that the dual optimization problem in Eq. (8) feature only very few (dual) design variables, whose number equals the number of constraints (which here is equal to the number of subdomains considered in the optimization problem, Eq. (4)). In Eq. (8), one has

$$L^k(\mathbf{X}^*(\boldsymbol{\lambda}),\boldsymbol{\lambda}) = f_0^k + \sum_{j=1}^{m} \lambda_j h_{0j}^k + \sum_{i=1}^{n} \left(a_i^k |x_i^* - c_i|^{\xi_i} + \sum_{j=1}^{m} \lambda_j b_{ji}^k |x_i^* - e_{ji}|^{\zeta_{ji}} \right) \tag{9}$$

where $\mathbf{X}^*(\boldsymbol{\lambda})=\{x_1^*(\boldsymbol{\lambda}),x_2^*(\boldsymbol{\lambda}),\ldots,x_n^*(\boldsymbol{\lambda})\}^T$ is a function of the dual variables, and can be obtained by solving the following equations:

$$\bar{a}_i^k \operatorname{sgn}\{x_i^* - c_i\}|x_i^* - c_i|^{\xi_i-1} + \sum_{j=1}^{m} \lambda_j \bar{b}_{ji}^k \operatorname{sgn}\{x_i^* - e_{ji}\}|x_i^* - e_{ji}|^{\zeta_{ji}-1} = 0$$

$$(\text{for } \underline{x}_i \leqslant x_i^* \leqslant \bar{x}_i, \text{ where } i = 1,2, \ldots ,n) \tag{10}$$

where $\bar{a}_i^k=\xi_i a_i^k$ and $\bar{b}_{ji}^k=\zeta_{ji} b_{ji}^k$. Note that $x_i^*=\underline{x}_i$ if $x_i^*\leqslant\underline{x}_i$, and $x_i^*=\bar{x}_i$ if $x_i^*\geqslant\bar{x}_i$.

Equation (10) is referred to as an updating equation. An explicit solution of Eq. (10) is called an updating rule, which in general can be written as

$$x_i^* = g_i(x_i^k) \ (i = 1,2, \ldots ,n) \tag{11}$$

where x_i^* denotes the updated value of the design variable x_i, and g_i is a function defined by the solution of Eq. (10).

Assuming $\boldsymbol{\lambda}^*$ is the optimum solution of the dual problem, Eq. (8), $\mathbf{X}=\mathbf{X}^*(\boldsymbol{\lambda}^*)$ gives an approximate optimum solution for the primary optimization problem, Eq. (4), at the kth step. By iterating, this approximate optimum solution will converge to the exact optimum of the primary optimization problem, provided the parameters in Eq. (5) are properly chosen. As mentioned previously, different SAO algorithms can be obtained by employing different intermediate variables in Eq. (5). As a special case, an extended CONLIN algorithm is obtained in this paper by assuming

$$\xi_i = \begin{cases} \xi_i^+ & \text{if } f_{,x_i}^k > 0 \\ \xi_i^- & \text{if } f_{,x_i}^k < 0 \end{cases}, \quad \zeta_{ji} = \zeta_i = \begin{cases} \xi_i^+ & \text{if } h_{j,x_i}^k > 0 \\ \xi_i^- & \text{if } h_{j,x_i}^k < 0 \end{cases}, \quad \text{and}$$

$$e_{ji} = c_i \ (j = 1,2, \ldots ,m) \tag{12}$$

where $\xi_i^- < 1 \leqslant \xi_i^+$. Then Eq. (11) becomes

$$x_i^* = c_i + w_i(x_i^k - c_i) \ (\text{for } \underline{x}_i \leqslant x_i^* \leqslant \bar{x}_i) \tag{13}$$

where

$$w_i = \begin{cases} \left(-\dfrac{q_i^-}{f_i^+ + q_i^+} \right)^{\eta_i} & \text{if } f_{,x_i}^k > 0 \\ \left(-\dfrac{f_i^- + q_i^-}{q_i^+} \right)^{\eta_i} & \text{if } f_{,x_i}^k < 0 \end{cases} \tag{14}$$

and

$$q_i^+ = \sum_{+} \lambda_j h_{j,x_i}, \quad q_i^- = \sum_{-} \lambda_j h_{j,x_i}, \quad f_i^+ = f_{,x_i}^k (f_{,x_i}^k > 0),$$

$$f_i^- = f_{,x_i}^k (f_{,x_i}^k < 0), \quad \text{and } \eta_i = \frac{1}{\xi_i^+ - \xi_i^-} \tag{15}$$

where Σ_+ represents the summation over the terms that satisfies $h_{j,x_i}^k>0$, and Σ_- is the summation over the other terms. Note that the CONLIN algorithm developed by Fleury and Braibant [21] is a special case of Eq. (13) when $c_i=0$, $\xi_i^+=1$, and $\xi_i^-=-1$.

If one assumes that $f_{,x_i}^k<0$ and $h_{j,x_i}^k>0$ are satisfied for the design variable x_i at the kth iteration step, then the updating rule in Eq. (13) can be further reduced (for the design variable and the iteration step) to

$$x_i^* = c_i + \left(-\frac{f_{,x_i}^k}{\sum_{j=1}^{m} \lambda_j h_{j,x_i}^k} \right)^{\eta_i} (x_i^k - c_i) \tag{16}$$

where

$$\eta_i = \frac{1}{\zeta_i - \xi_i} \ (\eta_i > 0 \text{ by definitions of } \zeta_i \text{ and } \xi_i) \tag{17}$$

Note that the OC algorithm used widely in topology optimization (for example, in [6]) is a special case of Eq. (16) for $c_i=0$ and $m=1$ ($m=1$ indicates a single domain topology optimization problem). Also note that parameter η_i in Eq. (17) (as well as in Eq. (14)) has the same meaning as in the standard OC algorithm, and it can be adjusted to affect the smoothness of the iteration process and the convergence speed. Further note that the conditions $f_{,x_i}^k<0$ and $h_{j,x_i}^k>0$ can always be satisfied for topology optimization problems with an objective to minimize the mean compliance. Therefore, updating rule Eq. (16) can be used to solve such problems. In general, updating rule Eq. (14) can be used when such conditions are not satisfied, which often happens when dealing with, for example, objectives related to eigenvalue and dynamic response. Updating rule Eq. (14) is generally recommended, since it can always satisfy the convexity conditions defined in [16] and it features a more robust convergence. Parameters $c_i(i=1,2,\ldots,n)$ in Eq. (13) can be set to $c_i=0$, $c_i=x_i^{k-1}$, or other values. In this paper, we simply suggest using $c_i=0$ for all i. The optimization algorithm proposed above has been extensively tested for many different topology optimization problems, and it has proven to be highly efficient and robust.

Additional updating rules can be obtained from the GSAO method using the updating Eq. (10). For instance, if one lets $\xi_i=\zeta_{ji}=-1$ and chooses c_i and e_{ji} as $c_i=L_i<x_i$ (when $f_{,x_i}^k<0$), $c_i=U_i>x_i$ (when $f_{,x_i}^k>0$), $e_{ji}=L_i$ (when $h_{j,x_i}^k<0$), and $e_{ji}=U_i$ (when $h_{j,x_i}^k>0$), then the GSAO algorithm reduces to the MMA algorithm [22]. If one lets $\xi_i=2$ and uses the second-order diagonal derivatives to determine c_i, then the GSAO algorithm becomes the DSQP algorithm [23]. Further discussion of this subject, however, is beyond the scope of this paper.

Note that, in practice, the use of moving boundaries is an effective way to avoid potential oscillations in the iterative process and to improve the convergence of the overall optimization procedure. The moving boundaries can be defined as

$$\underline{x}_i^k = \max\{(1-\mu)x_i^k,\underline{x}_i\} \text{ and } \bar{x}_i^k = \min\{(1+\mu)x_i^k,\bar{x}_i\} \tag{18}$$

where $\mu(0<\mu<1)$ is a given parameter. Then the updating rule can be modified as

$$x_i^{k+1} = \begin{cases} \underline{x}_i^k & \text{if } x_i^* \leqslant \underline{x}_i^k \\ x_i^* & \text{if } \underline{x}_i^k < x_i^* < \bar{x}_i^k \\ \bar{x}_i^k & \text{if } x_i^* \geqslant \bar{x}_i^k \end{cases} \tag{19}$$

where x_i^{k+1} denotes the updated design variable that is limited by the moving boundaries, and x_i^* is obtained from Eq. (11) (or its special cases, Eqs. (13) or (16)).

Substructuring Technique for MDTO

Substructuring and condensation methods provide an effective means of reducing significantly the size of structural models across the frequency range, which in turn permits MDTO problems to be solved in a more efficient and effective way.

State equations for the topology optimization problem, Eq. (4), usually feature a very large number of variables (typically the nodal displacements). Even for a simple structure, a finely discretized finite element model can involve from thousands to millions of nodal variables. In topology optimization, such fine meshes are usually required for obtaining smooth boundaries and interfaces. Substructuring methods provide an effective tool for condensing the analysis variables into a much smaller set, thus greatly reducing computer memory requirements and increase computational efficiency. This approach is particularly natural and useful for multidomain topology optimization problems, since it not only improves the computational efficiency, but also extends the applicability (in terms of problem solvability) of the methodology to a wider range of design problems, which will be detailed in this paper. While substructuring can be used in several ways to improve computational efficiency, including to condense out the design domains, in this paper the condensed analysis variables are those associated with the nondesign subdomains of the structure (see Fig. 2). The nodal displacement variables can be condensed into those at the boundaries of the subdomain when a static analysis is considered, and with an additional small number of generalized modal coordinates when a dynamic problem is considered. Note that the design and nondesign subdomains can be switched if necessary during the different design stages.

In the case of a static condensation, assuming that $\mathbf{u}_a$ denotes the vector of nodal displacements associated with the active nodes (usually the boundary nodes) of a nondesign subdomain, and $\mathbf{u}_o$ denotes the vector of the nodal displacements associated with the other nodes (usually the internal nodes) of the same subdomain, the state equation for the subdomain can be written as

$$\begin{bmatrix} \mathbf{k}_{oo} & \mathbf{k}_{oa} \\ \mathbf{k}_{ao} & \mathbf{k}_{aa} \end{bmatrix} \begin{Bmatrix} \mathbf{u}_o \\ \mathbf{u}_a \end{Bmatrix} = \begin{Bmatrix} \mathbf{f}_o \\ \mathbf{f}_a \end{Bmatrix} \tag{20}$$

where $\mathbf{k}_{ij}(i,j=o,a)$ are the blocks of the stiffness matrix of the subdomain associated with $\mathbf{u}_o$ and $\mathbf{u}_a$, respectively, and $\mathbf{f}_o$ and $\mathbf{f}_a$ are the corresponding nodal force vectors. Then Eq. (20) can be condensed into an equation that involves only $\mathbf{u}_a$, namely,

$$[\mathbf{k}^*_{aa}]\{\mathbf{u}_a\} = \{\mathbf{f}^*_a\} \tag{21}$$

where $\mathbf{k}^*_{aa}=\mathbf{k}_{aa}-\mathbf{k}_{ao}\mathbf{k}^{-1}_{oo}\mathbf{k}_{oa}$ and $\mathbf{f}^*_a=\mathbf{f}_a-\mathbf{k}_{ao}\mathbf{k}^{-1}_{oo}\mathbf{f}_o$.

Note that the coordinate transformation for the condensation is

$$\begin{Bmatrix} \mathbf{u}_o \\ \mathbf{u}_a \end{Bmatrix} = [\mathbf{C}]\{\mathbf{u}_a\} \tag{22}$$

where the transformation matrix is

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{oa} \\ \mathbf{C}_{oo} \end{bmatrix} = \begin{bmatrix} -\mathbf{k}^{-1}_{oo}\mathbf{k}_{oa} \\ \mathbf{I} \end{bmatrix} \tag{23}$$

Each column of $\mathbf{C}$ is called a "*static mode*" or a "*constraint mode*." Note that in practice, the static modes are calculated with a more efficient computational procedure without computing the inverse of $\mathbf{k}_{oo}$ in Eq. (23).

In the case of a dynamic condensation, assuming that the state equation is written as

$$\begin{bmatrix} \mathbf{m}_{oo} & \mathbf{m}_{oa} \\ \mathbf{m}_{ao} & \mathbf{m}_{aa} \end{bmatrix} \begin{Bmatrix} \ddot{\mathbf{u}}_o \\ \ddot{\mathbf{u}}_a \end{Bmatrix} + \begin{bmatrix} \mathbf{k}_{oo} & \mathbf{k}_{oa} \\ \mathbf{k}_{ao} & \mathbf{k}_{aa} \end{bmatrix} \begin{Bmatrix} \mathbf{u}_o \\ \mathbf{u}_a \end{Bmatrix} = \begin{Bmatrix} \mathbf{f}_o \\ \mathbf{f}_a \end{Bmatrix} \tag{24}$$

where $\mathbf{m}_{ij}(i,j=o,a)$ denote the blocks of the mass matrix, then the transformation equation, Eq. (22), is extended as

$$\begin{Bmatrix} \mathbf{u}_o \\ \mathbf{u}_a \end{Bmatrix} = [\mathbf{D}]\begin{Bmatrix} \mathbf{q}_n \\ \mathbf{u}_a \end{Bmatrix} \tag{25}$$

where $\mathbf{q}_n$ is a vector of so-called generalized modal coordinates. The transformation matrix $\mathbf{D}$ for the dynamic problem can then be obtained as

$$\mathbf{D} = \begin{bmatrix} \mathbf{\Phi}_{on} + \mathbf{\Psi}_{oa}\mathbf{\Phi}_{an} & \mathbf{\Psi}_{oa} \\ 0 & \mathbf{I} \end{bmatrix} \tag{26}$$

where $\mathbf{\Phi}=\left\{\begin{smallmatrix}\mathbf{\Phi}_{an}\\ \mathbf{\Phi}_{on}\end{smallmatrix}\right\}$ is the normal modes matrix, which is calculated by solving a free vibration eigenvalue problem associated with the subdomain. The matrix $\mathbf{\Psi}=\left\{\begin{smallmatrix}\mathbf{\Psi}_{oa}\\ \mathbf{\Psi}_{oo}\end{smallmatrix}\right\}$ contains the so-called "*quasi-static modes*" (QSM) [19,24,25], which can be calculated by solving the quasistatic (frequency response) problem associated with Eq. (24). Theoretically, $\mathbf{\Psi}$ can be written as

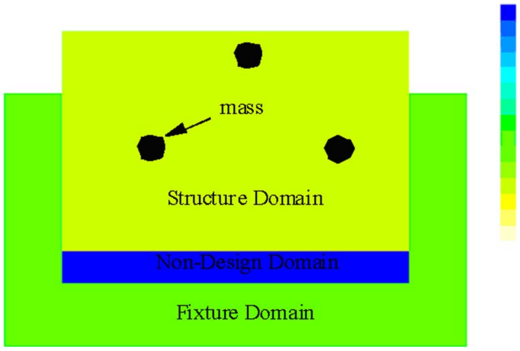


Fig. 3 A structure-fixture simultaneous design problem

$$\mathbf{\Psi} = \begin{bmatrix} -(\mathbf{k}_{oo}-\omega_c^2\mathbf{m}_{oo})^{-1}(\mathbf{k}_{oa}-\omega_c^2\mathbf{m}_{oa}) \\ \mathbf{I} \end{bmatrix} \tag{27}$$

where ω_c is the "*central frequency*," which is determined by the frequency range of interest [19]. Note that the size of $\mathbf{q}_n$ in Eq. (25) is usually far smaller than $\mathbf{u}_o$, therefore, the coordinate transformation Eq. (25) [as well as Eq. (22)] significantly reduces the size of the analysis problem. For a static analysis, Eq. (22) does not induce any error into the original analysis problem. But for a dynamic analysis problem, Eq. (25) is an approximation, and the error can be controlled by properly choosing the central frequency and/or considering additional modal coordinates in Eq. (25). Note that for $\omega_c=0$, the QSM defined in Eq. (27) reduce to the traditional *static (constraint) modes* proposed by Hurty [26] and Craig and Bampton [27], which are given in Eq. (23). Compared with the static modes, the QSM can significantly reduce both the size of the analysis problem and the error induced by the coordinate reduction process, particularly within the frequency range of interest. This is because not only the higher-frequency modes are truncated, but also the lower-frequency modes outside the frequency range of interest. Therefore, the QSM can handle dynamic response problems within higher frequency ranges, and they include the traditional static modes as a special case. Note that the central frequency is typically selected at the middle of the frequency domain of interest, and that multiple central frequencies can also be used [25].

Example Applications

Structure-Fixture Simultaneous Design Problem. This first example illustrates how the multidomain topology optimization technique can be applied to a "structure-fixture simultaneous design" problem. (The reader can find different treatments in [30,31] for similar design problems.) Figure 3 depicts a structure whose optimal topology is sought in a "structure domain," and which features a solid bar at the bottom of the domain (referred to as a nondesign domain). The structure needs to support three lumped masses distributed in the structure domain, as shown in Fig. 3. Meanwhile, a fixture design is sought in the "fixture domain," whose purpose is to attach the structure onto, for example, a fixed workplace. The fixture domain reflects the allowable space where the fixture can be placed between the structure and the workplace. In this example, the objective is to maximize the fundamental eigenfrequency of the structure-fixture system so as to limit its vibration response under certain operating conditions.

It is assumed that the total (normalized) amount of the material is 14, which represents a 35% material density over the whole design domain. Figure 4 depicts an optimal design obtained from the MDTO process when the material is evenly assigned to the structure and the fixture (namely 7:7). In this case, the optimized

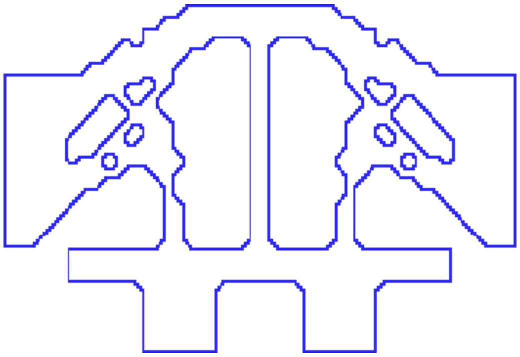


Fig. 4 Design 1 (7:7): optimized eigenfrequency=4.37

eigenfrequency of the coupled system is 4.37 (also on a normalized scale). Figure 5 shows another design obtained when a larger amount of the material is assigned to the structure domain and less material is assigned to the fixture domain (namely, 11 for the structure domain and 3 for the fixture domain). The optimized fundamental eigenfrequency becomes 5.01 in this case. Note that different materials could be assigned to the structure and to the fixture, and the use of multiple materials may allow the designer to make an even better design.

The results shown in Figs. 4 and 5 emonstrate a notable advantage of using the structure-fixture simultaneous design process. There is a smooth transition of the material between the structure and fixture subdomains, which would be difficult to achieve if the structure and fixture were designed separately. More detailed attachment conditions, for example, using bolts or welds, and other design objectives, for example, strength and stability, could be considered in the design process, but these are not discussed here.

Functionally Gradient Material Design Problem. The second example illustrates how the MDTO technique can be applied to a design problem for a so-called "functionally gradient material" (FGM) [28]. Here the "material" is interpreted as an "engineered material" or a "composite material" [29]. The objective of this example is to design a structure that can support a top layer of a material made of stiff tiles, as depicted in Fig. 6. The structure also needs to feature a gradually-reduced stiffness in order to match a soft material (skin) at the bottom. The top layer (tiles) and the skin layer are considered as nondesign domains, while the four supporting layers are optimized for their global stiffness subject to the graded material densities. The material density for each supporting layer is determined from the consideration of gradually reducing the stiffness of the structure. In this example, the material densities for the four layers are selected to be 45, 36, 27, and 18% from the top layer to the bottom layer, respectively (see Fig. 6). Figure 7(*a*) shows the optimum layout obtained from a multidomain topology optimization process. It is seen that the material density can be distributed (in this example along the vertical axis) in the desired way, and that a smooth stiffness transition can be obtained, thus allowing for the attachment of a very stiff material to a much softer material. Figure 7(*b*) further shows a polished design based on the optimum layout of Fig. 7(*a*) and obtained through postprocessing. Note that the objective function used in this optimization problem is to maximize the rigidity of the structure for the static loads shown in Fig. 6. Manufacturing constraints can also be considered in the design process so that the material pattern in Figs. 7(*a*) or 7(*b*) can be optimized to satisfy the fabrication requirements.

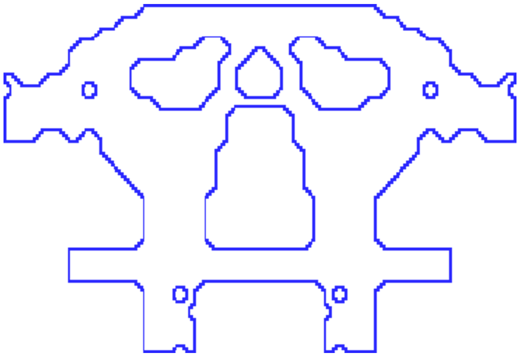


Fig. 5 Design 2 (11:3): optimized eigenfrequency=5.01

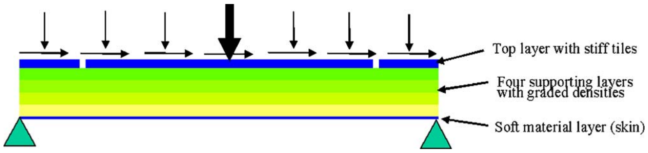

Top layer with stiff tiles
Four supporting layers with graded densities
Soft material layer (skin)

Fig. 6 Example FGM with a top layer made of stiff tiles, four supporting layers with graded densities, and a soft skin. The graded densities for the four supporting layers are, from the top layer to the bottom layer, 45%, 36%, 27%, and 18%, respectively.

Crush Energy Management Design Problem. This example illustrates the application of the MDTO technique to a "crush energy management" problem. Figure 8 depicts a sandwich beam under an axial load, which represents a major crush load, and under a bending load, which represents the load that the structure normally carries. Subdomains 2 and 4 are defined as the design areas for crush energy absorption, and they are the regions where buckling is desired during a crush event. Subdomains 1 and 3 are the regions where buckling is not allowed to happen during the crush event. The overall structure also needs to have a maximum stiffness for the given amount of material, so it can resist both the compressive and bending loads shown in Fig. 8. In order to obtain such a structure, the MDTO process is employed, in which less material is assigned to the subdomains 2 and 4, but more material is assigned to the subdomains 1 and 3, while the global stiffness of the structure is optimized with respect to the given loads.

Figure 9 shows the resulting optimized structure obtained by assigning a 52% material density to subdomains 1 and 3 and a 26% material density to subdomains 2 and 4. As more material has been distributed into subdomains 1 and 3, the structure is much stronger in these regions, while it is weaker in subdomains 2 and 4. Figure 9 also shows that a smooth material transition is obtained between the stronger areas and the weaker areas, which results in an optimum distribution for the global stiffness of the

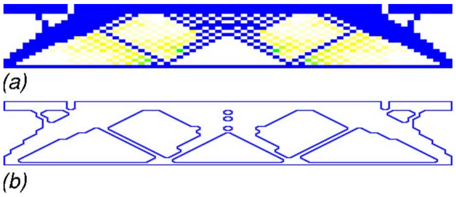


Fig. 7 Optimal design for the FGM example: (*a*) optimum layout; (*b*) finalized design

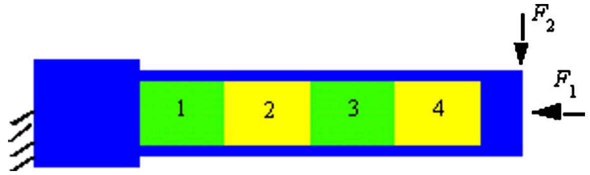
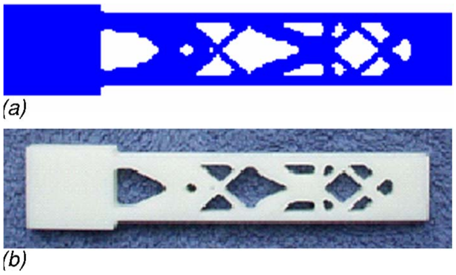


Fig. 8 Design problem for crush energy management


(a)

(b)

Fig. 9 Optimized design for the crush energy management problem; (*a*) optimum layout; (*b*) finalized design

structure.

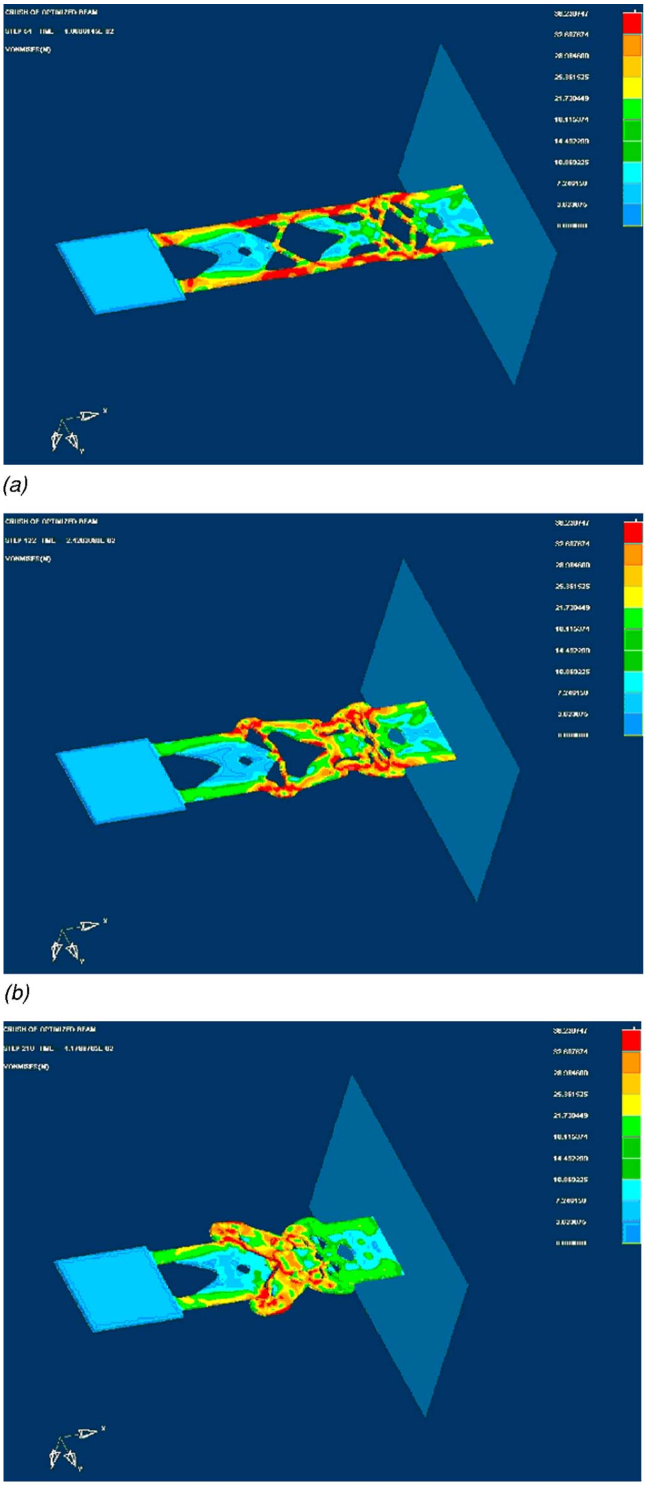
In order to verify the optimized design in terms of its crush energy management capability, a virtual prototype has been developed using LS-Dyna3D. Figure 10 shows snapshots from a time simulation in which the virtual prototype beam is crashed into a rigid wall at 13 m/s (48 km/h). As observed in Fig. 10(*a*), the first buckling occurs in Design Domain 4 (refer to Fig. 8), which was designed for energy absorption. Figure 10(*b*) shows that the second buckling occurs in Design Domain 2, which was also designed for absorbing energy. Figure 10(*c*) shows that the structure is completely crushed at the end of the crash process, with all the holes in the energy-absorbing areas being closed, while Design Domains 1 and 3 remain almost uncrushed as designed. Figure 11 illustrates the virtual crash test of a nominal beam design, where a homogeneous beam without open holes is crashed onto the same rigid wall at the same speed. Figure 12 further compares the crash forces obtained from the virtual crash tests of the optimum and the nominal beam designs, where the crash force is measured at the tip of the beam. As shown in Fig. 12, an almost-constant crash force is obtained for the optimum design, with an amplitude of less than one-third of that of the maximum crash force for the nominal design. This example demonstrates the effectiveness of the new methodology and the design process developed. These have significant potential applicability to the optimization of ground vehicles for crush energy management.

Note that, in this optimization problem, 72 iterations were required to satisfy the convergence criteria of $|f^{k+1}-f^k|/|f^k| \leq 1.0e-4$ when the original CONLIN algorithm [namely using $\xi_i^+=1$, and $\xi_i^-=-1$ in Eq. (13)] was used. The number of iterations was reduced by about half, to 35, for the same criteria when the extended CONLIN algorithm was used with the parameters $\xi_i^+=1$, and $\xi_i^-=-0.2$. Thus the new optimization algorithm proposed in this paper not only provides useful flexibility to deal with various topology optimization problems, but also features significantly greater convergence. Further discussion regarding how to determine the optimum parameters in the optimization algorithm will be discussed in a future paper.

Also note that the beam model considered was discretized with a 92×20 mesh, yielding 1840 finite elements and 1953 nodes, which include 1072 elements and 1288 nodes associated with the nondesign domains. By using the substructuring technique introduced in this paper, the 903 internal nodes in the nondesign domains were reduced, which resulted in an analysis problem of smaller size without loosing any fidelity. The savings on CPU time and memory requirement, however, will become meaningful only for larger and more complex structural systems.

Conclusions

The multidomain topology optimization technique introduced in this paper extends the standard topology optimization method to the realm of more realistic engineering design problems, including (1) the simultaneous design of multiple, coupled components, (2) the design of functionally graded and other composite materials, (3) the design of multifunctional structures, and (4) the design for crush energy management and other impact-related problems. The MDTO process developed allows the designer to distribute materials in a desired way as well as define different objectives and constraints among the different subdomains, en-



(a)

(b)

(c)

Fig. 10 Virtual crash test of the prototype: deformed structure is shown at (*a*) *t*=0.01 s, (*b*) *t*=0.02 s, (*c*) *t*=0.04 s. (*a*) Snapshot 1: first buckling occurs in Design Domain 4 (*t*=0.01 s). (*b*) Snapshot 2: second buckling occurs in Design Domain 2 (*t*=0.02 s). (*c*) Snapshot 3: third buckling occurs in Design Domain 3 (*t*=0.04 s)

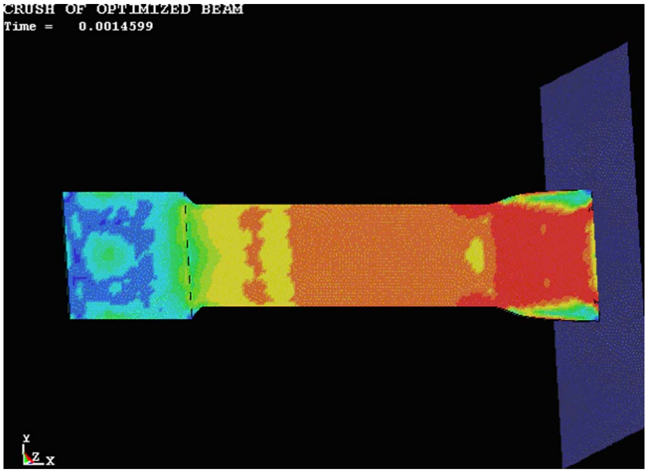


Fig. 11 Virtual crash test of a homogenous beam without holes

ables the use of multiple materials, and helps address issues such as multifunctionality, for example, by treating multidisciplinary physical processes in the different subdomains.

A new algorithm of Sequential Approximate Optimization has been proposed for multidomain topology optimization, which is an enhancement and generalization of previous SAO algorithms (including OC, CONLIN, MMA, and DSQP). An advanced substructuring method using quasi-static modes has also been introduced to condense the nodal variables associated with the multidomain topology optimization problem, especially for the nondesign subdomains. Several important applications have been presented to demonstrate the effectiveness of MDTO and its potential for a wide range of engineering design problems.

Acknowledgment

The authors would like to acknowledge the grant support provided by the U.S. Army RDECOM. They also thank Dr. Hui Wang for obtaining the crash simulation results in Figs. 10–12.

Nomenclature

$a(a_e)$ = design variable of the microstructure (with respect to element e)
$b(b_e)$ = design variable of the microstructure (with respect to element e)
a_i^k, b_{ji}^k = coefficients defined in Eq. (7)
$\bar{a}_i^k(\bar{b}_{ji}^k)$ = $\xi_i a_i^k(\zeta_{ji} b_{ji}^k)$
$\mathbf{C}$ = transformation matrix in Eq. (22) (also global damping matrix in Eq. (2))
c_i = controlling parameter in Eq. (5) for the objective function
$\mathbf{D}$ = transformation matrix in Eq. (25)
$\mathbf{D}^H$ = homogenized elasticity matrix
e_{ji} = controlling parameter in Eq. (5) for constraint functions
f = objective function
$f_{,x_i}^k$ = $\partial_f/\partial x_i|_{\mathbf{X}=\mathbf{X}^k}$
f_0^k = coefficients defined in Eq. (7)
$\mathbf{f}(\mathbf{f}_o,\mathbf{f}_a)$ = nodal force vector (components associated with $\mathbf{u}_o$ and $\mathbf{u}_a$)
$\mathbf{f}_a^*$ = condensed nodal force vector in Eq. (21)
h_j = jth constraint function
h_{j,x_i}^k = $\partial_{h_j}/\partial x_i|_{\mathbf{X}=\mathbf{X}^k}$
h_{0j}^k = coefficients defined in Eq. (7)
$\mathbf{k}_e,\mathbf{K}$ = stiffness matrix (element/global)
$\mathbf{k}_{ij}(i,j=o,a)$ = blocks of the stiffness matrix
$\mathbf{k}_{aa}^*$ = condensed stiffness matrix in Eq. (21)
L^k = Langrangian function in the kth iteration step
m = total number of subdomains
$\mathbf{m}_e,\mathbf{M}$ = mass matrix (element/global)
$\mathbf{m}_{ij}(i,j=o,a)$ = blocks of the mass matrix
n_{el} = total number of finite elements
n = total number of design variables
$\mathbf{q}_n$ = generalized modal coordinate vector
$\mathbf{u}(\mathbf{u}_o,\mathbf{u}_a)$ = nodal displacement vector (components associated with internal and interface nodes)
$\mathbf{X}(\mathbf{X}^k)$ = design variable vector (in the kth iteration)
$x_i(x_i^k)$ = ith design variable (in the kth iteration)
$\mathbf{X}^*(x_i^*)$ = updated design variable vector (updated ith design variable)
$\underline{x}_i$ = lower limit of the design variable x_i
$\bar{x}_i$ = upper limit of the design variable x_i
$\underline{x}_i^k$ = moving lower limit of the design variable x_i at the kth iteration step
$\bar{x}_i^k$ = moving upper limit of the design variable x_i at the kth iteration step
y_i = intermediate variable defined in Eq. (5) for objective function
z_{ji} = intermediate variable defined in Eq. (5) for constraint functions
$\lambda(\lambda_j)$ = Lagrange multiplier vector (jth component) corresponding to the volume constraints
ξ_i = controlling power parameter in Eq. (5) for the objective function
ζ_{ji} = controlling power parameter in Eq. (5) for the constraint functions
η_i = parameters defined in Eq. (15) or Eq. (17)
$\theta(\theta e)$ = orientation angle of microstructure (with respect to element e)
ρ^H = homogenized mass density
$\mathbf{\Phi}$ = normal mode matrix
$\mathbf{\Psi}$ = quasistatic mode matrix defined in Eq. (27)

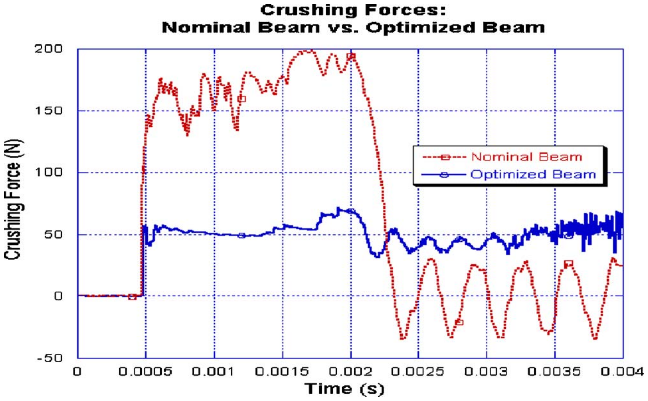


Fig. 12 Comparison of crash forces for the optimized design and the nominal design

References

[1] Rozvany, G. I. N., Zhou, M., and Sigmund, O., 1994, "Topology Optimization in Structural Design," in *Advances in Design Optimization*, H. Adeli, ed., Chapman and Hall, London, UK, pp. 340–399.

[2] Bendsøe, M. P., Diaz, A., and Kikuchi, N., 1992, "Topology and Generalized Layout Optimization of Elastic Structures," *Topology Design of Structures*, Bendsøe, M. P. and Soates, C. A. M., eds., NATO ASI Series, Kluwer Academic, Dordrecht, pp. 159–205.

[3] Bendsøe, M. P., 1995, *Optimization of Structural Topology, Shape, and Material*, Springer-Verlag, Berlin.

[4] Hassani, B., and Hinton, E., 1999, *Homogenization and Structural Topology*

Optimization: Theory, Practice, and Software, Springer-Verlag, Berlin.

[5] Bendsøe, M. P., 2003, *Topology Optimization: Theory, Methods and Applications*, Springer-Verlag, Berlin, Heidelberg.

[6] Bendsøe, M. P., and Kikuchi, N., 1988, "Generating Optimal Topologies in Structural Design Using a Homogenization Method," Comput. Methods Appl. Mech. Eng., **71**, pp. 197–24.

[7] Suzuki, K., and Kikuchi, N., 1991, "A Homogenization Method for Shape and Topology Optimization," Comput. Methods Appl. Mech. Eng., **93**, pp. 291–318.

[8] Kikuchi, N., Suzuki, K., and Fukushima, J., 1991, "Layout Optimization Using the Homogenization Method: Generalized Layout Design of Three-Dimensional Shells for Car Bodies," *Optimization of Large Structural Systems*, G. I. N. Rozvany, ed., NATO-ASI Series, Berchtesgaden, Vol. 3, pp. 110–126.

[9] Kikuchi, N., Nishiwaki, S., Fonseca, J. O., and Silva, E. C. N., 1998, "Design Optimization Method for Compliant Mechanisms and Material Microstructure," Comput. Methods Appl. Mech. Eng., **151**, pp. 401–417.

[10] Nishiwaki, S., Frecker, M., Min, S. J., and Kikuchi, N., 1998, "Topology Optimization of Compliant Mechanisms Using the Homogenization Method," Int. J. Numer. Methods Eng., **42**, pp. 535–559.

[11] Fecker, M., Ananthasuresh, G. K., Kikuchi, N., and Kota, S., 1997, "Topological Synthesis of Compliant Mechanisms Using Multi-Criteria Optimization," ASME J. Mech. Des., **119**(2), pp. 238–245.

[12] Diaz, A., and Kikuchi, N., 1992, "Solutions to Shape and Topology Eigenvalue Optimization Problems Using a Homogenization Method," Int. J. Numer. Methods Eng., **35**, pp. 1487–1502.

[13] Ma, Z.-D., Kikuchi, N., Cheng, H.-C., and Hagiwara, I., 1994, "Structural Design for Obtaining Desired Frequencies by Using the Topology and Shape Optimization Method," Comput. Syst. Eng., **5**(1), pp. 77–89.

[14] Ma, Z.-D., Kikuchi, N., Cheng, H.-C., and Hagiwara, I., 1995, "Topological Optimization Technique for Free Vibration Problems," ASME J. Appl. Mech., **62**, pp. 200–207.

[15] Cheng, H.-C., Kikuchi, N., and Ma, Z.-D., 1994, "Generalized Shape/Topology Designs of Plate/Shell Structures for Eigenvalue Optimization Problems," *Proceedings of ASME International Mechanical Engineering Congress and Exposition*, PED-Vol. 68-2, pp. 483–492.

[16] Ma, Z.-D., Kikuchi, N., and Hagiwara, I., 1993, "Structural Topology and Shape Optimization for a Frequency Response Problem," Comput. Mech., **13**(3), pp. 157–174.

[17] Ma, Z.-D., Kikuchi, N., and Cheng, H.-C., 1995, "Topological Design for Vibrating Structures," Comput. Methods Appl. Mech. Eng., **121**, pp. 259–280.

[18] Ma, Z.-D., and Kikuchi, N., 1995, "A New Method of the Sequential Approximate Optimization," Eng. Optimiz., **25**, pp. 231–253.

[19] Shyu, W.-H., Ma, Z.-D., and Hulbert, G. M., 1997, "A New Component Mode Synthesis Method: Quasi-Static Mode Compensation," Finite Elem. Anal. Design, **24**, pp. 271–281.

[20] Berke, L., and Khot, N. S., 1987, "Structural Optimization Using Optimality Criteria," *Computer Aided Optimal Design: Structural and Mechanical Systems*, C. A. Mota Soares, ed., Springer-Verlag, Berlin, pp. 271–311.

[21] Fleury, C., and Braibant, V., 1986, "Structural Optimization: a New Dual Method Using Mixed Variables," Int. J. Numer. Methods Eng., **23**, pp. 409–428.

[22] Svanberg, K., 1987, "The Method of Moving Asymptotes: A New Method for Structural Optimization," Int. J. Numer. Methods Eng., **24**, pp. 359–373.

[23] Fleury, C., 1987, "Efficient Approximation Concepts Using Second Order Information," Int. J. Numer. Methods Eng., **28**, pp. 2041–2058.

[24] Ma, Z.-D., and Hagiwara, I., 1991, "Improved Mode-Superposition Technique for Modal Frequency Response Analysis of Coupled Acoustic-Structural Systems," AIAA J., **29**(10), pp. 1720–1726.

[25] Shyu, W.-H., Gu, J., Hulbert, G. M., and Ma, Z.-D., 2000, "On the Use of Multiple Quasi-Static Mode Compensation Sets for Component Mode Synthesis of Complex Structures," Finite Elem. Anal. Design, **35**, pp. 119–140.

[26] Hurty, W. C., 1965, "Dynamic Analysis of Structural Systems Using Component Modes," AIAA J., **3**, pp. 678–685.

[27] Craig, R. R., and Bampton, M. C. C., 1968, "Coupling of Substructures for Dynamic Analysis," AIAA J., **6**, pp. 1313–1319.

[28] Cherradi, N., Kawasaki, A., and Gasik, M., 1994, "World Trends in Functional Gradient Materials Research and Development," Composites Eng., **4**(8), pp. 883–894.

[29] Wojciechowski, S., 2000, "New Trends in the Development of Mechanical Engineering Materials," J. Mater. Process. Technol., **106**, pp. 230–235.

[30] Chickermane, H., and Gea, H. C., 1997, "Design of Multi-Component Structural Systems for Optimal Layout Topology and Joint Locations," Eng. Comput., **13**, pp. 235–243.

[31] Thomas, Buhl, 2002, "Simultaneous Topology Optimization of Structure and Supports," Struct. Multidiscip. Optim., **23**(5), pp. 336–346.

[32] Ma, Z.-D., 2001, "Topology Optimization with Multiple Domains," oral presentation at the 6th U.S. National Congress on Computational Mechanics, Dearborn, Michigan, August, 2001.

M. Weiss[1]
e-mail: mweis@deakin.edu.au

B. F. Rolfe
e-mail: brolfe@deakin.edu.au

M. Dingle
e-mail: mdingle@deakin.edu.au

School of Engineering and Technology,
Deakin University,
Geelong, Vic. 3217,
Australia

J. L. Duncan
The University of Auckland,
Auckland, New Zealand
e-mail: jl.duncan@xtra.co.nz

Elastic Bending of Steel-Polymer-Steel (SPS) Laminates to a Constant Curvature

The shear strain of the interlayer in the elastic regime for a Steel-Polymer-Steel (SPS) laminate material has been studied during bending to a constant curvature. An analytical model is developed and the influence of process parameters are analyzed. The tension in the cover sheets is also determined and, finally, a moment diagram is calculated. The results show that the moment in the SPS laminate is nonuniform along the bent strip even though the curvature is constant because of the tension and compression forces introduced in the cover sheets by the shear reaction force of the interlayer material.

[DOI: 10.1115/1.2164512]

1 Introduction

In recent years, automobile manufacturers have developed Steel-Polymer-Steel (SPS) laminates for automobile parts such as fenders, doors and interior panels. Compared with homogeneous metal sheets, SPS laminates offer a significantly lower density [1] and better sound and vibration damping characteristics [2]. Compared to other lightweight solutions they also have the advantage of a good surface finish and can maintain a bending rigidity which is almost equal to that of a simple metal sheet of the same total thickness [3]. Thus, SPS laminates can have similar bending stiffness to homogeneous steels, but at lower weight.

Despite the superior mechanical properties of SPS laminates, they have not gained widespread use in the automobile industry because of geometric distortion and defects arising from the forming process. Several geometrical defects are observed which do not occur with homogeneous material, especially in bending. Here the two major concerns are unpredictable springback behaviour and the "gull-wing" bend, both being the result of large shear deformations in the interlayer, as the core material is weak compared with the metal cover sheets [4–6].

To ensure a safe and widespread usage of SPS laminates, it will be necessary to develop analytical models to predict their forming behavior in bending and determine ways to decrease geometrical defects like springback and the gull-wing bend.

Many authors have proposed first and high order shear theories to predict the behaviour of thick laminated composite plates [7–9], but all these investigations mainly focused on the elastic analysis of static or vibratory bending. Only a few investigations dealt with the elastic-plastic bending condition relevant to the press formability of thin sandwich plates with a soft core, or SPS laminates. Yoshida [10] analyzed the V-bending process of SPS laminates taking into account the transverse shear of the polymer core. Here the Kirchhoff-Love hypothesis for the deformation of the face layers was used. For the deformation of the core it was assumed that the transverse section, originally normal to the midsurface remains plane, but not normal to the midsurface during bending. The analytical model showed that the shear deformation in the core of SPS laminates and in that way the gull-wing bend can be decreased by using nonsymmetrical sandwich plates. Takiguchi and Yoshida [11] investigated the mechanism of shear deformation of the adhesive layer during bending to generate an idealized model of rate independent plasticity based on the Kirchhoff-Love hypothesis for the cover sheets and the shear flexible theory of Mindlin for the deformation of the adhesive material. They showed that the gull-wing defect of SPS laminates in the V-bending process can be decreased by using longer die-spans. In a second approach Takiguchi and Yoshida included a constitutive model of viscoplasticity for the adhesive layer showing that the gull wing bend can be restricted by forming under higher speed [12]. Li Liu and Jyhwen Wang [13] proposed an analytical solution based on the Euler-Bernoulli straight -and curved beam deflections to predict springback and side wall curl for SPS laminates in wiper die bending.

The present work presents a simple analytical solution for free bending of SPS laminates for the elastic case. In the first part an approximate analytical model for the elastic case is developed that includes the shear modulus of the core material and the transfer of shear stress to the cover sheets. Based on the developed model the tension acting on the cover sheets as a result of the transfer of shear stress is determined. The final part of this paper calculates the bending moment on the laminate sheet for a constant radius of curvature.

The shear strain evolution in the interlayer over the sheet length is calculated and the influence of process parameters, such as bend radius and shear modulus of the interlayer material, is determined.

2 Material Description

A SPS material is shown schematically in Fig. 1. The cover sheet is of cold rolled low carbon steel and the thickness of this material ranges from 0.2 mm to 0.3 mm. The polymer core can consist of several types of plastic; the authors have experience with using PolyVinyl Chloride (PVC) and PolyPropylene (PP) sheets. The thickness of the polymer core, or interlayer, ranges from 0.3 mm to 0.9 mm. A thermoplastic adhesive film, with a thickness of 0.05 mm, has been used to bond the interlayer to the cover sheets. The presence of this adhesive film is neglected in the present work.

---

[1] Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received February 22, 2005; final manuscript received September 22, 2005. Review conducted by A. Maniatty. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

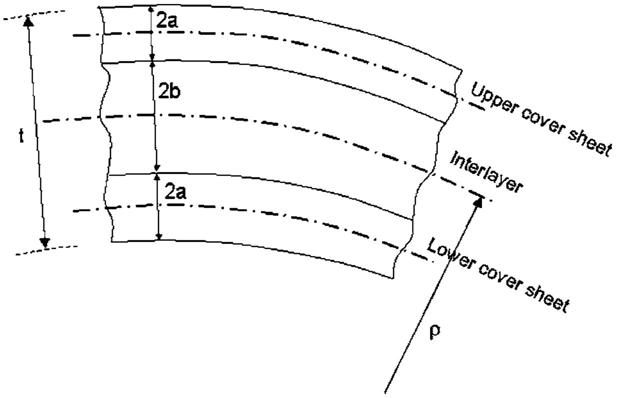


Fig. 1 Section of Steel Polymer Steel (SPS) material. The thickness of the SPS sheet is *t*, while the thickness of the cover sheets is 2*a*, and the thickness of the interlayer is 2*b*. The radius of curvature is *ρ*.

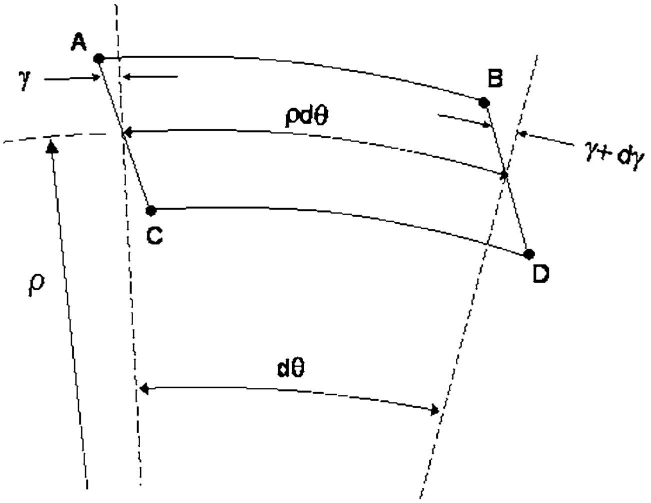


Fig. 3 Interlayer polymer under shear from the two cover sheets. Midsurface length is *ρdθ*, and the shear strain is *γ*, while the change in shear strain is *dγ*.

3 An Approximate Analytical Model for Pure Elastic Bending

This section develops the model for the strain and stress distribution along the length of a strip of continuous sheet bent in plane strain to the arc of a circle by a pure moment without any superimposed tension. In this model, the shear modulus of the core material is finite and the transfer of shear stress to the cover sheets is examined. Linear elastic behavior is assumed.

An idealized model of deformation of each layer for a bonded sheet is illustrated in Fig. 2. While the cover sheets deform in pure bending the interlayer material deforms in pure shear. We consider an initial element of length ds which after bending is as shown in Fig. 2. It is assumed that the midsurface of the core does not elongate, hence $ds = \rho d\theta$. The cover sheets may change in length and subtend an angle $d\phi$ as shown where $d\phi \neq d\theta$.

Figure 3 shows an element of the core bent to a midsurface radius ρ. As any influence of the inner and outer cover sheets are likely to be balanced, it is reasonable to assume that the length at the midsurface of the polymer core will not change and the initial and final length of the midsurface can be specified as $\rho d\theta$. The arc length between the upper corners of the polymer core, length AB, is therefore

$$\mathrm{AB} = (\rho + b)\mathrm{d}\theta - b\mathrm{d}\gamma \tag{1}$$

If the shear modulus of the core is greater than zero, a shear stress must exist at the interfaces between the core and cover sheets. To satisfy equilibrium, tension will build up in the outer cover sheet and compression in the inner. This results in elongation of the outer cover sheet and compression of the inner cover sheet. In Fig. 4 the strain distribution of an element of the outer cover sheet is shown. For $\rho \gg t$, the strain distribution in the cover sheet can be calculated as,

$$\varepsilon_1 = \varepsilon_a + \varepsilon_b = \varepsilon_a + \frac{y}{(\rho + a + b)} \approx \varepsilon_a + \frac{y}{\rho}, \tag{2}$$

where ε_a is the strain at the mid-surface. For the linear elastic case, the stress distribution is given by $\sigma_1 = \sigma_a + E' y/\rho$, where $\sigma_a = E' \varepsilon_a$.

The tension on the cover sheet is given by $\sigma_a = T/2a$, and during elastic deformation, the strain of the midsurface is $\varepsilon_a = (T/2a) \cdot (1/E')$.

The length of the midsurface of the cover sheet is therefore approximately,

$$\rho \mathrm{d}\theta \left\{ 1 + \frac{\mathrm{T}}{2a} \cdot \frac{1}{E'} \right\} = (\rho + a + b)\mathrm{d}\varphi. \tag{3}$$

The length of the interface AB can then be calculated as,

$$AB = (\rho + b)\mathrm{d}\varphi = \frac{\rho(\rho + b)}{(\rho + a + b)} \left\{ 1 + \frac{\mathrm{T}}{2a} \cdot \frac{1}{E'} \right\} \mathrm{d}\theta. \tag{4}$$

From Eqs. (1) and (4), the rate of change of shear strain is, after simplification:

$$\frac{\mathrm{d}\gamma}{\mathrm{d}\theta} = \frac{1}{b} \left\{ (a + b) - \frac{\rho}{2a} \cdot \frac{\mathrm{T}}{E'} \right\}, \tag{5}$$

where it is assumed $\rho \gg a$ or b, $\rho + a + b \approx \rho$, and $\rho + b \approx \rho$.

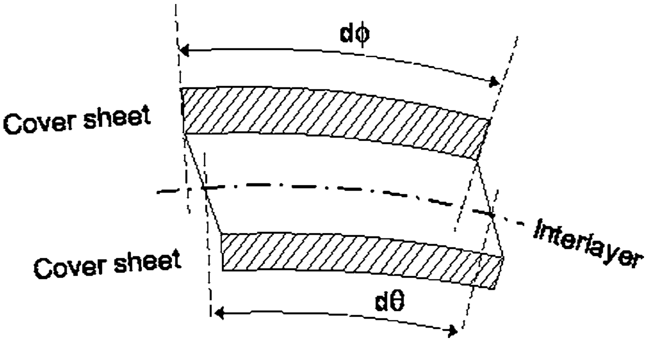


Fig. 2 An element of the laminate material bent at a constant radius of curvature

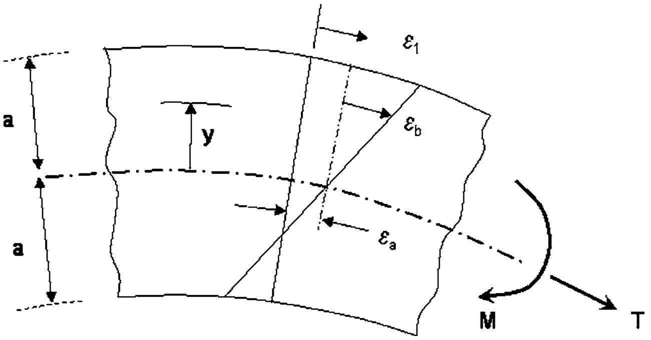


Fig. 4 Strain distribution in the cover sheet

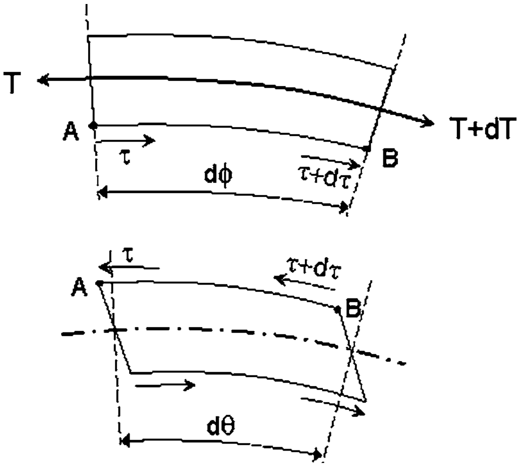


Fig. 5 Force equilibrium diagram of both the cover sheet and the interlayer

In Eq. (4), as indicated, the tension varies with the angle θ and this relationship must be determined. Figure 5 shows a schematic illustration of the shear stress acting at the interface AB due to the shear deformation of the interlayer and the resulting tension in the outer cover sheet. According to Fig. 5 the force equilibrium within the cover sheets is:

$$-T+T+\mathrm{d}T+\frac{(2\tau+\mathrm{d}\tau)}{2}AB=0, \tag{6}$$

where $(2\tau+\mathrm{d}\tau)/2\approx\tau$ is the average value of shear stress applied to the cover sheet surface. It is assumed that the change of the shear stress, $d\tau$, is very small when the subtended angle, $d\theta$, is very small. The shear stress τ applied by the interlayer material is equal to γG. Using Eqs. (1) and (6), the rate of change of the tension is

$$\frac{\mathrm{d}T}{\mathrm{d}\theta}=-\gamma G\left\{(\rho+b)-\frac{\mathrm{d}\gamma}{\mathrm{d}\theta}b\right\}. \tag{7}$$

Note that the change of the tension, dT, is negative as the shear angle is positive. It can be assumed that the change in shear strain term in Eq. (7) is much smaller than the bend radius term, $\mathrm{d}\gamma/\mathrm{d}\theta b \ll (\rho+b)\approx\rho$. Therefore Eq. (7) can be simplified to:

$$\frac{\mathrm{d}T}{\mathrm{d}\theta}=-\gamma G\rho. \tag{8}$$

A sensitivity analysis has shown that the error resulting from this simplification is less than 1% for typical material and processing values. Differentiating Eq. (5) with respect to θ, and substituting for $dT/d\theta$ using Eq. (8) gives:

$$\frac{\mathrm{d}^2\gamma}{\mathrm{d}\theta^2}-K\gamma=0 \tag{9}$$

where

$$K=\frac{\rho^2}{b}\cdot\frac{G}{2aE'}$$

This is an ordinary differential equation that can be solved in the following form:

$$\gamma=C_1\exp(\sqrt{K}\theta)+C_2\exp(-\sqrt{K}\theta). \tag{10}$$

The rate of change of shear strain can then be obtained as follows:

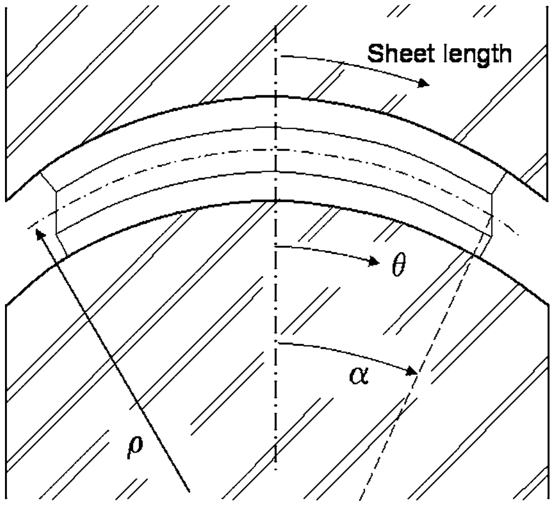


Fig. 6 Bending of the SPS laminate material to a constant radius of ρ for a frictionless clamped die

$$\frac{\mathrm{d}\gamma}{\mathrm{d}\theta}=\sqrt{K}C_1\exp(\sqrt{K}\theta)-\sqrt{K}C_2\exp(-\sqrt{K}\theta) \tag{11}$$

For pure bending the situation is illustrated schematically in Fig. 6. It is assumed that the sheet is bent to a circular arc in a frictionless die. The boundary conditions for this system are, by symmetry, $\gamma=0$, when $\theta=0$, and $T=0$, when $\theta=\alpha$.

The constants can then be determined using Eqs. (5), (10), and (11), and the boundary conditions to give

$$C_2=-C_1 \tag{12}$$

$$C_1=\frac{(a+b)}{b}(\sqrt{K}\{\exp(\sqrt{K}\alpha)+\exp(-\sqrt{K}\alpha)\})^{-1} \tag{13}$$

4 Bending Moment and Tension in the Strip

4.1 Calculation of Tension Over the Cover Sheet. The tension, T, acting on the cover sheet is calculated by rearranging Eq. (5); $\mathrm{d}\gamma/\mathrm{d}\theta$ can be determined from the linear approximation given in Eq. (11). This results in the following relationship:

$$T(\theta,\rho)=\frac{2aE'}{\rho}\left\{(a+b)-b\frac{\mathrm{d}\gamma}{\mathrm{d}\theta}\right\}. \tag{14}$$

4.2 Bending Moment of the Laminate Material. Assuming that the bending moment on the interlayer is small relative to the cover sheets and can be neglected, the bending moment on the laminate is composed of the moment to bend each cover sheet plus the moment due to the tensions on each cover sheet. Figure 7

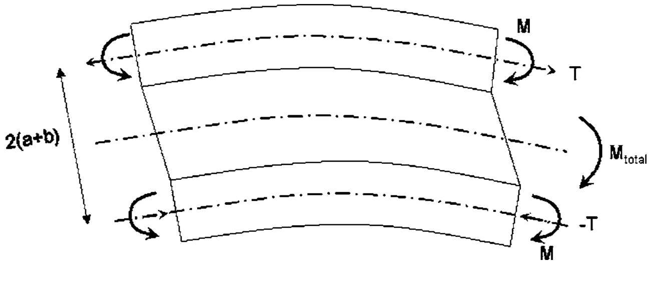


Fig. 7 Moments of the upper and lower cover sheets and the overall moment on an element section of the SPS laminate material

Table 1 Material and process conditions used in the analysis

Variable	Base Value	Range
Bend radius, ρ, mm	500	300–700
Interlayer shear modulus, G, MPa	300	180–420
Interlayer thickness, $2b$, mm	0.6	0.4–0.8
Cover sheet thickness, $2a$, mm	0.22	0.12–0.32
Fixed Values		
Cover Sheet Young's modulus, E', GPa	200	
Strip length, mm	150	

shows the bending moment of the laminate, where M_{upper} and M_{lower} are the bending moments of the cover sheets and the total bending moment is given by

$$M_{\text{total}} = M_{\text{upper}} + M_{\text{lower}} + 2(a+b)T(\theta,\rho) \tag{15}$$

The moments in the cover sheets about the midsurface of each sheet, which is also the line of action of the force T, is determined from the equilibrium equation,

$$M_{\text{cover}} = \int_{-a}^{a} \sigma_i y dy \tag{16}$$

which leads to

$$M_{\text{upper}} = 2\frac{E'}{(\rho+a+b)} \cdot \int_0^a y^2 dy \tag{17}$$

and

$$M_{\text{lower}} = 2\frac{E'}{(\rho-a-b)} \cdot \int_0^a y^2 dy \tag{18}$$

Assuming that $\rho \gg (a+b)$ and $(\rho+a+b) \approx (\rho-a-b) \approx \rho$, then the moment in each cover sheet is approximately

$$M_{\text{upper/lower}} = \frac{2E'a^3}{3\rho}, \tag{19}$$

and the overall moment is given by

$$M_{\text{total}} = 2M_{\text{upper/lower}} + 2(a+b)T(\theta,\rho) \text{ or}$$

$$M_{\text{total}} = 2\left(\frac{2E'a^3}{3\rho}\right) + 2(a+b)T(\theta,\rho) \tag{20}$$

5 An Example

The model is used to analyze the influence of the important process parameters of the bending process on the shear deformation of the interlayer material and in that way the shear stress at the interface interlayer/cover sheet. An arbitrary elastic bend was chosen as a base as shown in Table 1.

5.1 Shear Deformation. The shear angle of the interlayer over the sheet length for varying conditions indicated in Table 1 has been calculated using Eq. (10) and is shown in Fig. 8 for the change of bend radius and shear modulus. The shear deformation of the interlayer is very small towards the center of the bend and increases rapidly towards the ends of the SPS sheet. The reason for this rapid increase in angle is the decreasing tension in the top cover sheet or decreasing compression in the bottom cover sheet; the tension/compression in the cover sheets minimizes the shear deformation in the interlayer by reducing the in-plane distance between corresponding cover sheet points; for example, points A and C or points B and D in Fig. 3. The rate of change of the tension/compression in the cover sheets is primarily dependent on the shear modulus of the interlayer. A higher shear modulus of the interlayer material causes a decrease in the shear deformation because of increased shear stresses at the interface interlayer/cover sheet. This causes the cover sheets to deform in tension (compression) while the interlayer deforms less (Fig. 8). If the shear modulus of the interlayer is zero, then the shear angle is a linear function of the length along the strip from the center line. For this case, the interlayer is unable to create any shear stress on the steel-polymer interface to reduce the increase in shear deformation. A smaller bend radius results in an increased shear angle and with decreasing interlayer thickness the shear deformation decreases. Thicker cover sheets lead to higher shear deformation in the interlayer because the shear stresses necessary to deform the cover sheets are higher, which leads to more deformation of the interlayer.

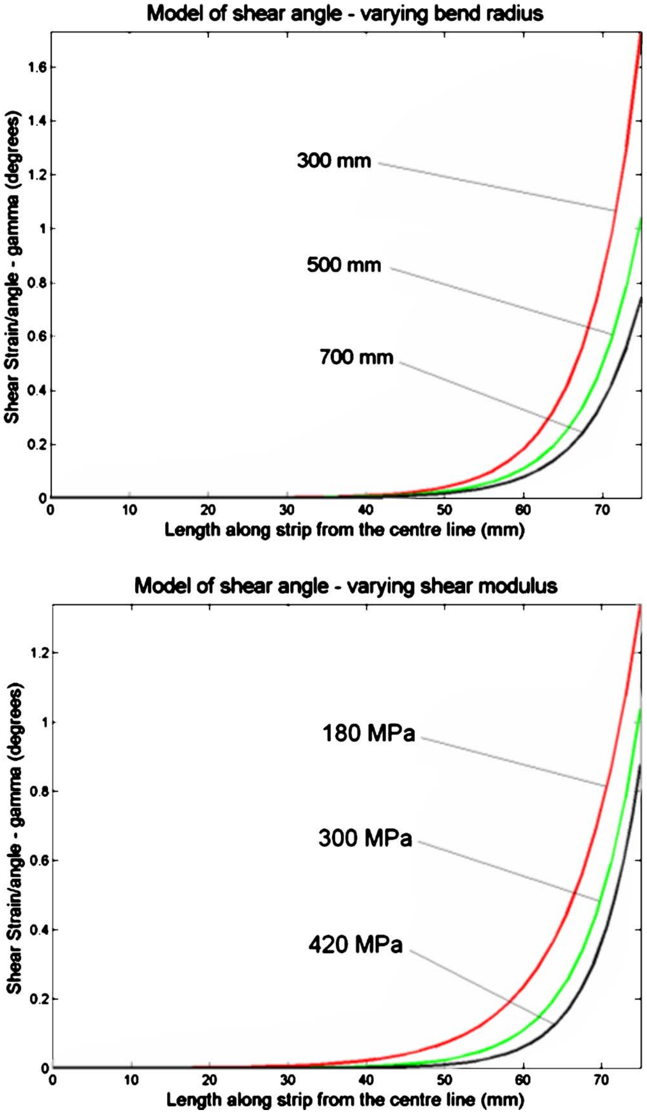


Fig. 8 Plots of the shear strain within the interlayer using the conditions specified in Table 1 for various bend radius and shear modulus

5.2 Tension. Based on the above model the tension force acting in the cover sheet as a result of the introduced shear stresses at the interlayer/cover sheet interface was determined to provide a moment diagram for the sheet. The tension in the cover sheet as a function of the sheet length is shown in Fig. 9. It is calculated from Eq. (14) using the base values shown in Table 1. To illustrate the influence of the shear deformation of the interlayer material on

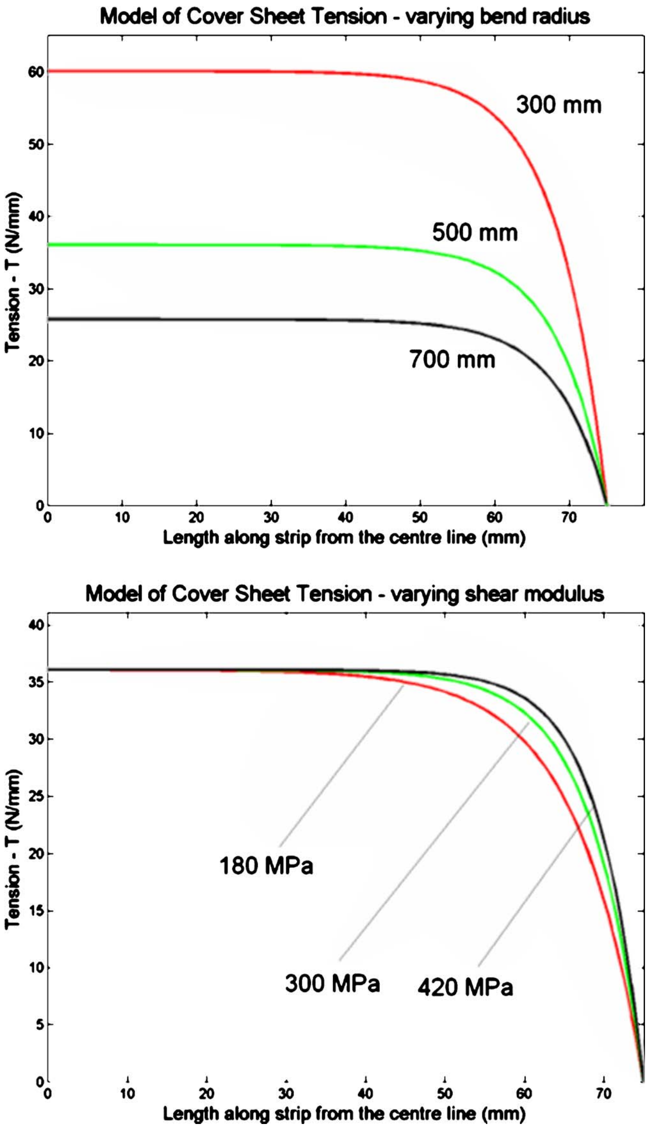


Fig. 9 Tension in the top cover sheet when varying the bend radius and shear modulus

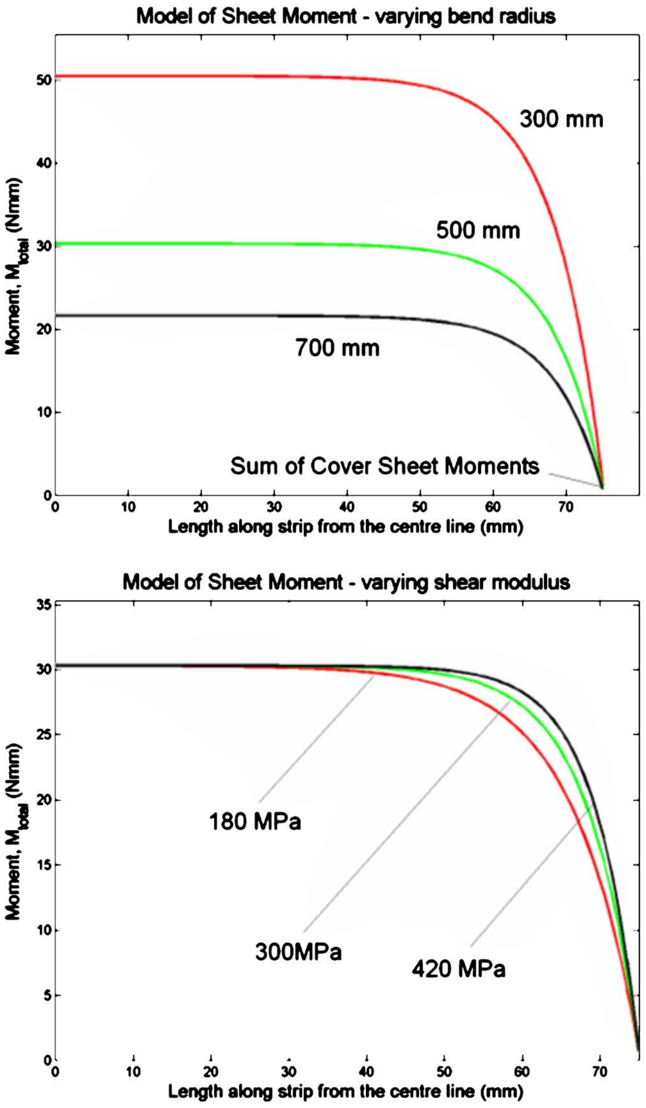


Fig. 10 The moment diagram of the laminate SPS material bent to a circular arc for different values of the bend radius and the shear modulus

tension in the cover sheets, the two major parameters influencing the shear deformation of the interlayer, bend radius and the shear modulus, were varied in the tension calculation. It can be seen that the tension force acting in the cover sheets has a maximum at $\theta=0$ and decreases to zero at the end of the SPS sheet. Therefore, as the tension along the sheet decreases, the shear deformation in the interlayer increases. The tension stretches the upper sheet and shrinks the lower cover sheet to decrease the difference between the corresponding points on the two cover sheets. This reduces the shear deformation where the tension is high, and increases the shear strain where the tension is low. Decreasing bend radius increases tension in the cover sheets. Compared to the change in shear deformation in the interlayer material (Fig. 8), the effect of bend radius on the tension is more significant. Here basically an increase in the maximum tension value can be seen. In contrast, as the shear modulus of the interlayer material decreases, the tension in the cover sheets decreased because of the lower shear stresses generated at the interface interlayer/cover sheet.

5.3 Bending Moment. To validate the assumption of a zero bending moment in the core material (Sec. 4.2) the material and process conditions of Table 1 have been used to determine the bending moment of the cover sheets and the core in the absence of any tension. The calculation results show that the moment resistance of the interlayer only accounts for approximately 4% of the moment required to bend the two cover sheets and can be therefore neglected in the calculation of the bending moment of the laminated strip. The calculated bending moment of the laminated strip as a function of the sheet length using Eq. (20) is shown in Fig. 10.

It can be seen that the bending moment has a maximum at the middle of the bend and it decreases towards to the end of the sheet. Notice that the bending moment reaches a constant value not equal to zero at the end of the strip, which equals the sum of the pure bending moment for both cover sheets. The change in bending moment over the sheet length is a result of the tension force acting on the cover sheets. The tension decreases to zero to the end of the sheet leaving only the pure moment to curve the cover sheets on the sheet end. This shows that in contrast to homogeneous sheet, for laminated sheet a bend to a constant radius of curvature does not result in a constant moment but a moment that changes along the sheet length. This result suggests that a pure moment applied on a SPS sheet would not lead to a perfect but an irregular circle of arc, which could lead to differences in shape after forming compared to homogeneous sheet material.

6 Conclusion

The use of laminated sheet is likely to increase in auto body construction and satisfactory design tools will be required. This paper only takes a very limited case of elastic bending to constant curvature, but it is hoped that the basic formulation is capable of further development. Within this limited scope it has been found that in SPS laminated sheet, the bending moment is not constant over the sheet length but decreases from a maximum in the center of the bend to a constant value towards the edge of the sheet. This change in bending moment is a result of the shear deformation of the interlayer material introducing shear stresses at the interface of the cover sheet and the interlayer. This causes tension forces in the cover sheets and results in an additional and nonuniform bending moment term. The key variables affecting the shear strain and resulting nonuniform moment appear to be the shear modulus of the interlayer or the core thickness. The results from our model also suggest that a constant moment applied on an SPS sheet would lead to a nonuniform curvature in contrast to that of a homogeneous sheet.

Although the equations developed are yet to be verified by experiments, the results could be useful for a better understanding of the forming behavior of SPS laminates in the bending process.

Acknowledgment

Many thanks must go to Professor Peter Hodgson of Deakin University, for his support of this work and advice in the preparation of the manuscript.

Nomenclature

E' = plane strain elastic modulus of the cover sheet (MPa)
G = shear modulus of the polymer core (MPa)
T = tension in the cover sheet (N/mm)
a = half the thickness of the cover sheet (mm)
b = half the thickness of the polymer core (mm)
ds = length of an element at the midsurface (mm)
ρ = midsurface radius of curvature of the bend (mm)
α = final bend angle (radians)
θ = angle subtended by part of the strip (radians)
$d\theta$ = angle subtended by an element across the midsurface of the polymer core
γ = shear strain of the polymer core (radians), it is a function of bend angle
$d\gamma$ = change in shear strain along an the element
$d\varphi$ = angle subtended by an element across the cover sheet
τ = shear stress between the cover sheet and the polymer core (MPa)
σ_a = stress at the midsurface on an element of the cover sheet (MPa)
ε_a = strain at the midsurface on an element of the cover sheet
ε_b = bending strain on the cover sheet
σ_1 = stress on an element along the sheet direction (MPa)
ε_1 = strain on an element along the sheet direction

References

[1] DiCello, J. A., 1980, "Steel-Polypropylene-Steel Laminate—A New Weight Reduction Material," SAE Technical Paper Series, 800078.
[2] Yang, B., Nunez, S. W., Welch, T. E., and Schwaegler, J. R., 2001, "Laminate Dash Ford Taurus Noise and Vibration Performance," SAE Technical Paper Series, 2001-01-1535.
[3] Link, T. M., 2001, "Formability and Performance of Steel-Plastic-Steel Laminated Sheet Materials," SAE Paper Series, 2001-01-0079, pp. 31–38.
[4] Kleiner, M., and Hellinger, V., 1999, "New Possibilities for Improved Bending of Vibration Damping Laminated Sheets," Ann. ICRP, **48**(1), pp. 217–220.
[5] You-Min, H., and Daw-Kwei, L., 1995, "Finite-Element Simulation of the Bending Process of Steel/Polymer/Steel Laminate Sheet," J. Mater. Process. Technol., **52**(2), pp. 319–337.
[6] Yin, S., Mason, J., Schmidt, S., and Corona, E., 2002, "Bending and Springback of Laminated Steel," JSME/ASME International Conference on Materials and Processing, pp. 159–163.
[7] Kim, J., and Swanson, S. R., 2001, "Design of Sandwich Structures for Concentrated Loading," Compos. Struct., **52**, pp. 365–373.
[8] Gosh, A. K., and Dey, S. S., 1992, "A Simple Finite Element for the Analysis of Laminated Plates," Comput. Struct., **44**(3), p. 585.
[9] Frostic, Y., 1992, "Higher-Order Theory for Sandwich-Beam Behaviour with Transversely Flexible Core," J. Eng. Mech., **118**, pp. 1026–1043.
[10] Yoshida, F., 1990, "Elastic-Plastic Analysis for Bending of Metal/Polymer/Metal Sandwich Plate," Advanced Technology of Plasticity, **3**, pp. 1513–1518.
[11] Takiguchi, M., and Yoshida, F., 2001, "Plastic Bending of Adhesive-Bonded Sheet Metals," J. Mater. Process. Technol., **113**(1–3), pp. 743–748.
[12] Takiguchi, M., and Yoshida, F., 2003, "Analysis of Plastic Bending of Adhesive-Bonded Sheet Metals Taking Into Account the Visco-Plasticity of Adhesive," J. Mater. Process. Technol., **140**(1–3), pp. 441–446.
[13] Liu, L., and Wang, J., 2004, "Modeling Springback of Metal-Polymer-Metal Laminates," J. Manuf. Sci. Eng., **126**, pp. 599–604.

Stability and Andronov-Hopf Bifurcation of Steady-State Motion of Rotor System Partly Filled With Liquid: Continuous and Discrete Models

N. V. Derendyaev
Nizhny Novgorod State University,
Faculty of Computational Mathematics and
Cybernetics,
Gagarin Avenue 23,
603022 Nizhny Novgorod,
Russia
e-mail: der@tudm.unn.ac.ru

A. V. Vostrukhov
e-mail: vostroukhov@yahoo.co.uk

I. N. Soldatov
e-mail: wvs@dynamo.nnov.ru

Mechanical Engineering Institute
of Russian Academy of Sciences,
Nizhny Novgorod branch,
Belinskogo 85,
603024 Nizhny Novgorod,
Russia

In this paper, a new method for investigation of dynamics of fluid-filled rotor systems is presented. The method consists of development of finite degrees-of-freedom (discrete) models for the rotor systems. The discrete models are physically justified and demonstrative. Being described by the system of ordinary differential equations, they allow one to employ powerful tools of the theoretical mechanics and oscillation theory. The method is applied to the case of the plane model of the rotor system partly filled with incompressible liquid. Both the continuous and discrete models are considered. The main attention is paid to the latter model. The discrete model consists of a disk symmetrically fixed on the shaft (Laval scheme), the ends of which are in viscoelastic bearings, and a ring sliding over the disk with friction. The centers of the disk and ring are elastically connected. The disk models the rotor, while the ring describes the liquid filling. When the ring is sliding over the disk surface, an interaction force arises that is diverted from the direction of the relative velocity at the contact points. It is demonstrated that an appropriate choice of the parameters of the discrete model allows one to determine the stability domain of the steady-state rotation of the rotor in the plane of the parameters of the shaft bearings with an excellent accuracy. It is found out that when the parameters overstep the limits of the stability domain, the Andronov-Hopf bifurcation occurs: a periodic motion of a kind of a circular precession arises from the steady-state rotation regime either "softly" or "hardly." [DOI: 10.1115/1.2164514]

Introduction

Nonconservative problems of dynamics of rotor systems partly or completely filled with a stratified viscous fluid belong to classic problems of mechanics. This class of problems was drawing attention of researchers for more than a century and a half being a fundamental issue in many practical applications, amongst which one can call rotors, turbines, gyroscopic devices, ultracentrifuges, flying apparatuses, and other transport means containing cavities with fluid (fuel or liquid freight).

In the 1950–1960s, the problems of the dynamics of hollow solids containing liquid received a significant theoretical and experimental treatments by a great number of authors. The basic reason for the developing was a rapid introduction of the space rocketry and new generation of flying apparatuses. Moiseev and Rumyantsev [1] and Mikishev and Rabinovich [2] published the overviews devoted to dynamics of solids containing cavities with fluid.

The instability of a rotor partially filled with fluid was first observed by Epishev [3], Crandall [4], and Kollmann [5]. These authors discovered experimentally that the rise of the instability of the steady-state rotation leads to a circular precession at a frequency close to the rotor whirl frequency.

Attempts to describe this fact by means of two-dimensional (2D) models for rotors were made by Ehrich [6] and Wolf Jr. [7]. Ehrich included viscosity of the fluid but assumed a restricted form for the fluid motion. Wolf Jr. used a plane model of a precessing rotor partly filled with ideal incompressible liquid. He found out that the instability is directly related to appearing of a perturbed motion in the form of circular precession. However, the frequency of the circular precession turned out quite different from that of the rotational velocity.

Daich and Bar [8] and Saito and Someya [9] extended the model of Wolf Jr. by accounting for the viscosity of the fluid. The former authors, having used the method of the boundary layer, showed that the introduction of the viscosity changes the results of Wolf Jr. insignificantly, at least where the method of the boundary layer is applicable. The latter authors came up to the analogous conclusion. They used an exact solution to the plane problem of the motion of a viscous incompressible liquid in a precessing revolving rotor. But they applied a conservative criterion of stability to the nonconservative problem. It was assumed that the rise of the instability is related to the appearing of a multiple imaginary root of the characteristic equation.

In 1982, Lichtenberg [10] extended the model of Wolf Jr. for the three-dimensional case. But he also could not obtain coincidence of the theoretical predictions with the experimental results. Holm-Christensen and Träger [11] directly used the full Navier-Stokes equations and solved them numerically. But the procedure was quite time-consuming and sensitive to the initial assumptions.

In 1982, Derendyaev and Sandalov [12] worked out a new method for investigation of stability of a steady rotation of the rotor systems with fluid. Making use of this method, they demonstrated theoretically that a perturbed motion in the form of a precession arises at the boundary of the stability domain in the space of the system parameters. If the external damping is small enough,

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 29, 2004; final manuscript received November 4, 2005. Review conducted by M. P. Mignolet. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

the frequency of the precession is close to the rotational velocity. The value of the critical speed of the rotor corresponds to that found by Epishev [3] and Crandall [4]. In 1988, Derendyaev [13] demonstrated that Andronof-Hopf bifurcation may occur in a liquid-filled rotor system. In his work [14], he analyzed the methodological mistakes made by Wolf Jr. and Saito and Someya.

From recent publications on the dynamics of rotors with fluid papers of Zhu Changsheng [15] and Derendyaev and Soldatov [16] should be mentioned. The former author experimentally investigated the instability of an overhung rigid centrifuge rotor with fluid making an emphasis on the study of stable and unstable regions in the plane of rotational speed and fluid-fill ratio. However, this experimental work was not founded theoretically. The latter authors made the next step in the study of dynamics of rotor systems with fluid. They showed that the distributed model of the rotor with fluid can be effectively replaced by a finite degrees-of-freedom one. On the basis of the developed finite degrees-of-freedom model that allows employing the classic methods of the oscillation theory, the authors investigated the stability of the rotor and showed that the discrete model gives the same results as the continuous model.

Analogous but separate development has been performed by Dosaev and Samsonov who proposed a discrete dynamic model for hollow solid body containing viscous liquid [17].

Summarizing all of the above-mentioned one can say that although the rotor systems with fluid were extensively studied, many important aspects of the rotor dynamics, such as self-excited vibrations, their stability, bifurcation processes, transition to chaotic oscillations, did not receive an exhaustive treatment. The basic reason is that the fluid is a multi degrees-of-freedom (distributed) system, for which it is very problematic to employ classic methods of the oscillation theory. The main disadvantage of the Galerkine's discretization method and the finite element methods (FEM) that are widely used nowadays is that both procedures are generally unfounded and not physically transparent enough. For instance, the direct calculation used by FEM is certainly not the best solution for the problem at hand. Besides this, it requires an immense amount of calculation time. Both methods have to be checked against experiments only, i.e., they are not reliable at the design stage. This implies that other methods, which would allow one to unveil the global dynamics of the rotor systems with fluid, are strongly necessary. Development of such methods could open wide perspectives in the analysis of the dynamics of fluid-filled rotor systems for researchers and engineers.

This paper is a development of work [16]. The novelty of the paper consists of the study of the Andronov-Hopf bifurcation and more detailed consideration of the instability phenomenon. The paper is structured as follows. First, the continuous model for the rotor partly filled with liquid is considered. The exact solution to the problem is presented. Then, the finite degrees-of-freedom analog of the system is described. The results obtained for both continuous and discrete models are compared. Third, on the base of the discrete model, the Andronov-Hopf bifurcation is investigated. Finally, conclusions and perspectives of the proposed method are addressed in the Discussion and Conclusions.

The Continuous Model

In this section, the continuous model of a rotor partly filled with incompressible liquid is shortly presented. The basic computations and particulars can be addressed to papers [12,18,19].

Let us consider the steady-state vibration of a cylindrical rotor partly filled with liquid. The rotor is symmetrically mounted on two rigid shafts the ends of which are viscoelastically supported in two roller bearings (see Fig. 1).

The system is analyzed under the following assumptions [12]:

(1) The system moves with constant angular velocity Ω around axis Ox_3 of the fixed right-handed Cartesian coordinate system $Ox_1x_2x_3$.

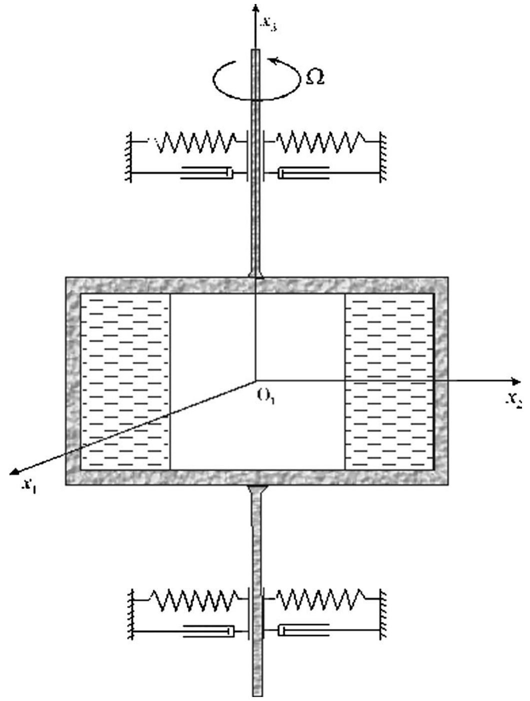


Fig. 1 Cross section of rotor partly filled with liquid

(2) The liquid is incompressible and concentrated in the layer of width h.
(3) On the strength of the high rotational speed the gravity force is neglected.
(4) The angular displacements of the shafts are negligibly small, i.e., all particles of the liquid move in the plane that is normal to the rotational axis.
(5) The cylinder is long enough, i.e., the edge effects are negligible.
(6) In the steady-state regime the liquid and the cylinder rotate as a single whole.

With all above-mentioned assumptions, the governing equations of motion and the boundary conditions can be written as follows:

The equations of motion for the rotor:

$$M_r\ddot{z} = F + f$$

$$\dot{\theta} = \Omega = \text{const} \tag{1}$$

with M the mass of the cylinder, $z=x_1^0+ix_2^0$ the complex variables ($i^2=-1$), x_1^0,x_2^0 the coordinates of the cross point of the rotor axis with plane Ox_1x_2, θ the rotational angle, $F=F_1+iF_2$ the complex force determined as $F_k=-\iint_S \sigma_{kj}n_j ds$, n_j the components of the outward unit vector $\mathbf{n}$ normal to the internal surface S of the rotor,

$$\sigma_{kj} = -p\delta_{kj} + \mu\left(\frac{\partial v_k}{\partial x_j} + \frac{\partial v_j}{\partial x_k}\right), \quad j,k = 1,2 \tag{2}$$

the components of the stress tensor, p the pressure, μ the coefficient of the dynamic viscosity of the liquid, v_k the components of the velocity vector $\mathbf{v}$ of the liquid particles, $f=-K_r z-H_r\dot{z}$ the reaction force of the bearings, where K_r and H_r are the stiffness and the damping coefficients of the bearings, respectively.

The equations of motion for the viscous incompressible liquid:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\,\nabla)\mathbf{v} = -\frac{\nabla p}{\rho} + \upsilon \Delta \mathbf{v}, \quad \operatorname{div} \mathbf{v} = 0 \tag{3}$$

with ρ the mass density of the liquid and $\nu = \mu/\rho$ the coefficient of kinematic viscosity.

The boundary conditions:

$$\left[\frac{\partial \phi}{\partial t} + (\mathbf{v}\,\nabla\,\phi)\right]_{\phi=0} = 0,$$

$$[\mathbf{v}]_S = [\mathbf{V}]_S,$$

$$\left[\sigma_{kj}\frac{\partial \phi}{\partial x_j}\right]_{\phi=0} = -p_*\frac{\partial \phi}{\partial x_k} \tag{4}$$

where $\phi(x_1, x_2, t) = 0$ is equation of the free surface of the liquid, $\mathbf{V}$ is the velocity of the rotor's surface at point $\mathbf{x} \in S$, p_* is the pressure at the free surface.

In 1949, Yuri I. Neimark [20] developed a powerful tool for parametric analysis of physical systems described by a set of linear differential equations. In literature, his method is referred to as the *D-decomposition method* (see [27], for instance). An interested reader may find a detail description of the method in reference [26]. Neimark demonstrated that the space of the system parameters can be decomposed into domains with different number of the "unstable" eigenvalues (eigenvalues with positive real parts that correspond to exponential growth). Knowing this number in any point of the parameter space allows for a direct determination of the number of "unstable" eigenvalues for any parameter set of the system.

Mathematically, the D-decomposition procedure is completely analogous to a conformal mapping. The idea of this method is to map the imaginary axis of the complex plane of a system eigenvalue (corresponding to the boundary between the stability and instability domains) into the space/plane of the system parameters. The mapping rule follows from the characteristic equation, which should be written to express the chosen parameters explicitly. Once the mapping is accomplished, a mapped line (*D-decomposition curve* or *D-curve*) is obtained, which divides the parameter plane into domains with different number of roots possessing a positive real part. It is customary to call this number *degree of instability*, whereas the corresponding domains are referred to as *domains with different degree of instability*.

Physically, the term "degree of instability" implies a certain dynamic regime. If the system parameters belong to a domain with zero degree of instability, then the system is stable (natural revolution of a rotor, for instance). If not, the system gets unstable. The degree of instability defines peculiarities of the oscillation build-up process. Crossing the D-curve (that is equal to a change in the system parameters) implies transition from one regime to another. This may be accompanied by a bifurcation process.

In work [12], it has been demonstrated that instability of the steady rotation is directly connected with the precession regime. The values of the parameters, at which the precession is possible, determine the boundaries of domains with different degree of instability in the space of the system parameters. In our paper, we will follow the methodology developed in Refs. [12,20].

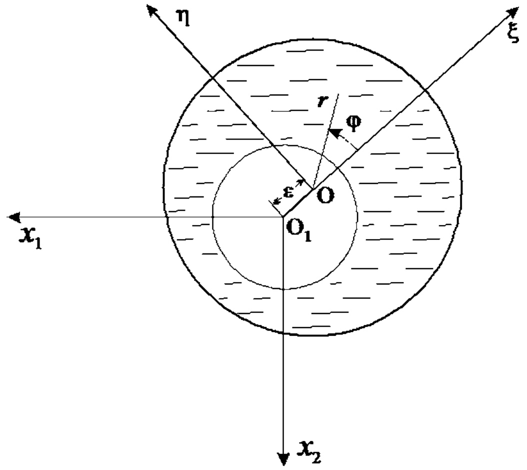


Fig. 2 Noninertial reference system

Let the rotor perform a precessing motion of small radius ε with frequency ω. We introduce the moving coordinate system $O\xi\eta$ in the following way: (1) the origin point O coincides with the center of the cross section of the cylinder; (2) axis $O\xi$ coincides with the junction line of the precession center O_1 and origin O (see Fig. 2). In the moving coordinate system we introduce the polar coordinate system $Or\varphi$. In the steady-state regime, the liquid rotates with the rotor as a single whole, i.e.,

$$\mathbf{v} = \mathbf{v}_0 = \omega_0 r \mathbf{e}_\varphi, \quad p = p_0 = p_* + \frac{\rho\Omega(r^2 - b^2)}{2} \tag{5}$$

where $\omega_0 = \Omega - \omega$ is the angular frequency of the natural rotation of the rotor, $b = a - h$ is the internal radius of the cylindrical liquid layer in the steady-state rotation regime, a is the internal radius of the cylinder.

The solution to the problems (1)–(4) linearized around condition (5) can be represented in the form [12]

$$u = 2\,\mathrm{Re}\left[\left(\frac{\omega_0 + 2\Omega}{\omega_0}c_1 - \frac{\omega_0 - 2\Omega}{\omega_0}\frac{c_2}{r^2} + i\frac{Z_1(kr)}{r}\right)e^{i\varphi}\right]$$

$$\upsilon = 2\,\mathrm{Re}\left[\left(i\frac{\omega_0 + 2\Omega}{\omega_0}c_1 + i\frac{\omega_0 - 2\Omega}{\omega_0}\frac{c_2}{r^2} - kZ_0(kr) + \frac{Z_1(kr)}{r}\right)e^{i\varphi}\right]$$

$$\frac{p}{\rho} = \frac{p_*}{\rho} + \frac{1}{2}\Omega^2(r^2 - b^2) + 2\,\mathrm{Re}\left[\frac{\omega^2\varepsilon r}{2} + i\frac{4\Omega^2 - \omega_0^2}{\omega_0}\left(c_1 r + \frac{c_2}{r}\right) - 2\Omega Z_1(kr)\right] \tag{6}$$

where u and υ are the radial and azimuth components of the velocity field, $Z_n(kr) = c_3 e^{-\kappa a}H_1^{(2)}(kr) + c_4 e^{\kappa b}H_1^{(1)}(kr)$, $H_n^{(1)}$ and $H_n^{(2)}$ are the Hankel functions of order n, $k = \kappa(i - (\omega_0/|\omega_0|))$, $\kappa = \sqrt{|\omega_0|/2\nu}$, and c_i, $i = \overline{1,4}$ are unknown constants that can be found from the boundary conditions:

$$\frac{3-\tau}{1-\tau}c_1 + \frac{1+\tau}{1-\tau}\frac{c_2}{a^2} + \frac{i}{a}Z_1(ka) = 0$$

$$i\frac{3-\tau}{1-\tau}c_1 - i\frac{1+\tau}{1-\tau}\frac{c_2}{a^2} - kZ_0(ka) + \frac{1}{a}Z_1(ka) = 0$$

$$\frac{1+\tau}{1-\tau}\frac{4i}{b^3}c_2 + \frac{2k}{b}Z_0(kb) + \left(k^2 - \frac{4}{b^2}\right)Z_1(kb) = 0$$

$$-\frac{\tau^2(3-\tau)}{(1-\tau)^2}ibc_1 + \frac{i}{b}(1+\tau)\left(\frac{\tau^2 - 4\tau + 2}{(1-\tau)^2} - \frac{4}{k^2b^2}\right)c_2 - 2\frac{1-\tau}{kb}Z_0(kb) + \left(\frac{2\tau - 1}{1-\tau} + 4\frac{1-\tau}{k^2b^2}\right)Z_1(kb) = -\frac{1}{2}\Omega\varepsilon b\tau^2 \tag{7}$$

with $\tau = \omega/\Omega$.

Integrating stresses (2) applied to the internal surface of the rotor, one obtains expressions for the components of the hydrodynamic force F per unit length:

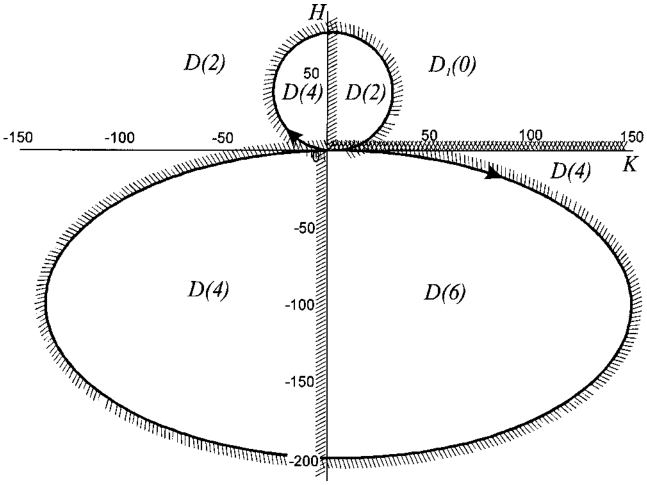


Fig. 3 ***D*-decomposition of the plane of the parameters of the bearings**

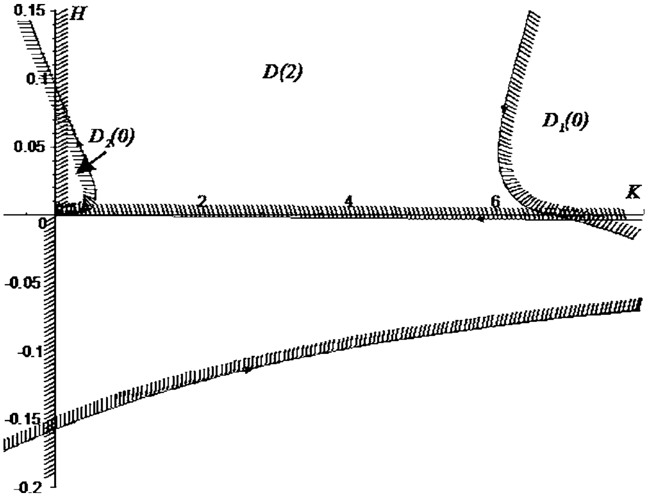


Fig. 4 ***D*-decomposition of the plane of the parameters of the bearings: expanded scale**

$$F_\xi = 2\pi a\rho \,\mathrm{Re}\left[\frac{1}{2}\omega^2\varepsilon a + 2i\frac{(\Omega+\omega)}{a}c_2\right]$$

$$F_\eta = -4\pi\rho(\Omega+\omega)\mathrm{Re}(c_2) \tag{8}$$

Substituting now expressions (8) into the first terms of Eqs. (1) and introducing the following dimensionless variables:

$$K = \frac{K_r}{m_l\Omega^2}, \quad H = \frac{H_r}{m_l\Omega}, \quad m = \frac{M}{M+m_l}, \quad E = \frac{\nu}{\Omega a^2},$$

$$F_\xi^* = \frac{F_\xi}{m_l\omega^2\varepsilon}, \quad F_\eta^* = \frac{F_\eta}{m_l\omega^2\varepsilon}, \quad \delta = \frac{b}{a} \tag{9}$$

with $m_l = \pi\rho(a^2 - b^2)$ the mass of the liquid filling, E the Ekman number, and δ the fluid-fill ratio, we obtain the relationships for the system parameters in the case of the circular precession:

$$K - \frac{m}{1-m}\tau^2 = F_\xi^*\tau^2$$

$$H\tau = F_\eta^*\tau^2 \tag{10}$$

The dimensionless components of the hydrodynamic force F_ξ^*, F_η^* depend on parameters τ, E, δ, and m only. When the latter three parameters are fixed, Eqs. (10) define a curve in the plane (K,H), the points of which correspond to such values of the parameters when the circular precession can occur. This curve, in accordance with the above-made notes, divides the plane (K,H) into domains with different degree of instability. Following Ref. [20] we will call it the D-curve.

Figure 3 represents the D-decomposition (solid curve with shading) of the plane (K,H) for $E=10^{-5}$, $\delta=0.9$, $m=0.6$. The arrows along the curve point out the direction of increase of the parameter τ. The D-curve is formed by the regular branch, along which the parameter τ changes from $-\infty$ to $+\infty$, and the singular straight line $K=0$ corresponding to $\tau=0$. The existence of the singular straight line is conditioned by the fact that the components of the hydrodynamic force are equal to zero when $\tau=0$ (no precession).

Due to the axial symmetry of the problem the degree of instability of the system is always an even number. This is why it is reasonable to shade the D-curve in the following way: a passing through the boundary of domain $D(n)$ from the unshaded side to the shaded one corresponds to an increase of the degree of instability by 2, i.e., it is equivalent to a passing to the domain $D(n+2)$. The stability domain is designated as $D(0)$. For the case at hand, there are two stability domains: $D_1(0)$ and $D_2(0)$. The first domain comprises the point corresponding to the infinitely large damping coefficient (see Fig. 3). The second one abuts upon the origin of the plane (K,H). It is not visible in Fig. 3 and shown in Fig. 4 in a bigger scale.

Figures 5 and 6 demonstrate the effect of the Ekman number E and the filling parameter δ on the sizes of the domains with different degree of instability. For the sake of convenience, the shading is skipped in Fig. 5. The dashed line corresponds to values $E=0.5\times10^{-5}$, $\delta=0.9$, $m=0.6$, while the dashed-dotted line corresponds to $E=10^{-5}$, $\delta=0.7$, $m=0.6$. As is seen from the pictures, the decrease of the Ekman number leads to the general distension of the figures that leads to a decrease of the stability domains $D_1(0)$ and $D_2(0)$. The decrease of the fluid-fill ratio distends mainly the lower part of the figure and somewhat draws up, slightly constricting, the upper oval.

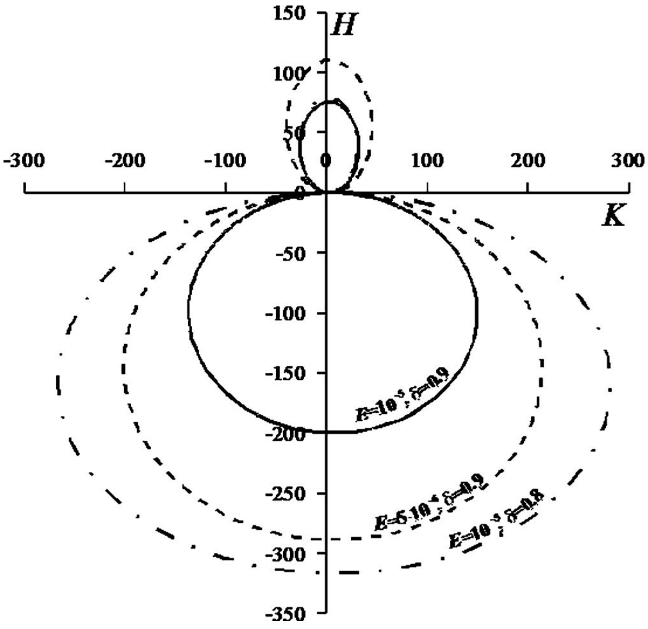


Fig. 5 **Effect of the Ekman number and the fluid-fill ratio on *D*-decomposition of the plane of the parameters of the bearings**

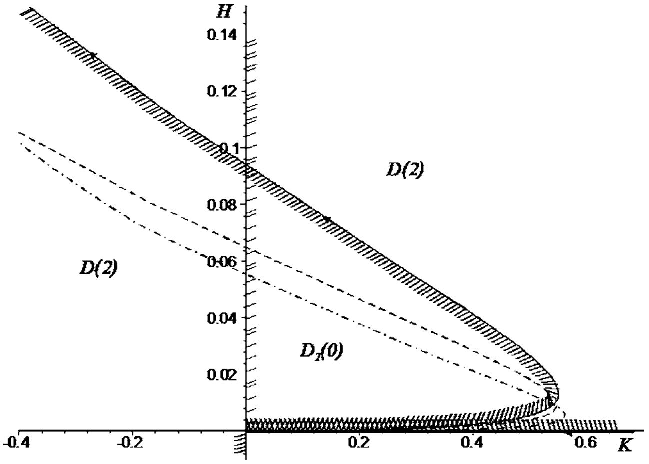


Fig. 6 Effect of the Ekman number and the fluid-fill ratio on the boundaries of $D_2(0)$ domain

The Finite Degrees-of-Freedom (Discrete) Model

This section presents the discrete analog for the model of the rotor partly filled with fluid.

The model consists of a disk of mass m_d and radius R_d, symmetrically mounted on an isotropic absolutely rigid shaft, which is supported into two viscoelastic roller bearings according to the Laval scheme, and of a ring of mass m_r and radius R_r sliding over the disk surface with friction (see Fig. 7). The centers of the disk and of the ring are connected elastically. The disk models the rotor, while the ring models the liquid filling.

The further assumptions adopted are:

(a) When the ring is sliding over the disk, in the contact points an interaction force arises that is proportional to the relative velocity and not collinear to it.

(b) The disk is in nonlinear viscoelastic fixation on the shaft AA' and rotates with constant angular velocity Ω in the horizontal plane.

(c) All points of the system move in the horizontal plane only, i.e., the velocity field does not depend on the coordinate along the axis of the steady-state rotation.

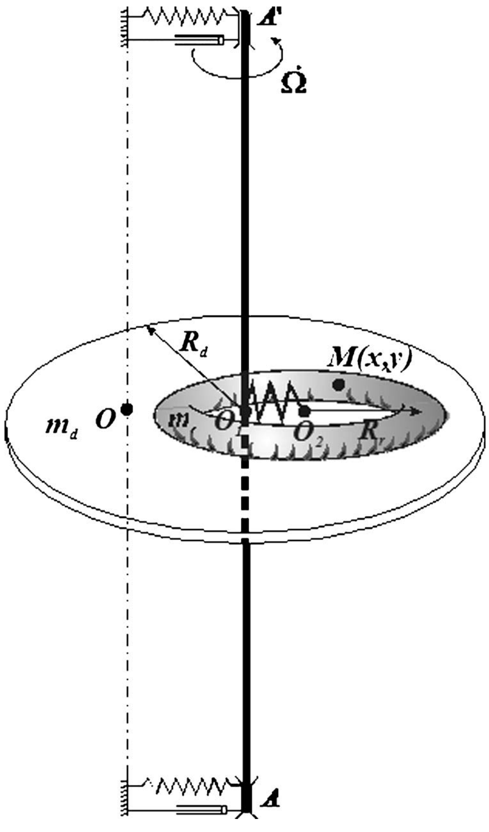


Fig. 7 Discrete model of the rotor with liquid

We introduce two right-handed reference frames: the fixed coordinate system $Oxyz$, the axis Oz of which is parallel to the shaft AA', and the moving coordinate system $O_1\xi\eta$ fixed on the disk. The origin of the fixed frame lays on the axis of the steady rotation. The point O_1 coincides with the disk center. We designate the center of the ring as O_2.

In accordance with all the above-mentioned assumptions, the disk experiences affection of the force:

$$\mathbf{f}_d = -(k_d + k_{2d}|\mathbf{r}_{01}|^{\alpha})\mathbf{r}_{01} - (\eta_d + \eta_{2d}|\mathbf{r}_{01}|^{\beta})\mathbf{r}_{01} + \mathbf{F}_{\text{int}} \qquad (11)$$

with $\mathbf{r}_{01}=x_1\mathbf{e}_x+y_1\mathbf{e}_y$ the radius vector of the point O_1 in the fixed reference frame; k_d, η_d and k_{2d}, η_{2d} the linear and nonlinear parts of the stiffness and the damping coefficients of the bearings, respectively; α and β the parameters of nonlinearity such that $\alpha>0$ and $\beta<2$ [21]. For instance, for the ball bearings, $\alpha=1/2$ (see [22]). $\mathbf{F}_{\text{int}}$ is the interaction force between the disk and the ring. This interaction can be represented as the resultant force

$$\mathbf{f}_r = \mathbf{F}_{\text{int}} = -k_r(\mathbf{r}_{02}-\mathbf{r}_{01}) - \sigma_r \oint (\mathbf{v}_r - \mathbf{v}_d)dl - \chi_r \oint [\mathbf{e}_z,(\mathbf{v}_r - \mathbf{v}_d)]dl \qquad (12)$$

and the moment about the point O_2

$$\mathbf{M}_r = -\oint [\mathbf{r}_2, \sigma_r(\mathbf{v}_r - \mathbf{v}_d) + \chi_r[\mathbf{e}_z,(\mathbf{v}_r - \mathbf{v}_d)]]dl, \qquad (13)$$

applied to the ring. Here $\mathbf{r}_{02}=x_2\mathbf{e}_x+y_2\mathbf{e}_y$ is the radius vector of the point O_2 in the fixed reference frame, k_r is the stiffness coefficient of the coupling of the centers of the disk and ring, σ_r is the coefficient of the Raleigh friction, χ_r is the coefficient of the gyroscopic force, $\mathbf{v}_d$ is the velocity of an arbitrary point $M(x,y)$ of the disk contiguous to the ring, $\mathbf{v}_r$ is the velocity of a point of the ring contiguous to the point M, $\mathbf{r}_2$ is the vector from the point O_2 to the point M. As it is seen from Eqs. (12) and (13), the interaction in the contact between the disk and the ring is represented by the frictional and the gyroscopic components. Integration in formulas (12) and (13) is taken over the ring's length, where dl is an arc element of the ring. Substituting into (12) and (13)

$$\mathbf{v}_d = \dot{\mathbf{r}}_{01} + [\boldsymbol{\omega}_1, \mathbf{r}_1], \quad \mathbf{v}_r = \dot{\mathbf{r}}_{02} + [\boldsymbol{\omega}_2, \mathbf{r}_2] \qquad (14)$$

with $\mathbf{r}_1=(x-x_1)\mathbf{e}_x+(y-y_1)\mathbf{e}_y$, $\mathbf{r}_2=(x-x_2)\mathbf{e}_x+(y-y_2)\mathbf{e}_y$, $\boldsymbol{\omega}_1=\dot{\varphi}_1\mathbf{e}_z$ the angular velocity of the disk, $\boldsymbol{\omega}_2=\dot{\varphi}_2\mathbf{e}_z$ the angular velocity of the ring, and integrating over the ring, one obtains:

$$\begin{aligned}\mathbf{f}_r = \{&-k_r(x_2-x_1) - \sigma_r L_r[(\dot{x}_2-\dot{x}_1) + (y_2-y_1)\dot{\varphi}_1] \\ &+ \chi_r L_r[(\dot{y}_2-\dot{y}_1) - (x_2-x_1)\dot{\varphi}_1]\}\mathbf{e}_x + \{-k_r(y_2-y_1) \\ &- \sigma_r L_r[(\dot{y}_2-\dot{y}_1) - (x_2-x_1)\dot{\varphi}_1] \\ &- \chi_r L_r[(\dot{x}_2-\dot{x}_1) + (y_2-y_1)\dot{\varphi}_1]\}\mathbf{e}_y\end{aligned} \qquad (15)$$

$$\mathbf{M}_r = -\mathbf{e}_z\sigma_r L_r R_r^2(\dot{\varphi}_2 - \dot{\varphi}_1), \qquad (16)$$

with $L_r=2\pi R_r$ the length of the ring.

We will use further the Lagrange formalism to describe the dynamic behavior of the system. The kinetic and the potential energies for the disk and for the ring have the form (index d designates the disk, while index r designates the ring):

$$T_d = \frac{m_d}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}I_d\dot{\varphi}_1^2,$$

$$U_d = \frac{k_d}{2}(x_1^2 + y_1^2) + \frac{k_{2d}}{(\alpha+2)}(x_1^2 + y_1^2)^{(\alpha+2)/2}, \tag{17}$$

$$T_r = \frac{m_r}{2}(\dot{x}_2^2 + \dot{y}_2^2) + \frac{1}{2}I_r\dot{\varphi}_2^2,$$

$$U_r = \frac{k_r}{2}[(x_2 - x_1)^2 + (y_2 - y_1)^2], \tag{18}$$

with I_d and I_r the central moments of inertia about axis Ox_3 of the disk and of the ring, respectively.

Expressions (17) and (18) allow one to compose the Lagrange function $L=T-U$, while Eqs. (11)–(13) give the expressions for the nonconservative generalized forces and the moment. Substituting the forces and the moment along with the Lagrangian into the Lagrange equation, one obtains:

$$\begin{aligned} m_d\ddot{x}_1 + k_dx_1 + k_r(x_1 - x_2) + k_{2d}(x_1^2 + y_1^2)^{\alpha/2}x_1 \\ = - \eta_d\dot{x}_1 - \eta_{2d}(\dot{x}_1^2 + \dot{y}_1^2)^{\beta/2}\dot{x}_1 + \sigma_rL_r((\dot{x}_2 - \dot{x}_1) + \dot{\varphi}_1(y_2 - y_1)) \\ - \chi_rL_r((\dot{y}_2 - \dot{y}_1) - \dot{\varphi}_1(x_2 - x_1)) \end{aligned} \tag{19}$$

$$\begin{aligned} m_d\ddot{y}_1 + k_dy_1 + k_r(y_1 - y_2) + k_{2d}(x_1^2 + y_1^2)^{\alpha/2}y_1 \\ = - \eta_d\dot{y}_1 - \eta_{2d}(\dot{x}_1^2 + \dot{y}_1^2)^{\beta/2}\dot{y}_1 + \sigma_rL_r((\dot{y}_2 - \dot{y}_1) - \dot{\varphi}_1(x_2 - x_1)) \\ + \chi_rL_r((\dot{x}_2 - \dot{x}_1) + \dot{\varphi}_1(y_2 - y_1)) \end{aligned} \tag{20}$$

$$\begin{aligned} m_r\ddot{x}_2 + k_r(x_2 - x_1) = - \sigma_rL_r((\dot{x}_2 - \dot{x}_1) + \dot{\varphi}_1(y_2 - y_1)) + \chi_rL_r((\dot{y}_2 - \dot{y}_1) \\ - \dot{\varphi}_1(x_2 - x_1)) \end{aligned} \tag{21}$$

$$\begin{aligned} m_r\ddot{y}_2 + k_r(y_2 - y_1) = - \sigma_rL_r((\dot{y}_2 - \dot{y}_1) - \dot{\varphi}_1(x_2 - x_1)) - \chi_rL_r((\dot{x}_2 - \dot{x}_1) \\ + \dot{\varphi}_1(y_2 - y_1)) \end{aligned} \tag{22}$$

$$I_d\ddot{\varphi}_1 = \sigma_rL_rR_r^2(\dot{\varphi}_2 - \dot{\varphi}_1) + M_{\mathrm{dr}} \tag{23}$$

$$I_r\ddot{\varphi}_2 = - \sigma_rL_rR_r^2(\dot{\varphi}_2 - \dot{\varphi}_1) \tag{24}$$

with M_{dr} the driving moment.

The system of Eqs. (19)–(24) has a particular solution:

$$x_1 = x_2 = 0; \quad y_1 = y_2 = 0; \quad \dot{\varphi}_1 = \dot{\varphi}_2 = \Omega \tag{25}$$

corresponding to the steady-state rotation regime around axis Oz.

We assume that the angular velocity of the disk $\dot{\varphi}_1$ is being maintained constant by the special driving moment M_{dr} and equal to Ω. Linearizing then system (19)–(24) near solution (25) and introducing complex variables

$$z_1 = x_1 + iy_1; \quad z_2 = x_2 + iy_2, \tag{26}$$

one can obtain the following system of equations:

$$\begin{cases} m_d\ddot{z}_1 + k_dz_1 + k_r(z_1 - z_2) = - \eta_d\dot{z}_1 + \eta_rL_r[(\dot{z}_2 - \dot{z}_1) - i\Omega(z_2 - z_1)] \\ m_r\ddot{z}_2 + k_r(z_2 - z_1) = - \eta_rL_r[(\dot{z}_2 - \dot{z}_1) - i\Omega(z_2 - z_1)] \\ I_r\delta\ddot{\varphi}_2 = - \sigma_rL_rR_r^2\delta\dot{\varphi}_2 \end{cases} \tag{27}$$

where $\delta\dot{\varphi}_2$ is the deviation of the angular velocity of the ring from the steady-state solution and $\eta_r=\sigma_r+i\chi_r$. The term η_r can be referred to as the *complex friction coefficient*. The real part of the complex friction describes the Raleigh friction, while the imaginary part corresponds to the gyroscopic force.

Note that the latter equation of system (27) is segregated. Its solution can be written as

$$\delta\dot{\varphi}_2 = C\exp\left(-\frac{\sigma_rL_rR_r^2}{I_r}t\right) \tag{28}$$

One can see from (28) that when time t tends to infinity the angular velocity of the ring $\dot{\varphi}_2$ approaches the angular velocity of the disk Ω.

Due to the symmetry properties of the system, Eqs. (27) are invariant with respect to the zero-time reference and a quarter-turn $(\pi/2)$ of the reference frame around the vertical axis. Thus, the first two equations in (27) admit the particular solutions of the kind

$$z_1 = \hat{z}_1\exp(\lambda t), \quad z_2 = \hat{z}_2\exp(\lambda t) \tag{29}$$

Introducing new variables

$$w_1 = \hat{z}_1, \quad w_2 = \hat{z}_2 - \hat{z}_1, \tag{30}$$

one yields

$$\begin{cases} m_dw_1\lambda^2 + k_dw_1 - k_rw_2 = - \eta_dw_1\lambda + \eta_rL_rw_2(\lambda - i\Omega) \\ m_rw_2\lambda^2 + m_rw_1\lambda^2 + k_rw_2 = - \eta_rL_rw_2(\lambda - i\Omega) \end{cases} \tag{31}$$

The characteristic equation of system (31) has the form

$$\begin{vmatrix} m_d\lambda^2 + \lambda\eta_d + k_d & -k_r - \eta_rL_r(\lambda - i\Omega) \\ m_r\lambda^2 & m_r\lambda^2 + k_r + \eta_rL_r(\lambda - i\Omega) \end{vmatrix} = 0 \tag{32}$$

Since λ is continuously dependent on the parameters of the problem, the change of the degree of instability can take place under appearance of an imaginary number $\lambda=i\omega$. Let us define those values of the damping coefficient η_d and the stiffness k_d of the bearings, which provide imaginary λ at given values of the other parameters.

Substituting $\lambda=i\omega$ into Eq. (32), introducing dimensionless parameters

$$K_d = \frac{k_d}{m_d\Omega^2}, \quad H_d = \frac{\eta_d}{m_d\Omega}, \quad \Pi_1 = \frac{\sigma_rL_r}{m_r\Omega}, \quad \Pi_2 = \frac{\chi_rL_r}{m_r\Omega},$$

$$\gamma = \frac{m_d}{m_d + m_r}, \quad K_r = \frac{k_r}{m_r\Omega^2}, \quad \tau = \frac{\omega}{\Omega}, \tag{33}$$

separating real and imaginary parts and solving the obtained equations with respect to K_d and H_d, one obtains:

$$K_d = \frac{\tau^2}{\gamma}\frac{(K_r - \Pi_2(\tau - 1) - \tau^2)(K_r - \Pi_2(\tau - 1) - \gamma\tau^2) + \Pi_1^2(\tau - 1)^2}{\Lambda}$$

$$H_d = (1 - \gamma^{-1})\frac{\Pi_1\tau^3(\tau - 1)}{\Lambda} \tag{34}$$

with $\Lambda=(K_r-\Pi_2(\tau-1)-\tau^2)^2+\Pi_1^2(\tau-1)^2$.

If the parameters γ, Π_1, Π_2, and K_r are fixed, Eqs. (34) define D-curve in the plane (K_d,H_d), which, alone with the singular straight line $H_d=0$, forms D-decomposition of the plane with domains of different degree of instability.

As is seen from (34), the D-curve crosses the straight line $H_d=0$ at the points

$$K_d = 0, \quad K_d = \frac{K_r - \gamma}{\gamma(K_r - 1)} \tag{35}$$

We assume that the parameters of the discrete model are analogous to those of the continuous one. This implies that, for instance, the parameter γ must be of the same order as the analogous parameter m of the continuous model, i.e., $\gamma \propto m$ (see Eq. (9)). We will choose the parameter K_r in such a way that the crossing of $H_d=0$ would take place at the same point (the same value $K_d=K_d^0$) as for the corresponding value of $K=K_0$ in the case of the continuous model. To this effect, we will use Eq. (35). One has to note that to reach the full equivalency of both models the value K_0 should be equal to $K_d^0\gamma/(1-\gamma)$, since in the case of the

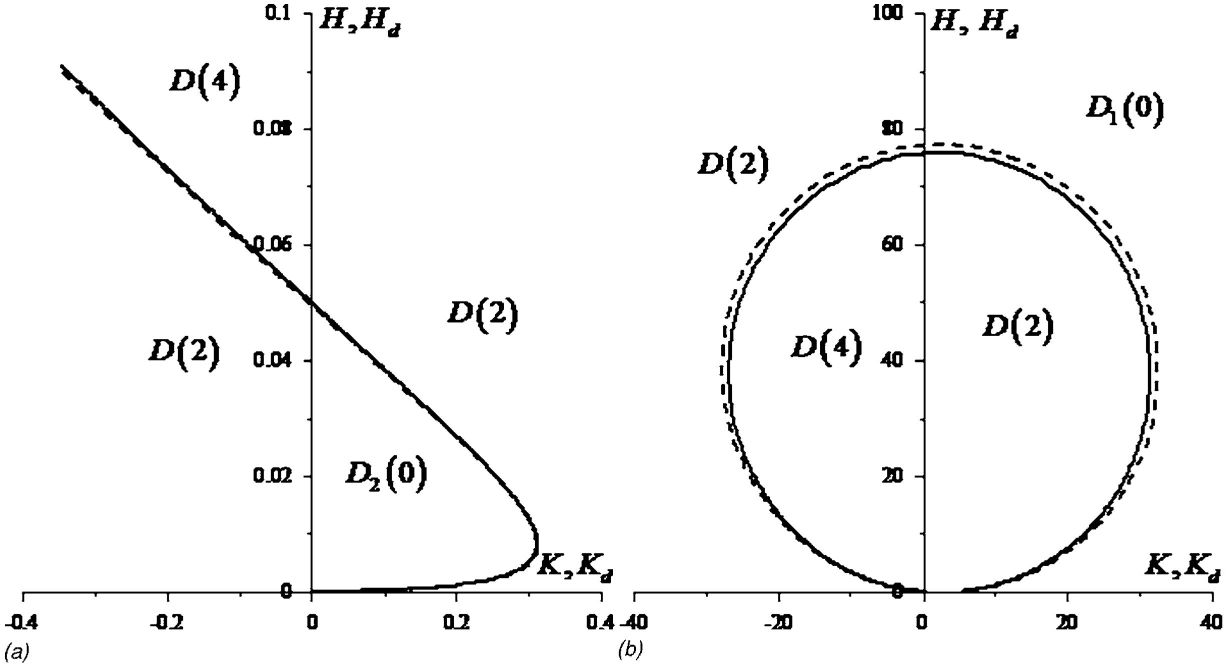


Fig. 8 ***D*-decomposition of the plane of the parameters of the bearings for the continuous (solid line) and the discrete (dashed line) models: (*a*) boundaries of $D_2(0)$ (expanded scale); (*b*) $D_1(0)$ domains**

continuous model the stiffness coefficient K_r is normalized by the mass of the liquid m_l, while in the case of the discrete model this coefficient (k_d) is normalized by the mass of the disk (rotor) M_d. The parameters Π_1 and Π_2 have to be taken to reach the maximal correspondence of the boundaries of the domains $D_2(0)$ in both cases.

In particular, for $E=10^{-5}$, $\delta=0.9$ and $m=0.4$ one yields $\Pi_1=0.021$, $\Pi_2=-2.6$, $K_r=1.18$, and $\gamma=0.57$. Figures 8(*a*) and 8(*b*) demonstrate *D*-decompositions of the plane (K_d,H_d) for the case of the continuous (the solid line) and the discrete (dashed line) models for the above-given parameters. As is seen from the figure, the curves coincide almost perfectly, at least, in the upper part.

It should be mentioned that in the whole plane (K_d,H_d), the *D*-curves differ quite noticeably (see Fig. 9). It is striking, however, that the difference between two upper parts of the "eights" is less than 5%. The difference between the lower parts of the "eights" has a transparent physical explanation. The point is that the upper part of the "eight" corresponds to the resonance excitation of the slow wave ($\tau<1$) propagating over the free surface of the liquid, while the lower part is associated with excitation of the faster mode ($\tau>1$). On the other hand, the viscosity of the liquid has significant influence on the rotor dynamics and on the resonances, in particular. This implies that the analogous discrete model must account for the effect of the viscosity appropriately. Further, the power of dissipation in the Navier-Stokes problem (3)

$$N_c=\frac{\mu}{2}\sum_{i,j=1}^{3}\int_D\left[\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i}\right]^2 dx_1dx_2dx_3, \tag{36}$$

to which the power of dissipation in the discrete model

$$N_d=\sigma_r\oint|\mathbf{v}_r-\mathbf{v}_d|^2dl \tag{37}$$

is to be compared, is a function of frequency. In Eq. (36) *D* is the liquid volume. Therefore, the friction coefficient σ_r must be taken as a convolution integral. In the model described above, σ_r is a constant. Thereby, having satisfied the first resonance, the discrete model cannot automatically satisfy the second one. This explanation is confirmed by the fact that with decrease of the Ekman number (the viscosity parameter) the curves become closer.

Andronov-Hopf Bifurcation: "Safe" and "Dangerous" Boundaries of the Stability Domain

In actual systems, slow changes of parameters can occur eventually. This evolution can result in that that the system will approach a boundary of the stability domain in the space of param-

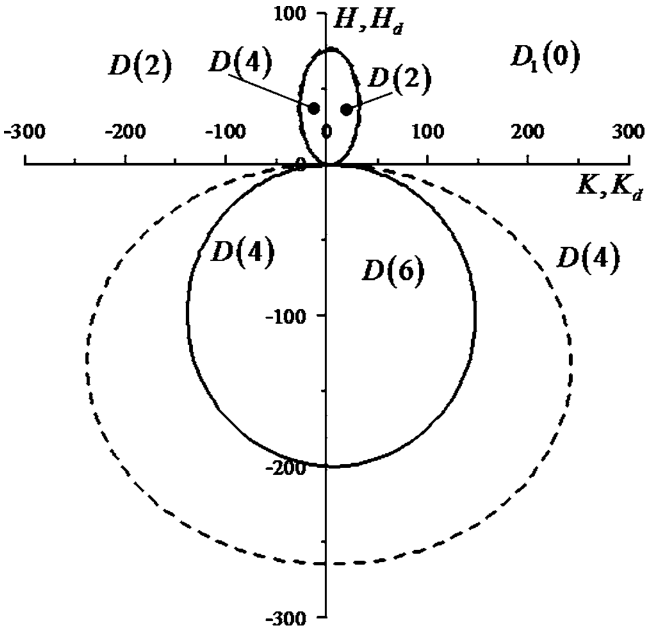


Fig. 9 General view of *D*-decomposition of the plane of the parameters of the bearings for the continuous (solid line) and the discrete (dashed line) models

eters. If this boundary is "dangerous," then perturbations of the system can lead to rise of self-excited vibrations even for the system parameters belonging to the stability domain. If the boundary is "safe," then the appearance of the self-excited vibrations can take place if the system has passed out the stability domain. This process is referred to in the literature as the Andronov-Hopf bifurcation or the bifurcation of generation of a limit cycle (see [23,24]). Let us demonstrate that in the considered discrete model, when the system parameters go in/out the stability domain, the Andronov-Hopf bifurcation can occur.

Using the complex variables (26), one can obtain from Eqs. (19)–(22)

$$\begin{cases} m_d\ddot{z}_1 + k_d z_1 + k_r(z_1 - z_2) + k_{2d}|z_1|^\alpha z_1 = -\eta_d\dot{z}_1 - \eta_{2d}|\dot{z}_1|^\beta\dot{z}_1 + \eta_r L_r[(\dot{z}_2 - \dot{z}_1) - i\Omega(z_2 - z_1)] \\ m_r\ddot{z}_2 + k_r(z_2 - z_1) = -\eta_r L_r[(\dot{z}_2 - \dot{z}_1) - i\Omega(z_2 - z_1)] \end{cases} \tag{38}$$

Introducing dimensionless variables

$$w_1 = z_1/R_d, \quad w_2 = (z_2 - z_1)/R_d, \quad t' = \Omega t, \tag{39}$$

we will seek the solution to the system (38) in the form

$$w_1 = \varepsilon_1 \exp(i\tau t'), \quad w_2 = \varepsilon_2 \exp(i\tau t') \tag{40}$$

Substituting this solution into Eqs. (38) one can come up to the following system of nonlinear equations with respect to ε_1 and ε_2:

$$\begin{cases} (K_d + i\tau H_d - \tau^2)\varepsilon_1 + (\gamma^{-1} - 1)[(\tau - 1)(\Pi_2 - i\Pi_1) - K_r]\varepsilon_2 = -K_{2d}|\varepsilon_1|^\alpha\varepsilon_1 - iH_{2d}|\tau\varepsilon_1|^\beta\tau\varepsilon_1 \\ \tau^2\varepsilon_1 + [\tau^2 + (\tau - 1)(\Pi_2 - i\Pi_1) - K_r]\varepsilon_2 = 0 \end{cases} \tag{41}$$

where

$$K_{2d} = \frac{k_{2d}R_d^\alpha}{m_d\Omega^2}, \quad H_{2d} = \frac{\eta_{2d}R_d^\beta\Omega^{\beta-1}}{m_d} \tag{42}$$

and the other denotations are taken from (33).

Note that if the system parameters belong to the D-curve, system (41) admits a nontrivial (with evanescent ε_1 and ε_2) solution such that

$$\det \mathbf{A} = 0 \tag{43}$$

where matrix

$$\mathbf{A} = \left\| \begin{matrix} K_d^* + i\tau H_d^* - \tau^{*2} & (\gamma^{-1} - 1)[(\tau^* - 1)(\Pi_2 - i\Pi_1) - K_r] \\ \tau^{*2} & \tau^{*2} + (\tau^* - 1)(\Pi_2 - i\Pi_1) - K_r \end{matrix} \right\| \tag{44}$$

and the star means that the parameters belong to the D-curve, i.e., satisfy to Eqs. (34).

Excluding ε_2 from (41), one obtains

$$(K_d + i\tau H_d - \tau^2)\varepsilon_1 - (\gamma^{-1} - 1)\frac{(\tau - 1)(\Pi_2 - i\Pi_1) - K_r}{\tau^2 + (\Pi_2 - i\Pi_1)(\tau - 1) - K_r}\tau^2\varepsilon_1 = -K_{2d}|\varepsilon_1|^\alpha\varepsilon_1 - iH_{2d}|\tau\varepsilon_1|^\beta\tau\varepsilon_1 \tag{45}$$

There always exists a simple root $\varepsilon_1 = 0$ corresponding to the steady-state rotation regime, or, to be more accurate, to the regime when only $\dot{\varphi}_2$ changes tending to Ω. Dividing (45) by ε_1, one yields an equation for nontrivial values of the precession radius (ε_1 can be always supposed to be real and positive), which, after separation of real and imaginary parts, can be reduces to the following system:

$$\begin{cases} (K_d - \tau^2) - (\gamma^{-1} - 1)\dfrac{\tau^2}{\Lambda}[(\Pi_2(\tau - 1) - K_r)(\tau^2 + \Pi_2(\tau - 1) - K_r) + \Pi_1^2(\tau - 1)^2] = -K_{2d}|\varepsilon_1|^\alpha \\ H_d + (\gamma^{-1} - 1)\dfrac{\tau^3}{\Lambda}\Pi_1(\tau - 1) = -H_{2d}|\tau\varepsilon_1|^\beta \end{cases} \tag{46}$$

with $\Lambda = (K_r - \Pi_2(\tau - 1) - \tau^2)^2 + \Pi_1^2(\tau - 1)^2$.

Introducing deviations of the dimensionless parameters K_d, H_d, and τ as

$$\delta K_d = K_d - K_d^*, \quad \delta H_d = H_d - H_d^*, \quad \delta\tau = \tau - \tau^*, \tag{47}$$

accounting for Eq. (43), and assuming δK_d, δH_d, and $\delta\tau$ to be small, one has

$$\begin{cases} K_{2d}|\varepsilon_1|^\alpha - \dfrac{dK_d(\tau^*)}{d\tau}\delta\tau = -\delta K_d \\ H_{2d}|\tau^*\varepsilon_1|^\beta - \dfrac{dH_d(\tau^*)}{d\tau}\delta\tau = -\delta H_d \end{cases} \tag{48}$$

Assuming now that $\alpha = \beta$, we obtain the following solution with respect to ε_1:

$$|\varepsilon_1|^\alpha = \frac{\dfrac{dK_d}{d\tau}\delta H_d - \dfrac{dH_d}{d\tau}\delta K_d}{\dfrac{dH_d}{d\tau}K_{2d} - |\tau|^\alpha H_{2d}\dfrac{dK_d}{d\tau}} \text{ or } |\varepsilon_1|^\alpha = \frac{\dfrac{dK_d}{dH_d}\delta H_d - \delta K_d}{K_{2d} - |\tau|^\alpha H_{2d}\dfrac{dK_d}{dH_d}} \tag{49}$$

To avoid jumps from one part of the D-curve to another, we take the deviations δK_d and δH_d normal to the boundary. Note that Eq. (49) can be rewritten as

$$|\varepsilon_1|^\alpha = \pm\frac{|\vec{l} \times \vec{n}|}{\dfrac{dH_d}{d\tau}K_{2d} - |\tau|^\alpha H_{2d}\dfrac{dK_d}{d\tau}} \tag{50}$$

with

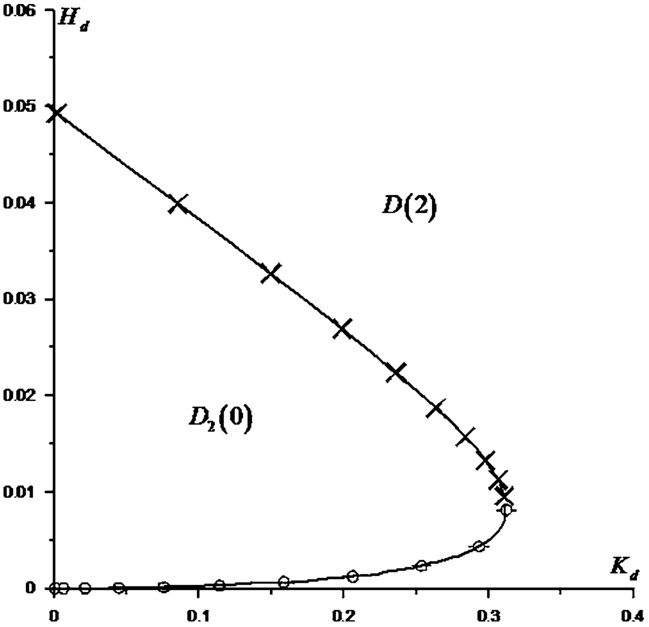


Fig. 10 "Dangerous" and "safe" intervals of the $D_2(0)$ stability domain

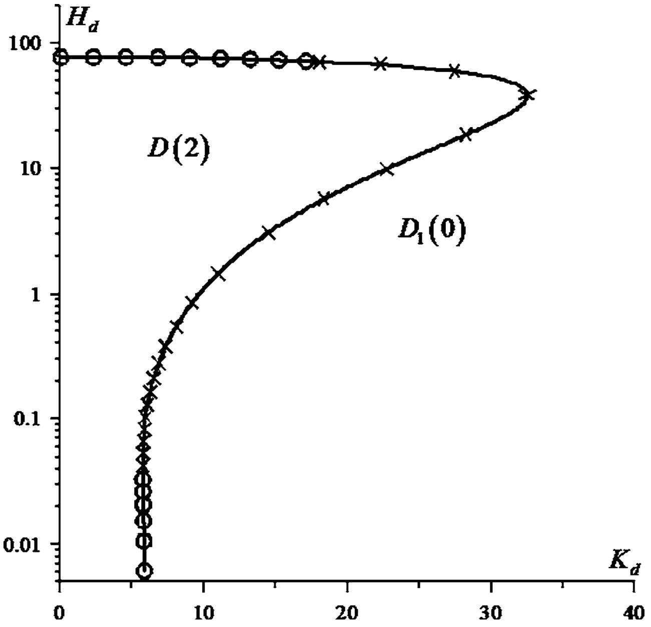


Fig. 11 "Dangerous" and "safe" intervals of the $D_1(0)$ stability domain

$$l=\left(\frac{dK_d}{d\tau},\frac{dH_d}{d\tau}\right)$$

the tangential vector to the D-curve and $\boldsymbol{n}=(\delta K_d,\delta H_d)$ the vector normal to the boundary. The plus/minus sign determines "entrance"/"exit" in/from the stability domain. Thus, the dynamic behavior in the vicinity of the boundary is defined by the sign of the denominator.

If generation of a periodic motion takes place when the system comes out of the stability domain, then this bifurcation is called "supercritical." In this case, it is conventional to call the excitation of self-excited vibrations "soft" and to denote the boundary as "safe." In Fig. 10 the "safe" intervals of the bifurcation curve are represented by the white circles. The "dangerous" part is denoted by the crosses. In the vicinity of the "dangerous" curve, the periodic motion of a kind of a circular precession of a small radius can exist in the stability domain of the steady-state rotation regime. This motion is generated when the system goes in the stability domain through the "dangerous" sector. In this case, the excitation of self-excited vibrations is called "hard" and the bifurcation as "subcritical."

In calculations, the following parameters were taken:

$$\gamma=0.57,\quad K_r=1.18,\quad \Pi_1=0.021,\quad \Pi_2=-2.6,$$
$$K_{2d}=-0.02,\quad H_{2d}=0.02,\quad \alpha=1 \tag{51}$$

Figure 10 corresponds to the case of a so-called soft or retrograding nonlinearity ($K_{2d}<0$). The curve consists of the "safe" interval that abuts upon the abscissa axis and the upper "dangerous" one. Formula (50) along with the D-curve demonstrates that the lower part of the curve is always "safe" for the retrograding nonlinearity. If one tests the case of the subcritical bifurcation for this interval, the denominator is always negative, i.e., there exists no positive roots. As far as the upper part of the curve is concerned, the "safe" interval extends as the parameter K_{2d} decreases. For $K_{2d}=-0.8$ the right-hand boundary of the "dangerous" branch corresponds to the point $K_d=0.16$, $H_d=0.03$. For $K_{2d}=-1.1$ this boundary is located at $K_d=0.05$, $H_d=0.044$. For $K_{2d}\leqslant-1.2$, all boundary becomes "safe."

In the case of a progressing nonlinearity, the picture reminds the case of the retrograding nonlinearity shown in Fig. 10. This is valid if $K_{2d}\leqslant 0.04$. When $K_{2d}>0.04$, the "dangerous" interval starts to extend to the lower part of the D-curve. For instance, for $K_{2d}=0.9$ the right-hand boundary of the "safe" interval is located at the point $K_d=0.2$, $H_d=0.001$, while for $K_{2d}=2$ this boundary lies at $K_d=0.05$, $H_d=5\times10^{-5}$. When $K_{2d}>3$ the all boundary becomes "dangerous."

Figure 11 shows "dangerous" and "safe" intervals of the boundary of the $D_1(0)$ domain for the parameters (51) (retrograding nonlinearity). As is seen from the figure, the boundary consists of one "dangerous" and two "safe" zones. With increase of K_{2d} the upper "safe" interval extends till the derivative dK_d/dH_d changes its sign (at the point $K_d=32$, $H_d=45$). It should be noted that the boundary is always "dangerous" in the interval $H_d=[0.04;45]$. If K_{2d} decreases, the "dangerous" part extends. When $K_{2d}=-0.18$, the lower "safe" part disappears, while the upper "safe" zone is confined by the interval $K_d=[0;3.2]$ on which the derivative dK_d/dH_d is positive. If the nonlinearity is progressive, the intervals with negative dK_d/dH_d are always "safe." The intervals with the positive dK_d/dH_d can be either "safe" or "dangerous" depending on the parameters K_{2d} and H_{2d}.

Discussion and Conclusions

In this paper, we have proposed new approach for investigation of fluid-filled rotor systems. The method consists in the development of finite degrees-of-freedom (discrete) analogs of the rotor systems. Although at the first time the development of the discrete model of the rotor requires also the development of the continuous models to identify the parameters of both models, one can obtain qualitatively and quantitatively complete correspondence of dynamic behavior of the systems. Furthermore, the method allows one to employ the tools of the theoretical mechanics and the oscillation theory. This implies that one can go farther in the study of the rotor systems with fluid and investigate such questions as bifurcations, transition to the chaotic oscillations, amplitude dispersion of the fluid, nonlinear effects of the bearings, wave effects in the shaft, etc.

Of course, the authors of this paper are aware of that that it is impossible to account for everything and that a problem, having been solved once by some method, can be also solved by means of a different procedure. However, the attractiveness of our method primarily consists of the simple mechanical model that provides clear physical interpretation of the results obtained. This can be very valuable in engineering practice. Potentially, the method presented in this paper may contribute to the solution and understanding of a much broader class of rotating machinery problems. In 1885, Zhukovskiy [25] demonstrated that the problem of motion of a solid, having a cavity completely filled with a homogeneous ideal fluid that performs potential motion, can be reduced to a finite degrees-of-freedom one where the fluid can be replaced by an equivalent solid. We are sure that an equivalent approach can be developed for the rotor systems partly/completely filled with viscous incompressible/compressible homogeneous/stratified fluid as well. Development of such a general approach is methodologically advantageous, for it will allow one to properly take into account such an important and complicated effect as combined fluid filling with weighted particles.

Acknowledgment

This work was financed by the Russian Foundation for Basic Research, Grants Nos. 03-01-00478, 03-01-00644, and 03-01-06184, and 03-02-16924. This support is highly appreciated.

References

[1] Moiseev, N. N., and Rumyantsev, V. V., 1965, *Dynamics of Solid With Cavities Containing Fluid*, Nauka, Moscow.

[2] Mikishev, G. N., and Rabinovich, B. I., 1968, *Dynamics of Solid With Cavities Filled with Liquid*, Mashinostroenine, Moscow.

[3] Epishev, L. V., 1959, "On the Dynamics Instability of a Revolving Rotor Partly Filled With a Liquid," Nauchnye Doklady vysshei shkoly. Series Mashinostroenie i priborostroenie, **2**, pp. 66–74 (in Russian).

[4] Crandall, S. H., 1962, "Rotating and Reciprocating Machines," *Handbook of Engineering Mechanics*, W. Flügge, ed., McGraw-Hill, New York, pp. 58.1–58.24.

[5] Kollmann, F. G., 1962, "Experimentelle und Theoretische Unterschungen uber die Kritischen Drehzalen Flussigkeitsgefullter Hohlkorper," Forsch. Geb. Ingenieurwes., **28**, pp. 115–123; 147–153.

[6] Ehrich, F. F., 1967, "The Influence of Trapped Fluids on High Speed Rotor Vibration," ASME J. Eng. Ind., **89**, pp. 806–812.

[7] Wolf, J. A., Jr., 1968, "Whirl Dynamics of a Rotor Partially Filled With Liquid," ASME J. Appl. Mech., **35**(4), pp. 676–682.

[8] Daich, I. M., and Bar, I. L., 1973, "Oscillations of a Rotating Solid Body With a Cavity Partly Filled With Viscous Fluid," Prikl. Mekh., **9**(5), pp. 64–69.

[9] Saito, S., and Someya, T., 1980, "Self-Excited Vibration of a Rotating Hollow Shaft Partially Filled With Liquid," ASME J. Mech. Des., **102**(1), pp. 185–192.

[10] Lichtenberg, G., 1982, "Vibrations of an Elastically Mounted Spinning Rotor Partially Filled With Liquid," ASME J. Appl. Mech., **104**, pp. 389–396.

[11] Holm-Christensen, O., and Träger, K., 1991, "A Note of Instability Caused by Liquid Motions," ASME J. Appl. Mech., **58**, pp. 804–811.

[12] Derendyeav, N. V., and Sandalov, V. M., 1982, "To Stability of Steady-State Rotation of Cylinder Partly Filled With Viscous Incompressible Liquid," Appl. Math. Mech., **46**(4), pp. 578–586.

[13] Derendyaev, N. V., 1988, "Andronov-Bopf Bifurcation in the Dynamics of a Liquid-Filled Rotor System," Sov. Phys. Dokl., **33**(8), pp. 592–594; [Dokl. Akad. Nauk Arm. SSR, **301**(4–6), pp. 798–801].

[14] Derendyaev, N. V., 1999, "Nonconservative Problems of Dynamics of Rotor Systems Containing Fluid," dissertation of Doctor of Science, NNSU, Nizhny Novgorod.

[15] Zhu, Changsheng, 2002, "Experimental Investigation Into the Instability of an Over-Hung Rigid Centrifuge Rotor Partially Filled With Fluid," ASME J. Vibr. Acoust., **124**, pp. 483–491.

[16] Derendyaev, N. V., and Soldatov, I. N., 2002, "Study of Stability and Self-Excited Vibrations of a Rotor With Liquid on the Basis of a Discrete Model," *Proceedings of the VI Scientific Conference on "Nonlinear vibrations of Mechanical Systems,"* Nizhny Novgorod, p. 66 (in Russian).

[17] Dosaev, M. Z., and Samsonov, V. A., 2002, "To Stability of Rotation of a Heavy Solid With Viscous Filling," Appl. Math. Mech., **66**(3), pp. 427–433.

[18] Derendyaev, N. V., 1983, "Stability of Steady Rotation of a Cylinder Filled With a Stratified Viscous Incompressible Liquid," Sov. Phys. Dokl., **28**(10), pp. 867–869; [Dokl. Akad. Nauk Arm. SSR, **272**(4–6), pp. 1073–1076].

[19] Derendyaev, N. V., and Senyatkin, V. A., 1984, "Stability Conditions for the Steady-State Rotation of a Cylinder Filled With a Stratified Nonuniform Viscous Incompressible Liquid," J. Appl. Mech., **25**(1), pp. 30–39; [Zhurnal Prikladnoi Mekhaniki i Tehknicheskoi Fiziki, **25**(1), pp. 34–44.

[20] Neimark, Y. I., 1949, "Stability of Linearized Systems," LKVVIA, Leningrad.

[21] Bautine, N. N., 1984, "Behaviour of Dynamic System Near the Boundaries of the Stability Domain," 2nd ed., Nauka, Moscow.

[22] Kapitsa, P. L., 1939, "Stability and Passing Through Critical Turnovers of High-Speed Revolving Rotors at the Presence of Friction," Zh. Tekh. Fiz., **9**(2), pp. 124–147.

[23] Andronov, A. A., and Lubina, A. G., 1935, "Application of the Poincaret Theory About the Bifurcation Points and the Change of Stabilities to Simplistic Self-Excited Systems," Zhurnal Eksperimentalnoi i Prikladnoi Physiki, **5**(3–4), pp. 296–309.

[24] Hopf, E., 1942, "Abzweigung Einer Periodischen Lösung von Einer Stazionaren Lözung eines Differential Systems," Ber. Math.-Phys. Sachsische Akademie der Wissensehaften Leipzig, **94**, pp. 3–22.

[25] Zhukovskiy, N. E., 1948, "Motion of a Solid Having Cavities Filled With a Homogeneous Liquid. Selected Works," Gostechizdat, Moscow-Leningrad (in Russian), Vol. 1, pp. 31–152.

[26] Neimark, J. I., 2003, *Mathematical Models in Natural Science and Engineering* (Foundations of Engineering Mechanics), Springer-Verlag, Berlin.

[27] Metrikine, A. V., Veritchev, S. N., and Blaauwendraad, J., 2005, "Stability of a Two-Mass Oscillator Moving on a Beam Supported by a Visco-Elastic Half-Space," Int. J. Solids Struct., **42**, pp. 1187–1207.

Tuhin Das

Ranjan Mukherjee
Associate Professor
e-mail: mukherji@egr.msu.edu

Department of Mechanical Engineering,
Michigan State University,
East Lansing, MI 48824

Reconfiguration of a Rolling Sphere: A Problem in Evolute-Involute Geometry

This paper provides a new perspective to the problem of reconfiguration of a rolling sphere. It is shown that the motion of a rolling sphere can be characterized by evolute-involute geometry. This characterization, which is a manifestation of our specific selection of Euler angle coordinates and choice of angular velocities in a rotating coordinate frame, allows us to recast the three-dimensional kinematics problem as a problem in planar geometry. This, in turn, allows a variety of optimization problems to be defined and admits infinite solution trajectories. It is shown that logarithmic spirals form a class of solution trajectories and they result in exponential convergence of the configuration variables. [DOI: 10.1115/1.2164515]

1 Introduction

The rolling sphere epitomizes the complexity of diverse problems in mechanics involving geometry of rotations and its reconfiguration has similarities with many engineering problems, such as spacecraft attitude reorientation and manipulation of rigid objects using robotic fingers. The configuration of a rolling sphere is described by two Cartesian coordinates of its center and three orientation coordinates, and reconfiguration refers to the task of designing a trajectory that enables it to roll from an arbitrary configuration to a desired configuration. This problem has seen a few solutions until date but new approaches and solutions to the problem should be welcome since they can provide new insight into the realm of rotational kinematics. In this paper, we show that the kinematics of a rolling sphere can be described by evolute-involute geometry and provide a fundamentally new approach towards solving the reconfiguration problem.

To the best of our knowledge, Hammersley [1] provided the first solution to the reconfiguration problem in 1983. A simpler solution to the problem in the form of a three-step algorithm was proposed by Li and Canny [2] in 1990. In the first step, the two Cartesian coordinates of the sphere are converged to their desired values. The second step generates an equatorial triangle on the surface of the sphere and converges two of the three orientation coordinates, and the third step uses a polhode to converge the third orientation coordinate. An optimal solution, based on minimization of integral of the kinetic energy of the sphere, was provided by Jurdjevic [3] in 1993. The results are elegant and indicate that the optimal trajectories are described by elliptic functions. In 2002, Mukherjee et al. [4] proposed two computationally efficient motion planning algorithms for the rolling sphere. The first algorithm is based on planar geometry whereas the second algorithm is based on spherical trigonometry.

In relation to the above papers, where the primary focus has been the solution to the reconfiguration problem, this paper establishes a fundamental property of the motion of rolling spheres, namely, the motion is equivalent to the action of wrapping and unwrapping a taut rope on a planar curve. This evolute-involute geometric characterization is a manifestation of our specific selection of Euler angle coordinates and choice of angular velocities in the rotating coordinate frame. The planar geometric formulation of the problem admits infinite solution trajectories and allows us to pose a variety of optimization problems with different objective functions, including the isoperimetric problem [5]. Clearly, the importance of the paper lies in the generality of the adopted approach rather than the specific solution to the problem provided on the basis of the approach.

In search of a class of solution trajectories for the sphere motion planning problem, discussed above, we investigate the Sweep-Tuck algorithm [6] which provides the first and only solution to the feedback stabilization problem. For our open-loop geometric problem, we show that the nonsmooth trajectories provided by the Sweep-Tuck algorithm transform into a class of smooth solution trajectories under limiting conditions. These solution trajectories form an evolute-involute pair of logarithmic spirals and result in exponential convergence of the configuration variables.

This paper is organized as follows: In Sec. 2 we present the kinematic model of the rolling sphere and give an overview of the Sweep-Tuck algorithm. In Sec. 3 we describe the motion of the sphere using evolute-involute geometry and pose the reconfiguration problem. In Sec. 4, we present an algorithm for partial reconfiguration of the sphere. The flexibility of the partial reconfiguration algorithm is exploited in developing the complete reconfiguration algorithm in Sec. 5. In Sec. 6 we present simulation results and concluding remarks are presented in Sec. 7.

2 Background

2.1 Kinematic Model. Consider an arbitrary configuration of the sphere, as shown in Fig. 1(*a*). We denote the Cartesian coordinates of the sphere center by $Q \equiv (x, y)$ and adopt the *z*-*y*-*z* Euler angle sequence (α, θ, ϕ) to represent the orientation of the sphere. As per the *z*-*y*-*z* Euler angle sequence, the inertially fixed xyz frame is first rotated about the positive z axis by angle α, $-\pi \leqslant \alpha \leqslant \pi$, to obtain the frame $x_1y_1z_1$. The frame $x_1y_1z_1$ is then rotated about the y_1 axis by angle θ, $0 \leqslant \theta \leqslant \pi$, to obtain frame $x_2y_2z_2$. Finally, the $x_2y_2z_2$ frame is rotated about the z_2 axis by angle ϕ to obtain frame $x_3y_3z_3$, which is fixed to the sphere. The points P and R are defined to be the intersection points of the sphere surface with the z_3 and x_3 axes, respectively. Assuming the sphere to have unity radius without any loss of generality, and denoting the angular velocities of the sphere about the x_1, y_1, z_1 axes as ω_x^1, ω_y^1, ω_z^1, respectively, the state equations for $\omega_z^1 = 0$ (the assumption $\omega_z^1 = 0$ is made to impose the constraint that the sphere cannot spin about the vertical axis) can be written as follows [4]:

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 23, 2005; final manuscript received October 28, 2005. Review conducted by N. Sri Namachchivaya. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

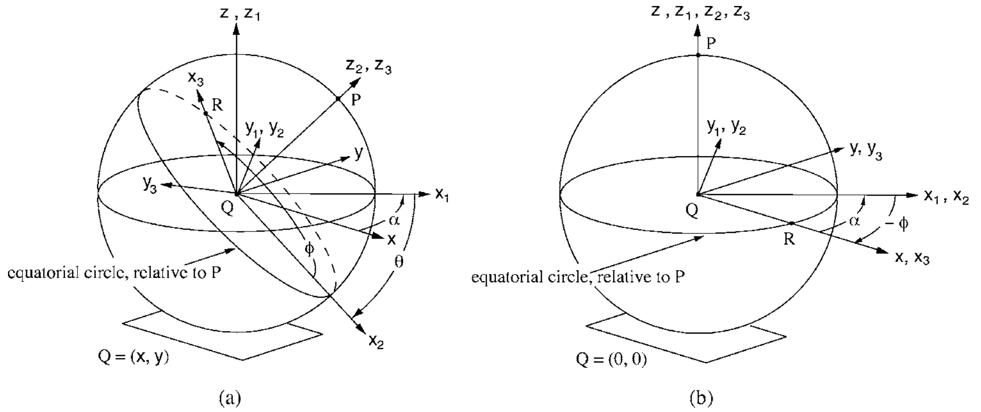


Fig. 1 (***a***) **An arbitrary configuration of the sphere; (*b*) desired configuration of the sphere**

$$\dot{x} = \omega_y^1 \cos \alpha + \omega_x^1 \sin \alpha \tag{1}$$

$$\dot{y} = \omega_y^1 \sin \alpha - \omega_x^1 \cos \alpha \tag{2}$$

$$\dot{\theta} = \omega_y^1 \tag{3}$$

$$\dot{\alpha} = -\,\omega_x^1 \cot \theta \tag{4}$$

$$\dot{\beta} = \omega_x^1 \tan(\theta/2) \tag{5}$$

where β is defined as follows:

$$\beta = \alpha + \phi \tag{6}$$

The reorientation of the sphere refers to the task of bringing P to the vertically upright position, and R, which then lies on the diametrical circle in the xy plane, to lie on the positive x axis. This can be achieved with $\theta=0$ and $\beta=0$, irrespective of the individual values of α and ϕ, as shown in Fig. 1(*b*). This is related to the notion of controllability and elaborated in [4]. Therefore, the sphere can be completely reconfigured by designing trajectories for ω_x^1 and ω_y^1 that result in

$$x=0, \quad y=0, \quad \theta=0, \quad \alpha+\phi=\beta=0 \tag{7}$$

2.2 Effect of Rotations About Moving Coordinates. Consider the motion of the sphere for the following actuations:

(A) $\omega_y^1 \neq 0,\ \omega_x^1 = 0$
(B) $\omega_x^1 \neq 0,\ \omega_y^1 = 0,\ \theta \neq 0$

The motion of the sphere for these actions is explained with the help of Fig. 2. Under action (A), the sphere moves along straight line CF as θ changes. Let F be the point on this straight line where the sphere would have $\theta=0$. Since the sphere rolls without slipping, this point remains invariant under action (A). For actuation (B), the instantaneous radius of the path traced by the sphere on the xy plane can be computed using Eqs. (1)–(4) as follows:

$$\rho = \left| \frac{(\dot{x}^2+\dot{y}^2)^{3/2}}{\dot{x}\ddot{y}-\dot{y}\ddot{x}} \right| = \tan \theta \tag{8}$$

Since $\omega_y^1=0$, θ is maintained constant. This implies that the contact point of the sphere, the center of the sphere, and points P and F move in the horizontal plane along circular arcs whose center lie on the vertical axis that passes through C. We can easily show that distance CF satisfies

$$CF = \tan \theta - \theta \tag{9}$$

The point C remains fixed under actuation (B), but under actuation (A) moves away from F, as θ increases, and converges to F, as θ converges to zero. The variables α and β in Eqs. (4) and (5) change during actuation (B) but remain constant during actuation (A). During actuation (B), the change in variable β is given by the expression

$$\Delta\beta = \Delta\alpha + \Delta\phi = \Delta\alpha(1-\sec\theta) \tag{10}$$

2.3 Partial Reconfiguration Using the Sweep-Tuck Algorithm. In this section we present the main results of the Sweep-Tuck algorithm detailed in [6]. With reference to Fig. 1, we define partial reconfiguration as the task of converging Q to the origin of the Cartesian coordinate frame and P to the vertically top position. This allows R to have an arbitrary location on the equatorial circle but requires us to satisfy

$$x=0, \quad y=0, \quad \theta=0 \tag{11}$$

Now consider an arbitrary configuration of the sphere as shown in Fig. 3. The points C and F in Fig. 3 were defined earlier in Sec. 2.2 using Fig. 2. It can be shown that

$$(CF, CO) \equiv (0,0) \Leftrightarrow (x,y,\theta) \equiv (0,0,0) \tag{12}$$

and this motivates the following remark:

Remark 1. The sphere in Fig. 3 will be partially reconfigured in the sense of Eq. (11) if and only if (CF, CO) *converge to* (0,0).

Towards the goal of partial reconfiguration, we now recall the following theorem from [6].

DUAL-POINT THEOREM. *Let C and F be two points in the xy plane with origin at O, as shown in Fig. 4. Suppose $\psi = \angle OCF$ is*

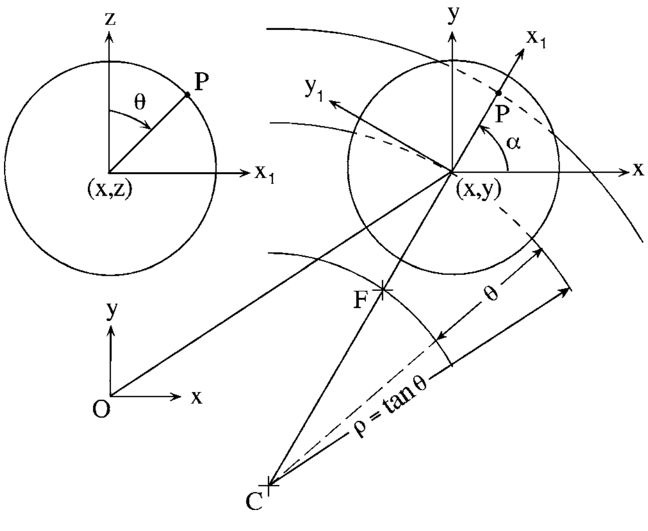


Fig. 2 **Actuations (A) and (B)**

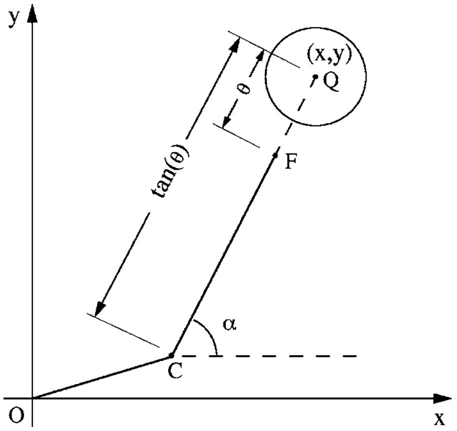


Fig. 3 An arbitrary configuration of the sphere

acute, and let $(CF/CO)=\lambda, \lambda \in (1,\infty)$. *If* ψ *satisfies the condition*

$$0 \leqslant \psi < \cos^{-1}(1/\lambda) \qquad (13)$$

then there exists a point C' *on the extended line* CF *such that for* $\psi' = \angle OC'F$, $0 \leqslant \psi' \leqslant \pi$,

$$(C'F/C'O) = \lambda \cdots (i) \quad 0 < (C'O/CO) < 1 \cdots \ (ii)$$

$$\psi' > \psi \cdots \ (iii) \qquad (14)$$

Remark 2. In [6] the Dual-Point Theorem is stated and proved for both cases $\lambda \in (1,\infty)$ and $\lambda \in (0,1)$. However, we restrict ourselves to the case $\lambda \in (1,\infty)$ in this paper.

The intermediate angle, ψ', can be expressed as follows [6]:

$$\tan \psi' = \frac{-|1-\lambda^2| \sin \psi}{(1+\lambda^2)\cos \psi - 2\lambda}, \quad \lambda \in (1,\infty) \qquad (15)$$

and it can be shown that Eq. (13) implies

$$\cos^{-1}\left(\frac{1}{\lambda}\right) < \psi' \leqslant \pi \qquad (16)$$

Now consider Fig. 5 where C and F define an arbitrary configuration of the sphere. Suppose $\psi = \angle OCF$ satisfies the conditions in Eq. (13). Let C' in Fig. 5 be the point on line CF that satisfies the conditions in Eq. (14). We now define three specific maneuvers of the sphere.

Definition 1 (dpt maneuver). *A "Dual-Point Tuck" (DPT) Maneuver is an actuation (A) that moves the sphere such that* C *moves to* C'.

From Dual-Point Theorem we know that a DPT maneuver results in $\psi' > \psi$. Therefore, ψ' can be restored to the value ψ in one of two ways as shown in Fig. 5. This motivates the next definition.

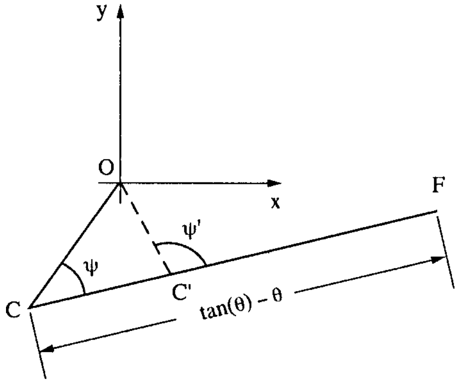


Fig. 4 The *C-C′* pair for the Dual-Point Theorem

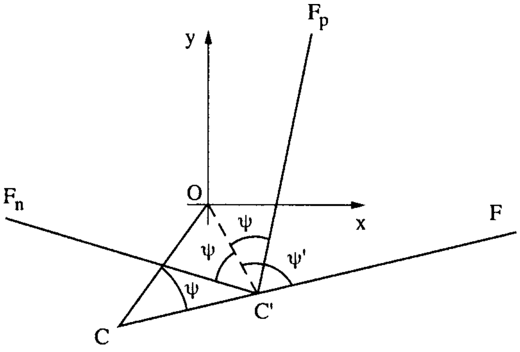


Fig. 5 RS and DPT maneuvers

Definition 2 (rs maneuver). *Following a DPT maneuver, an actuation (B) that moves the sphere to restore* ψ' *to* ψ *is defined as a "Restoring-Sweep" (RS) Maneuver.*

Definition 3 (ps maneuver). *An actuation (B) that moves the sphere at the initial time to bring* $\angle OCF$ *to* ψ' *is defined as a "Preliminary-Sweep" (PS) Maneuver.*

We now present the "Sweep-Tuck" algorithm [6].

Sweep-Tuck Algorithm. Consider a sphere whose partial configuration (x,y,θ), defined by the location of points C and F, initially satisfies $0 < \theta \leqslant (\pi/2-\epsilon)$ and $(CF/CO)=\lambda \in (1,\infty)$. If ψ is chosen in accordance with Eq. (13), partial reconfiguration in the sense of Eq. (11) can be achieved through a PS maneuver followed by repeated application of RS-DPT maneuvers.

The Sweep-Tuck algorithm utilizes the fact that alternate RS and DPT maneuvers decrease both CF and CO in geometric progression; the rate at which CF and CO decrease depends on n and ψ [6]. The distance traversed by C during each DPT maneuver also depends on n and ψ and is given by the relation

$$CC' = \frac{2\lambda CO(\lambda \cos \psi - 1)}{(\lambda^2 - 1)} \qquad (17)$$

3 Geometry of Reconfiguration

3.1 The Evolute-Involute Pair. We investigated the motion of C and F in Sec. 2.2, where the actuating inputs ω_x^1 and ω_y^1 were not applied simultaneously, i.e., $\omega_y^1=0$ when $\omega_x^1 \neq 0$, and vice versa. In this section we investigate the motion of C and F under simultaneous variation of ω_x^1 and ω_y^1. To this end we first note that the coordinates of C and F can be obtained from Fig. 3 as follows:

$$C_x = x - \tan\theta \cos\alpha \quad F_x = x - \theta\cos\alpha$$

$$C_y = y - \tan\theta \sin\alpha \quad F_y = y - \theta \sin\alpha \qquad (18)$$

By differentiating the above equations and substituting Eqs. (1)–(5) we get

$$\dot{C}_x = -\omega_y^1 \cos\alpha \tan^2\theta \quad \dot{F}_x = \omega_x^1 \sin\alpha(1-\theta\cot\theta)$$

$$\dot{C}_y = -\omega_y^1 \sin\alpha \tan^2\theta \quad \dot{F}_y = -\omega_x^1 \cos\alpha(1-\theta\cot\theta) \qquad (19)$$

From Eq. (19) we deduce

$$dC_y/dC_x = \tan\alpha \quad \text{for } \theta \neq 0 \text{ and } \omega_y^1 \neq 0$$

$$dF_y/dF_x = -\cot\alpha \quad \text{for } \theta \neq 0 \text{ and } \omega_x^1 \neq 0 \qquad (20)$$

We note from Figs. 2 and 3 that the instantaneous slope of CF is $\tan\alpha$. This motivates the following remark.

Remark 3. Under simultaneous application of actuating inputs ω_x^1 and ω_y^1, the instantaneous motion of C and F are tangential and perpendicular, respectively, to line CF.

Remark 3 outlines the qualitative nature of the trajectory of F,

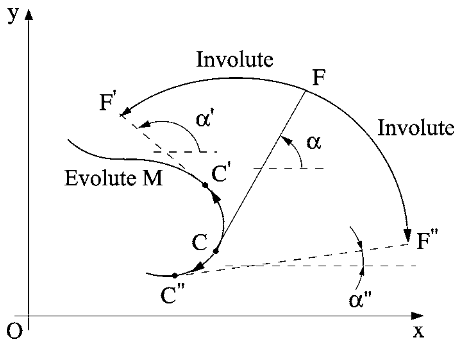
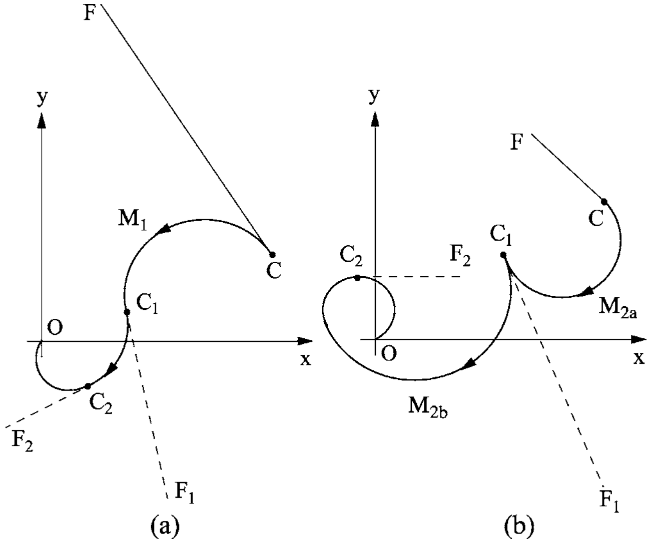


Fig. 6 Motion of C and F due to ω_x^1 and ω_y^1



(a) (b)

Fig. 7 Wrapping and unwrapping trajectories of the point C

given a trajectory of C. It indicates that if C is constrained to follow a desired curve, F will traverse a path such that for any instantaneous position of C, CF is tangential to the curve at that point. This is illustrated in Fig. 6 where C is constrained to follow the curve M. As C moves to C' or C'', F moves to F' or F'', respectively, such that $C'F'$ or $C''F''$ are tangential to the curve M at C' or C'', respectively. Using Eqs. (1)–(5) and (18), we can now show

$$\Delta C_x = -\cos\alpha\tan^2\theta\Delta\theta$$

$$\Delta C_y = -\sin\alpha\tan^2\theta\Delta\theta \tag{21}$$

which leads to

$$|\Delta C| = (\Delta C_x^2 + \Delta C_y^2)^{1/2} = \tan^2\theta|\Delta\theta| \tag{22}$$

Also, from Eq. (9) we obtain

$$CF = (\tan\theta - \theta) \Rightarrow |\Delta CF| = \tan^2\theta|\Delta\theta| \tag{23}$$

Equations (22) and (23) effectively imply that for an infinitesimal distance $\Delta\mathcal{L}_C$ traversed by C along its trajectory

$$|\Delta CF| = \Delta\mathcal{L}_C \tag{24}$$

Remark 3 and Eq. (24) together lead to the interesting geometric result that the distance traversed by C along its constrained trajectory M equals the change in length of CF. The result can be visualized with CF as a taut rope wrapping or unwrapping on a two-dimensional curve M. During wrapping or unwrapping, the rope always remains tangential to the curve, the point of contact being C and the other end being F as illustrated in Fig. 6. The action of wrapping is illustrated by the transition of C and F to C' and F', respectively, where distance $C'F' < CF$. Similarly, unwrapping is illustrated by the transition to C'' and F'', where distance $C''F'' > CF$. The observations made above immediately imply that the trajectories of C and F form a *Evolute-Involute* pair. The trajectory of F is an involute of the trajectory of C, which is the evolute. We summarize our observations in Remark 4 below.

Remark 4. If the point C is constrained to traverse a prescribed path by the inputs ω_x^1 and ω_y^1, the point F moves such that CF is tangential to the trajectory of C *and* CF "wraps" or "unwraps" on the curve followed by C. The trajectories of the points C and F form an Evolute-Involute pair.

The evolute and involute trajectories of points C and F provide a fundamentally new description of the motion of rolling spheres. This description is a direct manifestation of our specific selection of Euler angle coordinates and choice of angular velocities of the sphere in the rotating coordinate frame. In the next section we utilize the evolute-involute geometric description to pose the reconfiguration problem as a problem in planar geometry.

3.2 Geometry Based Problem Definition. For partial reconfiguration of the sphere, C and F must simultaneously converge to the origin—this follows from Remark 1. For a given trajectory of C, we infer from Sec. 3.1 that, to converge F simultaneously, the length of the trajectory of C should be equal to the initial length of CF. This is illustrated in Fig. 7(a), where the length of the path M_1 equals CF and hence C and F simultaneously converge to the origin. It may be argued that this approach is not applicable when $CF \leqslant CO$. However, this is not true as illustrated in Fig. 7(b). Here, C first follows a path M_{2a} that unwraps CF so that $C_1F_1 > CF$. The subsequent path M_{2b} then wraps CF to converge both C and F simultaneously to the origin. Note that C_1F_1 is tangential to both the curves M_{2a} and M_{2b} at the point C_1. This is in accordance with the characteristics of the motion of C and F, established in Remark 4. We shall now classify the trajectories of C, i.e. the evolutes, into two different categories

DEFINITION 4. *A "wrapping" evolute is one where the length of CF decreases as C traverses the trajectory.*

Some examples of wrapping trajectories are M_1 and M_{2b} given in Figs. 7(a) and 7(b), respectively. Any arbitrary point on a wrapping trajectory satisfies the condition

$$\hat{u}_t \cdot \frac{\overrightarrow{CF}}{|\overrightarrow{CF}|} = +1$$

where $\hat{u}_t$ is the unit tangent vector to the evolute.

DEFINITION 5. *An "unwrapping" evolute is one where the length of CF increases as C traverses the trajectory.*

An example of an unwrapping trajectory is M_{2a} in Fig. 7(b). An arbitrary point on an unwrapping trajectory satisfies the condition

$$\hat{u}_t \cdot \frac{\overrightarrow{CF}}{|\overrightarrow{CF}|} = -1$$

Remark 5. For partial reconfiguration, the trajectory of C must consist of either a single wrapping evolute, as in Fig. 7(a), or a sequence of alternate unwrapping and wrapping evolutes, such as in Fig. 7(b). Furthermore, the wrapping and unwrapping evolutes should be designed such that

$$\sum_{i=1}^{n}\mathcal{L}[\mathcal{C}_{wi}] - \sum_{j=0}^{m}\mathcal{L}[\mathcal{C}_{uj}] = CF(0) \tag{25}$$

where $CF(0)$ is the initial length of CF, $\mathcal{C}_{wi}$ and $\mathcal{C}_{uj}$ are the *ith* and *jth* wrapping and unwrapping evolutes, respectively, and $\mathcal{L}[\mathcal{C}_{wi}]$ and $\mathcal{L}[\mathcal{C}_{uj}]$ represent their lengths, respectively.

The existence of multiple solution trajectories is intuitive and is captured effectively in Remark 5 by the flexibility of the number and type of wrapping and unwrapping evolutes allowed for partial reconfiguration. Additionally Eq. (25) represents a fundamental constraint for partial reconfiguration of the rolling sphere in planar geometry. The above condition can also be written as

$$\int_M \left\{ \hat{u}_t \cdot \frac{\overrightarrow{CF}}{|\overrightarrow{CF}|} \right\} ds = CF(0) \qquad (26)$$

where $ds=|dx\hat{i}+dy\hat{j}|$, and M is the evolute and the path of integration. The flexibility in designing the evolute-involute pair while satisfying Eqs. (25) or (26) gives us the added freedom of posing a variety of optimization problems. For instance, we can define the partial reconfiguration problem as an isoperimetric problem [7], where the objective is to minimize the path length

$$J = \int_M ds \qquad (27)$$

subject to the integral constraint in Eq. (26). From Eq. (7) we know that complete reconfiguration additionally requires convergence of β to zero. This requires the following integral condition to be satisfied:

$$\int_M \left(\frac{\partial \beta}{\partial r} \right) ds = -\beta_0 \qquad (28)$$

where β_0 is the initial value of β. Eq. (26) alone, and together with Eq. (28), define the integral constraints for the isoperimetric problems for partial and complete reconfiguration, respectively. We do not solve the isoperimetric problem in this paper, instead we propose a class of solution trajectories that satisfy the integral constraints in Eqs. (26) and (28).

4 Partial Reconfiguration

In search of a class of solution trajectories for the problem posed in Sec. 3.2, we refer to the Sweep-Tuck algorithm in Sec. 2. With this algorithm, the sphere is reconfigured by a sequence of alternate circular arc and linear segments [6]. We will now show that under limiting conditions, the Sweep-Tuck algorithm yields a smooth trajectory of the sphere. Subsequently, we will establish that such a trajectory can be a solution for the problem posed by Eqs. (26) and (28).

4.1 Sweep-Tuck Algorithm With Smooth Motion. Consider the distance traversed by C during a DPT maneuver, as given in Eq. (17). For partial reconfiguration, we know that ψ must satisfy Eq. (13). As ψ approaches $\cos^{-1}(1/\lambda)$, we have

$$\lim_{\psi \to \cos^{-1}(1/\lambda)} CC' = \lim_{\psi \to \cos^{-1}(1/\lambda)} \frac{2\lambda CO(\lambda \cos \psi - 1)}{(\lambda^2 - 1)} = 0 \qquad (29)$$

Also, from Eq. (15) we conclude that

$$\lim_{\psi \to \cos^{-1}(1/\lambda)} \psi' = \psi \qquad (30)$$

From Definitions 1 and 2, we deduce that DPT and RS maneuvers become infinitesimally small and conclude that the points C and F, and hence the sphere, follow a smooth trajectory.

To corroborate the conclusion drawn above, we simulate the Sweep-Tuck algorithm for a general case and a case where $\psi \to \cos^{-1}(1/\lambda)$. The simulation results are shown in Figs. 8 and 9, respectively. Both simulations were performed with identical initial configuration of the sphere, given as follows:

$$\boldsymbol{(x\ y\ \theta\ \alpha\ \beta) = (7.0\ 1.0\ 1.4\ 0.5\ 2.5)} \qquad (31)$$

where the units are in meters and radians. The initial condition results in $\lambda=1.684$, and the range of ψ in Eq. (13) to be $0 \leqslant \psi < \psi_{\max} = \cos^{-1}(1/1.684) = 0.935$ rad. For the simulation in Fig. 8, ψ is chosen at 50% of $\psi_{\max}$ ($\psi_1=0.47$ rad), whereas, the choice of ψ is at 95% of $\psi_{\max}$ ($\psi_2=0.89$ rad) for the simulation in Fig. 9. In Figs. 8(a) and 9(a) the motion of F to F_1 is due to the PS and the first RS maneuver. The points F_1 mark the start of the first DPT maneuver. These points are shown by instances t_1 and t_2 in Figs. 8(c) and 9(c), respectively. Beyond these time instants C commences motion and θ starts decreasing due to the DPT maneuvers. The DPT maneuvers in Fig. 8(a) cause significant motion of the sphere but cause infinitesimal motion of the sphere in Fig. 9(a). This is evident in the motion of C which translates along the line

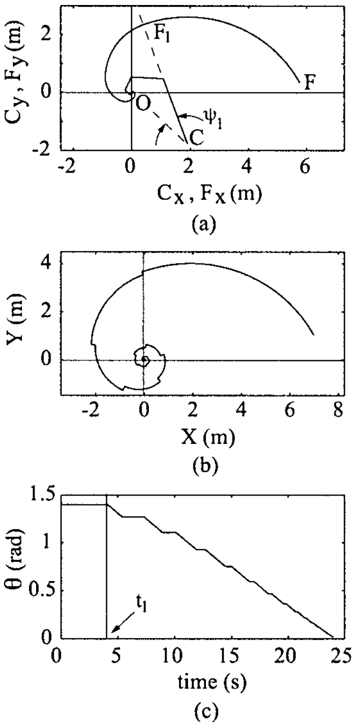


Fig. 8 Reconfiguration with $\psi < \cos^{-1}(1/\lambda)$

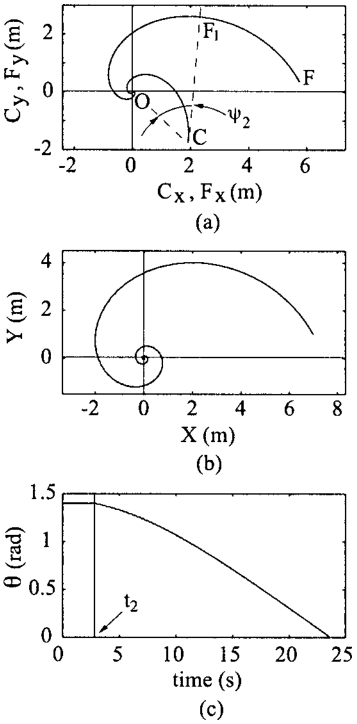


Fig. 9 Reconfiguration with $\psi \approx \cos^{-1}(1/\lambda)$

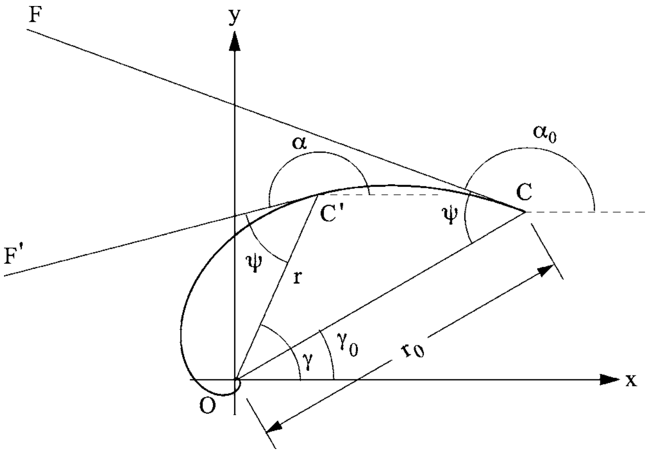


Fig. 10 A counterclockwise logarithmic spiral motion of *C*

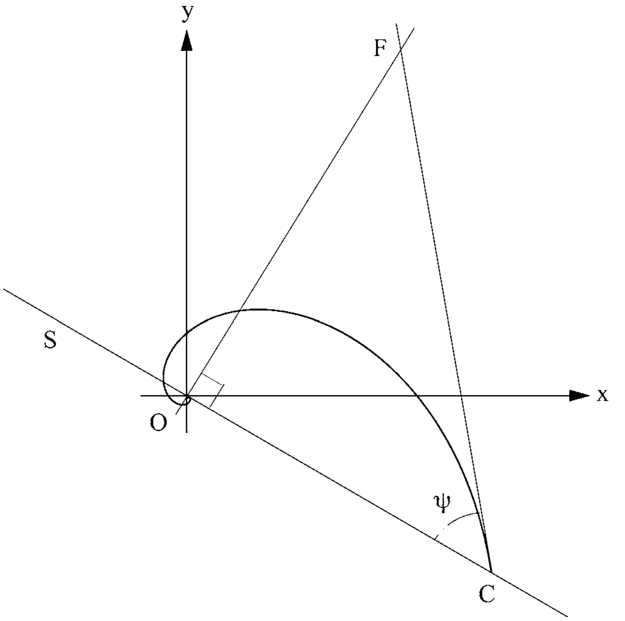


Fig. 11 Locus of *C* that allows partial reconfiguration

segments in Fig. 8(*a*), whereas in Fig. 9(*a*) it follows an almost smooth curve that is tangential to *CF*, thereby confirming Remark 4. It is clear from Fig. 8(*b*) and Fig. 9(*b*) that as ψ approaches $\cos^{-1}(1/\lambda)$, the trajectory of *C* and that of the sphere approach smooth curves. Also, θ decreases along a smooth curve in Fig. 9(*c*) whereas it decreases in alternate time intervals in Fig. 8(*c*).

4.2 Motion Along a Logarithmic Spiral. In a Sweep-Tuck sequence, both *CF* and *CO* decrease in geometric progression [6] at the end of every RS-DPT pair and finally converge to zero. As ψ approaches $\cos^{-1}(1/\lambda)$, the trajectory of *C* approaches a smooth curve with both *CF* and *CO* decreasing continuously to zero, as illustrated in Fig. 9(*a*). Since *CF* decreases, this smooth trajectory of *C* is a wrapping evolute curve according to Definition 4. Also, since Eq. (30) is satisfied, we deduce the following:

Remark 6. If the initial configuration of the sphere satisfies $\lambda \in (1,\infty)$, then, as ψ approaches the limiting value of $cos^{-1}(1/\lambda)$, the trajectory of the *C* generated by the Sweep-Tuck algorithm approaches a smooth wrapping evolute curve. Moreover, the tangent at any point on this curve makes a constant angle ψ with the radius vector $\vec{OC}$. Such a curve is the well known logarithmic spiral. Since any involute of a logarithmic spiral is also a logarithmic spiral [8], the trajectory of *F also approaches a logarithmic spiral motion*.

Consider a logarithmic spiral trajectory of the point *C* as shown in Fig. 10. The angle $\angle OC'F' = \psi$ is constant for any location *C* along its path and *r* represents the distance *OC'* which decreases with an increase in γ in this case. The spiral starts at $\gamma = \gamma_0$ where $r = r_0 = CO$. The mathematical expression of the logarithmic spiral can be derived easily from its definition, as

$$r = \begin{cases} r_0 e^{-(\gamma-\gamma_0)\cot\psi}, & \gamma_0 \leqslant \gamma < \infty \quad \text{for counter clockwise spiral} \\ r_0 e^{+(\gamma-\gamma_0)\cot\psi}, & -\infty < \gamma \leqslant \gamma_0 \quad \text{for clockwise spiral} \end{cases} \tag{32}$$

From Remark 6 it is evident that a logarithmic spiral is a potential solution to the geometric reconfiguration problem posed in Eq. (26). We will now establish that in the limiting case of $\psi = \cos^{-1}(1/\lambda)$, the length condition in Eq. (25) is satisfied. We consider a counterclockwise logarithmic spiral given in Eq. (32), the length of which can be computed using the expression

$$\mathcal{L}_C = \int_0^\infty \sqrt{r^2 + \left(\frac{dr}{d\gamma}\right)^2}\, d\gamma \tag{33}$$

where $\mathcal{L}_C$ is the length of the logarithmic spiral. Upon simplification, Eq. (33) yields

$$\mathcal{L}_C = \frac{r_0}{\cos\psi} = \frac{CO}{\cos\psi} = \lambda CO = CF \tag{34}$$

The same result is obtained upon considering a clockwise logarithmic spiral. This confirms that the logarithmic spiral satisfies Eq. (25) when $\psi = \cos^{-1}(1/\lambda)$ and is a solution trajectory of *C* for partial reconfiguration in the sense of Eq. (12). The choice of logarithmic spiral trajectory for the point *C*, the mathematical form in Eq. (32), and from the result above, we infer the following

Remark 7. A logarithmic spiral trajectory of *C*, converging to the origin, leads to exponential convergence of *CO* and *CF* to zero.

4.3 Partial Reconfiguration Using Logarithmic Spiral Motion of *C*. We concluded in Sec. 4.2 that the sphere can be partially reconfigured using a logarithmic spiral if the sphere configuration satisfies $\lambda \in (1,\infty)$ and $\angle OCF = \psi = \cos^{-1}(1/\lambda)$. Consider the second condition first

$$\cos\psi = (1/\lambda) \Rightarrow \cos\psi = \frac{CO}{CF} \Rightarrow FO \perp CO \tag{35}$$

An interesting observation here is that if *FO* is perpendicular to *CO*, then not only do we satisfy the second condition, but also the first condition. This is because

$$FO \perp CO \Rightarrow CF > CO \Rightarrow \lambda \in (1,\infty) \tag{36}$$

The result above implies that for a given point *F* defined by the initial conditions of the sphere, the point *C* should lie on a straight line perpendicular to *FO* and passing through *O* as shown in Fig. 11. Given *F*, the straight line *S* is the locus of the point *C* such that $FO \perp CO$ and $\cos\psi = 1/\lambda$. Therefore, for a given location of *F*, *S* is effectively the locus of the point *C* that allows the use of logarithmic spiral motion of *C* for partial reconfiguration.

We now devise a simple three step algorithm for achieving partial reconfiguration of the sphere from an arbitrary initial configuration. The three step algorithm is stated with the help of the next remark.

Remark 8. The rolling sphere can be partially reconfigured by the following three steps:

(1) Apply actuation (A) to make $\theta = 0$, i.e., make *C* and *F* coincident.

(2) Apply actuation (A) to move C to any point on the straight line that is perpendicular to FO and passing through O.
(3) Use logarithmic spiral motion of C with CF as the tangent to achieve partial reconfiguration.

The flexibility of choosing any point C along the line S will be utilized in the next section to achieve complete reconfiguration. We complete this section by deriving the expressions for the actuating inputs ω_x^1 and ω_y^1 that generate a logarithmic spiral motion of C. Differentiating the expression for x in Eq. (18) with respect to time, we have

$$\dot{x} = \dot{C}_x + \sec^2\theta\cos\alpha\dot{\theta} - \tan\theta\sin\alpha\dot{\alpha} \tag{37}$$

Considering a counterclockwise logarithmic spiral, from Fig. 10, C_x and $\dot{C}_x$ can be expressed as

$$C_x = r\cos\gamma \quad \dot{C}_x = -COe^{-(\gamma-\gamma_0)\cot\psi}(\cot\psi\cos\gamma + \sin\gamma) \tag{38}$$

where the expression for r is given in Eq. (32) and CO is the initial distance of C from the origin. Here we assume $\dot{\gamma}=1 \equiv \gamma = t+\gamma_0$ without any loss of generality. To obtain the expressions for ω_x^1 and ω_y^1 we note from Fig. 10, that for a counterclockwise logarithmic spiral

$$\alpha = \pi - (\psi - \gamma) \Rightarrow \dot{\alpha} = \dot{\gamma} = 1 \tag{39}$$

Substituting Eqs. (1), (3), (4), and (38) into Eq. (37) and using Eq. (39), we obtain the following expressions for ω_y^1 and ω_x^1

$$\omega_y^1 = \frac{-CO}{\tan^2\theta\sin\psi}e^{-(\gamma-\gamma_0)\cot\psi}, \quad \omega_x^1 = -\tan\theta \tag{40}$$

Similar expressions can be deduced for a clockwise logarithmic spiral. Thus, the actuating inputs ω_x^1 and ω_y^1 that specifically generate a logarithmic spiral motion of C are smooth functions of time and the states of the system.

5 Complete Reconfiguration

From Fig. 11 it is evident that different choice of ψ in the second step of Remark 8 can result in different C along the locus S. When $\psi=\pi/2$, C coincides with O, and as ψ reduces to zero C moves farther away from O. We now show that the different logarithmic spirals generated due to the different end points C result in different $\Delta\beta$. Consider a counterclockwise logarithmic spiral given by Eq. (32). From Eqs. (33), (23), and (24), and Remark 5, we write

$$\Delta\mathcal{L} = \sqrt{r^2 + \left(\frac{dr}{d\gamma}\right)^2}\Delta\gamma = |\Delta CF| = -\tan^2\theta\Delta\theta \tag{41}$$

Rewriting Eq. (41) as

$$r_0\csc\psi e^{-(\gamma-\gamma_0)\cot\psi}d\gamma = -\tan^2\theta d\theta \tag{42}$$

and integrating both sides of Eq. (42), we deduce the following:

$$e^{-(\gamma-\gamma_0)\cot\psi} = 1 - \frac{\cos\psi}{r_0}[(\tan\theta_0 - \theta_0) - (\tan\theta - \theta)] \tag{43}$$

Combining Eqs. (42) and (43) we have

$$d\gamma = -\frac{\tan^2\theta}{r_0\csc\psi - \cot\psi[(\tan\theta_0 - \theta_0) - (\tan\theta - \theta)]}d\theta \tag{44}$$

Also, from Eqs. (10), (39), and (44), we deduce the following:

$$d\alpha = d\gamma \Rightarrow d\beta = d\gamma(1 - \sec\theta) \Rightarrow d\beta = \tan\psi\frac{\tan^2\theta(\sec\theta - 1)}{(\tan\theta - \theta)}d\theta \tag{45}$$

Thus, the total change in β due a counter clockwise logarithmic spiral motion of C can be given by the following expression

$$\Delta\beta = -\tan\psi\int_0^{\theta_0}\frac{\tan^2\theta(\sec\theta - 1)}{(\tan\theta - \theta)}d\theta \tag{46}$$

where, ψ is the constant angle $\angle OCF$ and θ_0 is the value of θ at the beginning of the logarithmic spiral motion of C. Consider the actuation (A) in second step of the algorithm presented in Remark 8 in Sec. 4.3. Using Fig. 11, ψ and θ_0 can be related by the expression

$$\frac{OF}{\sin\psi} = CF = \tan\theta_0 - \theta_0 \tag{47}$$

Clearly, different choices of ψ will result in different values of θ_0 in Eq. (47). As $\Delta\beta$ is a function of θ_0, Eq. (46), this in turn will yield different values of $\Delta\beta$. This is the key to complete reconfiguration of the sphere. From Fig. 11 clearly $0<\psi\leqslant\pi/2$. Also, from Eq. (47) we have

$$0 < \psi \leqslant \pi/2 \Rightarrow \pi/2 > \theta_0 \geqslant \bar{\theta}_0 \ \forall \ OF \neq 0 \tag{48}$$

where $(OF/\sin\psi)|_{\psi=\pi/2}=OF=\tan\bar{\theta}_0-\bar{\theta}_0$. Note that, from Fig. 11 and Eq. (47), Eq. (46) can be rewritten as

$$\Delta\beta = -\frac{OF}{\sqrt{(\tan\theta_0 - \theta_0)^2 - OF^2}}\int_0^{\theta_0}\frac{\tan^2\theta(\sec\theta - 1)}{(\tan\theta - \theta)}d\theta \tag{49}$$

From the expression of $\Delta\beta$ in Eqs. (49) and (46), and from Eq. (48), one can deduce the following limits on $\Delta\beta$

$$\lim_{\theta_0\to\pi/2}\Delta\beta = -OF \Rightarrow \lim_{\psi\to 0}\Delta\beta = -OF$$

$$\lim_{\theta_0\to\bar{\theta}_0}\Delta\beta = -\infty \Rightarrow \lim_{\psi\to\pi/2}\Delta\beta = -\infty \tag{50}$$

It can be verified that the expression of $\Delta\beta$ in Eq. (49) is monotonic in θ_0. This leads to the following range of $\Delta\beta$;

$$0 < \psi \leqslant \pi/2 \Rightarrow \begin{cases} -OF \geqslant \Delta\beta > -\infty: & \text{counterclockwise logarithmic spiral} \\ OF \leqslant \Delta\beta < \infty: & \text{clockwise logarithmic spiral} \end{cases} \tag{51}$$

In Remark 8 note that β changes only during the logarithmic spiral motion in the third step and remains invariant during the first and second step of the partial reconfiguration algorithm. If β_0 is the initial value of β, then for complete reconfiguration,

$$\Delta\beta = -\beta_0 \tag{52}$$

Eq. (51) apparently puts a restriction based on the distance OF by imposing $-OF\leqslant\Delta\beta\leqslant OF$ to be an unachievable range of $\Delta\beta$. However, we can consider an equivalent β_0

$$\beta_{0\text{eq}} = \beta_0 \pm 2n\pi \tag{53}$$

where $n=1,2,\ldots$, such that the effective $\Delta\beta$, $\Delta\beta_{\text{eff}}$, given by

$$\Delta\beta_{\text{eff}} = -\beta_{0\text{eq}} \tag{54}$$

satisfies Eq. (51). This implies that any desired $\Delta\beta$ or its equivalent can be achieved, by appropriately choosing a point on the line S. Thus, while partial reconfiguration of the sphere can be achieved by following a logarithmic spiral trajectory of C starting from any point on the locus S as shown in Fig. 11, complete reconfiguration in the sense of Eq. (7) can be achieved only from specific points on S. These points are such that the corresponding effective $\Delta\beta$ satisfies the relation

$$\Delta\beta_{\text{eff}} = \begin{cases} -\beta_{0eq}: & \Delta\beta \in (-OF, OF) \\ -\beta_0: & \Delta\beta \in (-\infty, -OF] \cup [OF, \infty) \end{cases} \tag{55}$$

We now modify our algorithm in Remark 8 to incorporate complete reconfiguration of the sphere as follows:

Remark 9. The rolling sphere can be completely reconfigured by applying the following three step algorithm:

(1) Apply actuation (A) to make $\theta=0$, i.e., make C and F coincident.
(2) Apply actuation (A) to move C to a point on the straight line S where $\psi = \angle OCF$ is such that $\Delta\beta_{\text{eff}}$ satisfies Eq. (55).
(3) Use logarithmic spiral motion of C with CF as the tangent to achieve complete reconfiguration.

We have defined the line S to be perpendicular to OF passing through the origin O. Consider the case when initial configuration of the algorithm makes F and O coincident. Then, the line OF degenerates to the origin and the line S is undefined. Although this is a special case of the algorithm, it can be handled easily. Consider first the case when F coincides with the origin but C does not. In this case we first apply an actuation (B) that causes O, C, and F to lie on a straight line and in that order. Subsequently, the complete reconfiguration algorithm in Remark 9 can be applied. Next consider the case when both C and F coincide with O. Then we first apply an actuation (A) to move C away from the origin. The rest follow exactly the same steps as the first case.

6 Simulations

In this section we show simulation results of the complete reconfiguration algorithm presented in Remark 9. The initial conditions of the sphere for this simulation are as follows:

$$(x\ y\ \theta\ \alpha\ \beta) \equiv (5\ 1.5\ 1.3\ -\pi/2\ -2.5) \tag{56}$$

where the units are in meters and radians. The initial conditions yield $OF=5.5036$ m, which implies that the minimum $|\Delta\beta|$ achievable is 5.5036 rad, whereas the necessary $\Delta\beta=2.5$ rad. Hence we consider an effective $\Delta\beta$ as follows:

$$|\Delta\beta_{\text{eff}}| = |2.5 + 2m\pi| > 5.5036 \text{ rad}, \quad \text{where } m = \pm 1, \pm 2, \ldots \tag{57}$$

We choose $m=1$ which yields $\Delta\beta=8.7832$ rad and an equivalent $\beta_{0_{\text{eq}}}=-8.7832$ rad. The simulation results are given in Fig. 12.

Figure 12(a) illustrates the C and F trajectories. The first step of the complete reconfiguration algorithm, Remark 9, where actuation (A) is applied, causes motion of C to C_1 which is coincident with F. In Fig. 12(c) this corresponds to the linear decrease of θ from 1.3 rad to zero at t_1. This is followed by the second step where actuation (A) takes C to C_2. In Fig. 12(c) this refers to the linear increase of θ from zero to 1.47 rad at t_2. The equivalent β remains constant at -8.7832 rad from $t=0$ to $t=t_2$. The point C_2 is such that $\angle C_2OF=\pi/2$ and this allows partial reconfiguration by a subsequent logarithmic spiral motion of C_2 to the origin. Also, with this choice of C_2, $\psi=\angle OC_2F=0.7019$ rad and the subsequent logarithmic spiral will cause $\Delta\beta=8.7832$ rad and thereby guarantee complete reconfiguration by additionally converging β to the origin. These initial maneuvers results in $\lambda=1.31$ (the algorithm guarantees that λ will necessarily be greater than 1). The subsequent logarithmic spiral motion converges C_2 and F simultaneously to the origin. The trajectories C_2O and FO form an evolute-involute pair. The resulting convergence of the sphere-center to the origin is shown in Fig. 12(b). The convergence of θ and β to the origin are illustrated in Figs. 12(c) and 12(d).

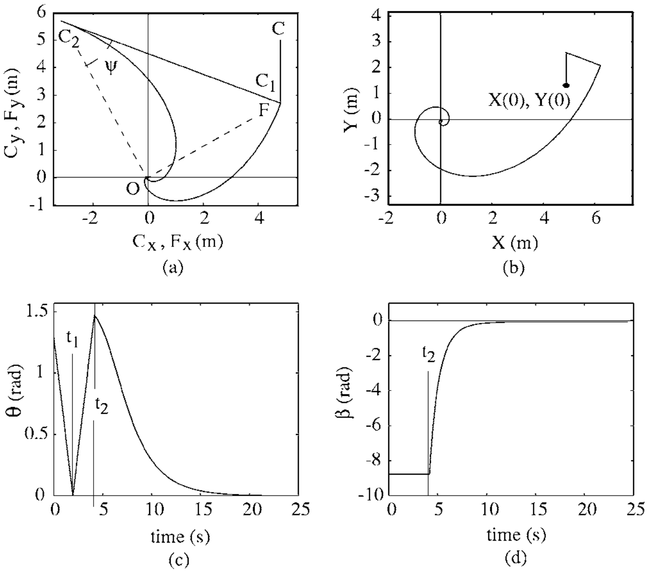


Fig. 12 Simulation showing complete reconfiguration using logarithmic spiral motion of *C*

7 Conclusion

In this paper we recast the classical problem of reconfiguration of a rolling sphere to a problem in planar geometry. We show that the rolling motion of a sphere is characterized by wrapping and unwrapping of a taut rope on a planar curve. The problem of reconfiguration therefore translates to that of designing an evolute-involute pair that originate at the initial configuration of the sphere and terminate at the desired configuration, while satisfying integral constraints. This geometric problem can be posed as an isoperimetric optimization problem, but rather than solving this problem directly, we obtain a class of solution trajectories where the evolute-involute pair are logarithmic spirals. It is shown that two preliminary maneuvers followed by a maneuver generated by the logarithmic spirals result in complete reconfiguration, and with exponential convergence of the configuration variables. We provide numerical simulations to illustrate the reconfiguration algorithm.

Acknowledgment

The authors gratefully acknowledge the support provided by National Science Foundation, NSF Grant No. CMS-9800343, in conducting this research.

References

[1] Hammersley, J., 1983, "Oxford Commemoration Ball," London Mathematical Society Lecture Notes, **79**, pp. 112–142.
[2] Li, Z., and Canny, J., 1990, "Motion of Two Rigid Bodies with Rolling Constraints," IEEE Trans. Rob. Autom., **6**(1), pp. 62–72.
[3] Jurdjevic, V., 1993, "The Geometry of the Plate-Ball Problem," Arch. Ration. Mech. Anal., **124**, pp. 305–328.
[4] Mukherjee, R., Minor, M., and Pukrushpan, J. T., 2002, "Motion Planning for a Spherical Mobile Robot: Revisiting the Classical Ball-Plate Problem," ASME J. Dyn. Syst., Meas., Control, **124**(4), pp. 502–511.
[5] Krener, A. J., and Nikitin, S., 1997, "Generalized Isoperimetric Problem," Journal of Mathematical Systems, Estimation, and Control, **7**(3), pp. 1–15.
[6] Das, T. , and Mukherjee, R., 2004, "Exponential Stabilization of the Rolling Sphere," Automatica, **40**, pp. 1877–1889.
[7] Courant, R., and Hilbert, D., 2003, *Methods of Mathematical Physics*, Wiley Interscience, New York, pp. 135–136.
[8] http://mathworld.wofram.com/Involute.html

Sourav Banerjee
e-mail: souravban@netscape.net

Tribikram Kundu[1]
e-mail: tkundu@email.arizona.edu

Department of Civil Engineering
and Engineering Mechanics,
University of Arizona,
Tucson, AZ 85721

Dominique Placko
Ecole Normale Superieure,
SATIE, 62, av. Du President Wilson,
F-94235 Cachan cedex,
France
e-mail: dominique.plcko@satie.ens-cachan.fr

Ultrasonic Field Modeling in Multilayered Fluid Structures Using the Distributed Point Source Method Technique

In the field of nondestructive evaluation (NDE), the newly developed distributed point source method (DPSM) is gradually gaining popularity. DPSM is a semi-analytical technique used to calculate the ultrasonic field (pressure and velocity fields) generated by ultrasonic transducers. This technique is extended in this paper to model the ultrasonic field generated in multilayered nonhomogeneous fluid systems when the ultrasonic transducers are placed on both sides of the layered fluid structure. Two different cases have been analyzed. In the first case, three layers of nonhomogeneous fluids constitute the problem geometry; the higher density fluid is sandwiched between two identical fluid half-spaces. In the second case, four layers of nonhomogeneous fluids have been considered with the fluid density monotonically increasing from the bottom to the top layer. In both cases, analyses have been carried out for two different frequencies of excitation with various orientations of the transducers. As expected, the results show that the ultrasonic field is very sensitive to the fluid properties, the orientation of the fluid layers, and the frequency of excitation. The interaction effect between the transducers is also visible in the computed results. In the pictorial view of the resulting ultrasonic field, the interface between two fluid layers can easily be seen. [DOI: 10.1115/1.2164516]

1 Introduction

Ultrasonic nondestructive evaluation (NDE) generally has two objectives: to detect defects in structures and to determine the material properties of the structure, without damaging it. Ultrasonic transducers are commonly used to generate ultrasonic waves used in NDE experiments. Transducers are used both as transmitters and receivers. Multilayered fluid systems excited by multiple transducers are modeled in this paper. Multilayered fluid structure is not that uncommon in nature. For example, in the early stage of pregnancy, in the human female body the embryo grows in layered fluid surroundings. The plasma of any cell, suspended in a fluid inside or outside of an animal body, is also an example of a layered fluid system. Human eye lenses also behave almost like layered fluid structures. Elastic properties of biological materials are needed for understanding their interaction with implant materials. Layers of crude oil in sea water, as seen after an oil spill in the ocean, are another example of layered fluid structure.

Ultrasonic nondestructive testing can be used to find the acoustical properties and thickness of different fluid layers in a multilayered fluid structure. With these applications in mind, an efficient semi-analytical tool has been developed in this paper to model the ultrasonic field in multilayered fluid structures. The method is based on the DPSM (distributed point source method) technique originally developed to model ultrasonic (or eddy current) fields, i.e., pressure and velocity fields (or magnetic fields) generated by ultrasonic (or eddy current) transducers. DPSM technique for ultrasonic field modeling was first developed by Placko and Kundu [1]. They used this technique to model ultrasonic fields in a homogeneous fluid medium, and then it was extended to include nonhomogeneous fluids with only one interface [2]. The technique is then extended by Lee et al. [3] to the inhomogeneous medium composed of one fluid half-space and one solid half-space following the Rayleigh-Sommerfeld theory. Lee et al. considered only 2D (two-dimensional) problem geometry. The interaction between two transducers, for different transducer arrangements and source strengths, placed in a homogeneous fluid has been studied by Ahmad et al. [4]. The scattered ultrasonic field generated by a solid scatterer of finite dimension placed in a homogeneous fluid has also been modeled by the DPSM technique [5]. Ahmad et al. [6] recently extended this technique to study the phased-array ultrasonic transducers in a homogeneous fluid. However, the ultrasonic field generated in a multilayered fluid system by multiple ultrasonic transducers of finite dimension has not been studied yet. In this paper, the multilayered fluid system is modeled. To study the interaction effect, we have considered two transducers on opposite sides of a multilayered fluid structure. When only one transducer is active, the second transducer acts as a scatterer. When both transducers are active they work as ultrasonic signal generators as well as scatterers. The scattering phenomenon is also known as the interaction effect between the two transducers. In our approach two transducers and various interfaces have been modeled by properly considering the interaction effect. In other words, the total field presented here is not simply the superposition of the ultrasonic fields generated by individual transducers working alone in absence of the other transducers.

2 Theory

2.1 Distributed Point Source Method (DPSM). The basic principle of distributed point source method for the modeling of ultrasonic and magnetic transducers has been described in detail in Refs. [7,8]. In this method arrays of point sources are distributed on the surface of the transducer and along interfaces. These surfaces act as wave-generating surfaces. Each point source is assumed to generate an ultrasonic field and the ultrasonic field at any point at a distance from the surface of the transducer and the

---

[1] Author to whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 23, 2005; final manuscript received October 20, 2005. Review conducted by H. D. Espinosa. Discussion on the paper should be addressed to the Editor, Professor Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

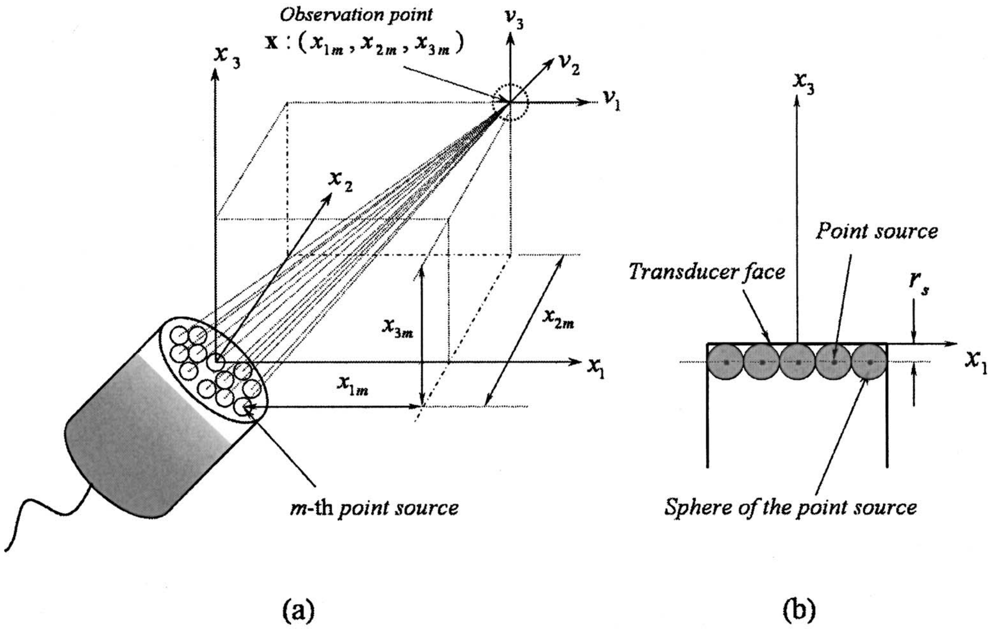


Fig. 1 (*a*) **Position of an observation point (target point) and its distance from the *m*th point source on the transducer surface. (*b*) Side view of a transducer and actual positions of the point sources.**

interface is obtained by superposition of the contributions of all point sources distributed along appropriate surfaces as shown in Fig. 1(*a*). More precisely, as discussed in Ref. [7], each point source is placed inside a small sphere and the spheres are placed just behind the active elastic wave-generating surface at a distance r_s from the surface; the apex of the spheres touches the transducer surface or interface as illustrated in Fig. 1(*b*). Restrictions on r_s have been discussed in Ref. [7]. Reliability and accuracy of DPSM generated results for some simpler cases for which other forms of solution are available have been already presented in Ref. [6] and are not repeated here.

2.2 Computation of Ultrasonic Field in a Homogeneous Fluid.

2.2.1 Computation of Velocity and Pressure Fields in a Fluid Generated by a Set of Point Sources. Ultrasonic transducers of finite dimension are generally used to generate the ultrasonic field in the fluid. In the DPSM technique the point sources are distributed near the transducer surface [1]. The combined effect of a large number of point sources distributed over a plane surface is the vibration of the surface points in a direction normal to the surface. Following the surface integral technique the pressure field at point $\mathbf{x}$ (see Fig. 1(*a*)) in front of the transducer can be obtained by integrating the spherical waves [1,9]

$$p(\mathbf{x}) = \int_S B \cdot \frac{e^{ik_f r}}{4\pi r} dS \tag{1}$$

where p is the pressure and B represents the source velocity. Therefore, B is a measure of the strength of the point sources, dS represents the elementary area on the surface of the source, k_f is the wave number in the fluid for a particular frequency of excitation, and r is the radial distance of the point $\mathbf{x}$ from the source. This integral equation can also be written in the summation form

$$p(\mathbf{x}) = \sum_{m=1}^{N} \left(\frac{B}{4\pi}\Delta S_m\right) \frac{\exp(ik_f r_m)}{r_m} = \sum_{m=1}^{N} A_m \frac{\exp(ik_f r_m)}{r_m} \tag{2}$$

Let the source strength of the mth point source be A_m, and r_m be the distance of the observation point or target point $\mathbf{x}$ from the mth point source. Therefore, pressure at any point at a distance r_m from the mth point source with source strength A_m can be written as

$$p_m(r) = A_m \frac{\exp(ik_f r_m)}{r_m} \tag{3}$$

For N number of point sources distributed on a surface, the pressure at the target point is given by

$$p(\mathbf{x}) = \sum_{m=1}^{N} p_m(r_m) = \sum_{m=1}^{N} A_m \frac{\exp(ik_f r_m)}{r_m} \tag{4}$$

From the pressure velocity relation [7], the velocity in the radial direction at a distance r from the mth point source can be written as

$$v_m(r) = \frac{A_m}{i\omega\rho}\frac{\partial}{\partial r}\left(\frac{\exp(ik_f r)}{r}\right) = \frac{A_m}{i\omega\rho}\left(\frac{ik_f \exp(ik_f r)}{r} - \frac{\exp(ik_f r)}{r^2}\right)$$

$$\therefore v_m(r) = \frac{A_m}{i\omega\rho}\frac{\exp(ik_f r)}{r}\left(ik_f - \frac{1}{r}\right) \tag{5}$$

The three components of velocity are

$$v_{1m}(r) = \frac{A_m}{i\omega\rho}\frac{\partial}{\partial x_1}\left(\frac{\exp(ik_f r)}{r}\right) = \frac{A_m}{i\omega\rho}\frac{x_1 \exp(ik_f r)}{r^2}\left(ik_f - \frac{1}{r}\right) \tag{6}$$

$$v_{2m}(r) = \frac{A_m}{i\omega\rho}\frac{\partial}{\partial x_2}\left(\frac{\exp(ik_f r)}{r}\right) = \frac{A_m}{i\omega\rho}\frac{x_2 \exp(ik_f r)}{r^2}\left(ik_f - \frac{1}{r}\right) \tag{7}$$

$$v_{3m}(r) = \frac{A_m}{i\omega\rho}\frac{\partial}{\partial x_3}\left(\frac{\exp(ik_f r)}{r}\right) = \frac{A_m}{i\omega\rho}\frac{x_3 \exp(ik_f r)}{r^2}\left(ik_f - \frac{1}{r}\right) \tag{8}$$

When the contributions of all N sources are added, the total velocity in x_1, x_2, and x_3 directions at point $\mathbf{x}$ can be written as

$$v_1(\mathbf{x}) = \sum_{m=1}^{N} v_{1m}(r_m) = \sum_{m=1}^{N} \frac{A_m}{i\omega\rho}\frac{x_{1m}\exp(ik_f r_m)}{r_m^2}\left(ik_f - \frac{1}{r_m}\right) \tag{9}$$

$$v_2(\mathbf{x}) = \sum_{m=1}^{N} v_{2m}(r_m) = \sum_{m=1}^{N} \frac{A_m}{i\omega\rho} \frac{x_{2m}\exp(ik_f r_m)}{r_m^2}\left(ik_f - \frac{1}{r_m}\right) \quad (10)$$

$$v_3(\mathbf{x}) = \sum_{m=1}^{N} v_{3m}(r_m) = \sum_{m=1}^{N} \frac{A_m}{i\omega\rho} \frac{x_{3m}\exp(ik_f r_m)}{r_m^2}\left(ik_f - \frac{1}{r_m}\right) \quad (11)$$

where x_{im} is the shortest distance along the x_i direction between the mth point source and the target point, as shown in Fig. 1(a). If the transducer surface velocity in the x_3 direction is given by v_0, then for all values of $\mathbf{x}$ on the transducer surface, the velocity should be equal to v_0. Therefore

$$v_3(\mathbf{x}) = \sum_{m=1}^{N} \frac{A_m}{i\omega\rho} \frac{x_{3m}\exp(ik_f r_m)}{r_m^2}\left(ik_f - \frac{1}{r_m}\right) = v_0 \quad (12)$$

If the transducer face is inclined at an angle θ, measured from x_3-axis, and is rotated about the x_2-axis (see Fig. 2), then the velocity of the transducer face can be expressed as

$$\begin{aligned} v_0 &= v_1(\mathbf{x})\sin\theta + v_3(\mathbf{x})\cos\theta \\ &= \sum_{m=1}^{N} \frac{A_m}{i\omega\rho}\left(ik_f - \frac{1}{r_m}\right) \\ &\quad \times \left(\frac{x_{1m}\exp(ik_f r_m)}{r_m^2}\sin\theta + \frac{x_{3m}\exp(ik_f r_m)}{r_m^2}\cos\theta\right) \quad (13) \end{aligned}$$

*2.2.2 **Matrix Formulation***. Velocity of the M target points distributed on the transducer face due to point sources distributed just behind the transducer surface at a distance r_s can be written in the matrix form as

$$\mathbf{V}_S = \mathbf{M}_{SS}\mathbf{A}_S \quad (14)$$

where $\mathbf{V}_S$ is the $(M\times 1)$ vector of the velocity components, perpendicular to the transducer surface. If the velocity of the transducer face is given by v_0, then $\mathbf{V}_S$ can be written as

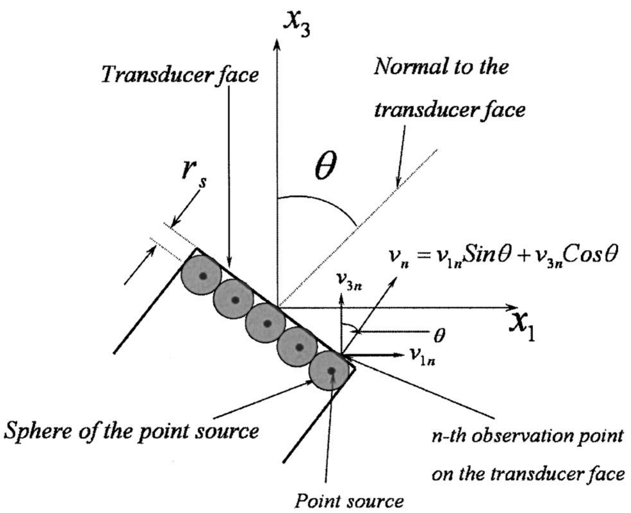

x_3
Transducer face
Normal to the transducer face
r_s
θ
$v_n = v_{1n}Sin\theta + v_{3n}Cos\theta$
v_{3n}
x_1
θ
v_{1n}
Sphere of the point source
n-th observation point on the transducer face
Point source

Fig. 2 Rotation of the transducer with respect to x_3-axis and velocity of the nth observation point adjacent to the transducer face

$$\{\mathbf{V}_S\}^T = [v_0^1\ v_0^2\ v_0^3\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ v_0^{M-1}\ v_0^M]^T \quad (15)$$

where v_0^n is the velocity of the nth target point. If $\mathbf{A}_S$ is the $(N\times 1)$ vector of the source strengths, then

$$\{\mathbf{A}_S\}^T = [A_1\ A_2\ A_3\ A_4\ A_5\ A_6\ \cdots\ A_{N-2}\ A_{N-1}\ A_N]^T \quad (16)$$

From the earlier discussion, we know that each point source is placed inside a sphere; hence, the number of apex points of the spheres will be the same as the number of point sources. When the target points are placed at the apex of the spheres of the point sources, then M is equal to N. Therefore, when the target points are located at the apex of the spheres of the point sources, the square matrix $\mathbf{M}_{SS}$ can be written as

$$\mathbf{M}_{SS} = \begin{bmatrix} f(x_{t1}^1, r_1^1) & f(x_{t1}^2, r_1^2) & f(x_{t1}^3, r_2^3) & f(x_{t1}^4, r_1^4) & \cdots & \cdots & f(x_{t1}^{N-1}, r_1^{N-1}) & f(x_{t1}^N, r_1^N) \\ f(x_{t2}^1, r_2^1) & f(x_{t2}^2, r_2^2) & f(x_{t2}^3, r_2^3) & f(x_{t2}^4, r_2^4) & \cdots & \cdots & f(x_{t2}^{N-1}, r_2^{N-1}) & f(x_{t2}^N, r_2^N) \\ f(x_{t3}^1, r_3^1) & f(x_{t3}^2, r_3^2) & f(x_{t3}^3, r_3^3) & f(x_{t3}^4, r_3^4) & \cdots & \cdots & f(x_{t3}^{N-1}, r_3^{N-1}) & f(x_{t3}^N, r_3^N) \\ f(x_{t4}^1, r_4^1) & f(x_{t4}^2, r_4^2) & f(x_{t4}^3, r_4^3) & f(x_{t4}^4, r_4^4) & \cdots & \cdots & f(x_{t4}^{N-1}, r_4^{N-1}) & f(x_{t4}^N, r_4^N) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ f(x_{tN}^1, r_N^1) & f(x_{tN}^2, r_N^2) & f(x_{tN}^3, r_N^3) & f(x_{tN}^4, r_N^4) & \cdots & \cdots & f(x_{tN}^{N-1}, r_N^{N-1}) & f(x_{tN}^N, r_N^N) \end{bmatrix}_{N\times N} \quad (17)$$

where

$$\begin{aligned} f(x_{tn}^m, r_n^m) &= \frac{x_{tn}^m \exp(ik_f r_n^m)}{i\omega\rho (r_n^m)^2}\left(ik_f - \frac{1}{r_n^m}\right) \\ &= \frac{\exp(ik_f r_n^m)}{i\omega\rho (r_n^m)^2}\left(ik_f - \frac{1}{r_n^m}\right)(x_{3n}^m \cos\theta + x_{1n}^m \sin\theta) \quad (18) \end{aligned}$$

and r_n^m is the distance between the mth point source and the nth target point.

For a general set of target points located on any surface, the velocity due to the transducer sources can be written as

$$\mathbf{V}_T = \mathbf{M}_{TS}\mathbf{A}_S \quad (19)$$

where $\mathbf{V}_T$, the $(M\times 1)$ velocity vector, contains the normal velocity components of the target points distributed on the surface. The matrix $\mathbf{M}_{TS}$ has elements that are similar to those of $\mathbf{M}_{SS}$, with different x_{tn}^m values and the size of the matrix is $(M\times N)$, where M is the number of target points and N is the number of source points. Following the same concept, the pressure at any M number of target points due to N number of source points can be written as

$$\underline{\mathbf{PR}}_T = \mathbf{Q}_{TS}\mathbf{A}_S \quad (20)$$

where $\underline{\mathbf{PR}}_T$ is the $(M\times 1)$ vector of pressure values at M target points, and $\mathbf{Q}_{TS}$ is an $(M\times N)$ matrix given below

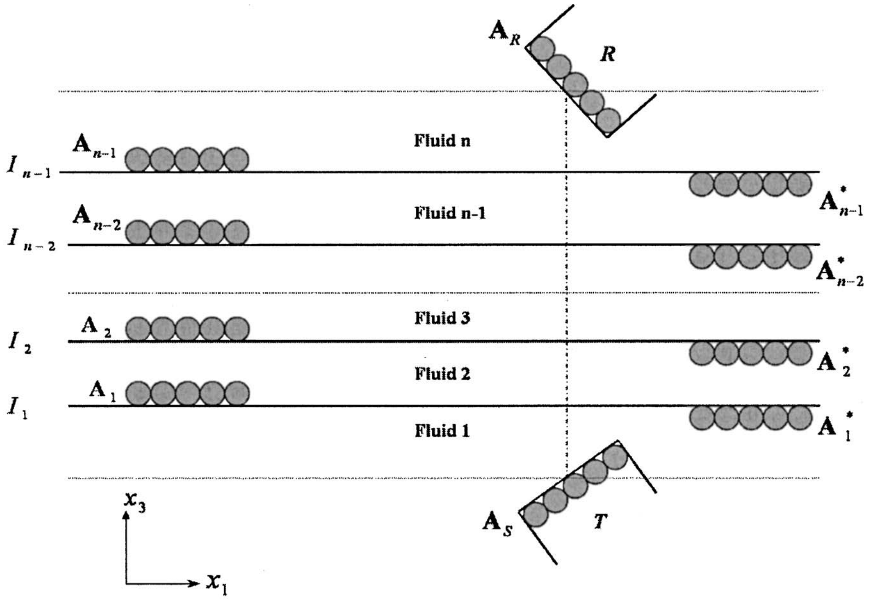


Fig. 3 Distribution of point sources in the multilayered fluid system with two transducers and n different fluids having $(n-1)$ interfaces

$$\mathbf{Q}_{TS}=\begin{bmatrix} \frac{\exp(ik_f r_1^1)}{r_1^1} & \frac{\exp(ik_f r_1^2)}{r_1^2} & \frac{\exp(ik_f r_1^3)}{r_1^3} & \cdots\ \cdots & \frac{\exp(ik_f r_1^M)}{r_1^M} \\ \frac{\exp(ik_f r_2^1)}{r_2^1} & \frac{\exp(ik_f r_2^2)}{r_2^2} & \frac{\exp(ik_f r_2^3)}{r_2^3} & \cdots\ \cdots & \frac{\exp(ik_f r_2^M)}{r_2^M} \\ \frac{\exp(ik_f r_3^1)}{r_3^1} & \frac{\exp(ik_f r_3^2)}{r_3^2} & \frac{\exp(ik_f r_3^3)}{r_3^3} & \cdots\ \cdots & \frac{\exp(ik_f r_3^M)}{r_3^M} \\ \cdots & \cdots & \cdots & \cdots\ \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots\ \cdots & \cdots \\ \frac{\exp(ik_f r_N^1)}{r_N^1} & \frac{\exp(ik_f r_N^2)}{r_N^2} & \frac{\exp(ik_f r_N^3)}{r_N^3} & \cdots\ \cdots & \frac{\exp(ik_f r_N^M)}{r_N^M} \end{bmatrix}_{N\times M} \tag{21}$$

When the target points are located at the apex of the spheres of the point sources, Eq. (20) takes the form

$$\underline{\mathbf{PR}}_S=\mathbf{Q}_{SS}\mathbf{A}_S \tag{22}$$

where $\mathbf{Q}_{SS}$ is an $(N\times N)$ matrix.

The definition of r_n^m is identical to Eq. (18). It is the distance between the mth point source and the nth target point.

2.3 Computation of Ultrasonic Fields in Multilayered Fluids.

2.3.1 Distribution of Point Sources. We are interested in computing the ultrasonic field in multilayered fluid systems. In the multilayered problem geometry several interfaces may be present. When fluids with different densities and acoustic properties form a multilayered system, the fluid density should monotonically vary from top to bottom. If we have n number of fluids in the system, we should have $(n-1)$ number of interfaces. Each interface acts as a transmitter as well as a reflector of elastic wave energy generated by the ultrasonic transducers. When the entire system is considered, several continuity conditions across the interfaces and boundary conditions at the transducer surface are to be satisfied. As shown in Fig. 3, a number of sources are introduced along the transducer surfaces T and R as well as along $(n-1)$ interfaces $I_1, I_2, \ldots, I_{n-1}$, for multilayered fluid geometry with n fluid layers. $\mathbf{A}_S$ and $\mathbf{A}_R$ denote the source strength vectors along the transducer surfaces T and R, respectively. Along each interface I_m two sets of source vectors are introduced. These two sets of source strength vectors are denoted by $\mathbf{A}_m$ (for sources located just above the mth interface) and $\mathbf{A}^*_m$ (for sources located just below the mth interface). The sources with source strength $\mathbf{A}_m$ generate the ultrasonic field in the fluid below it and the sources with source strength $\mathbf{A}^*_m$ generate the ultrasonic field in the fluid above it. The total ultrasonic field in each medium is obtained by superimposing the fields generated by two sets of sources as listed below:

Fluid 1: Summation of fields generated by $\mathbf{A}_S$ and $\mathbf{A}_1$.
Fluid 2: Summation of fields generated by $\mathbf{A}^*_1$ and $\mathbf{A}_2$.
Fluid 3: Summation of fields generated by $\mathbf{A}^*_2$ and $\mathbf{A}_3$.
Fluid $(n-1)$: Summation of fields generated by $\mathbf{A}^*_{n-2}$ and $\mathbf{A}_{n-1}$.Fluid n: Summation of fields generated by $\mathbf{A}^*_{n-1}$ and $\mathbf{A}_R$.

It should be mentioned here that, after independently developing the multilayered fluid modeling technique as discussed in this paper, the authors found a similar technique used in the seismological literature for modeling multilayered dipping layers in earth [10]. Authors in Ref. [10] used two layers of sources for modeling each interface, as done in this paper. Wong [11] proposed to model the displacement and stress fields by superimposing the fundamental solutions due to point sources that were placed slightly behind the surface to avoid the singularity in the integral equation; this approach was then followed by Dravinski and Mossessian [10]. However, there are also some major differences between our technique and the technique proposed in Refs. [10,11]. In both those references the half-space Green's function was taken to model the half-space problem. Therefore, those Green's functions must be changed when the half-space is changed to a full space or a quarter-space or any other geometry other than half-space. The Green's function used in our approach is independent of the problem geometry. In Refs. [10,11] the incident beam is assumed to be a plane wave (i.e., a beam of infinite width), which is not the case for our modeling. For the seismic applications the plane-wave assumption is reasonable; however, for the NDE applications often the striking wave is not a plane wave and the modeling should be carried out as outlined in this paper. Finally, in Refs. [10,11] the boundary and interface conditions were satisfied in a least-squares sense that is not the case here, as discussed in the following section.

2.3.2 Source Strength Determination From Boundary and Interface Conditions

Certain boundary and interface continuity conditions must be satisfied for this problem geometry. On the transducer surfaces T and R, the velocity fields are specified as $\mathbf{V}_{S0}$ and $\mathbf{V}_{R0}$, respectively. Across the n interfaces, the pressure $(\underline{\mathbf{PR}})$ and the x_3-direction velocity $(\mathbf{V})$ must be continuous.

Since for any set of sources and target points the velocity and pressure fields are given by $\mathbf{V}=\mathbf{M}\cdot\mathbf{A}$ and $\underline{\mathbf{PR}}=\mathbf{Q}\cdot\mathbf{A}$ (Refer to Eqs. (19) and (20)), the boundary and continuity conditions give rise to the following equations:

$$
\begin{gathered}
\mathbf{M}_{SS}\mathbf{A}_S+\mathbf{M}_{S1}\mathbf{A}_1=\mathbf{V}_{S0}\\
\mathbf{M}_{RR}\mathbf{A}_R+\mathbf{M}^*_{R(n-1)}\mathbf{A}^*_{n-1}=\mathbf{V}_{R0}\\
\mathbf{M}_{1S}\mathbf{A}_S+\mathbf{M}_{11}\mathbf{A}_1=\mathbf{M}_{12}\mathbf{A}_2+\mathbf{M}^*_{11}\mathbf{A}^*_1\\
\mathbf{Q}_{1S}\mathbf{A}_S+\mathbf{Q}_{11}\mathbf{A}_1=\mathbf{Q}_{12}\mathbf{A}_2+\mathbf{Q}^*_{11}\mathbf{A}^*_1\\
\mathbf{M}^*_{21}\mathbf{A}^*_1+\mathbf{M}_{22}\mathbf{A}_2=\mathbf{M}^*_{22}\mathbf{A}^*_2+\mathbf{M}_{23}\mathbf{A}_3\\
\mathbf{Q}^*_{21}\mathbf{A}^*_1+\mathbf{Q}_{22}\mathbf{A}_2=\mathbf{Q}^*_{22}\mathbf{A}^*_2+\mathbf{Q}_{23}\mathbf{A}_3\\
\cdots\\
\mathbf{M}^*_{(n-2)(n-3)}\mathbf{A}^*_{(n-3)}+\mathbf{M}_{(n-2)(n-2)}\mathbf{A}_{(n-2)}=\mathbf{M}^*_{(n-2)(n-2)}\mathbf{A}^*_{(n-2)}\\
+\mathbf{M}_{(n-2)(n-1)}\mathbf{A}_{n-1}\\
\mathbf{Q}^*_{(n-2)(n-3)}\mathbf{A}^*_{(n-3)}+\mathbf{Q}_{(n-2)(n-2)}\mathbf{A}_{(n-2)}=\mathbf{Q}^*_{(n-2)(n-2)}\mathbf{A}^*_{(n-2)}\\
+\mathbf{Q}_{(n-2)(n-1)}\mathbf{A}_{(n-1)}\\
\mathbf{M}^*_{(n-1)(n-2)}\mathbf{A}^*_{(n-2)}+\mathbf{M}_{(n-1)(n-1)}\mathbf{A}_{(n-1)}=\mathbf{M}^*_{(n-1)(n-1)}\mathbf{A}^*_{(n-1)}\\
+\mathbf{M}_{(n-1)T}\mathbf{A}_R\\
\mathbf{Q}^*_{(n-1)(n-2)}\mathbf{A}^*_{(n-2)}+\mathbf{Q}_{(n-1)(n-1)}\mathbf{A}_{(n-1)}=\mathbf{Q}^*_{(n-1)(n-1)}\mathbf{A}^*_{(n-1)}+\mathbf{Q}_{(n-1)T}\mathbf{A}_R
\end{gathered}
\tag{23}
$$

In the matrix form, the above equations can be written as

$$[\mathbf{MAT}]\{\Lambda\}=\{\mathbf{V}_0\} \tag{24}$$

The vectors $\{\Lambda\}$ and $\{\mathbf{V}_0\}$ are

$$\{\Lambda\}^T=\{\mathbf{A}_S\ \mathbf{A}_1\ \mathbf{A}^*_1\ \mathbf{A}_2\ \mathbf{A}^*_2\ \cdots\ \cdots\ \mathbf{A}_{n-2}\ \mathbf{A}^*_{n-2}\ \mathbf{A}_{n-1}\ \mathbf{A}^*_{n-1}\ \mathbf{A}_R\}^T \tag{25}$$

$$\{\mathbf{V}_0\}^T=\{\mathbf{V}_{S0}\ \mathbf{0}\ \mathbf{0}\ \mathbf{0}\ \cdots\ \cdots\ \mathbf{0}\ \mathbf{0}\ \mathbf{0}\ \mathbf{V}_{R0}\}^T \tag{26}$$

Replacing $(n-1)$ by $n1$, $(n-2)$ by $n2$, and so on, the matrix $[\mathbf{MAT}]$ can be written as

$$
[\mathbf{MAT}]=\begin{bmatrix}
M_{SS} & M_{S1} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
M_{1S} & M_{11} & -M^*_{11} & -M_{12} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
Q_{1S} & Q_{11} & -Q^*_{11} & -Q_{12} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & M^*_{21} & M_{22} & -M^*_{22} & -M_{23} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & Q^*_{21} & Q_{22} & -Q^*_{22} & -Q_{23} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & M^*_{32} & M_{33} & -M^*_{33} & -M_{34} & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & Q^*_{32} & Q_{33} & -Q^*_{33} & -Q_{34} & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & M^*_{n2n3} & M_{n2n2} & -M^*_{n2n2} & M_{n2n1} & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & Q^*_{n2n3} & Q_{n2n2} & -Q^*_{n2n2} & Q_{n2n1} & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & M^*_{n1n2} & M_{n1n1} & -M^*_{n1n1} & M_{n1T}\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & Q^*_{n1n2} & Q_{n1n1} & -Q^*_{n1n1} & Q_{n1R}\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & M^*_{Rn1} & M_{RR}
\end{bmatrix}
\tag{27}
$$

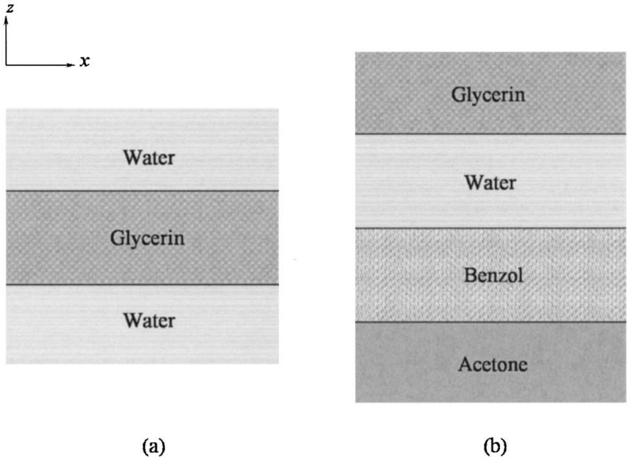


Fig. 4 Multilayered fluid structures considered: (*a*) case 1, water-glycerin-water; (*b*) case 2, acetone-benzol-water-glycerin.

2.3.3 Solution. The source strength vectors of the total system can be calculated by taking the inverse of $[\mathbf{MAT}]$ and multiplying it with the velocity vector $\{\mathbf{V}_0\}$,

$$\{\mathbf{A}\} = [\mathbf{MAT}]^{-1}\{\mathbf{V}_0\} \tag{28}$$

After calculating the source strengths, the pressure and velocity (ultrasonic field) at any point of the problem geometry can be calculated. For example, the pressure field in medium 3 can be written as

$$\underline{\mathbf{PR}}_{(F3)} = \mathbf{Q}_{(F3)2}\mathbf{A}_2^* + \mathbf{Q}_{(F3)3}\mathbf{A}_3 \tag{29}$$

where $F3$ is a set of target points inside the fluid medium 3. Similarly, in all n fluid layers, the pressure field can be calculated as

$$\underline{\mathbf{PR}}_{(F1)} = \mathbf{Q}_{(F1)S}\mathbf{A}_S + \mathbf{Q}_{(F1)1}\mathbf{A}_1$$

$$\underline{\mathbf{PR}}_{(F2)} = \mathbf{Q}_{(F2)1}\mathbf{A}_1^* + \mathbf{Q}_{(F2)2}\mathbf{A}_2$$

$$\cdots$$

$$\underline{\mathbf{PR}}_{(Fn-1)} = \mathbf{Q}_{(Fn-1)n-2}\mathbf{A}_{n-2}^* + \mathbf{Q}_{(Fn-1)n-1}\mathbf{A}_{n-1}$$

$$\underline{\mathbf{PR}}_{(Fn)} = \mathbf{Q}_{(Fn)n-1}\mathbf{A}_{n-1}^* + \mathbf{Q}_{(Fn)R}\mathbf{A}_R \tag{30}$$

where Fi is a set of target points inside the ith fluid medium.

Table 1 Fluid properties

Fluids and Properties	P-wave Speed (km/s)	Density (gm/cc)
Acetone	1.17	0.790
EthylBenzol	1.34	0.868
Water	1.48	1.00
Glycerine	1.92	1.26

3 Numerical Results and Discussions

MATLAB programs have been developed to model the ultrasonic field based on the DPSM formulation presented above. Two separate cases have been studied. In case 1 two different fluids form a three-layered structure as shown in Fig. 4(*a*), and in case 2 four different fluids form four layers of a second problem geometry as shown in Fig. 4(*b*). In the first case the higher density fluid has been placed in between two identical lower density fluid half-spaces. In the second case four layers of fluid have been arranged such that their density monotonically increases from the bottom to the top. For convenience from now onwards the x_1 axis is called the x axis and x_3 axis is called the z axis.

Four different orientations of the two transducers have been considered as illustrated in Fig. 5. In the first orientation, transducers are placed face to face; transducer faces are parallel to the x axis as shown in Fig. 5(*a*). In the second orientation the transducer R has been shifted horizontally and inclined at an angle θ with respect to the z axis, as shown in Fig. 5(*b*). Similarly in the third orientation the transducer T has been shifted and inclined at an angle θ with respect to the z axis as shown in Fig. 5(*c*). Figure 5(*d*) illustrates the fourth orientation where both transducers are inclined.

Densities and P-wave speeds of different fluids considered in this study are given in Table 1. Two different fluid structures and four different transducer orientations are considered to show the interference effect of the second transducer for different orientations and to produce symmetric and nonsymmetric ultrasonic field patterns in the fluid structure.

3.1 Ultrasonic Field Computation for Case 1. Three-layered fluid structure water-glycerin-water (WGW) as shown in Fig. 4(*a*) is studied first. This hypothetical structure is considered to test the computer program, such as whether symmetric ultrasonic field is generated when the two transducers are placed sym-

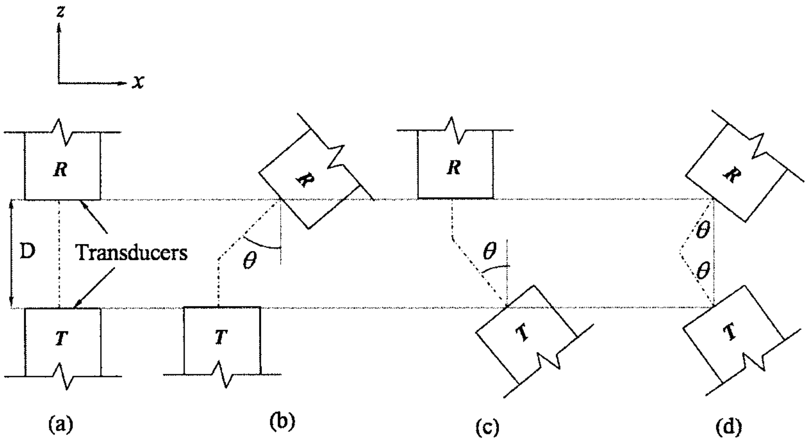


Fig. 5 Transducer orientations (*a*) orientation I; (*b*) orientation II; (*c*) orientation III; (*d*) orientation IV.

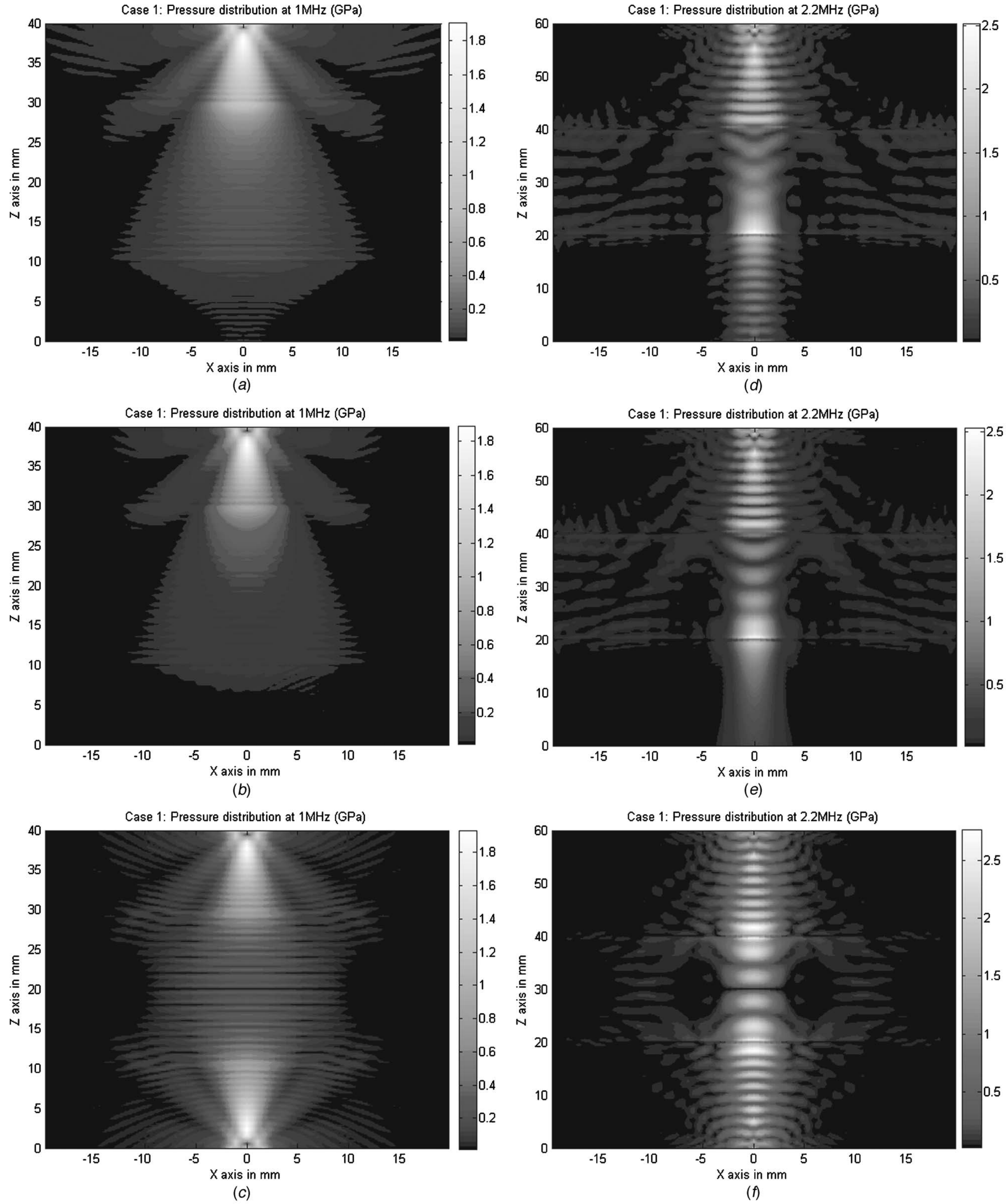


Fig. 6 Ultrasonic fields for case 1. (*a*) Orientation I for 1 MHz transducers with only *R* on. (*b*) Orientation III for 1 MHz transducers with only *R* on. (*c*) Orientation I for 1 MHz transducers with both *T* and *R* on. (*d*) Orientation I for 2.2 MHz transducers with only *R* on. (*e*) Orientation III for 2.2 MHz transducers with only *R* on. (*f*) Orientation I for 2.2 MHz transducers with both *T* and *R* on.

metrically. The middle layer is 20 mm thick and the transducers have 4 mm diameter. Figures 6 and 7 show the ultrasonic fields generated in the WGW structure for different transducer orientations. Ultrasonic fields generated in water-glycerin-water structure due to 1 MHz excitation have been presented in the left columns of Figures 6 and 7 (Figs. 6(*a*)–6(*c*) and 7(*a*)–7(*c*)) and ultrasonic fields due to 2.2 MHz excitation have been presented in right columns (Figs. 6(*d*)–6(*f*) and 7(*d*)–7(*f*)). In Figs. 6(*a*) and 6(*d*) the transducers *R* and *T* are placed face to face, orientation I as shown in Fig. 5(*a*), but only the transducer *R* is turned on, while the

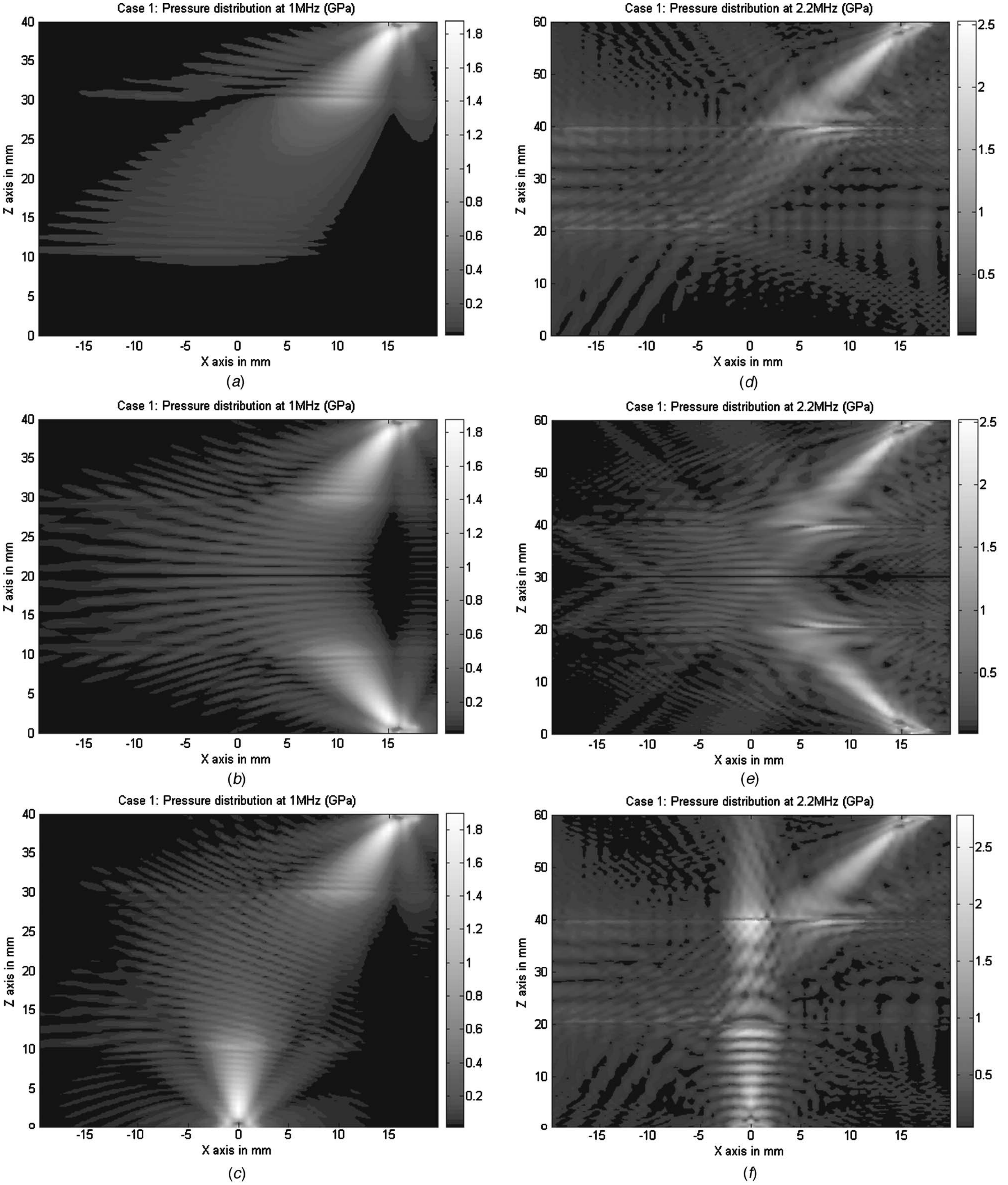


Fig. 7 Ultrasonic fields for case 1. (*a*) Orientation IV for 1 MHz transducers with only *R* on. (*b*) Orientation IV for 1 MHz transducers with both *T* and *R* on. (*c*) Orientation II for 1 MHz transducers with both *T* and *R* on. (*d*) Orientation IV for 2.2 MHz transducers with only *R* on. (*e*) Orientation IV for 2.2 MHz transducers with both *T* and *R* on. (*f*) Orientation II for 2.2 MHz transducers with both *T* and *R* on.

transducer T is kept off. The depth of the glycerin layer has been taken as 20 mm for both 1- and 2.2 MHz signals. One MHz transducers are kept at 10 mm distance from the water-glycerin interfaces but 2.2 MHz transducers are placed at 20 mm distance from the same interfaces. Additional distance for the 2.2 MHz transducers is necessary to make sure that the interface is not placed within the near-field region [7] of the transducers. Therefore, the total width (D; see Fig. 4) of the WGW structure is 40 mm for 1 MHz transducers but D is 60 mm for the 2.2 MHz transducers. It is well known that the pressure field generated at a point in front of a

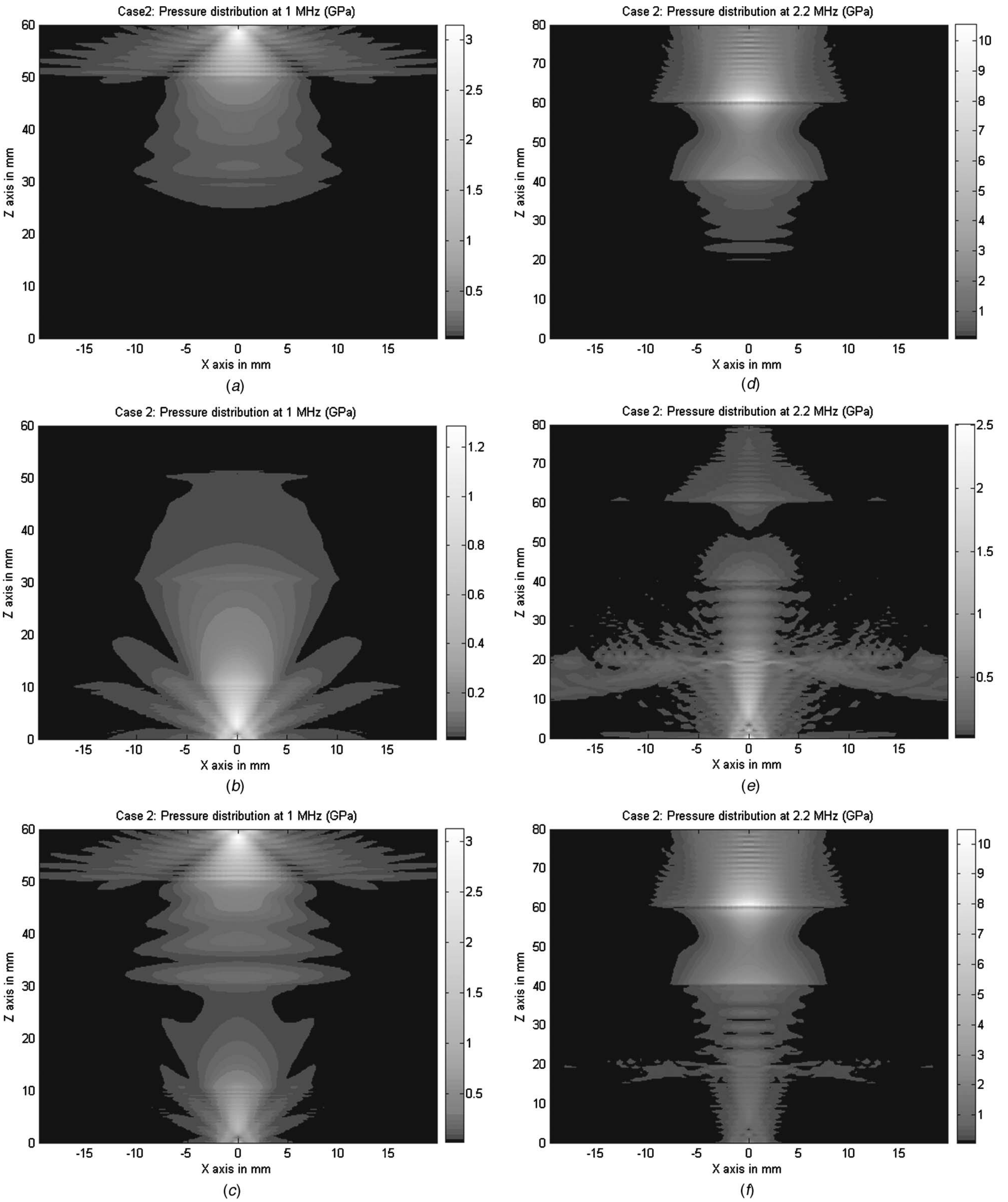
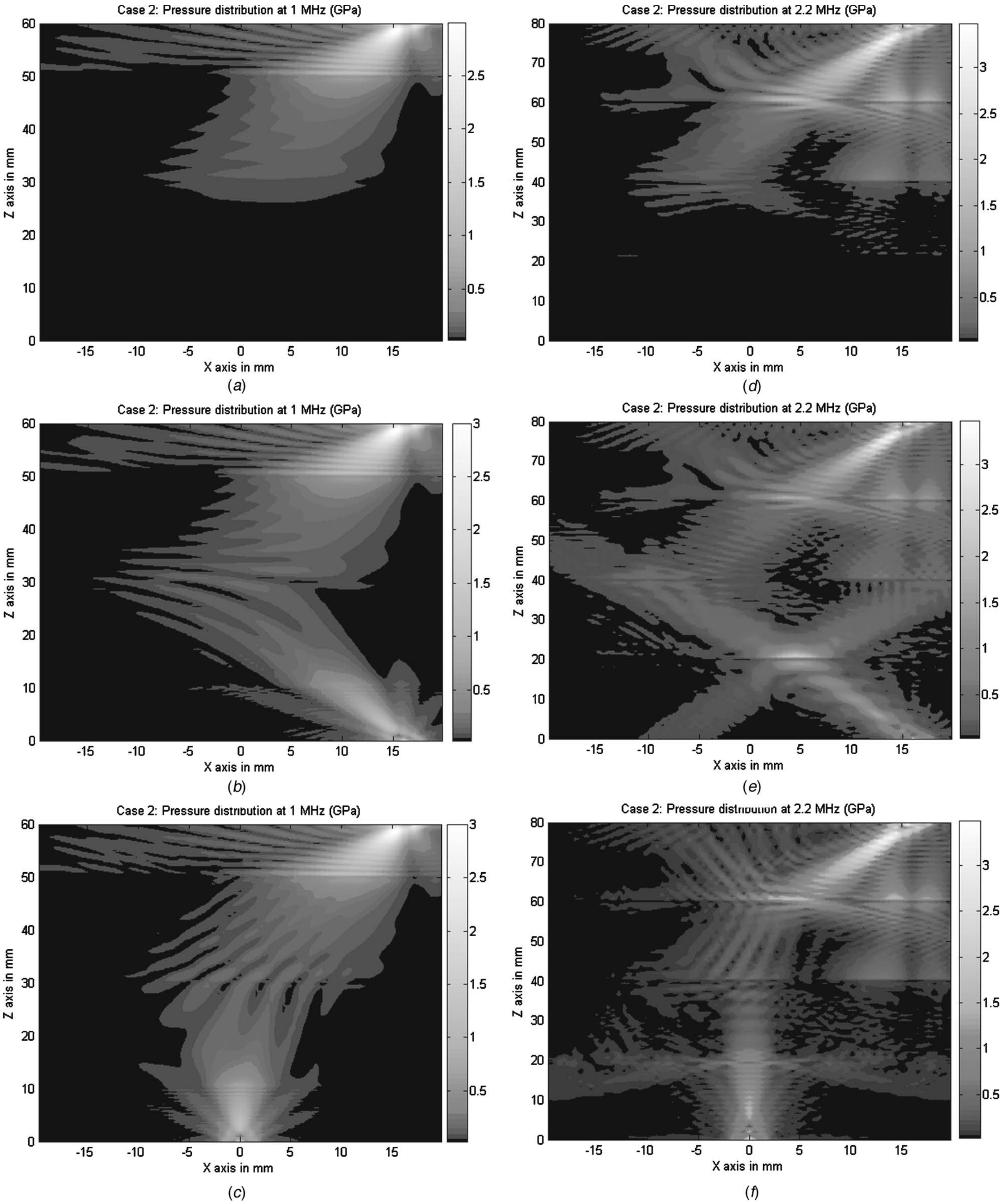


Fig. 8 Ultrasonic fields for case 2. (*a*) Orientation I for 1 MHz transducers with only *R* on. (*b*) Orientation I for 1 MHz transducers with only *T* on. (*c*) Orientation I for 1 MHz transducers with both *T* and *R* on. (*d*) Orientation I for 2.2 MHz transducers with only *R* on. (*e*) Orientation I for 2.2 MHz transducers with only *T* on. (*f*) Orientation I for 2.2 MHz transducers with both *T* and *R* on.

transducer depends on the frequency of excitation and the distance of the point from the transducer face [7]. As frequency increases the isobars gradually shift away from the transducer face. To generate approximately the same pressure value at the interface position, the distance *D* between the transducers has been varied when the transducer frequency is changed from 1 to 2.2 MHz.

From Figs. 6(*a*) and 6(*d*) it is clear that when transducer *T* is off, this transducer works as a reflector or scatterer and the scattering field is generated, near the central region of the bottom layer. At 2.2 MHz the scattering field becomes significantly stronger in comparison to that for 1 MHz transducers. Very strong pressure field is generated in glycerin at 2.2 MHz. It can easily be



(a) (d) (b) (e) (c) (f)

Fig. 9 Ultrasonic fields for case 2. (*a*) Orientation IV for 1 MHz transducers with only *R* on. (*b*) Orientation IV for 1 MHz transducers with both *T* and *R* on. (*c*) Orientation II for 1 MHz transducers with both *T* and *R* on. (*d*) Orientation IV for 2.2 MHz transducers with only *R* on. (*e*) Orientation IV for 2.2 MHz transducers with both *T* and *R* on. (*f*) Orientation II for 2.2 MHz transducers with both *T* and *R* on.

seen that for 2.2 MHz transducers the maximum pressure in glycerin is observed near the bottom of the glycerin layer, whereas for the 1 MHz transducer it is in the top water layer. In glycerin the alternate dips and peaks in the pressure field are due to the interference between the transmitted and the reflected fields and can be seen at both frequencies, but are more prominent near the upper interface at 2.2 MHz.

When the transducers are positioned as in orientation III with

$\theta=30$ deg (see Fig. 5(*c*)) the generated ultrasonic fields are as shown in Figs. 6(*b*) and 6(*c*). Note that the scattered fields almost disappear. However, a careful observation in a contour plot (not presented) shows a very weak scattered field on the right side of the bottom layer. Note that for 2.2 MHz transducers the transmitted field in the bottom layer (Fig. 6(*e*)) is uninterrupted and much more collimated compared to that in Fig. 6(*d*). Both transducers *T* and *R* are now turned on and placed in orientation I as shown in Fig. 5(*a*). The ultrasonic fields generated for this combination at 1 MHz and at 2.2 MHz are presented in Figs. 6(*c*) and 6(*f*), respectively. For 1 MHz excitation, the reflected and transmitted fields make strong alternate dips and peaks in glycerin. Similar phenomenon is also observed for 2.2 MHz transducers; the field is more collimated for 2.2 MHz signals.

In Figs. 7(*a*)–7(*f*) the inclined incidence of the wave field has been considered with inclination angle $\theta=30$ deg, measured from the vertical axis. As mentioned before, the left column is for 1 MHz and the right column is for 2.2 MHz signals. Figures 7(*a*) and 7(*d*) are generated with transducer orientation IV when only *R* is on. Figures 7(*b*) and 7(*e*) are generated for the same orientation but now both transducers are on. In Figs. 7(*b*) and 7(*e*), the ultrasonic fields in glycerin form a nice pattern due to interaction between the transmitted and reflected fields. Note that this pattern is dependent on the signal frequency. For case 1 when the transducers are oriented as in orientation II (Fig. 5(*b*)) with both transducers on, at 1 MHz there is a nice interaction pattern between the transmitted fields inside glycerin. A strong interaction between two transmitted fields is also visible for 2.2 MHz transducers, as shown in Fig. 7(*f*). The reflected field can also be seen in glycerin due to the reflection of the transmitted beam from the bottom interface.

3.2 Ultrasonic Field Computation for Case 2. In case 2 four fluids have been considered and they are placed with monotonically increasing density as shown in Fig. 4(*b*). Analyses have been carried out for transducer orientation I with only *R* on, only *T* on, and both transducers *R* and *T* on, and have been presented in Fig. 8. Ultrasonic fields generated in acetone-benzol-water-glycerin (ABWG) structure by 1 MHz transducers have been presented in the left column (Figs. 8(*a*)–8(*c*)) and the ultrasonic fields generated by 2.2 MHz transducers have been presented in the right column (Figs. 8(*d*)–8(*f*)). For both frequencies the thickness of benzol and water layers are taken as 20 mm each. For the same reason as discussed earlier the transducers are placed 10 mm away from the interfaces for 1 MHz transducers and 20 mm away from the interfaces for 2.2 MHz transducers. Therefore, *D* (see Fig. 5) is equal to 60 mm for 1 MHz and 80 mm for 2.2 MHz transducers.

In Fig. 8 only the normal incidence of the wave field from the transducers has been considered. Figures 8(*a*) and 8(*d*) show the ultrasonic fields when transducer *R* is on with 1- and 2.2 MHz signal frequency, respectively. The scattered field from the transducer *T* is visible near the central region of the bottom layer at 1 MHz but it is not very clear at 2.2 MHz. Relative pressure in glycerin is much higher than that in acetone at 2.2 MHz in comparison to 1 MHz. When *T* is on and *R* is turned off, the scattered field from transducer *R* can be clearly seen in the top layer at both frequencies. Figures 8(*c*) and 8(*f*) have been generated when both *T* and *R* are on. We can see that the ultrasonic field generated in acetone is weaker compared to that in glycerin due to higher impedance of glycerin. At 2.2 MHz the pressure field has several dips and peaks in benzol and it is not that prominent at 1 MHz. The pressure field in water is more collimated at 2.2 MHz compared to that at 1 MHz.

After analyzing the normal incidence cases the transducers are inclined as shown in orientations IV and II with the inclination angle $\theta=30$ deg. The ultrasonic fields generated when only *R* is on, at 1- and 2.2 MHz frequencies, are presented in Figs. 9(*a*) and 9(*d*), respectively. According to Snell's law the transmitted angle (angle between the transmitted field and the normal to the interface) in higher velocity fluid is greater than the incident angle (angle between the incident field and the normal to the interface) in the fluid that has lower velocity. Therefore, in Fig. 9(*b*) the transmitted pressure field in benzol, which has higher wave speed than that in acetone, has greater inclination with the normal to the interface. This phenomenon is also visible at 2.2 MHz in Fig. 9(*e*). Figures 9(*c*) and 9(*f*) show the ultrasonic fields with transducers oriented as shown in Fig. 5(*b*), orientation II.

It is clear from the above results that the algorithm presented is this paper is capable of producing ultrasonic fields in multilayered fluid structures in the presence of multiple transducers of finite dimensions.

4 Conclusion

In this paper the DPSM technique has been extended to model the multilayered fluid structure to study the interaction between the multilayered fluid structures and the bounded acoustic beams. For two different multilayered structures and four different orientations of the transducers, the numerical results have been presented to show the potential of this semi-analytical method in analyzing the wave-propagation problems. In the pressure contour plots presented in this paper the transmission, reflection, and different interactions of the ultrasonic beams in various fluids can be visually observed. Extending this work to incorporate solid layers, cracks, and inclusions is currently underway. Although the finite-element-based codes such as PZFLEX [12] are capable of generating these results, the DPSM technique has a major advantage over PZFLEX: it does not require three-dimensional discretization of the entire space. PZFLEX software is quite expensive and users of this software reported that "a major difficulty encountered was the extremely high storage requirement and runtime needed" [13]. The DPSM technique appears to be a useful alternative to the available finite-element codes.

Acknowledgment

Financial support from the NSF Grants CMS-9901221, OISE-0352680, and ENS-Cachan, France are gratefully acknowledged.

References

[1] Placko, D., and Kundu, T., 2001, "A Theoretical Study of Magnetic and Ultrasonic Sensors: Dependence of Magnetic Potential and Acoustic Pressure on the Sensor Geometry," *Advanced NDE for Structural and Biological Health Monitoring, Proceedings of SPIE*, T. Kundu, ed., SPIE's 6th Annual International Symposium on NDE for Health Monitoring and Diagnostics, March 4–8, 2001, Newport Beach, CA, 4335, pp. 52–62.

[2] Placko, D., Kundu, T., and Ahmad, R., 2002, "Theoretical Computation of Acoustic Pressure Generated by Ultrasonic Sensors in Presence of an Interface," *Smart NDE and Health Monitoring of Structural and Biological Systems, SPIE 7th Annual International Symposium on NDE and Health Monitoring and Diagnostics*, San Diego, CA; 4702, pp. 157–168.

[3] Lee, J. P., Placko, D., Alnuamaini, N., and Kundu, T., 2002, "Distributed Point Source Method (DPSM) for Modeling Ultrasonic Fields in Homogeneous and Non-Homogeneous Fluid Media in Presence of an Interface," *First European Workshop on Structural Health Monitoring, Ecole Normale Superieure de Cachan*, France, D. L. Balageas, ed., Pub. DEStech, PA, pp. 414–421.

[4] Ahmad, R., Kundu, T., and Placko, D., 2003, "Modeling of the Ultrasonic Field of Two Transducers Immersed in a Homogeneous Fluid Using Distributed Point Source Method," T. Kundu, guest ed., I2M (Instrumentation, Mesurement and Metrology) Journal, **3**, pp. 87–116.

[5] Placko, D., Kundu, T., and Ahmad, R., 2003, "Ultrasonic Field Computation in Presence of a Scatterer of Finite Dimension," *Smart NDE and Health Monitoring of Structural and Biological Systems, SPIE 8th Annual International Symposium on NDE and Health Monitoring and Diagnostics*, San Diego, CA, 5047, 169–179.

[6] Ahmad, R., Kundu, T., and Placko, D., 2005, "Modeling of Phased Array

Transducers," J. Acoust. Soc. Am., **117**, pp. 1762–1776.

[7] Placko, D., and Kundu, T., 2004, "Modeling of Ultrasonic Field by Distributed Point Source Method," in *Ultrasonic Nondestructive Evaluation: Engineering and Biological Characterization*, T. Kundu, ed., CRC Press, Boca Raton, FL, Chap. 2, pp. 144–201.

[8] Placko, D., Liebeaux, N., and Kundu, T., 2001, "Presentation d'une method generique pour la modelisation des capteurs de type ultrasons," D. Balageas, ed., Magnetiques at Electrostatiques, Instrumentation, Mesure, Metrologie (I2M Journal): Evaluation nondestructive, **1**, pp. 101–125.

[9] Schmerr, L. W., 1998, *Fundamental of Ultrasonic Nondestructive Evaluation-A Modeling Approach*, Plenum Press, New York.

[10] Dravinski, M., and Mossessian, T. K., 1987, "Scattering of Plane Harmonic P, SV, and Rayleigh Waves by Dipping Layers of Arbitrary Shape," Bull. Seismol. Soc. Am., **77**, pp. 212–235.

[11] Wong, H. L., 1982, "Diffraction of P, SV, and Rayleigh Waves by Surface Topography," Bull. Seismol. Soc. Am., **72**, pp. 1167–1184.

[12] PZFLEX Software, Version: 1-j.6, Weilinger Associates, Inc. 2001.

[13] Feng, F., Mal, A., Kabo, M., Wang, J. C., and Bar-cohen, Y., 2005, "The Mechanical and Thermal Effects of Focused Ultrasound in a Model Biological Material," J. Acoust. Soc. Am., **117**, pp. 2347–2355.

The Effects of Vibrations on Particle Motion Near a Wall in a Semi-Infinite Fluid Cell

Samer Hassan

Masahiro Kawaji
e-mail: kawaji@ecf.utoronto.ca

Department of Chemical Engineering
and Applied Chemistry,
University of Toronto,
Toronto, ON M5S 3E5,
Canada

Tatyana P. Lyubimova

Dmitry V. Lyubimov

Theoretical Physics Department,
Russian Academy of Sciences,
15, Bukirev str.,
Perm 614600,
Russia

The effects of small vibrations on a particle-fluid system relevant to material processing such as crystal growth in space have been investigated experimentally and theoretically. An inviscid model for a spherical particle of radius, R_0, suspended by a thin wire and moving normal to a cell wall in a semi-infinite liquid-filled cell subjected to external horizontal vibrations, was developed to predict the vibration-induced particle motion under normal gravity. The wall effects were studied by varying the distance between the equilibrium position of the particle and the nearest cell wall, H. The method of images was used to derive the equation of motion for the particle oscillating in an inviscid fluid normal to the nearest cell wall. The particle amplitude in a semi-infinite cell increased linearly with the cell vibration amplitude as expected from the results for an infinite cell, however, the particle amplitude also changed with the distance between the equilibrium position of the particle and the nearest wall. The particle amplitude was also found to increase or decrease depending on whether the cell vibration frequency was below or above the resonance frequency, respectively. The theoretical predictions of the particle amplitudes in the semi-infinite cell agreed well with the experimental data, where the effect of the wall proximity on the particle amplitude was found to be significant for ($H/R_0 < 2$) especially near the resonance frequency. Experiments performed at high frequencies well above the resonance frequency showed that the particle amplitude reaches an asymptotic value independent of the wire length. [DOI: 10.1115/1.2165229]

1 Introduction

Microgravity environment plays an important role for material processing in which buoyancy-induced convection process is nearly suppressed even when there is a density variation in a fluid system. Due to diffusion dominated environment, high quality materials such as protein crystals may be obtained under microgravity, provided the fluid system is not subjected to external excitations [1]. Trolinger et al. [2] used the particle image velocimetry (PIV) to detect the greater effect of small vibrations (g-jitter) aboard the Space Shuttle than expected. Recent experiments conducted on protein crystal growth by Gamache et al. [3] on the ground have shown that small vibrations can induce movements in protein crystals which in turn cause significant fluid motion around the growing crystal.

The objective of this paper is to better understand the vibration-induced motion of a small particle in a fluid-particle system, in particular the wall proximity effect on the particle motion, as illustrated in Fig. 1. To this end, the motion of a solid particle suspended in a semi-infinite fluid cell filled with an inviscid fluid and subjected to horizontal sinusoidal vibrations, has been investigated over a wide range of vibration conditions. To evaluate the proximity effect of a cell wall, the mean distance between the particle and the nearest cell wall was varied and the relative motion between the particle and the fluid cell was recorded and analyzed. A theoretical model was also developed and its predictions have been compared with the experimental data.

Previous studies on the fluid mechanics of vibrated particle-fluid systems will be reviewed first, including the motion of a particle subjected to a sinusoidal motion in an otherwise quiescent fluid, particle settling in a vertically oscillating liquid, and theoretical analysis of particle motion in an unsteady and uniform flow but without the wall effects.

Basset [4], Boussinesq [5], and Oseen [6] formulated the equation of motion, referred to as the BBO equation, and investigated theoretically the unsteady motion of a sphere in a stagnant viscous fluid. They eliminated the inertia terms in all their calculations. Oseen [7] linearized the nonlinear terms up to the first order of Reynolds number for steady flows and showed that the ratio of the inertial to the viscous terms cannot be negligible at distances (1/Re) as assumed by Stokes [8] for creeping flow no matter how small the particle diameter is. Maxey and Riley [9] derived the equation of motion for a sphere in a nonuniform creeping flow by modifying the equation of motion proposed previously by Tchen [10]. Unfortunately, in all the above-mentioned papers no analytical expression was derived for the instantaneous particle motion in an unsteady fluid flow.

Lamb [11] and Milne-Thomson [12] also obtained analytically an approximate expression for the force acting on a particle oscillating in a quiescent inviscid fluid near a wall. They used a Taylor series expansion for the ratio of particle diameter to particle-wall distance, however, they neglected terms higher than the first order. Eames et al. [13] studied the motion of an inviscid fluid induced by a sphere moving away from a wall at a constant speed. For vibrated fluid systems, Houghton [14] analyzed the nonlinear drag (Newton's law) on free particles in a sinusoidal velocity field leading to the Mathieu equation, where he found that stable particle trajectories may occur in certain ranges of vibration amplitude and frequency, however, Houghton did not include any finite cell effect in his study.

Li et al. [15] used a successive images method to determine the velocity potential and hence the interacting force on two spheres moving with steady velocities normal to the line joining their centers of mass. The motion of high Reynolds number bubbles in inhomogeneous flow has been predicted by Magnaudet et al. [16]

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received October 15, 2004; final manuscript received November 14, 2005. Review conducted by I. Mezic. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

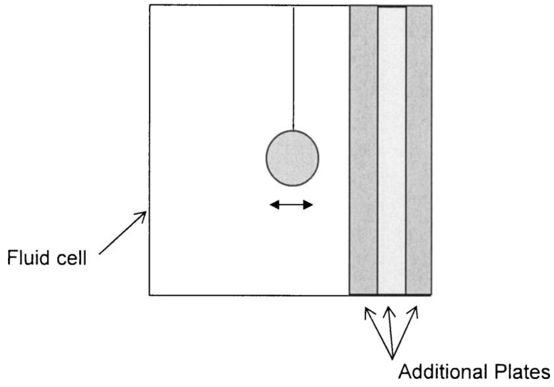


Fig. 1 A semi-infinite cell

but their analysis did not include any wall effect. Magnaudet [17] recently studied the motion of a particle near a wall at a finite Reynolds number, but their study did not incorporate any effect of vibrations near a cell wall.

Recently, Hassan et al. [18] developed a theoretical model to predict the vibration-induced particle motion in an infinite fluid cell assuming an inviscid fluid and irrotational flow. Their inviscid, infinite cell model was able to predict well the experimental data obtained with a solid particle attached to a wire in a large fluid cell which was subjected to small horizontal vibrations.

As clear from the above, to the authors' knowledge, no experimental results and theoretical expressions have been reported for the motion of a particle oscillating near a wall in a fluid cell. This is because particles with densities different from that of the fluid settle due to gravity and steady vibration experiments cannot be easily performed under normal gravity.

2 Experimental Apparatus and Procedure

The experimental apparatus consisted of a test section, a PC-controlled linear translation stage, and a video camera/recording system as shown in Fig. 2. The test cell and the translation stage were placed on a vibration-isolation optical table to suppress external vibrations, which were monitored using an accelerometer. Each of the major components is described in the following in detail.

The test section was a liquid filled rectangular container made of transparent acrylic plates with a height of 110 mm and an internal cross section of 50.8 mm(width) × 50.8 mm (length). A particle was placed at the center of the cell, so that the cell width of 50.8 mm was considered to be sufficiently large for the cell to be regarded as an infinite cell as shown experimentally in Hassan et al. [18]. To study the wall effect, three grooves were cut on a pair of side walls to allow insertion of up to three flat plates and change the cell width on one side as shown in Fig. 1. The distance from the particle at the center of the cell to the nearest wall could be changed to 7.5, 10, 13, 16.5, 18.5, or 25.4 mm, while the distances to the other three walls were always kept at 25.4 mm. This way the wall proximity effects on the particle motion could be determined systematically.

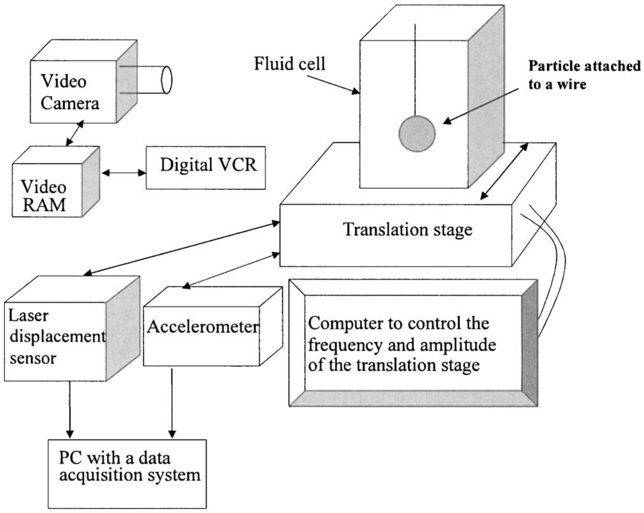

Video Camera
Fluid cell
Particle attached to a wire
Video RAM
Digital VCR
Translation stage
Laser displacement sensor
Accelerometer
Computer to control the frequency and amplitude of the translation stage
PC with a data acquisition system

Fig. 2 Experimental setup.

The cell was filled with distilled water which has relatively low viscosity ($\mu = 10^{-3}$ kg/m s at 20°C), and may be considered as inviscid from a theoretical point of view [18]. Inside the cell, a spherical particle of $D_0 = 12.7$ mm diameter was suspended using a thin platinum wire of 125 μm diameter and 45–85 mm length. The effects of the wire diameter as well as wire length on the particle motion had been studied previously by Hassan et al. [18]. Only spherical particles of different densities (acrylic, $\rho_S = 1.17$ g/cm^3; steel, $\rho_S = 7.83$ g/cm^3; and aluminum, $\rho_S = 2.7$ g/cm^3) have been used in the present experiments leaving other particle shapes for a future study.

A computer-controlled translation stage was used to vibrate the fluid cell with submicron resolution and repeatability. It was controlled to move horizontally with a specified amplitude and frequency in a near sinusoidal manner. A CCD video camera (Hitachi D.S.P VK C-370) with an interchangeable lens was used to capture the particle motion with sufficient magnification. The edge of the particle was captured at 30 frames/s with a shutter speed set at 1/1000 s^{-1} to obtain sharp images. A light source was placed far behind the fluid cell to avoid any heating effect. The particle edges recorded onto a video tape using a digital VCR (JVC AG-7355) were analyzed using an image analysis program. In total, 255 consecutive images of particle edges captured over 8.64 s were analyzed frame by frame, and the data were entered into a spreadsheet to calculate the particle vibration amplitude and frequency.

In the experiments, the cell vibration frequency was varied in small increments from 1 to 5 Hz at a constant cell vibration amplitude of 1.0, 0.5, or 0.25 mm. It should be noted that near the resonance frequency between 1.5 and 1.7 Hz, the cell vibration amplitude was reduced to avoid any possible collision of the steel particle with the cell walls. The same procedure was repeated for different particle densities (aluminum and acrylic) but under different vibration conditions depending on the particle amplitude.

The above procedure was repeated for different distances between the particle and the nearest cell wall to determine any wall effects. Before the data were collected, the system was run for at least 5 min to allow the particle and fluid motions to become stabilized. The video camera was used to record the particle motion for at least 2 min. A pixel-to-micron conversion factor was obtained by lowering a platinum wire of known diameter into the fluid cell and recording its image. Then by using the same image analysis program, the wire diameter in pixels was determined and a pixel-to-micron conversion factor was computed.

The particle amplitude data were reported by Hassan et al. [18] for the largest cell width, which could be regarded as an infinite cell since the data agreed well with the predictions of an infinite cell model. The model predicted a resonance phenomenon where the particle amplitude sharply increases and this prediction was observed experimentally. For a steel particle immersed in a large cell subjected to vibrations and filled with water of density $= 998$ kg/m^3, the particle amplitude was found to increase (decrease) with the wire length when the cell vibration frequency is below (above) the resonance frequency. The particle amplitude was also dependent on the cell vibration frequency and amplitude, and particle and fluid densities.

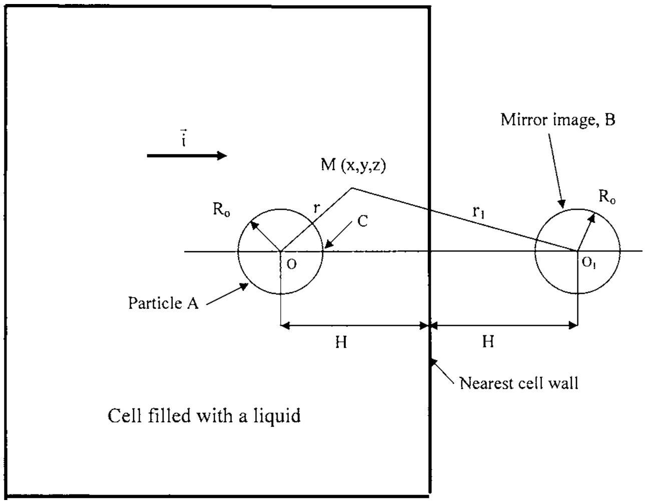


Fig. 3 Semi-infinite cell model

3 Theoretical Analysis of the Particle Motion

To study the wall proximity effect on the particle motion in an inviscid fluid, it is necessary to analytically derive an equation of particle motion that incorporates the particle distance to the nearest cell wall and the particle diameter. Here, the method of images is used for a particle moving in the direction normal to the nearest wall.

Method of Images. This method assumes a spherical particle A of radius R_0 near a wall inside the cell and its virtual image, B, with the same diameter existing on the outside of the cell as shown in Fig. 3. Particle A has an equilibrium position near the cell wall at a distance H from the wall. The other walls are considered to be sufficiently far away from the particle; thus, the particle motion analyzed would be that in a semi-infinite cell. The particle moves in the direction **i** normal to the nearest wall. The boundary conditions for the given problem to be satisfied are those for the so-called "mirror image" flow fields.

In the following theoretical analysis, the assumption of an inviscid model is justified because the Reynolds number is significantly high for different cell vibration frequencies. For a steel particle attached to a wire of length 76 mm and immersed in water, the values of Reynolds number for a cell vibration amplitude of 1 mm at vibration frequencies of 1.5 and 10 Hz are 530 and 668, respectively.

The velocity potential for a semi-infinite cell vibrated horizontally at a frequency, f (Hz) or ω ($=2\pi f$) in rad/s, and cell amplitude, a, can be decomposed into four components and is then given by applying the superposition principle

$$\varphi = \varphi_1 + \varphi_2 + \varphi_3 + \varphi_4 \tag{1}$$

where φ_1 is the velocity potential for a sphere in an infinite cell, φ_2 represents the pendulum motion in the liquid included as a potential flow, and φ_3 and φ_4 stand for the wall effects using the method of mirror image.

Each term of Eq. (1) satisfies Laplace's equation, as shown in Hassan et al. [18] and Lamb [11], and is given by

$$\varphi_1 = (a\omega\,.\,\mathbf{i}\cdot\mathbf{r})\cos\omega t + C_1\frac{\mathbf{i}\cdot\mathbf{r}}{r^3} \tag{2}$$

$$\varphi_2 = -\dot{\mathbf{R}}\cdot\mathbf{r} + C_2\frac{\dot{\mathbf{R}}\cdot\mathbf{r}}{r^3} \tag{3}$$

$$\varphi_3 = C_3\frac{\mathbf{i}\cdot\mathbf{r}_1}{r_1^3} \tag{4}$$

$$\varphi_4 = C_4\frac{\dot{\mathbf{R}}\cdot\mathbf{r}_1}{r_1^3} \tag{5}$$

where r and r_1 are the distances from the center of the particle and its image, respectively, to a given point M in the cell. **R** is the absolute position of the particle with respect to the inertial frame of reference, and the dot on the vector **R** indicates the time derivative. The constants C_1,C_2, C_3, and C_4 are to be determined later by applying the boundary conditions, and **i** is a unit vector in the direction of the cell motion.

In vector notation, the vectors,**i**, **r**, and $\mathbf{r}_1$ can be represented as $\mathbf{i}(1,0,0)$, $\mathbf{r}(x,y,z)$, and $\mathbf{r}_1(x\text{-}2H,y,z)$. The boundary condition at the wall is given by

$$(\mathbf{v}\cdot\mathbf{n})_W = \left.(\dot{\mathbf{R}} + a\omega\mathbf{i}\cos\omega t)\cdot\mathbf{n}\right|_W \tag{6}$$

The velocity field is then given by the gradient of the velocity potentials as

$$\mathbf{v} = \nabla\varphi = \nabla\varphi_1 + \nabla\varphi_2 + \nabla\varphi_3 + \nabla\varphi_4 \tag{7}$$

$$\text{where }\ \nabla\varphi_i = \frac{\partial\varphi_i}{\partial x}\mathbf{i} + \frac{\partial\varphi_i}{\partial y}\mathbf{j} + \frac{\partial\varphi_i}{\partial z}\mathbf{k} \tag{8}$$

By substituting Eq. (7) into the boundary condition (Eq. (6)) at the cell wall and using Eq. (1), the constants C_1 and C_2 are found to be related to C_3 and C_4 as follows

$$C_1 = -C_3, \quad C_2 = -C_4 \tag{9}$$

Equations (2)–(5) can be written in dimensionless form by introducing the dimensionless parameters for the fluid density, wire length, particle-to-wall distance, cell vibration frequency, particle and cell amplitudes, and fluid pressure as follows

$$\tilde{\rho} = \frac{\rho_L}{\rho_S} \quad \tilde{L} = \frac{L}{R_0} \quad \tilde{H} = \frac{H}{R_0} \quad \tilde{\omega} = \frac{\omega}{\sqrt{g/L}} \quad \tilde{A}_p = \frac{A_p}{R_0} \quad \tilde{a} = \frac{a}{R_0}$$

$$\tilde{X}_p = \frac{X_p}{R_0}$$

$$\tilde{t} = \omega t \quad \tilde{\varphi}_i = \frac{\varphi_i}{\omega R_0^2} \quad \tilde{r} = \frac{r}{R_0} \quad \tilde{r}_1 = \frac{r_1}{R_0} \quad \tilde{p} = \frac{p}{\rho\omega^2 R_0^2} \quad \tilde{g} = \frac{g}{\omega^2 R_0} \tag{10}$$

Using these dimensionless parameters, Eqs. (2)–(5) can be written as follows

$$\tilde{\varphi}_1 = (\tilde{a}\mathbf{i}\cdot\tilde{\mathbf{r}})\cos\tilde{t} + \tilde{C}_1\frac{\mathbf{i}\cdot\tilde{\mathbf{r}}}{\tilde{r}^3} \tag{11}$$

$$\tilde{\varphi}_2 = -\dot{\mathbf{R}}\cdot\tilde{\mathbf{r}} + \tilde{C}_2\frac{\tilde{\mathbf{R}}\tilde{r}}{\tilde{r}^3} \tag{12}$$

$$\tilde{\varphi}_3 = \tilde{C}_3\frac{\mathbf{i}\cdot\tilde{\mathbf{r}}_1}{\tilde{r}_1^3} \tag{13}$$

$$\tilde{\varphi}_4 = \tilde{C}_4\frac{\tilde{\mathbf{R}}\cdot\tilde{\mathbf{r}}_1}{\tilde{r}_1^3} \tag{14}$$

The potential field ϕ_1 can be expanded as

$$\tilde{\varphi}_1 = \tilde{a}\tilde{x}\cos\tilde{t} + \tilde{C}_1\frac{\tilde{x}}{\tilde{r}^3} \tag{15}$$

Hence, the velocity field $\tilde{\mathbf{v}}_1$ corresponding to $\tilde{\varphi}_1$ is given by

$$\tilde{\mathbf{v}}_1 = \tilde{a}\cos\tilde{t}\,\mathbf{i} + \frac{\tilde{C}_1}{\tilde{r}^3}\mathbf{i} - \frac{3\tilde{C}_1\tilde{x}^2}{\tilde{r}^5}\mathbf{i} - \frac{3\tilde{C}_1\tilde{x}\tilde{y}}{\tilde{r}^5}\mathbf{j} - \frac{3\tilde{C}_1\tilde{x}\tilde{z}}{\tilde{r}^5}\mathbf{k} \tag{16}$$

It can also be written as

$$\tilde{\mathbf{v}}_1 = \tilde{a}\cos\tilde{t}\,\mathbf{i} + \frac{\tilde{C}_1}{\tilde{r}^3}\mathbf{i} - \frac{3\tilde{C}_1(\mathbf{i}\cdot\tilde{\mathbf{r}})\tilde{\mathbf{r}}}{\tilde{r}^5} \tag{17}$$

by representing the position vector and its magnitude as

$$\tilde{\mathbf{r}}_1 = \tilde{x}\mathbf{i} + \tilde{y}\mathbf{j} + \tilde{z}\mathbf{k}, \quad |\tilde{\mathbf{r}}| = \sqrt{\tilde{x}^2 + \tilde{y}^2 + \tilde{z}^2} \tag{18}$$

Similarly from Eq. (12), the velocity potential, $\tilde{\varphi}_2$, can be written as

$$\tilde{\varphi}_2 = -\dot{\tilde{R}}_x\tilde{x} - \dot{\tilde{R}}_y\tilde{y} - \dot{\tilde{R}}_z\tilde{z} + \tilde{C}_2\left[\frac{\dot{\tilde{R}}_x\tilde{x} + \dot{\tilde{R}}_y\tilde{y} + \dot{\tilde{R}}_z\tilde{z}}{\tilde{r}^3}\right] \tag{19}$$

where the vectors $\tilde{\mathbf{R}}=(\tilde{R}_x,\tilde{R}_y,\tilde{R}_z)$ and $\dot{\tilde{\mathbf{R}}}=(\dot{\tilde{R}}_x,\dot{\tilde{R}}_y,\dot{\tilde{R}}_z)$ depend only on time $\tilde{t}$. The velocity field $\tilde{\mathbf{v}}_2$ corresponding to $\tilde{\varphi}_2$ is given by

$$\tilde{\mathbf{v}}_2 = -\dot{\tilde{\mathbf{R}}} + \tilde{C}_2\frac{\dot{\tilde{\mathbf{R}}}}{\tilde{r}^3} - \frac{3C_2(\dot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}})\tilde{\mathbf{r}}}{\tilde{r}^5} \tag{20}$$

The velocity potential $\tilde{\varphi}_3$ in Eq. (13) can be also written as

$$\tilde{\varphi}_3 = -\tilde{C}_1\frac{(\tilde{x} - 2\tilde{H})}{[(\tilde{x} - 2\tilde{H})^2 + \tilde{y}^2 + \tilde{z}^2]^{3/2}} \tag{21}$$

where the vector $\tilde{\mathbf{r}}$ and its magnitude are represented by

$$\tilde{\mathbf{r}}_1 = (\tilde{x} - 2\tilde{H})\mathbf{i} + \tilde{y}\mathbf{j} + \tilde{z}\mathbf{k} \quad |\tilde{\mathbf{r}}_1| = \sqrt{(\tilde{x} - 2\tilde{H})^2 + \tilde{y}^2 + \tilde{z}^2} \tag{22}$$

and the velocity field $\tilde{\mathbf{v}}_3$ corresponding to $\tilde{\varphi}_3$ is given by

$$\tilde{\mathbf{v}}_3 = -\frac{\tilde{C}_1}{\tilde{r}^3}\mathbf{i} + \frac{3\tilde{C}_1(\mathbf{i}\cdot\tilde{\mathbf{r}}_1)\tilde{\mathbf{r}}_1}{\tilde{r}_1^5} \tag{23}$$

The last dimensionless term $\tilde{\varphi}_4$ in Eq. (14) can be written as

$$\tilde{\varphi}_4 = -\tilde{C}_2\left[\frac{\dot{\tilde{R}}_x(x - 2H) + \dot{\tilde{R}}_y y + \dot{\tilde{R}}_z z}{\tilde{r}_1^3}\right] \tag{24}$$

and the velocity field $\tilde{\mathbf{v}}_4$ corresponding to $\tilde{\varphi}_4$ is given by

$$\tilde{\mathbf{v}}_4 = -\tilde{C}_2\frac{\dot{\tilde{\mathbf{R}}}}{\tilde{r}_1^3} + \frac{3\tilde{C}_2(\dot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}}_1)\tilde{r}_1}{\tilde{r}_1^5} \tag{25}$$

Hence, the velocity field $\tilde{\mathbf{v}}$ written in a dimensionless form for a semi-infinite cell is the sum of $\tilde{\mathbf{v}}_1$, $\tilde{\mathbf{v}}_2$, $\tilde{\mathbf{v}}_3$, and $\tilde{\mathbf{v}}_4$.

The boundary condition on the surface of a solid particle of radius R_0 is that the radial velocity has a zero component with respect to the cell. This results from the impermeability condition in an inviscid fluid,

$$\tilde{v}_r = \nabla\tilde{\varphi}\cdot\mathbf{n} = 0 \tag{26}$$

where $\tilde{v}_r$ and $\mathbf{n}$ are the radial velocity component of $\tilde{\mathbf{v}}$ and the unit outward normal on the surface of the solid sphere A, respectively.

The condition given by Eq. (26) is satisfied at any point on the particle surface, in particular at the point C on the same line joining the centers of the two spheres A and its mirror image, B, as shown in Fig. 3, H is taken as the particle-wall distance from the equilibrium position. Hence,

$$\mathbf{i}\cdot\mathbf{n} = 1, \quad \mathbf{i}\cdot\mathbf{n}_1 = -1, \quad \mathbf{n}\cdot\mathbf{n}_1 = -1 \tag{27}$$

where $\mathbf{n}_1$ is the unit outward normal to the image of the sphere. Setting the coefficient of $\mathbf{i}\,.\,\mathbf{n}$ in Eq. (26) to zero yields

$$\tilde{a}\cos\tilde{t} + \tilde{C}_1 - 3\tilde{C}_1 - \frac{\tilde{C}_1}{(2\tilde{H} - 1)^3} + \frac{3\tilde{C}_1}{(2\tilde{H} - 1)^3} = 0 \tag{28}$$

Solving for $\tilde{C}_1$ yields

$$\tilde{C}_1 = \frac{(\tilde{a}\cos\tilde{t})(2\tilde{H} - 1)^3}{2(8\tilde{H}^3 - 12\tilde{H}^2 + 6\tilde{H} - 2)} \tag{29}$$

Also, setting the coefficient of $\dot{\tilde{\mathbf{R}}}\cdot\mathbf{n}$ in Eq. (26) to zero results in

$$-1 - 3\tilde{C}_2 + C_2 + \frac{\tilde{C}_2}{(2\tilde{H} - 1)^3} - \frac{3\tilde{C}_2}{(2\tilde{H} - 1)^3} = 0 \tag{30}$$

Then, solving for $\tilde{C}_2$ yields

$$\tilde{C}_2 = \frac{-(2\tilde{H} - 1)^3}{2(8\tilde{H}^3 - 12\tilde{H}^2 + 6\tilde{H} - 2)} \tag{31}$$

If we let

$$\tilde{G} = (\tilde{W} - 1)^3 \tag{32}$$

and

$$\tilde{K} = 2(\tilde{W}^3 - 3\tilde{W}^2 + 3\tilde{W} - 2) \tag{33}$$

where $\tilde{W}$=2 $\tilde{H}$ is the dimensionless distance from the particle to its image through the nearest cell wall, expressions (29) and (31) can be written as

$$\tilde{C}_1 = \frac{(a\cos\tilde{t})\tilde{G}}{\tilde{K}} \tag{34}$$

$$\tilde{C}_2 = \frac{-\tilde{G}}{\tilde{K}} \tag{35}$$

and Eq. (1) in a dimensionless form can be written as

$$\tilde{\varphi} = \tilde{a}\cos\tilde{t}\,\mathbf{i}\cdot\tilde{\mathbf{r}} + (\tilde{a}\cos\tilde{t})\left(\frac{\tilde{G}}{\tilde{K}}\right)\left(\frac{\mathbf{i}\cdot\tilde{\mathbf{r}}}{\tilde{r}^3} - \frac{\mathbf{i}\cdot\tilde{\mathbf{r}}_1}{\tilde{r}_1^3}\right) - \dot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}} - \left(\frac{\tilde{G}}{\tilde{K}}\right) \times \left(\frac{\dot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}}}{\tilde{r}^3} - \frac{\dot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}}_1}{\tilde{r}_1^3}\right) \tag{36}$$

The pressure p at each point in the fluid can be calculated from the Bernoulli equation

$$\frac{\partial\varphi}{\partial t} + \frac{1}{2}(\nabla\varphi)^2 = -\frac{p}{\rho_L} - \ddot{\mathbf{R}}\cdot\mathbf{r} + \mathbf{g}\cdot\mathbf{r} \tag{37}$$

or in a dimensionless form as

$$\frac{\partial\tilde{\varphi}}{\partial\tilde{t}} + \frac{1}{2}(\nabla\tilde{\varphi})^2 = -\tilde{p} - \ddot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}} + \tilde{\mathbf{g}}\cdot\tilde{\mathbf{r}} \tag{38}$$

where the dimensionless particle acceleration is

$$\ddot{\tilde{\mathbf{R}}} = \frac{\ddot{\mathbf{R}}}{\omega^2 R_0} \tag{39}$$

Applying Newton' second law to the spherical particle of mass, m, gives

$$m\ddot{\mathbf{R}} = \mathbf{F}_p + m\mathbf{g} + \mathbf{f} \tag{40}$$

where $\mathbf{F}_p$ is the effective force due to the fluid pressure on the surface of the sphere, and $\mathbf{f}$ is the wire tension. Equation (40) can be written in a dimensionless form as

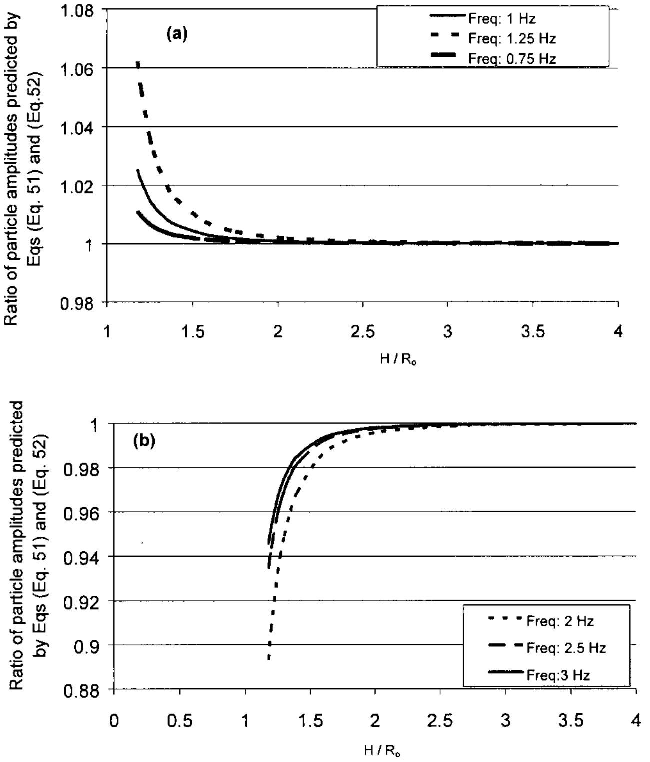


Fig. 4 Variation of the ratio of the particle amplitudes predicted by Eqs. (51) and (52) with the particle-wall distance to particle radius ratio, H/R_0, for (*a*) below-resonance frequencies and (*b*) above-resonance frequencies (steel particle in water, wire length=76 mm)

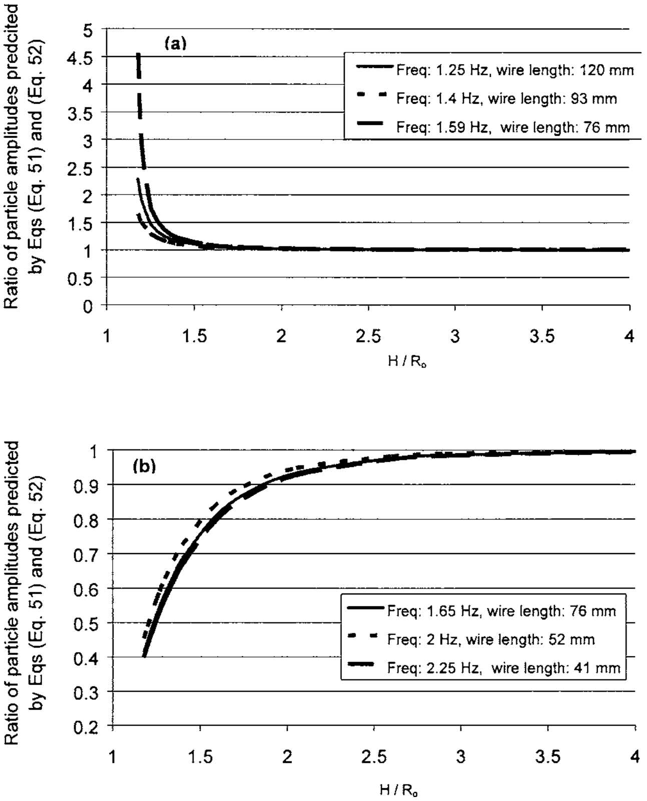


Fig. 5 Variation of the ratio of the particle amplitudes predicted by Eqs. (51) and (52) with the particle-wall distance to particle radius ratio, H/R_0, for (*a*) below-resonance frequencies and (*b*) above-resonance frequencies (steel particle in water, wire length=76 mm)

$$\frac{V\ddot{\tilde{\mathbf{R}}}}{\tilde{\rho}R_0^3} = \tilde{\mathbf{F}}_p + \frac{V\tilde{\mathbf{g}}}{\tilde{\rho}R_0^3} + \frac{\mathbf{f}}{\omega^2 R_0^4} \tag{41}$$

$$\text{where } \tilde{\mathbf{F}}_p = \frac{\mathbf{F}_p}{\rho_L\omega^2 R_0^4} = -\oint \tilde{p}\,\mathbf{d}\tilde{\mathbf{s}} \tag{42}$$

V is the volume of the spherical particle, and the dots denote the higher order derivatives.

By substituting Eq. (36) into Eq. (38) and using the fact that the velocity $\nabla\tilde{\varphi}$ is zero on the particle surface, the dimensionless pressure $\tilde{p}$ on the particle surface is given by

$$\tilde{p} = \left[\tilde{a}\sin\tilde{t}(\mathbf{i}\cdot\tilde{\mathbf{r}}) + \tilde{a}\sin\tilde{t}\left(\frac{\tilde{G}}{\tilde{K}}\right)\left(\frac{\mathbf{i}\cdot\tilde{\mathbf{r}}}{\tilde{r}^3} - \frac{\mathbf{i}\cdot\tilde{\mathbf{r}}_1}{\tilde{r}_1^3}\right) + \left(\frac{\tilde{G}}{\tilde{K}}\right)\left(\frac{\ddot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}}}{\tilde{r}^3} - \frac{\ddot{\tilde{\mathbf{R}}}\cdot\tilde{\mathbf{r}}_1}{\tilde{r}_1^3}\right) + \tilde{\mathbf{g}}\cdot\tilde{\mathbf{r}}\right] \tag{43}$$

Integrating Eq. (43) over the particle surface gives the dimensionless effective force as

$$\tilde{\mathbf{F}}_p = -\left[\frac{V}{R_0^3}\tilde{a}\sin\tilde{t}\mathbf{i} + \tilde{a}\sin\tilde{t}V\left(\frac{\tilde{G}}{\tilde{K}}\right)\left(\frac{1}{R_0^3} + \frac{2}{W^3}\right)\mathbf{i} + \left(\frac{\tilde{G}}{\tilde{K}}\right)V\left(\frac{1}{R_0^3} + \frac{2}{W^3}\right)\ddot{\tilde{\mathbf{R}}} + \tilde{\mathbf{g}}\frac{V}{R_0^3}\right] \tag{44}$$

Substituting Eq. (44) into Eq. (41) yields the acceleration in a dimensionless form as

$$\ddot{\tilde{\mathbf{R}}} = -\tilde{\rho}\left[\tilde{a}\sin\tilde{t}\mathbf{i} + \tilde{a}\sin\tilde{t}\left(\frac{\tilde{G}}{\tilde{K}}\right)\left(1 + \frac{2}{\tilde{W}^3}\right)\mathbf{i} + \left(\frac{\tilde{G}}{\tilde{K}}\right)\left(1 + \frac{2}{\tilde{W}^3}\right)\ddot{\tilde{\mathbf{R}}} + \tilde{\mathbf{g}}\right] + \tilde{\mathbf{g}} + \tilde{\mathbf{f}} \tag{45}$$

$$\text{where } \ddot{\tilde{\mathbf{R}}} = \ddot{\tilde{\mathbf{R}}}_p - \tilde{a}\sin\tilde{t}\mathbf{i} \tag{46}$$

and

$$\tilde{\mathbf{f}} = \frac{\tilde{\rho}\mathbf{f}}{\rho_L V\omega^2 R_0} \tag{47}$$

$\ddot{\tilde{\mathbf{R}}}_p$ is the dimensionless form of the particle position $\mathbf{R}_p$ in the cell frame of reference. Substituting Eq. (46) into Eq. (45) yields

$$\left[1 + \tilde{\rho}\left\{\frac{\tilde{G}}{\tilde{K}}\left(1 + \frac{2}{\tilde{W}^3}\right)\right\}\right]\ddot{\tilde{\mathbf{R}}}_p = (1-\tilde{\rho})\tilde{a}\sin\tilde{t}\mathbf{i} + (1-\tilde{\rho})\tilde{\mathbf{g}} + \tilde{\mathbf{f}} \tag{48}$$

Assuming small cell vibration amplitudes, the vertical component of Eq. (48) yields the wire tension as follows

$$\tilde{\mathbf{f}} = \tilde{\mathbf{g}}(1-\tilde{\rho}) \tag{49}$$

Substitution of Eq. (49) into Eq. (48) leads to the following differential equation for the horizontal dimensionless displacement, $\tilde{X}_p$, of the particle with respect to the equilibrium position

$$\left[1 + \tilde{\rho}\left\{\frac{\tilde{G}}{\tilde{K}}\left(1 + \frac{2}{\tilde{W}^3}\right)\right\}\right]\ddot{\tilde{X}}_p = (1-\tilde{\rho})\left[\tilde{a}\sin\tilde{t} - \tilde{\mathbf{g}}\frac{\tilde{X}_p}{\tilde{L}}\right] \tag{50}$$

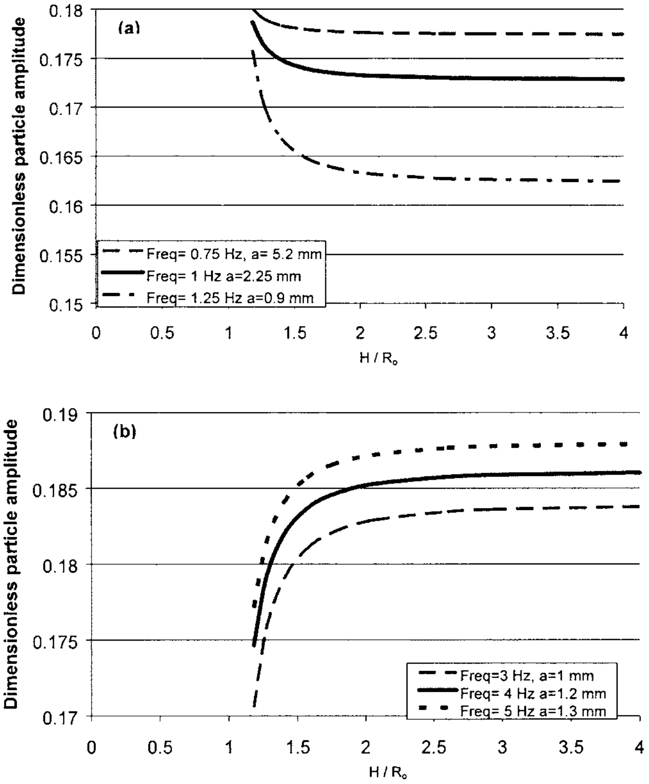


Fig. 6 Variation of dimensionless particle amplitudes with particle-wall distance to particle radius ratio predicted by Eq. (51) for (*a*) below-resonance frequencies and (*b*) above-resonance frequencies (steel particle in water, wire length =76 mm)

Solving this second-order differential equation yields the particle amplitude for the particle moving in the direction normal to the nearest cell wall,

$$\tilde{A}_p = \frac{(1-\tilde{\rho})\tilde{a}}{(1-\tilde{\rho})\dfrac{\tilde{g}}{\tilde{L}} - \left[\left(1+\tilde{\rho}\left\{\dfrac{\tilde{G}}{\tilde{K}}\left(1+\dfrac{2}{\tilde{W}^3}\right)\right\}\right)\right]}, \tag{51}$$

By expanding the second term of the denominator in Eq. (51) in a Taylor series in which $\tilde{W} \gg 1$, we can also obtain an approximate equation for the particle amplitude as follows

$$\tilde{A}_p = \frac{(1-\tilde{\rho})\tilde{a}}{(1-\tilde{\rho})\dfrac{\tilde{g}}{\tilde{L}} - \left[\left(1+\dfrac{\tilde{\rho}}{2}\left\{\left(1+\dfrac{3}{8\tilde{H}^3}\right)\right\}\right)\right]} \tag{52}$$

We note that in the limit $H \to \infty$, both Eqs. (51) and (52) reduce to the expression for the amplitude of a particle in an infinite fluid cell given by Hassan et al. [18]. For both semi-infinite and infinite cells, the inviscid model predicts the particle amplitude,A_p, to be linearly proportional to the cell vibration amplitude, a.

To assess the effect of the Taylor series expansion used in deriving Eq. (52) from Eq. (51), their ratio denoted by λ will be examined

$$\lambda = \frac{\tilde{A}_p}{\tilde{A}_{pa}} = \frac{\dfrac{g}{L\omega^2} - \dfrac{(\rho_S+\rho_L)}{(\rho_S-\rho_L)}\left(\dfrac{G}{K}\left(\dfrac{1}{R_0^3}+\dfrac{2}{W^3}\right)\right)}{\dfrac{g}{L\omega^2} - \dfrac{(2\rho_S+\rho_L)}{2(\rho_S-\rho_L)}\left(1+\dfrac{3}{8}\left(\dfrac{R_0}{H}\right)^3\right)} \tag{53}$$

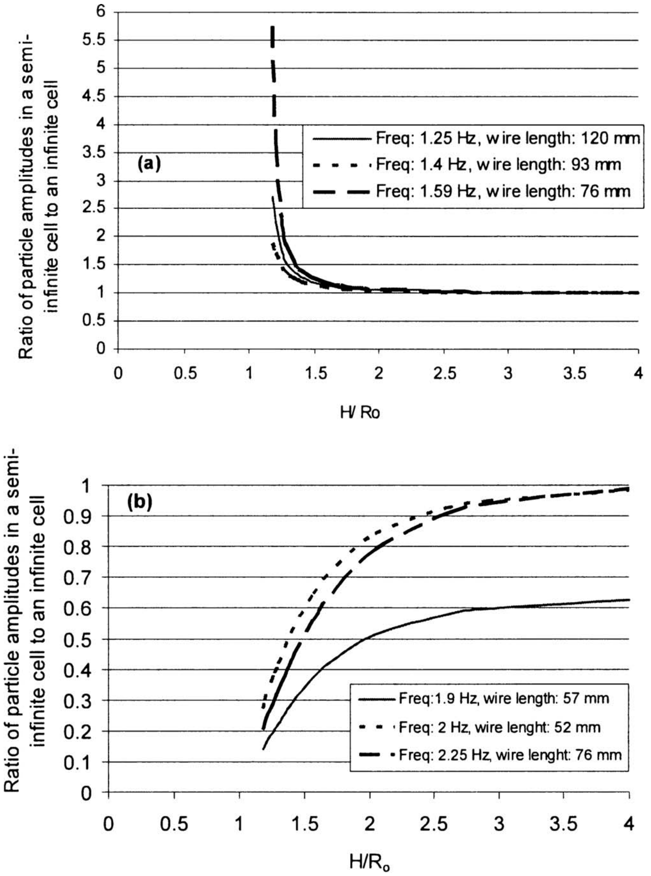


Fig. 7 Variation of the ratio of the particle amplitude for a semi-infinite cell (Eq. (51)) to that for an in infinite cell with the particle-wall distance to particle radius ratio, H/R_0, for (*a*) below-resonance frequencies and (*b*) above-resonance frequencies (cell amplitude=16 and 30 μm, respectively)

The resonance frequency can be obtained from Eq. (51) or Eq. (52) by equating the denominator to zero. This yields the following expressions for the dimensionless resonance frequency

$$\tilde{\omega}_{\mathrm{res}} = \sqrt{\frac{(1-\tilde{\rho})}{1+\tilde{\rho}\left\{\dfrac{\tilde{G}}{\tilde{K}}\left(1+\dfrac{2}{\tilde{W}^3}\right)\right\}}} \tag{54}$$

$$\tilde{\omega}_{\mathrm{res}} = \sqrt{\frac{(1-\tilde{\rho})}{1+\dfrac{\tilde{\rho}}{2}\left\{\left(1+\dfrac{3}{8\tilde{H}^3}\right)\right\}}} \tag{55}$$

Equation (54) is valid for any value of $\tilde{H}$. In the limit, $H \to \infty$, Eqs. (54) and (55) both reduce to the expression obtained by Hassan et al. [18] for an infinite cell. As the particle-wall distance is reduced, the resonance frequency also decreases but the degree of reduction depends on the particle and fluid densities.

4 Results and Discussion

The inviscid model predictions for a semi-infinite cell given by Eqs. (51) and (52) are now examined together with the experimental data to evaluate the wall proximity effect on the steel particle amplitude, when the particle is moving normal to the nearest cell wall. As seen from the above theoretical derivation, Eq. (51) is valid for any particle-to-nearest cell wall distance, while Eq. (52) is an approximate solution for the particle ampli-

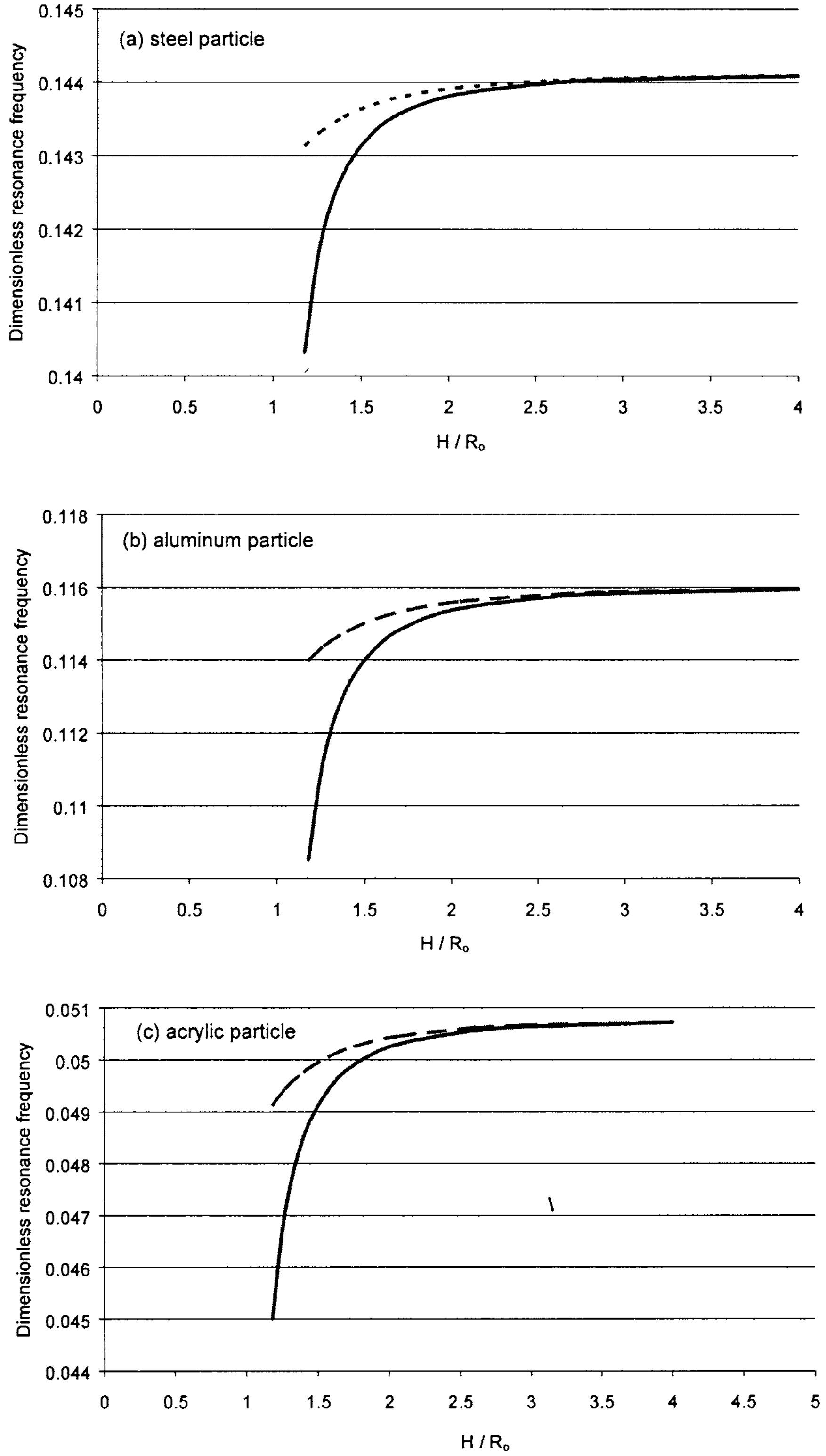


Fig. 8 Variations of the dimensionless resonance frequencies for different particles in a semi—infinite cell predicted by Eqs. (54) and (55) with H/R_0 (wire length =76 mm, (—) Eq. (54), (---) Eq. (55))

tude when $\tilde{W} \gg 1$. The two expressions are first compared by plotting their ratio, λ, given by Eq. (53) against $\tilde{H}$ in Fig. 4, for a steel particle suspended in water with a wire of length $L=76$ mm, under different vibration conditions. The resonance frequency for this case is about 1.60–1.64 Hz depending on the value of $\tilde{H}$, and Figs. 4(*a*) and 4(*b*) show the values of λ for vibration frequencies below and above the resonance frequency, respectively. It is clear that for $\tilde{H} > 1.5$, the two expressions give predictions that agree within 1%. At vibration frequencies well above the resonance frequency, for example, 2.5 and 3 Hz in Fig. 4(*b*), the approximate

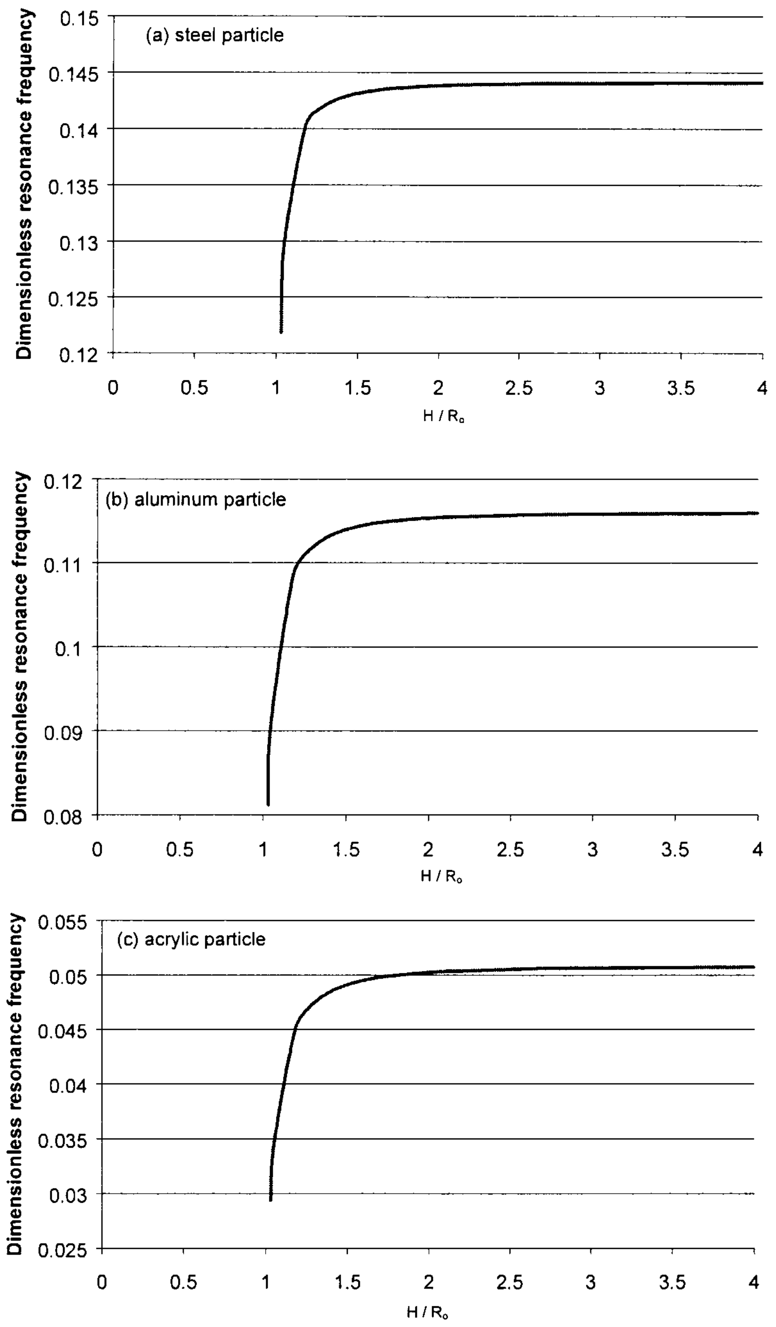


Fig. 9 Variation of the dimensionless resonance frequency predicted by Eq. (54) with the particle-wall distance to particle radius ratio, H/R_0 (wire length=76 mm)

Eq. (52) can be used to predict the particle amplitude even near the lowest value of $\tilde{H}=1.2$ shown in Fig. 4(*b*).

Much greater values of the ratio λ can, however, be obtained at near the resonance frequencies as shown in Fig. 5, for a steel particle in water with different wire lengths and a fixed cell amplitude. The ratio λ increases from 1 to 4.5 in Fig. 5(*a*) for below-resonance frequencies, but decreases from 1 to 0.4 in Fig. 5(*b*) for above-resonance frequencies. Thus, although the approximate Eq. (52) is simpler and easier to use, we will analyze the predictions of Eq. (51) in greater detail in the following.

4.1 Particle Amplitude in a Semi-Infinite Cell. The dimensionless particle amplitude in a semi-infinite cell for a given vibration condition is affected by the particle distance to the nearest cell wall, *H*, but the effect is different depending on whether the vibration frequency is above or below the resonance frequency. The predictions of Eq. (51) for a steel particle in water are shown against $\tilde{H}$ for vibration frequencies below and above the resonance frequency in Figs. 6(*a*) and 6(*b*), respectively. In both figures, the cell amplitudes have been chosen so that all the particle amplitudes are similar and can be plotted in the same figure. The particle amplitude for a constant cell vibration amplitude decreases slightly with the increasing distance of the particle to the nearest cell wall if the vibration frequency is below the resonance frequency (Fig. 6(*a*)). On the other hand, above the resonance frequency, the particle amplitude increases slightly with the increasing particle-wall distance (Fig. 6(*b*)), which is intuitively ex-

Table 1 Experimental particle amplitudes in microns for different particle densities and the widest cell width (cell vibration frequency: 0.5–2.5 Hz)

Frequency (Hz)	0.5	0.75	1	1.25	1.5	1.9	2	2.5
Steel particle (ρ=7.83 g/cm^3)	80	205	486	1130	2213	1500	2420	1390
Aluminum particle (ρ=2.7 g/cm^3)	94	265	760	2000	2100	990	900	716
Acrylic particle (ρ=1.17 g/cm^3)	250	450	150	128	118	105	110	10

Table 2 Experimental particle amplitudes in microns for different particle densities for the widest cell width (cell vibration frequency: 3.0–10.0 Hz)

Frequency (Hz)	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0
Steel particle	1180	980	925	870	860	855	840	830
Aluminum particle	656	580	560	540	530	520	510	505
Acrylic particle	95	100	110	100	100	110	105	95

pected due to the confinement effect of the cell wall.

The wall effect is amplified near the resonance frequency as illustrated in Figs. 7(*a*) and 7(*b*) which show the ratio of the steel particle amplitude in a semi-infinite cell as predicted by Eq. (51) to that in an infinite cell ($H/R_0 \to \infty$) for cell vibration frequencies below and above the resonance frequency, respectively. As the cell vibration frequency gets closer to the resonance frequency, the particle amplitude increases rapidly, so the wire length has been reduced in Figs. 7(*a*) and 7(*b*) to show three different cases that yield similar values of particle amplitude ratios.

The changes in the particle amplitudes shown in Fig. 6 and the ratio of amplitudes in semi-infinite to infinite cells in Fig. 7 are due to the wall proximity effect as well as a shift in the resonance frequency with the particle-wall distance as discussed next.

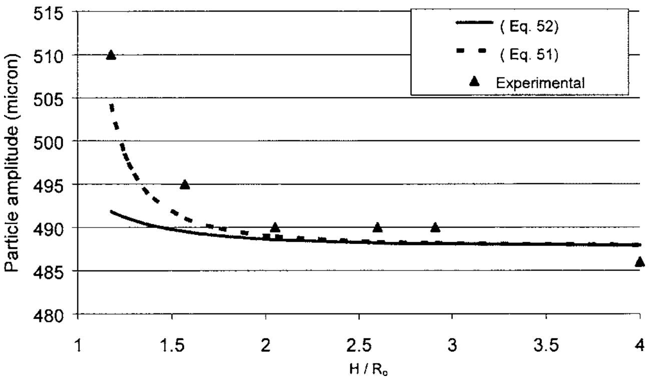


Fig. 10 Variation of particle amplitudes with particle-wall distance to particle radius ratio predicted by Eqs. (51) and (52) (steel particle in water, f=1.0 Hz, cell amplitude=1.0 mm, wire length=76 mm)

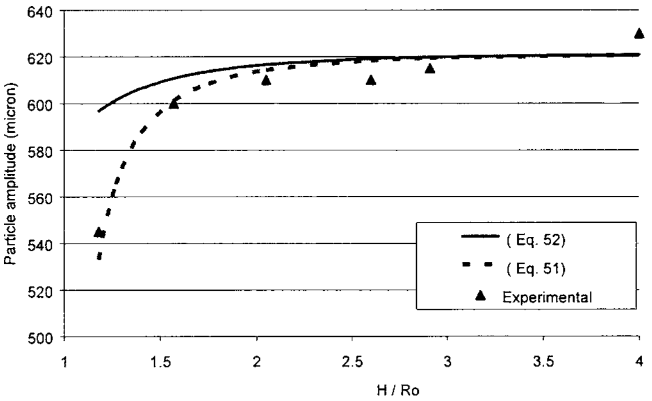


Fig. 12 Variation of particle amplitudes with particle-wall distance to particle radius ratio predicted by Eqs. (51) and (52) (steel particle in water, f=2.0 Hz, cell amplitude=0.25 mm, wire length=76 mm)

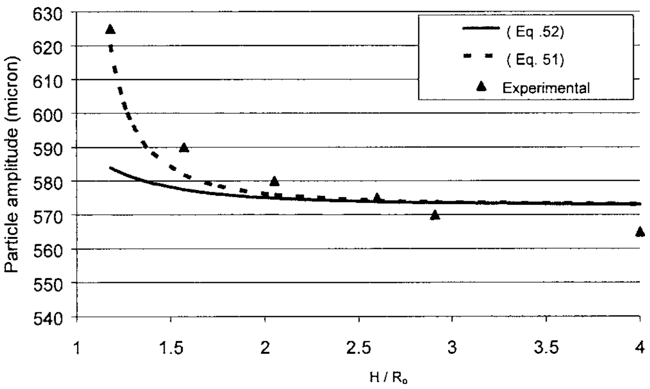


Fig. 11 Variation of particle amplitudes with particle-wall distance to particle radius ratio predicted by Eqs. (51) and (52) (steel particle in water, f=1.25 Hz, cell amplitude=0.50 mm, wire length=76 mm)

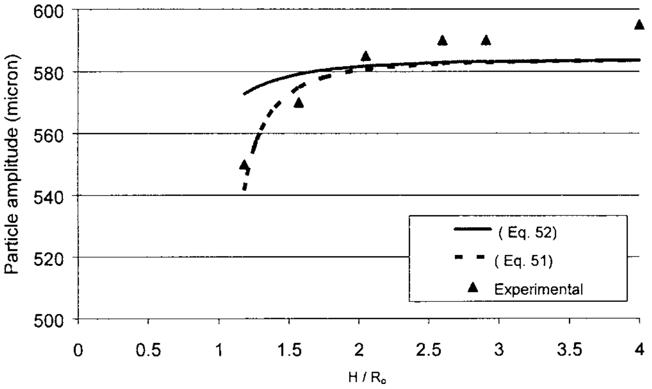


Fig. 13 Variation of particle amplitudes with particle-wall distance to particle radius ratio predicted by Eqs. (51) and (52) (steel particle in water, f=3.0 Hz, cell amplitude=0.50 mm, wire length=76 mm)

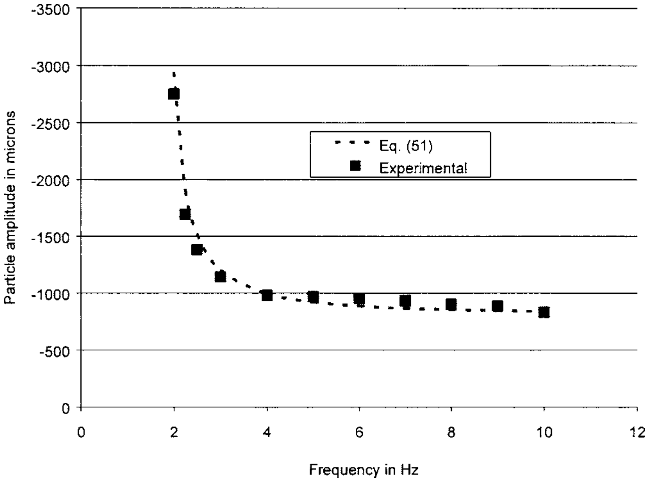


Fig. 14 Theoretical and experimental variations of particle amplitude with cell vibration frequency (steel particle in water, wire length=70 mm, widest cell tested, cell width=50.8 mm)

4.2 Effect of Particle-to-Wall Distance on Resonance Frequency. Figure 8 shows the variations in the resonance frequency with $\tilde{H}$ predicted by Eqs. (54) and (55) for different particle densities when the wire length is fixed at 76 mm. As shown in Fig. 8, Eq. (55) is valid only for $\tilde{H}>1.75$. If the ratio $\tilde{H}$ is below about 2, the resonance frequency is predicted by both equations to decrease, and the difference between the two predictions also increases. Figure 9 shows the dramatic reductions in the resonance frequency predicted by Eq. (54) as the value of $\tilde{H}$ gets closer to 1.0.

4.3 Comparison Between Experimental Data and Theory. The experimental data obtained for three different particle densities in the widest cell (half width of 25.4 mm), where the particle is in the center of the cell, are summarized in Tables 1 and 2 for different cell vibration frequencies below and above the resonance frequency, respectively.

The wire length used was 76 mm for steel and aluminum particles, and 70 mm for an acrylic particle. The cell amplitude for the steel particle was 1.0 mm for all the vibration frequencies, except at 1.5 and 1.9 Hz for which the cell amplitude was reduced to 0.5 mm because of their closeness to the resonance frequency. Similarly, the particle amplitudes obtained for aluminum and acrylic particles are for the cell amplitude of 1.0 mm except near the resonance at 1.25 Hz (aluminum particle), and 0.6 Hz (acrylic particle) where the cell amplitude was reduced to 0.5 mm.

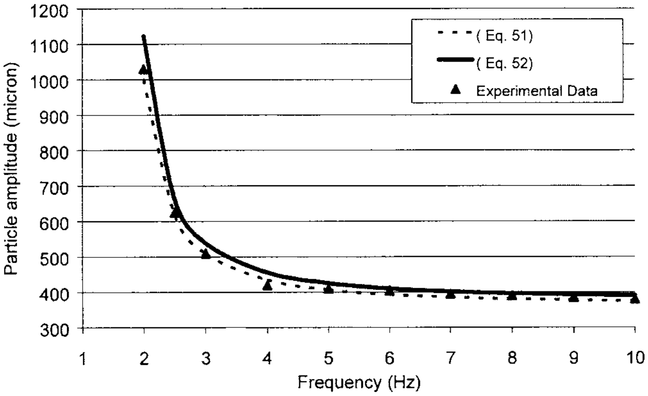


Fig. 15 Comparison of measured particle amplitudes with the predictions of Eqs. (51) and (52) for different vibration frequencies ($H/R_0=1.18$, cell amplitude=0.47 mm, wire length=76 mm)

For the steel particle in water, the particle amplitudes are plotted against $\tilde{H}$ in Figs. 10 and 11 for below-resonance frequencies and in Figs. 12 and 13 for above-resonance frequencies, respectively. The experimental data obtained for smaller particle-wall distances ($\tilde{H}<3$) were normalized by the particle amplitude data obtained in the widest cell with a half width of 25.4 mm. In comparison with the experimental data, while the particle amplitudes predicted by Eqs. (51) and (52) show consistent trends for the effect of the increasing particle-wall distance to particle radius ratio, $\tilde{H}$, Eq. (51) is seen to give better agreement with the experimental data.

The experimental values shown in Figs. 10–13 are the average values obtained from 15 to 20 separate measurements for each run. Thus, the experiments have been repeated at least 15 times

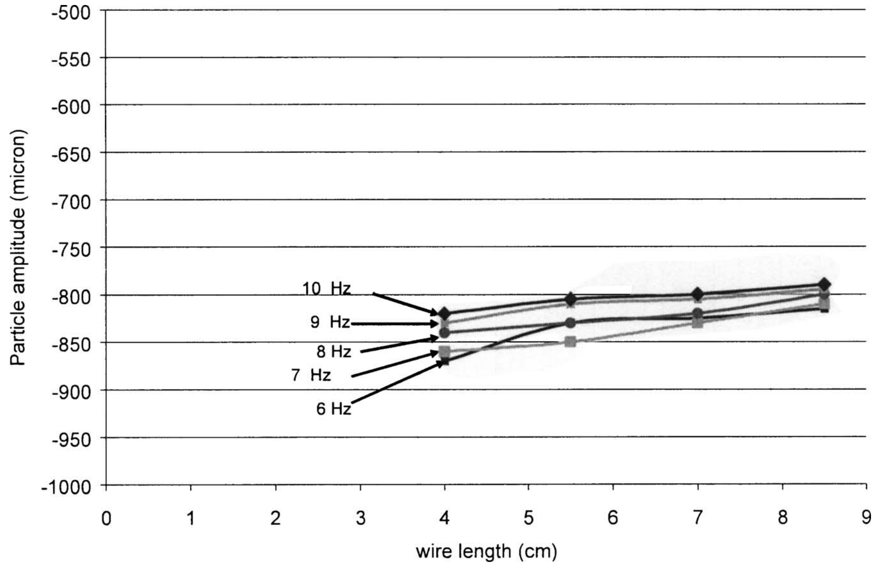


Fig. 16 Particle amplitude variation with wire length for different cell vibration frequencies (steel particle in water, cell amplitude=1 mm, $H/R_0=1.18$)

for a given condition. Error bars of ±3% were obtained for the data points which represent the maximum and minimum values for a given run.

4.4 Particle Motion at High Cell Vibration Frequencies in a Semi-Infinite Cell. As in the experiments conducted using an infinite cell [18], the particle amplitude remained nearly constant at high frequencies. This can be seen clearly in Figs. 14 and 15 where the particle amplitude data obtained for a steel particle located in the center of the widest cell (infinite cell) and at the minimum distance to the cell wall ($\tilde{H}=1.18$), respectively, are plotted against the cell vibration frequency. Theoretical predictions are also plotted in the same figures for vibration frequencies above the resonance frequency. In all the cases, the particle amplitude decreases rapidly with the cell vibration frequency, and reaches a nearly constant value for vibration frequencies above 5 Hz. The present theory predicts an asymptotic particle amplitude, independent of the vibration frequency and wire length at high cell vibration frequencies as given by Eqs (51) or (52). Although both theoretical predictions agree well with the experimental results even at $\tilde{H}=1.18$, Eq. (51) again shows better agreement with the data.

Finally, the lack of dependence of the particle amplitude on the wire length at high cell vibration frequencies is clearly shown in Fig. 16. For a cell vibration amplitude of 1.0 mm in the smallest semi-infinite cell ($\tilde{H}=1.18$) tested, the asymptotic particle amplitude was measured to be ~0.83 mm, which agrees well with the value predicted by Eq. (51) or Eq. (52) for a steel particle in water. Figure 16 shows that the particle amplitude remains nearly constant and independent of the wire length at high vibration frequencies. This indicates that for cell vibration frequencies well above the resonance frequency, the motion of a spherical particle suspended by a thin wire in a finite cell on the ground may correspond to the motion of a wire-free particle in a fluid cell under microgravity. This result establishes a firm basis on which the vibration-induced particle motion in microgravity could be studied under normal gravity using a particle suspended with a thin wire and vibrated at sufficiently high frequencies. Similar results were also obtained for aluminum and acrylic particles at vibration frequencies well above the resonance frequencies for different cell amplitudes and wire length.

5 Conclusion

A study of vibration-induced particle motion in a water-filled fluid cell has been conducted experimentally and theoretically. The effect of the distance between the particle and the nearest cell wall was investigated by measuring the amplitude of a solid particle suspended by a thin wire in the fluid cell. The semi-infinite cell with a varying distance between the particle and the nearest wall was vibrated horizontally using a linear translation stage at a certain amplitude and frequency. The effects of wire length and particle density have been investigated by using spherical particles made of steel, aluminum, or acrylic. An inviscid model was also developed using the method of images to analytically predict the particle amplitude for different particle-wall distances, wire lengths, and vibration conditions.

The inviscid model was found to be capable of predicting the measured particle amplitude, and its dependence on the particle-wall distance. Reducing the distance between the particle and the nearest wall was found to decrease (or increase) the particle amplitude when the cell vibration frequency is above (or below) the resonance frequency. The resonance frequency was also found to significantly decrease with a reduction in the particle-wall distance. Thus, the wall proximity effect on the particle amplitude could be attributed partly to the change in the resonance frequency.

At vibration frequencies well above the resonance frequency, the particle amplitude was found to reach a constant value, independent of both the wire length and cell vibration frequency. This suggests that the vibration-induced particle amplitudes measured on the ground with a particle suspended by a wire may be regarded as representative of vibration-induced amplitudes of a wire-free particle under microgravity.

Acknowledgment

This work was financially supported by a grant from the Canadian Space Agency. One of the authors (S. Hassan) also received a graduate fellowship from the Government of Ontario.

Nomenclature

a = cell vibration amplitude, m
A_p = particle amplitude, m
D_0 = particle diameter, m
f = cell vibration frequency, Hz
$\mathbf{f}$ = wire tension, N
$\mathbf{g}$ = gravitational acceleration, $\mathrm{m/s^2}$
H = distance between the particle and the nearest wall, m
$\tilde{H}$ = dimensionless distance of the particle to the nearest wall
$\mathbf{i}$ = unit vector in the direction of cell motion
L = length of the wire from the attachment point to the center of mass of the sphere, m
m = particle mass, kg
p = static pressure, $\mathrm{N/m^2}$
$\mathbf{r}$ = fluid element position with respect to the particle, m
$\mathbf{R}$ = absolute position of the particle, m
R_0 = particle radius, m
t = time, s
W = distance between the particle and its image ($=2H$), m
$\mathbf{v}$ = fluid velocity relative to the particle, m/s
X_p = horizontal displacement of the particle with respect to the cell, m

Greek Symbols

ρ_S = particle density, $\mathrm{kg/m^3}$
ρ_L = liquid density, $\mathrm{kg/m^3}$
φ = velocity potential, $\mathrm{m^2}/s$
ω = cell vibration frequency, rad/s

Subscripts

s = solid particle
l = liquid
w = wall

References

[1] Coimbra, C. F. M., and Rangel, R. H., 2001, "Spherical Particle Motion in Harmonic Stokes Flows," AIAA J., **39**(9), pp. 1673–1682.
[2] Trolinger, J. D., Rottenkolber, M., and Elandaloussi, F., 1997, "Development and Applications of Holographic Particle Image Velocimetry Techniques for Microgravity Applications," Meas. Sci. Technol., **8**, pp. 1573–1583.
[3] Gamache, O., and Kawaji, M., 2005, "Experimental Investigation of Marangoni Convection and Vibration-Induced Crystal Motion During Protein Crystal Growth," Microgravity Sci. Technol., **XVI-I**, pp. 342–347.
[4] Basset, A. B., 1888, *A Treatise on Hydrodynamics*, Deighton, Bell and Co., Cambridge, Vol. 2, Chap. 21. (Also New York: Dover Publications, Inc., 1961).
[5] Boussinesq, J. V., 1885, "Sur la Resistance Qu'oppose un Liquide Indefeni au Repos, sans Pesanteur, au Mouvememt d'une Sphere Solide qu'il Mouille sur toute sa Surface," C. R. Seances Acad. Sci. III, **100**, pp. 935–937.
[6] Oseen, C. W., 1910, "Uber die Stokes'sche Formel und Uber eine Verwandte Aufgabe in der Hydrodynamik," Ark. Mat., Astron. Fys., **6**(29), pp. 1–20.
[7] Oseen, C. W., 1927, *Hydromechanik*, Akademicshe, Leipzig.
[8] Stokes, C. G., 1851, *Mathematical and Physical Papers*, Johnson Reprint Corp., New York, Vol. 3, pp. 25–35.
[9] Maxey, M., and Riley, J., 1982, "Equation for a Small Rigid Sphere in a Non-Uniform Flow," Phys. Fluids, **26**, pp. 883–889.

[10] Tchen, C. M., 1947, “Mean Value and Correlation Problems Connected With the Motion of Small Particles Suspended in a Turbulent Fluid,” Ph.D. thesis, Delft University.

[11] Lamb, H., 1932, *Hydrodynamics*, Cambridge University Press, London.

[12] Milne-Thomson, L. M., 1968, *Theoretical Hydrodynamics*, 5th ed., Macmillan, London.

[13] Eames, I., Hunt, J. C., and Belcher, S. E., 1996, “Displacement of Inviscid Fluid by a Sphere Moving Away From a Wall,” J. Fluid Mech., **234**, pp. 33–53.

[14] Houghton, G., 1961, “The Behaviour of Particles in a Sinusoidal Vector Field,” Proc. R. Soc. London, Ser. A, **272**, pp. 33–43.

[15] Li, L., Schultz, W. W., and Merte, H., 1993, “The Velocity Potential and the Interacting Force of Two Spheres Moving Perpendicularly to the Line Joining Their Centers,” J. Eng. Math., **27**, pp. 147—160.

[16] Magnaudet, J., and Eames, I., 2000, “The Motion of High Reynolds-Number Bubbles in Inhomogeneous Flows,” Annu. Rev. Fluid Mech., **32**, pp. 659–708.

[17] Magnaudet, J., 2003, “Small Inertial Effects on a Spherical Bubble, Drop or Particle Moving Near a Wall in a Time-Dependent Linear Flow,” J. Fluid Mech., **485**, pp. 115–142.

[18] Hassan, S., Kawaji, M., Lyubimova, T. P., and Lyubimov, D. V., 2006, “Motion of a Sphere Suspended in a Vibrating Liquid-Filled Container,” ASME J. Appl. Mech., **73**, pp. 72–78.

Exact Solutions for the Functionally Graded Plates Integrated With a Layer of Piezoelectric Fiber-Reinforced Composite

M. C. Ray

H. M. Sachade

Department of Mechanical Engineering,
Indian Institute of Technology,
Kharagpur 721 302,
India

This paper deals with the derivation of exact solutions for the static analysis of functionally graded (FG) plates integrated with a layer of piezoelectric fiber reinforced composite (PFRC) material. The layer of the PFRC material acts as the distributed actuator of the FG plates. The Young's modulus of the FG plate is assumed to vary exponentially along the thickness of the plate while the Poisson's ratio is assumed to be constant over the domain of the plate. The numerical values of the exact solutions are presented for both thick and thin smart FG plates and indicate that the activated PFRC layer potentially counteracts the deformations of the FG plates due to mechanical load. The through-thickness behavior of the plates revealed that the coupling of bending and extension takes place in the FG plates even if the PFRC layer is not subjected to the applied voltage. The solutions also revealed that the activated PFRC layer is more effective in controlling the deformations of the FG plates when the layer is attached to the surface of the FG plate with minimum stiffness than when it is attached to the surface of the same with maximum stiffness. The solutions of this benchmark problem may be useful for verifying the other approximate and numerical models of the smart functionally graded plates for which exact solutions cannot be derived. [DOI: 10.1115/1.2165230]

1 Introduction

Recently, a new class of materials has been emerged which exhibit variation of material properties particularly across the thickness direction. Such class of materials is called functionally graded materials (FGM). In an endeavor to develop the super heat resistant materials for using in space plane, Koizumi [1] first proposed the concept of FGM. A general introduction to these FGMs encompassing the fabrication, characterization, and design of these materials has been presented by Suresh and Mortensen [2]. The laminated composite structures can be tailored to design advanced structures with high stiffness to weight ratios, high strength to weight ratios, and better thermal and transport properties. However, the sharp demarcation of the properties of each layer at the interface between the two adjacent layers of laminated composite structures causes large interlaminar shear stresses which eventually may give rise to the initiation of delamination. Such detrimental effect can be circumvented if the properties are smoothly varied across the thickness direction and thus the use of FGM may become an important issue for advanced structural applications. During the past few years a great amount of research has been reported to analyze the behavior of functionally graded materials. For example, Noda and Jin [3] derived a model to determine the stress intensity factors of a functionally graded body having a crack subjected to prescribed temperature field. Teymur et al. [4] carried out the thermomechanical analysis of materials which are functionally graded in two directions and demonstrated that the onset of delamination could be prevented by tailoring the microstructures of the composite plies. Feldman and Aboudi [5] studied the elastic bifurcational buckling of functionally graded plates under in-plane compressive load. They concluded that with optimal nonuniform distribution of reinforcing phases, the buckling load can be significantly improved for the FG plate over the plate with uniformly distributed reinforcing phase. Gu and Asaro [6] analytically investigated the behavior of functionally graded material in the presence of a semi-infinite crack. Mian and Spencer [7] derived the exact solutions for functionally graded plates with zero surface traction. Praveen and Reddy [8] investigated the nonlinear thermoelastic behavior of functionally graded ceramic metal plates. Lee and Yu [9] derived the two-dimensional governing equations for electroded piezoelectric plates with general symmetry and graded properties across the thickness from the three-dimensional equation of linear piezoelectricity. Loy et al. [10] studied the vibration of cylindrical shells made of a functionally graded material which is composed of stainless steel and nickel. Aboudi et al. [11] further developed a more general higher-order theory for functionally graded materials and illustrated the utility of functionally graded microstructures in tailoring the behavior of structural components in various applications. In 2000, Wang et al. [12] proposed a method to determine the transient and steady state thermal stress intensity factors of graded composite plate containing noncollinear cracks subjected to dynamic thermal loading. Yang [13] presented an analytical solution for computing the time-dependent stresses in FGM undergoing creep. Yang and Shen [14] studied the dynamic response of initially stressed functionally graded thin plates subjected to partially distributed impulsive loads. An elasticity solution for functionally graded beams is provided by Shankar [15] in which the beam properties are graded in the thickness direction according to an exponential law. The exact solutions for thermoelastic deformations of thick FG plates subjected to both thermal and mechanical loads have been presented by Vel and Batra [16]. Woo and Meguid [17] presented an analytical solution for the large deflections of plates and shallow shells made of FGMs under the combined action of thermal and

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 15, 2005; final manuscript received April 12, 2005. Review conducted by Z. Suo. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

mechanical loads. Shen [18] also carried out nonlinear bending analysis for simply supported functionally graded plate subjected to mechanical and thermal loadings using a higher-order shear deformation theory. Considering time-dependent thermal loads, Vel and Batra [19] presented an analytical solution for three-dimensional thermomechanical deformation of a simply supported FG plate. Zhong and Shang [20] presented an exact three-dimensional analysis of a simply supported functionally graded piezoelectric plate considering that the mechanical and electrical properties are the same exponential function of thickness coordinate. More recently, Najafizadeh and Hedayati [21] studied the thermoelastic stability of functionally graded circular plates.

In the quest for developing lightweight high performing flexible structures a concept was emerged to develop the structures with self-controlling and self-monitoring capabilities. Expediently, it was discovered that if the piezoelectric materials are used as distributed actuators and sensors which are either mounted on or embedded in the structure then the structure attains these self-controlling and self-monitoring capabilities [22,23]. Such structures are customarily called "smart structures." Piezoelectric materials induce an electric potential/charge when they are subjected to a mechanical load by virtue of the direct piezoelectric effect and are deformed due to the externally applied voltage/charge by virtue of the converse piezoelectric effect. Use of piezoelectric materials as distributed sensors and actuators for developing smart structures is attributed to these two phenomena. The concept of developing smart structures has been extensively employed for the active vibration control of flexible structures during the past decade [24–31]. Recently, a considerable interest has also been generated to analyze the functionally graded structures integrated with piezoelectric actuators. For example, Ootao and Tanigawa [32] carried out a theoretical analysis of simply supported functionally graded plate integrated with a piezoelectric plate subjected to transient thermal loading. They subsequently studied the control of this smart functionally graded plate subjected to nonuniform heat supply [33]. Reddy and Cheng [34] presented the three-dimensional solution for functionally graded plates coupled with piezoelectric actuator layer employing transfer matrix and asymptotic expansion techniques. Wang and Noda [35] analyzed a smart functionally graded composite structure composed of a layer of metal, a layer of piezoelectric material (PZT), and a functionally graded layer between the metallic layer and the piezoelectric layer. They showed that by the introduction of this functionally graded layer between the piezoelectric actuator and the metallic layer, both stress discontinuity and edge local stresses can be essentially reduced. They also investigated the thermally induced fracture in a smart functionally graded structure [36]. Yang et al. [37] performed an analysis of prestressed functionally graded laminated plates integrated with piezoelectric actuator to predict the large amplitude vibration behavior of the plates.

Although the monolithic piezoelectric materials are being widely used as the distributed actuators for smart structures, the main drawback of these existing monolithic piezoelectric materials is that the control authority of these materials is very low as their piezoelectric stress/strain coefficients are of very small magnitudes. As the active damping of smart structures depends on the control authority of the piezoelectric materials, tailoring of the piezoelectric stress/strain coefficients may improve their control authority and hence the damping characteristics of lightweight smart structures can be improved. In an endeavor to tailor the piezoelectric properties, Mallik and Ray [38] newly proposed the concept of longitudinally piezoelectric fiber reinforced composite (PFRC) materials and investigated the effective mechanical and piezoelectric properties of these composites. The main concern of the investigations was to determine the effective piezoelectric coefficient (e_{31}) of these new concept PFRC materials which quantifies the induced normal stress in the fiber direction due to the applied electric field in the direction transverse to the fiber direction. They predicted that this effective piezoelectric coefficient becomes significantly larger than the corresponding coefficient of piezoelectric material of the fibers. Note that the significant achievement of active control of flexural vibration of smart structures depends mainly on the magnitude of this piezoelectric coefficient.

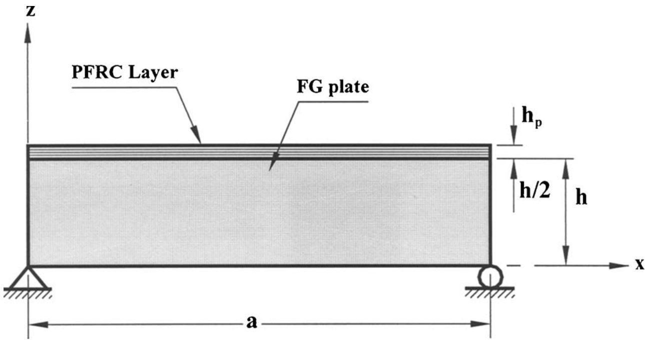


Fig. 1 Functionally graded plate integrated with a PFRC layer

In an attempt toward the development of new functionally graded smart structures, the authors intend to carry out the static analysis of functionally graded plates integrated with a layer of this new piezoelectric fiber reinforced composite (PFRC) material and this paper is concerned with the derivation of exact solutions for this analysis. The PFRC layer is considered to act as the distributed actuator of the FG plate. Pagano's method [39] has been employed to derive the exact solutions for displacements and stresses of these smart FG plates due to the combined action of the mechanical load and the activated PFRC layer. The exact solutions of this benchmark problem could serve for the development of new high performing smart functionally graded structures and may be used for verifying the other approximate theories and numerical models.

2 Methodology

In what follows, the method of deriving the exact solutions for static behavior of the FG plate coupled with the PFRC layer will be discussed. Figure 1 illustrates a simply supported plate made of a FG material. Although the top surface of the plate has been shown to be integrated with a layer of PFRC material which acts as the distributed actuator of the FG plate, the PFRC layer can also be attached to the bottom surface of the plate. The length, width, and thickness of plate are denoted by a, b, and h, respectively, while the thickness of the PFRC layer is denoted by h_p. The microstructural feature of the PFRC material has been described in Ref. [38] and is not repeated here. The piezoelectric fibers in the PFRC layer are considered to be aligned longitudinally in the plane of the layer which is parallel to the reference plane. When the PFRC layer is subjected to an applied electric field, the in-plane actuation of the activated layer can be exploited to control the deformations of the substrate FG plate and the resulting overall plate becomes a new smart functionally graded plate. The bottom surface of the substrate FG plate is considered as the reference plane and the origin of the reference coordinate system (x,y,z) is located at one corner of the bottom surface of the plate such that $x=0,a$ and $y=0,b$ represent the edges of the substrate FG plate. In order to obtain the exact solutions for static analysis of this smart FG plate subjected to both mechanical and electrical load, the expressions for displacements and stresses are to be derived for the substrate FG plate and the PFRC layer satisfying all the governing field equations, boundary conditions, surface tractions, and the interface continuity conditions between the FG plate and the PFRC layer. Since the coupling of electric field and elastic field occur in the PFRC material, the expressions for electric displacements and potential are also required to be derived for the PFRC layer satisfying all the electrical boundary conditions for obtaining the exact solutions of the smart FG plate.

The constitutive equations for the material of the elastic FG plate are given by [15]

$$\{\sigma^k\} = [C^k]\{\in^k\} \quad k=1 \tag{1}$$

and those for the PFRC material are [38]

$$\{\sigma^k\} = [C^k]\{\in^k\} - [e]\{\mathbf{E}\} \quad k=2 \tag{2}$$

$$\{\mathbf{D}\} = [e]^T\{\in^k\} + [\varepsilon]\{\mathbf{E}\} \quad k=2 \tag{3}$$

in which the vectors $\{\sigma\}$ and $\{\epsilon\}$ represent the state of stress and state of strain at any point in the continuum denoted by the superscript k $[C]$ is the matrix of elastic constant, $[e]$ is the matrix of piezoelectric stress constants, $[\varepsilon]$ is the dielectric constant matrix, $\{\mathbf{E}\}$ is the electric field vector, and $\{\mathbf{D}\}$ is the electric displacement vector. The components of the stress and strain vectors are given by

$$\{\sigma\} = [\sigma_x \ \sigma_y \ \sigma_z \ \sigma_{yz} \ \sigma_{xz} \ \sigma_{xy}]^T \quad \{\in\} = [\in_x \ \in_y \ \in_z \ \in_{yz} \ \in_{xz} \ \in_{xy}]^T \tag{4}$$

where σ_x, σ_y, and σ_z are the normal stresses in the x, y, and z directions, respectively, σ_{xz} and σ_{yz} are the transverse shear stresses; σ_{xy} is the in-plane shear stress; ϵ_x, ϵ_y, ϵ_z, ϵ_{xz}, ϵ_{yz}, and ϵ_{xy} are the corresponding strains. The piezoelectric coefficient matrix $[e]$ and the dielectric constant matrix of the PFRC material are given by

$$[e]^T = \begin{bmatrix} 0 & 0 & 0 & 0 & e_{15} & 0 \\ 0 & 0 & 0 & e_{24} & 0 & 0 \\ e_{31} & e_{32} & e_{33} & 0 & 0 & 0 \end{bmatrix} \quad [\varepsilon] = \begin{bmatrix} \varepsilon_{11} & 0 & 0 \\ 0 & \varepsilon_{22} & 0 \\ 0 & 0 & \varepsilon_{33} \end{bmatrix} \tag{5}$$

The electric field vector $\{\mathbf{E}\}$ and the electric displacement vector $\{\mathbf{D}\}$ appearing in Eqs. (2) and (3) are given by

$$\{\mathbf{E}\} = [\mathbf{E}_x \ \mathbf{E}_y \ \mathbf{E}_z]^T \quad \{\mathbf{D}\} = [\mathbf{D}_x \ \mathbf{D}_y \ \mathbf{D}_z]^T \tag{6}$$

where $\mathbf{E}_x$, $\mathbf{E}_y$, and $\mathbf{E}_z$ are the electric fields along x, y, and z directions, respectively, and $\mathbf{D}_x$, $\mathbf{D}_y$, and $\mathbf{D}_z$ are the corresponding electric displacements.

The assumption regarding the material properties of the functionally graded materials employed by several researchers [3,15,20,40,41] which is not only simple for mathematical manipulations but also provides essential features of FGM is that the properties of FGM are exponential functions of thickness coordinate. Thus in view of the need for simplicity in solving the elasticity problem, it may be assumed that the FG material considered in this study is isotropic at any point but its Young's modulus E is an exponential function of z and is given by [3,15,20,40,41]

$$E = E_0 e^{\lambda z} \tag{7}$$

where E_0 refers to the Young's modulus of the material located at the bottom surface of the FG plate and λ is a parameter describing the inhomogeneity of the FG material across the thickness. It is also assumed that the Poisson's ratio is constant over the domain of the FG plate [15,20,41]. Thus the elastic coefficients of the material of the FG plate are given by

$$C_{ij}^1 = C_{ij}^0 e^{\lambda z} \quad (i,j = 1,2,3, \ldots, 6) \tag{8}$$

in which C_{ij}^0 are the elastic constants of the isotropic material located at the bottom surface of the FG plate. Thus based on the value of λ, either the top or the bottom surface of the FG plate will be softest. In the absence of body forces, the governing equilibrium equations for an elastic medium are given by

$$\sigma_{ij,j} = 0 \quad i,j = 1,2, \text{ and } 3 \tag{9}$$

in which the comma in the subscript followed by the subscript j indicates the partial differentiation with respect to the coordinate denoted by j and repeated index implies summation over the range of the index. Also, it should be noted that the values (1, 2, 3) of the subscript i, j refer to the x, y, and z coordinates, respectively. In addition to satisfying the governing equilibrium equations, the charge equilibrium equation must be satisfied for obtaining the exact solutions for the PFRC layer and is given by

$$D_{i,i} = 0 \quad i = x,y, \text{ and } z \tag{10}$$

In the present investigation, the simply supported boundary conditions as considered by Pagano [39] will be used for obtaining the exact solutions and are given by

$$\sigma_x^k = v^k = w^k = 0 \quad \text{at } x = 0,a \quad \sigma_y^k = u^k = w^k = 0 \quad \text{at } y = 0,b$$

$$k = 1 \text{ and } 2 \tag{11}$$

in which u, v, and w are the mechanical displacement components at any point in the continuum denoted by the superscript k along the x, y, and z directions, respectively. In order to derive the exact solution for electric potential (ϕ) in the PFRC layer, it is assumed that the edges of the PFRC layer are suitably grounded such that the electric potential at these edges become zero. Therefore the essential electric boundary conditions for the PFRC layer are [20,42]

$$\phi = 0 \text{ at } x = 0,a \quad y = 0,b \tag{12}$$

For a particular mode of deformation, the admissible functions for the displacements and electric potential which satisfy the boundary conditions given by Eqs. (11) and (12) can be expressed as [20,39,42]

$$\begin{Bmatrix} u^k(x,y,z) \\ v^k(x,y,z) \\ w^k(x,y,z) \\ \phi(x,y,z) \end{Bmatrix} = \begin{bmatrix} \cos px \sin qy & 0 & 0 & 0 \\ 0 & \sin px \cos qy & 0 & 0 \\ 0 & 0 & \sin px \sin qy & 0 \\ 0 & 0 & 0 & \sin px \sin qy \end{bmatrix} \begin{Bmatrix} U^k(z) \\ V^k(z) \\ W^k(z) \\ \Phi(z) \end{Bmatrix} \quad k = 1,2 \tag{13}$$

in which $p = m\pi/a$ and $q = n\pi/b$ and m, n denote the mode numbers. The functions $U^k(z)$, $V^k(z)$, $W^k(z)$ appearing in Eq. (13) describe the variations of the axial displacements across the thickness of the substrate FG plate and the PFRC layer while the function $\Phi(z)$ represents the variation of the electric potential across the thickness of the PFRC layer.

2.1 Solutions of FG Plate. In case of functionally graded plates, the functions $U^1(z)$, $V^1(z)$, and $W^1(z)$ can be assumed as [39]

$$[U^1(z) \ V^1(z) \ W^1(z)] = [U^{01} \ V^{01} \ W^{01}]e^{tz} \tag{14}$$

where U^{01}, V^{01}, W^{01} and t are the unknown constants to be determined. Substituting Eq. (1) into Eq. (9) and then using the strain-displacement relations [39] in conjunction with Eqs. (7), (8), (13), and (14) a set of homogeneous algebraic equations are obtained for the solutions of the FG plate which can be expressed in a matrix form as follows

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{Bmatrix} U^{01} \\ V^{01} \\ W^{01} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix} \tag{15}$$

The various coefficients appearing in Eq. (15) are given by

$$a_{11} = C_{66}^0(t^2 + \lambda t) - \bar{a}_{11}, \quad a_{22} = C_{66}^0(t^2 + \lambda t) - \bar{a}_{22},$$

$$a_{33} = C_{11}^0(t^2 + \lambda t) - \bar{a}_{33}, \quad a_{12} = -pqC_{s1}^0$$

$$C_{s1}^0 = C_{12}^0 + C_{66}^0, \quad a_{13} = p(tC_{s1}^0 + \lambda C_{66}^0), \quad a_{23} = q(tC_{s1}^0 + \lambda C_{66}^0)$$

$$a_{31} = -p(tC_{s1}^0 + \lambda C_{12}^0), \quad a_{32} = -q(tC_{s1}^0 + \lambda C_{12}^0),$$

$$\bar{a}_{11} = C_{11}^0 p^2 + C_{66}^0 q^2$$

$$\bar{a}_{22} = C_{66}^0 p^2 + C_{11}^0 q^2 \quad \bar{a}_{33} = C_{66}^0(p^2 + q^2)$$

Using the last two equations of (15), the unknown constants V^{01} and W^{01} can be expressed in terms of the unknown constant U^{01} as follows

$$V^{01} = F^1 U^{01} \quad W^{01} = F^2 U^{01} \tag{16}$$

in which the coefficients F^1 and F^2 are functions of the characteristic parameter t and are given by

$$F^1 = \frac{a_{23}a_{31} - a_{21}a_{33}}{a_{22}a_{33} - a_{23}a_{32}} \quad F^2 = \frac{a_{21}a_{32} - a_{31}a_{22}}{a_{22}a_{33} - a_{23}a_{32}} \tag{17}$$

In order to find the nontrivial solutions of the unknown constant U^{01}, Eqs. (16) and (17) are substituted into the first equation of (15) and the following characteristic equation for the FG plate is obtained

$$\sum_{n=1}^{6} A_n t^n = 0 \tag{18}$$

where

$$A_6 = C_{11}^0 C_{66}^{0\,2}, \quad A_5 = 3\lambda C_{11}^0 C_{66}^{0\,2}$$

$$A_4 = 3\lambda^2 C_{11}^0 C_{66}^{0\,2} + C_{s1}^{0\,2} C_{66}^0(p^2 + q^2) - C_{66}^0(C_{11}^0\bar{a}_{11} + C_{11}^0\bar{a}_{22} + C_{66}^0\bar{a}_{33})$$

$$A_3 = -2\lambda C_{66}^0(C_{11}^0\bar{a}_{11} + C_{11}^0\bar{a}_{22} + C_{66}^0\bar{a}_{33}) + 2\lambda C_{s1}^{0\,2} C_{66}^0(p^2 + q^2) + C_{11}^0 C_{66}^{0\,2}\lambda^3$$

$$A_2 = C_{66}^0\bar{a}_{33}(\bar{a}_{11} + \bar{a}_{22}) + C_{11}^0\bar{a}_{11}\bar{a}_{22} - \lambda^2 C_{66}^0(C_{11}^0\bar{a}_{11} + C_{11}^0\bar{a}_{22} + C_{66}^0\bar{a}_{33}) + C_{s1}^{0\,2}p^2q^2(2C_{s1}^0 - C_{11}^0) + \lambda^2(p^2 + q^2)(C_{12}^0 C_{66}^{0\,2} + C_{12}^0 C_{s1}^0 C_{66}^0 + C_{s1}^0 C_{66}^{0\,2}) - C_{s1}^{0\,2}(\bar{a}_{11}q^2 + \bar{a}_{22}p^2)$$

$$A_1 = C_{66}^0\lambda\bar{a}_{33}(\bar{a}_{11} + \bar{a}_{22}) + C_{11}^0\lambda\bar{a}_{11}\bar{a}_{22} + C_{s1}^{0\,2}p^2q^2\lambda(2C_{s1}^0 - C_{11}^0) + \lambda^3(p^2 + q^2)C_{12}^0 C_{66}^{0\,2} - \lambda C_{s1}^{0\,2}(\bar{a}_{11}q^2 + \bar{a}_{22}p^2)$$

$$A_0 = 2C_{s1}^0 C_{66}^0 C_{12}^0\lambda^2 p^2 q^2 - C_{12}^0 C_{66}^0\lambda^2(\bar{a}_{11}q^2 + \bar{a}_{22}p^2) + \bar{a}_{33}(C_{s1}^{0\,2}p^2q^2 - \bar{a}_{11}\bar{a}_{22})$$

Denoting the roots of Eq. (18) by t_j ($j=1,2,3,\ldots,6$), the expression for the axial displacement u^1 at any point in the FG plate along x direction can be derived as

$$u^1 = \sum_{j=1}^{6} \gamma_j e^{t_j z} \cos px \sin qy \tag{19}$$

where γ_j ($j=1,2,3,\ldots,6$) are the unknown constants and can be determined by satisfying the prescribed boundary and interface conditions between the FG plate and the PFRC layer. Next, using Eqs. (16) and (17), the solutions for the displacements v^1 and w^1 in the y and z directions, respectively, can be obtained as

$$v^1 = \sum_{j=1}^{6} \gamma_j F_j^1 e^{t_j z} \sin px \cos qy \tag{20}$$

$$w^1 = \sum_{j=1}^{6} \gamma_j F_j^2 e^{t_j z} \sin px \sin qy \tag{21}$$

in which F_j^1 and F_j^2 are the values of F^1 and F^2 for $t=t_j$, respectively. Finally, using the constitutive relations given by Eq. (1) in conjunction with Eqs. (8), (19), and (21) the solutions for the stresses at any point in the substrate FG plate are derived as follows

$$\sigma_x^1 = \sum_{j=1}^{6} (-pC_{11}^0 - qC_{12}^0 F_j^1 + C_{13}^0 t_j F_j^2)\gamma_j e^{t_j z} \sin px \sin qy \tag{22}$$

$$\sigma_y^1 = \sum_{j=1}^{6} (-pC_{12}^0 - qC_{22}^0 F_j^1 + C_{23}^0 t_j F_j^2)\gamma_j e^{t_j z} \sin px \sin qy \tag{23}$$

$$\sigma_z^1 = \sum_{j=1}^{6} (-pC_{13}^0 - qC_{23}^0 F_j^1 + C_{33}^0 t_j F_j^2)\gamma_j e^{t_j z} \sin px \sin qy \tag{24}$$

$$\sigma_{yz}^1 = \sum_{j=1}^{6} C_{44}^0(t_j F_j^1 + qF_j^2)\gamma_j e^{t_j z} \sin px \cos qy \tag{25}$$

$$\sigma_{xz}^1 = \sum_{j=1}^{6} C_{55}^0(t_j + pF_j^2)\gamma_j e^{t_j z} \cos px \sin qy \tag{26}$$

$$\sigma_{xy}^1 = \sum_{j=1}^{6} C_{66}^0(q + pF_j^1)\gamma_j e^{t_j z} \cos px \cos qy \tag{27}$$

2.2 Solutions of the PFRC Layer.

In case of the PFRC layer, the functions $U^k(z)$, $V^k(z)$, $W^k(z)$, and $\Phi(z)$ appearing in the admissible functions given by Eq. (13) can be expressed as

$$[U^2(z)\ V^2(z)\ W^2(z)\ \Phi(z)] = [U^{02}\ V^{02}\ W^{02}\ \Phi^0]e^{rz} \tag{28}$$

where U^{02}, V^{02}, W^{02}, Φ^0, and r are the unknown constants to be determined. Using the governing equations (9) and (10), the constitutive relations (2) and (3), the strain-displacement relations [39], the electric field-potential relations [26], and Eqs. (13) and (28) another set of algebraic equations for the PFRC layer are obtained as follows

$$\begin{bmatrix} b_{11} & b_{12} & b_{13}r & b_{14}r \\ b_{12} & b_{22} & b_{23}r & 0 \\ -b_{13}r & -b_{23}r & b_{33} & 0 \\ -b_{14}r & 0 & 0 & b_{44} \end{bmatrix} \begin{Bmatrix} U^{02} \\ V^{02} \\ W^{02} \\ \Phi^0 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{29}$$

The various elements of the above matrix are given by

$$b_{11} = C_{55}^2 r^2 - \bar{b}_{11}, \quad b_{12} = -pq(C_{12}^2 + C_{66}^2), \quad b_{13} = p(C_{13}^2 + C_{55}^2),$$

$$b_{14} = pe_{31}, \quad b_{22} = C_{44}^2 r^2 - \bar{b}_{22}$$

$$b_{23} = q(C_{23}^2 + C_{44}^2), \quad b_{33} = C_{33}^2 r^2 - \bar{b}_{33}, \quad b_{44} = \bar{b}_{44} - \varepsilon_{33} r^2,$$

$$\bar{b}_{11} = C_{11}^2 p^2 + C_{66}^2 q^2$$

$$\bar{b}_{22} = C_{66}^2 p^2 + C_{22}^2 q^2, \quad \bar{b}_{33} = C_{55}^2 p^2 + C_{44}^2 q^2, \quad \bar{b}_{44} = \varepsilon_{11} p^2 + \varepsilon_{22} q^2$$

At this point, it should be noted that the useful and practical range of interest for the fiber volume fraction in a composite material is

0.3–0.7. Within this range of fiber volume fraction, the values of the effective piezoelectric constants e_{32}, e_{33}, e_{24}, and e_{15} of the piezoelectric fiber reinforced composite material are negligibly small [38,43] in comparison to the value of the effective piezoelectric constant e_{31}. Also, within the said useful range of fiber volume fraction, the effective value of the piezoelectric constant e_{31} of the PFRC material being considered here is much larger than that of the constituent fiber material. Hence, for the sake of simplicity in the analysis, the piezoelectric coefficients e_{32}, e_{33}, e_{24}, and e_{15} of the PFRC material are not considered for deriving the above noted algebraic equations given by Eq. (29). From Eq. (29), the unknown constants V^{02}, W^{02}, and Φ^0 can be expressed in terms of the unknown constant U^{02} as follows

$$V^{02} = F^3 U^{02}, \quad W^{02} = F^4 U^{02}, \quad \Phi^0 = F^5 U^{02} \tag{30}$$

in which the coefficients F^3, F^4, and F^5 are given by

$$F^3 = -\frac{b_{13}b_{23}r^2 + b_{12}b_{33}}{b_{22}b_{33} + (b_{23})^2 r^2}, \quad F^4 = \frac{r(b_{22}b_{13} - b_{12}b_{23})}{b_{22}b_{33} + (b_{23})^2 r^2}, \quad F^5 = \frac{b_{14}r}{b_{44}} \tag{31}$$

In order to find the nontrivial solution of U^{02}, substitution of Eqs. (30) and (31) into the first equation of (29) leads to the following characteristic equation

$$\sum_{n=1}^{5} B_n r^{2(n-1)} = 0 \tag{32}$$

where

$$B_5 = -C_{55}^2 C_{44}^2 C_{33}^2 \varepsilon_{33}, \quad B_4 = g_2\varepsilon_{33} + g_3 C_{33}^2 C_{44}^2,$$

$$B_3 = g_6\varepsilon_{33} + g_7 \quad B_2 = g_8\varepsilon_{33} + g_9$$

$$B_1 = -g_8\bar{b}_{44}, \quad g_1 = \bar{b}_{33}C_{44}^2 + \bar{b}_{22}C_{33}^2 - (b_{23})^2,$$

$$g_2 = g_1 C_{55}^2 + \bar{b}_{11}C_{33}^2 C_{44}^2 - (b_{13})^2 C_{44}^2$$

$$g_3 = \bar{b}_{44}C_{55}^2 + (b_{14})^2, \quad g_4 = b_{23}b_{13} + b_{12}C_{33}^2, \quad g_5 = \bar{b}_{22}b_{13} + b_{12}b_{23}$$

$$g_6 = g_4 b_{12} + g_5 b_{13} - g_1\bar{b}_{11} - \bar{b}_{22}\bar{b}_{33}C_{55}^2, \quad g_7 = (b_{13})^2\bar{b}_{44}C_{44}^2 - g_1 g_3 - \bar{b}_{11}\bar{b}_{44}C_{44}^2 C_{33}^2$$

$$g_8 = \bar{b}_{11}\bar{b}_{22}\bar{b}_{33} - \bar{b}_{33}(b_{12})^2,$$

$$g_9 = g_3\bar{b}_{22}\bar{b}_{33} + \bar{b}_{44}(g_1\bar{b}_{11} - g_4 b_{12} - g_5 b_{13})$$

For the PFRC layer considered here, the roots of Eq. (32) are four pairs of real quantities. Denoting these roots by $\pm r_j$ ($j=1$, 2, 3, and 4), the expressions for the axial displacements (u^2, v^2, and w^2) and the electric potential (ϕ) at any point in the PFRC layer can be derived as

$$u^2 = \sum_{j=1}^{4} \{\alpha_j \cosh(r_j z) + \beta_j \sinh(r_j z)\}\cos px \sin qy \tag{33}$$

$$v^2 = \sum_{j=1}^{4} \{\alpha_j \cosh(r_j z) + \beta_j \sinh(r_j z)\}F_j^3 \sin px \cos qy \tag{34}$$

$$w^2 = \sum_{j=1}^{4} \{\alpha_j \sinh(r_j z) + \beta_j \cosh(r_j z)\}F_j^4 \sin px \sin qy \tag{35}$$

$$\phi = \sum_{j=1}^{4} \{\alpha_j \sinh(r_j z) + \beta_j \cosh(r_j z)\}F_j^5 \sin px \sin qy \tag{36}$$

where α_j, and β_j ($j=1$, 2, 3, and 4) are the unknown constants to be determined from the prescribed boundary and interface continuity conditions; r_j is the magnitude of the jth pair of roots of the characteristic equation given by Eq. (32); F_j^3, F_j^4, and F_j^5 are the values of the constants F^3, F^4, and F^5 when $r=r_j$. Finally, using Eqs. (33)–(36) and the constitutive relations given by Eqs. (2) and (3), the expressions for the stresses in the PFRC layer which exactly satisfy the governing equilibrium equations (9) and (10) can be derived and are given by

$$\sigma_x^2 = \sum_{j=1}^{4} (-pC_{11}^2 - qC_{12}^2 F_j^3 + C_{13}^2 F_j^4 r_j + e_{31}F_j^5 r_j)\{\alpha_j \cosh(r_j z) + \beta_j \sinh(r_j z)\}\sin px \sin qy \tag{37}$$

$$\sigma_y^2 = \sum_{j=1}^{4} (-pC_{12}^2 - qC_{22}^2 F_j^3 + C_{23}^2 F_j^4 r_j)\{\alpha_j \cosh(r_j z) + \beta_j \sinh(r_j z)\}\sin px \sin qy \tag{38}$$

$$\sigma_z^2 = \sum_{j=1}^{4} (-pC_{13}^2 - qC_{23}^2 F_j^3 + C_{33}^2 F_j^4 r_j)\{\alpha_j \cosh(r_j z) + \beta_j \sinh(r_j z)\}\sin px \sin qy \tag{39}$$

$$\sigma_{yz}^2 = \sum_{j=1}^{4} C_{44}^2 (F_j^3 r_j + qF_j^4)\{\alpha_j \sinh(r_j z) + \beta_j \cosh(r_j z)\}\sin px \cos qy \tag{40}$$

$$\sigma_{xz}^2 = \sum_{j=1}^{4} C_{55}^2 (r_j + pF_j^4)\{\alpha_j \sinh(r_j z) + \beta_j \cosh(r_j z)\}\cos px \sin qy \tag{41}$$

$$\sigma_{xy}^2 = \sum_{j=1}^{4} C_{66}^2 (pF_j^3 + q)\{\alpha_j \cosh(r_j z) + \beta_j \sinh(r_j z)\}\cos px \cos qy \tag{42}$$

3 Numerical Results

In order to assess the performance of the PFRC layer as a distributed actuator for the FG plate, numerical values of the exact solutions derived in the previous section are to be evaluated. It should be noted that the expressions for the solutions of the substrate FG plate contain six unknowns and those for the PFRC layer contain eight unknowns. Thus, all together 14 unknown constants are required to be evaluated. This can be accomplished by satisfying the prescribed boundary conditions as well as the interface continuity conditions as described in the following.

On the exposed surface of the PFRC layer, the following prescribed boundary conditions are considered

$$\sigma_z^2(x,y,z) = f_0 \sin px \sin qy, \quad \sigma_{xz}^2(x,y,z) = 0, \quad \sigma_{yz}^2(x,y,z) = 0$$

$$\phi(x,y,z) = V \sin px \sin qy \tag{43}$$

where $f_0(\mathrm{N/m^2})$ and $V(\mathrm{V})$ are the amplitudes of externally applied sinusoidal surface traction and electric potential functions, respectively. Note that the value of z coordinate to be used in the above-prescribed boundary conditions should be either $h+h_p$ or $-h_p$ according to whether the PFRC layer is bonded to the top or bottom surface of the substrate FG plate, respectively.

Table 1 Comparison of the center deflection and normal stress of the functionally graded plate ($s=100$) with that of the homogeneous isotropic plate ($E=E_0$) when the thickness of the PFRC layer is negligible

Functionally graded plate					
$\lambda=\ln(E_h/E_0)/h$	$\bar{w}$	$\bar{\sigma}_x$ $(a/2,b/2,h)$	$\lambda=\ln(E_h/E_0)/h$	$\bar{w}$	$\bar{\sigma}_x$ $(a/2,b/2,h)$
2	−1.9905	−0.2484	0.5	−3.9809	−0.1564
1.5	−2.2926	−0.2260	0.7	−3.3549	−0.1753
1.1	−2.6737	−0.2040	0.9	−2.9560	−0.1908
1.01	−2.7901	−0.1983	0.99	−2.8181	−0.1969
1.001	−2.8026	−0.1977	0.999	−2.8054	−0.1975
1.0001	−2.8039	−0.1976	0.9999	−2.8042	−0.1976
Homogenous isotropic plate[a]					
N.A.	−2.8040	−0.1976	N.A	−2.8040	−0.1976

[a] Reference [32].

The surface of the PFRC layer being attached to the substrate FG plate is grounded such that electric potential at this surface becomes zero [20,26,42], i.e.

$$\phi(x,y,z)=0 \tag{44}$$

At the interface between the FG plate and the PFRC layer, the continuity conditions for displacements and interlaminar stresses must be satisfied. These conditions are provided by

$$u^1(x,y,z)=u^2(x,y,z), \quad v^1(x,y,z)=v^2(x,y,z)$$

$$w^1(x,y,z)=w^2(x,y,z), \quad \sigma_z^1(x,y,z)=\sigma_z^2(x,y,z)$$

$$\sigma_{yz}^1(x,y,z)=\sigma_{yz}^2(x,y,z), \quad \sigma_{xz}^1(x,y,z)=\sigma_{xz}^2(x,y,z) \tag{45}$$

At the exposed surface of the substrate FG plate the following natural boundary conditions must be satisfied

$$\sigma_z^1(x,y,z)=0, \quad \sigma_{xz}^1(x,y,z)=0 \quad \sigma_{yz}^1(x,y,z)=0 \tag{46}$$

In Eqs. (44) and (45), the values of z should be considered as h or 0, according to whether the top or bottom surface of the substrate FG plate is integrated with the PFRC layer, respectively, while in Eq. (46) the corresponding values of z will be 0 or h, respectively. It is now obvious that the satisfaction of the boundary and the interface continuity conditions, given by Eqs. (45)–(47), results into 14 simultaneous algebraic equations with same number of unknown constants and thus these 14 unknown constants can be uniquely determined yielding the numerical values of the exact solutions for displacements and stresses in the overall smart FG plate.

The following material properties for the isotropic material located at the bottom surface of the substrate FG plate are considered for the numerical results

$$E_0=200\text{ GPa} \quad \nu=0.3$$

in which ν is the Poison's ratio of the FGM. The thickness of the FG plate is considered as 3 mm. The piezoelectric fiber and the matrix of the PFRC layer are made of PZT5H and epoxy, respectively. Considering 40% fiber volume fraction, the following elastic and piezoelectric co-efficients of the PFRC layer are obtained by using the micromechanics model derived earlier by one of the authors [38] of this paper and are used for computing the numerical results

$$C_{11}=32.6\text{ GPa}, \quad C_{12}=4.3\text{ GPa}, \quad C_{13}=4.76\text{ GPa},$$

$$C_{22}=C_{33}=7.2\text{ GPa}$$

$$C_{23}=3.85\text{ GPa}, \quad C_{44}=1.05\text{ GPa}, \quad C_{55}=C_{66}=1.29\text{ GPa},$$

$$e_{31}=-6.76\text{ C/m}^2$$

$$\varepsilon_{11}=\varepsilon_{22}=0.037e-9\text{ C/V m} \quad \varepsilon_{33}=10.64e-9\text{ C/V m}$$

Considering $m=n=1$ in the definition of p and q, the numerical results are evaluated with and without applying the electrical potential distribution on the actuator surface for different values of length to thickness ratios $s(=a/h)$ of the substrate FG plates. The following non-dimensional parameters are used for presenting the numerical results

$$(\bar{\sigma}_x,\bar{\sigma}_y,\bar{\sigma}_{xy})=\frac{1}{f_0s^2}(\sigma_x,\sigma_y,\sigma_{xy}), \quad (\bar{\sigma}_{xz},\bar{\sigma}_{yz})=\frac{1}{f_0s}(\sigma_{xz},\sigma_{yz}),$$

$$(\bar{\sigma}_z)=\frac{1}{f_0}(\sigma_z) \tag{47}$$

$$\bar{u}=\frac{E_0}{f_0s^3h}u, \quad \bar{w}=\frac{100E_0}{f_0s^4h}w$$

It should be noted here that if the value of the inhomogeneity parameter (λ) tends to zero then the FG plate turns out to be a homogeneous isotropic plate. It is obvious from Eq. (7) that if $E_h/E_0 \to 1$ then the inhomogeneity parameter tends to zero with E_h being the Young's modulus of the material located at the top surface of the FG plate. Hence, considering that the PFRC layer is of negligible thickness and not subjected to the applied voltage, the center deflection of the FG plate has been computed using the expressions for the exact solutions derived in the previous section when $\lambda \to 0$. These are compared with the existing exact solutions of homogeneous isotropic plate [39] with Young's modulus E_0 and Poisson's ratio ν as shown in Table 1. It may be observed from this table that the results excellently converge to that of the homogeneous isotropic plate when $\lambda \to 0$. This comparison may be considered as a method for validating the present exact solutions. Next, assuming the thickness of the PFRC layer as 250 μm, and the amplitude of distributed sinusoidal mechanical load as $f_0=40\text{ N/m}^2$ (downward), the mechanical displacements and stresses in the substrate FG plate with large gradient ($E_h/E_0=10$) of properties are computed with and without applying the externally applied voltage to the PFRC layer being bonded to the top surface of the plate. These are presented in Table 2 for different aspect ratios (s). The results presented in Table 2 clearly reveal that when the PFRC layer is subjected to externally applied positive voltage, it counteracts the deformations caused by the vertically applied downward mechanical load while for a negative voltage the PFRC layer increases the deformation caused by the

Table 2 Responses of the FG substrate plates ($E_h/E_0=10$) with and without the applied voltage to the PFRC layer bonded to the top surface of the plates

s	V (V)	$\bar{u}$ $(0,b/2,0)$ $(0,b/2,h)$	$\bar{w}$ $(a/2,b/2,h/2)$	$\bar{\sigma}_x$ $(a/2,b/2,0)$ $(a/2,b/2,h)$	$\bar{\sigma}_y$ $(a/2,b/2,0)$ $(a/2,b/2,h)$	$\bar{\sigma}_z$ $(a/2,b/2,h/2)$	$\bar{\sigma}_{xy}$ (0,0,0) $(0,0,h)$	$\bar{\sigma}_{yz}$ $(a/2,0,h/2)$	$\bar{\sigma}_{xz}$ $(0,b/2,h/2)$
10	0	−0.0194 0.0093	−0.9553	0.0871 −0.4201	0.0870 −0.4213	−0.3537	−0.0469 0.2242	−0.2215	−0.2218
	100	−0.0307 −5.9373	186.8222	−5.8052 203.9846	−19.6720 57.9752	43.9487	6.8593 −70.4771	43.9606	2.7211
	−100	−0.0081 5.9558	−188.7329	5.9794 −204.8249	19.8460 −58.8178	−44.6562	−6.9530 70.9256	−44.4035	−3.1646
20	0	−0.0195 0.0093	−0.9252	0.0873 −0.4171	0.0872 −0.4180	−0.3554	−0.0470 0.2243	−0.2223	−0.2225
	100	−0.0337 −1.4670	45.5393	−1.3742 50.3911	−4.9333 14.3338	10.8462	1.6982 −17.4232	10.9839	0.4628
	−100	−0.0053 1.4855	−47.3897	1.5489 −51.2253	5.1078 −15.1698	−11.5570	−1.7922 17.8717	−11.4284	−0.9079
100	0	−0.0195 0.0093	−0.9155	0.0874 −0.4161	0.0873 −0.4170	−0.3559	−0.0470 0.2243	−0.2225	−0.2228
	100	−0.0202 −0.0497	0.9368	0.0291 1.6124	−0.1145 0.1751	0.0938	0.0230 −0.4813	0.2278	−0.1960
	−100	−0.0188 0.0682	−2.7678	0.1457 −2.4446	0.2892 −1.0090	−0.8056	−0.1171 0.9298	−0.6729	−0.2495

downward mechanical load. Thus the PFRC layer can act as a distributed actuator for smart functionally graded plates. It may be noted from the constitutive equations of the piezoelectric material given by Eq. (2) that the magnitude of the induced normal stress depends on the magnitude of the applied voltage and the piezoelectric coefficients while the stresses and deformations in the substrate as well as in the PFRC layer due to mechanical load only depend on the aspect ratio of the substrate, material properties of the substrate layers, and the intensity of the applied mechanical load. It is known that if the aspect ratio decreases, i.e., the substrate plate becomes thick, the stresses and deflection of the substrate plate decrease [39]. Hence, for a particular applied voltage and intensity of applied mechanical load, the PFRC layer has to counteract the less amount of deformation in case of thick substrate ($s=10$) than in case of thin substrate ($s=100$) as can be observed from Table 2. Thus the PFRC layer becomes more effective in the case of thick plate FG substrate than the thin one. Note that similar results were found for laminated composite substrate [26]. Table 3 illustrates the behavior of the FG plate for which $E_h/E_0=0.1$ under the combined action of mechanical load and the activated PFRC layer being bonded to the top surface of the plate. It may be observed from this table that as the maximum value of E for this plate is much lower than that of the plate with $E_h/E_0=10$, the deformations of this plate with and without the applied voltage is much larger than that in case of the FG plate with $E_h/E_0=10$. The results presented in Table 4 illustrate the comparison of the performance of the PFRC layer when it is at-

Table 3 Responses of the FG substrate plate ($E_h/E_0=0.1$) with and without the applied voltage to the PFRC layer bonded to the top surface of the plates

s	V (V)	$\bar{u}$ $(0,b/2,0)$ $(0,b/2,h)$	$\bar{w}$ $(a/2,b/2,h/2)$	$\bar{\sigma}_x$ $(a/2,b/2,0)$ $(a/2,b/2,h)$	$\bar{\sigma}_y$ $(a/2,b/2,0)$ $(a/2,b/2,h)$	$\bar{\sigma}_z$ $(a/2,b/2,h/2)$	$\bar{\sigma}_{xy}$ (0,0,0) $(0,0,h)$	$\bar{\sigma}_{yz}$ $(a/2,0,h/2)$	$\bar{\sigma}_{xz}$ $(0,b/2,h/2)$
10	0	−0.0914 0.1686	−8.7397	0.4042 −0.0807	0.3900 −0.0825	−0.6177	−0.2138 0.0416	−0.2114	−0.2248
	100	−3.1845 −114.0117	3253.3	−41.0915 44.3075	−170.3199 28.1032	115.6052	56.9184 −19.4500	79.9749	−41.5579
	−100	3.0017 114.3489	−3270.8	41.8999 −44.4689	171.0999 −28.2681	−116.8407	−57.3461 19.5333	−80.3977	41.1083
20	0	−0.0912 0.1735	−8.5437	0.4031 −0.0796	0.3883 −0.0812	−0.6184	−0.2131 0.0427	−0.2117	−0.2255
	100	−0.9839 −27.0971	813.7105	−9.8557 10.6491	−43.1562 7.1105	28.8665	14.2724 −4.7790	20.2434	−10.8859
	−100	0.8015 27.4441	−830.7979	10.6619 −10.8083	43.9327 −7.2729	−30.1034	−4.6986 4.8645	−20.6668	10.4348
100	0	−0.0912 0.1751	−8.4806	0.4027 −0.0793	0.3877 −0.0808	−0.6186	−0.2128 0.0431	−0.2118	−0.2258
	100	−0.1285 −0.8991	24.4974	−0.0061 0.3451	−1.3654 0.2088	0.5663	0.3692 −0.1491	0.6118	−0.6565
	−100	−0.0539 1.2493	−41.4587	0.8115 −0.5036	2.1408 −0.3704	−1.8036	−0.7948 0.2353	−10.0353	0.2049

Table 4 Comparison of effectiveness of the PFRC layer when placed at the top and bottom surfaces of the FG plate

			PFRC layer on the top of the FG plate			PFRC layer on the bottom of the FG plate		
E_h/E_0	a/h		$V=0$	$V=100$	$V=-100$	$V=0$	$V=100$	$V=-100$
10	10	$\bar{w}$ $(a/2,b/2,h/2)$	−0.9553	186.8222	−188.7329	−0.9532	359.0874	−360.9938
	100	$\bar{w}$ $(a/2,b/2,h/2)$	−0.9155	0.9368	−2.7678	−0.9100	2.7147	−4.5346
0.1	10	$\bar{w}$ $(a/2,b/2,h/2)$	−8.7397	3253.3	−3270.8	−9.2761	1789.0	−1807.5
	100	$\bar{w}$ $(a/2,b/2,h/2)$	−8.4806	24.4974	−41.4587	−8.9838	8.7604	−26.7281

tached either on the top or at the bottom surface of the FG plate. It may be observed from this table that for both thick and thin FG plates, the activated PFRC layer causes maximum change in center deflection with respect to the center deflection for $V=0$ when it is placed on that surface of the FG plate which has the minimum value of Young's modulus (E). Note that this maximum change in center deflection is a measure of control authority of the PFRC layer. Therefore, as a benchmark result, it may be concluded that if the PFRC layer is attached with the surface of the FG plate where the stiffness of the material is minimum then the activated PFRC layer becomes most effective in controlling the elastic deformation of the FG plate. The through-thickness behavior of thin FG substrate plates due to the action of the activated PFRC layer being attached at the top surface of the substrates has been investigated and are presented here for a thin FG plate with $E_h/E_0=10$. Figures 2 and 3 illustrate the variation of axial ($\bar{u}$) and transverse ($\bar{w}$) displacements across the thickness of the substrate FG plate. It may be observed from these figures that similar to the homogeneous isotropic plate, as the substrate is thin, $\bar{u}$ varies linearly across the thickness of the substrate while the transverse displacement is constant through the thickness. But, unlike the homogeneous isotropic plate, the axial deformation ($\bar{u}$) of the FG substrate across its thickness indicates that the coupling of bend-

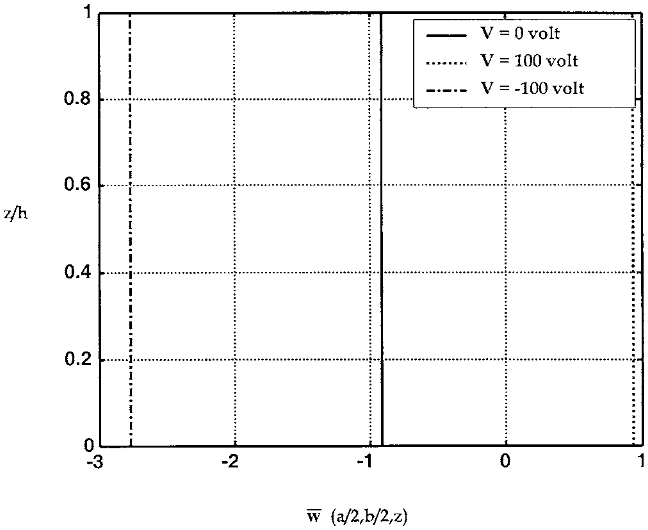


Fig. 3 **Distribution of transverse displacement ($\bar{w}$) across the thickness of thin ($s=100$) FG plate ($E_h/E_0=10$) with and without applied voltage to the PFRC layer**

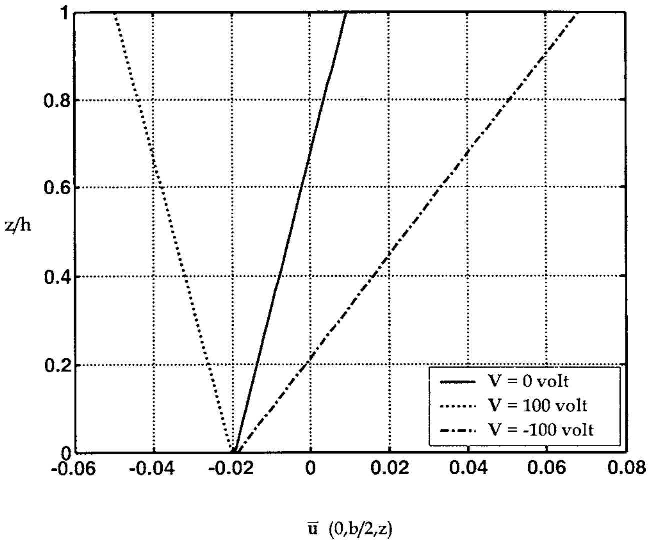


Fig. 2 **Distribution of axial displacement ($\bar{u}$) across the thickness of thin ($s=100$) FG plate ($E_h/E_0=10$) with and without applied voltage to the PFRC layer**

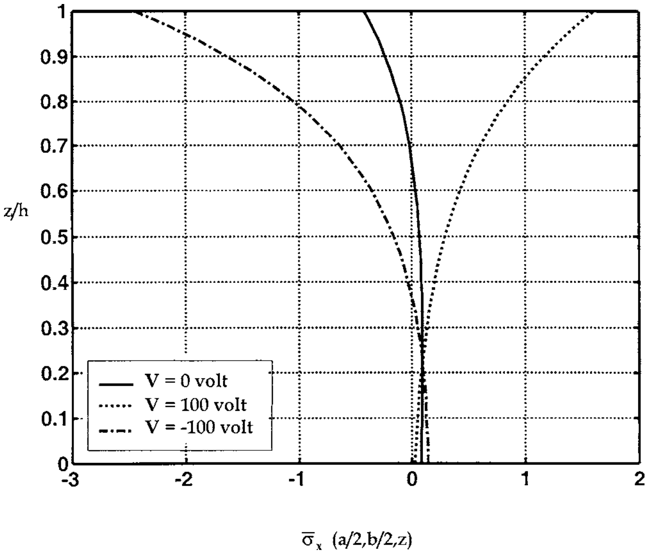


Fig. 4 **Distribution of inplane normal stress ($\bar{\sigma}_x$) across the thickness of thin ($s=100$) FG plate ($E_h/E_0=10$) with and without applied voltage to the PFRC layer**

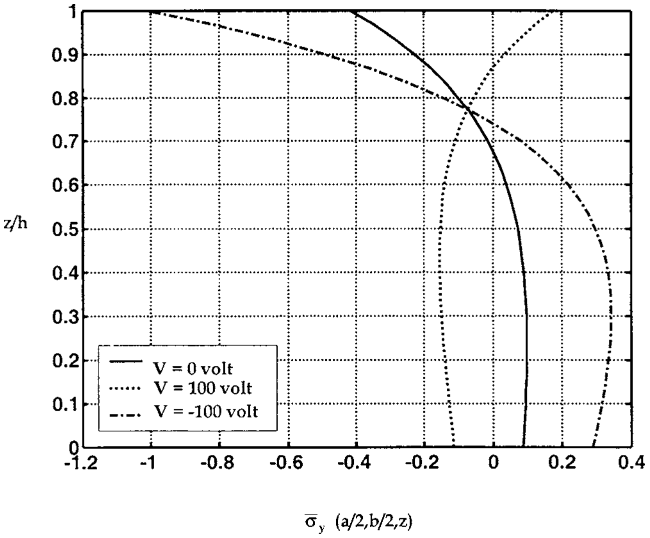


Fig. 5 Distribution of in-plane normal stress ($\bar{\sigma}_y$) across the thickness of thin ($s=100$) FG plate ($E_h/E_0=10$) with and without applied voltage to the PFRC layer

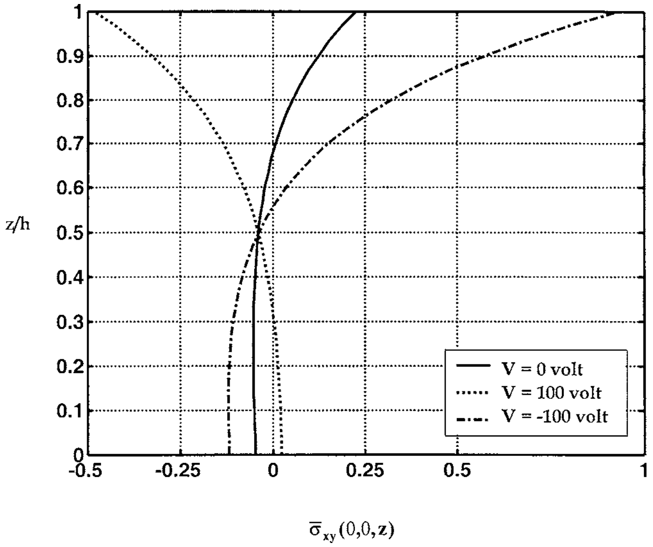


Fig. 7 Distribution of in-plane shear stress ($\bar{\sigma}_{xy}$) across the thickness of thin ($s=100$) FG plate ($E_h/E_0=10$) with and without applied voltage to the PFRC layer

ing and extension takes place in the FG plate even if the PFRC layer is not subjected to any voltage. This may be attributed to the fact that the material properties of the substrate plate are graded along the thickness direction. When the PFRC layer is subjected to a positive applied voltage, both extension and bending of the plate occur in unison as opposed to that caused by the mechanical load. Also, it can be observed from Figs. 2 and 3 that if the polarity of the voltage changes, the PFRC layer causes reversal of the axial and transverse displacements of the FG substrate. The variations of the axial normal stresses (σ_x, σ_y, and σ_z), in-plane shear stress (σ_{xy}), and transverse shear stresses (σ_{xz},σ_{yz}) across the thickness of this plate have been plotted in Figs. 4–9, respectively. Since the Young's modulus of the substrate FG plate is having nonzero gradient with respect to the thickness coordinate, the stresses are nonlinear across the thickness of the plate. It can also be observed from Figs. 4, 5, 7, and 8 that the axial normal stresses (σ_x,σ_y) and the shear stresses (σ_{xy},σ_{xz}) become maximum at the interface between the PFRC layer and the FG plate. Although it may be expected that the interlaminar stress (σ_{xz}) will be higher at the top of the FG plate because of the transfer of induced stress from the activated PFRC layer, large interlaminar shear stress may cause detrimental effect such as delamination. As the mechanical load acts along the vertically downward direction, the transverse normal stress (σ_z) is negative across the thickness of the FG plate. When the PFRC layer is subjected to positive voltage of 100 V, the transverse normal stress (σ_z) becomes positive across a part of the thickness of this FG plate while it is negative across the remaining part of the thickness and the maximum value of the negative part is less than that due to mechanical load only as shown in Fig. 6. When the PFRC layer is subjected to a negative voltage of magnitude 100 V, the transverse normal stress (σ_z) increases throughout the thickness of the FG plate over its mechanical counterpart.

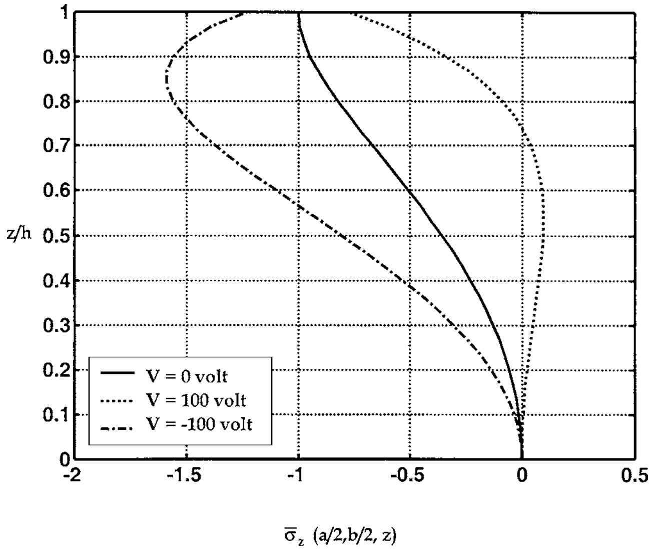


Fig. 6 Distribution of transverse normal stress ($\bar{\sigma}_z$) across the thickness of thin ($s=100$) FG plate ($E_h/E_0=10$) with and without applied voltage to the PFRC layer

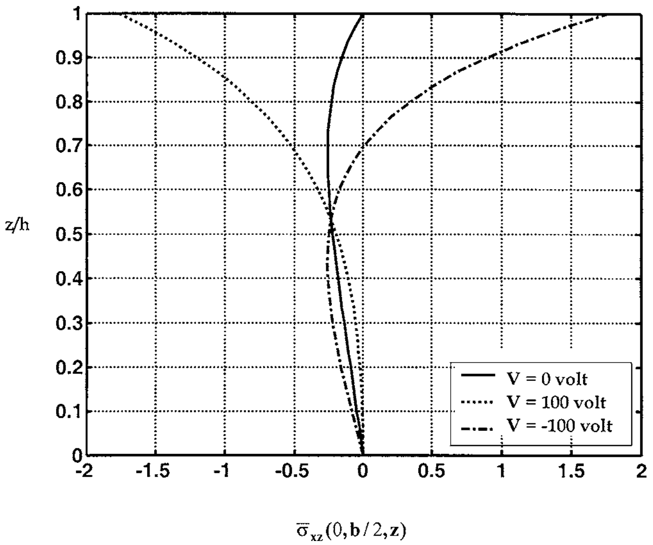


Fig. 8 Distribution of transverse shear stress ($\bar{\sigma}_{xz}$) across the thickness of thin ($s=100$) FG plate ($E_h/E_0=10$) with and without applied voltage to the PFRC layer

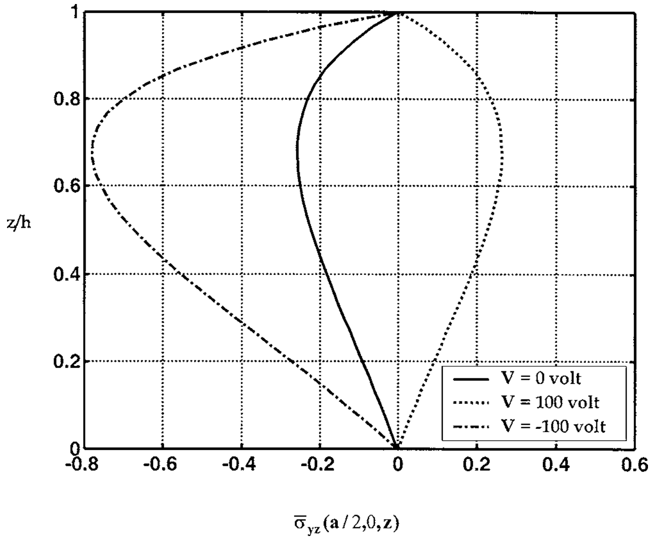


Fig. 9 Distribution of transverse shear stress ($\bar{\sigma}_{yz}$) across the thickness of thin (s=100) FG plate (E_h/E_0=10) with and without applied voltage to the PFRC layer

4 Conclusions

This paper deals with the exact solutions for static analysis of functionally graded plates integrated with a layer of PFRC material which acts as the distributed actuator for the plates. The PFRC material considered here is a new smart material with enhanced effective piezoelectric coefficient as compared to its constituent monolithic counterpart and the piezoelectric fibers are oriented longitudinally along the length of the substrate FG plates. When the PFRC layer is subjected to a positive voltage, it counteracts the deformations caused by the vertically downward mechanical load and vice versa. The activated PFRC layer significantly affects the distribution of all in-plane and transverse stresses across the thickness of the FG substrate plate and is efficiently able to actuate both thick and thin plates. The through-thickness behavior of the FG plate indicates that similar to homogeneous isotropic plates axial displacements are linear across the thickness of the plate while the transverse displacement is constant across the thickness. But in contrast to the homogeneous isotropic plates, coupling of bending and extensional deformations occurs in the FG plate even if the PFRC layer is not subjected to any voltage. Also, unlike the thin homogeneous isotropic plates, the axial normal stresses and the in-plane shear stress are nonlinear across the thickness of the thin substrate FG plate. The numerical values of the exact solutions also revealed that the performance of the activated PFRC layer becomes maximum when the PFRC layer is attached with the softest surface of the substrate FG plate. The benchmark results presented in this paper may be useful for developing new functional smart structures and may serve the purpose of verifying the numerical models of functionally graded smart structures for which exact solutions are not possible.

References

[1] Koizumi, M., 1993, "Concept of FGM," Ceram. Trans., **34**, pp. 3–10.
[2] Suresh, S., and Mortensen, A., 1998, *Fundamentals of Functionally Graded Materials*, Maney, London.
[3] Noda, N., and Jin, Z. H., 1993, "Thermal Stress Intensity Factors for a Crack in a Strip of a Functionally Gradient Material," Int. J. Solids Struct., **30**, pp. 1039–1056.
[4] Teymur, M., Chitkara, N. R., Yohngjo, K., Aboudi, J., Pindera, M. J., and Arnold, S. M., 1996, "Thermoelastic Theory for the Response of Materials Functionally Graded in Two Directions," **33**, pp. 931–966.
[5] Feldman, E., and Aboudi, J., 1997, "Buckling Analysis of Functionally Graded Plates Subjected to Uniaxial Loading," Compos. Struct., **38**, pp. 29–36.
[6] Gu, P., and Asaro, J., 1997, "Crack in Functionally Graded Materials," Int. J. Solids Struct., **34**, pp. 1–17.
[7] Mian, A. M., and Spencer, A. J. M., 1998, "Exact Solutions for Functionally Graded and Laminated Elastic Materials," J. Mech. Phys. Solids, **46**, pp. 2283–2295.
[8] Praveen, G. N., and Reddy, J. N., 1998, "Nonlinear Transient Thermoelastic Analysis of Functionally Graded Ceramic Metal Plates," Int. J. Solids Struct., **35**, pp. 4457–4476.
[9] Lee, P. C. Y., and Yu, J. D., 1998, "Governing Equations for a Piezoelectric Plate With Graded Properties Across the Thickness," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **45**, pp. 236–250.
[10] Loy, C. T., Lam, K. Y., and Reddy, J. N., 1999, "Vibration of Functionally Graded Cylindrical Shells," Int. J. Mech. Sci., **41**, pp. 309–324.
[11] Aboudi, J., Pindera, M. J., and Arnold, S. M., 1999, "Higher-Order Theory for Functionally Graded Materials," Composites, Part B, **30**, pp. 777–832.
[12] Wang, B. L., Han, J. C., and Du, S. Y., 2000, "Crack Problems for Functionally Graded Materials Under Transient Thermal Loading," J. Polym. Sci. A, **23**, pp. 143–168.
[13] Yang, Y. Y., 2000, "Time-Dependent Stress Analysis in Functionally Graded Materials," Int. J. Solids Struct., **37**, pp. 7593–7608.
[14] Yang, J., and Shen, H. S., 2001, "Dynamic Response of Initially Stressed Functionally Graded Rectangular Thin Plates," Compos. Struct., **54**, pp. 497–508.
[15] Sankar, B. V., 2001, "An Elasticity Solution for Functionally Graded Beams," Compos. Sci. Technol., **61**, pp. 689–696.
[16] Batra, R. C., and Vel, S. S., 2001, "Exact Solution for Thermoelastic Deformations of Functionally Graded Thick Rectangular Plates," AIAA J., **40**, pp. 1421–1433.
[17] Woo, J., and Meguid, S. A., 2001, "Nonlinear Analysis of Functionally Graded Plates and Shallow Shells," Int. J. Solids Struct., **38**, pp. 7409–7421.
[18] Shen, H. S., 2002, "Nonlinear Bending Response of Functionally Graded Plates Subjected to Transverse Loads and in Thermal Environments," Int. J. Mech. Sci., **44**, pp. 561–584.
[19] Vel, S. S., and Batra, R. C., 2003, "Three-Dimensional Analysis of Transient Thermal Stresses in Functionally Graded Plates," Int. J. Solids Struct., **40**, pp. 7181–7196.
[20] Zhong, Z., and Shang, E. T., 2003, "Three-Dimensional Exact Analysis of a Simply Supported Functionally Gradient Piezoelectric Plate," Int. J. Solids Struct., **40**, pp. 5335–5352.
[21] Najafizadeh, M. M., and Hedayati, B., 2004, "Refined Theory for Thermoelastic Stability of Functionally Graded Circular Plates," J. Therm. Stresses, **27**, pp. 857–880.
[22] Bailey, T., and Hubbard, J. E., 1985, "Distributed Piezoelectric Polymer Active Vibration Control of a Cantilever Beam," J. Guid. Control Dyn., **8**, pp. 605–611.
[23] Miller, S. E., and Hubbard, J. E., 1987, "Observability of a Bernoulli-Euler Beam Using, PVF$_2$, as a Distributed Sensor," MIT Draper Laboratory Report, July 1, 1987.
[24] Crawley, E. F., and Luis, J. D., 1987, "Use of Piezoelectric Actuators as Elements of Intelligent Structures," AIAA J., **25**, pp. 1373–1385.
[25] Baz, A., and Poh, S., 1988, "Performance of an Active Control System With Piezoelectric Actuators," J. Sound Vib., **126**, pp. 327–343.
[26] Ray, M. C., Bhattacharyya, R., and Samanta, B., 1993, "Exact Solutions for Static Analysis of Intelligent Structures," AIAA J., **31**, pp. 1684–1691.
[27] Clark, R. L., Fleming, M. R., and Fuller, C. R., 1997, "Piezoelectric Actuators for Distributed Excitation of Thin Plates; A Comparison Between Theory and Experiment," ASME J. Vibr. Acoust., **115**, pp. 332–339.
[28] Baz, A., and Poh, S., 1996, "Optimal Vibration Control With Modal Positive Position Feedback," Opt. Control Appl. Methods, **17**, pp. 141–149.
[29] Dong, S., and Tong, L., 2001, "Vibration Control of Plates Using Discretely Distributed Piezoelectric Quasi-Modal Actuators/Sensors," AIAA J., **39**, pp. 1766–1772.
[30] Ray, M. C., 2003, "Optimal Control of Laminated Shells With Piezoelectric Sensor and Actuator Layers," AIAA J., **41**, pp. 1151–1157.
[31] Ray, M. C., and Reddy, J. N., 2004, "Optimal Control of Thin Circular Cylindrical Shells Using Active Constrained Layer Damping Treatment, Smart Materials and Structures," Smart Mater. Struct., **13**, pp. 64–72.
[32] Ootao, Y., and Tanigawa, Y., 2000, "Three-Dimensional Transient Thermoelasticity in Functionally Graded Rectangular Plate Bonded to a Piezoelectric Plate," Int. J. Solids Struct., **37**, pp. 4377–4401.
[33] Ootao, Y., and Tanigawa, Y., 2001, "Control of the Transient Thermoelastic Displacement of a Functionally Graded Rectangular Plate Bonded to a Piezoelectric Plate due to Nonuniform Heating," Acta Mech., **148**, pp. 17–33.
[34] Reddy, J. N., and Cheng, Z. Q., 2001, "Three-Dimensional Solutions of Smart Functionally Graded Plates," ASME J. Appl. Mech., **68**, pp. 234–241.
[35] Wang, B. L., and Noda, N., 2001, "Design of a Smart Functionally Graded Thermopiezoelectric Composite Structure," Smart Mater. Struct., **10**, pp. 189–193.
[36] Wang, B. L., and Noda, N., 2001, "Thermally Induced Fracture of a Smart

Functionally Graded Composite Structure," Theor. Appl. Fract. Mech., **35**, pp. 93–109.

[37] Yang, J., Kitipornchi, S., and Liew, K. M., 2003, "Large Amplitude Vibration of Thermoelectro-Mechanically Stressed FGM Laminated Plates," **192**, pp. 3861–3885.

[38] Mallik, N., and Ray, M. C., 2003, "Effective Coefficients of Piezoelectric Fiber Reinforced Composites," AIAA J., **41**, pp. 704–710.

[39] Pagano, N. J., 1970, "Exact Solution for Rectangular Bidirectional Composites and Sandwich Plates," J. Compos. Mater., **4**, pp. 20–34.

[40] Erdogan, F., 1985, "The Crack Problem for Bonded Nonhomogeneous Materials Under Antiplane Shear Loading," ASME J. Appl. Mech., **52**, pp. 823–828.

[41] Delale, F., and Erdogan, F., 1988, "On the Mechanical Modeling of the Interfacial Region in Bonded Half Planes," ASME J. Appl. Mech., **55**, pp. 317–324.

[42] Heyliger, P., 1997, "Exact Solutions for Simply-Supported Laminated Piezoelectric Plates," ASME J. Appl. Mech., **64**, pp. 299–306.

[43] Tan, P., and Tong, L., 2001, "Micro-Electromechanics Models for Piezoelectric Fiber Reinforced Composite Materials," Compos. Sci. Technol., **61**, pp. 759–769.

On the Crashworthiness of Shear-Rigid Sandwich Structures

Dirk Mohr
Tomasz Wierzbicki
Impact and Crashworthiness Laboratory,
Massachusetts Institute of Technology,
Cambridge, MA 02139

This paper deals with the evaluation of the crashworthiness of thin-walled sandwich box structures for automotive applications. Quasi-static crushing simulations are carried out to estimate the energy absorption of prismatic box columns made from sandwich sheets. The sandwich sheets have perforated cores of different densities with staggered holes perpendicular to the panel faces. It is found that the specific energy absorption of columns made of sandwich sheets is approximately the same as that of conventional columns composed of homogeneous sheets of the same total wall thickness. Furthermore, theoretical analysis indicates that by increasing the core thickness, sandwich structures could be up to 50% lighter while providing the same mean crushing force. However, these gains may not be achieved in practical applications since increasing the core thickness also increases the likelihood of premature face sheet fracture during crushing.

[DOI: 10.1115/1.2165232]

1 Introduction

Various technologies have been proposed in the past to design lightweight and crashworthy automotive body structures. Concepts include the use of fiber reinforced plastics (e.g., Jacob et al. [1]), crash-optimized cross-sections (Kim and Wierzbicki [2]), foam fillings (e.g., Reid et al. [3], Seitzberger et al. [4], Santosa et al. [5]), and sandwich construction. The latter concept has been successfully employed in various transportation vehicles, notably in aircrafts and high-speed trains, but sandwich technology is seldom used in passenger cars. Sandwich panels exhibit an exceptionally high flexural stiffness per unit weight, but in automotive applications, their crushing performance must be considered as well. In passenger cars, a substantial amount of the frontal impact energy is absorbed by deforming the longitudinal rail which is typically designed as a thin-walled box structure. When subject to axial loading, thin-walled box structures respond under the formation of folds (Fig. 1). As illustrated in Fig. 2 and discussed by Mohr and Wierzbicki [7], the folding of a box structure made from sandwich sheets is strongly influenced by the shear strength of the core material. Figures 2(*b*) and 2(*c*) each show a longitudinal cut through a folding sandwich wall (see dashed lines in Fig. 2(*a*)) for so-called "shear-soft" and "shear-rigid" core materials, respectively. Throughout the folding, the core material is subject to out-of-plane shear stresses, τ_{TX} (where the x-axis is aligned with the column axis, and the T-axis corresponds to the thickness direction, see coordinate system in Fig. 2(*b*). If the shear stresses exceed the core yield strength, $|\tau_{TX}| > s_{TX}$, the sandwich cross section experiences large shear deformation, which is referred to as shear-soft behavior. Conversely, the folding mode is referred to as "shear rigid" when the shear strength is sufficiently large, i.e. $|\tau_{TX}| < s_{TX}$, and the shear deformation in the core material is small. In the shear-rigid case, initially perpendicular cross sections remain perpendicular which requires substantial stretching and compressing of the respective sandwich face sheets (Fig. 2(*c*)).

State-of-the-art sandwich panels which are used in aerospace and civil engineering are usually too thick for the manufacture of automotive components. Their thicknesses are typically in the tens of millimeter range whereas common sheet thicknesses in car bodies range from about 0.5 to 2.0 mm. Therefore, new technologies have been developed which led to the design of so-called *thin* sandwich sheets with thicknesses of the order of 1 mm. Gustafsson [8] proposed the 1.2-mm-thick hybrid stainless steel assembly (HSSA). This sandwich sheet material comprises a 1-mm-thick 10% relative density core material made of 20-μm-thick steel fibers which are oriented perpendicular to the face sheets. In order to improve the shear strength of the HSSA fiber core, Markaki and Clyne [9] developed a random network fiber core of similar density but different fiber orientations and solid joints between contact fibers. With automotive applications in mind, Mohr and Straza [10] presented an all-metal sandwich sheet of 2.5 mm thickness with a fine cell honeycomb core that could withstand the high shear loads in a forming operation. Mohr and Wierzbicki [7] investigated the crushing behavior of box columns made of the 1.2-mm-thick HSSA and found that the crushing response is dominated by the "shear-soft" behavior of the fibrous core material (Fig. 2(*b*)). The crushed sandwich columns featured a considerably short folding wavelength due to the pronounced strength disparity between the face sheets and the core material. At the same time, the results have shown that the specific energy absorption of box structures made from shear-soft sandwich sheets may be up to 60% higher than that of very-thin monocoque structures. Santosa and Wierzbicki [11] performed a numerical and theoretical analysis on sandwich sheets and concluded that for the same energy absorption, sandwich columns may be 40%–60% lighter than their solid-section counterparts.

The overall objective is to evaluate the weight specific energy absorption of sandwich structures and compare their performance with conventional thin-walled designs. Thin sandwich sheet technology is still at an early stage of development. Estimating their weight savings potential in crashworthiness applications is of great importance in order to guide future sandwich material developments. The main focus of this study is to assess the weight savings potential of so-called "shear-rigid" sandwich constructions. First, a finite element study is carried out to determine the energy absorption behavior of 1.6-mm-thick sandwich columns. Perforated sandwich core materials of relative densities ranging from 25% to 60% are considered (Fig. 3(*a*)) since these may be manufactured in a low cost mass production process. Theoretical arguments are developed to explain the simulation results. It is shown that the aforementioned conclusions regarding the specific energy absorption of sandwich structures must be tempered from a theoretical point of view. Sandwich construction may lead to sub-

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 20, 2004; final manuscript received November 7, 2005. Review conducted by K. Ravi-Chandar. Discussion on the paper should be addressed to the Editor Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

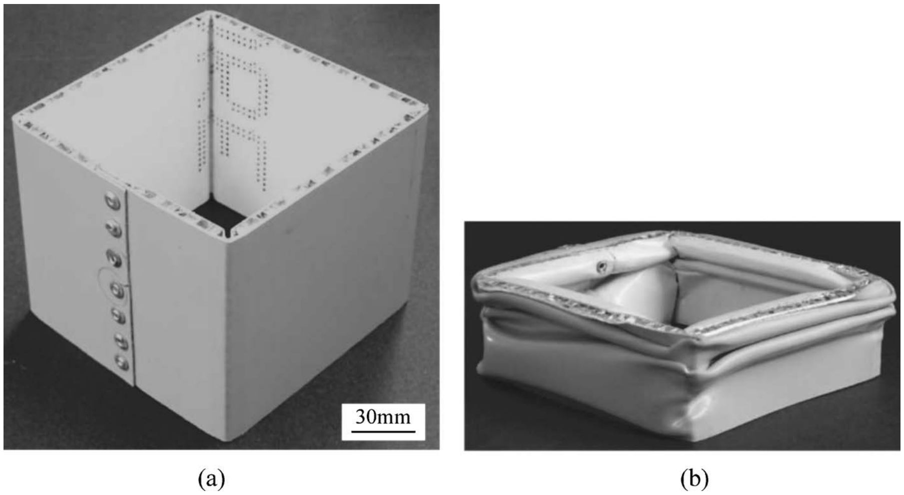


Fig. 1 Example of a 120-mm-wide and 100-mm-high square box column made from an aluminum honeycomb sandwich sheet (*a*) before and (*b*) after quasi-static crushing (Mohr and Wierzbicki [6])

stantial weight savings in crashworthiness applications, but these are fairly small when realistic configurations are considered.

2 Finite Element Simulations

The crushing behavior of square sandwich box columns (Fig. 2) is studied using finite element simulations. It is assumed that both the face sheets and the cellular sandwich core are made of the same metallic base material. In a first step, we determine the relationship between the elasto-plastic properties of the core material and its relative density from the finite element analysis of characteristic microstructures. Subsequently, at a larger length scale, numerical simulations are carried out to investigate the crushing response of prismatic columns made from sandwich materials of different relative densities.

2.1 Mechanical Properties of the Perforated Core Material. Figure 3(*b*) shows the two-dimensional staggered hole pattern which characterizes the microstructure of the perforated core material. Denoting the hole diameter and spacing by D and d, respectively, the relative density of the core material, ρ^* (that is the ratio of the core material density ρ_c to the density of the basis material ρ_f) may be expressed as

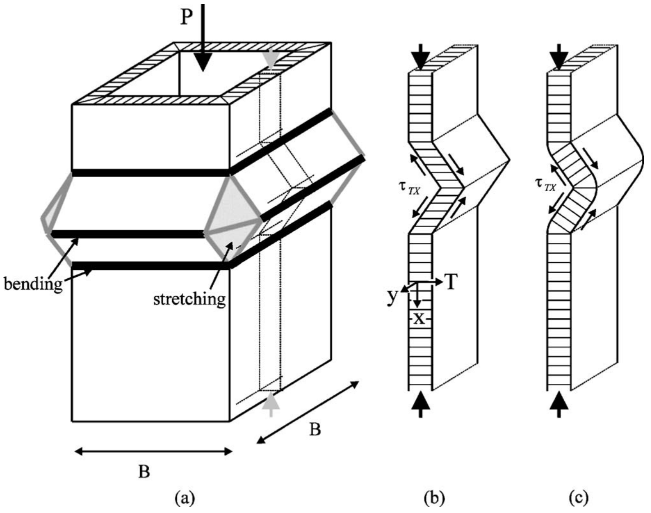


Fig. 2 (*a*) Schematic of the crushing of a square sandwich box column. In the column corners the energy is mostly dissipated through stretching (shaded area) whereas cell wall bending dominates along the hinge lines (thick lines) (*b*) and (*c*) Longitudinal cut through a cell wall; in the case of shear-soft sandwich sheets (*b*), the core material undergoes large shear deformations as the sandwich is bent, whereas in the case of shear-rigid sandwich sheets (*c*), initially perpendicular cross sections remain perpendicular throughout bending while the two face sheets are respectively compressed or stretched (Mohr and Wierzbicki [7]).

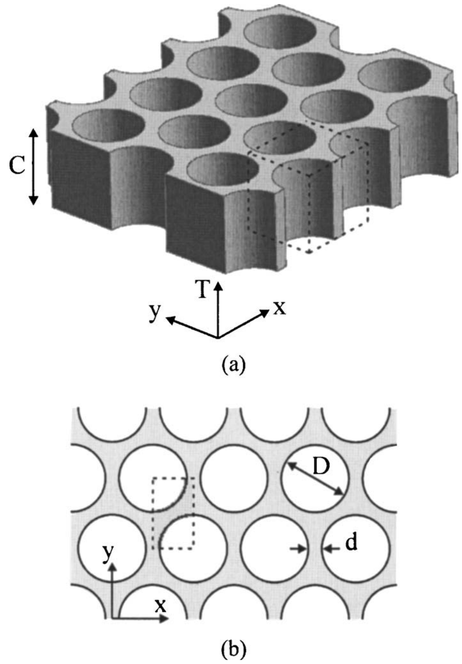


Fig. 3 Microstructure of the perforated sandwich core material. The in-plane coordinate axes have been labeled "*x*" and "*y*," while "*T*" indicates the out-of-plane direction. The dashed wire frame in (*a*) and the rectangle in (*b*) highlight the mechanical unit cells which are chosen for numerical analysis of the out-of-plane and in-plane properties, respectively.

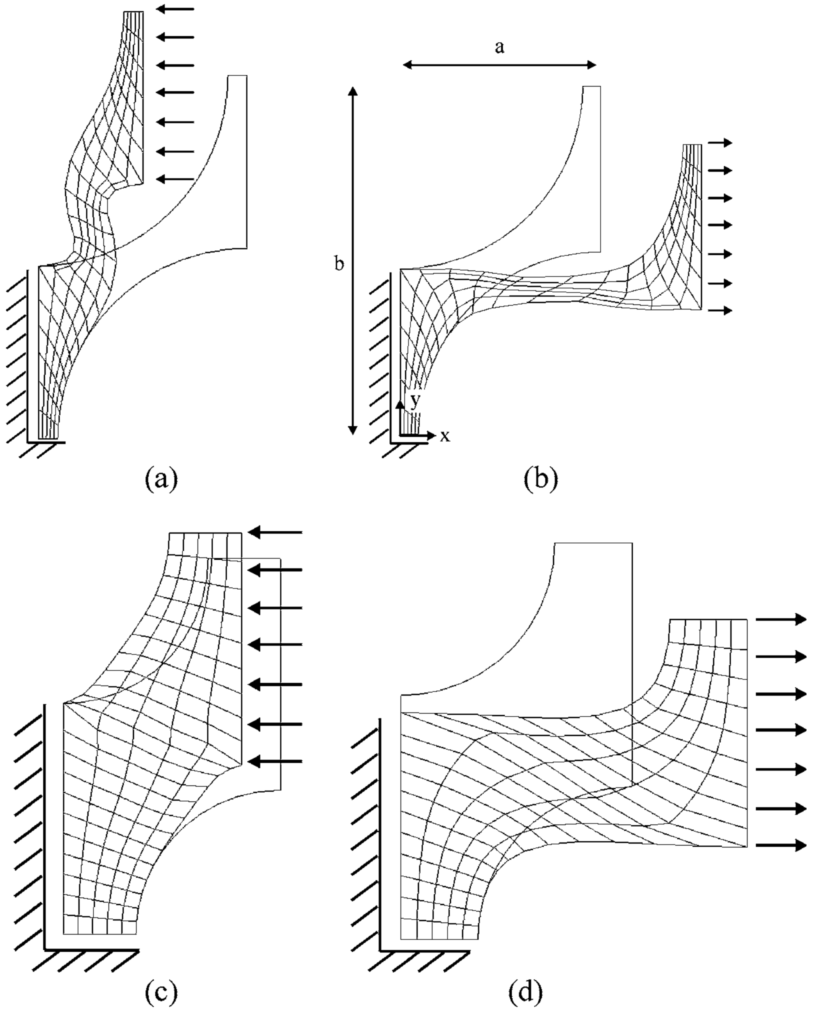


Fig. 4 In-plane loading in the x-direction: (a) compression and (b) tension for $\rho^*=25\%$, (c) compression and (d) tension for $\rho^*=60\%$

$$\rho^* = \frac{\rho_c}{\rho_f} = 1 - \frac{\pi}{2\sqrt{3}} \frac{1}{\left(1 + \frac{d}{D}\right)^2} \tag{1}$$

In the limiting case of zero spacing between the holes ($d=0$), the relative density is $\rho^*=9.3\%$, which corresponds to the theoretically lightest perforated core material. Note that standard perforated plates have a significantly larger ρ^*. When characterizing orthotropic sandwich core materials, it is useful to distinguish between the in-plane and out-of-plane properties. The two in-plane directions x and y are aligned with the symmetry axes of the characteristic two-dimensional core microstructure. The out-of-plane direction which is aligned with the direction of the cylindrical holes is referred to as T-direction. The mechanical behavior of the perforated core materials depends on their relative density. Here, finite element simulations are carried out for four different core densities. We chose $D=1$ mm and varied the hole spacing d. The core configurations are as follows: $d=0.1$ mm ($\rho^*=25\%$), $d=0.2$ mm ($\rho^*=37\%$), $d=0.3$ mm ($\rho^*=46\%$), and $d=0.5$ mm ($\rho^*=60\%$). The base material is modeled as elasto-plastic Levy-Mises solid with linear isotropic hardening behavior: Young's modulus $E_f=210$ GPa, elastic Poisson's ratio $\upsilon=0.33$, initial yield stress $\sigma_0=300$ MPa, hardening modulus $H=1000$ MPa, and mass density $\rho_f=7.8\ \mathrm{g/cm^3}$.

2.1.1 In-Plane Properties. Due to the two-dimensional periodicity of the perforated microstructure, the in-plane properties may be determined from the analysis of a suitable mechanical unit cell (see dashed rectangle in Fig. 3(b)). The geometry is discretized by four-node plane stress elements (type CPS4R, Abaqus [12]). The deformed meshes for the lightest and heaviest core structure are shown in Fig. 4. The following displacement conditions have been chosen along the unit cell boundaries

$$\text{at } x=0 \quad u_x = 0 \tag{2}$$

$$\text{at } x=a \quad u_x = \varepsilon_{xx} a \tag{3}$$

$$\text{at } y=0 \quad u_y = 0 \tag{4}$$

$a=(D+d)/2$ and $b=\sqrt{3}(D+d)/2$ denote the unit cell width along the x- and y-direction, respectively (Fig. 4(b)). It follows from symmetry that

$$\text{at } y=b \quad u_y = \langle u_y \rangle_{y=b} \tag{5}$$

where the term $\langle u_y \rangle_{y=b}$ denotes the average displacement of all nodes along the boundary $y=b$. For each core configuration, the unit cell is subjected to compression and tension along the

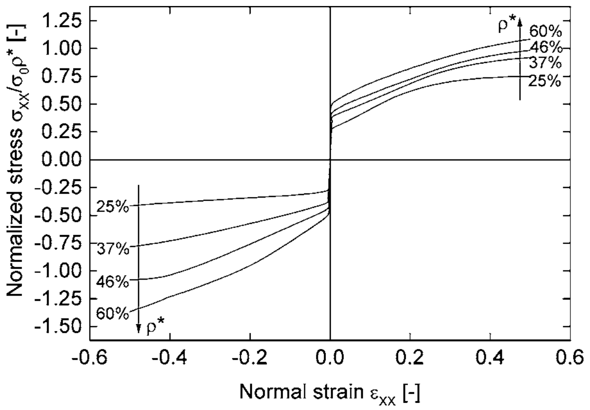


Fig. 5 Engineering stress-strain curves for uniaxial in-plane loading along the *x*-direction. The results are shown for four different cores densities: ρ^*=25%, 37%, 46%, and 60%.

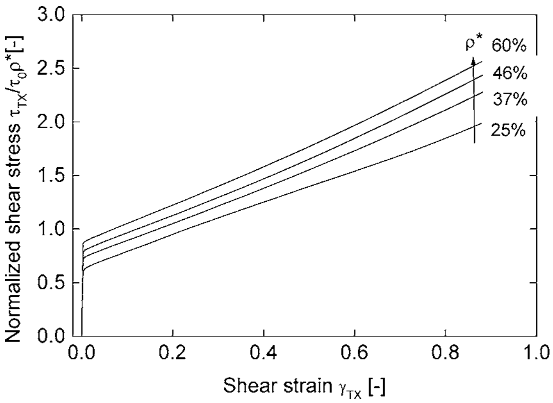


Fig. 7 Engineering stress-strain curves for out-of-plane shear loading along the *T*-*x*-plane

x-direction as prescribed by ε_{xx}. The respective macroscopic engineering stresses σ_{xx} are calculated and plotted as a function of the work-conjugate engineering strain ε_{xx} (Fig. 5). Note that the depicted macroscopic stresses are normalized by $\sigma_0\rho^*$. Irrespective of the core configuration and loading direction, the stress-strain curves increase monotonically. All curves are in hierarchical order with respect to the relative density which indicates an increase in microstructural efficiency the higher the relative density. The material response to in-plane shear loading has not been investigated because it is only of minor importance for the crushing response of sandwich structures. Furthermore, it is assumed that the material response to uniaxial loading along the *y*-axis may be approximated by the macroscopic response curves for loading along the *x*-axis, i.e., $\sigma_{yy}(\varepsilon_{yy}) \cong \sigma_{xx}(\varepsilon_{xx})$.

2.1.2 Out-Of-Plane Properties. The out-of-plane properties of the perforated core material describe its behavior under shear and normal loading in the *T*-*x*- and *T*-*y*-planes. A fully three-dimensional finite element model is used to study the out-of-plane response (Fig. 6). Again, we make use of the periodicity of the microstructure and limit our attention to the behavior of the mechanically representative unit cell of the core material. The boundary conditions at the top and bottom boundaries of the three-dimensional unit cell are

$$\text{at } T=0 \quad u_x = u_y = u_T = 0 \tag{6}$$

$$\text{at } T=C \quad u_x = \gamma_{TX}C \quad u_y = 0 \quad u_T = \langle u_T \rangle_{T=C} \tag{7}$$

The term $\gamma_{TX}C$ corresponds to the shear displacement which is applied to the top boundary. Plane strain conditions are prescribed in the *y*-direction, i.e.,

$$\text{at } y=0 \quad u_y = 0 \tag{8}$$

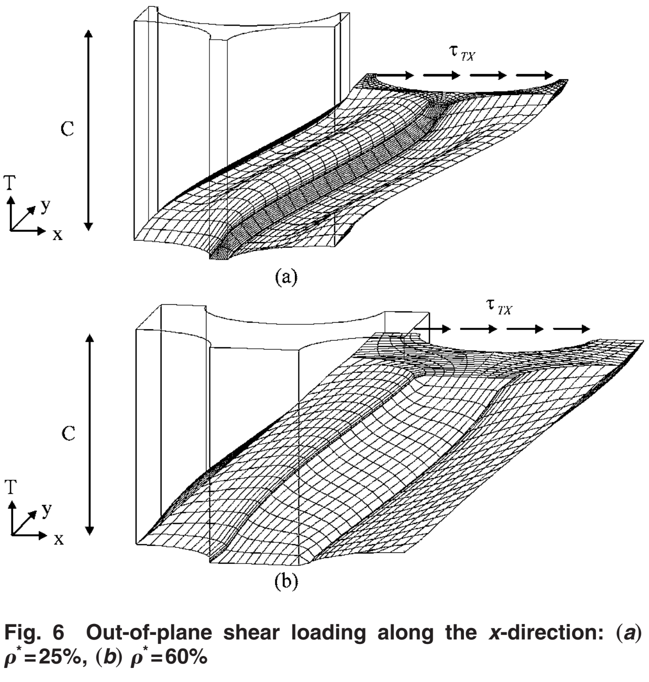
$$\text{at } y=b \quad u_y = 0 \tag{9}$$

Furthermore, we impose three periodicity conditions at the boundaries $x=0$ and $x=2a$

$$u_x(x=0, y=y_i, T=T_j) = u_x(x=2a, y=y_i, T=T_j) \tag{10}$$

$$u_y(x=0, y=y_i, T=T_j) = u_y(x=2a, y=y_i, T=T_j) \tag{11}$$

$$u_T(x=0, y=y_i, T=T_j) = u_T(x=2a, y=y_i, T=T_j) \tag{12}$$



(a)

(b)

Fig. 6 Out-of-plane shear loading along the *x*-direction: (*a*) ρ^*=25%, (*b*) ρ^*=60%

Figure 6(*a*) shows the unit cell for the 25% relative density core material subject to shear loading in *x*-direction. The boundary conditions at the top are chosen such that the macroscopic stress in the T-direction is zero. In return, we observe considerable compaction along the T-direction as shear loading is applied. Similar behavior is observed for higher density cores, but the shear-induced compaction is less pronounced. As for the in-plane loading, we assume that the response to shear loading in the *T*-*y* plane (Fig. 7) may be approximated by the results for the *T*-*x*-plane, i.e., $\tau_{TY}(\varepsilon_{TY}) \cong \tau_{TX}(\varepsilon_{TX})$.

The core material response to uniaxial loading along the T-direction is calculated analytically. For this loading condition, the microstructural stress state is uniaxial and the rule of mixtures applies

$$\sigma_{TT} = \rho^*(\sigma_0 + H\varepsilon_{TT}) \tag{13}$$

2.1.3 Macroscopic Modeling of the Core Material. Fully three-dimensional mechanism-based constitutive models for perforated core materials must be developed in the future. In this study, we calibrate the heuristic constitutive model 126 of the LS-DYNA material model library (LSTC [13]) such as to approximate the mechanical behavior of the core material. According to the formulation of model 126, the elasto-plastic material behavior is defined by a six-dimensional cube in the six-dimensional stress space of σ_{xx}, σ_{yy}, σ_{TT}, τ_{Tx}, τ_{Ty}, and τ_{xy} along with an associated flow rule. Strain-hardening/softening changes the size of the cube, but it does not alter the orientation of the cube faces. All stress-

Table 1 Input data for LS-DYNA material model 126

	Core configuration			
ρ^* [−]	S1 25.0%	S2 37.0%	S3 46.3%	S4 59.7%
E_{xx} (MPa)	10,025	24,380	39,155	65,420
s_{xx}^0 (MPa)	22	44	58	92
H_{xx}^c (MPa)	84	143	195	251
H_{xx}^t (MPa)	13	62	131	216
E_{xT} (MPa)	11,812	16,350	22,209	31,850
s_{xT}^0 (MPa)	28	48	65	92
H_{xT}^c (MPa)	69	114	151	200
H_{xT}^t (MPa)	69	114	151	200
E_{TT} (MPa)	52,604	77,744	97,308	125,356
s_{TT}^0 (MPa)	75	111	139	179
H_{TT}^c (MPa)	250	370	463	597
H_{TT}^t (MPa)	250	370	463	597

strain curves are approximated by linear hardening functions. Thus, for each component σ_{ij} of the Cauchy stress tensor, the interval of admissible stress states is given by

$$-s_{ij}^0 + H_{ij}^c \varepsilon_{ij}^p \leq \sigma_{ij} \leq s_{ij}^0 + H_{ij}^t \varepsilon_{ij}^p \tag{14}$$

where the logarithmic plastic strains are defined as

$$\varepsilon_{ij}^p = \varepsilon_{ij} - \frac{\sigma_{ij}}{E_{ij}} \tag{15}$$

In sum, four material properties must be specified for each stress component σ_{ij}: (1) the elastic modulus, E_{ij}, (2) the initial yield stress, s_{ij}^0, (3) the hardening modulus for tensile loading, H_{ij}^t, and (4) the hardening modulus for compressive loading, H_{ij}^c. Table 1 summarizes these material model parameters after calibration for the perforated core materials. Note that we assume the same material response to loading in the *T*-*x* and *T*-*y*-planes. The key deficiency of the core material model is that the yield and hardening responses of the six stress components are decoupled. In the present structural application, coupling effects would weaken the core material whereas the local interaction of the face sheets and core would increase the core strength. In sum, both opposite effects are expected to result in an overall small error in the simulation results.

2.2 Crushing Simulations. Thin-walled box columns with square cross sections (Fig. 1) of $B=80$ mm width are considered as representative for energy absorbing members in standard car bodies. The crushing simulations are carried out for columns made from sandwich sheets with different core materials but the same face sheet and core thickness. With the exception of the relative density of the sandwich core material, all other model parameters are kept constant:

(i) cross-section width, $B=80$ mm,
(ii) column height of $l=156$ mm,
(iii) face sheet thickness, $t=0.2$ mm,
(iv) total sandwich sheet thickness, $h=1.6$ mm,
(v) core thickness, $C=h-2t=1.2$ mm.

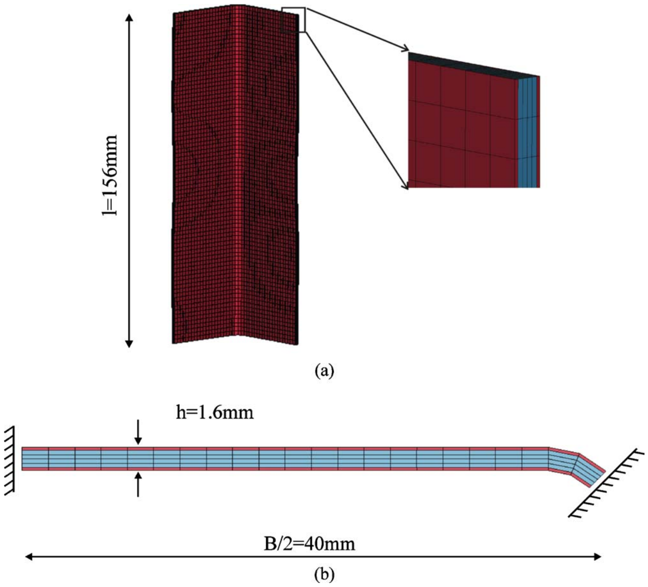


Fig. 8 FE-mesh: (*a*) side view of the quarter model, (*b*) top view of the eighth model. The red elements represent the $t=0.2$-mm-thick face sheets, the blue elements discretized the $C=1.2$-mm-thick core material. The total sandwich sheet thickness is $h=C+2t=1.6$ mm.

Solid elements are chosen to discretize the sandwich column (Fig. 7). It is important to note that high-density core materials contribute to the bending resistance of the sandwich cross section and thus, the central core layer has to be represented by several solid elements along the out-of-plane direction. The 0.2-mm-thick face sheets are discretized by a single layer of solid elements, whereas four reduced-integration solid elements represent the 1.2-mm-thick core material (Fig. 8(*b*)). Regarding the modeling of the box columns, we make use of the symmetry of the square cross section: the quarter model (Fig. 8(*a*)) may be used with suitable symmetry boundary conditions to represent the extensional and inextensional folding modes. Moreover, since preliminary simulations indicated the dominance of the fully extensional mode, an eighth model is chosen for computational efficiency (Fig. 8(*b*)). The axial loading is quasi-statically applied to the top of the column by a flat moving rigid wall. Contact is defined between all nodes and elements of the face sheets. In sum, the model includes about 10,000 reduced-integration brick elements. 180,000 explicit time steps are performed throughout the crushing over a distance of 80 mm. The simulations are carried out for quasi-static conditions. Under dynamic loading, the crushing mode may be affected by both the strain rate sensitivity and the lateral inertia throughout the fold formation. The problem of strain rate sensitivity has been extensively studied in literature also in relation to the crashworthiness problems (see for example, Abramowicz and Jones [14]). It is relatively easy to replace the plastic constitutive equations by the viscoplastic behavior and rerun the calculations. This has not been done because this will not bring any new qualities. It should be noted that the loading velocities in car crashes typically do not exceed 15 m/s. In this range, lateral inertia effects are still considered to be small for conventional thin-walled structures. It was shown by Jamjiam et al. [15] that inclusion of lateral inertial will introduce only 1% change in resistance in dynamically loaded honeycomb blocks. The same argument may also apply to sandwich structures since the weight of the laterally accelerated material is of the same order of magnitude as that of conventional homogeneous cross-sections.

2.3 Results. The computed force-displacement curves are shown in Fig. 9. The curves are in hierarchical order with respect to the relative density of the core material. The mean crushing force has been evaluated for each simulation result by dividing the integral of each force-displacement curve $P(u)$ by the final crushing distance of $u_f=80$ mm,

$$P_m = \frac{1}{u_f}\int_0^{u_f} P(u)du \tag{16}$$

The lowest curve, which is the result for the structure with the 25% relative density core material, has a mean force level of

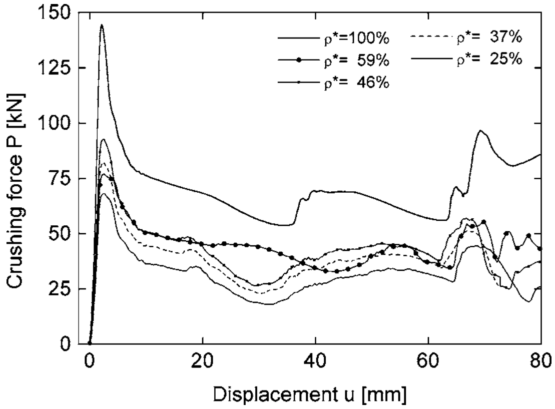


Fig. 9 Force-displacement curves for the crushing of square sandwich box columns

about 32 kN. The curve for the solid cross section (100% relative density) shows a mean crushing force of 71 kN. The force-displacement curves for the remaining specimens are sandwiched between these two results.

The folding wave length is remarkably similar for all specimens (Fig. 10). Furthermore, it is noted that none of the crushed sandwich structures exhibits the shear-folding mode which has been reported for fibrous core materials (Mohr and Wierzbicki [7]). This clearly indicates that the shear strength of the perforated core materials has been sufficiently high, i.e., the core behaves in a shear-rigid manner. In return, the normal strength of the core material in the T-direction seems to be more important during the crushing of sandwich structures with shear-rigid cores. Observe that the lightest sandwich structure with a core density of 25% shows some partial through-the-thickness crushing of the central layer in the regions of the smallest bending radii (Fig. 10(*a*)). On the other hand, core materials of higher densities can withstand the pressure along the local T-direction while the sandwich sheet is bent plastically.

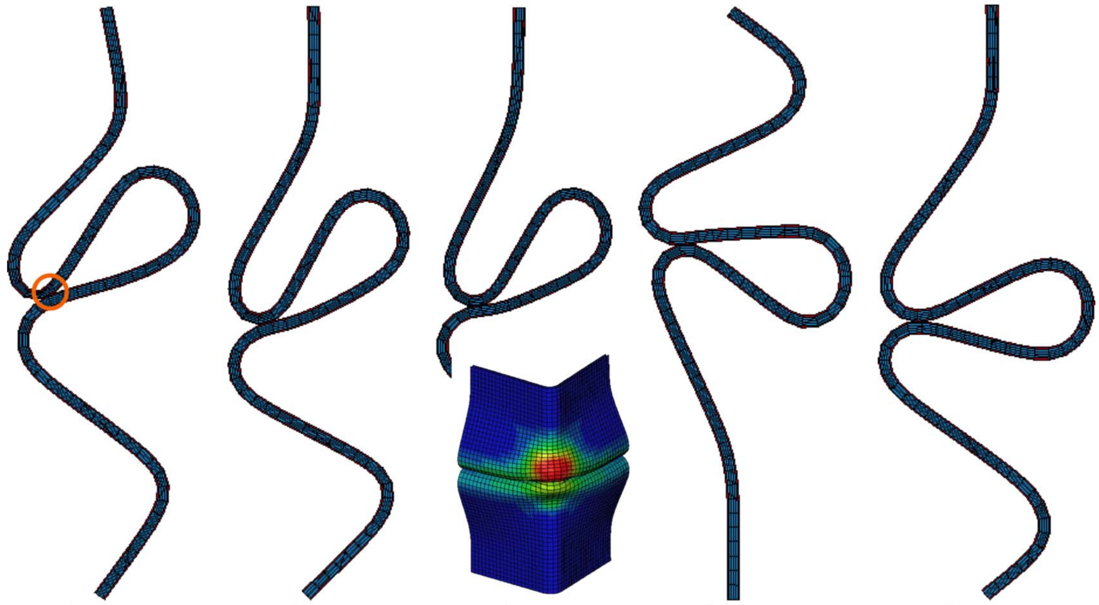
The total mass of each specimen, m, depends on the relative density of the core material. For the all-metal sandwich columns with square cross sections of width B, the total mass is

$$m \cong 4Bl(C\rho^* + 2t)\rho_f \tag{17}$$

As the data in Table 2 show, the column mass varies linearly from 273 g for the lightest column ($\rho^*=25\%$) to 623 g for the solid-section profile ($\rho^*=100\%$). The specific energy absorption (SEA) is calculated as

$$\text{SEA} = \frac{P_m u_f}{m} \tag{18}$$

Upon evaluation, we find that the specific energy absorption is approximately the same for all specimens (Table 2). In other



(a) $\rho^* = 25\%$, (b) $\rho^* = 37\%$, (c) $\rho^* = 46\%$, (d) $\rho^* = 59\%$, (e) $\rho^* = 100\%$

Fig. 10 Folded sandwich cross sections for different core densities. The encircled region in (*a*) highlights the area of partial through-the-thickness crushing of the core material. The insert in (*c*) shows the 3D view of the bending and stretching of the profile walls (compare with Fig. 1) as obtained from FEA; the color indicates the equivalent plastic strain in the outer face sheet.

Table 2 Results obtained from crushing simulations

	Specimen				
	S1	S2	S3	S4	S5
Relative core density, ρ^*	25.0%	37.0%	46.3%	59.7%	100.0%
Mass (g)	273	329	372	435	623
Mean crushing force, P_m (kN)	32.0	38.5	42.8	44.5	70.8
Specific energy absorption, SEA (kJ/kg)	13.7	13.7	13.5	12.0	13.3
Energy absorbed by core	37%	47%	48%	45%	67%
Energy absorbed by face sheets	63%	53%	52%	55%	33%

words, the choice of the sandwich core density has only little influence on the specific energy absorption when keeping the total wall thickness constant.

3 Theoretical Analysis

The purpose of the following analysis is to support the above observations through theoretical arguments. No attempt is made to make quantitative estimates of the mean crushing force or the folding wavelength. The reader is referred to Santosa and Wierzbicki [11] and Mohr and Wierzbicki [7] for the comparison of theory with simulations and experiments.

3.1 Energy Partitioning. In their crushing analysis of sandwich structures, Mohr and Wierzbicki [7] identified three potential energy absorption mechanisms: plastic bending, stretching, and core shearing. Using the symbol Ψ_{tot} to denote the energy absorbed throughout the formation of a single fold (such as that depicted in Fig. 1), we have the sum

$$\Psi_{\text{tot}} = \Psi_{\text{bend}} + \Psi_{\text{stretch}} + \Psi_{\text{shear}} \tag{19}$$

where Ψ_{bend}, Ψ_{stretch}, and Ψ_{shear} are the contributions of the aforementioned mechanisms. Core shear crushing is only relevant for shear-soft sandwich structures, whereas in the present case of shear-rigid sandwich structures, we have $\Psi_{\text{shear}}=0$. In other words, the bending and stretching of the sandwich profile walls (Fig. 2) are the only energy absorption mechanisms. The theoretical analysis by Mohr and Wierzbicki [7] also indicates that if $\Psi_{\text{shear}}=0$, the two remaining mechanisms contribute equally to the overall energy absorption. That is

(1) one half of the energy is dissipated by plastic bending along hinge lines

$$\Psi_{\text{bend}} = \tfrac{1}{2}\Psi_{\text{tot}} \tag{20}$$

(2) the other half is dissipated by the stretching of the sheet midplane in the corners of the thin-walled column (Fig. 1)

$$\Psi_{\text{stretch}} = \tfrac{1}{2}\Psi_{\text{tot}} \tag{21}$$

The energy dissipated during plastic bending, here denoted as Ψ_{bend}, is proportional to the fully plastic bending moment M_0 of the cell walls

$$\Psi_{\text{bend}} \propto M_0 \tag{22}$$

Similarly, the energy dissipated during stretching, Ψ_{stretch}, is proportional to the fully plastic normal force, N_0

$$\Psi_{\text{stretch}} \propto N_0 \tag{23}$$

For the sandwich sheets, the fully plastic bending moment of the cell walls reads

$$M_0 = \underbrace{(C+t)t\sigma_f}_{\text{facings}} + \underbrace{\frac{1}{4}C^2\sigma_C}_{\text{core}} \tag{24}$$

where σ_f and σ_C are the energy-equivalent yield strengths of the face sheet and core material, respectively. Recall that C is the core thickness, t is the face sheet thickness. Furthermore, the fully plastic normal force reads

$$N_0 = \underbrace{2t\sigma_f}_{\text{facings}} + \underbrace{C\sigma_C}_{\text{core}} \tag{25}$$

Using Eqs. (8)–(14), the partitioning of the total energy between the core and face sheets may be estimated

$$\Psi_{\text{tot}} = \Psi_{\text{core}} + \Psi_{\text{face}}. \tag{26}$$

As far as plastic bending is concerned, the core material absorbs, Ψ^b_{core}

$$\Psi^b_{\text{core}} = \frac{1}{2}\Psi_{\text{tot}}\frac{\frac{1}{4}C^2\sigma_C}{M_0} \tag{27}$$

while the stretching contribution, Ψ^s_{core}, is

$$\Psi^s_{\text{core}} = \frac{1}{2}\Psi_{\text{tot}}\frac{C\sigma_C}{N_0} \tag{28}$$

Thus, in sum, the energy absorbed by the core material reads

$$\Psi_{\text{core}} = \Psi^b_{\text{core}} + \Psi^s_{\text{core}} = \frac{1}{2}\Psi_{\text{tot}}\left[\frac{\frac{1}{4}C^2\sigma_C}{M_0} + \frac{C\sigma_C}{N_0}\right] \tag{29}$$

Similarly, we find the energy absorbed by the face sheets, Ψ_{face},

$$\Psi_{\text{face}} = \Psi_{\text{tot}} - \Psi_{\text{core}} = \frac{1}{2}\Psi_{\text{tot}}\left[\frac{(C+t)t\sigma_f}{M_0} + \frac{2t\sigma_f}{N_0}\right] \tag{30}$$

In the following, Eqs. (18) and (19) are used to compare the distribution of the absorbed energy with the distribution of the mass within the sandwich structure.

3.2 Specific Energy Absorption for the Same Wall Thickness. Conventional solid-section box structures may be considered as sandwich columns with a fictitious core material of 100% relative density, that is $\sigma_f=\sigma_C$ and $\rho^*=1$. Upon evaluation of Eq. (29), we find that the 1.2-mm-thick solid core absorbs about two-thirds of the total energy

$$\frac{\Psi_{\text{core}}}{\Psi_{\text{tot}}} = 66\% \tag{31}$$

This result is in good agreement with the corresponding numerical simulation (67%, see Table 2). For solid cores ($\rho^*=100\%$), the mass distribution is given by the ratio of the layer thickness to the total sheet thickness. Here, we have $C=1.2$ mm and $h=1.6$ mm and thus

$$\frac{m_{\text{core}}}{m_{\text{tot}}} = \frac{C}{h} = 75\% \tag{32}$$

In other words, 75% of the mass is attributed to the fictitious core material, 25% to the face sheets. In sum, it may be concluded that the partitioning between face sheets and core material is fairly similar as far as the energy dissipation and mass are concerned.

Recall that the success of the sandwich concept in stiffness applications relies on the redistribution of material within the cross section such as to obtain a more homogeneous distribution of the elastic strain energy per unit mass. For illustration, consider the flexural stiffness per unit width, D_{tot}, for a sandwich sheet

$$D_{\text{tot}} = \underbrace{\frac{E_f[(C+2t)^3 - C^3]}{12}}_{\text{facings}} + \underbrace{\frac{E_c C^3}{12}}_{\text{core}} \tag{33}$$

Here, E_f and E_C denote the Young's moduli for the face sheet and core material. In the case of the 1.6-mm-thick homogeneous sheet, it is found that the core contributes about 42% of the total bending stiffness

$$\frac{D_{\text{core}}}{D_{\text{tot}}} = \frac{1}{\left(1 + 2\frac{t}{C}\right)^3} = 42\% \tag{34}$$

Observe that there is a significant difference between the elastic strain energy (which is proportional to D) and the mass distribution (cf. Eq. (21)) in stiffness applications. Clearly, in order to increase the flexural stiffness per unit weight, reducing the mass of the central layer is beneficial as it will reduce the amount of inefficiently used material within the cross section. In the case of the crushing of box columns, the energy distribution is far more homogeneous. Here, the contribution of the fictitious core layer to the total energy dissipation is almost the same as its mass percentage. Thus, unlike for elastic bending, there is only little room left for optimization of the sheet cross section as far as the specific energy absorption is concerned. This conclusion which is based on theoretical considerations is in accordance with the simulation results.

3.3 Specific Energy Absorption for the Same Mean Crushing Force.

A challenging task in automotive engineering is the design of lightweight longitudinal rails which crush under a given mean crushing force P_m. According to the simple analytical model of Mohr and Wierzbicki [7], the mean crushing force of shear-rigid sandwich structures may be related to M_0 and N_0

$$P_m \propto \sqrt{M_0 N_0} \qquad (35)$$

Assume that the baseline design be given by a solid-section profile $(\rho^*=1)$ of wall thickness $h=h_0$. The question of interest to the automotive engineer is as to whether there may exist a sandwich design which provides the same mean crushing force but at a lower weight. To provide an answer, we introduce the relative mean crushing force, $\tilde{p}$

$$\tilde{p}(C,t,\rho^*) := \sqrt{\frac{M_0(C,t,\rho^*)N_0(C,t,\rho^*)}{\hat{M}_0\hat{N}_0}} \qquad (36)$$

with the baseline solutions

$$\hat{M}_0 = M_0(C=h_0-2t,t,\rho^*=1)=\tfrac{1}{4}h_0^2\sigma_f \qquad (37)$$

$$\hat{N}_0 = N_0(C=h_0-2t,t,\rho^*=1)=h_0\sigma_f \qquad (38)$$

and the relative mass, $\tilde{m}$

$$\tilde{m}(C,t,\rho^*) := \frac{C\rho^*+2t}{h_0} \qquad (39)$$

In a first approximation, it is assumed that the energy-equivalent yield strength of the core material is proportional to its relative density

$$\sigma_c = \rho^*\sigma_f \qquad (40)$$

Furthermore, we limit our attention to sandwich constructions with the minimum possible face sheet thickness, i.e. $t=t_{\min}$ $=$const. This thickness is typically given by technical requirements such as the surface denting resistance or manufacturing limits. When keeping the mean crushing force constant,

$$\tilde{p}(C,t_{\min},\rho^*)=1 \qquad (41)$$

defines the relationship between the required core thickness C and the relative core material density. Upon evaluation, we find

$$\left[\left(\frac{C}{t_{\min}}+1\right)+\frac{1}{4}\left(\frac{C}{t_{\min}}\right)^2\rho^*\right]\left[2+\left(\frac{C}{t_{\min}}\right)\rho^*\right]=\frac{1}{4}\left(\frac{h_0}{t_{\min}}\right)^3 \qquad (42)$$

We solved Eq. (42) for $h_0/t_{\min}=1.6/0.2=8$ and made use of the solution $C(\rho^*)$ to evaluate $\tilde{m}$ as a function of ρ^*. In addition, the required relative total wall thickness, $\tilde{h}$ is of interest

$$\tilde{h}(\rho^*) := \frac{C(\rho^*)+2t_{\min}}{h_0} \qquad (43)$$

Both the change in column mass and profile wall thickness are plotted in Fig. 11. The basis design corresponds to the points at $\rho^*=1$, where both $\tilde{m}$ and $\tilde{h}$ equal unity. The total column mass decreases monotonically in the relative density which reveals that weight savings may be achieved by the use of sandwich structures. At the same time, while decreasing the mass, the total wall thickness must be increased in order to maintain the mean crushing force. From a theoretical point of view, weight savings as high as 50% may be achieved through the use of sandwich structures in crashworthiness applications (Fig. 11). However, in view of practical application this conclusion must be tempered for three reasons:

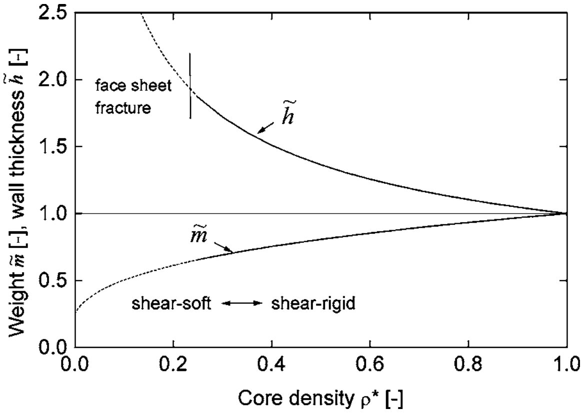


Fig. 11 Design for a constant mean crushing force. Relative weight and relative wall thickness of the equivalent sandwich structure as a function of the relative core material density (see Eqs. (39) and (43)).

(1) The likelihood of face sheet fracture increases for thicker profile walls. Thus, the full weight savings potential may not be exploited due to face sheet fracture throughout crushing.

(2) For low core densities, the assumption of shear-rigid behavior is expected to break down. Instead, shear-soft crushing may become more dominant. As discussed in Mohr and Wierzbicki [7], large shear deformation of the core material reduces substantially the energy absorption of the face sheets. Therefore, the sandwich structure would no longer provide the same mean crushing force.

(3) The energy-equivalent flow stress of the core may be lower than predicted by Eq. (40). Note that the normalized stress strain curves in Fig. 5 are usually below 1, which indicates that the perforated core material would absorb less energy than assumed throughout the discussion.

4 Conclusions

In an earlier study, Mohr and Wierzbicki [7] investigated the crush behavior of shear-soft sandwich columns and found that these absorb about 60% more energy than conventional monocoque structures of the same weight (wall thickness 0.4 mm). However, for standard passenger cars the baseline design typically uses solid sections of about 1.6 mm wall thickness. Box columns made of these materials provide a mean crushing force that cannot be obtained by existing shear-soft sandwich materials. In this paper, shear-rigid sandwich materials are considered as an alternative to shear-soft designs. It is found from numerical simulations that 1.6-mm-thick sandwich columns with perforated cores exhibit almost the same specific energy absorption as solid-section profiles of the same wall thickness. This has been explained by the observation that the energy distribution in conventional thin-walled structures is fairly homogeneous over the wall thickness. Thus, the basic concept of sandwich design, namely the redistribution of the material within the wall cross section, does not improve the specific energy absorption when the profile wall thickness is kept constant.

The theoretical analysis also reveals that the weight specific crushing performance may be improved by increasing the total wall thickness. From a theoretical point of view, sandwich structures may be up to 50% lighter than conventional designs while providing the same mean crushing force. However, in practice, these gains may be considerably smaller since face sheet fracture and core shear failure are likely to limit the performance of sandwich box columns with thick core materials. Therefore, as for most existing sandwich applications, the high bending stiffness-to-weight ratio is seen as the main advantage of sandwich technology in automotive design.

Acknowledgment

Financial support of the Cambridge-MIT Institute (CMI) is gratefully acknowledged.

References

[1] Jacob, G. C., Fellers, J. F., Simunovic, S., and Starbuck, J. M., 2002, “Energy Absorption in Polymer Composites for Automotive Crashworthiness,” J. Compos. Mater., **36**(7), pp. 813–850.

[2] Kim, H. S., and Wierzbicki, T., 2001, “Effect of the Cross-Sectional Shape of Hat-Type Cross-Sections on Crash Resistance of an “S”-Frame,” Thin-Walled Struct., **39**, pp. 535–554.

[3] Reid, S. R., Reddy, T. Y., and Gray, M. D., 1986, “Static and Dynamic Crushing of Foam-filled Sheet Metal Tubes,” Int. J. Mech. Sci., **28**(5), pp. 295–322.

[4] Seitzberger, M., Rammerstorfer, F. G., Gradinger, R., Degischer, H. P., Blaimschein, M., and Walch, C., 2000, “Experimental Studies on the Quasi-Static Axial Crushing of Steel Columns Filled With Aluminum Foam,” Int. J. Solids Struct., **37**, pp. 4125–4147.

[5] Santosa, S. P., Wierzbicki, T., Hanssen, A. G., and Langseth, M., 2000, “Experimental and Numerical Studies of Foam-Filled Sections,” Int. J. Impact Eng., **24**(5), pp. 509–534.

[6] Mohr, D., and Wierzbicki, T., 2001, “Shear Folding of Sandwich Box Columns Under Axial Loading,” MIT Impact and Crashworthiness Lab Report #58.

[7] Mohr, D., and Wierzbicki, T., 2003, “Shear Folding of Soft-Core Sandwich Profiles,” Int. J. Mech. Sci., **45**(2), pp. 253–271.

[8] Gustafsson, R. N. G., 2000, “Ultralight Stainless Steel Sandwich Materials—HSSA,” in *Sandwich Construction* Vol. 5, H. R. Meyer-Piening and D. Zenkert, ed., EMAS, pp. 169–176.

[9] Markaki, A. E., and Clyne, T. W., 2003, “Mechanics of Thin Ultra-light Stainless Steel Sandwich Sheet Material Part I. Stiffness,” Acta Mater., **51**, pp. 1341–1350.

[10] Mohr, D., and Straza, G., 2005, “Development of Formable All-Metal Sandwich Sheets for Automotive Applications,” Adv. Eng. Mater., **7**(4), pp. 243–246.

[11] Santosa, S. P., and Wierzbicki, T., 1999, “The Concept of Double-Walled Sandwich Columns for Energy Absorption,” Int. J. Crashworthiness, **4**(2), pp. 175–197.

[12] Abaqus, 2003, *Abaqus Reference Manuals*, Abaqus Inc, Providence, RI.

[13] LSTC, 2003, *LS-DYNA Reference Manuals*, Livermore Software Technology Corporation, Livermore, CA.

[14] Abramowicz, and Jones, N. A., 1986, “Dynamic Progressive Buckling of Circular and Square Tubes,” Int. J. Impact Eng., **4**(4), pp. 243–270.

[15] Jamjiam, M., Sackman, J. L., and Goldsmith, W., 1994, “Response of an Infinite Plate on a Honeycomb Foundation to a Rigid Cylindrical Impactor,” Int. J. Impact Eng., **15**(3), pp. 183–200.

Transient Growth and Stick-Slip in Sliding Friction

Norbert Hoffmann
Institute of Mechanics and Ocean Engineering,
Hamburg University of Technology,
Eissendorfer Strasse 42,
D-21073 Hamburg, Germany
e-mail: norbert.hoffmann@tuhh.de

In linearly stable dynamical systems the non-normality of the corresponding linear operator is known to lead to transient growth of characteristic quantities, like, e.g., energy. For sliding friction systems it is shown in the present paper that this transient growth also applies to the sliding velocity and that under certain conditions this mechanism can lead to stick-slip limit-cycles in linearly stable system configurations.
[DOI: 10.1115/1.2165233]

1 Introduction

In fluid dynamics one of the most puzzling questions is still the mechanism for transition to turbulence in systems which do not show linear instability of the laminar state. The most prominent representatives of this type of systems are surprisingly two of the most well-known flow configurations: Poiseuille flow due to a constant pressure gradient in a circular pipe and plane Couette flow due to the shearing of parallel plates. For these flow configurations transition from laminar to turbulent flow has experimentally been well mapped out. From a stability point of view, however, both laminar flows are stable, which means that there is no bifurcation sequence characterizing the initial stages leading finally to the highly irregular turbulent flow situations encountered experimentally.

Caused by this lack of understanding starting in the 1990s a number of investigations (e.g., Refs. [1–4]) focused on properties of the linearized systems' operators. It turned out that due to the non-normality of the operators, i.e., basically the eigenvectors cannot be chosen orthogonally onto each other, the linearized equations do show characteristic and marked transients of, e.g., the system's total energy content. As a consequence the hypothesis was posed that these transients generically cause small perturbations to grow to such a size that nonlinearities can become active and a transition from the laminar state to dynamical states dominated by nonlinearity is accomplished. This route to turbulence has then been discussed very controversially, since of course strictly speaking there always is a basin of attraction for whatever fixed point, limit-cycle, or generic solution, and the boundaries of this basin of attraction of course cannot be correlated in a strict mathematical sense to the rather intuitive energy-growth reasoning brought into discussion. On the other hand, from a practical point of view basin boundaries are at least very hard to map out, such that the transient growth transition can at least be taken as an engineering approach to understanding and predicting when the laminar flow gives way to more complex flows, up to turbulence.

Rather recently this transient growth phenomenon has been investigated in systems from fluid-structure interaction [5] and also first investigations of transiently energy amplifying beatings in sliding friction systems have already been conducted [6]. It is the purpose of the present text to investigate in a generic fashion what role transient growth plays in sliding friction systems that are prone to mode-coupling instability (refer to Refs. [7–9] for details on the instability mechanisms themselves), especially in relation to the occurrence of stick-slip limit-cycles. The present work is organized as follows: First a minimal two-degree-of-freedom model is set up to allow theoretical and numerical investigation of mode-coupling instability and the corresponding transition to stick-slip limit-cycles. Then typical properties of stick-slip limit-cycles in the system are presented, with a special note on subcritical limit-cycles in the case of the kinetic coefficient of friction exceeding the static one. Consequently exemplary transition to a stick-slip limit cycle in a linearly stable, subcritical system configuration is shown and discussed. Finally a technique for determining the worst case amplification of the system's tangential velocity component is presented.

2 The Model Problem

Since we are to investigate effects of systems which are prone to mode-coupling instability, the simplest model to set up is a two-degree-of-freedom lumped mass model. A graphical interpretation of the model used is given in Fig. 1. The model may be thought of as a single point mass sliding over a conveyor belt moving with constant speed v. The mass is mainly held in position by two linear springs K_x and K_z parallel and normal to the belt surface. Note that K_z may be regarded as the physical contact stiffness between the objects in relative sliding motion. Moreover, there is another linear spring k (oriented at an oblique angle of 45 deg relative to the normal direction) leading to off-diagonal entries in the model's stiffness matrix. The mass is loaded with a constant normal load F_n. For the friction an Amontons-Coulomb model is assumed, where in the case of sliding the frictional force F_t may be (choosing the coordinate system such that $z=0$ corresponds to the system just first making contact to the belt) directly derived from the compression of the vertical spring as $F_t = -\mu_k K_z z\,\mathrm{sign}(1-\dot{x}/v)$ with μ_k as kinetic coefficient of friction taken to be constant.

In the case of static friction the friction law reads $|F_t| \leqslant \mu_s K_z |z|$, where μ_s denotes the coefficient of static friction. Assuming that in the course of the dynamical evolution contact always remains closed the equations of motion in the sliding case read

$$\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}\begin{pmatrix} \ddot{x} \\ \ddot{z} \end{pmatrix} + \begin{bmatrix} C_x & 0 \\ 0 & C_z \end{bmatrix}\begin{pmatrix} \dot{x} \\ \dot{z} \end{pmatrix} + \begin{bmatrix} K_x + \frac{1}{2}k & -\frac{1}{2}k \\ -\frac{1}{2}k & K_z + \frac{1}{2}k \end{bmatrix}\begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} -\mu K_z z\,\mathrm{sign}\left(1-\frac{\dot{x}}{v}\right) \\ F_n \end{pmatrix}, \tag{1}$$

where linear viscous damping has been assumed. The equations for the sticking case can be derived analogously by taking into account the restraint $\dot{x}=v$. The system may be nondimensionalized in the usual way by using a time-scale T and a length-scale l to

---

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received December 11, 2004; final manuscript received November 18, 2005. Review conducted by I. Mezic. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of applied Mechanics, Department of Mechanical and Environmental Engineering, University of California – Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME Journal of Applied Mechanics.

Copyright © 2006 by ASME

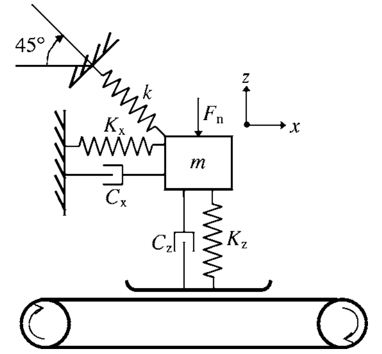


Fig. 1 Two-degree-of-freedom model

obtain dimensionless variables $t=T\hat{t}$ and $x=l\hat{x}$, $z=l\hat{z}$. Inserting these variables into Eq. (1), using $T=\sqrt{m/K_z}$ and $l=vT$ and omitting the carets of the non-dimensionalized variables for convenience yields—again for the sliding case—the nondimensionalized equations

$$\begin{pmatrix} x'' \\ z'' \end{pmatrix} + \begin{bmatrix} \gamma_x & 0 \\ 0 & \gamma_z \end{bmatrix} \begin{pmatrix} x' \\ z' \end{pmatrix} + \begin{bmatrix} \Omega_x^2 & -\Omega^2 \\ -\Omega^2 & \Omega_z^2 \end{bmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} -\mu_k z\,\mathrm{sign}(1-x') \\ \alpha \end{pmatrix} \tag{2}$$

with $\gamma_x=TC_x/m$, $\gamma_z=TC_z/m$, $\Omega_x=T\sqrt{(K_x+k/2)/m}$, $\Omega_z=T\sqrt{(K_z+k/2)/m}$, $\Omega=T\sqrt{k/2m}$, $\alpha=T^2F_n/ml$, where the prime denotes the derivative with respect to the new time scale and again the equations for the sticking case can be derived analogously.

From Eq. (2) the static solution corresponding to the steady sliding state may be calculated—by setting all temporal derivatives to zero—in a straightforward manner as

$$x_0 = \frac{\alpha[\Omega^2-\mu_k]}{\Omega^2[-\Omega^2+\mu_k]+\Omega_x^2\Omega_z^2}, \quad z_0 = \frac{\alpha\Omega_x^2}{\Omega^2[-\Omega^2+\mu_k]+\Omega_x^2\Omega_z^2} \tag{3}$$

Now Eq. (2) may be linearized around this state (3) to obtain the linear equations valid for (small) perturbations of the steady sliding state. All linear terms of Eq. (2) of course remain unchanged and the sign$(1-x')$ term reduces to 1 since for small oscillations around the steady sliding state the point mass has such a low velocity that it does not reach the belt velocity.

$$\begin{pmatrix} x'' \\ z'' \end{pmatrix} + \begin{bmatrix} \gamma_x & 0 \\ 0 & \gamma_z \end{bmatrix} \begin{pmatrix} x' \\ z' \end{pmatrix} + \begin{bmatrix} \Omega_x^2 & -\Omega^2+\mu_k \\ -\Omega^2 & \Omega_z^2 \end{bmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = 0 \tag{4}$$

One should note that although Eq. (4) has been obtained by a formal linearization procedure, Eq. (4) is a full representation of Eq. (2) except for stick-slip effects. As a consequence results of Eq. (4) are exact solutions of Eq. (2) up to the point when stick-slip effects take place.

In the following we will consider Eq. (4) to determine the linear stability of the steady sliding state and to investigate transiently growing quantities up to the point when the system nonlinearities, i.e., in this case stick-slip, sets in. To capture the stick-slip limit-cycles, time integrations of the piecewise linear system composed of the equations given above and the equations for the sticking regime are conducted. Since we are mainly interested in the fundamental effects, we do not intend to perform parameter studies and for simplicity the parameters will be set to the arbitrary but fixed values

$$\gamma_x = 0.02, \quad \gamma_z = 0.04, \quad \Omega_x = 0.9, \quad \Omega_z = 1.1, \quad \Omega = 0.5,$$
$$\alpha = -10.0 \tag{5}$$

Subsequently these values will be used for all numerical investigations and it will turn out that all phenomena of the type to be studied can be investigated with them.

3 Stability of Steady Sliding and Subcritical Stick-Slip Limit-Cycles

This section serves to show that whenever $\mu_s > \mu_k$ stick-slip limit-cycles do exist for parameter values for which the steady sliding state is linearly stable. To better understand the setting of the problem solutions of the homogenous equations (4) can be obtained using the exponential ansatz

$$\begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} \exp(st), \quad s = \sigma + i\omega \tag{6}$$

which leads to a generalized quadratic eigenvalue problem. The results of the corresponding eigenvalue analysis are shown in Fig. 2. Obviously the system has two distinct eigenvalues that approach each other with increasing μ_k. At $\mu_k=\mu_k^c=0.40$ the growth rate σ becomes positive for the less damped mode and for $\mu_k \geqslant \mu_k^c=0.40$ instability results. This is the typical behavior for what is usually called mode-coupling instability (refer to Refs. [8,9] for further aspects of the mode-coupling instability).

Figure 3 shows typical time-series results from time integrations in the linearly unstable regime with given initial conditions. It is clear that after a period of exponential growth sticking limits

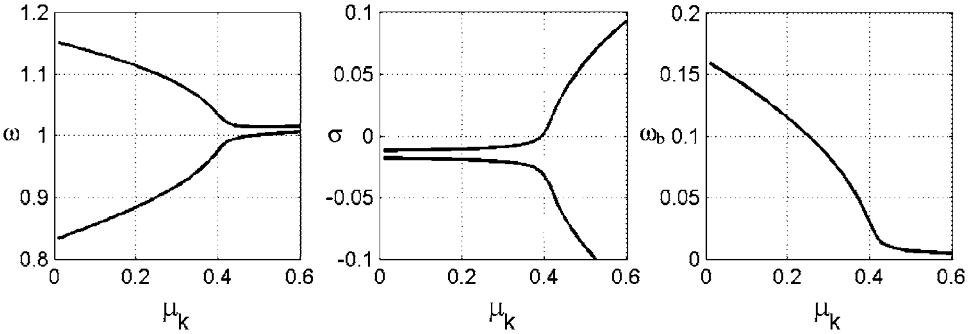


Fig. 2 Results of eigenvalue analysis of the model system. Oscillation frequencies ω (left), growth rates σ (middle), and beating frequency $(\omega_1-\omega_2)/2$ (right) as functions of the friction coefficient μ_k.

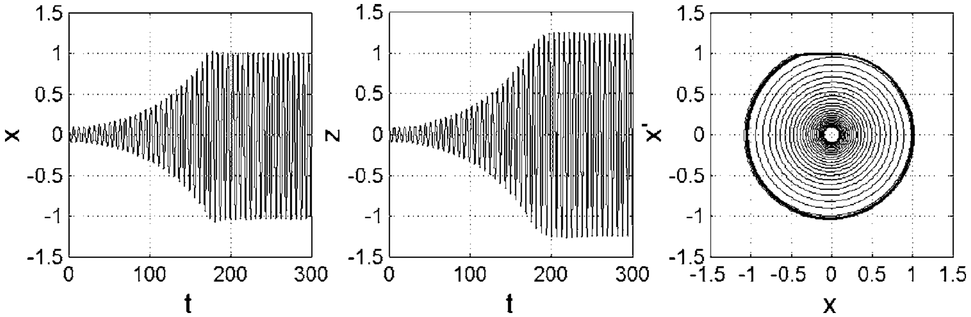


Fig. 3 Results of direct time-integration of the nonlinear equations in the unstable regime at $\mu_k=0.42$ with $(x,z,x',z')=(0.1,0.1,0,0)$ as initial conditions. Both $x(t)$, $z(t)$ and a phase space plot of x' versus x are shown.

the amplitude of the vibration and a typical stick-slip limit-cycle results (refer to Ref. [10] for further aspects of supercritical limit-cycles in systems like the one described).

Now in the case of $\mu_s > \mu_k$ numerically the limit-cycles can be traced in the space of the control parameters. To characterize the limit-cycles, the average vibrational energy, defined by $E = 1/T\int_0^T E(t)dt$ with

$$2E(t) = (x',z')\begin{pmatrix}x'\\ z'\end{pmatrix} + (x,z)\begin{bmatrix}\Omega_x^2 & -\Omega^2\\ -\Omega^2 & \Omega_z^2\end{bmatrix}\begin{pmatrix}x\\ z\end{pmatrix} \qquad (7)$$

has been used. Figure 4 shows the results. For computational convenience the ratio of μ_s and μ_k has been kept constant, and for several plausible values of this ratio the limit-cycles have been traced into the regime of linear stability of the steady sliding state, i.e., toward values of μ_k with $\mu_k < 0.4$. The result is very clear: When $\mu_k = \mu_s$, limit-cycles exist only supercritically for friction coefficients μ_k larger than the critical friction coefficient μ_k^c. When $\mu_s > \mu_k$ however, stable (otherwise they would not be accessible by time-integration) stick-slip limit-cycles exist far into the subcritical regime of $\mu_k < 0.40$.

First of all this result shows very impressively the limitations of linear stability analysis: only in the case of $\mu_k = \mu_s$ linear stability analysis is able to find the separating line between existence and nonexistence of finite-amplitude states. Whenever $\mu_s > \mu_k$, the linear stability analysis of the steady sliding state does not assure the absence of limit-cycle dynamics. Then one should note that this result is—from an intuitive point of view—not that surprising as one could expect. It is rather obvious that in the case of $\mu_s > \mu_k$ when the sticking phase is terminated, the system has gained some potential energy due to being held "in stick". Even if the following vibration dissipates some energy, as long as this energy dissipation is not too large, after almost one oscillation period, stick should be reached again. Only when the dissipation during sliding is larger than the energy input due to the sticking phase should steady sliding asymptotically be approached. This is of course exactly what can be seen in Fig. 4.

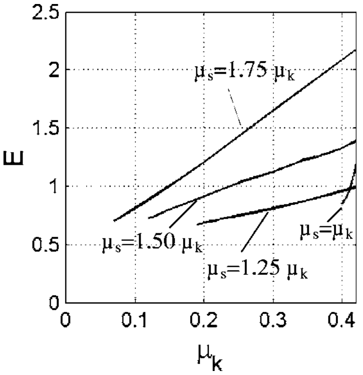


Fig. 4 Average vibrational energy E of stick-slip limit-cycles in the subcritical regime for different ratios of μ_s/μ_k. All solutions below $\mu_k=0.4$ are subcritical.

Finally one should now come to the central question of the present paper: in the subcritical regime, where the steady sliding is linearly stable, i.e., stable with respect to small perturbations, what could lead the system into the limit-cycle mode? How large does an initial perturbation have to be? The answer to these questions might—at least to some extent—be found in the phenomenon of transient growth.

4 Transient Growth Leading to Stick-Slip

In Fig. 5 the middle and the right columns show time-integrations of the linearized and the full nonlinear equations (with $\mu_s = 1.5\mu_k$) for initial conditions leading to maximal amplification of the initial tangential velocity component, that for maximal amplification just exceeds the belt velocity (it will be shown in the following how this initial condition can be obtained numerically). For the linear equations obviously for small μ_k the temporal evolution clearly resembles a decaying exponential. For μ_k approaching $\mu_k^c = 0.4$ however substantial transient amplification of the tangential velocity component can be observed. Using these same initial conditions in the full nonlinear equations, the transient amplification leads to a stick-slip limit-cycle. This shows clearly that even for initial conditions with rather small components in the tangential velocity for slightly subcritical parameter values, transient growth can lead to stick modes being activated.

To quantitatively determine the size of initial conditions that could potentially trigger limit-cycles, a mathematical technique to determine those initial conditions (called optimal in the following) that yield extremal transient amplification in the tangential velocity component after a certain period of time is being derived in the following. It should be noted that the present technique is sort of an extension or modification of an analogous technique to determine extremal amplification of the system's total vibrational energy content, compare Ref. [6]. In fluid dynamics transition to turbulence has traditionally—due to the quadratic nonlinearity of the advection term—been attributed to some sort of energy level in the flow being reached. In the present case of stick-slip oscillations however it seems much more plausible to assume limit-cycles to set in as soon as the tangential velocity of the oscillating mass first reaches the belt velocity, which yields a necessary condition for a sticking mode being possible. Therefore the technique has to devise a solution to the question how much the tangential velocity can transiently grow maximally before asymptotic decay. An approach to this problem is given in the following.

To simplify the mathematical treatment we first rewrite Eq. (4) as a system of first order in temporal derivatives

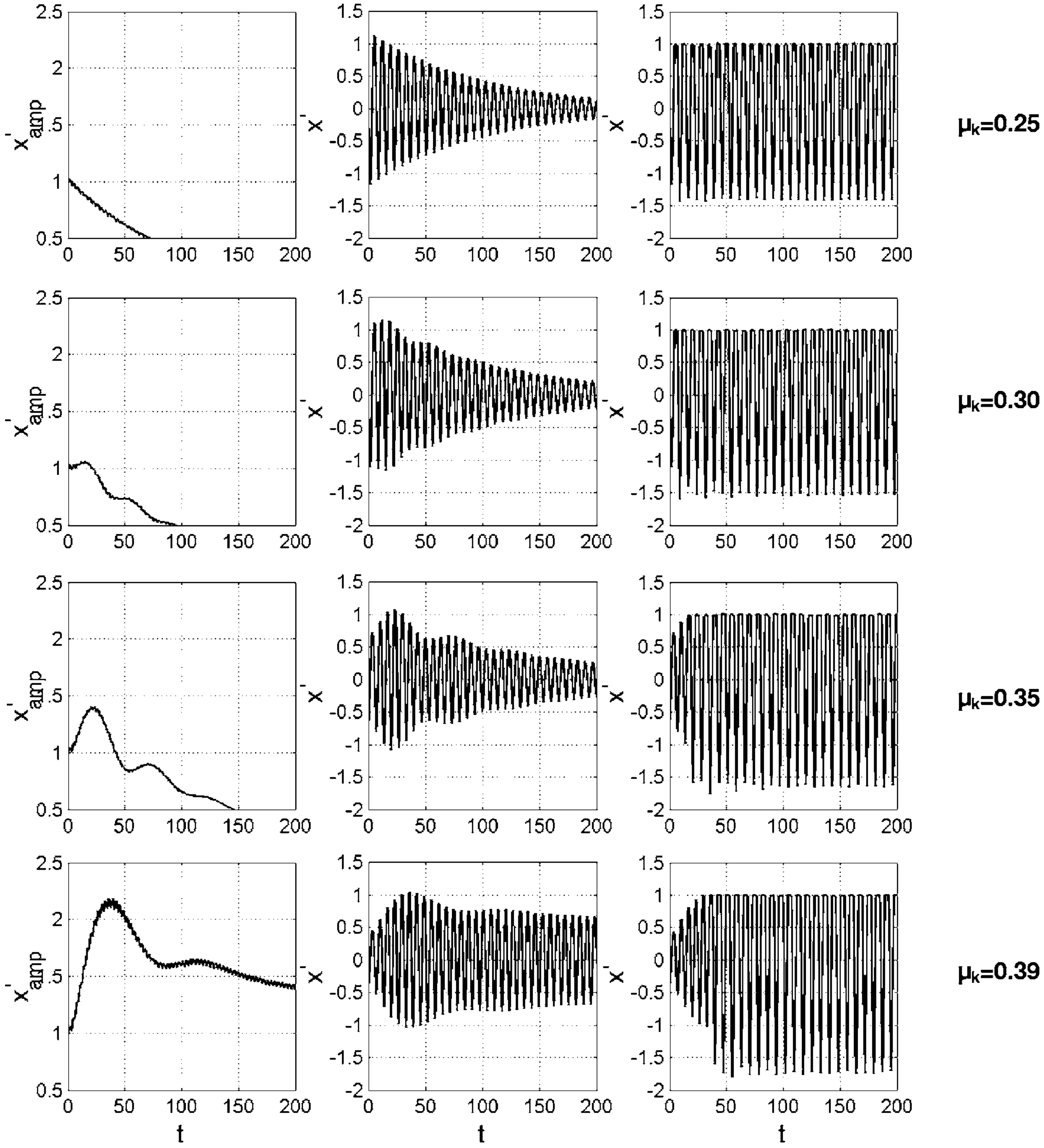


Fig. 5 Maximum amplification x'_{amp} versus time t for different values of μ_k (left column). Time integration results for optimal initial conditions, i.e., initial conditions leading to extremal amplification of x', scaled such that there is a time for which x' slightly exceeds 1. Middle column: results for linearized equations, right column: results for fully nonlinear equations.

$$\frac{d}{dt}\begin{pmatrix} x \\ z \\ x' \\ z' \end{pmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\Omega_x^2 & \Omega^2-\mu_k & -\gamma_x & 0 \\ \Omega^2 & -\Omega_z^2 & 0 & -\gamma_z \end{bmatrix}\begin{pmatrix} x \\ z \\ x' \\ z' \end{pmatrix} \tag{8}$$

for which we will use the obvious abbreviated notation

$$\mathbf{x}' = \mathbf{A}\mathbf{x} \tag{9}$$

from now on. Since the system is linear, homogeneous, and time-invariant a complete set of solutions can be obtained by the usual exponential ansatz $\mathbf{x}=\mathbf{x}^0\exp(st)$ resulting in the eigenvalue problem

$$\mathbf{A}\mathbf{x}^0 = s\mathbf{x}^0 \tag{10}$$

The general solution of Eq. (9) may then be written down as

$$\mathbf{x}(a,t) = \sum_i a_i \mathbf{x}_i^0 \exp(s_i t) = \boldsymbol{\psi}_t \mathbf{a}, \tag{11}$$

where $\mathbf{a}$ is the vector of expansion coefficients defined by decomposing the initial conditions into the eigenvectors arising from the eigenproblem and s_i and $\mathbf{x}_i^0$ represent eigenvalues and eigenvectors, respectively. Note that for the ease of notation the temporal evolution of the system may be represented as the product of the matrix $\boldsymbol{\psi}_t$ containing the eigenvectors marched forward in time and the vector $\mathbf{a}$, which for simplicity we will call the initial conditions in the following.

Now consider a mapping that maps a given state of the system onto the real numbers, e.g., to determine the system's content of vibrational energy or some other scalar at a given time

$$S^t = \mathbf{x}^{*T}\mathbf{N_S}\mathbf{x} \tag{12}$$

Here $\mathbf{N_S}$ is a hermitian matrix ($\mathbf{N_S}=\mathbf{N_S}^{*T}$) to ensure that S^t is a real quantity. Using Eq. (11) to express $\mathbf{x}$, Eq. (12) may be rewritten as

$$S^t = \mathbf{a}^{*T}\boldsymbol{\psi}_t^{*T}\mathbf{N_S}\boldsymbol{\psi}_t\mathbf{a} = \mathbf{a}^{*T}\mathbf{N}_\mathbf{S}^t\mathbf{a} \tag{13}$$

where $\mathbf{N}_\mathbf{S}^t=\boldsymbol{\psi}_t^{*T}\mathbf{N_S}\boldsymbol{\psi}_t$ has been introduced to collect all temporal dependencies.

Now it is the mathematical task to maximize—for a given time t—the value of S^t by varying the initial conditions $\mathbf{a}$ appropriately. Since the problem considered is linear, we of course have to add an amplitude restraint to this optimization task. For that purpose we additionally require a unit magnitude of the initial vibrational energy E^t at $t=0$, i.e., $E^0=1$, where E^t can of course again be expressed in the form of a scalar product as $E^t=\mathbf{x}^{*T}\mathbf{N_E}\mathbf{x}$. Formally the optimization task then reads

$$S^t_{\max} = \max_{\mathbf{a}}\{S^t - \lambda(E^0 - 1)\} = \max_{\mathbf{a}}\{\mathbf{a}^{*T}\mathbf{N_S}{}^t\mathbf{a} - \lambda(\mathbf{a}^{*T}\mathbf{N_E}{}^0\mathbf{a} - 1)\} \tag{14}$$

where the Lagrange parameter λ has been introduced to take the restraint $E^0=1$ into account. This optimization problem is obviously equivalent to the generalized eigenvalue problem

$$\mathbf{N_S}{}^t\mathbf{a} = \lambda\mathbf{N_E}{}^0\mathbf{a} \tag{15}$$

If $\mathbf{N_E}{}^0$ can be inverted, the problem may be transformed into the standard eigenvalue problem

$$(\mathbf{N_E}{}^0)^{-1/2}\mathbf{N_S}{}^t(\mathbf{N_E}{}^0)^{-1/2}\mathbf{b} = \lambda\mathbf{b} \quad (\mathbf{N_E}{}^0)^{1/2}\mathbf{a} = \mathbf{b} \tag{16}$$

The eigenvector corresponding to the largest eigenvalue will then give those initial conditions ("optimal initial conditions") which yield maximal values of S^t after the specified time span t and the eigenvalue itself gives nothing but the ratio of S^t and E^0, as can readily be seen from Eqs. (15) and (16) when the eigensolutions are put in. If E^0 is taken to be of unit amplitude then the eigenvalues of course just give directly the values of S^t.

After we have now set up the mathematical framework for determining optimal initial conditions for general mappings (12), we will subsequently consider the specific mappings yielding the instantaneous total vibrational energy E^t, as defined in Eq. (7), and the square of the instantaneous tangential velocity

$$2\mathbf{N_E} = \begin{bmatrix} \Omega_x^2 & -\Omega^2 & 0 & 0 \\ -\Omega^2 & \Omega_z^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{N_S} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{17}$$

Exemplary results are shown in the left column of Fig. 5. x'_{amp} here denotes the maximal values that the tangential velocity can reach at a given time t, assuming that the initial condition at $t=0$ did have unit amplitude in vibrational energy E^0 and that the initial conditions were "optimal". It turns out that for kinetic friction coefficients μ_k smaller than about 0.2, $x'_{\max}(t)$ decays monotonically over time, whereas when μ_k approaches the border of instability at μ_k^c there is transient growth in x'_{amp} before asymptotic decay; in these cases the maximum value of x'_{amp} is reached at a certain time $t_{\max}$ and Fig. 6 shows these overall maximal values $max(x'_{\mathrm{amp}})$ as a function of the kinetic friction coefficient μ_k. It shows that for $\mu_k<0.2$ the potential tangential velocity has its largest value for $t=0$ and decays for larger times. When μ_k approaches μ_k^c however, the maximal tangential velocity is up to a factor of 3.5 higher than the tangential velocity in the initial condition. This result conforms well to the results from direct time-integrations corresponding to "optimal initial conditions".

5 Summary and Conclusions

First of all it has been shown that in sliding friction systems with friction laws of the Amontons-Coulomb type subcritical stick-slip limit-cycles do exist when the static friction coefficient exceeds the kinetic one, i.e., even though a linear stability analysis of the steady sliding state indicates system stability, stable limit-cycles do exist. This observation alone has significant implications for the application side: when the borderline between steady sliding and stick-slip limit-cycle dynamics has to be determined, in the described cases the existing bistability of solutions has to be taken into account. A linear stability analysis of the steady sliding state alone does not yield sufficient conditions for the nonexistence of stick-slip. In fact the full nonlinear dynamics has to be taken into account.

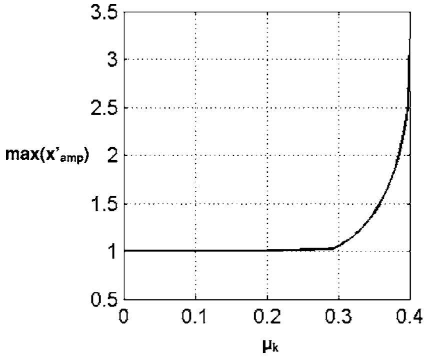


Fig. 6 Maximum amplification of x', max(x'_{amp}) versus kinetic friction coefficient μ_k

Next it has been shown that for subcritical parameter combinations transient growth of small initial conditions can lead to strong amplification of the inherent tangential velocity. When this tangential velocity reaches the belt velocity, typically stick-slip limit-cycles are initiated. The given mathematical technique to determine maximal amplification of the tangential velocity, given an initial condition with unit amplitude vibrational energy, can therefore be of substantial help in judging if the system will eventually—due to the permanent presence of disturbances evolving in time—end up in a stick-slip limit-cycle, even though the steady sliding state is linearly stable. It turns out that by calculating the maximal amplifications as a function of the control parameters, regimes with and without transient growth can be identified.

At this point a comment on the relation of the present approach to the approach of a basin of attraction seems appropriate. Of course strictly speaking the present approach is rather informal. In fact each attractor, i.e., also the stick-slip limit-cycle considered here, has a basin of attraction. Therefore for all the initial conditions considered, strictly speaking one would only have to check whether the initial condition is within the basin of attraction; then of course stable limit-cycle dynamics will be reached asymptotically. From a practical point of view however the boundaries of basins cannot usually be mapped out, merely due to the dimensionality of the system's phase space. It is this problem that might make the present approach useful to engineering applications: although the transient amplification up to the sticking condition will of course always remain a necessary condition for the onset of stick-slip only, nevertheless an investigation of the strength of the temporal growth might give valuable indications, e.g., with respect to system design.

Future work on the topic should therefore focus on the following fields: first of all experimental evidence and plausibility should be gained for stick-slip oscillations excited through transients. Especially since the approach as a whole is somewhat heuristic, the experiment should serve as the best test on the possibilities and limitations of the approach. In addition a more conceptual investigation should take into account the stochasticity of the ubiquitous perturbations (i.e., noise) and focus on the probabilities for reaching stick-slip limit-cycles in the spirit of a stochastically excited first-time-transit problem, complicated through the phenomenon of transient growth. Here the basic question has to be: how long does it take, given a certain noise level or quality, to reach stick-slip through transient growth in subcritical parameter situations. After this experimental and conceptual work it should become more clear, whether the techniques developed in the present paper might—especially due to the low numerical effort necessary—form an attractive tool for the everyday engineering application in the context of friction self-excited oscillations.

References

[1] Farrell, B. F., 1988, "Optimal Excitation of Perturbations in Viscous Shear Flow," Phys. Fluids, **31**, pp. 2093–2102.

[2] Waleffe, F., 1995, "Transition in Shear Flows. Nonlinear Normality Versus Non-Normal Linearity," Phys. Fluids **7**, pp. 3060–3066.

[3] Grossmann, S., 1995, "Wie Entsteht Eigentlich Turbulenz?," Phys. Bl., **51**, pp. 641–646.

[4] Gebhardt, T., and Grossmann, S., 1994, "Chaos Transition Despite Linear Stability," Phys. Rev. E, **50**, pp. 3705–3711.

[5] Schmid, P. J., and de Langre, E., 2003, "Transient Growth Before Coupled-Mode Flutter," ASME J. Appl. Mech. **70**, pp. 894–901.

[6] Hoffmann, N., and Gaul, L., 2004. "Non-Conservative Beating in Sliding Friction Affected Systems: Transient Amplification of Vibrational Energy and a Technique to Determine Optimal Initial Conditions," Mech. Syst. Signal Process., **18**, pp. 611–623.

[7] Popp, K., Rudolph, M., Krger, M., and Lindner, M., 2002, "Mechanisms to Generate and to Avoid Friction Induced Vibrations," VDI-Ber. **1736**, pp. 1–15.

[8] Hoffmann, N., Fischer, M., Allgaier, R., and Gaul, L., 2002, "A Minimal Model for Studying Properties of the Mode-Coupling Type Instability in Friction Induced Oscillations," Mech. Res. Commun. **29**, pp. 197–205.

[9] Hoffmann, N., and Gaul, L., 2003, "Effects of Damping on Mode-Coupling Instability in Friction Induced Oscillations," Z. Angew. Math. Mech. **83**, pp. 524–534.

[10] Hoffmann, N., Bieser, S., and Gaul, L., 2004, "Harmonic Balance and Averaging Techniques for Stick-Slip Limit-Cycle Determination in Mode-Coupling Friction Self-Excited Systems," Techn. Mech. **24**, pp. 185–197.

Complex Flow Dynamics in Dense Granular Flows—Part I: Experimentation

Piroz Zamankhan

Mohammad Hadi Bordbar

Laboratory of Computational Fluid
and BioFluid Dynamics,
Lappeenranta University of Technology,
Lappeenranta 53851,
Finland

By applying a methodology useful for analysis of complex fluids based on a synergistic combination of experiments, computer simulations, and theoretical investigation, a model was built to investigate the fluid dynamics of granular flows in an intermediate regime where both collisional and frictional interactions may affect the flow behavior. In Part I, the viscoelastic behavior of nearly identical sized glass balls during a collision have been studied experimentally using a modified Newton's cradle device. Analyzing the results of the measurements, by employing a numerical model based on finite element methods, the viscous damping coefficient was determined for the glass balls. Power law dependence was found for the restitution coefficient on the impact velocity. In order to obtain detailed information about the interparticle interactions in dense granular flows, a simplified model for collisions between particles of a granular material was proposed to be of use in molecular dynamic simulations, discussed in Part II. [DOI: 10.1115/1.2165234]

1 Introduction

Complex fluids refer to those fluids that do not exhibit simple fluid mechanical behavior. In contrast, simple fluids are fluids with a straightforward description, including many common fluids such as water, which may all be described quite elegantly with the same equation [1,2]. The understanding of complex fluid flows is important both from a fundamental and industrial point of view. Complex fluids comprise a rich variety of systems; including granular materials, porous media, fractures, colloidal suspensions, surfactant mixtures, polymeric liquids, bimolecular assemblies, electro-hydrodynamics flow, and many others.

Collisions between grains in a granular flow, which have been studied since Newton's time, still pose many difficulties and unanswered questions. One of the simplest of these problems is the binary collision between two monosized, spherical grains that involves repulsive rigid elastic interactions as well as dissipative frictional contacts. A binary collision represents the simplest of the multibody interactions in a dense granular flow. The frictional interaction represents the most fundamental difference between a granular system and a molecular system. Note that even a highly polished grain has surface roughness on many different length scales.

Many diverse examples of granular flows exist, ranging from flow of grains in a silo [3] to planetary rings [4]. Flow dynamics of granular materials may be categorized in the complex fluid families. As an example, in these systems solid-like and fluid-like qualities may exist, side-by-side [5].

This behavior could be due to a complex nonlinear elasticity exhibited collectively by an aggregate of grains that governs the flow properties of the system.

In this light, developing realistic models for the granular collisions, by which slow collisions for rough as well as smooth grains can be described, would be invaluable in the understanding of the physics of deformation and flow of granular materials, especially in industrial systems.

A contact force model such as $F=k\delta^{\gamma}+c\delta^{\xi}d\delta/dt$ is often adopted to capture the key features of granular inelastic interactions when the impact velocity is much less than the speed of sound in the grain material [6]. In discussing the aforementioned expression, it has been assumed that the elastic restoring force is proportional to the displacement, and the dissipative force is proportional to the velocity.

In light of the above discussion, the exponent γ was found to be $3/2$ [7,8], consistent with that obtained from Hertzian quasi-static force [9]. However, the $\delta^{1/4}$ dependence of contact force on viscous dissipation observed in experiments [10] appears to be at odds with the $\delta^{1/2}$ scaling predicted by model developed in Refs. [7,8]. This observation might question justification for making the assumptions that the contact forces developed in solids such as that used in Ref. [10] are the sum of two terms, one proportional to the strain and the other proportional to the rate of change of strain with $\delta^{1/2}$ scaling factor.

Brilliantov et al. [8] compared the computed values of the normal restitution coefficients with those measured by Bridges, Hatez, and Lin [4], at different impact velocities. It has been shown in Ref. [8] that excellent agreement with the measurement data [4] can be obtained, given that a model with a spring and dashpot in parallel is used to describe viscoelastic behavior of the ice particles with little or no ductility. Note that the model predictions could be improved using a larger value than $1/2$ for ξ with the exponent of the contact force model given above.

The numerical results reported in Ref. [8] suggested a nonlinear dependence of the coefficient of restitution on the impact velocity, consistent with the findings of Bridges et al. [4] and Gugan [11]. Surprisingly, the aforementioned important feature of inelastic collisions remained completely undiscovered by Foester et al. [12].

Note that the aforementioned model for viscoelastic materials is quite useful in describing how internal friction may arise from processes of a viscous nature. However, it is too simplified to be quantitatively useful to predict a collision in which a number of different relaxation processes may take place simultaneously, in solids. Hence, a more complicated model has to be considered in order to predict collisions of the real solids having a continuous spectrum of relaxation times.

The objective of the current work is to provide an overview of impact analysis useful for developing a practical methodology for large-scale three-dimensional systems. To measure the viscoelastic coefficients of glass particles the classical collision experiment

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 3, 2005; final manuscript received November 11, 2005. Review conducted by K. M. Liechti. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

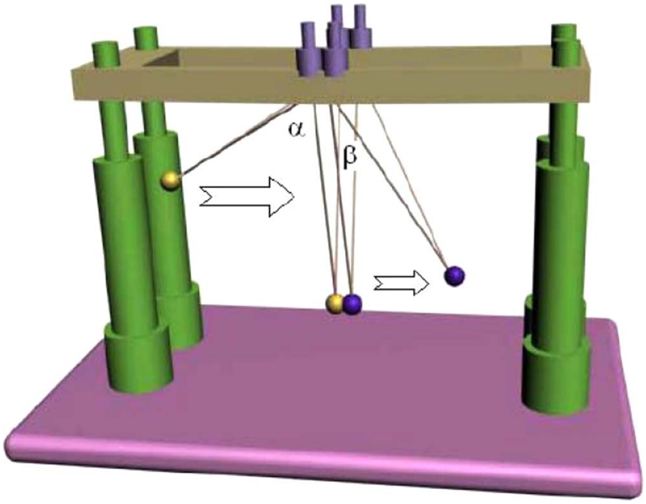


Fig. 1 Schematic of apparatus used for the Newton's cradle experiments. The balls used were nearly identical, spherical glass particles suspended from very thin threads. Here, the positions of the balls before, at, and after a collision are shown.

known as Newton's cradle was used. These quantities cannot be measured directly; therefore, a numerical approach based on finite element methods was developed to analyze collision processes in the Newton's cradle device. A numerical model based on Boltzmann's approach for viscoelastic solids was used to investigate the properties of impact processes, including the estimation of the viscoelastic coefficients. In addition, the results obtained were used to develop a simplified but more realistic model for the collision of grains, capable of reproducing the observed behavior of granular flows discussed in the subsequent paper.

2 Experimental Studies

Experimental setup. As stated in the proceeding section, granular materials may be treated as a collection of inelastic hard spheres. There are no attractive forces between particles but in each collision energy is lost. One of the simplest of these problems is the binary collision between spherical grains, which involves repulsive, rigid, elastic interactions as well as dissipative frictional contacts. A binary collision represents the simplest of the multibody interactions in a dense granular flow.

One aim of this study is to further investigate collisions between inelastic spherical particles. To this end the classical collision experiment known as Newton's Cradle was used, as shown schematically in Fig. 1. As illustrated in Fig. 1, two nearly identical, spherical glass particles of mass 13.3×10^{-3} kg were arranged by suspending each particle from a frame using two very thin threads. In order to ensure the elimination of the rotation of the balls their centers of mass were aligned very carefully. By releasing the first ball from an elevation an impact was generated which was captured using a high speed digitalized video camera (JVC GR-DVL9800 with a recording speed of 100 Hz) connected to a personal computer. Reflective tapes were used to mark the center of each ball. Note that it is important to have very bright, focused light sources closely aligned with the camera axis. Using the acquired video images the dropping height and the corresponding inclination angle with respect to the equilibrium position of the first ball, α, and the maximum post impact height attained by the other ball as well as its corresponding inclination angle, β, were determined.

To generate collisions at different impact velocities, the experiments were repeated using different elevations for the first ball. Moreover, to examine the effect of particle size on collision, the aforementioned balls were replaced by the lighter glass particles of mass 5.61×10^{-3} kg and similar experiments were conducted to determine the normal coefficient of restitution whose magnitude may be approximated by

$$e_n = 2\frac{\sin \beta/2}{\sin \alpha/2} - 1 \tag{1}$$

Table 1 Material properties

Material	Properties	Symbol	Value
Glass	Elastic modulus	E	6.3×10^{10} Pa
	Density	ρ	2390 kg/m^3
	Poisson's ratio	ν	0.244
	Chemical compositions	SiO_2	66.0%
		Nd_2O	16.0%
		CaO	7.0%
		Al_2O_3	5.0%
		Bi_2O_3	3.0%

In this case the impact velocity is given as $V_{\text{imp}} = 2\sqrt{gl}\sin \alpha/2$. In the first set of experiments, the sphere has a diameter of 22 mm and its physical properties as well as chemical compositions (provided by Sigmund Linder GmbH, Germany) are listed in Table 1. The diameter of lighter glass balls used in the second set of experiments was 16.5 mm. The sample results presented in Table 2 represent the degree of plasticity of the collisions in terms of the impact velocity. Note that values of $e_n = 1$ and $e_n = 0$ denote the idealized cases of perfectly elastic and plastic impact, respectively. The present results suggest a nonlinear dependence of the coefficient of restitution on the impact velocity, consistent with the findings in previous attempts [4–11].

Further examination of the impact problem studied in this section requires a quantitative description of the brittle behavior of glass particles. In the following section, a viscoelastic model will be presented characterized by a time dependent relation between stress and strain. Models for glass balls may be constructed by suitable combinations of spring and dashpots, where the dashpots introduce a viscous type resistance associated with internal friction in solids.

Note that neither the Maxwell nor Kelvin model represents the behavior of most viscoelastic materials, including the glass balls used in Newton's cradle device illustrated in Fig. 1. For example the Maxwell model [13] predicts that the stress asymptotically approaches zero when the strain is kept constant. On the other hand, the Kelvin model does not describe a permanent strain after unloading. In this study, a second spring is placed in series with the Kelvin model used in Ref. [8], to develop a model of linear viscoelastic material equivalent to a Maxwell model with a spring in series. The schematic of the three-parameter model used in this study is shown in Fig. 2.

The success of theoretical approach such as that discussed in Ref. [8] is limited for nonlinear problems because of difficulties of analysis. However, a numerical approach has become popular because collision processes can be studied under realistic situations, where the action of loads suddenly applied to a body such as a

Table 2 Measured values for the coefficient of restitution at different impact velocity, for a glass ball with a diameter of 22 mm

V_{imp} (m/s)	Coefficient of restitution
0.554	0.868
0.599	0.866
0.610	0.852
0.682	0.836
0.896	0.827
0.954	0.816

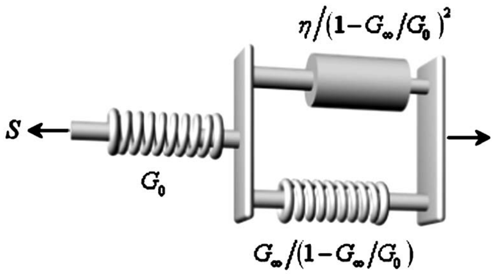


Fig. 2 The model consists of auxiliary spring in series with the Kelvin model

spherical particle may be propagated in a wave-like manner and its reflection at bounding surfaces could lead to vibrations. In addition, it is easy to control situations to investigate the properties of impact processes as well as to analyze nonlinear problems as detailed in the following.

3 Analysis

Rigid body motion in Newton's cradle. As depicted in Fig. 1, two nearly identical, spherical glass particles of mass m are arranged by suspending each particle from a frame using two very thin threads. The first glass ball was pulled aside and allowed to swing as a pendulum, resulting in its colliding with the second glass ball. The motion of the moving glass ball having diameter σ and mass m suspended from a weightless thread with a fixed length l in the field of gravitation, and subject to a drag force may be described as follows.

Suppose the mass is given a displacement θ_0, φ_0 in a spherical coordinates system (r,θ,φ) from the equilibrium position, namely, $\theta_{\mathrm{eq}}=\pi$, $\varphi_{\mathrm{eq}}=0$ and then released. The mass then approaches the equilibrium position with acceleration due to the action of resultant of the force of tension of thread T and the forces of gravity mg, and drag D due to air resistance. The components of acceleration of the glass ball in a Cartesian coordinates system (x,y,z) are given as [14]

$$A_x=-(l\dot{\theta}^2+l\dot{\phi}^2)\sin\theta\cos\phi+(l\ddot{\theta})\cos\theta\cos\phi-(l\ddot{\phi})\sin\theta\sin\phi$$
$$-2l\dot{\theta}\dot{\varphi}\cos\theta\sin\phi,$$

$$A_y=-(l\dot{\theta}^2+l\dot{\phi}^2)\sin\theta\sin\phi+(l\ddot{\theta})\cos\theta\sin\phi+(l\ddot{\phi})\sin\theta\cos\phi$$
$$+2l\dot{\theta}\dot{\varphi}\cos\theta\cos\phi,$$

$$A_z=-(l\dot{\theta}^2)\cos\theta-(l\ddot{\theta})\sin\theta. \tag{2}$$

Likewise, the components of resultant of the tension, gravity, and drag forces acting on the ball in the Cartesian coordinates system (x,y,z) are given as

$$F_x=-T_p\sin\theta\cos\phi-\frac{\pi}{8}\rho_f(\sigma l)^2(\dot{\theta}^2+\dot{\varphi}^2\sin^2\theta)^{1/2}$$
$$\times(\dot{\theta}\cos\theta\cos\varphi-\dot{\varphi}\sin\theta\sin\varphi)C_D$$

$$F_y=-T_p\sin\theta\sin\phi-\frac{\pi}{8}\rho_f(\sigma l)^2(\dot{\theta}^2+\dot{\varphi}^2\sin^2\theta)^{1/2}$$
$$\times(\dot{\theta}\cos\theta\sin\varphi+\dot{\varphi}\sin\theta\cos\varphi)C_D$$

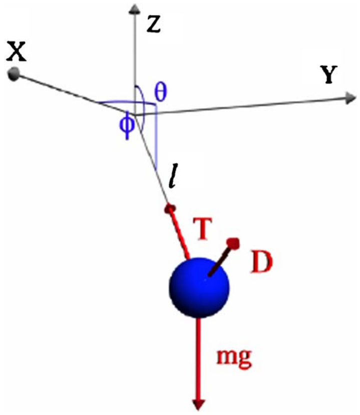


Fig. 3 Free-body diagram of a single pendulum in a spherical coordinates (r,θ,φ)

$$F_z=-mg-T_p\cos\theta+\frac{\pi}{8}\rho_f(\sigma l)^2(\dot{\theta}^2+\dot{\varphi}^2\sin^2\theta)^{1/2}\times(\dot{\theta}\sin\theta)C_D \tag{3}$$

where C_D represents the coefficient of drag for a spherical particle defined as [15]

$$C_D=\frac{24}{\mathrm{Re}_p}(1+0.1862\,\mathrm{Re}_p)+\frac{0.4373\,\mathrm{Re}_p}{7185.4+\mathrm{Re}_p} \tag{4}$$

Therefore, using the classical mechanics theory the equation of motion of the single pendulum as illustrated in Fig. 3 may be written as

$$\ddot{\theta}=\sin\theta\left(-\frac{D_z}{ml}+\frac{g}{l}\right)+\cos\theta\left(\frac{D_x}{ml}\cos\varphi+\frac{D_y}{ml}\sin\varphi+\dot{\varphi}^2\sin\theta\right)$$
$$\ddot{\varphi}=-\csc\theta\left(2\dot{\varphi}\dot{\theta}\cos\theta-\frac{D_y}{ml}\cos\varphi+\frac{D_x}{ml}\sin\varphi\right) \tag{5}$$

The parameters D_x, D_y, and D_z in Eq. (5) are given as

$$D_x=-\frac{\pi}{8}\rho_f(\sigma l)^2(\dot{\theta}^2+\dot{\varphi}^2\sin^2\theta)^{1/2}$$
$$\times(\dot{\theta}\cos\theta\cos\varphi-\dot{\varphi}\sin\theta\sin\varphi)C_D$$

$$D_y=-\frac{\pi}{8}\rho_f(\sigma l)^2(\dot{\theta}^2+\dot{\varphi}^2\sin^2\theta)^{1/2}$$
$$\times(\dot{\theta}\cos\theta\sin\varphi+\dot{\varphi}\sin\theta\cos\varphi)C_D \tag{6}$$

$$D_z=\frac{\pi}{8}\rho_f(\sigma l)^2(\dot{\theta}^2+\dot{\varphi}^2\sin^2\theta)^{1/2}\times(\dot{\theta}\sin\theta)C_D$$

Note that Eq. (5) is nonlinear and thus capable of producing chaotic behavior.

4 Particle Collision in Newton's Cradle

When two spherical particles are brought into contact they touch initially at a single point. As the particles deform, they come into contact in the vicinity of the initial contact point over an area that is small compared with the dimensions of the particles.

As depicted in Fig. 1, two nearly identical, spherical glass particles of mass m are arranged by suspending each particle from a frame using two very thin threads. In order to prevent the balls from rotating, their centers of mass were aligned very carefully. The first glass ball was pulled aside and allowed to swing as a pendulum, resulting in its colliding with the second glass ball. When two spherical particles are brought into contact they touch initially at a single point. As the particles deform, they come into contact in the vicinity of the initial contact point, over an area that is small compared with the dimensions of the particles. Note that the action of loads suddenly applied to a body is propagated from one place to another place in a wave-like manner. Any intensive impact loading could be accompanied by plastic deformation [16]. This category of impact problems will be discussed briefly in the following.

Numerical simulation of a collision in Newton's cradle—Elastic collisions. In the following, a numerical treatment is presented based on the Galerkin weighted residual method applied to the governing equations and boundary conditions to develop velocity-pressure finite element contact formulations. In order to validate the aforementioned numerical method, first the elastic impact of spherical particles are discussed for which analytical solutions, such as those of Hertz, are available for conditions for which the duration of contact is large compared to the period of the lowest natural frequency of either spheres where the bodies may be considered to be at the state of quasiequilibrium.

The general equations of motion in referential coordinates for an elastic continuum in absence of body forces may be given as [17]

$$\rho \ddot{u}_i = \sigma^{(\mathrm{el})}_{ij,j} \tag{7}$$

The displacements in a fixed rectangular Cartesian coordinate system $X_j (j=1,2,3)$ are denoted by $u_i = x_i - X_i = u_i(X_j, t)$ with kinematical relations $\ddot{u}_i = \partial / \partial t^2 x_i(X_j, t)$. Equation (7) with constitutive equation $\sigma^{(\mathrm{el})}_{ij} = \lambda(E_{kk})\delta_{ij} + 2GE_{ij}$ and geometric equations $E_{ij} = 1/2(S_{ij} + S_{ji})$ where S_{ij} represents the replacement gradient provide 15 equations for the 15 unknowns, comprising six stresses $\sigma^{(\mathrm{el})}_{ij}$, six strains E_{ij}, and three displacements u_i.

The set of equations given above has to be solved for appropriate boundary and initial conditions listed in the following

$$\begin{gathered} \sigma_{ij} n_j = T_i(t), \quad \text{on } \partial b_1 \\ (\sigma_{ij}|_{x_i^+} - \sigma_{ij}|_{x_i^-}) n_j = 0, \quad \text{on } \partial b_2 \\ x_i(X_j, 0) = X_i^0 \\ \dot{x}_i(X_j, 0) = V_i(X_j) \end{gathered} \tag{8}$$

The second condition in Eq. (8) represents the contact discontinuity given along an interior boundary ∂b_2 when $x_i^+ = x_i^-$. The Galerkin formulation may be obtained from Eqs. (7) and (8) by using the Gauss theorem. That is

$$\delta W = \int_{V_b} \rho \ddot{u}_i \delta u_i dV_b + \int_{V_b} \sigma_{ij} \delta u_{i,j} dV_b - \int_{\partial b_1} T_i \delta u_i dS_b = 0 \tag{9}$$

This equation is the weak form of the equilibrium equations. Spatial discretization in numerical analyses such as binary collision of spheres could be typically based on a single method such as Lagrange, Euler, a mixture of Lagrange and Euler (arbitrary Lagrange Euler), or mesh-free Lagrangian (smooth particle hydrodynamics). For the numerical simulation of the binary collision of spheres, each of the different solutions mentioned above has unique advantages and there is no single ideal numerical method that would be appropriate to the various regimes of a collision. In the present study the Lagrange method of space discretization is used, for which the numerical grid moves and deforms with the material.

The advantages of the aforementioned method are computational efficiency and the ease of incorporating complex material models. The disadvantage of the method used is that the numerical grids may become severely distorted in an extremely deformed region. This can cause adverse effects on the integration time step and accuracy. However, these problems can be somewhat overcome by applying numerical techniques such as rezoning.

The spatial discretization is performed by representing the fields using computational points in space connected with each other through grids [18]. Thus

$$u_i(X_j, t) = \sum_{p=1}^{N} \phi_p(\xi, \zeta, \eta) u_i^p(t) \tag{10}$$

Using M elements, the discretized form of Eq. (10) becomes

$$\sum_{k=1}^{M} \left\{ \int_{V_k} \rho \ddot{u}_i \Phi_i^k dV + \int_{V_k} \sigma_{ij}^k \Phi_{i,j}^k dV_b - \int_{\partial b_1} T_i \Phi_i^k dS \right\} = 0 \tag{11}$$

where $\Phi_i^k = \{\phi_1, \phi_2, \ldots, \phi_N\}_i^k$. These lead to matrix form of Eq. (11) given as

$$\sum_{k=1}^{M} \left\{ \int_{V_k} \rho \Pi^t \Pi a dV + \int_{V_k} B^t \sigma dV - \int_{\partial b_1} \Pi^t T dS \right\}_k = 0 \tag{12}$$

To demonstrate the application of the above-detailed numerical method, numerical simulations of the impact problem of two equally sized, smooth spheres are illustrated in Fig. 4(*a*). The sphere has a diameter of 22 mm and its physical properties as well as chemical compositions (provided by Sigmund Linder GmbH, Germany) are listed in Table 1. In the case illustrated in Fig. 4, the right sphere collides with the left one, which is initially stationary, at a velocity of $V_{\mathrm{imp}} = 0.56$ m/s. The impact velocity was found solving Eq. (5) for the moving sphere. Due to symmetry only one half of the circular cross section of the symmetry plane passing through point of contact is modeled. The grid consists of 5456 triangular axisymmetric elements comprising a total of 2940 nodes. As illustrated in Fig. 4(*b*), in order to assure the accurate modeling of the collision the mesh is built sufficiently fine in regions of high gradients of stresses and strains. A rezoning mechanism is applied to the Lagrange to eliminate the elements that become severely distorted or tangled. Using time step of 5×10^{-11} s, the unsteady calculations were performed until the velocity of the spherical glass ball on the left reached to that of the right ball before the collision. Figures 4(*c*)–4(*f*) represent a gradual change in the velocities of balls, from initial to terminal velocity states during the collision process. The numerical model was further verified by comparing its output with the analytical results of the Hertz theory. The verification included the maximum contact pressure, and the collision time. The variations of the contact pressure experienced by a node located at the point of initial contact as a function of time for several impact velocities are depicted in Fig. 5. As illustrated in Fig. 6(*a*), the difference between the numerical and analytical results for the maximum contact pressure was always less than 3.5%, which increases with the impact velocity. Since the impact does not produce a permanent deformation, the law of conservation of mechanical energy requires that the ratio of final to initial relative velocity components of colliding balls in the direction. The numerical results were recorded after every 10^{-6} s. In this light, the small difference between the numerical and analytical results for the collision time, as illustrated in Fig. 6(*b*), may be an effect of the limit of uncertainty of numerical calculations. Using Arbitrary Lagrange Euler (ALE) grids did not reduce the difference, illustrated in Fig. 6. Note that ALE method of space discretization is a hybrid of the Lagrange and Euler method that allows redefining the grid continuously in arbitrary, as well as predefined ways, as the calculation proceeds. This method effectively provides a continuous rezoning by which the elements that become highly distorted can be eliminated.

Another verification of the model was carried out by refining the mesh in the area of contact. However, by doubling the number of nodes in the vicinity of the area of contact, the largest difference in the contact pressure was found to be less than 2%. These

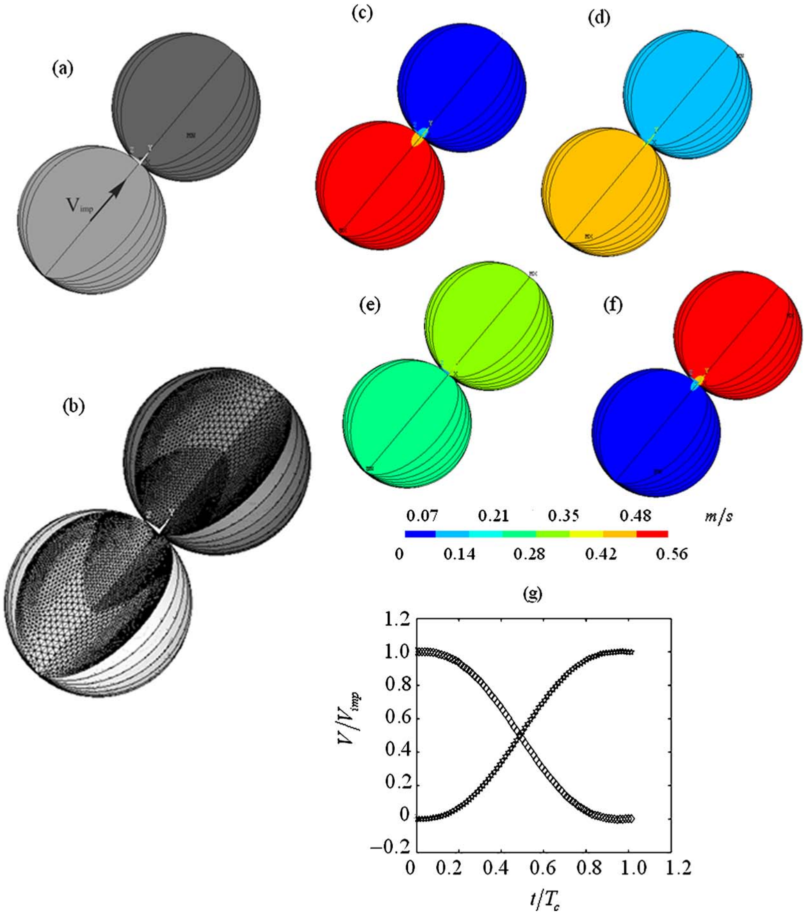


Fig. 4 Three-dimensional view of a collinear collision of two identical spheres. (*a*) The left sphere with velocity of 0.56 m/s approaches the right sphere, which is initially stationary. To obtain better visualizations one-fourth of the particles are removed. (*b*) Grid used in the numerical treatments. Note that finer meshes are used in the vicinity of contact area where the gradient of stresses and strains are high. (*c*)–(*f*) Time evolutions of the local velocities of the balls during the collision process. (*g*) Variations of velocities of center of mass of the particles with time. Here velocity and time are normalized with impact velocity and collision time, respectively. Squares and diamonds represent the right side and the left side particles, respectively.

verifications establish the validity of the numerical model with the mesh, depicted in Fig. 6(*b*), to study the collision processes in colliding spherical particles.

During impact, due to elastic deformation, the centers of the spheres approach each other resulting in a compressive stress, whose distribution in the vicinity of initial point of contact is demonstrated in Fig. 7(*a*). After the instant of maximum compression whose stress distribution is illustrated in Fig. 7(*b*), the spheres expand and a restitution period begins during which the compressive stress in the contact area decreases as shown in Fig. 7(*c*).

The restitution period is lasting to the instant of separation at which stresses are practically removed as depicted in Fig. 7(*d*).

Since the particles are perfectly elastic and frictionless, and energy absorbed in wave motion is neglected, the deformation is perfectly reversible. From Fig. 8 it can be seen that the stresses decrease rapidly with increased distance from the area of contact. In addition, from Fig. 7 the contact area and its radius may be roughly estimated [19].

In order to solve the elastic-plastic contact of spherical balls, a yielding criterion should be adopted, such as the von Mises criterion [20]. It is possible to present the yield condition as a function of the stress deviator invariants. Hence, it is useful to demonstrate

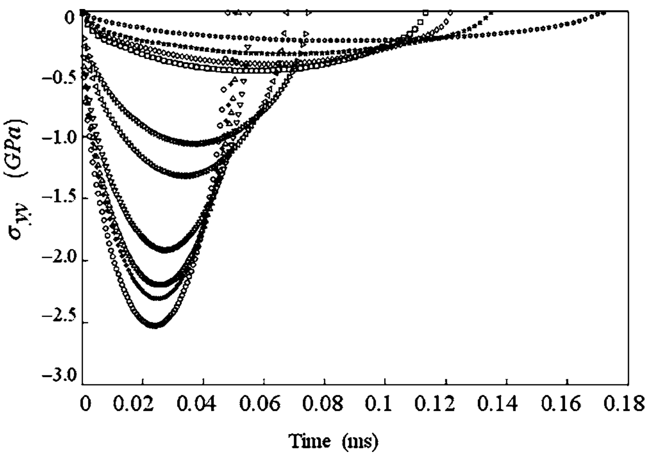


Fig. 5 Variations of the component of normal stress σ_{yy} (contact pressure) of the node located at the point of initial contact as a function of time. Six point star, five point star, diamond, box, left triangle, right triangle, inverted triangle, triangle, asterisk, and circle represent collisions of impact velocities of 0.01, 0.03, 0.05, 0.07, 0.5544, 0.9538, 2.5, 3.5, 4.0, and 5.0 m/s, respectively.

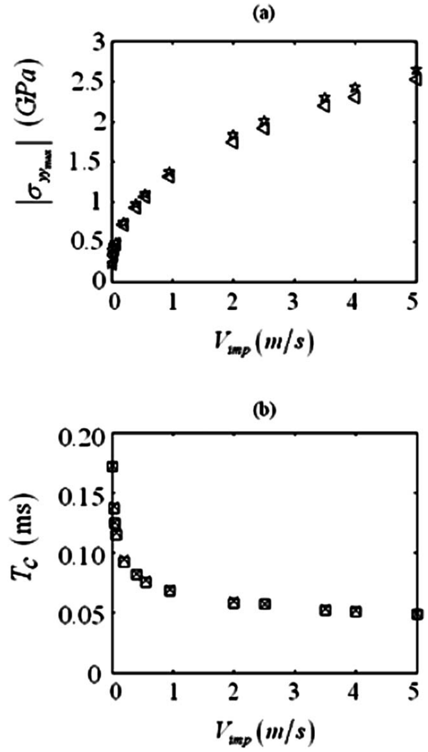


Fig. 6 (***a***) **Maximum of contact pressure $|\sigma_{yy}|$ at the initial point of contact as a function of impact velocity. Stars and triangles are the analytical and the numerical results, respectively. (*b*) Variations of collision time with impact velocity. Crosses and squares represent the analytical and the numerical results.**

the distribution of the effective stress, defined as $\sigma_{\text{eff}}=\sqrt{3\Pi_\Sigma}$, in the area of contact. Here, Π_Σ represents the second invariant of the deviatoric stress tensor. Figure 7(*b*) presents the contour plot of σ_{eff} at the end of approach period when the interference is the largest. From Fig. 8(*b*) it can be seen if a transition from the elastic to the elastic-plastic regime occurs at moderately high velocity impacts such as discussed above, it is likely that an elastic core remains locked between the plastic regions where the highest effective stress, σ_{eff}, is located. The present model may offer an accurate numerical solution for the inelastic impacts, which will be discussed in the following.

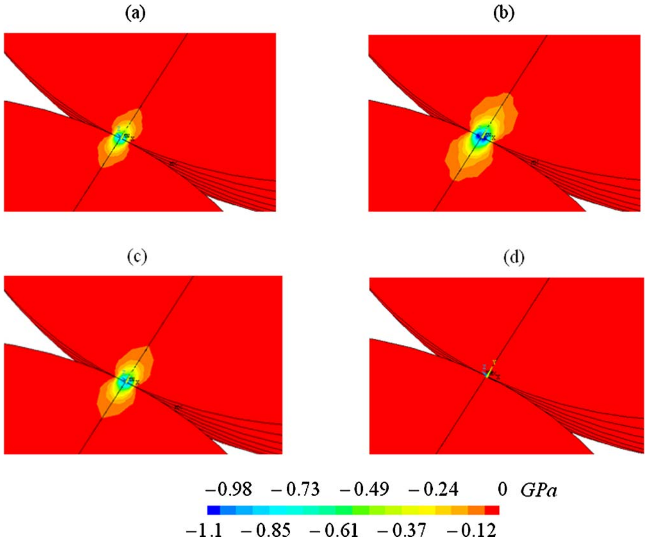


Fig. 7 Contour plots of σ_{yy} for two spheres during a collinear impact. (*a*) At the middle of compression period. (*b*) At the maximum compression. (*c*) At the middle of restitution period. (*d*) At separation. The impact velocity is 0.55 m/s.

Inelastic Collisions. Here, a three-parameter model, namely two springs with one dashpot, was used in order to generalize the numerical approach described in the preceding section. The schematic of the three-parameter model is shown in Fig. 2.

One of the central problems in devising descriptions of viscoelastic materials is the question of how to describe the manifestation of both elastic and viscous effects. Mase [20] suggested that in developing the three-dimensional theory for viscoelasticity, distortional and volumetric effects must be treated independently. To this end, the stress tensor may be resolved into deviatoric and spherical parts, given as

$$\sigma_{ij} = S_{ij} + \tfrac{1}{3}\delta_{ij}\sigma_{kk} \tag{13}$$

where $S_{ij}=\int_0^t \phi_s(t-t')de_{ij}/dt'dt'$, and the relaxation function used in the model is given by $\phi_s(t)=G_\infty+(G_0-G_\infty)e^{-t/\tau}$. Note that in the present study the so-called summation convention is used. Moreover, it is assumed that the stress and strain vanish for times $-\infty < t' < 0$.

The volumetric part of stress tensor would have a similar form, but with different relaxation functions. Therefore, the governing filed equations for an isotropic, viscoelastic continuum body such as the glass balls used in Newton's cradle experiments of the preceding section take the form of

$$\rho\ddot{u}_i = (\sigma_{ij})_{,j}$$

$$\varepsilon_{ij} = (u_{i,j} + u_{j,i})/2 \tag{14}$$

$$\dot{S}_{ij} + S_{ij}/\tau = (G_0 + G_\infty)\dot{e}_{ij} + G_\infty e_{ij}/\tau$$

The first equation in Eq. (14) is the equation of motion, the second equation is strain-displacement expressions, and the third equation is the stress-strain relation.

To demonstrate the use of above-mentioned viscoelastic model, a numerical simulation of an impact problem of two spherical glass balls, whose physical properties are listed in Table 1, is conducted using the Galerkin weighted, residual method described in the preceding section. Consider the collision of two spherical balls, as depicted in Fig. 9, of diameter $\sigma_p=2.2$ mm at a point on the line connecting their centers of mass. In this case, a translating smooth sphere on the left with velocity 0.954 m/s collides with the stationary smooth sphere on the right.

The value of G_0 is found from that of Young's modulus presented in Table 1 using an expression given for shear modulus, namely $G_0=E/2(1+\nu)=2.53\times 10^{10}$ kg/ms. Suggestion on selecting the long-time shear modulus, G_∞, was made by the representative of Sigmund Linder GmbH, Germany. Apparently, one can use $G_0/G_\infty \approx 4$. Hence, the value of G_∞ is set to 0.61×10^{10} kg/ms. Moreover, an initial guess was made for the relaxation time, which is an unknown in the present model. Gogun (2000) [11] reported that the collision time for a case in which about 40% of the kinetic energy is lost on collision was slightly longer than that predicted by Hertz theory for an elastic collision with the same impact velocity. Using the measured value for the coefficient of restitution listed in Table 2, namely $e=0.816$, the trial value for the relaxation time may be found to be $\tau=1.676\times 10^{-5}$ s.

Figure 9(*b*) illustrates the variations of kinetic energies of the spherical balls, shown in Fig. 9(*a*), as a function of time for $\tau=1.676\times 10^{-5}$ s. In this case, the value of $e=0.887$ is found at the impact velocity of $V_{\text{imp}}=0.954$ m/s, which is larger than that reported in Table 2. Note that the viscoelastic model used in this study captures the essential idea that the response to slow disturbances is that characteristic of a viscous fluid. In light of this, the viscous type resistance increases by selecting a shorter relaxation

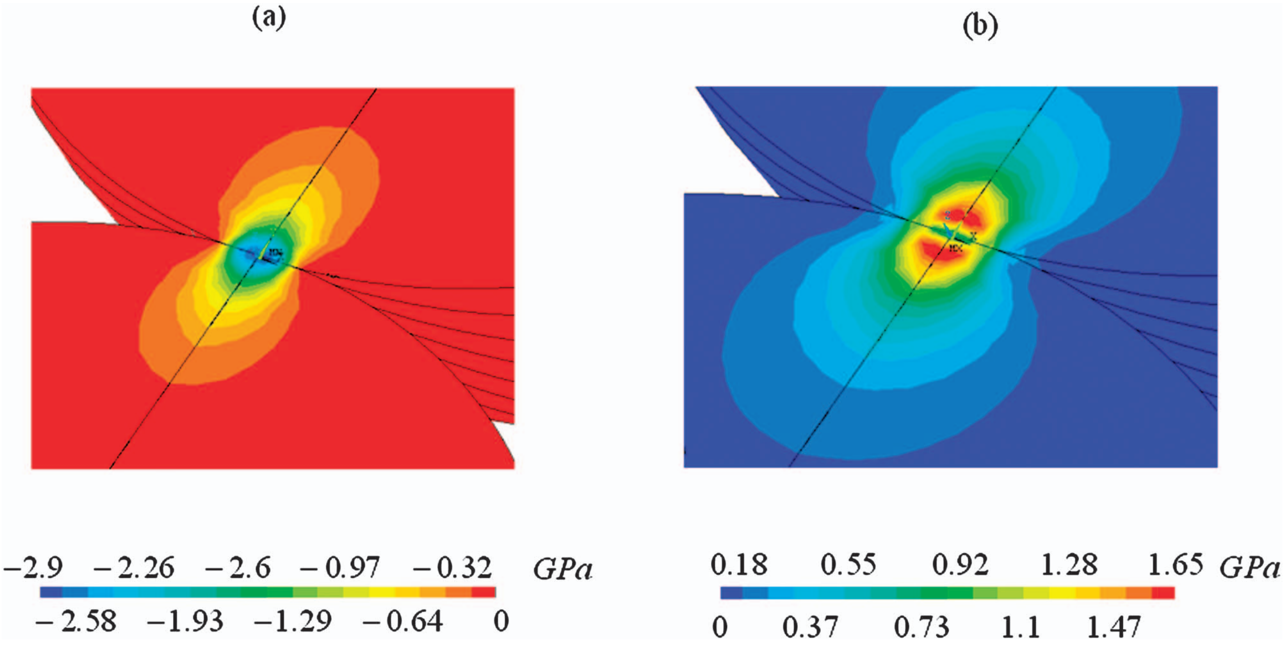


Fig. 8 Contour plots of (*a*) σ_{yy} and (*b*) σ_{eff} (Von Mises stress) at the maximum compression of a collision with an impact velocity of 5 m/s

time compared with the initial guess of 1.676×10^{-5} s. The correct value for the relaxation time can be determined using a trial and error approach. From Fig. 9(*c*) it can be seen that by using a value of 9.87×10^{-6} s for the relaxation time, the calculated value of the coefficient of restitution approaches that measured by using the Newton's cradle device at an impact velocity of 0.954 m/s. In this case, the balls lost about 20% of their kinetic energy at the impact velocity of 0.954 m/s, as illustrated in Fig. 9(*c*). A detailed comparison of the profiles of kinetic energy of particles in a partially elastic and an elastic collision is illustrated in Fig. 9(*c*). Note that for the above-mentioned collision, T_c is one order of magnitude larger than the relaxation time, namely $T_c/\tau\approx7$.

Figure 9(*d*) presents the time evolution of σ_{yy} for a node located at the initial point of contact of the viscoelastic balls, as well as that for an elastic collision with the same impact velocity for comparison. From Fig. 9(*d*) it can be seen that a considerable part of the initial kinetic energy is dissipated for the case of inelastic collision in the restitution period. In this case, the contact pressure vanishes earlier than that of the corresponding elastic collision. Moreover, the numerical results presented in Fig. 9(*d*) imply that the maximum value of contact pressure at the point of initial contact decreases as the dissipation increases. From results presented in Fig. 9(*d*) it can also be seen that the dissipation could occur during the processes of conversion of energies in both approach and restitution periods. However, it appears that the dissipation is more enhanced in the latter period.

The deformation history as shown in Fig. 9(*e*) appears to be very similar to that presented by Goldsmith [21], as an assumed scenario for deformation history in Stereomechanical impacts. The only difference is that the calculated time for the approach period is shorter than that proposed by Goldsmith, as depicted in Fig. 9(*f*). The results presented in Fig. 9(*e*) imply that the energy dissipation can also occur during approach period, which suggests that the elastic potential energy function used by Hertz might be valid only for the short period of time in the beginning of approach period. As illustrated in Fig. 9(*e*), the approach period followed by a somewhat longer restitution resulting the collision time, which is surprisingly close to that of elastic collision, where the difference may be at the limit of numerical errors.

Note that it appears that the results presented in Fig. 9(*e*) predict a shorter restitution period for partially elastic impact compared to that of elastic impact at the same impact velocity. The reason is that at the end of restitution period for a partially elastic impact is the instant of separation, which is not exactly the same as the instant at which the contact pressure at the point of impact diminishes. Note that for a perfectly plastic collision no separation may be observed even though the contact pressure at the point of impact diminishes at the end of the approach period.

As evidenced by Fig. 9(*e*), a small residual deformation appears to exist at the instant in which the contact pressure at the point of impact diminishes for a partially elastic collision.

Using the numerical model described in the above general behavior of the coefficient of restitution, e, as a function of impact velocity V_{imp} is predicted for the range of the impact velocity from 5 mm/s to 3 m/s. The results can be qualitatively divided into two categories. At relatively large impact velocity, the coefficient of restitution changes little over a wide range of impact velocity. Below $V_{\text{imp}}\approx10$ cm/s, the coefficient of restitution increases rapidly with decreasing the impact velocity.

The results presented in Fig. 10(*a*) are consistent with the classical theory of nearly elastic deformation discussed in Landau and Lifshitz [22] at the low velocity limit. The present results suggest that at a finite and sufficiently small velocity the coefficient of restitution approaches unity. Note that the calculated value for the coefficient of restitution at impact velocity of 5 mm/s is found to be 0.997. As illustrated in Fig. 10(*b*), the numerical results presented in Fig. 10(*a*) can be fitted with a power law

$$e_n = 1 - 0.74904V^{*0.1}_{\text{imp}} + 4.4873V^{*0.2}_{\text{imp}} - 9.7068V^{*0.3}_{\text{imp}} + 9.11395V^{*0.4}_{\text{imp}} - 3.3367V^{*0.5}_{\text{imp}} \tag{15}$$

where V^{*}_{imp} is dimensionless impact velocity.

The present numerical model provides a route from the details of physical properties of the system to macroscopic properties of experimental interest, such as coefficient of restitution. Note that as well as being of academic interest, the results provided in the section are technologically useful. It may be difficult to carry out experiments at very low or very high impact velocities, while a numerical simulation of an impact at any impact velocity would

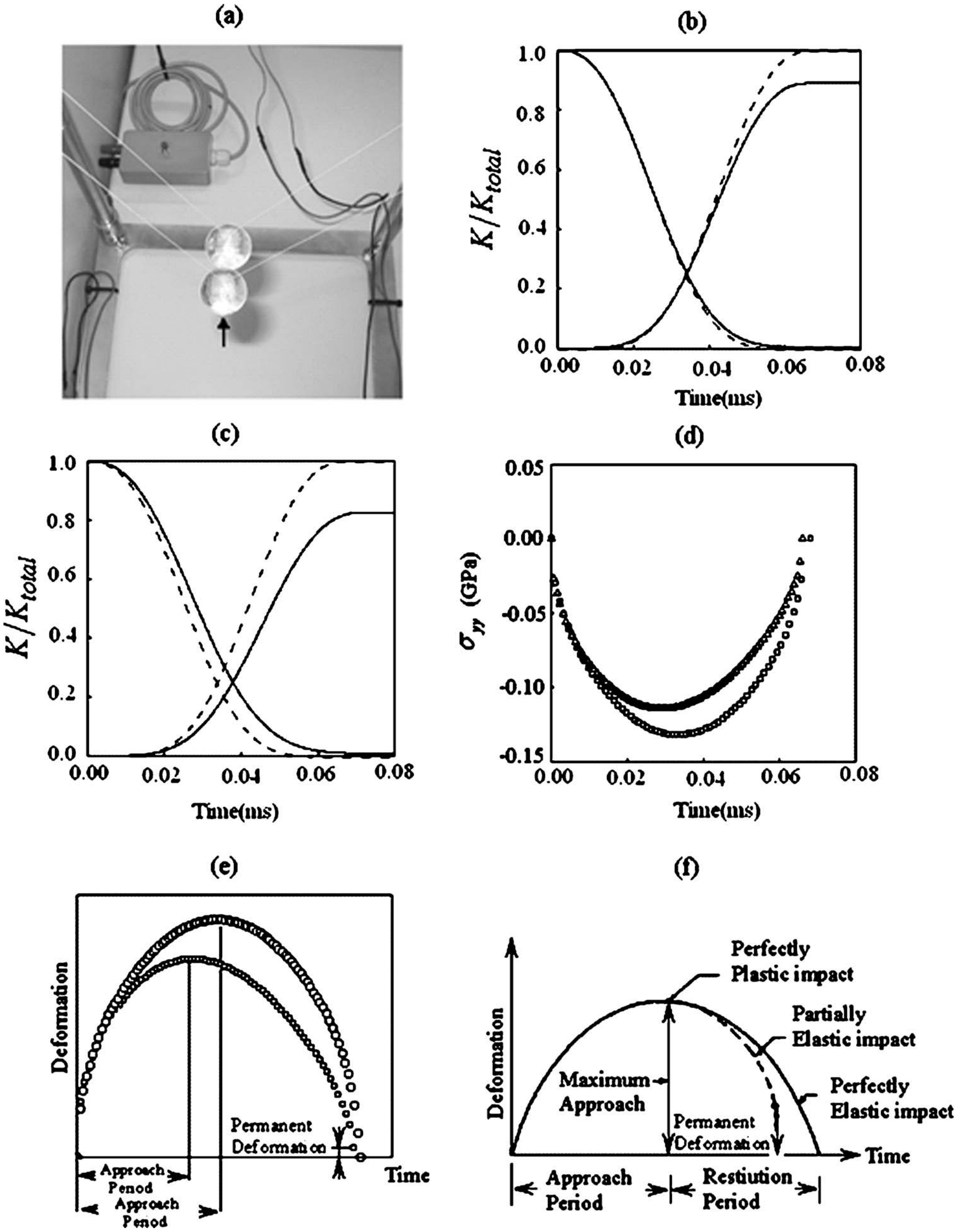


Fig. 9 (*a*) Three-dimensional view of a collinear collision of two identical glass balls used in experiments. (*b*) Normalized kinetic energy for the balls versus time. The relaxation time is set to $\tau=1.676\times10^{-5}$ s. The solid and dashed lines represent the viscoelastic and elastic collision, respectively. K_{total} is the total kinetic energy of the system. (*c*) The same as (*b*) but using a value of 9.87×10^{-6} s for the relaxation time. (*d*) Time evolution of contact pressure experienced by the node located at the initial point of contact. Box and triangle represent elastic and viscoelastic collision, respectively. (*e*) Numerical results for deformation history of a node located at the initial point of contact in a collision between two glass balls. Spheres and squares represent elastic and viscoelastic collision, respectively. (*f*) Typical deformation history proposed by Goldsmith [21].

be perfectly feasible. A wide range of physical phenomena, from elastic contact between rough surfaces to elastic-plastic impact, may be studied reliably using the generalized form of the proposed numerical model in this section.

One limitation of the scheme presented above is the relatively small number of particles that can be studied in a system. In this light, a more simplified scheme would be required to treat the interactions between a large numbers of glass balls in a system, such as a spinning bucket of sand. To this end, using the detailed results obtained in this section a simplified model for collisions between particles will be presented in the following section to be of use in molecular dynamic type simulations.

5 Simplified Model for Collision of Viscoelastic Spheres

Following the Hertz theory [19], the binary collision between two monosized, spherical particles in Fig. 9(*a*), involves repulsive

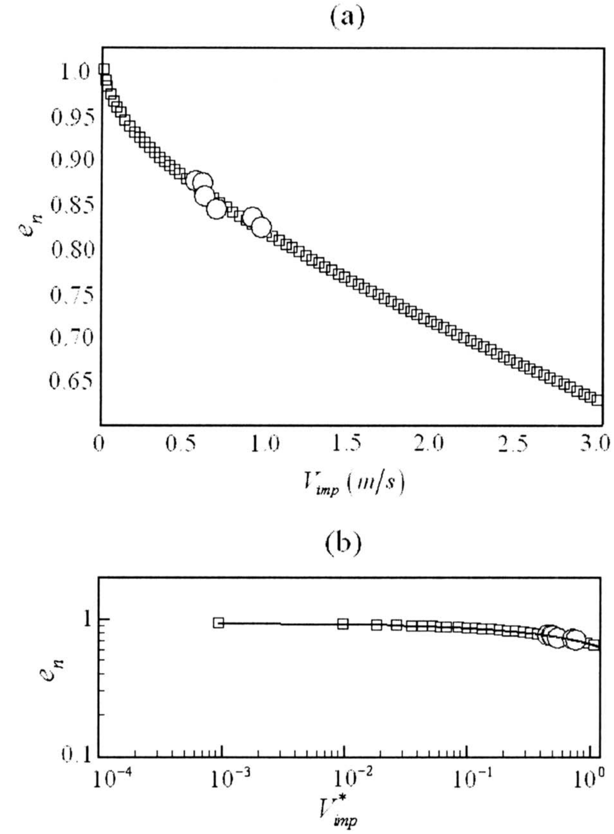


Fig. 10 (*a*) The coefficient of restitution versus impact velocity for viscoelastic glass beads with the properties given in Table 1. Squares indicate the numerical results obtained using finite element approach and circles represent experimental data obtained using Newton's cradle device depicted in Fig. 9(*a*). (*b*) The results in (*a*) are presented on a log-log scale. The solid line indicates the power law (15). Distribution of e_n is seen to follow the power law distribution.

rigid elastic interactions as well as dissipative plastic contacts. As stated earlier, a contact force model such as $F=2E/3(1-\nu^2)\sigma_p^{1/2}\delta^{3/2}+c\delta^\eta d\delta/dt$ (in which the first term represents the elastic interactions and the second term models plastic contact) is often adopted to capture the key features of inelastic interactions [7,8].

Employing dimensional analysis, physical quantities can be determined that significantly influence the viscous behavior of glass balls, shown in Fig. 9(*a*) characterized by the coefficient c in the expression for contact force model. Thus the coefficient c can be expressed as

$$c = K\tau G_0^2/(G_0 - G_\infty)\sigma_p^{1-\eta} \tag{16}$$

Combining Eq. (16) with the above-mentioned expression for the contact force, an expression for δ may be found by solving the following ordinary differential equation

$$\frac{d^2\delta}{dt^2} = -(2E/3m(1-\nu^2))\sigma_p\delta^{3/2} - K\tau(G_0^2/m(G_0-G_\infty))\sigma_p^{1-\eta}\delta^\eta d\delta/dt \tag{17}$$

Applying the physical properties listed in Table 1, Eq. (17) was integrated using fourth- and fifth-order embedded formulas from Dormand and Prince [23] with $\Delta t=3\times10^{-9}$ s. The initial conditions for Eq. (17) are $\delta(0)=0$ and $\dot{\delta}(0)=2\sqrt{gl}\sin\alpha/2$. Given that the coefficient of restitution is defined as $e_n=-\dot{\delta}(t_c)/\dot{\delta}(0)$, the value of the model parameter K in Eq. (17) can be estimated by fitting the results of the simplified model for a binary collision with those presented in Fig. 10.

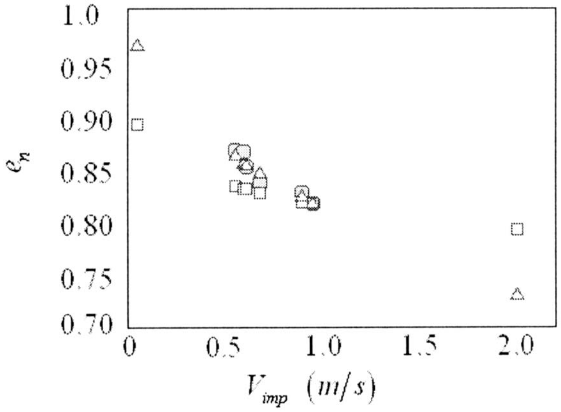


Fig. 11 The coefficient of restitution versus impact velocity for a binary collision of viscoelastic glass beads with the properties as given in Table 1. Deltas and squares indicate the numerical results obtained using simplified model (17) for $\eta=1$, and $\eta=1/2$, respectively. Gray circles represent experimental data obtained using Newton's cradle device. The values obtained for the parameter *K* for $\eta=1$ and $\eta=1/2$ are 143 and 1.67, respectively. The Kelvin-type model for $\eta=1/2$ does not seem to predict the behavior of glass balls during a collision at the low velocity limit accurately.

Figure 11(*a*) illustrates two sets of numerical results of the coefficient of restitution in a binary collision of glass balls. Those obtained using different values for the exponent η where deltas and squares represent the coefficient of restitution as a function of impact velocity for $\eta=1$ and $\eta=1/2$, respectively. The circles are the experimental data as presented in Table 2. The results presented in Fig. 11(*a*) show that δ dependence of contact force on viscous dissipation appears to provide a better representation of the measurements. The model can also predict the behavior of glass balls during a collision at the low velocity limit accurately. However, the Kelvin-type model (for which $\delta^{1/2}$ dependence of contact force on viscous dissipation was obtained [8]) can predict the coefficient of restitution neither at the low nor at the high impact velocity accurately. This observation implies that a Kelvin-type model may not be quite appropriate to predict the glassy-type viscoelastic behavior of glass balls during a collision.

The simplified model (17) will be tested further in the subsequent paper to predict the observed behavior of sand particles in a rotating bucket. There is a need for further experimental investigation of collisions to refine the simplified model (17) and to obtain a more generalized form useful for describing a wide range of materials.

6 Conclusions

The collisions of viscoelastic, nearly identical sized glass balls have been studied experimentally using a modified Newton's cradle device. A numerical model based on finite element methods was developed to analyze the measurement data. By employing the comprehensive model, a better understanding of the processes during a collision was made possible. For example, a quantity such as the viscous damping coefficients, which is lacking in the literature, was determined for the glass balls. Power law dependence was found for the normal restitution coefficient on the impact, velocity was found to be given as $e_n=1-0.74904V_{\text{imp}}^{*0.1}+4.4873V_{\text{imp}}^{*0.2}-9.7068V_{\text{imp}}^{*0.3}+9.11395V_{\text{imp}}^{*0.4}-3.3367V_{\text{imp}}^{*0.5}$, where V_{imp}^* is dimensionless impact velocity.

To obtain detailed information about the interparticle interactions in dense granular flows, a simplified model based on dimensional analysis has been developed for collisions between glass particles. The model will be examined through its application to predict unexplained hydrodynamic phenomena that are observed in spinning buckets of sand.

Nomenclature

(A_x, A_y, A_z) = components of acceleration of the glass ball in its rigid body motion
a = acceleration vector
B = displacement matrix
C_D = coefficient of drag
c = coefficient characterizing the viscoelastic behavior of the grains
(D_x, D_y, D_z) = components of drag force due to air
E = Young modulus
e_n = normal coefficient of restitution
e_{ij} = deviatoric part of strain tensor $(e_{ij} = \varepsilon_{ij} - 1/3\varepsilon_{kk}\delta_{ij})$
F = contact force
(F_x, F_y, F_z) = components of force acting on the glass ball (rigid body motion)
G_0 = instantaneous (glassy) shear modulus
G_∞ = long time shear modulus
g = acceleration due to gravity
K = coefficient characterizing the inelastic behavior of the grains
l = length of pendulum
m = mass of the glass ball (see Fig. 1)
N = number of nodal points defining the element
Re_p = particle Reynolds number $(\mathrm{Re}_p = \sigma_p V/\nu_f)$
(r, θ, φ) = spherical coordinates system
S = the replacement gradient
T_p = tension of thread
T = traction loads vector
T_c = collision time
t = time
u_i^p = nodal position of the pth node in the ith direction.
V = particle velocity (rigid body motion) $V = l(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta)^{1/2}$
V_{imp} = impact velocity
V_{imp}^* = dimensionless impact velocity $(V_{\mathrm{imp}}^* = V_{\mathrm{imp}}/V_{\mathrm{ref}})$
V_{ref} = reference velocity $(V_{\mathrm{ref}} = 1\ \mathrm{m/s})$.
(x, y, z) = Cartesian coordinates system

Greek

α = inclination angle (see Fig. 1)
β = inclination angle (see Fig. 1)
δ = overlapping of a pair spherical particles with diameter σ_p at positions r_1 and r_2 $(\delta = \sigma_p - |r_1 - r_2|)$
δ_{ij} = Kronecker delta
e_{ij} = deviatoric part of strain tensor defined as $e_{ij} = \varepsilon_{ij} - 1/3\varepsilon_{kk}\delta_{ij}$
λ = Lame coefficient $(\lambda = \nu E/(1+\nu)(1-2\nu))$
σ_p = particle diameter
$\sigma_{ij}^{(\mathrm{el})}$ = stress tensor in a rectangular coordinate
σ = stress vector whose transpose is given as $\sigma^t = (\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{yz}, \sigma_{zx})$
σ_{eff} = von Mises stresses defined as $\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 - (\sigma_{xx}\sigma_{yy} + \sigma_{xx}\sigma_{zz} + \sigma_{yy}\sigma_{zz}) +, 3(\tau_{xy}^2 + \tau_{xz}^2 + \tau_{yz}^2)$
γ = exponent in the expression for contact force
η = coefficient of viscosity
ρ = density
ρ_f = density of air
ν = Poisson's ratio
ν_f = air kinematic viscosity
Π = interpolation matrix
ξ = exponent in the expression for contact force
∂b_1 = the boundary of spheres
ϕ_p = interpolation functions of parametric coordinates (ξ, ζ, η)
θ_0, φ_0 = initial displacements
$\theta_{\mathrm{eq}}, \varphi_{\mathrm{eq}}$ = equilibrium position
τ = relaxation time
(ξ, ζ, η) = parametric coordinates

Superscript

dot = the derivative with respect to time

References

[1] White, F. M., 1974, *Viscous Fluid Flow*, McGraw-Hill, New York.
[2] Schlichting, H., 1979, *Boundary-Layer Theory*, McGraw-Hill, New York.
[3] Nedderman, R. M., 1992, *Statics and Kinematics of Granular Materials*, Cambridge University Press, Cambridge, UK.
[4] Bridges, F. G., Hatzes, A., and Lin, D. N. C., 1984, "Structure, Stability, and Evolution of Saturn's Rings," Nature (London), **309**, pp. 333–335.
[5] Jaeger, H. M., Nagel, S. R., and Behringer, R. P., 1996, "Granular Solids, Liquids, and Gases," Rev. Mod. Phys., **68**, pp. 1259–1272.
[6] Kuwabara, G., and Kono, K., 1987, "Restitution Coefficient in a Collision Between Two Spheres," Jpn. J. Appl. Phys., Part 1, **26**(8), pp. 1230–1233.
[7] Morgado, W. A. M., and Oppenheim, I., 1997, "Energy Dissipation for Quasielastic Granular Particle Collisions," Phys. Rev. E, **55**, pp. 1940–1945.
[8] Brilliantov, N. V., Spahn, F., Hertzsch, J.-M., and Poschel, T., 1996, "Model for Collisions in Granular Gases," Phys. Rev. E, **53**, pp. 5382–5392.
[9] Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, Dover, New York.
[10] Falcon, E., Laroche, C., Fauve, S., and Coste, C., 1998, "Behavior of One Inelastic Ball Bouncing Repeatedly off the Ground," Eur. Phys. J. B, **3**, pp. 45–47.
[11] Gugan, D., 2000, "Inelastic Collision and the Hertz Theory of Impact," Am. J. Phys., **68**(10), pp. 920–924.
[12] Foester, S. F., Louge, M. Y., Chang, H., and Allia, K., 1994, "Measurements of the Collision Properties of Small Spheres," Phys. Fluids, **6**(3), pp. 1108–1115.
[13] Findley, W. N., Lai, J. S., and Onaran, K., 1989, *Creep and Relaxation of Nonlinear Viscoelastic Materials (With an Introduction to Linear Viscoelasticity)*, Dover, New York.
[14] Hughes, W. F., Gaylord, E. H., and Hughes, G. W., 1964, *Basic Equations of Engineering*, McGraw-Hill, New York.
[15] Haider, A., and Levenspiel, O., 1989, "Drag Coefficient and Terminal Velocity of Spherical and Nonspherical Particles," Powder Technol., **58**, pp. 63–70.
[16] Kolsky, H., 1963, *Stress Waves in Solids*, Dover, New York.
[17] Billington, E. W., and Tate, A., 1981, *The Physics of Deformation and Flow*, McGraw-Hill, New York.
[18] Hughes, T. J. R., 2000, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Dover, Mineola, NY.
[19] Johnson, K. L., 1987, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.
[20] Mase, G. E., 1970, *Continuum Mechanics*, McGraw-Hill, New York.
[21] Goldsmith, W., 2001, *Impact, the Theory and Physical Behaviour of Colliding Solids*, Dover, New York.
[22] Landau, L. D., and Lifshitz, E. M., 1970, *Theory of Elasticity*, Pergamon, New York.
[23] Dormand, J. R., and Prince, P. J., 1980, "A Family of Embedded Runge-Kutta Formulae," J. Comput. Appl. Math., **6**, pp. 19–26.

Isotropic Clamped-Free Thin Annular Circular Plate Subjected to a Concentrated Load

Ajayi O. Adewale
Department of Mechanical Engineering,
University of North Florida,
Jacksonville, FL 32224
e-mail: aadewale@unf.edu

The problem of an isotropic annular plate clamped along one edge and free at the other and subjected to a concentrated load is solved by a series approximation. The continuity conditions of deflection, slope, shear and radial moments at the radius of load application are satisfied. Variations of deflection coefficient, radial moment coefficients and shear coefficients with radius and angle are presented. [DOI: 10.1115/1.2165235]

1 Introduction

Annular plate problems occur in engineering application, for example in the design of structures where a load is supported by a circular overhang. Some of the early attempts to solve the annular plate problems include the work of Conway [1] who considered an annular plate with linearly varying thickness subjected to a uniformly distributed load and a line load uniformly distributed along the edge of the hole. The nature of the problem solved by Conway [1] ensures that variations along the circumference varnish and that simplifies the governing equations greatly. Sherborne and Murthy [2] considered the elastic bending of an anisotropic annular plate of variable thickness. But like the work of Conway [1], the solution is only valid for symmetrical loading. Minguez and Vogwell [3] had solved the problem of an isotropic clamped-free annular plate subjected to a uniform pressure. Lord and Yousef [4] had attempted a similar problem by using numerical methods. Bird and Steele [5] presented an elegant treatment of a circular plate with arbitrary number of circular holes subjected to loading along boundaries.

Recently, Sharafutdinov [6] has considered the problem of an annular plate subjected to concentrated load along its edges, using the theory of functions of a complex variable. Sharafutdinov [6] obtained stress distribution along the contour of the circular aperture. Frequently, however, the applied load in neither uniform nor symmetrical but concentrated. Furthermore, a more common mode of concentrated load application is not along the edge of an annular plate but normal to it. A common engineering design is an annular plate loaded by a load-bearing member, transmitting a concentrated load.

2 The Annular Plate Subjected to a Concentrated Load

Consider an isotropic annular circular plate clamped at the outer edge and free at the inner edge, such as shown in Fig. 1. The plate is subjected to normal concentrated load, P applied at point A at distance b from the center O of the plate. Timoshenko [7] solved a similar problem for the circular plate clamped along its edge. In this paper, an approximate solution is obtained for the annular problem. The solutions are obtained by dividing the circular plate into the inner and outer parts. The separate solutions are required to satisfy continuity relationship along the radius of the load application. The radial moment, shear and deflections variations with radius and angle are presented. This information can be used in predicting the failure in this type of structure.

3 Governing Equations

The general theory of plate deformations is well documented [7,8]. Consider an isotropic annular circular plate that is loaded at point A at a distance b from the center of the plate (Fig. 1). The differential equation describing the deformation of the plate may be written in polar coordinates as [8]

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)\left(\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}+\frac{1}{r^2}\frac{\partial^2 w}{\partial\theta^2}\right)=P/Drr \tag{1}$$

Following an approach used by Timoshenko [7], we may divide the plate into two parts by the cylindrical section of radius b as shown in the figure by the dashed line. We then apply a series solution of the form

$$w=\sum_{m=0}^{\infty}R_m\cos(m\theta) \tag{2}$$

to the homogeneous equation each of the portions of the plate. Once an equation for deflection is known, other structural quantities like shear and moment may be readily obtained for the deflection equation. This approach avoids the existence of degenerate solutions that may exist when different series representations are introduced for the different structural quantities, as may become necessary in employing a numerical method [9]. Chen, Wu, Chen, and Lee [9] showed that by using different series representations for moment and shear forces, and by examining the resulting matrices, degenerate cases may result when the boundary integral equation and boundary element methods are employed. Equation (2) is convergent provided that R_m is convergent. Substituting Eq. (2) into Eq. (1) gives

$$\begin{aligned}&m^2(m^2-2)R_m(r)+rR'_m(r)-r^2R''_m(r)\\&\quad-2m^2\{R_m(r)-rR'_m(r)+r^2R''_m(r)\}+2r^3R'''_m(r)+r^4R''''_m(r)=0\end{aligned} \tag{3}$$

Now substituting

$$R_m=r^n \tag{4}$$

into Eq. (3) gives

$$\begin{aligned}&r^n\{(-2+n)(-1+n)^2n-2m^2(-1+n)^2+(-2m^2+m^4+2n-n^2)\}\\&\quad=0\end{aligned} \tag{5}$$

Solving Eq. (5) for n gives

$$\{n\}=\{m,m+2,-m+2,-m\} \tag{6}$$

Hence, in general, the equation describing the inner plate is

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 8, 2005; final manuscript received October 7, 2005. Review conducted by S. Mukherjee. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

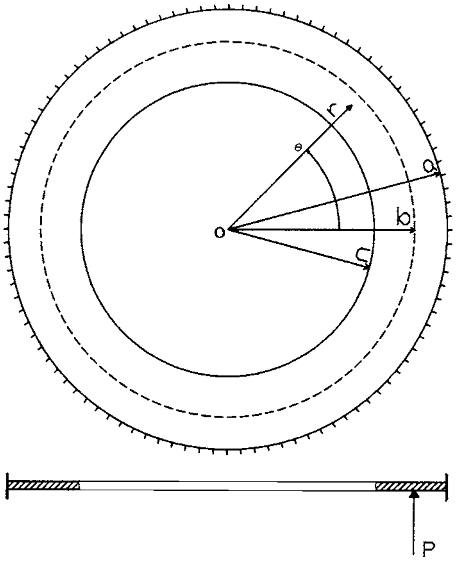


Fig. 1 Annular circular plate subjected to a concentrated load

$$R_m = A_m r^m + B_m r^{m+2} + C_m r^{-m+2} + D_m r^{-m} \tag{7}$$

Similar expressions to those given in Eq. (6) may be written for the functions R'_0, R'_1, and R'_m corresponding to the inner portion of the plate. In the same sense, we employ the symbols A'_m, B'_m and so on in the place of $A_m, B_m, \ldots$ for the constants of the outer part of the plate.

Specifically, the equation describing the outer plate becomes

$$R'_m = A'_m r^m + B'_m r^{m+2} + C'_m r^{-m+2} + D'_m r^{-m} \tag{8}$$

In this way, the basis for the inner and outer sections emerges naturally from the solution of the governing equations. This method of arriving at the basis has numerously been applied by Timoshenko [7], Gupta [10], and Carrier [11,12] in the solution of this type of problems. Recently, Chen, Wu, and Lee [13] have shown how different bases may be selected for the inner and outer sections of the plate of a circular plate. Using different bases for the inner and outer regions of a circular plate, as done for example by Chen et al. [13], averts the existence of singularity at the center of the plate, as r tends to zero. Such a situation does not arise in an annular plate.

4 Boundary Conditions

For each of the terms in Eqs. (7) and (8), we have to determine four constants for the outer portion of the plate and four for the inner portion. Hence, a total of eight unknowns are involved in the solution, requiring eight independent equations. Four of these equations are obtained from the inner and outer boundary conditions at the edges of the plate. Four additional equations are obtained from the continuity conditions along the circle of radius b. The applied concentrated load is readily expanded as Fourier series,

$$P \approx P\left[\frac{1}{2} + \sum_{m=1}^{\infty} \frac{2\sin(m\pi/2)}{m\pi}\cos(m\theta)\right], \; m = 1,3,5,\ldots, \; 0 \leqslant \theta \leqslant \pi/2. \tag{9}$$

Thus, the boundary conditions involved are obtained as follows. Since the plate is clamped at the outer radius $r=a$, deflection and slope are zero on this boundary. Hence

$$R_m(r=a) = 0 \tag{10}$$

$$\frac{\partial R_m}{\partial r}(r=a) = 0 \tag{11}$$

Since the inner boundary of the plate is free, the requirements that the shear and moment be zero on this boundary lead, respectively, to

$$\left(\frac{\partial^3 R_m}{\partial r^3} - \frac{1}{r^2}\frac{\partial R_m}{\partial r} + \frac{1}{r}\frac{\partial^2 R_m}{\partial r^2} - \frac{m^2}{r^2}\frac{\partial R_m}{\partial r}\right)\Bigg|_{r=c} = 0 \tag{12}$$

$$\left(\frac{\partial^2 R_m}{\partial r^2} + \nu\left(\frac{1}{r}\frac{\partial R_m}{\partial r} - \frac{m^2}{r^2}R_m\right)\right)\Bigg|_{r=c} = 0 \tag{13}$$

Along the circle $r=b$ where the concentrated load is placed, the following continuity equations are imposed [7],

$$R_m(r=b) = R'_m(r=b) \tag{14}$$

$$\frac{\partial R_m}{\partial r}(r=b) = \frac{\partial R'_m}{\partial r}(r=b) \tag{15}$$

$$\frac{\partial^2 R_m}{\partial r^2}(r=b) = \frac{\partial^2 R'_m}{\partial r^2}(r=b) \tag{16}$$

$$N_r(R_0)|_{r=b} - N_r(R'_0)|_{r=b} = \frac{P}{4\pi b} \tag{17}$$

$$N_r(R_m)|_{r=b} - N_r(R'_m)|_{r=b} = \frac{P\sin(m\pi/2)}{\pi^2 mb} m = 1,3,5\ldots \tag{18}$$

where

$$N_r(R) = -D\left\{\frac{\partial^3 R}{\partial r^3} + \frac{1}{r}\frac{\partial^2 R}{\partial r^2} - \frac{1}{r^2}\frac{\partial R}{\partial r} - \frac{m^2}{r^2}\frac{\partial R}{\partial r}\right) \tag{19}$$

Also, from Eq. (9),

$$R_m = 0, m = 2,4,8,\ldots \tag{20}$$

and R_m converges as the coefficients of the Fourier series (9).

5 The Solution

Using the boundary conditions in the governing equations gives the desired solutions. The solution of the eight simultaneous equations is both tedious and prone to algebraic errors. For simplicity, the external radius of the annular circular plate is taken to be unity, since this does not affect the generalization of the results. Other dimensions are normalized with respect to the outer radius. Furthermore, a symbolic tool, Mathematica [14], has been used to carry out the necessary algebraic simplifications. Depending on the value of m, different solutions are obtained. For example, for $m=0$, Eqs. (6) yield the solutions,

$$R_0 = A_0 r^2 + B_0 r^2 \ln r + C_0 + D_0 \ln r \tag{21}$$

$$R'_0 = A'_0 r^2 + B'_0 r^2 \ln r + C'_0 + D'_0 \ln r \tag{22}$$

Similarly, for $m=1$, Eq. (6) yields the solutions

$$R_1 = A_1/r + B_1 r + C_1 r \ln r + D_1 r^3 \tag{23}$$

$$R'_1 = A'_1/r + B'_1 r + C'_1 r \ln r + D'_1 r^3 \tag{24}$$

For $m>1$, the solution are written consistently as

$$R_m = A_m r^{m+2} + B_m r^{-m+2} + C_m r^{-m} + D_m r^m \tag{25}$$

$$R'_m = A'_m r^{m+2} + B'_m r^{-m+2} + C'_m r^{-m} + D'_m r^m \tag{26}$$

Using the boundary conditions expressed in Eqs. (10)–(19), the constants A_n, B_n, C_n, D_n, A'_n, B'_n, C'_n, D'_n, for $\{n=0,1,m\}$ are obtained. The expressions for the coefficients are given in the Appendix . Using Mathematica, the coefficients for $m=0$ and $m=1$ are found to reduce to the coefficients of the isotropic circular

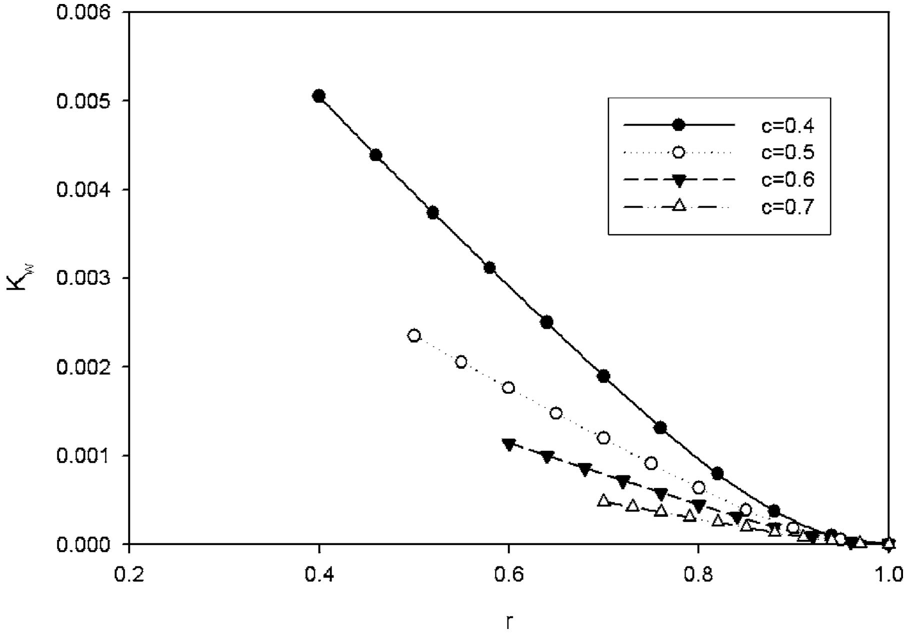


Fig. 2 Variation of deflection coefficient versus $r[\theta=0, b=(a+c)/2]$

plate in the limit as the inner radius becomes vanishingly small.

6 Results and Discussions

It is convenient to express the deflection, the radial moment, and shear, respectively, as

$$w = \frac{k_w P}{Drr} \tag{27}$$

$$M_r = \frac{k_{\mathrm{mr}} P}{Drr} \tag{28}$$

$$M_s = \frac{k_s P}{Drr} \tag{29}$$

Using a Poisson ratio of 0.3, the following results are obtained. Figure 2 shows the variation of deflection coefficient with radius for different annular radii, *c*, for a load placed at the center of the plate. The smaller the inner radius is, the less the deflection for the range of *c*, considered. Figure 3 shows the variation of radial moment with radius. Between the inner radius and the point of load application, moment is vanishingly small but increases suddenly from the point of load application to a maximum at the clamped outer radius. Figure 4 shows the variation of shear with radius. Shear is zero at the inner free boundary. From the point of load application, shear is suddenly finite and decreases in magnitude towards the outer clamped end. Figure 5 shows the variation of deflection with the circumferential angle. As expected, deflection is maximum along the meridian of load application and diminishes at the angle increases.

7 Conclusion

The problem of the point-loaded annular plate problem with one edge clamped and the other free has been solved using a series approximation. The approach had divided the plate into the inner and outer regions, based on the radius at which the concentrated load is located. The continuity of deflection, slope, shear, and radial moments at the radius of load application are satisfied. Variations of deflection coefficient, radial moment coefficients, and shear coefficients with radius and angle have been presented.

Nomenclature

a = outer radius of the annular plate
c = inner radius of the annular plate
b = distance of point load from the plate's center
r, θ = Cartesian coordinates axes
w = deflection
M_r = moment radial component
Drr = uniform flexural rigidity of plate
v = Poisson's ratio
P = concentrated load
K_w = deflection coefficient
K_{mr} = radial moment coefficient
K_s = shear coefficient
m = series index

Appendix

$$A_0 = [1 + b^2 - (1 + b^2 - 2c^2)v + 2c^2(1 + v2 \log b)]/\Delta_0$$

$$B_0 = 1/(8\pi\Delta_0)$$

$$C_0 = [v_1 + b^2 v_1 - 2c^2 v_2(1 + \log b)]/\Delta_0$$

$$D_0 = 2[b^2 v_1 - c^2 v_2 - 2c^2 v_2 \log b]/\Delta_0$$

$$A_0' = [v_1 - b^2 v_1 + 2v_1 \log b]/\Delta_0$$

$$B_0' = 0$$

$$C_0' = [(b^2 - 1)(v_1 - 2c^2 v_2) + 2(-c^2 v_2 + b^2(v_1 - c^2 v_2))\log b]/\Delta_0$$

$$D_0' = 2c^2 v_2[-1 + b^2 - 2 \log b]/\Delta_0$$

$$\begin{aligned} A_1 = [&- b^4(1 - 2v) - 3c^4(2 + v) + b^2c^2(- b^2(- 3 + v) \\ &+ c^2(9 + 5v) + v_2) + 2b^2c^2(- c^2(3 + v) + v_2)\log b - (- (- 1 \\ &+ 2b^2)c^4(3 + v) + b^4 v_1)\log c]/\Delta_1 \end{aligned}$$

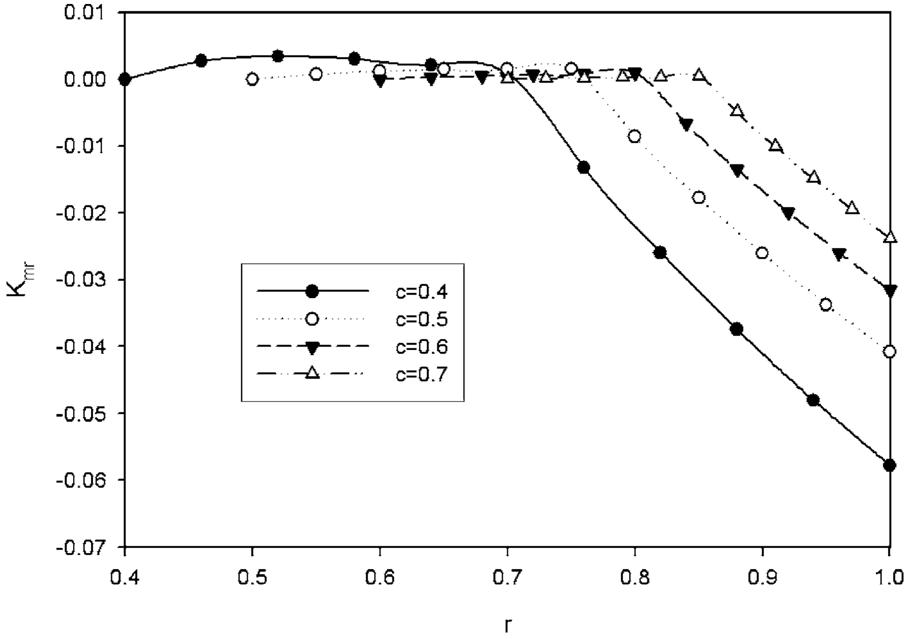


Fig. 3 Variation of radial moment coefficient versus $r[\theta=0, b=(a+c)/2]$

$$B_1=[-2b^2(-1+c^2)(c^2(3+v)-v_1)\log b$$
$$+(-1+b^2)(-9c^4-5c^4v+b^2(-1+c^2(-3+v)+3v)$$
$$+X_5-2(c^4(3+v)-b^2v_1)\log c]/\Delta_1$$

$$C_1=2[-(b^2+3c^2)(1+b^2+(-1+3b^2)c^2)$$
$$+(1+c^2)(b^4+c^2+b^2(3-5c^2))v+2b^2(-c^4(3+v)$$
$$+v_1)\log(c/b)]/\Delta_1$$

$$D_1=[b^4v+c^4(3+2v)+b^2(1-3v+c^2v_2)-X_5$$
$$+2b^2(v_1-c^2v_2)\log b+(c^4(3+v)+b^2(-2+b^2)v_1)\log(c)]/\Delta_1$$

$$A_1'=[c^2[2b^2(c^2(3+v)-v_2)\log b+(-1+b^2)\{c^2+(-3(2+v)$$
$$+b^2(3+2v))+b^2v_2+(-1+b^2)c^2(3+v)\log c\}]]/\Delta_1$$

$$B_1'=[-[2b^2\log b(1-3v+c^4(9+5v)-2a^2X_5+2(c^4(3+v)$$
$$-v_1)\log c)+(-1+b^2)(-9c^4-5c^4v+b^2(-1+c^2(-3+v)$$
$$+3v)+X_5-2(c^4(3+v)-b^2v_1)\log c)]]/\Delta_1$$

$$C_1'=2[(-1+b^2)(-c^4(3+v)+b^2(c^2(-3+v)+v_1)-X_5)$$
$$+2b^2(c^4(3+v)-v_1)\log b]/\Delta_1$$

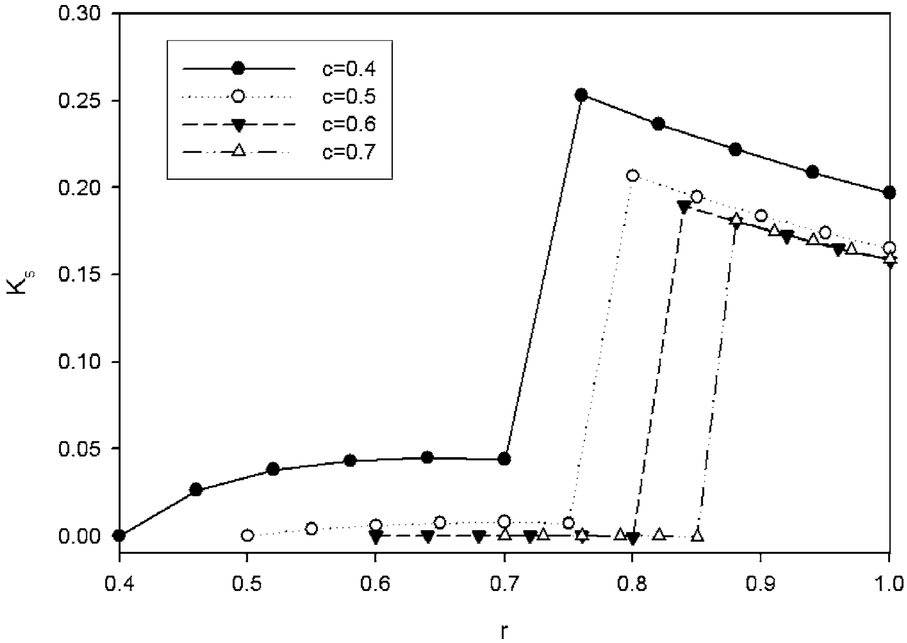


Fig. 4 Variation of shear moment coefficient versus $r[\theta=0, b=(a+c)/2]$

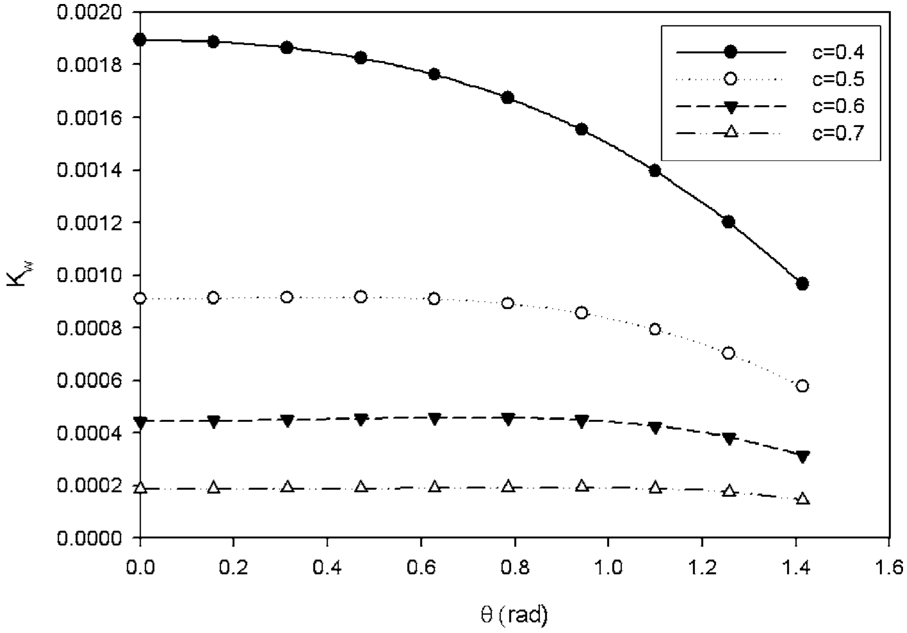


Fig. 5 Variation of deflection coefficient versus $\theta[r=a+c)/2, b=(a+c)/2]$

$$\begin{aligned} D_1' = {} & [2b^2(v_1 - c^2v_2)\log b + (-1 + b^2)(1 - 2v + b^2v + c^2(1 + v) + (\\ & -1 + b^2)v_1 \log c)]/\Delta_1 \end{aligned}$$

$$\begin{aligned} A_m = {} & [b^{-m}[c^{2+4m}(-mn_2 + (-2 + m^2)v) - b^{2m}c^2(b^2m - n_1)(-2v \\ & + mX_4) + c^{2m}(-b^2m(b^{2m} - n_1)v_1 - c^4(m^2(1 + 3v) - 4v_2) \\ & + c^2n_1(m_1X_3 + (b^{2m} - b^2m)X_4))]]/\Delta_2 \end{aligned}$$

$$\begin{aligned} B_m = {} & [b^{-m}[c^{2+4m}(b^2m - m_1)(-mn_2 + (-2 + m^2)v) - b^{2m}c^2(2v \\ & - mX_4) + c^{2m}\{b^2mv_1 + c^2m_1X_3 + b^{2m}\{-c^4(-4(1 + v) + m^2(1 \\ & + 3v)) + b^2mm_1v_1 + c^2(b^2m(2 - mn_1 + (-1 + m)mv)\}\}]]/\Delta_2 \end{aligned}$$

$$\begin{aligned} C_m = {} & [b^{-m}[c^{2+4m}m(m - b^2n_1)(-mn_2 + (-2 + m^2)v) - c^{4+2m}(-1 \\ & + b^{2m}n_1) + (-4(1 + v) + m^2(1 + 3v)) - b^{2+2m}c^{2m}m^3v_1 \\ & - b^{2+2m}c^2m(-2v + mX_4) + c^{2+2m}mn_1(-b^2X_4 + b^{2m}(b^2(-2 \\ & + m(n_1 + v - mv)) + mX_4))]]/\Delta_3 \end{aligned}$$

$$\begin{aligned} D_m = {} & [b^{-m}[-b^2c^{2+4m}m + (-mn_2 + (-2 + m^2)v) - b^{2m}c^2m(m \\ & - b^2m_1)(-2v + mX_4) + c^{2m}\{-b^{2+2m}c^2m(2 + m^2v - m(n_1 + v)) \\ & + c^4m_1(m^2 + 3m^2v - 4v_2) + b^{2m}c^4(m^2(1 + 3v) - 4v_2) \\ & - c^2m^2m_1X_3 + b^2(-m^3v_1 + c^2mm_1n_1)\}]]/\Delta_3 \end{aligned}$$

$$\begin{aligned} A_m' = {} & [b^{-m}[-b^2c^{2+4m}m + (-mn_2 + (-2 + m^2)v) - b^{2m}c^2m(m \\ & - b^2m_1)(-2v + mX_4) + c^{2m}\{-b^{2+2m}c^2m(2 + m^2v - m(n_1 + v)) \\ & + c^4m_1(m^2 + 3m^2v - 4v_2) + b^{2m}c^4(m^2(1 + 3v) - 4v_2) \\ & - c^2m^2m_1X_3 + b^2(-m^3v_1 + c^2mm_1n_1)\}]]/\Delta_3 \end{aligned}$$

$$\begin{aligned} B_m' = {} & [b^{-m}c^{2m}[-c^{2+2m}(b^{2m} - b^2m + m_1)(-mn_2 + (-2 + m^2)v) \\ & + m(b^2 + b^{2m}(-m + b^2m_1))v_1 + c^2(2 + m^2v - m(n_1 + v) \\ & + b^{2m}m_1(b^2m - n_1)X_3)]]/\Delta_2 \end{aligned}$$

$$\begin{aligned} C_m' = {} & [b^{-m}c^{2+2m}[c^{2m}m(b^{2+2m} + m - b^2n_1)(-mn_2 + (-2 + m^2)v) \\ & + c^2\{4(1 + v) - m^2(1 + 3v) + b^{2m}\{n_1(-4 + m^2 - 4v + 3m^2v) \\ & - b^2(m^3 + 3m^3v - 4mv_2)\}\} - m(b^2 + b^{2m}(-m + b^2m_1))]]/\Delta_3 \end{aligned}$$

$$\begin{aligned} D_m' = {} & [b^{-m}c^2[b^{2m}(c^{2+2m}(-4 + m^2 - 4v + 3m^2v) + m^2(2v + m(m_2 \\ & - mv))) + c^{2m}m_1(c^2(-4 + m^2 - 4v + 3m^2v) - m^2X_3) \\ & + b^{2+2m}mm_1(-2v - c^{2m}X_3 + mX_4) + b^2m(-2v + c^{2m}(-c^2(m^2 \\ & + 3m^2v - 4v_2) + m_1n_1X_3) + mX_4)]]/\Delta_3 \end{aligned}$$

$$\Delta_0 = 32\pi(a^2v_1 - c^2v_2)D_{rr}/P$$

$$\Delta_1 = 16b\pi[1 + 4c^2 - 2v + c^4(3 + 2v) - (-c^4(3 + v) + v_1)\ln c]D_{rr}/P$$

$$\begin{aligned} \Delta_2 = {} & 8m^2\pi^2m_1[c^{2+4m}(-mn_2 + (-2 + m^2)v) + c^{2m}(-4c^2(-1 + m^2) \\ & + m^2v_1 - c^4(m^2(1 + 3v) - 4v_2)) + c^2(-2v + mX_4)]D_{rr}/P \end{aligned}$$

$$\begin{aligned} \Delta_3 = {} & m\Delta_1n_1/m_1m_1 = m - 1, m_2 = m - 2, n_1 = m + 1, n_2 = m + 2, v_1 \\ & = v - 1, v_2 = v + 1, X_1 = 2m - v - 1 \end{aligned}$$

$$\begin{aligned} X_2 = {} & 2m + v + 1, X_3 = m(v - 1) - 2, X_4 = 2 + m(v - 1), X_5 = c^2(v \\ & - 3) \end{aligned}$$

References

[1] Conway, H. D., 1948, J. Appl. Mech., **15**, pp. 1–5.

[2] Sherbourne, A. N., and Murthy, D. N. S., 1970, "Elastic Bending of Anisotropic Circular Plates of Variable Thickness," Int. J. Mech. Sci., **12**, pp. 1023–1035.

[3] Minguez, J. M., and Vogwell, J., 1998, "Plates With Holes Under Lateral Load Pressure," Eng. Failure Anal., **5**(4), pp. 299–315.

[4] Lord, H. W., and Yousef, S. S., 1970, "Elastic Bending of Circular Plates of Variable Thickness: An Analytical and Experimental Study," Int. J. Mech. Sci., **12**, pp. 417–434.

[5] Bird, M. D., and Steele, C. R., 1991, "Separated Solution Procedure for Bending of Circular With Circular Holes," Appl. Mech. Rev., **44**, pp. 27–35.

[6] Sharafutdinov, G. Z., 2004, "Stress and Concentrated Forces in Thin Annular Plates," J. Appl. Math. Mech., **68**, pp. 39–51.

[7] Timoshenko, S., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, McGraw–Hill, New York, p. 206.

[8] Clebsch, A., 1862, "Theorie der Elasticitat Fester Korper."

[9] Chen, J. T., Wu, C. S., Chen, K. H., and Lee, Y. T., 2005, "Degenerate Scale for the Analysis of Circular Thin Plate Using the Boundary Integral Equation Method and Boundary Element Methods," Comput. Math. Appl. (to be published).

[10] Sen Gupta, A. M., 1952, "Bending of a Cylindrically Aeolotropic Circular Plate With Eccentric Load," J. Appl. Mech., **19**(1), pp. 9–12.

[11] Carrier, G. F., 1944, "The Bending of the Cylindrically Aeolotropic Plate," J. Appl. Mech., **11**(3), pp. A129–A133.

[12] Carrier, G. F., 1943, "Stress Distributions in Cylindrically Aeolotropic Plates," Trans. ASME, **65**, pp. A117–A122.

[13] Chen, J. T., Wu, C. S., and Lee, Y. T., 2005, "On the Equivalence of the Trefftz Method and Method of Fundamental Solutions for Laplace and Biharmonic Equations," Comput. Math. Appl. (to be published).

[14] Mathematica, 2004, "A System of Doing Mathematics by Computer," Version 5.0.

S. T. Santillan

L. N. Virgin
Mem. ASME
e-mail: l.virgin@duke.edu

Department of Mechanical Engineering and
Materials Science,
Duke University,
Durham, North Carolina 27708-0300

R. H. Plaut
Fellow ASME
Department of Civil and Environmental
Engineering,
Virginia Polytechnic Institute and State
University,
Blacksburg, Virginia 24061-0105

Post-buckling and Vibration of Heavy Beam on Horizontal or Inclined Rigid Foundation

A slender, straight beam resting on a flat, rigid foundation does not buckle when subjected to a compressive load, since the load cannot overcome the effect of the beam's weight. However, it buckles if its ends are moved toward each other. Post-buckling of such a beam is examined, both theoretically and experimentally, for horizontal and inclined foundations. The beam is modeled as an elastica, and equilibrium states with large deflections are computed, including cases in which self-contact occurs. Frequencies and mode shapes for small vibrations about equilibrium are also determined. Agreement between the theoretical and experimental results is very good. [DOI: 10.1115/1.2165237]

1 Introduction

A beam is called "heavy" if its self-weight is included in the analysis. Upheaval buckling of a heavy beam (actually a beam-column) resting on a rigid horizontal foundation has been investigated for various situations. One involves pipelines, which may be buried or rest on the seabed; then the weight of soil, water, or other material is added to the self-weight. Another deals with the handling of fabric, paper, or thin plastic or metal sheets. Other applications involve railroad tracks, concrete roads and runways, floating ice sheets, and rock strata. The compressive loads are sometimes caused by thermal stresses. Papers on these problems include [1–17] and earlier publications cited therein.

Most of these studies consider equilibrium configurations and perform a linear analysis. Often end conditions are not involved, and local buckling occurs in an internal segment of the beam. In Plaut and Mróz [5,6], large displacements were analyzed and the beams were subjected to normal pressure. If the system is perfect (horizontal foundation, straight beam, and concentric loading), buckling does not occur for a finite compressive load (i.e., there is no bifurcation from the flat configuration). Displacements may occur due to imperfections, or to displacement control in which the ends are pushed together.

An inclined foundation was considered by Bogy and Paslay [18] in relation to buckling of a drill pipe in an inclined hole. The bottom end of the heavy beam was simply supported, and the beam was subjected to a compressive load. A linear analysis was presented, and approximate equilibrium shapes were determined. There was no finite buckling load. A similar configuration was treated by Sampaio and Hundhausen [19], except that the beam was not supported by a foundation and it simply deflected downward.

Here the heavy beam (called a strip) is treated as an inextensible elastica. Its ends are fixed (no transverse displacement or rotation), and one end is pushed toward the other. A nonlinear analysis is performed to compute equilibrium shapes. Both horizontal and inclined (including vertical) foundations are considered. In addition to the static analysis, small vibrations about the buckled equilibrium states are investigated (the references cited above do not examine dynamic behavior). A shooting method is applied to solve the governing equations numerically. In addition, experiments are carried out with polycarbonate strips. The correlations between the experimental and analytical equilibrium shapes, vibration modes, and vibration frequencies are very good.

The analytical formulation is presented in Sec. 2. In Sec. 3, the experiments are described. Analytical and experimental results are presented in Sec. 4 for a horizontal foundation, and in Sec. 5 for inclined foundations. Finally, concluding remarks are given in Sec. 6.

2 Analytical Formulation

The thin, uniform strip is depicted in Fig. 1(*a*) in an inclined configuration. It has length L, constant bending stiffness EI, and constant weight W per unit length. Points on the strip have coordinates $X(S,T)$ and $Y(S,T)$, and rotation $\theta(S,T)$ with respect to the X axis, where S is the arc length and T is time. The inclination angle of the foundation with respect to the horizontal is β, and the axial end-shortening of the strip is Δ.

The internal forces in the strip are denoted $P(S,T)$ and $Q(S,T)$ parallel to the X and Y axes, respectively, and the bending moment is $M(S,T)$. The governing equations, based on geometry, moment-curvature relation, and dynamic equilibrium, are [20]

$$\partial X/\partial S = \cos\theta, \quad \partial Y/\partial S = \sin\theta,$$

$$\partial\theta/\partial S = M/EI, \quad \partial M/\partial S = Q\cos\theta - P\sin\theta,$$

$$\partial P/\partial S = -W\sin\beta - (W/g)\partial^2 X/\partial T^2,$$

$$\partial Q/\partial S = -W\cos\beta - (W/g)\partial^2 Y/\partial T^2. \tag{1}$$

Damping is neglected.

The following nondimensional quantities are defined:

$$w = WL^3/EI, \quad x = X/L, \quad y = Y/L, \quad s = S/L, \quad \delta = \Delta/L,$$

$$p = PL^2/EI, \quad q = QL^2/EI, \quad m = ML/EI,$$

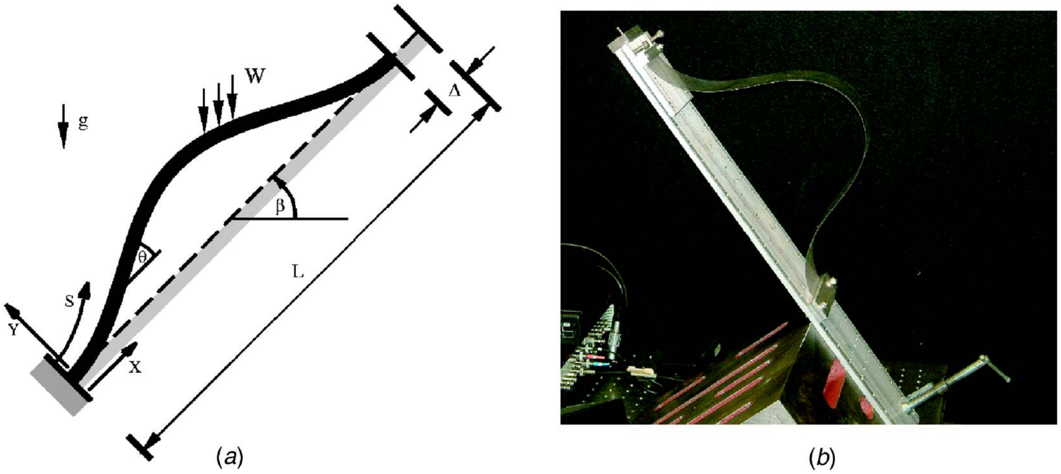
$$t = (T/L^2)\sqrt{EIg/W}, \quad \Omega = \omega L^2\sqrt{W/EIg} \tag{2}$$

where ω is a dimensional vibration frequency. In nondimensional terms, Eqs. (1) become

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 14, 2005; final manuscript received October 27, 2005. Review conducted by N. Triantafyllidis. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME



(*a*) (*b*)

Fig. 1 (*a*) Schematic of heavy inclined strip (*b*) Photo of experimental system

$$\partial x/\partial s = \cos\theta, \quad \partial y/\partial s = \sin\theta,$$

$$\partial\theta/\partial s = m, \quad \partial m/\partial s = q\cos\theta - p\sin\theta,$$

$$\partial p/\partial s = -w\sin\beta - \partial^2 x/\partial t^2, \quad \partial q/\partial s = -w\cos\beta - \partial^2 y/\partial t^2. \quad (3)$$

The variables are written in the form

$$x(s,t) = x_e(s) + x_d(s)\sin\Omega t, \quad y(s,t) = y_e(s) + y_d(s)\sin\Omega t,$$

$$\theta(s,t) = \theta_e(s) + \theta_d(s)\sin\Omega t, \quad m(s,t) = m_e(s) + m_d(s)\sin\Omega t,$$

$$p(s,t) = p_e(s) + p_d(s)\sin\Omega t, \quad q(s,t) = q_e(s) + q_d(s)\sin\Omega t \quad (4)$$

where subscripts e and d denote equilibrium and dynamic quantities, respectively. At equilibrium, the equations are given by

$$x'_e = \cos\theta_e, \quad y'_e = \sin\theta_e,$$

$$\theta'_e = m_e, \quad m'_e = q_e\cos\theta_e - p_e\sin\theta_e, \quad (5)$$

where the internal forces can be written as

$$p_e(s) = p_o - sw\sin\beta, \quad q_e(s) = q_o - sw\cos\beta \quad (6)$$

with p_0 and q_0 representing values at $s=0$.

After determining the equilibrium shape with the use of Eqs. (5), small vibrations about equilibrium are considered. Equations (4) are substituted into Eqs. (3), and Eqs. (5) are utilized. The resulting linear equations in the dynamic variables are

$$x'_d = -\theta_d\sin\theta_e, \quad y'_d = \theta_d\cos\theta_e,$$

$$\theta'_d = m_d, \quad m'_d = (q_d - p_e\theta_d)\cos\theta_e - (p_d + q_e\theta_d)\sin\theta_e,$$

$$p'_d = \Omega^2 x_d, \quad q'_d = \Omega^2 y_d. \quad (7)$$

The strip is called "short" if it does not touch the foundation between its ends. If there is a flat section resting on the foundation at both ends, the strip is called "long" [20]. For inclined foundations, the strip may have a flat section at its upper end but not at its lower end, and then it will be called "short-long." For sufficiently high values of end-shortening, self-contact may occur, in which two points on the strip contact each other and the segment between has a teardrop shape [21,22].

3 Experiments

Thin polycarbonate strips were used for the experiments. The specific weight was 11.2 kN/m^3 and Young's modulus was 2.4 GPa. The strips were 76.2 mm wide. Two thicknesses (0.508 and 1.016 mm) and two lengths (0.532 and 0.832 m) were used in combinations to yield three different values for the nondimensional weight w.

Each end of the strip was clamped under an aluminum bar, and these bars were attached to a long plate, allowing one bar to move toward the other in 6.35-mm increments to create the end-shortening (see Fig. 1(*b*)). The strip was deflected beyond the transition from long to short or short-long equilibrium before taking any measurements for both the static and dynamic experiments.

For equilibrium measurements, the vertical deflection of the strip's midpoint was recorded for varying end-shortening values. Deflection measurements were continued for a few end-shortening values larger than that for which self-contact occurred.

A point-to-point laser vibrometer (*Ometron VH300+*) was used to measure the velocity at certain points along the strip during the small-vibration experiments. For δ values ranging from 0.021 to 0.917, the first four frequencies were obtained by exciting the strip manually in varying directions and at different locations along the strip. Frequency measurements also were acquired by measuring the beam response to forced excitation. The long plate supporting the strip was attached to a shaker (*MB Dynamics PM50A Vibration Exciter*) that moved horizontally. The shaker frequency was configured for a swept sine wave with frequencies from 0.008 to 50 Hz.

For the modal analysis, the strip was excited with a modal impact hammer (*Endevco 2302-50*). This forcing and the beam response data acquired by the vibrometer were analyzed using ME'scope VES (*Vibrant Technology, Inc.*) to generate a frequency response function for each measurement point. The hammer was utilized to excite the strip at the same location and in the same direction for every measurement at a given end-shortening, and the vibrometer was adjusted to measure the response at 30 different points, in a specified direction, along the strip. Vibration modes associated with the first three frequencies were constructed.

4 Horizontal Strip

4.1 Equilibrium. For a horizontal foundation ($\beta=0$) and a long strip, extensive equilibrium results have been computed by Domokos et al. [15]. The weight of the strip was incorporated into the nondimensionalization, so that the case of negligible weight was not included. Dimensional lengths were divided by the "bending length" $b=(EI/W)^{1/3}$, so that, in the present analysis, $w=(L/b)^3$. Naturally the boundary conditions at the ends of the strip are irrelevant for long strips. At the lift-off (separation) points

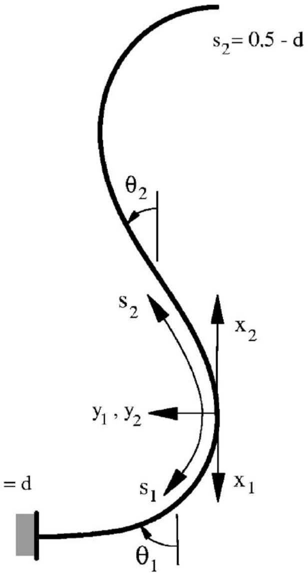


Fig. 2 Coordinate systems for symmetric equilibrium with self-contact

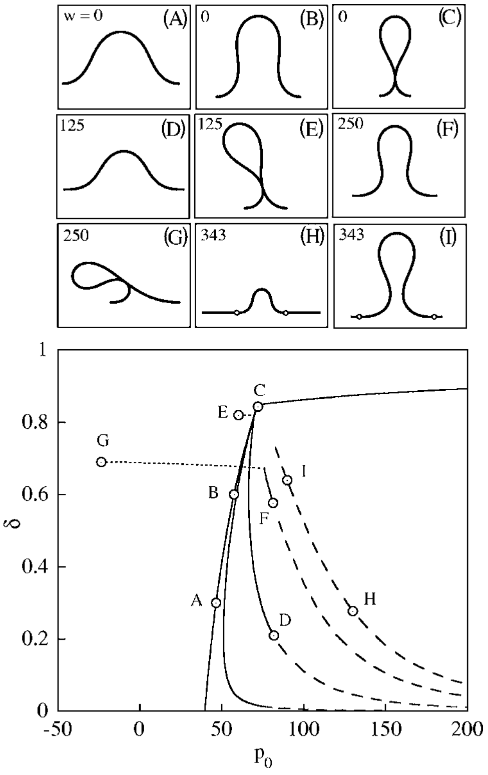


Fig. 3 Equilibrium shapes and end-shortening as a function of axial load for horizontal strip with weights (from left to right) w=0, 25, 125, 250, and 343

from the foundation, the deflection, rotation, and bending moment are zero. The authors showed that the primary stable equilibrium state is symmetric until shortly before self-contact would occur, and with further end-shortening the central part of the uplifted segment tilts to one side and then self-contact occurs later. The location of the self-contact point changes as the ends of the strip are pushed together more, and finally the central portion drops low enough to touch the foundation. In the present work, both long and short strips will be studied.

To obtain numerical results for the equilibrium shapes of strips on a horizontal foundation ($\beta=0$ in Fig. 1(a)), a shooting method is used, utilizing some of the advanced capabilities of Matlab and Mathematica. For equilibrium shapes of short strips with no self-contact, the known conditions at $s=0$ are $x_e=y_e=\theta_e=0$. The weight w and end-shortening δ are specified, the quantities $m_e(0)$, p_o, and q_o are guessed, Eqs. (5) are integrated numerically, and the three quantities are varied until the conditions $x_e=1-\delta$ and $y_e=\theta_e=0$ are satisfied at $s=1$ with sufficient accuracy. Alternatively, for symmetric shapes, one can set $q_o=w/2$ and vary $m_e(0)$ and p_o until two of the three end conditions are satisfied, or combine the three conditions into two (e.g., $x_e=1-\delta$ and $y_e^2+\theta_e^2=0$).

To compute symmetric shapes of short strips with self-contact, the origin is moved to the contact point and half the strip is considered, as shown in Fig. 2 in nondimensional terms. The arc length of the lower segment is denoted d. Equations (4) and (5) are valid in each segment for the appropriate variables. The weight w and length d are specified. The value of q_o is $(0.5-d)w$ for each segment, and the bending moments are equal at the contact point. The three variables are p_o for each segment and $m_1(0)$, where a subscript 1 or 2 indicates the segment. The variables are changed until $\theta_1(d)=\pi/2$, $\theta_2(0.5-d)=-\pi/2$, and $y_2(0.5-d)=0$. In the numerical procedure, the arc lengths are scaled, with s_1 divided by d and s_2 by $0.5-d$, so that all end conditions occur at the scaled arc length of unity. The end-shortening is given by $\delta=1-2y_1(d)$.

To determine equilibrium shapes of long strips, the origin of the coordinate system is moved to the left lift-off point (i.e., the end of the flat portion on the left side). The unknown length of the central uplifted portion is denoted c. The weight w is specified. At $s=0$, the known conditions are $x_e=y_e=\theta_e=m_e=0$, and c, p_0, and q_0 are varied until the conditions $y_e=\theta_e=m_e=0$ are satisfied at $s=c$. For symmetric equilibrium configurations, $q_o=wc/2$ and there are only two variables. The end-shortening is given by $\delta=c-x_e(c)$. (In [20], for symmetric shapes, the origin is placed on the strip at the midpoint rather than at a lift-off point.)

Equilibrium paths for the horizontal foundation ($\beta=0$) are depicted in Fig. 3 in the plane of end-shortening δ versus axial load p_o, along with some corresponding equilibrium shapes, for weights $w=0$, 25, 125, 250, and 343. (In [20], curves of $b=2y_e(0.5)$ and 2δ are plotted versus $F=p_o/4$ for $B=w/8=0$, 0.5, 1, 2, and 5.) For $w=0$ (no weight), the fixed-fixed strip buckles at $p_o=4\pi^2$, and shapes at points A, B, and C along the post-buckling path are shown. Self-contact occurs at C, when $\delta=0.849$ and $p_o=72.18$ [21], with midpoint deflection $y(0.5)=0.403$. Upon further increase of δ, the path to the right of C is followed, with large increases in p_0 associated with small increases in δ. (Analysis of a pinched loop similar to the central portion of the shape at C is presented in [22].)

For heavy strips ($w>0$), as δ is increased from zero, the strip is initially long. For weights $w=25$, 125, and 250, respectively, the strip becomes short when $\delta=0.010, 0.209$ (point D), and 0.577 (point F), with corresponding axial load values $p_o=80.60$, 81.79, and 81.36, and midpoint deflections $y_e(0.5)=0.060$, 0.255, and 0.354. For $w=25$, symmetric self-contact occurs when $\delta=0.845$, $p_o=71.65$, and $y_e(0.5)=0.401$. For $w=125$, the symmetric shape becomes unstable when $\delta=0.820$ and $p_o=69.74$, and the stable equilibrium associated with tilted (asymmetric) shapes bifurcates leftward in Fig. 3 toward point E, where self-contact occurs with $\delta=0.820$ and $p_o=60.33$. For $w=250$, the symmetric shape becomes unstable when $\delta=0.673$ and $p_o=75.46$, the path bifurcates from the solid curve to the leftward dashed curve, and self-contact occurs in the asymmetric shape at point G where $\delta=0.690$ and $p_o=-23.58$ (a tensile force at $s=0$). For $w=343$, the strip remains long until the symmetric shape becomes unstable and then the asymmetric shape exhibits self-contact when $\delta=0.73$ and p_o

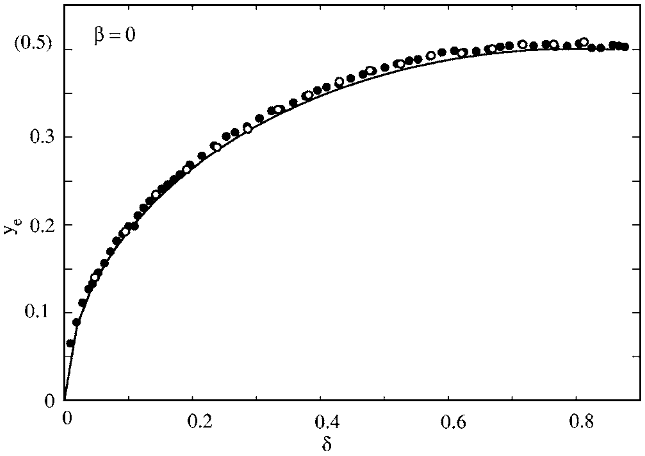


Fig. 4 Midpoint deflection as a function of end-shortening for horizontal strip. Continuous line and ○, $w=32.56$; ●, $w=9.74$.

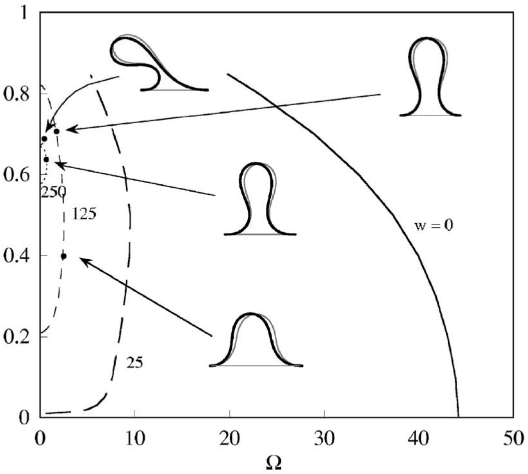


Fig. 5 End-shortening as a function of fundamental frequency for short horizontal strips with $w=0$, 25, 125, and 250, and mode shapes; equilibrium shapes are shown in gray

$=83.0$. Long equilibrium shapes are sketched for points H and I on the path, in which the open circles indicate the locations at which lift-off occurs.

The relationship of the vertical midpoint deflection $y_e(0.5)$ to the end-shortening δ is shown in Fig. 4. The curve presents numerical results for the case $w=32.56$, and the open circles represent experimental data for a strip having length 0.532 m and thickness 0.508 mm (and also $w=32.56$). There is excellent agreement between the theoretical and experimental results. The transition from a long to short strip occurs at $\delta=0.017$ according to the analysis. Self-contact occurs theoretically at $\delta=0.844$ and experimentally at $\delta=0.86$. The closed circles correspond to test results for a shorter strip, with length 0.356 m ($w=9.74$). For the two lengths considered, the nondimensional midpoint deflection $y_e(0.5)$ (i.e., ratio of dimensional vertical midpoint deflection to length of the strip) is very similar for the same nondimensional end-shortening (i.e., ratio of dimensional end-shortening to strip length).

4.2 Vibration. Small vibrations about equilibrium are considered. For long strips, the fundamental (i.e., lowest) frequency is zero, since the equilibrium shape can be moved to the side (i.e., there is an adjacent equilibrium state). Long equilibrium shapes are unstable to asymmetric perturbations. Only vibrations of short strips are examined in this study.

A similar shooting method is applied to obtain numerical solutions to Eqs. (7), making use of the equilibrium values. The initial conditions at $s=0$ are $x_d=y_d=\theta_d=0$. The quantities $m_d(0)$, $p_d(0)$, and $q_d(0)$ are varied until $x_d=y_d=\theta_d=0$ at $s=1$.

Figure 5 shows the fundamental frequency for $w=0$, 25, 125, and 250, along with the mode shapes for four cases. For symmetric equilibrium shapes, the corresponding mode is antisymmetric with just a node at the midpoint, so the central portion of the strip sways (rocks) from side to side during the motion. As the end-shortening δ is increased, the fundamental frequency is zero at the transition from long to short equilibrium shape, and then is zero again when the symmetric shape becomes unstable (as seen for weights w of 125 and 250).

For $w=0$, the frequency at $\delta=0$ is $\Omega=44.36$, corresponding to the second frequency of a fixed-fixed column subjected to an axial load $p_o=4\pi^2$. When $\Omega=19.81$ for $w=0$, and also when $\Omega=5.35$ for $w=25$, symmetric self-contact occurs, and these two curves in Fig. 5 are ended. For $w=125$, as δ is increased beyond the value 0.820 where the symmetric shape becomes unstable, the strip tilts and the frequency increases until self-contact occurs when $\Omega=0.155$ (and $\delta=0.820$ still). For $w=250$, as δ is increased beyond 0.673 and the strip tilts, the frequency increases to $\Omega=0.501$ (with $\delta=0.685$), and then decreases to $\Omega=0.343$ where asymmetric self-contact occurs (with $\delta=0.690$). An asymmetric mode along this path is depicted in Fig. 5.

In Fig. 6, similar curves are plotted to compare experimental and analytical results, with frequencies in Hz. The open circles are associated with tests on a strip of length 0.532 m and thickness 1.016 mm, giving $w=8.145$. For the data point with the largest end-shortening, the strip exhibits self-contact. The dark curve with slightly smaller frequencies is from the analysis with this weight. Since the ratio of the strip's width to its thickness is 7.5 for this case (and 15 for the other two cases in this figure), platelike effects might be significant. Therefore, the light curves are also plotted, in which the beam bending stiffness is divided by $1-\nu^2$, where Poisson's ratio, ν, is taken to be 0.37 (e.g., $w=7.030$ for this case). The experimental frequencies tend to lie between the curves for the beam and plate analyses.

The solid circles in Fig. 6 correspond to tests on a strip of length 0.532 m again, but a thickness of 0.508 mm, so that $w=32.56$ for the beam model and 28.10 for the plate model. Several data points at high values of end-shortening are associated with

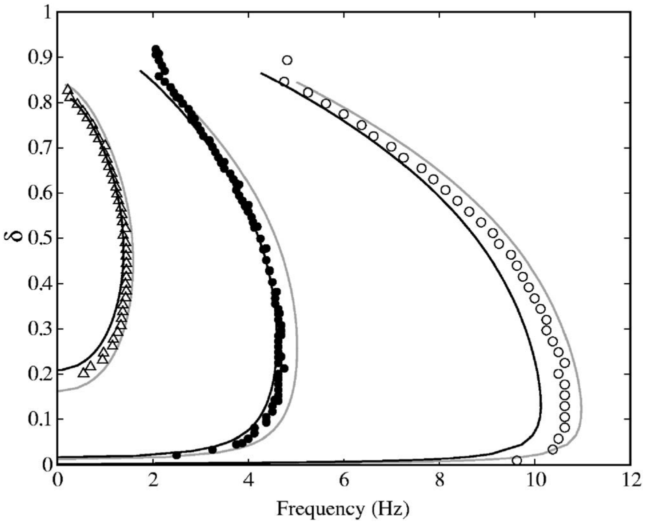


Fig. 6 End-shortening as a function of fundamental frequency for horizontal strips. Dark curves are (from right to left) for $w=8.145$, 32.56, and 124.7; light curves are for $w=7.030$, 28.12, and 107.6. ○, ●, and △, experiment.

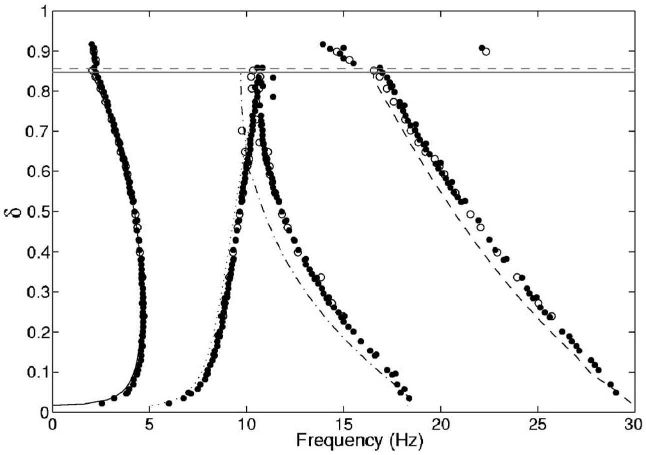


Fig. 7 End-shortening as a function of lowest four frequencies for horizontal strip with $w=32.56$. Experiment: ○ (forced) and ● (free); various lines, theory. Self-contact is indicated by the gray lines: Continuous, theory; dashed, experiment.

strips having self-contact. For the leftmost results, the open triangles were obtained experimentally for a longer strip, with length 0.832 m and thickness 0.508 mm. In this case, $w=124.7$ for the dark curve (beam) and 107.6 for the light curve (plate).

For $\beta=0$ and $w=32.56$, the first four frequencies (in Hz) are plotted in Fig. 7. (Results for the fundamental frequency are also contained in Fig. 6.) The experimental data points were obtained for a strip of length 0.532 m and thickness 0.508 mm. Open circles denote results obtained by forced vibration. The agreement is excellent for the first two frequencies, and very good for the third and fourth.

The fundamental frequency is zero when the strip becomes short at $\delta=0.017$ (with $p_o=80.86$). As the end-shortening is increased, the fundamental frequency increases and then decreases. Self-contact occurs when $\delta=0.844$ and $p_o=71.47$. As mentioned before, the first mode shape is antisymmetric with one central node. Vibration in this mode is shown at the top of Fig. 8 for $\delta=0.117$ and at the top of Fig. 9 for $\delta=0.821$, with the analytical shape on the left and the experimental shape on the right.

The second computed frequency at $\delta=0.017$ is 5.14 Hz ($\Omega=42.92$), corresponding to a symmetric mode with no nodes. This is possible even though the strip is inextensible, since the equilibrium shape has some inflection points [23]. If plotted with upward vertical displacement, the mode shape has a local minimum at the midpoint and a local maximum on each side. The central parts of Figs. 8 and 9 show vibration in this mode, with the vibrating strip above the equilibrium shape (i.e., the mode shape is added to the equilibrium configuration). The frequency for this mode increases as δ increases, and (according to the analysis) is equal to 10.70 Hz ($\Omega=89.35$) when self-contact occurs.

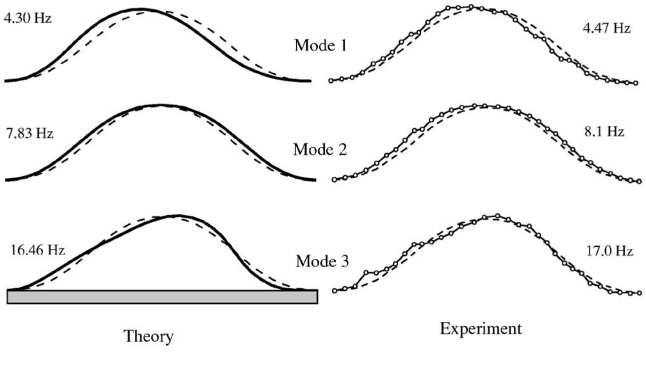


Fig. 8 First three vibration modes from analysis and experiment for horizontal strip with $w=32.56$ and $\delta=0.117$; equilibrium shape is dashed

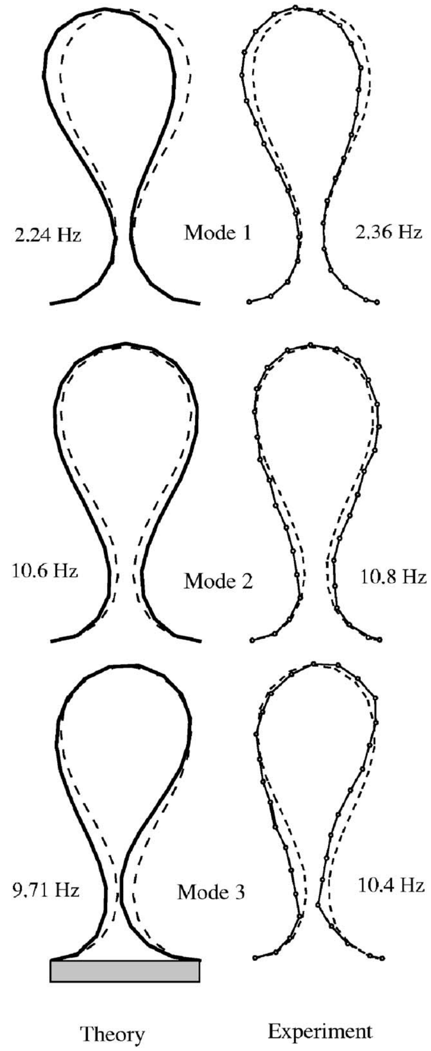


Fig. 9 First three vibration modes from analysis and experiment for horizontal strip with $w=32.56$ and $\delta=0.821$; equilibrium shape is dashed

The third computed frequency at $\delta=0.017$ is 18.27 Hz ($\Omega=152.5$), and the mode is antisymmetric with three nodes. Vibration in this mode is shown at the bottom of Figs. 8 and 9. In this case, the frequency decreases as δ increases, becomes lower than that for the no-node symmetric mode, and is equal to 9.71 Hz ($\Omega=81.09$) when self-contact occurs. (The shapes in Fig. 9 are labeled according to their mode number in Fig. 8, and not in the order of the frequencies.) The fourth frequency also decreases as δ increases, reaching 16.55 Hz ($\Omega=138.2$) at self-contact, and the mode is symmetric with four nodes.

5 Inclined Strip

Inclined strips with $0<\beta<\pi/2$ (Fig. 1) are considered first. There are no symmetric equilibrium shapes. For small end-shortening, the strip is short-long: it has a flat section resting on the foundation at the higher end, but not at the lower end. For short inclined strips, the numerical solution technique is similar to that for short horizontal strips. To compute short-long equilibrium shapes, the unknown length of the uplifted portion is again denoted c, and the origin is at the lower end of the strip. At $s=0$, the

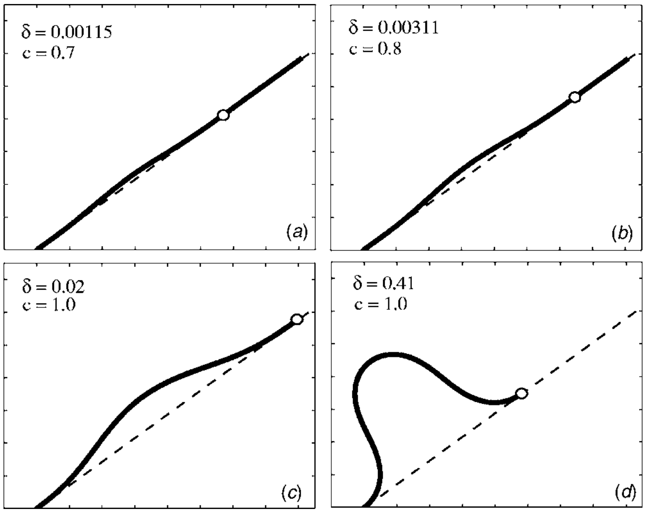


Fig. 10 Equilibrium shapes for the inclined strip with $\beta=0.627$. The strip is short-long in (a) and (b), and short in (c) and (d).

known conditions are $x_e=y_e=\theta_e=0$. For given w, β, and $c<1$, the values of $m_e(0)$, p_o, and q_o are varied until the conditions $y_e=\theta_e=m_e=0$ are satisfied at $s=c$. The remaining length, $1-c$, lies flat against the base, and the end-shortening is given by $\delta=c-x_e(c)$. For inclination angle $\beta=0.627$ (35.9°), computed equilibrium shapes for $c=0.7$ and $c=0.8$ are depicted in Fig. 10, along with two short shapes $(c=1)$. The open circles indicate the lift-off points.

Figure 11 depicts end-shortening versus fundamental frequency (in Hz) for $w=32.56$ and four values of β. The leftmost curve and solid circles are for the horizontal strip $(\beta=0)$ and are the same as the dark curve and solid circles in the middle of Fig. 6. The next curve (dot-dashed) and open circles correspond to $\beta=0.627$ (35.9°). For this inclination, the transition from short-long to short equilibrium occurs when $\delta=0.000862$, and self-contact occurs when $\delta=0.844$, at a fundamental frequency of 2.14 Hz $(\Omega=17.82)$. To the right of those results in Fig. 11, the dashed curve and open diamonds are for the case $\beta=0.944$ (54.1°). The strip becomes short when $\delta=0.00444$, and self-contact occurs again at $\delta=0.844$, with a higher fundamental frequency of 2.24 Hz $(\Omega=18.72)$.

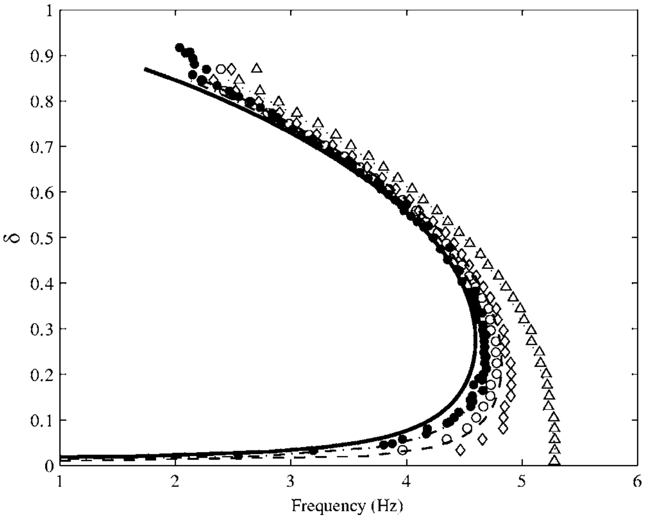


Fig. 11 End-shortening as a function of fundamental frequency for inclined strips with $w=32.56$. Solid curve and ●, $\beta=0$; dot-dashed curve and ○, $\beta=0.627$; dashed line and ◇, $\beta=0.944$; dotted line and △, $\beta=\pi/2$.

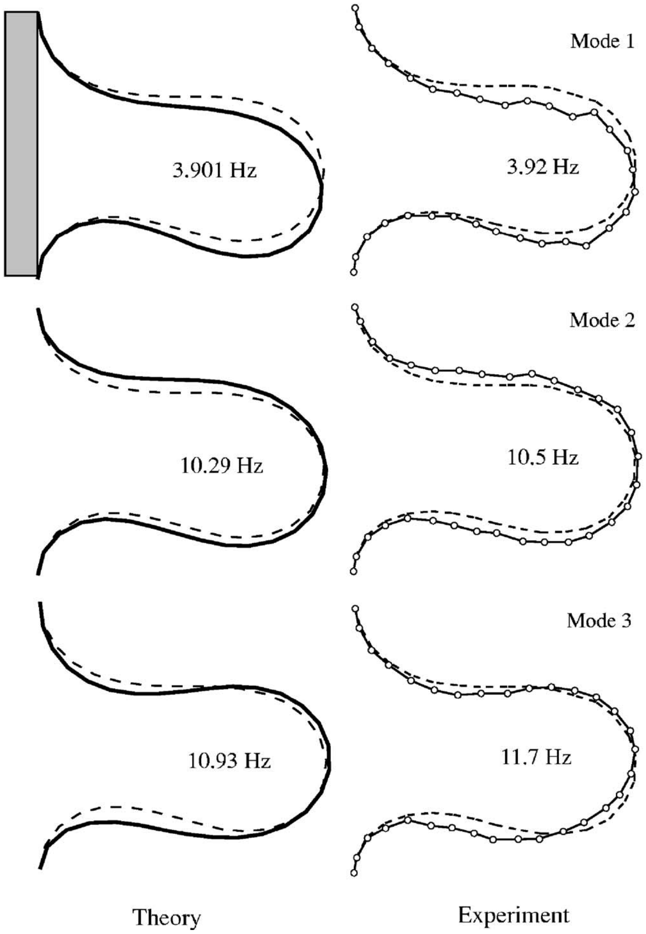


Fig. 12 First three vibration modes from analysis and experiment for vertical strip with $w=32.56$ and $\delta=0.630$, along with equilibrium shape (dashed curve)

On the right of Fig. 11 (dotted line and open triangles) are the results for $\beta=\pi/2$. This vertical case corresponds to a heavy column with fixed ends, and only has short post-buckled shapes. In fact, the foundation is irrelevant. Under the assumption of an inextensible elastica, the end-deflection δ for the vertical strip remains zero until the column buckles when the axial force p_o at the base reaches 55.4, and the initial fundamental frequency on the post-buckling path is 5.29 Hz $(\Omega=44.20)$. Self-contact occurs when $\delta=0.847$, with a fundamental frequency of 2.45 Hz $(\Omega=20.42)$. For a given end-shortening δ, as the inclination angle β increases from 0 to $\pi/2$, it is seen in Fig. 11 that the fundamental frequency increases.

Equilibrium shapes and modes are shown in Fig. 12 for the vertical strip $(\beta=\pi/2)$ with $w=32.56$ and $\delta=0.630$. The equilibrium shape is shown as a dashed curve, and the vibration shapes for the first three modes are depicted as solid curves, with analytical results on the left and experimental shapes on the right. Again the first mode has one node, the second mode has no nodes, and the third mode has three nodes.

6 Concluding Remarks

If a slender elastic beam, unstrained when straight, rests on a flat, rigid foundation that is horizontal or inclined (but not vertical), the weight of the beam will prevent it from buckling under any finite, concentric, compressive end load. However, the beam

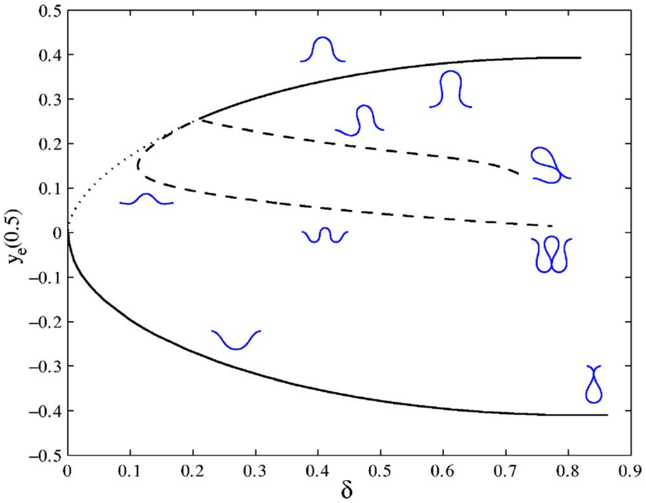


Fig. 13 Midpoint deflection as a function of end-shortening with no foundation, $\beta=0$, and $w=125$

will buckle if the ends are moved toward each other (i.e., displacement control, with applied end-shortening). Applications occur in the handling of fabrics and other materials, upheaval buckling of pipelines and roads, and drilling.

This problem was analyzed under the assumption that the strip is an inextensible elastica of finite length with fixed ends. Equilibrium shapes with large deformations were determined numerically. Small vibrations about the equilibrium shapes were investigated, and frequencies and modes were obtained. Experiments with polycarbonate strips also were carried out, and the test data correlated very well with the theoretical results. Most of the cases examined were associated with a horizontal foundation. The inclined cases involved two inclination angles as well as a vertical strip.

For the case of a horizontal foundation, as the ends of the strip are moved together, an internal segment buckles into a symmetric equilibrium shape and the strip is called ″long.″ This shape is unstable to asymmetric perturbations. If the weight of the strip is sufficiently large, the buckled segment tilts at a certain value of end-shortening, then two sides of the segment contact each other at a point (self-contact), and then the segment continues tilting until it hits the foundation [15].

If the weight is not too large, the length of the internal buckled segment increases as the end-shortening increases, until the whole strip is buckled, and the strip is then called "short." The first mode of the buckled short strip is antisymmetric with one node, the second mode is symmetric with no nodes, the third mode is antisymmetric with three nodes, and the fourth mode is symmetric with four nodes. For small weights, self-contact occurs with a symmetric shape, but for intermediate weights it occurs after the equilibrium shape of the short strip has tilted to one side.

For strips with inclination angles between 0 and $\pi/2$, as the higher end is initially moved downward toward the lower end, a flat segment exists adjacent to the higher end but not the lower end, and the strip is called "short-long." If the weight is not too large, the buckled length expands until the whole strip is raised from the foundation, i.e., the strip becomes short. After this, the fundamental frequency increases and then decreases. For sufficiently large end-shortening, self-contact occurs.

If the foundation is horizontal ($\beta=0$) and the beam is buckled upward by moving the ends together, and then the foundation is removed and the ends are pulled apart, how does the beam behave? Equilibrium states have been computed for the case $w=125$ with the use of Eqs. (5) and no restriction that $y_e(s)$ be positive or symmetric. The resulting midpoint deflection $y_e(0.5)$ is plotted as a function of the end-shortening δ in Fig. 13.

Here solid curves correspond to stable equilibrium states, and dashed curves to unstable states. As the ends of the upward-buckled beam are pulled apart, the upper solid curve is followed from right to left. At $\delta=0.209$, there is a bifurcation point with an asymmetric equilibrium path. With further pulling of the ends, the beam snaps downward and (if damping is present) settles into a downward symmetric shape on the lower solid curve, and then follows that curve to the left, approaching the origin where the beam is straight.

On the unstable branch passing downward to the left from the bifurcation point, the configurations are symmetric and the beam hangs below the x axis near its ends. The limit point occurs at $\delta=0.112$, and continuing on that branch, self-contact occurs at $\delta=0.774$. On the unstable branch of asymmetric shapes, self-contact occurs at $\delta=0.720$. The shapes depicted on that branch show the upper part of the beam bending to the left, and there are also mirror-image equilibrium configurations with the upper part bending to the right.

The lower solid curve is associated with hanging equilibrium states, as considered previously in [19] for an inclined elastica. If the ends are pushed together, self-contact occurs at $\delta=0.863$. On the upper solid curve, results for high end-shortening were discussed previously, and the tilted self-contacting shape at point E of Fig. 3 is reached when $\delta=0.820$, where the curve ends. The dotted curve from the bifurcation point to the origin in Fig. 3 corresponds to the long beam with a rigid foundation preventing it from snapping downward (as considered previously in this paper). Similar results for $w=32.56$ and 9.74 were plotted in Fig. 4.

The equilibrium equations for the horizontal beam possess many solutions that have multiple upward humps (i.e., at least two local maxima). Some of these states exist above a rigid foundation, as was demonstrated in [13] for long beams, including the case of two humps with a local minimum at the midpoint. Most are associated with high axial loads, and all the cases for short beams that were computed during the present study, with and without a foundation, are unstable and are not plotted in any of the figures.

Acknowledgment

This research was supported by the U.S. National Science Foundation under Grant No. CMS-0301084.

References

[1] Clapp, T. G., and Peng, H., 1990, "Buckling of Woven Fabrics, Part I: Effect of Fabric Weight," Text. Res. J., **60**, pp. 228–234.
[2] Clapp, T. G., and Peng, H., 1990, "Buckling of Woven Fabrics, Part II: Effect of Weight and Frictional Couple," Text. Res. J., **60**, pp. 285–292.
[3] Clapp, T. G., and Peng, H., 1990, "Buckling of Woven Fabrics, Part III: Experimental Validation of Theoretical Models," Text. Res. J., **60**, pp. 525–533.
[4] Gershon, D., and Grosberg, P., 1993, "The Buckling of Fabrics During Feeding Into Automatic Sewing Stations," J. Text. Inst., **83**, pp. 35–44.
[5] Plaut, R. H., and Mróz, Z., 1992, "Uni-Directional Buckling of a Pinned Elastica With External Pressure," Int. J. Solids Struct., **29**, pp. 1241–1253.
[6] Plaut, R. H., and Mróz, Z., 1994, "Upheaval Buckling of a Mechanism Under External Pressure and Compressive Loading," Int. J. Mech. Engrg. Educ., **22**, pp. 79–82.
[7] Maltby, T. C., and Calladine, C. R., 1995, "An Investigation Into Upheaval Buckling of Buried Pipelines. II. Theory and Analysis of Experimental Observations," Int. J. Mech. Sci., **37**, pp. 965–983.
[8] Maltby, T. C., and Calladine, C. R., 1995, "An Investigation Into Upheaval Buckling of Buried Pipelines. I. Experimental Apparatus and Some Observations," Int. J. Mech. Sci., **37**, pp. 943–963.
[9] Croll, J. G. A., 1997, "A Simplified Model of Upheaval Thermal Buckling of Subsea Pipelines," Thin-Walled Struct., **29**, pp. 59–78.
[10] Hunt, G. W., and Blackmore, A., 1997, "Homoclinic and Heteroclinic Solutions of Upheaval Buckling," Philos. Trans. R. Soc. London, Ser. A, **355**, pp. 2185–3295.
[11] Mørk, K. J., Bjørnsen, T., Venås, A., and Thorkildsen, F., 1997, "A Reliability-Based Calibration Study for Upheaval Buckling of Pipelines," ASME J. Offshore Mech. Arct. Eng., **119**, pp. 203–208.
[12] Kim, H. O., and Chang, D. C., 1999, "Method Yields Download Force to Arrest Upheaval Buckling in Offshore Lines," Oil Gas J., **97**, pp. 78–83.

[13] Kopp, C., Rahn, C. D., and Paul, F. W., 2000, "Measuring Deformations of Limp Fabrics for Material Handling," Text. Res. J., **70**, pp. 920–932.

[14] Andreuzzi, F., and Perrone, A., 2001, "Analytical Solution for Upheaval Buckling in Buried Pipeline," Comput. Methods Appl. Mech. Eng., **190**, pp. 5081–5087.

[15] Domokos, G., Fraser, W. B., and Szeberényi, I., 2003, "Symmetry-Breaking Bifurcations of the Uplifted Elastic Strip," Physica D, **185**, pp. 67–77.

[16] Palmer, A. C., and Williams, P. J., 2003, "Frost Heave and Pipeline Upheaval Buckling," Can. Geotech. J., **40**, pp. 1033–1038.

[17] Abduljauwad, S. N., Al-Ghamedy, H. N., Siddiqui, J. A., Asi, I. M., and Al-Shayea, N. A., 2004, "Stability of Vertically Bent Pipelines Buried in Sand," ASME J. Pressure Vessel Technol., **126**, pp. 382–390.

[18] Bogy, D. B., and Paslay, P. R., 1964, "Buckling of Drill Pipe in an Inclined Hole," ASME J. Eng. Ind., **86**, pp. 214–220.

[19] Sampaio, Jr., J. H. B., and Hundhausen, J. R., 1998, "A Mathematical Model and Analytical Solution for Buckling of Inclined Beam-Columns," Appl. Math. Model., **22**, pp. 405–421.

[20] Wang, C. Y., 1984, "On Symmetric Buckling of a Finite Flat-Lying Heavy Sheet," ASME J. Appl. Mech., **51**, pp. 278–282.

[21] Flaherty, J. E., and Keller, J. B., 1973, "Contact Problems Involving a Buckled Elastica," SIAM J. Appl. Math., **24**, pp. 215–225.

[22] Santillan, S., Virgin, L. N., and Plaut, R. H., 2005, "Equilibria and Vibration of a Heavy Pinched Loop," J. Sound Vib., **288**, pp. 81–90.

[23] 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.

The Arithmetic Mean Theorem of Eshelby Tensor for Exterior Points Outside the Rotational Symmetrical Inclusion

Min-Zhong Wang
e-mail: wangminz@pku.edu.cn

Bai-Xiang Xu
e-mail: xubaixiang@sohu.com

State Key Laboratory for Turbulence and Complex Systems
and Department of Mechanics and Engineering Science,
Peking University,
Beijing 100871, China

In 1957, Eshelby proved that the strain field within a homogeneous ellipsoidal inclusion embedded in an infinite isotropic media is uniform, when the eigenstrain prescribed in the inclusion is uniform. This property is usually referred to as the Eshelby property. Although the Eshelby property does not hold for the non-ellipsoidal inclusions, in recent studies we have successfully proved that the arithmetic mean of Eshelby tensors at N rotational symmetrical points inside an N-fold rotational symmetrical inclusion is constant and equals the Eshelby tensor for a circular inclusion, when $N \geqslant 3$ and $N \neq 4$. The property is named the quasi-Eshelby property or the arithmetic mean theorem of Eshelby tensors for interior points. In this paper, we investigate the elastic field outside the inclusion. By the Green formula and the knowledge of complex variable functions, we prove that the arithmetic mean of Eshelby tensors at N rotational symmetrical points outside an N-fold rotational symmetrical inclusion is equal to zero, when $N \geqslant 3$ and $N \neq 4$. The property is referred to as the arithmetic mean theorem of Eshelby tensors for exterior points. Due to the quality of the Green function for plane strain problems, the fourfold rotational symmetrical inclusions are excluded from possessing the arithmetic mean theorem. At the same time, by the method proposed in this paper, we verify the quasi-Eshelby property which has been obtained in our previous work. As corollaries, two more special properties of Eshelby tensor for N-fold rotational symmetrical inclusions are presented which may be beneficial to the evaluation of effective material properties of composites. Finally, the circular inclusion is used to test the validity of the arithmetic mean theorem for exterior points by using the known solutions.
[DOI: 10.1115/1.2165238]

1 Introduction

Eshelby [1] proved that the strain field within a homogeneous ellipsoidal inclusion embedded in an infinite isotropic media is uniform, when the eigenstrain prescribed in the inclusion is uniform. This property is usually referred to as the Eshelby property. Since then, the inclusion problem with an inclusion in an infinitely extended media has been widely investigated by many authors [1–5]. These works are fundamental in the study of material science. However, due to the absence of the Eshelby property, the study of non-ellipsoidal inclusions remains a challenging problem [6–9]. Efforts once were made to find the uniform elastic fields for other shapes of inclusions for prescribed uniform eigenstrains [10], but later [11] it was proved that the non-ellipsoidal inclusion cannot have a uniform field.

In spite of this, the non-ellipsoidal inclusion problem has become increasingly important since more and more non-ellipsoidal reinforcements are used in composite materials. Furthermore, in recent years the Eshelby problems with arbitrarily shaped inclusions have been successfully introduced to analyze the strained semiconductor quantum devices [12–16], where the strain-induced quantum wire and quantum dot growth is critical to the semiconductor nanostructure design. Among the non-ellipsoidal inclusions are the common rotational symmetrical inclusions. As we know, the cross section of most reinforcements in composites and most quantum wires in strained semiconductor devices has a boundary composed of uniform curves with specified rotational symmetry, e.g., the carbon nanotubes which display hexagonal symmetry [17], a special kind of rotational symmetry. Hence, rotational symmetrical inclusion problems are gaining interest from more and more researchers. For instance, the authors in Ref. [18] considered a two-dimensional isotropic oscillator model of a single cylindrical quantum dot with rotational symmetry.

Although the Eshelby property does not hold for non-ellipsoidal inclusions, special properties have been found for the rotational symmetrical inclusions in recent studies. In 1997, when investigating the elastic field in an arbitrary convex polygonal inclusion within an infinite matrix, Nozaki and Taya [19] obtained numerically that for any regular polygonal inclusion, except for a square, both the Eshelby tensor at the center and the average Eshelby tensor over the inclusion domain are equal to the Eshelby tensor for a circular inclusion and independent of the orientation of the inclusion. In 2001, Kawashita and Nozaki [20] verified mathematically these remarkable results. Then, in the previous works [21], we initiated the topic on the Eshelby problem with the common two-dimensional rotational symmetrical inclusion and discovered the quasi-Eshelby property, also called the arithmetic mean theorem of Eshelby tensors for interior points. Namely, for an N-fold rotational symmetrical inclusion, when $N \geqslant 3$ and $N \neq 4$, the arithmetic mean of the Eshelby tensors at N rotational symmetrical points in the inclusion is the same as the Eshelby tensor for a circular inclusion and independent of the orientation of the inclusion. Meanwhile, as a corollary of the quasi-Eshelby property, we presented that the average Eshelby tensor over the

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 21, 2005; final manuscript received November 15, 2005. Review conducted by H. Gao. Discussion the paper should be addressed to the Editor, Prof. Robert M. McMeeking, JOURNAL OF APPLIED MECHANICS, Department of Mechanical and Environmental Engineering, University of California – Santa Barbara, Santa Barbara, CA 93106–5070 and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Copyright © 2006 by ASME

N-fold rotational symmetrical inclusion domain was a constant and equal to the Eshelby tensor for a circular inclusion. These results can be used in micromechanics and composite mechanics based on the concept of equivalent eigenstrains, like the renowned Eshelby property.

In this paper, we still discuss the Eshelby problem with a two-dimensional rotational symmetrical inclusion of uniform eigenstrains. However, our purpose is to find special properties of Eshelby tensors for the points outside the rotational symmetrical inclusion, which is also an important topic in the study of rotational symmetrical inclusions. Most of the formal works [20–23] on the special properties of Eshelby tensors are focused on the points inside the inclusion. It can be expected that the Eshelby tensors at the exterior points also possess some similar special properties since the exterior field shares the same Green function with the interior field.

By making use of the Green formula and introducing the complex plane, we prove successfully that the arithmetic mean of Eshelby tensors at N rotational symmetrical points outside the N-fold rotational symmetrical inclusion is equal to zero, when $N \geqslant 3$ and $N \neq 4$. This property is referred to as the arithmetic mean theorem of Eshelby tensors for exterior points. As a by-product, this paper also proposes an alternative method to prove the quasi-Eshelby property obtained in Ref. [21]. Furthermore, two corollaries are presented to show more special properties of the Eshelby tensors for the rotational symmetrical inclusion. The first corollary shows that the line integral average of Eshelby tensors along a concentric circle $\Gamma_c(r,\mathbf{o})$, which is located outside the inclusion, equals zero. The second corollary indicates that the area average of Eshelby tensors over the domain bounded by the circle $\Gamma_c(r,\mathbf{o})$ is in direct proportion to the area of the inclusion and in inverse proportion to the area of the circle. Hopefully, these properties may be used in the evaluation of effective material properties of composites. It should be noticed that due to the quality of the Green function for plane strain problems, the fourfold rotational symmetrical inclusions are deprived of the arithmetic mean theorem.

In order to validate the arithmetic mean theorem for exterior points, the Eshelby problem with a circular inclusion is considered by using the solution of the elastic field for exterior points in Ref. [2]. Since a circle can be treated as an arbitrary fold rotational symmetrical figure, the arithmetic mean theorem implies that the arithmetic mean of Eshelby tensors at any N rotational symmetrical points outside the circular inclusion should be equal to zero, as long as $N \geqslant 3$ and $N \neq 4$. Taking advantage of the elastic solution of the exterior field for the Eshelby problem with an ellipsoidal inclusion by Ferrers [24] and Dyson [25], we obtained the explicit solution for the circular inclusion problem. Then, by the explicit solution, we check the arithmetic mean of Eshelby tensors at any N rotational symmetrical points, and it turns out that the specified arithmetic means are all equal to zero, under the condition of $N \geqslant 3$ and $N \neq 4$. These results totally conform to the arithmetic mean theorem of Eshelby tensors for exterior points.

The paper is organized as follows: By the Green formula we first express the Eshelby tensors for any-shaped inclusions by two kinds of line integrals along the boundary of the inclusion in Sec. 2. Then in Sec. 3, by introducing the complex plane, we transform the two contour integrals into the complex integral representation and obtain some universal relations. In Sec. 4, on the basis of the complex integral representation, we discuss the rotational symmetrical inclusion and prove the arithmetic mean theorems for both exterior points and interior points. Two corollaries are also provided in this section. In Sec. 5, we test the arithmetic mean theorem for exterior points by considering a circular inclusion and using the available solutions. Finally, a brief summary and some remarks are given in Sec. 6.

2 Basic Equations

Consider an infinite, homogeneous, and isotropic matrix consisting of a homogeneous inclusion Ω^+ with the boundary Γ. The domain outside the inclusion is denoted as Ω^-. It is assumed that the inclusion shares the same elastic moduli with the matrix and is prescribed a uniform eigenstrain ε_{ij}^*. The Cartesian coordinate system (x_1,x_2,x_3) is adopted and the plane strain condition for $x_1 - x_2$ is assumed in this paper. Consequently, all subscripts in the following equations i,j,k,l,m,n range from 1 to 2. We denote by ν and μ the Poisson's ratio and shear modulus, respectively. The elastic stiffness tensor C_{ijkl} is of the form

$$C_{ijkl} = \lambda \delta_{ij}\delta_{kl} + \mu(\delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl}) \tag{1}$$

where $\lambda = 2\mu\nu/(1-2\nu)$ and δ_{ij} is the Kronecker's delta. The displacement field induced by the uniform eigenstrain ε_{ij}^* can be written as (see Ref. [2])

$$u_i(\mathbf{x}) = -C_{lkmn}\varepsilon_{mn}^* \int_{\Omega^+} \partial_{x_k} G_{il}(\mathbf{x}-\mathbf{x}')d\mathbf{x}' \tag{2}$$

where $\mathbf{x}$ is the observation point, inside or outside the inclusion, and $\mathbf{x}'$ is the point inside Ω^+. In this paper, the repeated indices imply summation. $G_{ij}(\mathbf{x}-\mathbf{x}')$ is the Green function for plane strain problems and takes the form

$$G_{il}(\mathbf{x}-\mathbf{x}') = \frac{1}{8\pi(1-\nu)\mu}\left[(3-4\nu)\delta_{il}\log\frac{1}{|\mathbf{x}-\mathbf{x}'|} + \frac{(x_i - x_i')(x_l - x_l')}{|\mathbf{x}-\mathbf{x}'|^2}\right] \tag{3}$$

By the Green formula, Eq. (2) yields

$$u_i(\mathbf{x}) = C_{lkmn}\varepsilon_{mn}^* \oint_{\Gamma} G_{il}(\mathbf{x}-\mathbf{x}')n_k ds_{\mathbf{x}'} \tag{4}$$

where n_k is the kth component of unit outer normal of Γ. Substituting the last equation into the geometric differential relations between strains and displacements, we obtain the perturbed strain field by the eigenstrains ε_{ij}^*

$$\varepsilon_{ij}(\mathbf{x}) = S_{ijmn}(\mathbf{x})\varepsilon_{mn}^* \tag{5}$$

with

$$S_{ijmn}(\mathbf{x}) = \frac{C_{lkmn}}{2}\oint_{\Gamma} n_k[\partial_{x_j}G_{il}(\mathbf{x}-\mathbf{x}') + \partial_{x_i}G_{jl}(\mathbf{x}-\mathbf{x}')]ds_{\mathbf{x}'} \tag{6}$$

$S_{ijmn}(\mathbf{x})$ is known as the Eshelby tensor. Substituting the Green function in Eq. (3) into the Eshelby tensor, one has

$$S_{ijmn}(\mathbf{x}) = \frac{C_{lkmn}}{8\pi\mu(1-\nu)}\{2J_{ijlk}(\mathbf{x}) - \delta_{ij}J_{lk}(\mathbf{x}) + (1-2\nu)[\delta_{il}J_{jk}(\mathbf{x}) + \delta_{jl}J_{ik}(\mathbf{x})]\} \tag{7}$$

where

$$J_{ij}(\mathbf{x}) = \oint_{\Gamma} \frac{\rho_i}{\rho^2} n_j ds_{\mathbf{x}'} \tag{8}$$

$$J_{ijlk}(\mathbf{x}) = \oint_{\Gamma} \frac{\rho_i\rho_j\rho_l}{\rho^4} n_k ds_{\mathbf{x}'} \tag{9}$$

with $\rho_i = x_i' - x_i$ and $\rho = |\mathbf{x}'-\mathbf{x}|$.

By the Green formula, the strain field for the Eshelby problem of an arbitrary inclusion with uniform eigenstrains is given in the form of the integral expressions. Equation (7) indicates the Eshelby tensor only depends on the two integrals $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$. Hence the discussion of the Eshelby tensor, and thus the strain field, is reduced to the evaluation of $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$.

3 The Complex Integral Representation of $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$ for Arbitrary Inclusions

We need further discussions on $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$ shown in Eqs. (8) and (9), respectively, since the Eshelby tensors are fully dependent on these two kinds of integrals. In this section, we will obtain the complex integral expressions of $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$ which will be used in the next section to derive the arithmetic mean theorem of Eshelby tensor for the rotational symmetrical inclusions.

Taking advantage of the identities $n_1 ds_{\mathbf{x}'} = dx_2'$ and $n_2 ds_{\mathbf{x}'} = -dx_1'$, we can rewrite $J_{ij}(\mathbf{x})$ from Eq. (8) as follows

$$J_{12}(\mathbf{x}) = -\oint_\Gamma \frac{\rho_1}{\rho^2} dx_1', J_{11}(\mathbf{x}) = \oint_\Gamma \frac{\rho_1}{\rho^2} dx_2' \tag{10}$$

and

$$J_{22}(\mathbf{x}) = -\oint_\Gamma \frac{\rho_2}{\rho^2} dx_1', J_{21}(\mathbf{x}) = \oint_\Gamma \frac{\rho_2}{\rho^2} dx_2' \tag{11}$$

Introduce the complex plane $z = x_1 + ix_2$ with $i^2 = -1$. Then in the complex plane, z and z' represent the points $\mathbf{x}$ and $\mathbf{x}'$, respectively. Thus, Eqs. (10) and (11) lead to

$$-J_{12}(\mathbf{x}) + iJ_{11}(\mathbf{x}) = \oint_\Gamma \frac{\rho_1}{\rho^2} dz' \tag{12}$$

$$-J_{22}(\mathbf{x}) + iJ_{21}(\mathbf{x}) = \oint_\Gamma \frac{\rho_2}{\rho^2} dz' \tag{13}$$

where $z' = x_1' + ix_2'$ is the point on the boundary of the inclusion. Equations (12) and (13) are called the complex integral representation of $J_{ij}(\mathbf{x})$. Similarly, one can obtain the complex integral representation of $J_{ijlk}(\mathbf{x})$ from Eq. (9) as follows

$$-J_{1112}(\mathbf{x}) + iJ_{1111}(\mathbf{x}) = \oint_\Gamma \frac{\rho_1^3}{\rho^4} dz' \tag{14}$$

$$-J_{2222}(\mathbf{x}) + iJ_{2221}(\mathbf{x}) = \oint_\Gamma \frac{\rho_2^3}{\rho^4} dz' \tag{15}$$

$$-J_{1122}(\mathbf{x}) + iJ_{1121}(\mathbf{x}) = \oint_\Gamma \frac{\rho_1^2\rho_2}{\rho^4} dz' \tag{16}$$

$$-J_{2212}(\mathbf{x}) + iJ_{2211}(\mathbf{x}) = \oint_\Gamma \frac{\rho_1\rho_2^2}{\rho^4} dz' \tag{17}$$

Other components of $J_{ijlk}(\mathbf{x})$ can be given using the four obvious relations:

$$J_{1122}(\mathbf{x}) = J_{1212}(\mathbf{x}) = J_{2112}(\mathbf{x}) \tag{18}$$

$$J_{1121}(\mathbf{x}) = J_{1211}(\mathbf{x}) = J_{2111}(\mathbf{x}) \tag{19}$$

$$J_{2212}(\mathbf{x}) = J_{2122}(\mathbf{x}) = J_{1222}(\mathbf{x}) \tag{20}$$

$$J_{2211}(\mathbf{x}) = J_{2121}(\mathbf{x}) = J_{1221}(\mathbf{x}) \tag{21}$$

The complex integral representation of $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$ enables us to take advantage of complex variable function theory in the analysis of Eshelby problems. Next, we will find the universal relations for $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$ in virtue of the complex integral representation.

By $z' - z = \rho_1 + i\rho_2$, we note that

$$\oint_\Gamma \frac{1}{z' - z} dz' = \oint_\Gamma \frac{(\bar{z}' - \bar{z})}{(z' - z)(\bar{z}' - \bar{z})} dz' = \oint_\Gamma \frac{\rho_1 - i\rho_2}{\rho^2} dz' \tag{22}$$

Replacing Eq. (22) by Eqs. (12) and (13) and simplifying the result, one obtains

$$\oint_\Gamma \frac{1}{z' - z} dz' = [J_{21}(\mathbf{x}) - J_{12}(\mathbf{x})] + i[J_{11}(\mathbf{x}) + J_{22}(\mathbf{x})] \tag{23}$$

From the Cauchy integral formula [26], Eq. (23) yields

$$[J_{21}(\mathbf{x}) - J_{12}(\mathbf{x})] + i[J_{11}(\mathbf{x}) + J_{22}(\mathbf{x})] = 2\pi i\delta(\mathbf{x}) \tag{24}$$

where $\delta(\mathbf{x})$ is the characteristic function of the inclusion

$$\delta(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega^+ \\ 0, & \mathbf{x} \in \Omega^- \end{cases} \tag{25}$$

In the same way, we have

$$\oint_\Gamma \frac{1}{z' - z} dz' = \oint_\Gamma \frac{(z' - z)(\bar{z}' - \bar{z})^2}{(z' - z)^2(\bar{z}' - \bar{z})^2} dz' = \oint_\Gamma \frac{(\rho_1 + i\rho_2)(\rho_1 - i\rho_2)^2}{\rho^4} dz' \tag{26}$$

By expanding Eq. (26), replacing it by Eqs. (14)–(17), and then simplifying the result, one has

$$\oint_\Gamma \frac{1}{z' - z} dz' = [-J_{1112}(\mathbf{x}) - J_{2212}(\mathbf{x}) + J_{1121}(\mathbf{x}) + J_{2221}(\mathbf{x})] + i[J_{1111}(\mathbf{x}) + J_{2211}(\mathbf{x}) + J_{1122}(\mathbf{x}) + J_{2222}(\mathbf{x})] \tag{27}$$

By the Cauchy integral formula, Eq. (27) leads to

$$[-J_{1112}(\mathbf{x}) - J_{2212}(\mathbf{x}) + J_{1121}(\mathbf{x}) + J_{2221}(\mathbf{x})] + i[J_{1111}(\mathbf{x}) + J_{2211}(\mathbf{x}) + J_{1122}(\mathbf{x}) + J_{2222}(\mathbf{x})] = 2\pi i\delta(\mathbf{x}) \tag{28}$$

Equations (24) and (28) hold for arbitrarily shaped inclusion, as long as it is a simply connected domain. Hence we refer to Eq. (24) as the universal relation of $J_{ij}(\mathbf{x})$, and Eq. (28) as the universal relation of $J_{ijlk}(\mathbf{x})$.

Before ending this section, we provide several equations which will be used in Sec. 4. Introduce three complex variable functions of the form

$$L_1(z) = \oint_\Gamma \frac{(z' - z)}{(z' - z)(\bar{z}' - \bar{z})} dz' = \oint_\Gamma \frac{(z' - z)^2(\bar{z}' - \bar{z})}{(z' - z)^2(\bar{z}' - \bar{z})^2} dz' \tag{29}$$

$$L_2(z) = \oint_\Gamma \frac{(\bar{z}' - \bar{z})^3}{(z' - z)^2(\bar{z}' - \bar{z})^2} dz' \tag{30}$$

$$L_3(z) = \oint_\Gamma \frac{(z' - z)^3}{(z' - z)^2(\bar{z}' - \bar{z})^2} dz' \tag{31}$$

Considering $z' - z = \rho_1 + i\rho_2$ and $\bar{z}' - \bar{z} = \rho_1 - i\rho_2$, and then employing Eqs. (12) and (13) one gets from the first equation in Eq. (29)

$$L_1(z) = [-J_{12}(\mathbf{x}) - J_{21}(\mathbf{x})] + i[J_{11}(\mathbf{x}) - J_{22}(\mathbf{x})] \tag{32}$$

Similarly, by Eqs. (14)–(17), we get from the second equation in Eq. (29) and Eqs. (30) and (31), respectively,

$$L_1(z) = [-J_{1112}(\mathbf{x}) - J_{2212}(\mathbf{x}) - J_{2221}(\mathbf{x}) - J_{1121}(\mathbf{x})] + i[J_{1111}(\mathbf{x}) + J_{2211}(\mathbf{x}) - J_{2222}(\mathbf{x}) - J_{1122}(\mathbf{x})] \tag{33}$$

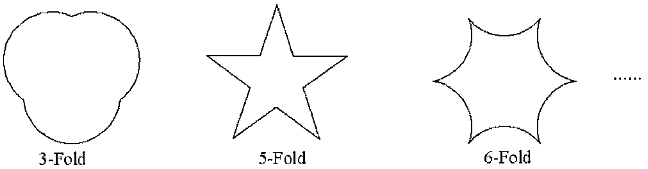


Fig. 1 Rotational symmetrical inclusions

$$L_2(z) = [-J_{1112}(\mathbf{x}) + 3J_{2212}(\mathbf{x}) - J_{2221}(\mathbf{x}) + 3J_{1121}(\mathbf{x})] + i[J_{1111}(\mathbf{x}) - 3J_{2211}(\mathbf{x}) - J_{2222}(\mathbf{x}) + 3J_{1122}(\mathbf{x})] \quad (34)$$

$$L_3(z) = [-J_{1112}(\mathbf{x}) + 3J_{2212}(\mathbf{x}) + J_{2221}(\mathbf{x}) - 3J_{1121}(\mathbf{x})] + i[J_{1111}(\mathbf{x}) - 3J_{2211}(\mathbf{x}) + J_{2222}(\mathbf{x}) - 3J_{1122}(\mathbf{x})] \quad (35)$$

The definition of $L_i(z), i=1,2,3$ shows that the Cauchy integral formula is not valid for the three functions. However, they contain all the components of $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$. As we will see in Sec. 4, $L_1(z)$, $L_2(z)$, and $L_3(z)$ play critical roles in the calculation of the arithmetic mean of Eshelby tensors for the rotational symmetrical inclusions.

4 The Arithmetic Mean of Eshelby Tensors for the Rotational Symmetrical Inclusions

In this section, we consider an N-fold ($N \geqslant 3$ and $N \neq 4$) rotational symmetrical inclusion, see Fig. 1. That is, Ω^+ in this section is a rotational symmetrical inclusion. This is a new topic with special consideration on the shape of inclusion and of practice interests since the cross section of most reinforcements in composites and most quantum wires in strained semiconductor devices has a boundary composed of uniform curves with specified rotational symmetry. Especially, the regular polygon inclusions, a special kind of planar rotational symmetrical figure, have been found in many nanostructures and studied by many researchers.

Suppose that $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$ are N rotational symmetrical points which can be located inside or outside the inclusion, see Fig. 2. We denote by $\bar{S}^a_{ijmn}(\mathbf{x})$ the arithmetic mean of the Eshelby tensors at $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$, i.e.,

$$\bar{S}^a_{ijmn}(\mathbf{x}) = \frac{1}{N}\sum_{p=1}^{N} S_{ijmn}(\mathbf{x}^p) \quad (36)$$

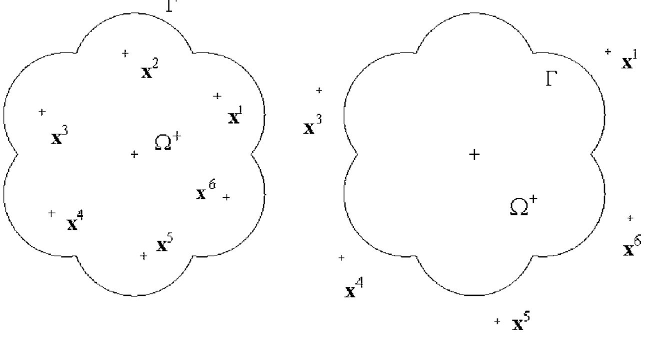


Fig. 2 Rotational symmetrical points

where $p=1,2,\ldots,N$, and

$$\begin{pmatrix} x_1^p \\ x_2^p \end{pmatrix} = \begin{bmatrix} \cos\dfrac{2\pi}{N}(p-1) & \sin\dfrac{2\pi}{N}(p-1) \\ -\sin\dfrac{2\pi}{N}(p-1) & \cos\dfrac{2\pi}{N}(p-1) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (37)$$

Equation (37) shows $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$ are N points which are defined from $\mathbf{x}$. Also, Eq. (37) tells us that $\mathbf{x}$ is the same point as $\mathbf{x}^1$. In fact, due to the symmetry, $\mathbf{x}$ can represent any point of $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$. By substituting Eq. (7) into Eq. (36), $\bar{S}^a_{ijmn}(\mathbf{x})$ can be written as

$$\bar{S}^a_{ijmn}(\mathbf{x}) = \frac{C_{lkmn}}{8\pi\mu(1-\nu)}\{2\bar{J}^a_{ijlk}(\mathbf{x}) - \delta_{ij}\bar{J}^a_{lk}(\mathbf{x}) + (1-2\nu)[\delta_{il}\bar{J}^a_{jk}(\mathbf{x}) + \delta_{jl}\bar{J}^a_{ik}(\mathbf{x})]\} \quad (38)$$

where

$$\bar{J}^a_{ij}(\mathbf{x}) = \frac{1}{N}\sum_{p=1}^{N} J_{ij}(\mathbf{x}^p) \quad (39)$$

$$\bar{J}^a_{ijlk}(\mathbf{x}) = \frac{1}{N}\sum_{p=1}^{N} J_{ijlk}(\mathbf{x}^p) \quad (40)$$

In the complex plane, the rotational symmetrical points $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$ are denoted by $z_1, z_2, \ldots, z_N$, respectively. Then, the rotational symmetry (37) is transformed to

$$z_k = \omega^{k-1} z, \quad k = 1, 2, \ldots, N \quad (41)$$

where $\omega = \exp(i(2\pi/N))$. Employing the definition of $L_1(z)$, the integral variable substitution $z' = \omega^{k-1}\hat{z}$ and considering the rotational symmetry of the inclusion boundary, we have

$$L_1(z_k) = \oint_\Gamma \frac{1}{\bar{z}' - \bar{\omega}^{k-1}\bar{z}} dz' = \omega^{2(k-1)} \oint_\Gamma \frac{1}{\bar{\hat{z}} - \bar{z}} d\hat{z} = \omega^{2(k-1)} L_1(z) \quad (42)$$

Following the similar procedure, one has

$$L_2(z_k) = \bar{\omega}^{2(k-1)} L_2(z) \quad (43)$$

$$L_3(z_k) = \omega^{4(k-1)} L_3(z) \quad (44)$$

Thus, when $N \geqslant 3$ and $N \neq 4$, the following equations hold

$$\frac{1}{N}\sum_{k=1}^{N} L_1(z_k) = \frac{1}{N}[1 + \omega^2 + \cdots + \omega^{2(N-1)}] L_1(z) = 0 \quad (45)$$

$$\frac{1}{N}\sum_{k=1}^{N} L_2(z_k) = \frac{1}{N}[1 + \bar{\omega}^2 + \cdots + \bar{\omega}^{2(N-1)}] L_2(z) = 0 \quad (46)$$

$$\frac{1}{N}\sum_{k=1}^{N} L_3(z_k) = \frac{1}{N}[1 + \omega^4 + \cdots + \omega^{4(N-1)}] L_3(z) = 0 \quad (47)$$

By using the formula of $L_1(z)$ in Eq. (32), Eq. (45) yields

$$[-\bar{J}^a_{12}(\mathbf{x}) - \bar{J}^a_{21}(\mathbf{x})] + i[\bar{J}^a_{11}(\mathbf{x}) - \bar{J}^a_{22}(\mathbf{x})] = 0 \quad (48)$$

Similarly, utilizing the formulas of $L_i(z), i=1,2,3$ in Eqs. (33)–(35), we obtain from Eqs. (45)–(47), respectively

$$[-\bar{J}^a_{1112}(\mathbf{x}) - \bar{J}^a_{2212}(\mathbf{x}) - \bar{J}^a_{2221}(\mathbf{x}) - \bar{J}^a_{1121}(\mathbf{x})] + i[J^a_{1111}(\mathbf{x}) + \bar{J}^a_{2211}(\mathbf{x}) - \bar{J}^a_{2222}(\mathbf{x}) - \bar{J}^a_{1122}(\mathbf{x})] = 0 \quad (49)$$

$$[-\bar{J}^a_{1112}(\mathbf{x}) + 3\bar{J}^a_{2212}(\mathbf{x}) - \bar{J}^a_{2221}(\mathbf{x}) + 3\bar{J}^a_{1121}(\mathbf{x})] + i[J^a_{1111}(\mathbf{x}) - 3\bar{J}^a_{2211}(\mathbf{x}) - \bar{J}^a_{2222}(\mathbf{x}) + 3\bar{J}^a_{1122}(\mathbf{x})] = 0 \quad (50)$$

$$[-\bar{J}^a_{1112}(\mathbf{x})+3\bar{J}^a_{2212}(\mathbf{x})+\bar{J}^a_{2221}(\mathbf{x})-3\bar{J}^a_{1121}(\mathbf{x})] + i[\bar{J}^a_{1111}(\mathbf{x})-3\bar{J}^a_{2211}(\mathbf{x})+\bar{J}^a_{2222}(\mathbf{x})-3\bar{J}^a_{1122}(\mathbf{x})]=0 \quad (51)$$

Furthermore, using the universal relations of $J_{ij}(\mathbf{x})$ and $J_{ijlk}(\mathbf{x})$, i.e., Eqs. (24) and (28), one obtains, respectively

$$[\bar{J}^a_{21}(\mathbf{x})-\bar{J}^a_{12}(\mathbf{x})]+i[\bar{J}^a_{11}(\mathbf{x})+\bar{J}^a_{22}(\mathbf{x})]=2\pi i\delta(\mathbf{x}) \quad (52)$$

and

$$[-\bar{J}^a_{1112}(\mathbf{x})-\bar{J}^a_{2212}(\mathbf{x})+\bar{J}^a_{1121}(\mathbf{x})+\bar{J}^a_{2221}(\mathbf{x})] + i[\bar{J}^a_{1111}(\mathbf{x})+\bar{J}^a_{2211}(\mathbf{x})+\bar{J}^a_{1122}(\mathbf{x})+\bar{J}^a_{2222}(\mathbf{x})]=2\pi i\delta(\mathbf{x}) \quad (53)$$

Combining Eqs. (48) and (52), we obtain all components of $\bar{J}^a_{ij}(\mathbf{x})$

$$\bar{J}^a_{ij}(\mathbf{x})=\pi\delta_{ij}\delta(\mathbf{x}) \quad (54)$$

Similarly, the combination of Eqs. (49)–(51) and (53) gives all components of $\bar{J}^a_{ijlk}(\mathbf{x})$

$$\bar{J}^a_{ijlk}(x)=\frac{\pi}{4}(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})\delta(\mathbf{x}) \quad (55)$$

Substituting Eqs. (54) and (55) into Eq. (38), we obtain the arithmetic mean of the Eshelby tensors at $\mathbf{x}^1,\mathbf{x}^2,\ldots,\mathbf{x}^N$.

THEOREM. *When* Ω^+ *is an N-fold* ($N\geqslant 3$ *and* $N\neq 4$) *rotational symmetrical inclusion*

$$\bar{S}^a_{ijmn}(\mathbf{x})=\left[\frac{4\nu-1}{8(1-\nu)}\delta_{ij}\delta_{mn}+\frac{3-4\nu}{8(1-\nu)}(\delta_{in}\delta_{jm}+\delta_{im}\delta_{jn})\right]\delta(\mathbf{x}) \quad (56)$$

where $\bar{S}^a_{ijmn}(\mathbf{x})$ *is defined in Eq.* (36), *and* $\delta(\mathbf{x})$ *in Eq.* (25). ■

The theorem is given the name of the arithmetic mean theorem of Eshelby tensors, which shows that although the elastic field induced by an *N*-fold rotational symmetrical inclusion with uniform eigenstrains is nonuniform, the arithmetic mean of the strains at *N* rotational symmetrical points have attractive properties. It shows that the arithmetic mean of the Eshelby tensors at *N* rotational symmetrical interior points is the same as the Eshelby tensor for a circular inclusion. This arithmetic mean theorem of Eshelby tensors for interior points, also referred to as the quasi-Eshelby property, has been acquired in a different method in our previous works [21–23].

Remarkably, Eq. (56) indicates that for any *N*-fold ($N\geqslant 3$ and $N\neq 4$) rotational symmetrical inclusion, the arithmetic mean of the Eshelby tensors at rotational symmetrical points outside the inclusion is equal to zero. As we know, while the elastic field within the ellipsoidal inclusion with uniform eigenstrains is uniform, the elastic field outside the inclusion does not possess any kind of special property, according to the existing literature. Thus, in some sense, the discovery of arithmetic mean theorem for exterior points makes *N*-fold ($N\geqslant 3$ and $N\neq 4$) rotational symmetrical inclusions more preferable than ellipsoidal inclusions. In the next section, the Eshelby problem with a circular inclusion will be used to check the arithmetic mean theorem for exterior points. Here, we present two corollaries from the theorem.

Let $\Gamma_c(r,\mathbf{o})$ be a circle which is totally outside the inclusion and concentric with the inclusion, and the domain bounded by $\Gamma_c(r,\mathbf{o})$ is denoted by Ω_c, see Fig. 3. Parameter r is the radius and $\mathbf{O}$ represents the center of the rotational symmetrical inclusion domain Ω^+. From the theorem, it is not difficult to obtain the first corollary:

COROLLARY 1. *For any N-fold* ($N\geqslant 3$ *and* $N\neq 4$) *rotational symmetrical inclusion, the line integral average of Eshelby tensors along* $\Gamma_c(r,\mathbf{o})$ *is equal to zero, i.e.*

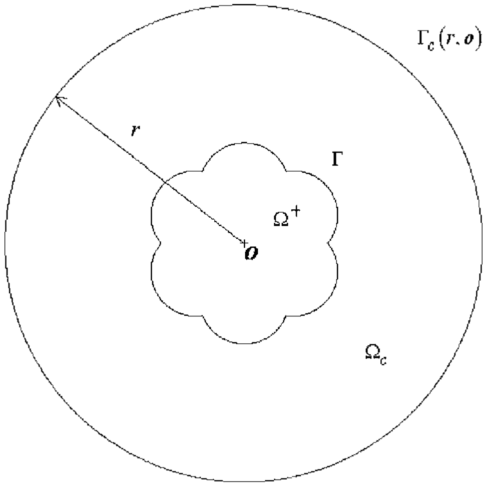


Fig. 3 A concentric circle outside the inclusion

$$\bar{S}^c_{ijmn}=\frac{1}{2\pi r}\oint_{\Gamma_c(r,\mathbf{O})}S_{ijmn}(\mathbf{x})ds=0 \quad (57)$$

where $\Gamma_c(r,\mathbf{o})$ *is a concentric circle which is totally outside the inclusion, as shown in Fig.* 3. ■

It is easy to know from the theorem that the average Eshelby tensor over the inclusion domain Ω^+ is equal to the Eshelby tensor for the circular inclusion. On the other hand, the arithmetic mean theorem for exterior points implies that the area integral of Eshelby tensors over the domain $\Omega_c-\Omega^+$ is equal to zero. Taking advantage of these two facts, we obtain the area average of Eshelby tensor over Ω_c.

COROLLARY 2. *For any N-fold* ($N\geqslant 3$ *and* $N\neq 4$) *rotational symmetrical inclusion, the following equations hold*

$$\bar{S}_{ijmn}(r)=\frac{\Sigma}{\pi r^2}\left[\frac{4\nu-1}{8(1-\nu)}\delta_{ij}\delta_{mn}+\frac{3-4\nu}{8(1-\nu)}(\delta_{in}\delta_{jm}+\delta_{im}\delta_{jn})\right] \quad (58)$$

where $\bar{S}_{ijmn}(r)$ *is the area average of Eshelby tensors over* Ω_c, *and* Σ *is the area of the inclusion* Ω^+. ■

Corollary 1 shows that the line integral average of Eshelby tensors along the concentric circle $\Gamma_c(r,\mathbf{o})$ is equal to zero. Corollary 2 implies that the area average of Eshelby tensors over the domain bounded by the circle $\Gamma_c(r,\mathbf{o})$ is in direct proportion to the area of the inclusion and in inverse proportion to the area of the circle. The two corollaries are presented here to encourage the possibility of application of the arithmetic mean theorem. For example, on the basis of the equivalent inclusion method [2] and the various approximate schemes [27–30], these special properties can be expected to be useful in the estimation of macroscopic mechanics properties of the composites.

5 Arithmetic Mean of Eshelby Tensors for Circular Inclusions

In the last section, we have proved that for any *N*-fold ($N\geqslant 3$ and $N\neq 4$) rotational symmetrical inclusion, the arithmetic mean of Eshelby tensors at *N* rotational symmetrical points presents special properties. The detailed discussions and remarks have been given on the interior points in our former work [21]. However, the result for exterior points, which shows that arithmetic mean of Eshelby tensors at *N* rotational symmetrical points outside the inclusion equals zero, has not been paid enough attention. In order to remove the doubts from the readers, in this section the circular inclusion is used to test the validity of the arithmetic mean theorem for the exterior points. It is clear that a circle can be treated as an arbitrary fold rotational symmetrical figure. Hence, the arithmetic mean theorem implies that the arithmetic mean of Eshelby tensors at any N rotational symmetrical points outside the circular inclusion should be equal to zero, as long as $N \geqslant 3$ and $N \neq 4$.

To make the testing process convincing, we will take the following steps: We start from the elastic solution of the exterior field for the Eshelby problem with an ellipsoidal inclusion, which has been obtained by Ferrers [24] and Dyson [25] and shown in Mura's monograph [2]. From the elastic solution for ellipsoidal inclusions, we obtain the explicit solution for circular inclusions. Then, by this explicit solution, we consider any N ($N \geqslant 3$ and $N \neq 4$) rotational symmetrical points outside the inclusion and calculate the arithmetic mean of Eshelby tensors at these points. According to the arithmetic mean theorem, this arithmetic mean has to equal zero. Otherwise, the arithmetic mean theorem is not valid.

For the sake of the completeness, we first introduce the elastic solution of the exterior field for the Eshelby problem with an ellipsoidal inclusion by Ferrers and Dyson. Let a_1, a_2, a_3 be the principal half axes of the ellipsoidal inclusion. Ferrers and Dyson gave the following elliptic integrals

$$I_i(v) = 2\pi a_1 a_2 a_3 \int_v^\infty \frac{ds}{(a_i^2 + s)\,\triangle(s)} \tag{59}$$

$$I_{ij}(v) = 2\pi a_1 a_2 a_3 \int_v^\infty \frac{ds}{(a_i^2 + s)(a_j^2 + s)\,\triangle(s)} \tag{60}$$

where $\triangle(s) = \sqrt{(a_1^2+s)(a_2^2+s)(a_3^2+s)}$, and v is the largest positive root of the equation

$$\frac{x_1^2}{(a_1^2 + \lambda)} + \frac{x_2^2}{(a_2^2 + \lambda)} + \frac{x_3^2}{(a_3^2 + \lambda)} = 1 \tag{61}$$

As shown in Chapter 2 in Ref. [2], the induced strain field outside the ellipsoidal inclusion can be written as

$$\varepsilon_{ij}(\mathbf{x}) = S_{ijkl}(\mathbf{x})\varepsilon_{kl}^{*} \tag{62}$$

where

$$\begin{aligned} 8\pi(1-\nu)S_{ijkl}(\mathbf{x}) = {} & 8\pi(1-\nu)D_{ijkl}(\lambda) + 2\nu\delta_{kl}x_i I_{I,j}(\lambda) + (1-\nu) \\ & \times\{\delta_{il}x_k I_{K,j}(\lambda) + \delta_{jl}x_k I_{K,i}(\lambda) + \delta_{ik}x_l I_{L,j}(\lambda) \\ & + \delta_{jk}x_l I_{L,i}(\lambda)\} - \delta_{ij}x_k[I_K(\lambda) - a_I^2 I_{KI}(\lambda)]_{,l} \\ & - (\delta_{ik}x_j + \delta_{jk}x_i)[I_J(\lambda) - a_I^2 I_{IJ}(\lambda)]_{,l} - (\delta_{il}x_j \\ & + \delta_{jl}x_i)[I_J(\lambda) - a_I^2 I_{IJ}(\lambda)]_{,k} - x_i x_j[I_J(\lambda) \\ & - a_I^2 I_{IJ}(\lambda)]_{,lk} \end{aligned} \tag{63}$$

with

$$\begin{aligned} 8\pi(1-\nu)D_{ijkl}(\lambda) = {} & \delta_{ij}\delta_{kl}[2\nu I_I(\lambda) - I_K(\lambda) + a_I^2 I_{KI}(\lambda)] \\ & + (\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il})\{a_I^2 I_{IJ}(\lambda) - I_J(\lambda) \\ & + (1-\nu)[I_K(\lambda) + I_L(\lambda)]\} \end{aligned} \tag{64}$$

In Eqs. (63) and (64), upper case indices take on the same numbers as the corresponding lower case ones but are not summered.

Next, we reduce the ellipsoidal inclusion to a cylindrical inclusion by supposing $a_1 = a_2 = a$ and $a_3 \to +\infty$. Therefore, Eq. (61) yields

$$\lambda = x_1^2 + x_2^2 - a^2 = r^2 - a^2 \tag{65}$$

Correspondingly, Eqs. (59) and (60) give

$$I_1(\lambda) = I_2(\lambda) = 2\pi\frac{a^2}{r^2}, \quad I_3(\lambda) = 0 \tag{66}$$

$$I_{11}(\lambda) = I_{22}(\lambda) = I_{12}(\lambda) = \pi\frac{a^2}{r^4} \tag{67}$$

$$I_{13}(\lambda) = I_{23}(\lambda) = I_{33}(\lambda) = 0 \tag{68}$$

$$a_3^2 I_{13}(\lambda) = I_1(\lambda), \quad a_3^2 I_{23}(\lambda) = I_2(\lambda), \quad a_3^2 I_{33}(\lambda) = 0 \tag{69}$$

By Eqs. (65)–(69), Eqs. (63) and (64) can be reduced to the solution for circular inclusions. Take $S_{1111}(\mathbf{x})$ as an example

$$\begin{aligned} 8\pi(1-\nu)S_{1111}(\mathbf{x}) = {} & 2\pi(1-2\nu)\frac{a^2}{r^2} + 3\pi\frac{a^4}{r^4} + (8\nu+8)\frac{a^2x_1^2}{r^4} \\ & - 24\pi\frac{a^4x_1^2}{r^6} - 16\pi\frac{a^2x_1^4}{r^6} + 24\pi\frac{a^4x_1^4}{r^8} \end{aligned} \tag{70}$$

By using the polar coordination system, $S_{1111}(\mathbf{x})$ can be rewritten as

$$\begin{aligned} 8\pi(1-\nu)S_{1111}(\mathbf{x}) = {} & 2\pi(1-2\nu)\frac{a^2}{r^2} + 3\pi\frac{a^4}{r^4} + (8\nu+8)\pi\frac{a^2\cos^2\theta}{r^2} \\ & - 24\pi\frac{a^4\cos^2\theta}{r^4} - 16\pi\frac{a^2\cos^4\theta}{r^2} \\ & + 24\pi\frac{a^4\cos^4\theta}{r^4} \end{aligned} \tag{71}$$

Taking advantage of two trigonometric identities

$$\cos^2\theta = \tfrac{1}{2}(1 + \cos 2\theta) \tag{72}$$

$$\cos^4\theta = \tfrac{3}{8} + \tfrac{1}{2}\cos 2\theta + \tfrac{1}{8}\cos 4\theta \tag{73}$$

we write $S_{1111}(\mathbf{x})$ as

$$\begin{aligned} 8\pi(1-\nu)S_{1111}(\mathbf{x}) = {} & 4(\nu-1)\pi\frac{a^2\cos 2\theta}{r^2} - 2\pi\frac{a^2\cos 4\theta}{r^2} \\ & + 3\pi\frac{a^4\cos 4\theta}{r^4} \end{aligned} \tag{74}$$

As previously mentioned, $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$ are N rotational symmetrical points defined from $\mathbf{x}$ by Eq. (37). Employing Eq. (74), we have

$$8\pi(1-\nu)\bar{S}^a_{1111}(\mathbf{x}) = 4(\nu-1)\pi\frac{a^2}{r^2}H_1(\theta) - 2\pi\frac{a^2}{r^2}H_2(\theta) + 3\pi\frac{a^4}{r^4}H_2(\theta) \tag{75}$$

where

$$H_1(\theta) = \frac{1}{N}\sum_{p=1}^{N}\cos\left[2\theta + \frac{4\pi}{N}(p-1)\right] \tag{76}$$

$$H_2(\theta) = \frac{1}{N}\sum_{p=1}^{N}\cos\left[4\theta + \frac{8\pi}{N}(p-1)\right] \tag{77}$$

For any angle ϕ, it is clear that

$$\begin{aligned} & \sum_{p=1}^{M}\cos\left[\phi + \frac{2L\pi}{M}(p-1)\right] + i\sum_{p=1}^{M}\cos\left[\phi + \frac{2L\pi}{M}(p-1)\right] \\ & = e^{i\phi}\left(1 + \exp\left(i\frac{2L\pi}{M}\right) + \cdots + \exp\left(i\frac{2L\pi}{M}(M-1)\right)\right) \\ & = e^{i\phi}\frac{1 - e^{i2L\pi}}{1 - \exp\left(\dfrac{i2L\pi}{M}\right)} = 0 \end{aligned} \tag{78}$$

where M, L are integers while L/M is not an integer. When $N \geqslant 3$ and $N \neq 4$, from Eq. (78) we observe $H_1(\theta) = H_2(\theta) = 0$ for any

θ. Therefore, Eq. (75) gives $\bar{S}^a_{1111}(\mathbf{x})=0$. Similarly, we can prove other components of $\bar{S}^a_{ijkl}(\mathbf{x})$ also vanish. In other words, for circular inclusions, the arithmetic mean of Eshelby tensors at N rotational symmetrical points outside the inclusion equals zero as long as $N\geqslant 3$ and $N\neq 4$. Consequently, the validity of the arithmetic mean theorem is checked safely.

6 Conclusions

The Eshelby problems with two-dimensional N-fold ($N\geqslant 3$ and $N\neq 4$) rotational symmetrical inclusions are considered in this paper. By making use of the Green formula and introducing the complex plane, we have successfully proved that the arithmetic mean of Eshelby tensors at N rotational symmetrical points, either inside or outside the inclusion, possesses some special properties. For the rotational symmetrical interior points, the arithmetic mean of Eshelby tensors is a constant and equal to the Eshelby tensor for the circular inclusion. This property also has been obtained in our former work [21] in a different way and is referred to as the quasi-Eshelby property, or the arithmetic mean theorem of Eshelby tensors for interior points. For the rotational symmetrical exterior points, the arithmetic mean of Eshelby tensors is proved to be zero. This property is the major conclusion of this paper and referred to as the arithmetic mean theorem of Eshelby tensors for exterior points.

By the arithmetic mean theorem for exterior points, we present two corollaries at the end of Sec. 4. The first corollary shows that the line integral average of Eshelby tensors along a circle $\Gamma_c(r,\mathbf{o})$, which is totally outside the inclusion and concentric with the inclusion, is equal to zero. The second corollary gives the area average of Eshelby tensors over the domain bounded by the circle $\Gamma_c(r,\mathbf{o})$. We find this area average of Eshelby tensor is in direct proportion to the area of the inclusion and in inverse proportion to the area of the circle. The arithmetic mean theorems and their corollaries are expected to be useful in the study of composites and strained semiconductor devices.

To make the results of the arithmetic mean theorem more acceptable to the readers, we have considered a circular inclusion, a special kind of rotational symmetrical inclusion, to test the theorem. We have checked the arithmetic mean of Eshelby tensors at any N ($N\geqslant 3$ and $N\neq 4$) rotational symmetrical points outside the circular inclusion by the available solutions and they turn out to be zero, just as the theorem suggests.

Acknowledgment

Financial support of the National Natural Science Foundation of China No. 10372003 is greatly acknowledged.

References

[1] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems," Proc. R. Soc. London, Ser. A, **241**, pp. 376–396.

[2] Mura, T., 1987, *Micromechanics of Defects in Solids*, 2nd ed., Nijhoff, Dordrecht.

[3] Willis, J. R., 1981, "Variational and Related Methods for the Overall Properties of Composites," Adv. Appl. Mech., **21**, pp. 1–78.

[4] Reid, A. C. E., and Gooding, R. J., 1992, "Inclusion Problem in a Two-Dimensional Nonlocal Elastic Solid," Phys. Rev. B, **46**(10), pp. 6045–6049.

[5] Sun, Y. Q., Gu, X. M., and Hazzledine, P. M., 2002, "Displacement Field Inside a Spherical Dislocation Cage and the Eshelby Tensor," Phys. Rev. B, **65**(22), pp. 220103–1–220103-4.

[6] Ru, C. Q., 1999, "Analytic Solution for Eshelby's Problem of an Inclusion of Arbitrary Shape in a Plane or Half-Plane," ASME J. Appl. Mech., **66**(2), pp. 315–322.

[7] Owen, D. R. J., 1972, "Analysis of Fibre-Reinforced Materials by an Initial Strain Method," Fibre Sci. Technol., **5**, pp. 37-59.

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

[9] Rodin, G., 1996, "Eshelby's Inclusion Problem for Polygons and Polyhedra," J. Mech. Phys. Solids, **44**, pp. 1977–1995.

[10] Mura, T., Shodja, H. M., Lin, T. Y., Safadi, A., and Makkawy, A., 1994, "The Determination of the Elastic Field of a Pentagonal Star Shaped Inclusion," Bull. Technical University of Istanbul, **47**, pp. 267–280.

[11] Lubarda, V. A., and Markenscoff, X., 1998, "On the Absence of Eshelby Property for Non-ellipsoidal Inclusions," Int. J. Solids Struct., **35**, pp. 3405–3411.

[12] Downes, J. R., and Faux, D. A., 1995, "Calculation of Strain Distributions in Multiple-Quantum-Well Strained-Layer Structures," J. Appl. Phys., **77**(6), pp. 2444–2447.

[13] Faux, D. A., Downes, J. R., and O'Reilly, E. P., 1996, "A Simple Method for Calculating Strain Distributions in Quantum-Wire Structures," J. Appl. Phys., **80**(4), pp. 2515–2517.

[14] Faux, D. A., Downes, J. R., and O'Reilly, E. P., 1997, "Analytic Solutions for Strain Distributions in Quantum-Wire structures," J. Appl. Phys., **82**(8), pp. 3754–3762.

[15] Faux, D. A., and Pearson, G. S., 2000, "Green's Tensors for Anisotropic Elasticity: Application to Quantum Dots," Phys. Rev. B, **62**(8), pp. R4798–R4801.

[16] Andreev, A. D., Downes, J. R., Faux, D. A., and O'Reilly, E. P., 1999, "Strain Distributions in Quantum Dots of Arbitrary Shape," J. Appl. Phys., **86**(1), pp. 297–305.

[17] Bergman, D. J., Strelniker, Y. M., and Sarychev, A. K., 1997, "Recent Advances in Strong Field Magneto-Transport in a Composite Medium," Physica A, **241**, pp. 278–283.

[18] Coffey, M. W., 2002, "Logarithmically Perturbed Two Dimensional Oscillator Model of a Quantum-Dot Nanostructure," Appl. Phys. Lett., **80**(7), pp. 1219–1221.

[19] Nozaki, H., and Taya, M., 1997, "Elastic Fields in a Polygon-Shaped Inclusion with Uniform Eigenstrains," ASME J. Appl. Mech., **64**, pp. 495–502.

[20] Kawashita, M., and Nozaki, H., 2001, "Eshelby Tensor of a Polygonal Inclusion and Its Special Properties," J. Elast., **64**, pp. 71–84.

[21] Wang, M. Z., and Xu, B. X., 2004, "The Arithmetic Mean Theorem of Eshelby Tensor for a Rotational Symmetrical Inclusion," J. Elast., **77**, pp. 13–23.

[22] Xu, B. X., and Wang, M. Z., 2005, "The Quasi Eshelby Property for Rotational Symmetrical Inclusions of Uniform Eigencurvatures within an Infinite Plate," Proc. R. Soc. London, Ser. A, **461**, pp. 2899–2910.

[23] Xu, B. X., and Wang, M. Z., 2005, "Special Properties of Eshelby Tensor for a Regular Polygonal Inclusion," Acta Mech. Sin. (in English), **21**, pp. 267–271.

[24] Ferrers, N. M., 1877, "On the Potentials of Ellipsoids, Ellipsoidal Shells, Elliptic Laminae and Elliptic Rings of Variable Densities," Q. J. Pure Appl. Math., **14**, pp. 1–22.

[25] Dyson, F. W., 1891, "The Potentials of Ellipsoids of Variable Densities," Q. J. Pure Appl. Math., **25**, pp. 259–288.

[26] Holland, A. S. B., 1980, *Complex Function Theory*, Elsevier, New York.

[27] Aboudi, J., 1991, *Mechanics of Composite Materials: A Unified Micromechanical Approach*, Elsevier, Amsterdam.

[28] Nemat-Nasser, S., and Hori, M., 1999, *Micromechanics: Overall Properties of Heterogeneous Elastic Solids*, 2nd ed., North-Holland, Amsterdam.

[29] Milton, G. W., 2002, *The Theory of Composites*, Cambridge University Press, Cambridge, UK.

[30] Torquato, S., 2002, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*, Springer, New York.

Null-Field Integral Equation Approach for Plate Problems With Circular Boundaries

Jeng-Tzong Chen
e-mail: jtchen@mail.ntou.edu.tw

Chia-Chun Hsiao

Department of Harbor and River Engineering,
National Taiwan Ocean University,
No. 2, Pei-Ning Road,
Keelung 20224, Taiwan

Shyue-Yuh Leu
Hydraulic Engineering Department,
Sinotech Engineering Consultants,
171 Nanking E. Road, Sec. 5,
Taipei 10570, Taiwan

In this paper, a semi-analytical approach for circular plate problems with multiple circular holes is presented. Null-field integral equation is employed to solve the plate problems while the kernel functions in the null-field integral equation are expanded to degenerate kernels based on the separation of field and source points in the fundamental solution. The unknown boundary densities of the circular plates are expressed in terms of Fourier series. It is noted that all the improper integrals are transformed to series sum and are easily calculated when the degenerate kernels and Fourier series are used. By matching the boundary conditions at the collocation points, a linear algebraic system is obtained. After determining the unknown Fourier coefficients, the displacement, slope, normal moment, and effective shear force of the plate can be obtained by using the boundary integral equations. Finally, two numerical examples are proposed to demonstrate the validity of the present method and the results are compared with the available exact solution, the finite element solution using ABAQUS software and the data of Bird and Steele. [DOI: 10.1115/1.2165239]

1 Introduction

The boundary element method (BEM) by discretizing the boundary integral equation (BIE) has been extensively applied to engineering problems recently more than domain type methods, e.g., finite element method (FEM) or finite difference method. It is noted that improper integrals on the boundary should be handled particularly when BEM is used. In the past, many researchers proposed several regularization techniques to deal with the singularity and hypersingularity. To determine the Cauchy principal value and the Hadamard principal value in the singular and hypersingular integrals is a critical issue in BEM/BIEM [1]. The technique of the integration by parts to reduce the order of singularity [2] is an alternative. One order of singularity is shifted to the density function from the kernel. In this paper, instead of using the previous concepts, the kernel function is described in an analytical form on each side (interior and exterior) by employing the separable kernel since the potential is discontinuous across the boundary. Therefore, degenerate kernel, namely separable kernel, is a vital tool to study the perforated plate which satisfies the biharmonic equation.

BIEs for the plate problems were acquired from the Rayleigh-Green identity [3,4] and the null-field integral equations were derived by collocating the field point outside the domain. Null-field integral equation in conjunction with degenerate kernel is proposed to solve the biharmonic problems with circular boundaries. It is well known that Fourier series is always incorporated to formulate the solution for problems with circular boundaries [5–8]. Bird and Steele [5] presented a Fourier series procedure to solve circular plate problems containing multiple circular holes in a similar way to the Trefftz method by adopting the interior and exterior T-complete sets. Either the interior or exterior bases in the Trefftz method are embedded in degenerate kernels [9]. A bridge to connect the Trefftz method and method of fundamental solution was constructed by using the degenerate kernels [9]. The Fourier series procedure can solve the circular plate problems regardless of the number, location, and size of circular holes. Also, Crouch and Mogilevskaya [6] presented a method for solving problems with circular boundaries. Their formulation is based on real-variable approach. Mogilevskaya and Crouch [8] have used the Galerkin method instead of collocation technique. Our approach can be extended to the Galerkin formulation only for the circular and annular cases. However, it may encounter difficulty for the eccentric example. Two requirements are needed: degenerate kernel expansion must be available and distinction of interior and exterior expression must be separated. Therefore, the collocation angle of f is not in the range 0 to $2p$ in our adaptive observer system. This is the reason why we cannot formulate in terms of Galerkin formulation using orthogonal properties twice. Free of worrying how to choose the collocation points, uniform collocation along the circular boundary yields a well-posed matrix. On the other hand, Bird and Steele [5] have also used separated solution procedure for bending of circular plates with circular holes in a similar way to the Trefftz method and addition theorem. They used the so-called method of series and addition theorem. Addition theorems are re-expansion formulas for the special functions (e.g., Bessel function, Legendre functions, surface spherical harmonics, etc.) in a transferred coordinate system (see, e.g., Gradshteyn and Ryzhik Table of Integrals [7]). Bird and Steele sought the solution of the original problem as the superposition of the solutions for a single hole problem. After taking the limit to the boundary of each hole, Bird and Steele got the equation that linked the Fourier series with the known coefficients used to approximate the boundary condition with the Fourier series with the unknown coefficients obtained from the solution process. The Fourier series from both sides of the equation were written in the same coordinate system, so the unknown coefficients were found by using orthogonal properties of the terms involved in the Fourier series [10]. To the authors' best knowledge, null-field integral equations and degenerate kernels were not employed to fully capture the circular boundaries although Fourier series expansion was used in previous research [5,6,8,11]. Jeffery [12] and Ling [13] adopted the bipolar coordinate system to derive the solution of stress for the plate problems since it is recognized as the best treatment for analyzing the biharmonic problem with two circular boundaries. Nevertheless, an analytical approach may be hindered

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received August 5, 2005; final manuscript received October 18, 2005. Review conducted by S. Mukherjee. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, JOURNAL OF APPLIED MECHANICS, Department of Mechanical and Environmental Engineering, University of California – Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

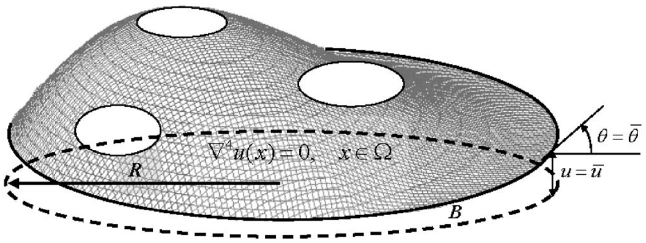


Fig. 1 The perforated Kirchhoff plate subject to the essential boundary conditions

for the complicated problems with more than two holes. Developing a systematic method to solve problems with several holes is not trivial.

The purpose of this paper is to study biharmonic problems with circular boundaries by using the null-field integral formulation in conjunction with degenerate kernels and Fourier series. According to the degenerate kernels, null-field integral formulation and Fourier series in the adaptive coordinate system, a linear algebraic system is constructed by matching the boundary conditions at the collocation points. After determining the Fourier coefficients, the displacement, slope, moment, and shear force of the plate with circular boundaries can be obtained by using the boundary integral equations for the domain point. In the polar coordinate system, the calculation of potential gradients in the normal and tangential directions for the non-concentric domain must be determined with care. Therefore, the technique of vector decomposition is adopted to deal with the problem for the non-concentric plate. Finally, several examples are presented to show the validity of the present method and some conclusions are made.

2 Problems Statement for a Plate

Consider a Kirchhoff plate for the two-dimensional domain under the distributed load $w(x)$, the governing equation is written as follows

$$\nabla^4 u^*(x)=\frac{w(x)}{D}, \quad x\in\Omega \tag{1}$$

where $u^*(x)$ is the lateral displacement, Ω is the domain of the thin plate, D is the flexural rigidity of the plate which is expressed as

$$D=\frac{Eh^3}{12(1-\nu^2)} \tag{2}$$

in which E is Young's modulus, ν denotes the Poisson ratio, and h is the plate thickness. For simplicity, the clamped case is considered

$$u^*(x)=0, \quad \theta^*(x)=0, \quad x\in B \tag{3}$$

where B is the boundary of the domain and $\theta^*(x)$ is the slope. Since the governing equation contains the body force, Eq. (1) can be reformulated to the homogeneous equation by using the splitting method as follows

$$\nabla^4 u(x)=0, \quad x\in\Omega \tag{4}$$

and the essential boundary conditions are changed to

$$u(x)=\bar{u}(x), \quad \theta(x)=\bar{\theta}(x), \quad x\in B \tag{5}$$

as shown in Fig. 1, where $u(x)$ is the displacement and $\theta(x)$ is the slope of the plate.

3 Formulation

3.1 Integral Equation for the Collocation Point in the Domain. The boundary integral equations for the domain point can be derived from the Rayleigh-Green identity [3,4] as follows

$$8\pi u(x)=-\int_B U(s,x)v(s)dB(s)+\int_B \Theta(s,x)m(s)dB(s) -\int_B M(s,x)\theta(s)dB(s)+\int_B V(s,x)u(s)dB(s), \quad x\in\Omega \tag{6}$$

$$8\pi \theta(x)=-\int_B U_\theta(s,x)v(s)dB(s)+\int_B \Theta_\theta(s,x)m(s)dB(s) -\int_B M_\theta(s,x)\theta(s)dB(s)+\int_B V_\theta(s,x)u(s)dB(s), \quad x\in\Omega \tag{7}$$

$$8\pi m(x)=-\int_B U_m(s,x)v(s)dB(s)+\int_B \Theta_m(s,x)m(s)dB(s) -\int_B M_m(s,x)\theta(s)dB(s)+\int_B V_m(s,x)u(s)dB(s), \quad x\in\Omega \tag{8}$$

$$8\pi v(x)=-\int_B U_v(s,x)v(s)dB(s)+\int_B \Theta_v(s,x)m(s)dB(s) -\int_B M_v(s,x)\theta(s)dB(s)+\int_B V_v(s,x)u(s)dB(s), \quad x\in\Omega \tag{9}$$

where B is the boundary of the domain Ω, $u(x)$, $\theta(x)$, $m(x)$, and $v(x)$ are the displacement, slope, moment, and shear force, s and x mean the source and field points, respectively. The kernel functions U, Θ, M, V, U_θ, Θ_θ, M_θ, V_θ, U_m, Θ_m, M_m, V_m, U_v, Θ_v, M_v, V_v in Eqs. (6)–(9), which are expanded to degenerate kernels by using the separation of source and field points, will be elaborated on later. The kernel function $U(s,x)$ in Eq. (6) is the fundamental solution which satisfies

$$\nabla^4 U(s,x)=8\pi\delta(s-x) \tag{10}$$

where $\delta(s-x)$ is the Dirac-delta function. Therefore, the fundamental solution can be obtained

$$U(s,x)=r^2\ln r \tag{11}$$

where r is the distance between the source point s and field point x. The relationship among $u(x)$, $\theta(x)$, $m(x)$, and $v(x)$ is shown as follows

$$\theta(x)=K_{\theta,x}(u(x))=\frac{\partial u(x)}{\partial n_x} \tag{12}$$

$$m(x)=K_{m,x}(u(x))=\nu\nabla_x^2 u(x)+(1-\nu)\frac{\partial^2 u(x)}{\partial^2 n_x} \tag{13}$$

$$v(x)=K_{v,x}(u(x))=\frac{\partial\nabla_x^2 u(x)}{\partial n_x}+(1-\nu)\frac{\partial}{\partial t_x}[n_i(x)t_j(x)u_{,ij}(x)] \tag{14}$$

where $K_{\theta,x}(\cdot)$, $K_{m,x}(\cdot)$, $K_{v,x}(\cdot)$ are the slope, moment, and shear force operators with respect to the point x, $\partial/\partial n_x$ is the normal derivative with respect to the field point x, $\partial/\partial t_x$ is the tangential derivative with respect to the field point x, ∇_x^2 means the Laplacian operator, and ν is the Poisson ratio.

3.2 Null-Field Integral Equations. The null-field integral equations are obtained by collocating the field point x outside the domain as follows

$$0 = -\int_B U(s,x)v(s)dB(s) + \int_B \Theta(s,x)m(s)dB(s) - \int_B M(s,x)\theta(s)dB(s) + \int_B V(s,x)u(s)dB(s), \quad x \in \Omega^C \tag{15}$$

$$0 = -\int_B U_\theta(s,x)v(s)dB(s) + \int_B \Theta_\theta(s,x)m(s)dB(s) - \int_B M_\theta(s,x)\theta(s)dB(s) + \int_B V_\theta(s,x)u(s)dB(s), \quad x \in \Omega^C \tag{16}$$

$$0 = -\int_B U_m(s,x)v(s)dB(s) + \int_B \Theta_m(s,x)m(s)dB(s) - \int_B M_m(s,x)\theta(s)dB(s) + \int_B V_m(s,x)u(s)dB(s), \quad x \in \Omega^C \tag{17}$$

$$0 = -\int_B U_v(s,x)v(s)dB(s) + \int_B \Theta_v(s,x)m(s)dB(s) - \int_B M_v(s,x)\theta(s)dB(s) + \int_B V_v(s,x)u(s)dB(s), \quad x \in \Omega^C \tag{18}$$

where Ω^C is the complementary domain of Ω. Since the four equations of Eqs. (15)–(18) in the plate formulation are provided, there are six (C_2^4) options for choosing any two equations to solve the problems. For simplicity, Eqs. (15) and (16) are used to analyze the plate problems. In the real implementation, the point in the null-field integral equation is moved to the boundary from Ω^C such that the kernel functions can be expressed in terms of appropriate forms of degenerate kernels. Novelly, all the improper integrals disappear and transform to series sum in the BIEs since the potential across the boundary can be determined in both sides by using degenerate kernels.

3.3 Expansion of Fourier Series for Boundary Densities. The displacement $u(s)$, slope $\theta(s)$, moment $m(s)$, and shear force $v(s)$ along the circular boundaries in the null-field integral equations are expanded in terms of Fourier series, which are expressed as follows

$$u(s) = c_0 + \sum_{n=1}^{M} (c_n \cos n\theta + d_n \sin n\theta), \quad s \in B \tag{19}$$

$$\theta(s) = g_0 + \sum_{n=1}^{M} (g_n \cos n\theta + h_n \sin n\theta), \quad s \in B \tag{20}$$

$$m(s) = a_0 + \sum_{n=1}^{M} (a_n \cos n\theta + b_n \sin n\theta), \quad s \in B \tag{21}$$

$$v(s) = p_0 + \sum_{n=1}^{M} (p_n \cos n\theta + q_n \sin n\theta), \quad s \in B \tag{22}$$

where a_0, a_n, b_n, c_0, c_n, d_n, g_0, g_n, h_n, p_0, p_n, and q_n are the Fourier coefficients and M is the number of Fourier series terms.

3.4 Expansion of Kernels. By employing the separation technique for the source and field points, the kernel function $U(s,x)$ can be expanded in terms of degenerate kernel in a series form [14] as shown in the following

$$U(s,x) = r^2 \ln r = \begin{cases} U^I(s,x) = \rho^2(1+\ln R) + R^2 \ln R - \left[R\rho(1+2\ln R) + \dfrac{1}{2}\dfrac{\rho^3}{R}\right]\cos(\theta-\phi) \\ \quad - \displaystyle\sum_{m=2}^{\infty}\left[\frac{1}{m(m+1)}\frac{\rho^{m+2}}{R^m} - \frac{1}{m(m-1)}\frac{\rho^m}{R^{m-2}}\right]\cos[m(\theta-\phi)], \quad R \geqslant \rho & (23a) \\ U^E(s,x) = R^2(1+\ln \rho) + \rho^2 \ln \rho - \left[\rho R(1+2\ln \rho) + \dfrac{1}{2}\dfrac{R^3}{\rho}\right]\cos(\theta-\phi) \\ \quad - \displaystyle\sum_{m=2}^{\infty}\left[\frac{1}{m(m+1)}\frac{R^{m+2}}{\rho^m} - \frac{1}{m(m-1)}\frac{R^m}{\rho^{m-2}}\right]\cos[m(\theta-\phi)], \quad \rho > R & (23b) \end{cases}$$

where the superscripts "I" and "E" denote the interior and exterior cases of $U(s,x)$ kernel depending on the geometry as shown in Fig. 2. The other kernels in the boundary integral equations can be obtained by utilizing the operators of Eqs. (12)–(14) with respect to the $U(s,x)$ kernel. The degenerate kernels U, Θ, M, V, U_θ, Θ_θ, M_θ, and V_θ in Eqs. (15) and (16) are listed in Appendix A. It is noted that the interior and exterior cases of U, Θ, M, U_θ, and Θ_θ are the same when they both approach to the boundary ($\rho=R$), since the degenerate kernels are continuous functions across the boundary. Then, the kernel function with the superscript "I"is chosen while the field point is inside the circular region; otherwise, the kernels with the superscript "E" are chosen.

4 Adaptive Observer System and Vector Decomposition for the Slope

4.1 Adaptive Observer System. Consider a plate problem with circular boundaries as shown in Fig. 3. Since the boundary integral equations are frame indifferent, i.e., rule of objectivity is obeyed. Adaptive observer system is chosen to fully employ the circular property by expanding the kernels into degenerate forms. The origin of the observer system can be adaptively located on the center of the corresponding boundary contour under integration. The dummy variable in the circular contour integration is the angle (θ) instead of radial coordinate (R). By using the adaptive

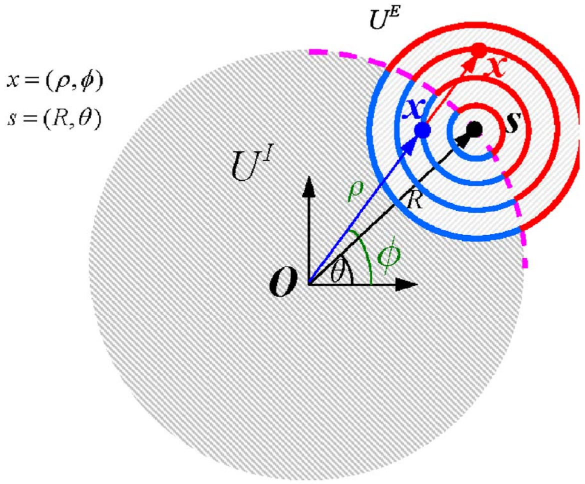


Fig. 2 Degenerate kernel for $U(s,x)$

system, all the boundary integrals can be determined analytically free of principal value senses.

4.2 Vector Decomposition. Since the higher-order singular equation is also one alternative to deal with the plate problem, potential gradient or higher-order gradients is required to calculate carefully. For the non-concentric case, special treatment for the potential gradient should be given as the source and field points locate on different circular boundaries. As shown in Fig. 4, the true normal direction with respect to the collocation point x on the B_i boundary can be superimposed by using the radial direction $\underset{\sim}{e}_{\rho}$ and angular direction $\underset{\sim}{e}_{\phi}$ on the B_j boundary. The degenerate kernels in Eq. (16) for the higher-order singular equation are changed to

$$U_n(s,x) = \frac{\partial U(s,x)}{\partial n_x}\cos(\phi-\phi') + \frac{\partial U(s,x)}{\partial t_x}\cos\left(\frac{\pi}{2}-\phi+\phi'\right) \tag{24}$$

$$\Theta_n(s,x) = \frac{\partial \Theta(s,x)}{\partial n_x}\cos(\phi-\phi') + \frac{\partial \Theta(s,x)}{\partial t_x}\cos\left(\frac{\pi}{2}-\phi+\phi'\right) \tag{25}$$

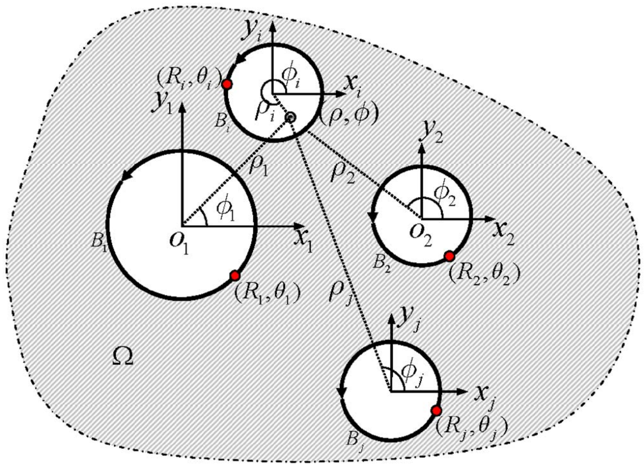


Fig. 3 Adaptive observer system when integrating the corresponding circular boundaries

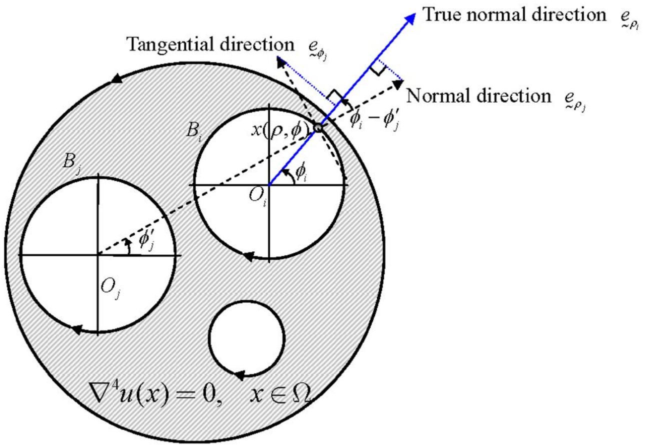

True normal direction $\underset{\sim}{e}_{\rho_i}$
Tangential direction $\underset{\sim}{e}_{\phi_j}$
Normal direction $\underset{\sim}{e}_{\rho_j}$
B_i
$x(\rho,\phi)$
$\phi_i-\phi'_j$
B_j
O_i
ϕ_i
ϕ'_j
O_j
$\nabla^4 u(x)=0,\quad x\in\Omega$

Fig. 4 Vector decomposition (collocation on x and integration on B_j)

$$M_n(s,x) = \frac{\partial M(s,x)}{\partial n_x}\cos(\phi-\phi') + \frac{\partial M(s,x)}{\partial t_x}\cos\left(\frac{\pi}{2}-\phi+\phi'\right) \tag{26}$$

$$V_n(s,x) = \frac{\partial V(s,x)}{\partial n_x}\cos(\phi-\phi') + \frac{\partial V(s,x)}{\partial t_x}\cos\left(\frac{\pi}{2}-\phi+\phi'\right) \tag{27}$$

The tangential derivative $\partial/\partial t_x$ with respect to the field point x for the four kernels needs to be additionally derived and is listed in Appendix A, where the normal derivative $\partial/\partial n_x$ is $\partial/\partial\rho$, and has been derived in the U_θ, Θ_θ, M_θ, and V_θ kernels. We call this treatment "vector decomposition technique." By approaching the collocation point from Ω^C to B_i and integrating circle B_j using the adaptive observer system of origin O_j, the normal and tangent derivatives can be superimposed as follows

$$\frac{\partial}{\partial\rho_i} = \frac{\partial}{\partial\rho_j}\cos(\phi_i-\phi'_j) + \frac{1}{\rho_j}\frac{\partial}{\partial\phi_j}\cos\left(\frac{\pi}{2}-\phi_i+\phi'_j\right) \tag{28}$$

$$\frac{1}{\rho_i}\frac{\partial}{\partial\phi_i} = \frac{\partial}{\partial\rho_j}\cos\left(\frac{\pi}{2}-\phi_i+\phi'_j\right) + \frac{1}{\rho_j}\frac{\partial}{\partial\phi_j}\cos(\phi_i-\phi'_j) \tag{29}$$

5 Linear Algebraic System

Consider the plate problem with circular domain containing N_h randomly distributed circular holes centered at the position vector $\underset{\sim}{c}_j$ $(j=1,2,\ldots,N)$, $(N=N_h+1$ and $\underset{\sim}{c}_1$ is the position vector of the outer circular boundary for the plate), as shown in Fig. 5 in which R_j denotes the radius of the jth circular region and B_j is the boundary of the jth circular hole. By uniformly collocating the $2M+1$ points x on each circular boundary in Eqs. (15) and (16), we have

$$0 = \sum_{j=1}^{N}\int_{B_j}\{-U(s,x)v(s)+\Theta(s,x)m(s)-M(s,x)\theta(s) + V(s,x)u(s)\}dB_j(s), \quad x\in\Omega^C \tag{30}$$

$$0 = \sum_{j=1}^{N}\int_{B_j}\{-U_\theta(s,x)v(s)+\Theta_\theta(s,x)m(s)-M_\theta(s,x)\theta(s) + V_\theta(s,x)u(s)\}dB_j(s), \quad x\in\Omega^C \tag{31}$$

It is noted that we select the null-field point on the boundary in the real computation. The selection of interior or exterior degenerate kernels depends on $r<R$ or $r>R$, respectively, according to the

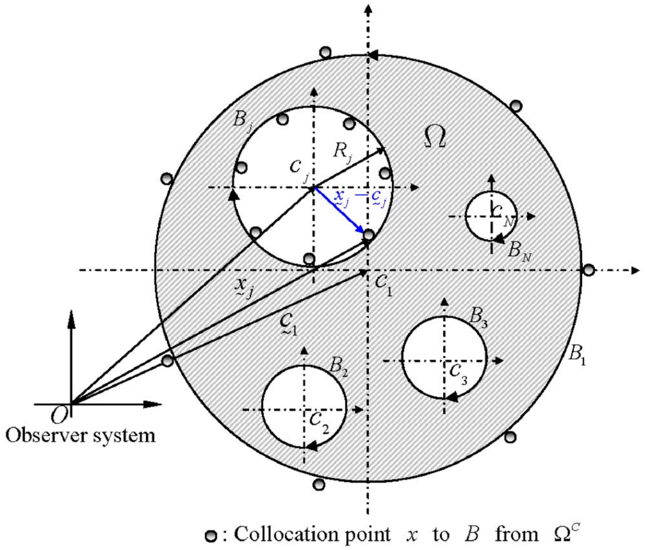


Fig. 5 Collocation point and boundary contour integration in the null-field integral equation

observer system. Besides, the path is counterclockwise for the outer circle; otherwise, it is clockwise. For the integral of the circular boundary, the degenerate kernels of $U(s,x)$, $\Theta(s,x)$, $M(s,x)$, $V(s,x)$, $U_\theta(s,x)$, $\Theta_\theta(s,x)$, $M_\theta(s,x)$, and $V_\theta(s,x)$ are utilized while the boundary densities of $u(s)$, $\theta(s)$, $m(s)$, and $v(s)$ along the circular boundary are substituted by using the Fourier series of Eqs. (19)–(22), respectively. In the B_j integration, the origin of the observer system is adaptively set to collocate at the center c_j to novelly utilize the degenerate kernels and Fourier series. A linear algebraic system

$$\begin{bmatrix} \mathbf{U}11 & \mathbf{\Theta}11 & \mathbf{U}12 & \mathbf{\Theta}12 & \cdots & \mathbf{U}1N & \mathbf{\Theta}1N \\ \mathbf{U}11_\theta & \mathbf{\Theta}11_\theta & \mathbf{U}12_\theta & \mathbf{\Theta}12_\theta & \cdots & \mathbf{U}1N_\theta & \mathbf{\Theta}1N_\theta \\ \mathbf{U}21 & \mathbf{\Theta}21 & \mathbf{U}22 & \mathbf{\Theta}22 & \cdots & \mathbf{U}2N & \mathbf{\Theta}2N \\ \mathbf{U}21_\theta & \mathbf{\Theta}21_\theta & \mathbf{U}22_\theta & \mathbf{\Theta}22_\theta & \cdots & \mathbf{U}2N_\theta & \mathbf{\Theta}2N_\theta \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{U}N1 & \mathbf{\Theta}N1 & \mathbf{U}N2 & \mathbf{\Theta}N2 & \cdots & \mathbf{U}NN & \mathbf{\Theta}NN \\ \mathbf{U}N1_\theta & \mathbf{\Theta}N1_\theta & \mathbf{U}N2_\theta & \mathbf{\Theta}N2_\theta & \cdots & \mathbf{U}NN_\theta & \mathbf{\Theta}NN_\theta \end{bmatrix} \begin{Bmatrix} \mathbf{v}_1 \\ \mathbf{m}_1 \\ \mathbf{v}_2 \\ \mathbf{m}_2 \\ \vdots \\ \mathbf{v}_N \\ \mathbf{m}_N \end{Bmatrix} = \begin{bmatrix} \mathbf{M}11 & \mathbf{V}11 & \mathbf{M}12 & \mathbf{V}12 & \cdots & \mathbf{M}1N & \mathbf{V}1N \\ \mathbf{M}11_\theta & \mathbf{V}11_\theta & \mathbf{M}12_\theta & \mathbf{V}12_\theta & \cdots & \mathbf{M}1N_\theta & \mathbf{V}1N_\theta \\ \mathbf{M}21 & \mathbf{V}21 & \mathbf{M}22 & \mathbf{V}22 & \cdots & \mathbf{M}2N & \mathbf{V}2N \\ \mathbf{M}21_\theta & \mathbf{V}21_\theta & \mathbf{M}22_\theta & \mathbf{V}22_\theta & \cdots & \mathbf{M}2N_\theta & \mathbf{V}2N_\theta \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{M}N1 & \mathbf{V}N1 & \mathbf{M}N2 & \mathbf{V}N2 & \cdots & \mathbf{M}NN & \mathbf{V}NN \\ \mathbf{M}N1_\theta & \mathbf{V}N1_\theta & \mathbf{M}N2_\theta & \mathbf{V}N2_\theta & \cdots & \mathbf{M}NN_\theta & \mathbf{V}NN_\theta \end{bmatrix} \times \begin{Bmatrix} \boldsymbol{\theta}_1 \\ \mathbf{u}_1 \\ \boldsymbol{\theta}_2 \\ \mathbf{u}_2 \\ \vdots \\ \boldsymbol{\theta}_N \\ \mathbf{u}_N \end{Bmatrix} \tag{32}$$

is obtained, where N denotes the number of circular boundaries (including inner and outer circular boundaries). For brevity, a unified form $[\mathbf{U}ij]$ ($i=1,2,3,\ldots,N$ and $j=1,2,3,\ldots,N$) denote the response of $U(s,x)$ kernel at the ith circle point due to the source at the jth circle. Otherwise, the same definition for $[\mathbf{\Theta}ij]$, $[\mathbf{M}ij]$, $[\mathbf{V}ij]$, $[\mathbf{U}ij_\theta]$, $[\mathbf{\Theta}ij_\theta]$, $[\mathbf{M}ij_\theta]$, and $[\mathbf{V}ij_\theta]$ cases. The submatrices of $[\mathbf{U}ij]$, $[\mathbf{\Theta}ij]$, $[\mathbf{M}ij]$, $[\mathbf{V}ij]$, $[\mathbf{U}ij_\theta]$, $[\mathbf{\Theta}ij_\theta]$, $[\mathbf{M}ij_\theta]$, and $[\mathbf{V}ij_\theta]$ are defined as follows

$$[\mathbf{U}ij] = \begin{bmatrix} Uij^{0c}(\phi_1) & Uij^{1c}(\phi_1) & Uij^{1s}(\phi_1) & \cdots & Uij^{Mc}(\phi_1) & Uij^{Ms}(\phi_1) \\ Uij^{0c}(\phi_2) & Uij^{1c}(\phi_2) & Uij^{1s}(\phi_2) & \cdots & Uij^{Mc}(\phi_2) & Uij^{Ms}(\phi_2) \\ Uij^{0c}(\phi_3) & Uij^{1c}(\phi_3) & Uij^{1s}(\phi_3) & \cdots & Uij^{Mc}(\phi_3) & Uij^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Uij^{0c}(\phi_{2M}) & Uij^{1c}(\phi_{2M}) & Uij^{1s}(\phi_{2M}) & \cdots & Uij^{Mc}(\phi_{2M}) & Uij^{Ms}(\phi_{2M}) \\ Uij^{0c}(\phi_{2M+1}) & Uij^{1c}(\phi_{2M+1}) & Uij^{1s}(\phi_{2M+1}) & \cdots & Uij^{Mc}(\phi_{2M+1}) & Uij^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{33}$$

$$[\mathbf{\Theta}ij] = \begin{bmatrix} \Theta ij^{0c}(\phi_1) & \Theta ij^{1c}(\phi_1) & \Theta ij^{1s}(\phi_1) & \cdots & \Theta ij^{Mc}(\phi_1) & \Theta ij^{Ms}(\phi_1) \\ \Theta ij^{0c}(\phi_2) & \Theta ij^{1c}(\phi_2) & \Theta ij^{1s}(\phi_2) & \cdots & \Theta ij^{Mc}(\phi_2) & \Theta ij^{Ms}(\phi_2) \\ \Theta ij^{0c}(\phi_3) & \Theta ij^{1c}(\phi_3) & \Theta ij^{1s}(\phi_3) & \cdots & \Theta ij^{Mc}(\phi_3) & \Theta ij^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \Theta ij^{0c}(\phi_{2M}) & \Theta ij^{1c}(\phi_{2M}) & \Theta ij^{1s}(\phi_{2M}) & \cdots & \Theta ij^{Mc}(\phi_{2M}) & \Theta ij^{Ms}(\phi_{2M}) \\ \Theta ij^{0c}(\phi_{2M+1}) & \Theta ij^{1c}(\phi_{2M+1}) & \Theta ij^{1s}(\phi_{2M+1}) & \cdots & \Theta ij^{Mc}(\phi_{2M+1}) & \Theta ij^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{34}$$

$$[\mathbf{M}ij] = \begin{bmatrix} Mij^{0c}(\phi_1) & Mij^{1c}(\phi_1) & Mij^{1s}(\phi_1) & \cdots & Mij^{Mc}(\phi_1) & Mij^{Ms}(\phi_1) \\ Mij^{0c}(\phi_2) & Mij^{1c}(\phi_2) & Mij^{1s}(\phi_2) & \cdots & Mij^{Mc}(\phi_2) & Mij^{Ms}(\phi_2) \\ Mij^{0c}(\phi_3) & Mij^{1c}(\phi_3) & Mij^{1s}(\phi_3) & \cdots & Mij^{Mc}(\phi_3) & Mij^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Mij^{0c}(\phi_{2M}) & Mij^{1c}(\phi_{2M}) & Mij^{1s}(\phi_{2M}) & \cdots & Mij^{Mc}(\phi_{2M}) & Mij^{Ms}(\phi_{2M}) \\ Mij^{0c}(\phi_{2M+1}) & Mij^{1c}(\phi_{2M+1}) & Mij^{1s}(\phi_{2M+1}) & \cdots & Mij^{Mc}(\phi_{2M+1}) & Mij^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{35}$$

$$[\mathbf{V}ij]=\begin{bmatrix} Vij^{0c}(\phi_1) & Vij^{1c}(\phi_1) & Vij^{1s}(\phi_1) & \cdots & Vij^{Mc}(\phi_1) & Vij^{Ms}(\phi_1) \\ Vij^{0c}(\phi_2) & Vij^{1c}(\phi_2) & Vij^{1s}(\phi_2) & \cdots & Vij^{Mc}(\phi_2) & Vij^{Ms}(\phi_2) \\ Vij^{0c}(\phi_3) & Vij^{1c}(\phi_3) & Vij^{1s}(\phi_3) & \cdots & Vij^{Mc}(\phi_3) & Vij^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Vij^{0c}(\phi_{2M}) & Vij^{1c}(\phi_{2M}) & Vij^{1s}(\phi_{2M}) & \cdots & Vij^{Mc}(\phi_{2M}) & Vij^{Ms}(\phi_{2M}) \\ Vij^{0c}(\phi_{2M+1}) & Vij^{1c}(\phi_{2M+1}) & Vij^{1s}(\phi_{2M+1}) & \cdots & Vij^{Mc}(\phi_{2M+1}) & Vij^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{36}$$

$$[\mathbf{U}ij_\theta]=\begin{bmatrix} Uij_\theta^{0c}(\phi_1) & Uij_\theta^{1c}(\phi_1) & Uij_\theta^{1s}(\phi_1) & \cdots & Uij_\theta^{Mc}(\phi_1) & Uij_\theta^{Ms}(\phi_1) \\ Uij_\theta^{0c}(\phi_2) & Uij_\theta^{1c}(\phi_2) & Uij_\theta^{1s}(\phi_2) & \cdots & Uij_\theta^{Mc}(\phi_2) & Uij_\theta^{Ms}(\phi_2) \\ Uij_\theta^{0c}(\phi_3) & Uij_\theta^{1c}(\phi_3) & Uij_\theta^{1s}(\phi_3) & \cdots & Uij_\theta^{Mc}(\phi_3) & Uij_\theta^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Uij_\theta^{0c}(\phi_{2M}) & Uij_\theta^{1c}(\phi_{2M}) & Uij_\theta^{1s}(\phi_{2M}) & \cdots & Uij_\theta^{Mc}(\phi_{2M}) & Uij_\theta^{Ms}(\phi_{2M}) \\ Uij_\theta^{0c}(\phi_{2M+1}) & Uij_\theta^{1c}(\phi_{2M+1}) & Uij_\theta^{1s}(\phi_{2M+1}) & \cdots & Uij_\theta^{Mc}(\phi_{2M+1}) & Uij_\theta^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{37}$$

$$[\mathbf{\Theta}ij_\theta]=\begin{bmatrix} \Theta ij_\theta^{0c}(\phi_1) & \Theta ij_\theta^{1c}(\phi_1) & \Theta ij_\theta^{1s}(\phi_1) & \cdots & \Theta ij_\theta^{Mc}(\phi_1) & \Theta ij_\theta^{Ms}(\phi_1) \\ \Theta ij_\theta^{0c}(\phi_2) & \Theta ij_\theta^{1c}(\phi_2) & \Theta ij_\theta^{1s}(\phi_2) & \cdots & \Theta ij_\theta^{Mc}(\phi_2) & \Theta ij_\theta^{Ms}(\phi_2) \\ \Theta ij_\theta^{0c}(\phi_3) & \Theta ij_\theta^{1c}(\phi_3) & \Theta ij_\theta^{1s}(\phi_3) & \cdots & \Theta ij_\theta^{Mc}(\phi_3) & \Theta ij_\theta^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \Theta ij_\theta^{0c}(\phi_{2M}) & \Theta ij_\theta^{1c}(\phi_{2M}) & \Theta ij_\theta^{1s}(\phi_{2M}) & \cdots & \Theta ij_\theta^{Mc}(\phi_{2M}) & \Theta ij_\theta^{Ms}(\phi_{2M}) \\ \Theta ij_\theta^{0c}(\phi_{2M+1}) & \Theta ij_\theta^{1c}(\phi_{2M+1}) & \Theta ij_\theta^{1s}(\phi_{2M+1}) & \cdots & \Theta ij_\theta^{Mc}(\phi_{2M+1}) & \Theta ij_\theta^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{38}$$

$$[\mathbf{M}ij_\theta]=\begin{bmatrix} Mij_\theta^{0c}(\phi_1) & Mij_\theta^{1c}(\phi_1) & Mij_\theta^{1s}(\phi_1) & \cdots & Mij_\theta^{Mc}(\phi_1) & Mij_\theta^{Ms}(\phi_1) \\ Mij_\theta^{0c}(\phi_2) & Mij_\theta^{1c}(\phi_2) & Mij_\theta^{1s}(\phi_2) & \cdots & Mij_\theta^{Mc}(\phi_2) & Mij_\theta^{Ms}(\phi_2) \\ Mij_\theta^{0c}(\phi_3) & Mij_\theta^{1c}(\phi_3) & Mij_\theta^{1s}(\phi_3) & \cdots & Mij_\theta^{Mc}(\phi_3) & Mij_\theta^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Mij_\theta^{0c}(\phi_{2M}) & Mij_\theta^{1c}(\phi_{2M}) & Mij_\theta^{1s}(\phi_{2M}) & \cdots & Mij_\theta^{Mc}(\phi_{2M}) & Mij_\theta^{Ms}(\phi_{2M}) \\ Mij_\theta^{0c}(\phi_{2M+1}) & Mij_\theta^{1c}(\phi_{2M+1}) & Mij_\theta^{1s}(\phi_{2M+1}) & \cdots & Mij_\theta^{Mc}(\phi_{2M+1}) & Mij_\theta^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{39}$$

$$[\mathbf{V}ij_\theta]=\begin{bmatrix} Vij_\theta^{0c}(\phi_1) & Vij_\theta^{1c}(\phi_1) & Vij_\theta^{1s}(\phi_1) & \cdots & Vij_\theta^{Mc}(\phi_1) & Vij_\theta^{Ms}(\phi_1) \\ Vij_\theta^{0c}(\phi_2) & Vij_\theta^{1c}(\phi_2) & Vij_\theta^{1s}(\phi_2) & \cdots & Vij_\theta^{Mc}(\phi_2) & Vij_\theta^{Ms}(\phi_2) \\ Vij_\theta^{0c}(\phi_3) & Vij_\theta^{1c}(\phi_3) & Vij_\theta^{1s}(\phi_3) & \cdots & Vij_\theta^{Mc}(\phi_3) & Vij_\theta^{Ms}(\phi_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Vij_\theta^{0c}(\phi_{2M}) & Vij_\theta^{1c}(\phi_{2M}) & Vij_\theta^{1s}(\phi_{2M}) & \cdots & Vij_\theta^{Mc}(\phi_{2M}) & Vij_\theta^{Ms}(\phi_{2M}) \\ Vij_\theta^{0c}(\phi_{2M+1}) & Vij_\theta^{1c}(\phi_{2M+1}) & Vij_\theta^{1s}(\phi_{2M+1}) & \cdots & Vij_\theta^{Mc}(\phi_{2M+1}) & Vij_\theta^{Ms}(\phi_{2M+1}) \end{bmatrix} \tag{40}$$

where ϕ_k $(k=1,2,3,\ldots,2M+1)$ is the kth collocation angle of the collocation points on each boundary and the element of the submatrices are defined as follows

$$Uij^{nc}(\phi_k)=\int_{B_j} U(s,x_k)\cos(n\theta_j)dB_j(s), \quad n=0,1,2,3,\ldots,M \tag{41}$$

$$Uij^{ns}(\phi_k)=\int_{B_j} U(s,x_k)\sin(n\theta_j)dB_j(s), \quad n=1,2,3,\ldots,M \tag{42}$$

$$\Theta ij^{nc}(\phi_k)=\int_{B_j} \Theta(s,x_k)\cos(n\theta_j)dB_j(s), \quad n=0,1,2,3,\ldots,M \tag{43}$$

$$\Theta ij^{ns}(\phi_k)=\int_{B_j} \Theta(s,x_k)\sin(n\theta_j)dB_j(s), \quad n=1,2,3,\ldots,M \tag{44}$$

$$Mij^{nc}(\phi_k)=\int_{B_j} M(s,x_k)\cos(n\theta_j)dB_j(s), \quad n=0,1,2,3,\ldots,M \tag{45}$$

$$Mij^{ns}(\phi_k)=\int_{B_j} M(s,x_k)\sin(n\theta_j)dB_j(s), \quad n=1,2,3,\ldots,M \tag{46}$$

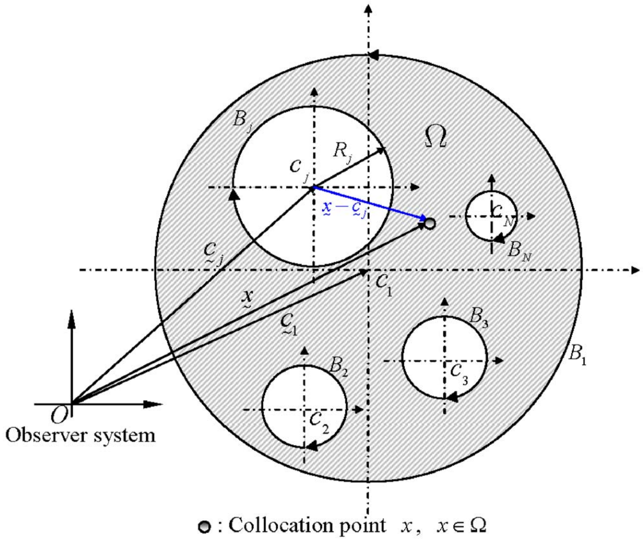

B_j
Ω
c_j
R_j
$x - c_j$
c_N
B_N
c_j
c_1
x
c_1
B_3
B_2
B_1
c_3
c_2
O
Observer system
: Collocation point x, $x \in \Omega$

Fig. 6 Boundary integral equation for the domain point

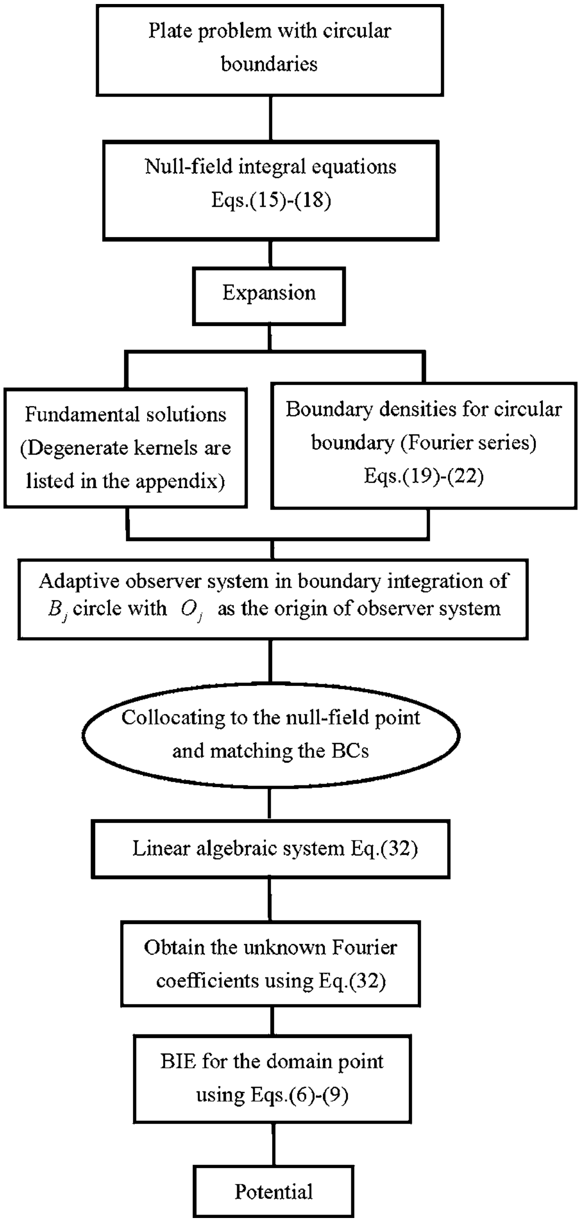
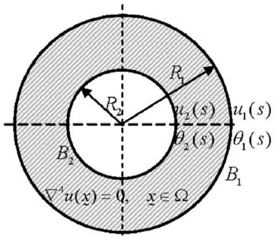


Plate problem with circular boundaries
Null-field integral equations Eqs.(15)-(18)
Expansion
Fundamental solutions (Degenerate kernels are listed in the appendix)
Boundary densities for circular boundary (Fourier series) Eqs.(19)-(22)
Adaptive observer system in boundary integration of B_j circle with O_j as the origin of observer system
Collocating to the null-field point and matching the BCs
Linear algebraic system Eq.(32)
Obtain the unknown Fourier coefficients using Eq.(32)
BIE for the domain point using Eqs.(6)-(9)
Potential

Fig. 7 Flowchart of the present method


R_1
R_2
$u_2(s)$
$\theta_2(s)$
$u_1(s)$
$\theta_1(s)$
B_2
B_1
$\nabla^4 u(x) = 0, \quad x \in \Omega$

Geometric data:
$R_1 = 2, \; R_2 = 1$
Essential boundary conditions:
$u_1(s) = 0$ and $\theta_1(s) = 2\sin\theta$ on B_1
$u_2(s) = -3\sin\theta$ and $\theta_2(s) = 5\sin\theta$ on B_2

Fig. 8 An annular plate subject to the essential boundary conditions

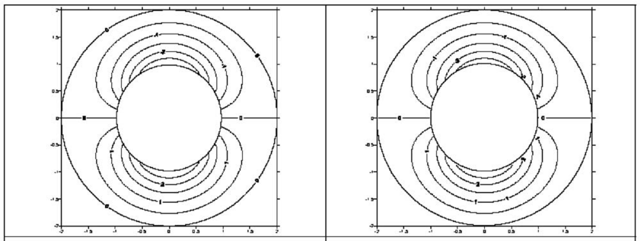
$$Vij^{nc}(\phi_k) = \int_{B_j} V(s,x_k)\cos(n\theta_j)dB_j(s), \quad n = 0,1,2,3, \ldots ,M \tag{47}$$

$$Vij^{ns}(\phi_k) = \int_{B_j} V(s,x_k)\sin(n\theta_j)dB_j(s), \quad n = 1,2,3, \ldots ,M \tag{48}$$

$$Uij^{nc}_{\theta}(\phi_k) = \int_{B_j} U_{\theta}(s,x_k)\cos(n\theta_j)dB_j(s), \quad n = 0,1,2,3, \ldots ,M \tag{49}$$

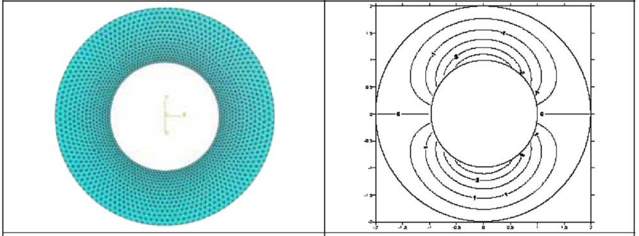
$$Uij^{ns}_{\theta}(\phi_k) = \int_{B_j} U_{\theta}(s,x_k)\sin(n\theta_j)dB_j(s), \quad n = 1,2,3, \ldots ,M \tag{50}$$

$$\Theta ij^{nc}_{\theta}(\phi_k) = \int_{B_j} \Theta_{\theta}(s,x_k)\cos(n\theta_j)dB_j(s), \quad n = 0,1,2,3, \ldots ,M \tag{51}$$



(*a*) Present method (M=10)

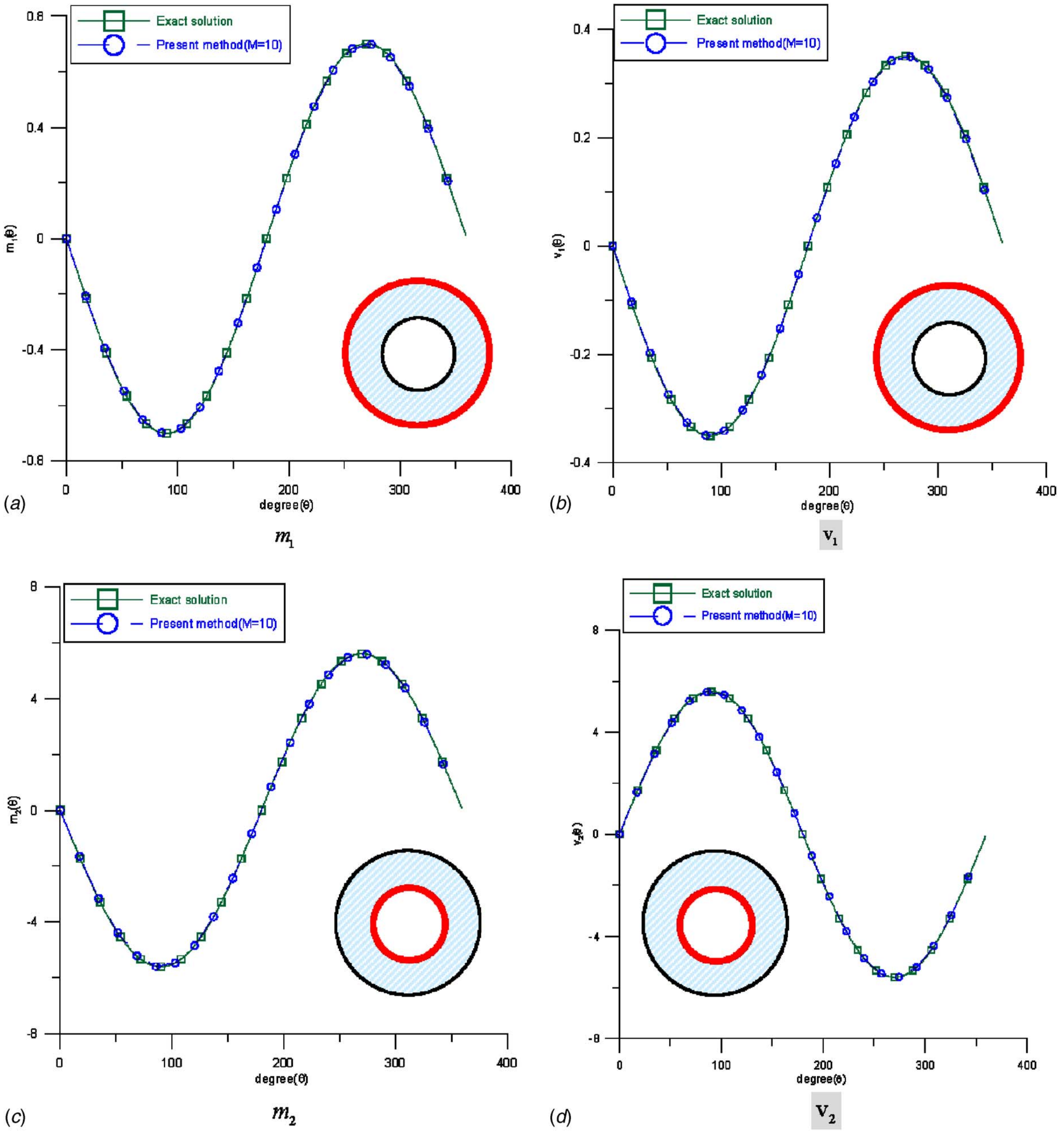
(*b*) Exact solution



(*c*) Finite element mesh for ABAQUS program (No. of nodes=1,920, No. of elements=3,600)

(*d*) FEM (ABAQUS)

Fig. 9 The contour plot of displacement for the annular plate subject to the essential boundary conditions by using a different method



(*a*) m_1 (*b*) $\mathbf{v_1}$

(*c*) m_2 (*d*) $\mathbf{v_2}$

Fig. 10 Error estimation of the moment and shear force on the boundaries for the concentric circular domain

$$\Theta ij_{\theta}^{ns}(\phi_k)=\int_{B_j}\Theta_{\theta}(s,x_k)\sin(n\theta_j)dB_j(s),\quad n=1,2,3,\ldots,M \tag{52}$$

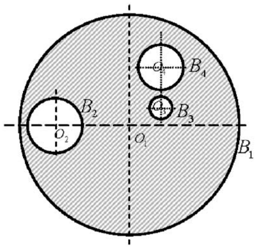
$$Mij_{\theta}^{nc}(\phi_k)=\int_{B_j}M_{\theta}(s,x_k)\cos(n\theta_j)dB_j(s),\quad n=0,1,2,3,\ldots,M \tag{53}$$

$$Mij_{\theta}^{ns}(\phi_k)=\int_{B_j}M_{\theta}(s,x_k)\sin(n\theta_j)dB_j(s),\quad n=1,2,3,\ldots,M \tag{54}$$

$$Vij_{\theta}^{nc}(\phi_k)=\int_{B_j}V_{\theta}(s,x_k)\cos(n\theta_j)dB_j(s),\quad n=0,1,2,3,\ldots,M \tag{55}$$

$$Vij_{\theta}^{ns}(\phi_k)=\int_{B_j}V_{\theta}(s,x_k)\sin(n\theta_j)dB_j(s),\quad n=1,2,3,\ldots,M \tag{56}$$

where the interior degenerate kernels are used for $j=1$, $i=2,3,\ldots,N$ and $i=j=2,3,4,\ldots,N$; otherwise, exterior degenerate kernels are used. The explicit forms of the boundary integral for U kernel are listed in Appendix B. Finite value of singularity is obtained after introducing the degenerate kernel. Besides, the limiting case across the boundary ($R^{-}=r=R^{+}$) is also addressed. In-



Geometric data:
$O_1=(0,0), R_1=20$; $O_2=(-14,0)$, $R_2=5$;
$O_3=(5,3)$, $R_3=2$; $O_4=(5,10)$, $R_4=4$
Essential boundary conditions:
$u(s)=0$ and $\theta(s)=0$ on B_1
$u(s)=\sin\theta$ and $\theta(s)=0$ on B_2
$u(s)=-1$ and $\theta(s)=0$ on B_3
$u(s)=1$ and $\theta(s)=0$ on B_4

Fig. 11 A circular plate containing three circular holes subject to the essential boundary conditions

stead of boundary data in BEM, the Fourier coefficients become the new unknown degree of freedom in the formulation. By rearranging the known and unknown sets, the Fourier coefficients can be obtained. Since the boundary data are determined, the displacement, slope, normal moment, and effective shear force of the plate can be solved by using the boundary integral equations for the domain point as shown in Fig. 6. The procedure of solution is described in a flowchart as shown in Fig. 7.

6 Numerical Results and Discussions

Case 1: An annular plate. An annular circular plate subject to the essential boundary conditions is considered as shown in Fig. 8. The unknown boundary densities of the plate are expressed in terms of Fourier series and the numerical result using fewer bases of Fourier series terms ($M=10$) is shown in Fig. 9(a). The annular case was also solved by using the FEM software (ABAQUS) [15] with 3,600 triangle elements as shown in Fig. 9(c). Good agreement is made after comparison with the exact solution,

$$u(\rho,\phi)=\rho\sin\phi-\frac{4}{\rho}\sin\phi, \quad 1<\rho<2, \quad 0\leqslant\phi<2\pi \tag{57}$$

as shown in Fig. 9(b) and the FEM result is shown in Fig. 9(d).

Since the exact solution is known, error estimation can be established. The boundary densities of the annular circular plate, normal moment, and effective shear force can be obtained by using the operators of Eqs. (13) and (14) with respect to the field point x. By substituting $\rho_1=2$ and $\rho_2=1$ into the two equations, the moment and the shear force on the boundaries of the annular plate are

$$m_1(x)=(\nu-1)\sin\phi, \quad x\in B_1, \quad 0\leqslant\phi<2\pi \tag{58}$$

$$v_1(x)=\frac{\nu-1}{2}\sin\phi \quad ,x\in B_1, \quad 0\leqslant\phi<2\pi \tag{59}$$

$$m_2(x)=8(\nu-1)\sin\phi, \quad x\in B_2, \quad 0\leqslant\phi<2\pi \tag{60}$$

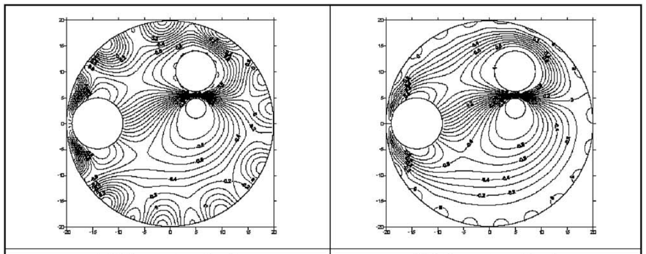
$$v_2(x)=8(1-\nu)\sin\phi, \quad x\in B_2, \quad 0\leqslant\phi<2\pi \tag{61}$$

The numerical results of the moment and the shear force expanded in fewer bases of Fourier series ($M=10$) agree well with Eqs. (58)–(61) and are shown in Fig. 10.

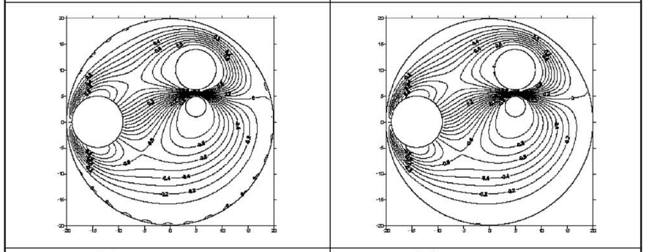
Case 2: A circular plate with three holes [5]. A circular plate with the three circular holes which had been solved by Bird and Steele [5] is revisited by using the present method. The geometric data and the essential boundary conditions are shown in Fig. 11. The contour plots of displacement by using different numbers of terms in the Fourier series ($M=10,20,30,40,50$) are shown in Figs. 12(a)–12(e). It also shows that convergence is good with increasing of the terms of Fourier series. The case was also solved by using the ABAQUS software with 6,606 triangle elements as shown in Fig. 12(g). Good agreement among the data of Bird and Steele [5] as shown in Fig. 12(f), ABAQUS software results as shown in Fig. 12(h), and present solutions is obtained. To discuss the convergence of Fourier series, Parseval sum

$$\int_0^{2\pi} f^2(\theta)d\theta=2\pi a_0^2+\pi\sum_{n=1}^{\infty}(a_n^2+b_n^2) \tag{62}$$

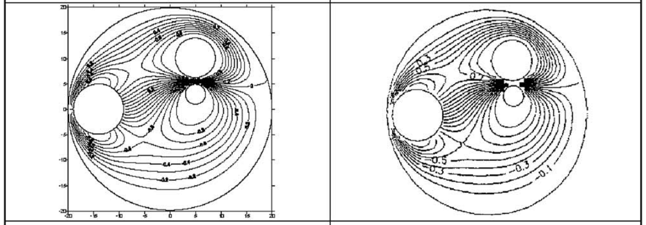
versus terms of Fourier series with respect to m_1, v_1, m_2, v_2, m_3, v_3, m_4, v_4 (determined moment and shear force on the boundaries) are shown in Fig. 13, where $f(\theta)$ is the expansion of Fourier series and a_0, a_n, and b_n are the Fourier coefficients.



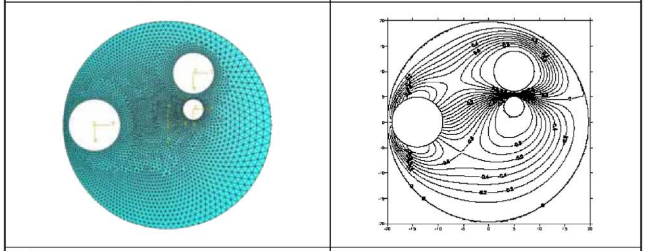
(a) Present method (M=10) | (b) Present method (M=20)



(c) Present method (M=30) | (d) Present method (M=40)



(e) Present method (M=50) | (f) Bird and Steele [2]

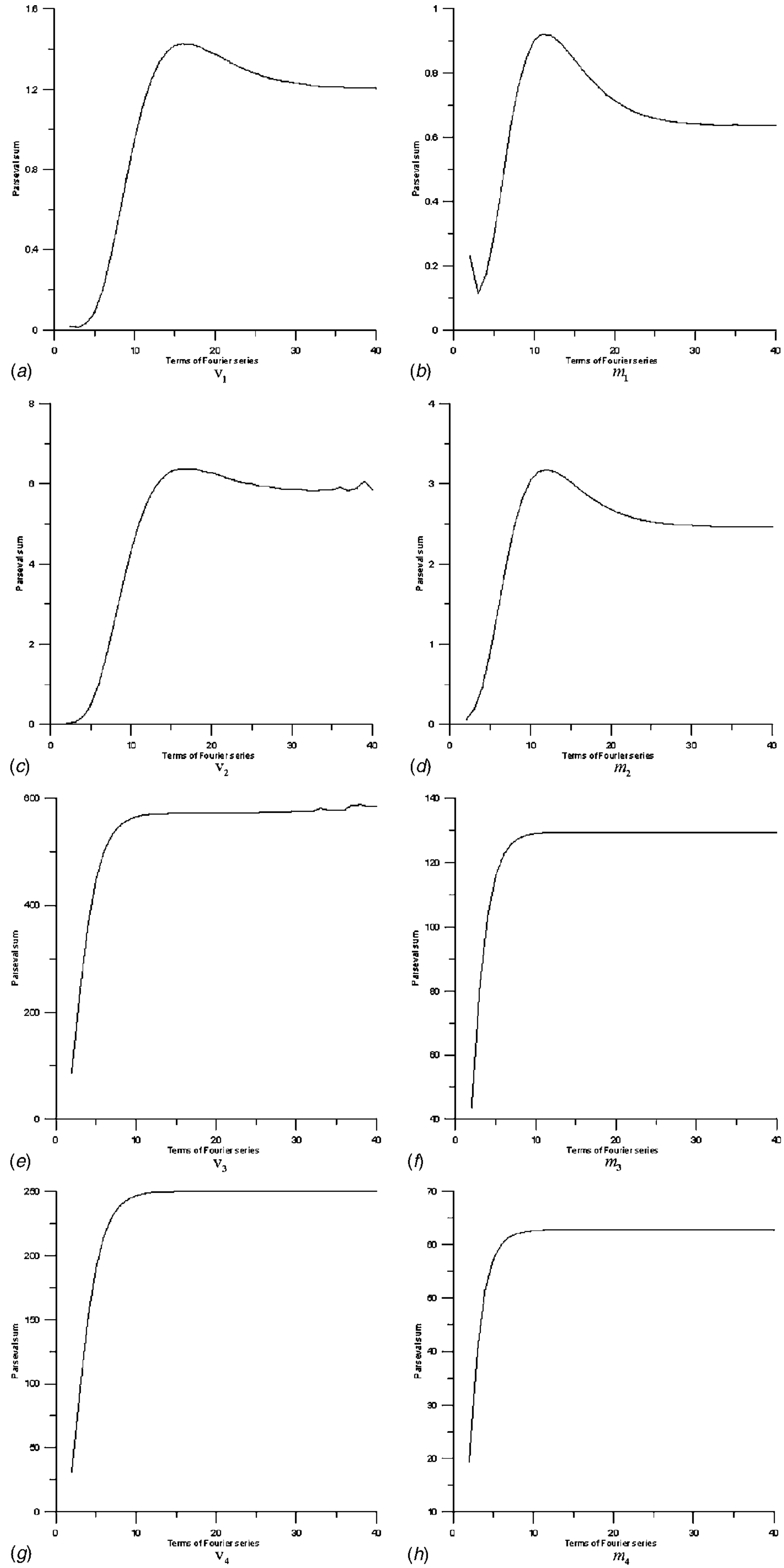


(g) Finite element mesh for ABAQUS program (No. of nodes=3,462, No. of elements=6,606) | (h) FEM (ABAQUS)

Fig. 12 The contour plots of displacement for the plate containing three circular holes subject to the essential boundary conditions by using different methods

7 Conclusions

For plate problems with circular boundaries, a semi-analytical solution by using degenerate kernels, null-field integral equation, and Fourier series in an adaptive observer system was obtained. The main advantage of the present method over BEM is that all the improper integrals are transformed to series sum and can be



(*a*) v_1 (*b*) m_1

(*c*) v_2 (*d*) m_2

(*e*) v_3 (*f*) m_3

(*g*) v_4 (*h*) m_4

Fig. 13 Parseval sum versus terms of Fourier series

easily calculated when degenerate kernels are used. The potential across the boundary can be described explicitly from both sides (interior and exterior). Also, discretization of boundaries is not required. Once the Fourier coefficients of the unknown boundary densities were determined, the displacement, slope, moment, and shear force of the circular plate can be easily determined by substituting the boundary densities into the boundary integral equations for the domain point. Not only the annular plate but also the plate problems with multiple holes have been solved easily and effectively by using the present method in comparison with available exact solution and FEM results. The present method can be applied to plate containing arbitrary number of circular holes as well as various sizes and positions of circular holes. Finally, convergence study on the Fourier series was also done.

Appendix A: Degenerate Kernels

Degenerate kernels for U, Θ, M, V in the first boundary integral equation

$$U(s,x)=\begin{cases} U^I(s,x)=\rho^2(1+\ln R)+R^2\ln R-\left[R\rho(1+2\ln R)+\dfrac{1}{2}\dfrac{\rho^3}{R}\right]\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[\frac{1}{m(m+1)}\frac{\rho^{m+2}}{R^m}-\frac{1}{m(m-1)}\frac{\rho^m}{R^{m-2}}\right]\cos[m(\theta-\phi)], \quad R\geqslant\rho \\ U^E(s,x)=R^2(1+\ln\rho)+\rho^2\ln\rho-\left[\rho R(1+2\ln\rho)+\dfrac{1}{2}\dfrac{R^3}{\rho}\right]\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[\frac{1}{m(m+1)}\frac{R^{m+2}}{\rho^m}-\frac{1}{m(m-1)}\frac{R^m}{\rho^{m-2}}\right]\cos[m(\theta-\phi)], \quad \rho>R \end{cases}$$

$$\Theta(s,x)=\begin{cases} \Theta^I(s,x)=\dfrac{\rho^2}{R}+R(1+2\ln R)-\left[\rho(3+2\ln R)-\dfrac{1}{2}\dfrac{\rho^3}{R^2}\right]\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[\frac{1}{m+1}\frac{\rho^{m+2}}{R^{m+1}}-\frac{m-2}{m(m-1)}\frac{\rho^m}{R^{m-1}}\right]\cos[m(\theta-\phi)], \quad R\geqslant\rho \\ \Theta^E(s,x)=2R(1+\ln\rho)-\left[\rho(1+2\ln\rho)+\dfrac{3}{2}\dfrac{R^2}{\rho}\right]\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[\frac{m+2}{m(m+1)}\frac{R^{m+1}}{\rho^m}-\frac{1}{m-1}\frac{R^{m-1}}{\rho^{m-2}}\right]\cos[m(\theta-\phi)], \quad \rho>R \end{cases}$$

$$M(s,x)=\begin{cases} M^I(s,x)=(\nu-1)\dfrac{\rho^2}{R^2}+(\nu+3)+2(\nu+1)\ln R-\left[(\nu+1)-\dfrac{2\rho}{R}-(\nu-1)\dfrac{\rho^3}{R^3}\right]\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[(\nu-1)\frac{\rho^{m+2}}{R^{m+2}}+\frac{m(1-\nu)-2(1+\nu)}{m}\frac{\rho^m}{R^m}\right]\cos[m(\theta-\phi)], \quad R\geqslant\rho \\ M^E(s,x)=2(1+\nu)(1+\ln\rho)-(\nu+3)\dfrac{R}{\rho}\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[\frac{m(\nu-1)-2(\nu+1)}{m}\frac{R^m}{\rho^m}+(1-\nu)\frac{R^{m-2}}{\rho^{m-2}}\right]\cos[m(\theta-\phi)] \quad \rho>R \end{cases}$$

$$V(s,x)=\begin{cases} V^I(s,x)=\dfrac{4}{R}+\left[\dfrac{2\rho}{R^2}-(3-\nu)\dfrac{\rho^3}{R^4}(1-\nu)\right]\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[m(1-\nu)\frac{\rho^{m+2}}{R^{m+3}}-(4+m(1-\nu))\frac{\rho^m}{R^{m+1}}\right]\cos[m(\theta-\phi)], \quad R>\rho \\ V^E(s,x)=(-3-\nu)\dfrac{1}{\rho}\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[(m(1-\nu)-4)\frac{R^{m-1}}{\rho^m}-m(1-\nu)\frac{R^{m-3}}{\rho^{m-2}}\right]\cos[m(\theta-\phi)], \quad \rho>R \end{cases}$$

Degenerate kernels for U_θ, Θ_θ, M_θ, V_θ in the second boundary integral equation

$$U_\theta(s,x)=\begin{cases} U^I_\theta(s,x)=2\rho(1+\ln R)-\left[R(1+2\ln R)+\dfrac{3}{2}\dfrac{\rho^2}{R}\right]\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[\frac{m+2}{m(m+1)}\frac{\rho^{m+1}}{R^m}-\frac{1}{m-1}\frac{\rho^{m-1}}{R^{m-2}}\right]\cos[m(\theta-\phi)], \quad R\geqslant\rho \\ U^E_\theta(s,x)=\dfrac{R^2}{\rho}+\rho(1+2\ln\rho)-\left[R(3+2\ln\rho)-\dfrac{1}{2}\dfrac{R^3}{\rho^2}\right]\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[\frac{1}{m+1}\frac{R^{m+2}}{\rho^{m+1}}-\frac{m-2}{m(m-1)}\frac{R^m}{\rho^{m-1}}\right]\cos[m(\theta-\phi)], \quad \rho>R \end{cases}$$

$$\Theta_\theta(s,x)=\begin{cases} \Theta^I_\theta(s,x)=\dfrac{2\rho}{R}-\left[(3+2\ln R)-\dfrac{3}{2}\dfrac{\rho^2}{R^2}\right]\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[\frac{m+2}{m+1}\frac{\rho^{m+1}}{R^{m+1}}-\frac{m-2}{m-1}\frac{\rho^{m-1}}{R^{m-1}}\right]\cos[m(\theta-\phi)], \quad R\geqslant\rho \\ \Theta^E_\theta(s,x)=\dfrac{2R}{\rho}-\left[(3+2\ln\rho)-\dfrac{3}{2}\dfrac{R^2}{\rho^2}\right]\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[\frac{m+2}{m+1}\frac{R^{m+1}}{\rho^{m+1}}-\frac{m-2}{m-1}\frac{R^{m-1}}{\rho^{m-1}}\right]\cos[m(\theta-\phi)], \quad \rho>R \end{cases}$$

$$M_\theta(s,x)=\begin{cases} M^I_\theta(s,x)=\dfrac{2\rho}{R^2}(\nu-1)-\left[\dfrac{2}{R}(\nu+1)-3(\nu-1)\dfrac{\rho^2}{R^3}\right]\cos(\theta-\phi) \\ \qquad +\displaystyle\sum_{m=2}^{\infty}\left[(m+2)(\nu-1)\frac{\rho^{m+1}}{R^{m+2}}+(m(1-\nu)-2(1+\nu))\frac{\rho^{m-1}}{R^m}\right]\cos[m(\theta-\phi)], \quad R>\rho \\ M^E_\theta(s,x)=\dfrac{2(1+\nu)}{\rho}+(\nu+3)\dfrac{R}{\rho^2}\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[(m(\nu-1)-2(\nu+1))\frac{R^m}{\rho^{m+1}}+(m-2)(1-\nu)\frac{R^{m-2}}{\rho^{m-1}}\right]\cos[m(\theta-\phi)], \quad \rho>R \end{cases}$$

$$V_\theta(s,x)=\begin{cases} V^I_\theta(s,x)=\left[\dfrac{2}{R^2}(3-\nu)-3(1-\nu)\dfrac{\rho^2}{R^4}\right]\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[m(m+2)(1-\nu)\frac{\rho^{m+1}}{R^{m+3}}-m(4+m(1-\nu))\frac{\rho^{m-1}}{R^{m+1}}\right]\cos[m(\theta-\phi)], \quad R>\rho \\ V^E_\theta(s,x)=(3+\nu)\dfrac{1}{\rho^2}\cos(\theta-\phi) \\ \qquad -\displaystyle\sum_{m=2}^{\infty}\left[m(m(1-\nu)-4)\frac{R^{m-1}}{\rho^{m+1}}-m(m-2)(1-\nu)\frac{R^{m-3}}{\rho^{m-1}}\right]\cos[m(\theta-\phi)], \quad \rho>R \end{cases}$$

where U_θ, Θ_θ, M_θ, V_θ are equal to $\partial U(s,x)/\partial n_x$, $\partial\Theta(s,x)/\partial n_x$, $\partial M(s,x)/\partial n_x$, and $\partial V(s,x)/\partial n_x$, respectively.

Tangential derivative with respect to the field point

$$U_{,t}(s,x)=\begin{cases} U^{I}_{,t}(s,x)=-\left[R(1+2\ln R)+\dfrac{1}{2}\dfrac{\rho^2}{R}\right]\sin(\theta-\phi)-\displaystyle\sum_{m=2}^{\infty}\left[\dfrac{1}{m+1}\dfrac{\rho^{m+1}}{R^{m}}-\dfrac{1}{m-1}\dfrac{\rho^{m-1}}{R^{m-2}}\right]\sin[m(\theta-\phi)], & R>\rho \\ U^{E}_{,t}(s,x)=-\left[R(1+2\ln \rho)+\dfrac{1}{2}\dfrac{R^3}{\rho^2}\right]\sin(\theta-\phi)-\displaystyle\sum_{m=2}^{\infty}\left[\dfrac{1}{m+1}\dfrac{R^{m+2}}{\rho^{m+1}}-\dfrac{1}{m-1}\dfrac{R^{m}}{\rho^{m-1}}\right]\sin[m(\theta-\phi)], & \rho>R \end{cases}$$

$$\Theta_{,t}(s,x)=\begin{cases} \Theta^{I}_{,t}(s,x)=-\left(3+2\ln R-\dfrac{1}{2}\dfrac{\rho^2}{R^2}\right)\sin(\theta-\phi)+\displaystyle\sum_{m=2}^{\infty}\left[\dfrac{m}{m+1}\dfrac{\rho^{m+1}}{R^{m+1}}-\dfrac{m-2}{m-1}\dfrac{\rho^{m-1}}{R^{m-1}}\right]\sin[m(\theta-\phi)], & R>\rho \\ \Theta^{E}_{,t}(s,x)=-\left(1+2\ln \rho+\dfrac{3}{2}\dfrac{R^2}{\rho^2}\right)\sin(\theta-\phi)-\displaystyle\sum_{m=2}^{\infty}\left[\dfrac{m+2}{m+1}\dfrac{R^{m+1}}{\rho^{m+1}}-\dfrac{m}{m-1}\dfrac{R^{m-1}}{\rho^{m-1}}\right]\sin[m(\theta-\phi)], & \rho>R \end{cases}$$

$$M_{,t}(s,x)=\begin{cases} M^{I}_{,t}(s,x)=-\left[\dfrac{2(\nu+1)}{R}-(\nu-1)\dfrac{\rho^2}{R^3}\right]\sin(\theta-\phi) \\ \qquad+\displaystyle\sum_{m=2}^{\infty}\left[m(\nu-1)\dfrac{\rho^{m+1}}{R^{m+2}}+(m(1-\nu)-2(1+\nu))\dfrac{\rho^{m-1}}{R^{m}}\right]\sin[m(\theta-\phi)], & R>\rho \\ M^{E}_{,t}(s,x)=-(\nu+3)\dfrac{R}{\rho^2}\sin(\theta-\phi) \\ \qquad+\displaystyle\sum_{m=2}^{\infty}\left[(m(\nu-1)-2(\nu+1))\dfrac{R^{m}}{\rho^{m+1}}+m(1-\nu)\dfrac{R^{m-2}}{\rho^{m-1}}\right]\sin[m(\theta-\phi)], & \rho>R \end{cases}$$

$$V_{,t}(s,x)=\begin{cases} V^{I}_{,t}(s,x)=\left[\dfrac{2(3-\nu)}{R^2}-\dfrac{\rho^2}{R^4}(1-\nu)\right]\sin(\theta-\phi)-\displaystyle\sum_{m=2}^{\infty}\left[m^2(1-\nu)\dfrac{\rho^{m+1}}{R^{m+3}}-m(4+m(1-\nu))\dfrac{\rho^{m-1}}{R^{m+1}}\right]\sin[m(\theta-\phi)], & R>\rho \\ V^{E}_{,t}(s,x)=(-3-\nu)\dfrac{1}{\rho^2}\sin(\theta-\phi)+\displaystyle\sum_{m=2}^{\infty}\left[m(m(1-\nu)-4)\dfrac{R^{m-1}}{\rho^{m+1}}-m^2(1-\nu)\dfrac{R^{m-3}}{\rho^{m-1}}\right]\sin[m(\theta-\phi)], & \rho>R \end{cases}$$

Appendix B: Analytical Evaluation of the Integral and its Limit for $U(s,x)$ Kernel

	$U(s,x)$ and $\int_B U(s,x)t(s)dB(s)$
Degenerate kernel	$U(s,x)=\begin{cases} U^{I}(s,x)=\rho^2(1+\ln R)+R^2\ln R-\left[R\rho(1+2\ln R)+\frac{1}{2}\frac{\rho^3}{R}\right]\cos(\theta-\phi)-\sum_{m=2}^{\infty}\left[\frac{1}{m(m+1)}\frac{\rho^{m+2}}{R^{m}}-\frac{1}{m(m-1)}\frac{\rho^{m}}{R^{m-2}}\right]\cos[m(\theta-\phi)], & R\geqslant\rho \\ U^{E}(s,x)=R^2(1+\ln \rho)+\rho^2\ln \rho-\left[\rho R(1+2\ln \rho)+\frac{1}{2}\frac{R^3}{\rho}\right]\cos(\theta-\phi)-\sum_{m=2}^{\infty}\left[\frac{1}{m(m+1)}\frac{R^{m+2}}{\rho^{m}}-\frac{1}{m(m-1)}\frac{R^{m}}{\rho^{m-2}}\right]\cos[m(\theta-\phi)], & \rho>R \end{cases}$

Orthogonal process	$R \geqslant \rho$	$\int_0^{2x} U^I \cos(n\theta) R d\theta = \begin{cases} 0, n = \{0 \\ -\pi\left[R^2\,\rho(1+2\ln R)+\frac{1}{2}\rho^3\right]\cos\phi, n=1 \\ -\pi\left[R^2\,\rho(1+2\ln R)+\frac{1}{2}\rho^3\right]\cos\phi-\pi\left[\frac{1}{n(n+1)}\frac{\rho^{n+2}}{R^{n+1}}-\frac{1}{n(n-1)}\frac{\rho^n}{R^{n-1}}\right]\cos(n\phi), n\geqslant 3 \end{cases}$ $\int_0^{2\pi} U^I \sin(n\theta) R d\theta = \begin{cases} 0, n = 1 \\ -\pi\left[R^2\rho(1+2\ln R)+\tfrac{1}{2}\rho^3\right]\sin\phi, n=1 \\ -\pi\left[R^2\rho(1+2\ln R)+\tfrac{1}{2}\rho^3\right]\sin\phi-\pi\left[\frac{1}{n(n+1)}\frac{\rho^{n+2}}{R^{n+1}}-\frac{1}{n(n-1)}\frac{\rho^n}{R^{n-1}}\right]\sin(n\phi), n\geqslant 2 \end{cases}$
	$R < \rho$	$\int_0^{2\pi} U^E \cos(n\theta) R d\theta = \begin{cases} 0, n = 1 \\ -\pi\left[\rho R^2(1+2\ln\rho)+\frac{1}{2}\frac{R^4}{\rho}\right]\cos\phi, n=1 \\ -\pi\left[\rho R^2(1+2\ln\rho)+\frac{1}{2}\frac{R^4}{\rho}\right]\cos\phi-\pi\left[\frac{1}{n(n+1)}\frac{R^{n+3}}{\rho^n}-\frac{1}{n(n-1)}\frac{R^{n+1}}{\rho^{n-2}}\right]\cos(n\phi), n\geqslant 2 \end{cases}$
	$R < \rho$	$\int_0^{2\pi} U^E \sin(n\theta) R d\theta = \begin{cases} 0, n = 1 \\ -\pi\left[\rho R^2(1+2\ln\rho)+\frac{1}{2}\frac{R^4}{\rho}\right]\sin\phi, n=1 \\ -\pi\left[\rho R^2(1+2\ln\rho)+\frac{1}{2}\frac{R^4}{\rho}\right]\sin\phi-\pi\left[\frac{1}{n(n+1)}\frac{R^{n+3}}{\rho^n}-\frac{1}{n(n-1)}\frac{R^{n+1}}{\rho^{n-2}}\right]\sin(n\phi), n\geqslant 2 \end{cases}$
Limiting behavior $R^- = \rho = R^+$		$\lim_{\rho\to R}\int_0^{2\pi} U^I \cos(n\theta) R d\theta = \begin{cases} 0, n = 0 \\ -\pi\left[R^3(1+2\ln R)+\frac{1}{2}R^3\right]\cos\phi, n=1 \\ -\pi\left[R^3(1+2\ln R)+\frac{1}{2}R^3\right]\cos\phi-\pi\left[\frac{1}{n(n+1)}R-\frac{1}{n(n-1)}R\right]\cos(n\phi), n\geqslant 2 \end{cases}$ $\lim_{\rho\to R}\int_0^{2\pi} U^I \sin(n\theta) R d\theta = \begin{cases} 0, n = 1 \\ -\pi\left[R^3(1+2\ln R)+\tfrac{1}{2}R^3\right]\sin\phi, n=1 \\ -\pi\left[R^3(1+2\ln R)+\tfrac{1}{2}R^3\right]\sin\phi-\pi\left[\frac{1}{n(n+1)}R-\frac{1}{n(n-1)}R\right]\sin(n\phi), n\geqslant 2 \end{cases}$ $\lim_{\rho\to R}\int_0^{2\pi} U^E \cos(n\theta) R d\theta = \begin{cases} 0, n = 1 \\ -\pi\left[R^3(1+2\ln R)+\tfrac{1}{2}R^3\right]\cos\phi, n=1 \\ -\pi\left[R^3(1+2\ln R)+\tfrac{1}{2}R^3\right]\cos\phi-\pi\left[\frac{1}{n(n+1)}R^3-\frac{1}{n(n-1)}R^3\right]\cos(n\phi), n\geqslant 2 \end{cases}$ $\lim_{\rho\to R}\int_0^{2\pi} U^E \sin(n\theta) R d\theta = \begin{cases} 0, n = 1 \\ -\pi\left[R^3(1+2\ln R)+\tfrac{1}{2}R^3\right]\sin\phi, n=1 \\ -\pi\left[R^3(1+2\ln\rho)+\tfrac{1}{2}R^3\right]\sin\phi-\pi\left[\frac{1}{n(n+1)}R^3-\frac{1}{n(n-1)}R^3\right]\sin(n\phi), n\geqslant 2 \end{cases}$ (Continuous for $R^- < \rho < R^+$)

References

[1] Tanaka, M., Sladek, V., and Sladek, J., 1994, “Regularization Technique Applied to Boundary Element Methods,” Appl. Mech. Rev., **47**, pp. 457–499.

[2] Sladek, J., and Sladek, V., 1982, “Three-Dimensional Crack Analysis for Anisotropic Body,” Appl. Math. Model., **6**, pp. 374–380.

[3] Chen, J. T., Wu, C. S., and Chen, K. H., 2005, “A Study of Free Terms for Plate Problems in the Dual Boundary Integral Equations,” Eng. Anal. Boundary Elem., **29**, pp. 435–446.

[4] Chen, J. T., Wu, C. S., Chen, K. H., and Lee, Y. T., 2005, “Degenerate Scale for Plate Analysis Using the Boundary Integral Equation Method and Boundary Element Method,” Comput. Mech. (to be published).

[5] Bird, M. D., and Steele, C. R., 1991, “Separated Solution Procedure for Bending of Circular Plates with Circular Holes,” Appl. Mech. Rev., **44**, pp. 27–35.

[6] Crouch, S. L., and Mogilevskaya, S. G., 2003, “On the Use of Somigliana’s Formula and Fourier Series for Elasticity Problems with Circular Boundaries,” Int. J. Numer. Methods Eng., **58**, pp. 537–578.

[7] Gradshteyn, I. S., and Ryzhik, I. M., 1996, *Table of Integrals, Series, and Products*, 5th ed., Academic, New York.

[8] Mogilevskaya, S. G., and Crouch, S. L., 2001, “A Galerkin Boundary Integral Method for Multiple Circular Elastic Inclusions,” Int. J. Numer. Methods Eng., **52**, pp. 1069–1106.

[9] Chen, J. T., Wu, C. S., Lee, Y. T., and Chen, K. H., 2005, “On the Equivalence of the Trefftz Method and Method of Fundamental Solutions for Laplace and Biharmonic Equations,” Comput. Math. Appl. (to be published).

[10] Bird, M. D., and Steele, C. R., 1992, “A Solution Procedure for Laplace’s Equation on Multiple Connected Circular Domains,” ASME J. Appl. Mech., **59**, pp. 398–404.

[11] Chen, J. T., Shen, W. C., and Wu, A. C., 2005, “Null-Field Integral Equation for Stress Field Around Circular Holes Under Antiplane Shear,” Eng. Anal. Boundary Elements (to be published).

[12] Jeffery, G. B., 1921, “Plane Stress and Plane Strain in Bipolar Coordinates,” Philos. Trans. R. Soc. London, Ser. A, **221**, pp. 265–293.

[13] Ling, C. B., 1948, “On the Stresses in a Plate Containing Two Circular Holes,” J. Appl. Phys., **19**, pp. 77–82.

[14] Hsiao, C. C., 2005, “A Semi-Analytical Approach for Stokes Flow and Plate Problems With Circular Boundaries,” Master’s thesis, Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung, Taiwan.

[15] ABAQUS/CAE 6.5, 2004. Hibbitt, Karlsson and Sorensen, Inc., RI.

Journal of Applied Mechanics

Technical Briefs

A Brief Evaluation of Approximation Methods for Microcrack Shielding Problems

Xiangqiao Yan
Research Laboratory on Composite Materials,
Harbin Institute of Technology, Harbin 150001, China
e-mail: yanxiangqiao@hotmail.com

This note can be regarded as a supplement of a paper entitled "On the Use of Approximation Methods for Microcrack Shilding Problems" (H. Cai and K. T. Faber, 1992, ASME J. Appl. Mech. ***59****, 497–501). In this note, an effective approximation method for microcrack shielding problems is introduced briefly and is further used to study the examples given in the Cai and Faber paper for the interaction of a macrocrack with microcracks. It is found that the applicability range of the present approximation method in which the errors are acceptably small is much larger than that of the approximation methods listed in the Cai and Faber paper.* [DOI: 10.1115/1.2127957]

1 Introduction

It is well known that microcrack toughening has been the subject of numerous modeling studies. The two basic approaches are continuum modeling [1–3] and discrete modeling [4–9]. This note concerns with discrete modeling methods. Discrete methods require consideration of the interaction of a macrocrack with microcracks. For many microcracks necessary to treat the toughening problem, it is essential to use approximation methods to keep the amount of computation to a tractable level. Under certain conditions, the results of these approximation methods are close to the exact solution. However, when the macrocrack-microcrack and microcrack-microcrack distances are small, the results of the approximation methods can be significally different from the exact solution. Indiscriminate use of these results could inevitably lead to incorrect conclusions. In Cai and Faber paper [10], three approximation methods were briefly described and estimated by means of the numerical solution (based on the exact formulation). They are Kachanov and Montagut, iterative-average traction, and iterative-point representation. In this note, the displacement discontinuity method with crack-tip elements proposed recently by the author is used to study the examples given in Cai and Faber paper for the interaction of a macrocrack with microcracks. It is found that the applicability range of the present approximation method in which the errors are acceptably small is much larger than the approximation methods listed in Cai and Faber paper.

2 Brief Description of the Present Approximation Method

In the present approximation method, by extending Bueckner's principle [11] suited for a single crack to a general system containing multiple interacting cracks, the original problem (the multiple interacting crack problem) is divided into a homogeneous problem (the one without cracks) subjected to remote loads and a multiple-crack problem in an unloaded body with applied tractions on the crack surfaces. Thus the results in terms of stress intensity factors (SIFs) can be obtained by considering the latter problem, which is analyzed easily by means of the displacement discontinuity method with crack-tip element (a boundary element method) proposed recently by the author [12].

The boundary element method consists of the constant displacement discontinuity element presented by Crouch and Starfied [13] and the crack-tip displacement discontinuity elements due to the author [12]. In the boundary element implementation the left or the right crack-tip element is placed locally at the corresponding left or right each crack tip on top of the ordinary nonsingular displacement discontinuity elements that cover the entire crack surface and the other boundaries. The author [12] used the boundary element method to analyze the SIFs of branched cracks. It was found that the numerical approach is very efficient and accurate for analyzing the branched crack problems.

3 A Brief Evaluation of Approximation Methods

The numerical solution (based upon the exact formulation) obtained by the iterative method for a collinear microcrack of length $2c$ ahead of a macrocrack with an associated applied stress intensity of K_I^∞ (see Fig. 1(*a*)) was verified by Cai and Faber [10] by comparing it with the analytical solution presented by Rubinstein [7] and Rose [8]; see Table 1. Here, the cracked configuration with the finite macrocrack of half-length $a=20c$, as shown in Fig. 1(*b*), is used to model the interaction of a half-infinite macrocrack with a collinear microcrack shown in Fig. 1(*a*). As done by Cai and Faber [10], cases of $L/(2c)$ from 0.05 to 0.5 in steps of 0.05 were considered. Regarding discretization, 30 and 600 elements with an equal size, $2c/30$, are discretized on the microcrack and the mac-

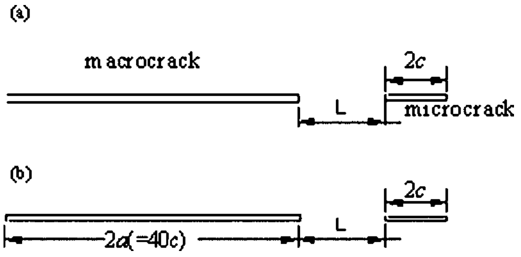


Fig. 1 Macrocrack interaction with a collinear microcrack

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 19, 2005; final manuscript received September 28, 2005. Review conducted by Z. Suo.

Table 1 Comparison of $(K_I - K_I^\infty)/K_I^\infty$ for the macrocrack interaction with a collinear microcrack

	$(K_I - K_I^\infty)/K_I^\infty$				
	Cai and Faber [10]			Present paper	
$L/2c$	Numerical	Analytical	Error (%)	Present	Error (%)
0.05	0.6497	0.6539	0.65	0.6092	−6.8
0.10	0.3859	0.3873	0.40	0.3691	−4.7
0.15	0.2729	0.2737	0.31	0.2640	−3.5
0.20	0.2087	0.2092	0.26	0.2033	−2.8
0.25	0.1671	0.1675	0.23	0.1636	−2.3
0.30	0.1379	0.1382	0.21	0.1355	−2.0
0.35	0.1164	0.1166	0.20	0.1147	−1.6
0.40	0.0999	0.1001	0.19	0.0987	−1.4
0.45	0.0870	0.0871	0.17	0.0860	−1.3
0.50	0.0765	0.0766	0.17	0.0758	−1.0

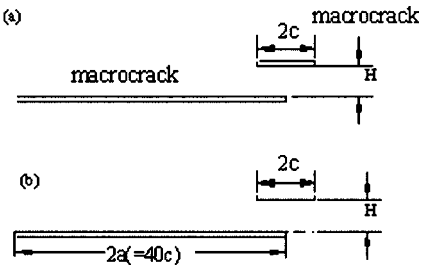


Fig. 2 Macrocrack interaction with a parallel microcrack

rocrack, respectively. The numerical results of SIFs at the macrocrack tip A are also listed in Table 1. It can be seen that the present approximation method has very good accuracy for the interaction of a half-infinite macrocrack with a collinear microcrack shown in Fig. 1(*a*).

The approximation methods: iterative method with average traction (iterative-average), the iterative method with point representation of microcracks (iterative-point) and the approximation method by Kachanov and Montagut [6] were compared with the numerical solution for the two configurations shown in Fig. 1(*a*) (half infinite macrocrack interaction with a collinear microcrack) and in Fig. 2(*a*) (half infinite macrocrack interaction with a parallel microcrack, which centered just above the macrocrack tip). The main emphasis is placed upon the mode I stress intensity, as the mode I shielding is of particular interest. Comparison of the results given by Cai and Faber [10] are shown in Figs. 3 and 4 for the collinear microcrack and the horizontal microcrack cases, respectively. By using the conventional definition of the relative error for the normalized change in the stress intensity $(k_I - K_I^\infty)/K_I^\infty$ and the errors of ten percent or less reasonable for ap-

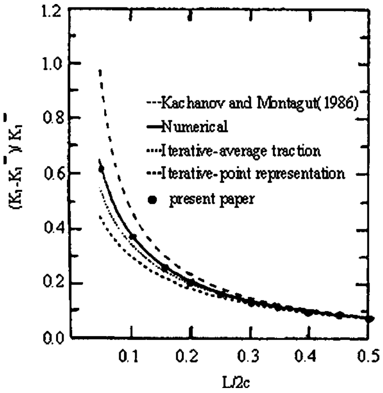


Fig. 3 Comparison of the change in the mode I stress intensity for the collinear microcrack

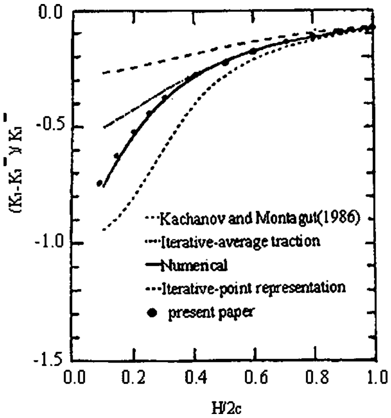


Fig. 4 Comparison of the change in the mode I stress intensity for the horizontal microcrack

Table 2 $(K_I - K_I^\infty)/K_I^\infty$ for the macrocrack interaction with a parallel microcrack

$H/2c$												
0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
−0.889	−0.743	−0.619	−0.518	−0.439	−0.376	−0.283	−0.219	−0.173	−0.140	−0.115	−0.096	−0.081

Table 3 Range of applicability for approximation methods

	Approximation method	Collinear microcrack	Horizontal microcrack
Cai and Faber	Iterative-average	$L/2c > 0.1$	$H/2c > 0.3$
	Iterative-point	$L/2c > 0.3$	$H/2c > 0.9$
	Kachanov-Montagut	$L/2c > 0.2$	$H/2c > 1.0$
Present paper	Present	$L/2c > 0.04$	$H/2c > 0.05$

plicability of the approximation methods, Cai and Faber [10] determined the ranges with which the approximation methods are applicable; see Table 3.

For the interaction of a half-infinite macrocrack with a parallel microcrack shown in Fig. 2(*a*), the configuration shown in Fig. 2(*b*) is used. Cases of $H/(2c)$ from 0.05 to 1.0 were considered. The present approximation results are given in Table 2. In order to compare the present approximation results with those obtained by Cai and Faber [10], the present approximation results are also pictured in Figs. 3 and 4 for the collinear microcrack and the horizontal microcrack, respectively. By comparing the present approximation results with the numerical solutions, which are based on the exact formulation, the applicability range of the present approximation method can be determined; see Table 3. It is found that the applicability range of the present approximation method is much larger than that of the approximation methods mentioned above.

4 Concluding Remarks

This note presents an effective approximation method for microcrack shielding problems. The examples given in the Cai and Faber paper were analyzed to illustrate that the applicability range of the present approximation method in which the errors are acceptably small is much larger than that of the approximation methods listed in the Cai and Faber paper.

By the way, it is pointed out that the macrocrack interaction with a parallel microcrack shown in Fig. 2 is a mixed-mode crack problem, which is analyzed easily by using the present approximate method (e.g., see Yan [12] for a branch crack problem). Because this note concentrated specifically on the approximate method for analyzing microcrack shielding problems, as done by Cai and Faber [12]; here, the mode II SIFs at the macrocrack tip due to the microcrack were not considered.

Acknowledgment

Special thanks are due to the Natural Science Foundation (10272037) of China for supporting the present work.

References

[1] Evans, A. G., and Faber, K. T., 1981, "Toughening of Ceramics by Circumferential Microcracking," J. Am. Ceram. Soc., **67**, 394–398.

[2] Charalambides, P. G., and McMeeking, R. M., 1987, "Finite Element Method Simulation of Crack Propagation in a Brittle Microcracking Solids," Mech. Mater., **6**, 71–87.

[3] Hutchinson, J. W., 1987, "Crack Tip Shielding by Micro-Cracking in Brittle Solids," Acta Metall., **35**, 1605–1619.

[4] Hoagland, R. G., and Embury, J. D., 1980, "A Treatment of Inelasticdeformation Around a Crack Tip Due to Microcracking," J. Am. Ceram. Soc., **63**, 404–410.

[5] Bowling, G. D., Faber, K. T., and Hoagland, R. G., 1987, "Computer Simulation of R-Curve in Microcracking Materials," J. Am. Ceram. Soc., **70**, 849–854.

[6] Kachanov, M., and Montagut, E., 1986, "Interaction of a Crack with Certain Microcrack Arrays," Eng. Fract. Mech., **25**, 625–636.

[7] Rubinstein, A. A., 1985, "Macrocrack Interaction with Semi-Infinite Microcrack Array," Int. J. Fract., **27**, 113–119.

[8] Rose, L. R.F., 1986, "Microcrack Interaction with a Main Crack," Int. J. Fract., **31**, 233–242.

[9] Gong, S.-X., and Horii, H., 1989, "General Solution to the Problem of Microcracks Near the Tip of a Main Crack," J. Mech. Phys. Solids, **37**, 27–46.

[10] Cai, H., and Faber, K. T., 1992, "On the Use of Approximate Methods for Microcrack Shielding," ASME J. Appl. Mech., **59**, 497–501.

[11] Buckner, H. F., 1958, "The Propagation of Cracks and the Energy of Elastic Deformation," Trans. ASME, **80E**, 1225–1230.

[12] Yan, X., 2005, "An Efficient and Accurate Numerical Method of SIFs Calculation of a Branched Crack," ASME J. Appl. Mech., **72**(3), 330–340.

[13] Crouch, S. L., and Starfied, A. M., 1983, *Boundary Element Method in Solid Mechanics, with Application in Rock Mechanics and Geological Mechanics*, London, Geore Allon & Unwin, Bonton, Sydney.

Ice or Snow in the Tempel 1 Comet?

Nicola M. Pugno
Department of Structural Engineering,
Politecnico di Torino, Corso Duca degli Abruzzi 24,
10129 Torino, Italy
e-mail: nicola.pugno@polito.it

An estimation of the Tempel 1 comet strength is deduced as intermediate between that of snow and ice. [DOI: 10.1115/1.2127958]

On July 4, 2005 NASA celebrated the success of the mission *Deep Impact*, culminated with the impact of a ballistic projectile on the *Tempel 1* comet. The projectile, with a mass of $m \sim 370$ Kg, hit the target at a (relative) velocity of $v \sim 37\,000$ Km/h creating a crater of $r \sim 100$ m in radius and $h \sim 40$ m in height.

Material scientists would be happy to measure the mechanical strength σ_C of the comet material, as well as the other solid heavenly bodies. The extraction of specimens for such a purpose has not been realized until now. In spite of this, an estimation of the comet strength can be easily deduced by the following procedure. The kinetic energy K available in the collision is proportional to a power of the crater volume V with a fractional exponent [1]. For example, according to the Gault's scaling [2], assuming geometrical self-similarity (i.e., $h \propto r$) we deduce 1.19 from the lunar crater data. In particular, for large-sized fragmentations the exponent is close to the unity and the constant of proportionality to the mechanical strength of the comminuted material [3]. Thus, from the previous reported experimental observations we can estimate the mechanical strength of the *Tempel 1* comet, to be around

$$\sigma_C \approx \frac{K}{V} = \frac{1/2mv^2}{\pi h/2(r^2 + h^2/3)} \approx 30 \text{ kPa} \tag{1}$$

This value is in between that of terrestrial snows (~1–10 kPa) and ice (~1000 kPa).

To have an idea of such strengths, note that the last value corresponds to the fracture of an ice cubelet just under the weight of a middle-sized man, as can be easily verified in our own home.

References

[1] Carpinteri, A., and Pugno, N., 2002, "One-, two- and Three-Dimensional Universal Laws for Fragmentation due to Impact and Explosion," ASME J. Appl. Mech., **69**, pp. 854–856.
[2] Gault, D., 1974, "Impact Cratering," *Primer of Lunar Geology*, R. Greeley and P. Schultz, eds., NASA Ames Report, pp. 137–175.
[3] Carpinteri, A., and Pugno, N., 2002, "A Fractal Comminution Approach to Evaluate the Drilling Energy Dissipation," Int. J. Numer. Analyt. Meth. Geomech., **26**, pp. 499–513.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 9, 2005; final manuscript received September 22, 2005. Review conducted by R. M. McMeeking.

Copyright © 2006 by ASME

Relation Between Stochastic Failure Location and Strength in Brittle Materials

Sefi Givli
e-mail: mesefi@tx.technion.ac.il

Eli Altus

Faculty of Mechanical Engineering, Technion,
Israel Institute of Technology,
Haifa 32000, Israel

Statistical characteristics of failure location and their relation to strength in brittle materials are studied. One-dimensional rod and bending of a beam under arbitrary distributed loads are studied as examples. The analysis is based on the weakest link approach, and is not confined to specific strength distributions (such as Weibull, Gaussian, etc.). It is found that the statistical moments of the failure location (average, variance, etc.) are directly related to the area moments (centroid, inertia, etc.) of a simple function of the stress field. Therefore, important information related to material strength can be experimentally obtained based on measuring failure locations. Such experiments do not require the measurement of stresses, strains, or displacements, and are very attractive for MEMS/NEMS applications. The approach is general and can be applied to other types of testing specimens.
[DOI: 10.1115/1.2150501]

Introduction and Motivation

The probabilistic nature of strength in brittle materials has been studied extensively [1]. Large dispersion, coupled with size effects were observed in ceramics, concrete, rocks, and other materials. In most cases the strength probability was found to follow the Weibull distribution, and size effects were attributed to the weakest link concept [2]. Similar characteristics have been also observed in micro/nano-structures [3–7], focusing on silicon/polysilicon, which are the prime materials used in MEMS. This probabilistic nature dramatically affects the device reliability and must not be ignored [8–10]. To obtain the statistical properties of the strength, a large number of experiments must be performed, involving the measurement of strains and stresses. This task becomes extremely complex on the micro-nano-scale.

It is shown in this paper that the statistical properties of the failure location can be used to obtain important information regarding the strength distribution. The procedure proposed is based on relatively simple tests which only involve measurements of the failure location. A similar approach, yet based on a different concept, has been applied in Ref. [11]. It should be noted that due to the inherent probabilistic nature of strength, the failure location does not necessarily take place at the location of the highest stress level.

Basic Relations of Random Strength

The analysis of random strength (failure probability) based on the weakest link approach and the Weibull distribution is well established (e.g., Ref. [2]). In this section the emphasis is on less-common relations that include important approximations for cases of low failure probabilities, which leads to a general approach that is not confined to the Weibull distribution. These relations are essential for the analytical derivations in the next section.

Consider a rod of length L under a stress field $\sigma(x)$. Strength is controlled by surface defects, microstructure heterogeneity, etc. Fracture mechanics tools are impractical, since defect geometries and stress concentration fields are too complex to be measured. Therefore, failure probability measures are taken. We rely on some reference (experimental) data, which correspond to the failure probability of an element of a standard, arbitrarily chosen, length ℓ, subjected to a uniform stress $\bar{\sigma}$, i.e.

$$F(\bar{\sigma},\ell)=\int_0^{\bar{\sigma}} f(\sigma',\ell)\,d\sigma' = 1-G(\bar{\sigma},\ell) \tag{1}$$

F and G are failure and survival probabilities, respectively, and f is the failure probability density.

In order to calculate the failure probability of the rod, a failure criterion must be applied. A common approach for brittle materials is based on the weakest link principle, in which any part of the structure leads to the failure of the whole structure. Hence, dividing the rod into N equal elements of length $\ell = L/N$ and considering a "weakest link" failure criterion

$$G(\sigma(x),L)=\prod_{i=1}^{N} G(\bar{\sigma}_i,\ell) \tag{2}$$

Here G is a functional of $\sigma(x)$. ℓ is small enough (N large), such that a uniform stress ($\bar{\sigma}_i$) inside each basic element (i) is assumed. For any G, define [12]

$$g(\bar{\sigma},\ell) = -\ln G(\bar{\sigma},\ell) \rightarrow F(\bar{\sigma},\ell) = 1-\exp(-g(\bar{\sigma},\ell)) \tag{3}$$

where g is a nonnegative, nondecreasing function of σ. From Eq. (3), rewrite Eq. (2) as

$$g(\sigma(x),L)=\sum_{i=1}^{N} g(\bar{\sigma}_i,\ell). \tag{4}$$

In a continuous form

$$g(\sigma(x),L)=\frac{1}{\ell}\int_0^L g(\sigma(x),\ell) \tag{5a}$$

$$g(\sigma(x),L)x=\frac{L}{\ell}\int_0^1 g(\sigma(x),\ell)dx \tag{5b}$$

x in Eq. (5b) is normalized by L. The above is a form of the very well known dependency of the survival probability of a structure on the stress field, used frequently for ceramics [2] and MEMS applications [3,4]. For the particular case of a uniform stress, Eq. (5) reduces to

$$g(\bar{\sigma},L)=\frac{L}{\ell}g(\bar{\sigma},\ell) \tag{6}$$

demonstrating the size-effect in brittle materials.

Equation (5) indicates that when $L \rightarrow 0$, $g \rightarrow 0$ and $G \rightarrow 1$, i.e., a vanishingly small element never fails. This limit possesses no practical difficulty, since measurable elements always have a finite size and strength. Thus, in a structure consisting of a large number of elements, the failure probability of a single element is much smaller compared to the failure probability of the entire structure. This observation is of great importance in obtaining analytical results, as shown in the sequel.

Consider the case of a 1D structure of length L, made from a very large number of elements, under a uniform stress field. Since $\ell \ll L$, the failure probability of each element, associated with the failure stress of the structure, is very small. In this case we have [1]

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 22, 2004; final manuscript received October 19, 2005. Review conducted by K. Ravi-Chandar.

Copyright © 2006 by ASME

$$F(\bar{\sigma},\ell) \cong g(\bar{\sigma},\ell) \tag{7}$$

Moreover, g can be approximated as a power function [8,13]

$$g(\bar{\sigma},\ell) = \left(\frac{\bar{\sigma}}{\sigma_\ell}\right)^\beta . \tag{8}$$

Inserting Eq. (8) in Eq. (3), we obtain the Weibull distribution function

$$F_W(\bar{\sigma},\ell) = 1 - \exp\left(-\left(\frac{\bar{\sigma}}{\sigma_\ell}\right)^\beta\right) \tag{9}$$

For the Weibull distribution, σ_ℓ is approximately the average strength of the element, and β is associated with the dispersion (higher β indicates lower dispersion). It is evident that Weibull distribution is a good approximation for a large class of probabilistic strength, provided that the stresses are low enough compared to some reference value (average strength for example). Combining Eqs. (7) and (8) we have simply

$$1 \gg F(\sigma,\ell) \cong \left(\frac{\sigma}{\sigma_\ell}\right)^\beta \tag{10}$$

Note that Eq. (10) is not confined to the Weibull distribution. Nevertheless, since the Weibull probability function is commonly used for brittle materials and MEMS applications, interpretation of the results will be with reference to the Weibull distribution. Note that the Weibull distribution has no local maximum point (most probable value) for values of $\beta \leqslant 1$. Therefore, practical values of β are expected to be greater than 1, a conclusion which will be used in the following. Indeed, values of $3 < \beta < 25$ are found for most brittle materials [2–6]. Values of $\beta < 1$ are usually associated with "infant mortality" type of failures [14] related to problems such as assembly, production, etc., and not to strength.

Failure Location in a Tensile Test

The failure location of a homogeneous specimen loaded slowly up to failure is at the point of maximum stress. However, in materials with random strength, an element with a lower stress may fail first. Therefore, failure is only more probable at locations of higher stresses. Thus, failure location is a random parameter, which is a functional of the whole stress field.

Consider a structure of N elements, each of size ℓ under unidirectional (not necessarily uniform) stress field, loaded slowly up to failure. The probability of failure at a *specific* element k is equal to the probability of failure of the kth, divided by the probability of failure at *any* element, i.e.

$$p_k = \frac{F_k \cdot \left(\dfrac{\prod_{j=1}^{N}(1-F_j)}{1-F_k}\right)}{\sum_{m=1}^{N}\left[F_m \cdot \left(\dfrac{\prod_{j=1}^{N}(1-F_j)}{1-F_m}\right)\right]} = \frac{[\prod_{j=1}^{N}(1-F_j)] \cdot \dfrac{F_k}{1-F_k}}{\sum_{m=1}^{N}\left[(\prod_{j=1}^{N}(1-F_j)) \cdot \dfrac{F_m}{1-F_m}\right]} = \frac{\phi_k}{\sum_{m=1}^{N}\phi_m} \tag{11}$$

where

$$\phi_j = \frac{F_j}{1-F_j} \tag{12}$$

and F_j is the failure probability of the jth element. Hence, the cumulative probability for the failure location, or the probability that the structure will fail at a location lower (or equal) than the kth element is simply

$$P_k = \frac{\sum_{m=1}^{k}\phi_m}{\sum_{m=1}^{N}\phi_m} . \tag{13}$$

In a continuous form

$$F_x = \frac{\displaystyle\int_0^x \phi(x_1)dx_1}{\displaystyle\int_0^1 \phi(x_1)dx_1} \tag{14}$$

Differentiation yields

$$f_x = \frac{d}{dx}F_x = \frac{\phi(x)}{\displaystyle\int_0^1 \phi(x_1)dx_1} \tag{15}$$

F_x and f_x are the probability and probability density functions, respectively, for the failure location, and $\phi(x_1)$ is associated with Eq. (12) for an element located at x_1. For the particular case of a uniform stress distribution, $\phi(x_1)$ is constant (independent of x_1), i.e.

$$f_{x(\text{uniform})} = 1 \tag{16}$$

as expected. The average failure location in this case is at the middle point

$$\langle x_{(\text{uniform})}\rangle = \int_0^1 x \cdot f_x dx = \frac{1}{2} . \tag{17}$$

Now consider the particular case where $\{ \ll L$. Since the failure probability is positive and smaller than 1, we write

$$\phi(x) = \frac{F(\sigma(x),\ell)}{1-F(\sigma(x),\ell)} = F(\sigma(x),\ell) + [F(\sigma(x),\ell)]^2 + [F(\sigma(x),\ell)]^3 + \cdots \tag{18}$$

As discussed above, when $l \ll L$ the failure stress level corresponds to a very low failure probability of each element (see the demonstration in Ref. [8]). Thus, using Eq. (10) we approximate Eq. (18) by considering its first term only

$$\phi(x) \cong \left(\frac{\sigma(x)}{\sigma_\ell}\right)^\beta \tag{19}$$

The probability density function of the failure location (15) reduces to

$$f_x = \frac{\sigma(x)^\beta}{\displaystyle\int_0^1 \sigma(x_1)^\beta dx_1} \tag{20}$$

From Eq. (21) it is seen that f_x is invariant to the properties of the basic element (σ_l, l), and depends solely on the stress field and on β. Since β is related to the strength dispersion of *both* the basic element and the entire structure (see Eq. (6)), β can be obtained by measuring statistical properties of the failure location (such as average and variance).

Using Eq. (21), the average failure location is

$$\langle x\rangle = \int_0^1 x \cdot f_x dx = \frac{\displaystyle\int_0^1 x \cdot \sigma(x)^\beta dx}{\displaystyle\int_0^1 \sigma(x_1)^\beta dx_1} \tag{21}$$

Denoting ψ as

$$\psi(x) = [\sigma(x)]^\beta \tag{22}$$

we see that the average failure location is at the centroid of ψ. Furthermore, all higher statistical moments of failure location are associated with the corresponding moments of $\psi(x)$, i.e.

$$\langle x'^n \rangle = \int_0^1 x'^n \cdot f_x dx = \frac{\int_0^1 x'^n \cdot \psi(x) dx}{\int_0^1 \psi(x) dx} = \int_0^1 x'^n \cdot \tilde{\psi}(x) dx$$

$$\tilde{\psi}(x) = \frac{\psi(x)}{\int_0^1 \psi(x) dx} \tag{23}$$

x' is the deviation from the average failure location $\langle x \rangle$. For the extreme case of $\beta \to 1$ (large dispersion in strength), the average failure location is found at the centroid of the stress field itself and for $\beta \to \infty$ (nonrandom strength), the average failure location is just at the location where the highest stress level exists whereas the variance ($n=2$) vanishes, as expected. Thus, β is obtained based on the failure location.

Since finding β this way requires only simple instrumentation and inexpensive technology (no need to measure stresses, displacements, or strains) it introduces a new alternative for obtaining important mechanical characteristics of the structure. This feature is extremely important in MEMS/NEMS applications, where measurements of stresses and strains are complex. Moreover, the suggested method can be used in MEMS mass-productions for detecting process anomalies by unexpected changes in β.

Failure Location in a Bending Test

In the previous section, the statistical characteristics of the failure location were obtained for a 1D rod under a tensile stress field. In practice, especially with micro-nano-specimens, it is easier to perform a bending test. The failure location analysis of a beam under bending differs from the unidirectional tensile case since the stress distribution through the thickness of the beam is nonuniform. Moreover, part of the cross section is under tensile stress, while the other is under compression.

Another important issue related to the analysis of beams is whether failure occurs at the surface of the beam or in its interior. To demonstrate the difference, consider the size effect on the strength of a beam: if failure occurs in the interior of the beam, the size effect on strength (as well as the failure probability) is related to the *volume* of the beam. However, if failure begins at the surface of the beam the size effect of strength corresponds to the area of the beam surface. Both types of failure modes can be found in brittle materials [2–7]; yet, the latter, which is usually associated with surface defects and stress concentrations due to surface roughness, is more common. Interestingly, assuming an Euler-Bernoulli behavior, it can be shown (details are not given here) that in the case of a symmetric cross section, the failure location analysis of a beam under a bending moment field leads to identical expressions for both failure modes (failure at the surface of the beam or in its interior). The reason lays in the normalized expression (e.g., Eqs. (11) and (14)) associated with the probability that failure will occur at a certain cross-section along the length of the beam.

For brittle materials, the compressive strength is much higher than the tensile strength. Therefore, the probability of failure due to compressive stress is usually neglected [2,4,8]. As a result, the statistical moments of the failure location obtained in Eqs. (23) and (21) are valid also for bending, except that the definition of ψ in Eq. (22) should be replaced by

$$\psi(x) = |M(x)|^\beta \tag{24}$$

where $M(x)$ is the bending moment field. In many common tests, such as bending of a cantilever with a point-force, or a three-point bending specimen, M does not change its sign, so the absolute value in Eq. (24) is not necessary. Also, it should be noted that for the three point bending test (as for any test involving a symmetric stress-field), the average failure location is expected to be at the midspan, regardless of the strength dispersion. Therefore, finding β in this case requires higher statistical moments (variance, etc.) of the failure location.

For demonstration, consider an ensemble of cantilever beams loaded up to failure by a point force at its free end. We have

$$M(x) = M_0 \cdot x \tag{25}$$

where M_0 is proportional to the magnitude of the force. From Eqs. (21) and (23) the average and variance of the failure location are, respectively

$$\langle x \rangle = \frac{\beta+1}{\beta+2} \quad \langle x'^2 \rangle = \frac{\beta+1}{(\beta+3)(\beta+2)^2} \tag{26}$$

As expected, for very high values of β, the average failure location approaches 1, corresponding to the location of the highest bending moment, and the variance vanishes. For very large strength dispersions ($\beta \to 1$) the average failure location is found at the centroid of the bending moment field ($x=2/3$), and the coefficient of variation (COV, the standard deviation normalized by the average) approaches 35%.

Conclusions

- A simple relation between the failure location statistics and the statistical properties of the failure were obtained for tensile and bending tests. The approach is valid for other testing methods.
- Measuring failure location is simple, inexpensive, and does not require measurements of stresses, displacements, or strains. It is especially important to micro-nano-applications, and can be used to detect process anomalies during manufacturing.

Acknowledgment

This research was supported in part by the Israeli Science Foundation.

References

[1] Ashby, M. F., and Jones, D. E. H., 1986, *Engineering Materials 2*, Pergamon, New York.

[2] Davidge, R. W., 1979, *Mechanical Behavior of Ceramics*, Cambridge University Press, Cambridge, UK.

[3] Greek, S., Ericson, F., Johansson, S., Furtsch, M., and Rump, A., 1999, “Mechanical Characterization of Thick Polysilicon Films: Young’s Modulus and Fracture Strength Evaluated with Microstructures,” J. Micromech. Microeng., **9**, pp. 245–151.

[4] Jones, P. T., Johnson, G. C., and Howe, R. T., 1999, “Statistical Characterization of Fracture of Brittle MEMS Materials,” *Proceedings MEMS Reliability for Critical and Space Applications*, Soc. Of Photo-Optical Instr. Eng., Bellingham WA, pp. 20–29.

[5] Sharpe, W. N., Yuan, B., Vaidynathan, R., and Edwards, R., 1997, “Measurements of Young Modulus, Poisson Ratio and Tensile Strength of Polysilicon,” *Proceedings of the 10th IEEE International Workshop on Microelectromechanical Systems*, IEEE, Piscataway, NJ, pp. 424–429.

[6] Elbrecht, L., and Binder, J., 1999, “The Mechanical Properties of Thin Polycrystalline Silicon Films as Function of Deposition and Doping Conditions,”

Sens. Mater., **11**(3), pp. 163–179.

[7] Namazu, T., Isono, Y., and Tanaka, T., 2000, "Evaluation of Size Effect on Mechanical Properties of Single Crystal Silicon by Nanoscale Bending Test Using AFM," J. Micromechanical Syst., **9**(4), pp. 450–459.

[8] Altus, E., and Givli, S., 2003, "Strength Reliability of Statistical Heterogeneous Microbeams," Int. J. Solids Struct., **40**(9), pp. 2069–2083.

[9] Givli, S., and Altus, E., 2003, "Effect of Modulus-Strength Correlation on the Reliability of Randomly Heterogeneous Beams," Int. J. Solids Struct., **40**(24), pp. 6703–6722.

[10] Ostoja-Starzewski, M., 2004, "Fracture of Brittle Microbeams," J. Appl. Mech., **71**, pp. 424–427.

[11] Elata, D., and Hirshberg, A., 2005, "A Novel Test Device for Measuring the Strength of Microbeams," J. Microelectromech. Syst., (to be published).

[12] Weibull, W., 1951, "A Statistical Distribution Function of Wide Applicability," J. Appl. Mech., **18**(3), pp. 293–297.

[13] Ballarini, R., Mullen, R. L., and Heuer, A. H., 1999, "The Effects of Heterogeneity and Anisotropy on the Size Effect in Cracked Polycrystalline Films," Int. J. Fract., **95**, pp. 19–39.

[14] Abernethy, R. B., 2000, *The New Weibull Handbook*, SAE International, North Palm Beach, FL.

Continuum Damage Mechanics and the Life-Fraction Rule

U. Stigh
Division of Mechanical Engineering,
University of Skövde,
P. O. Box 408, SE-541 28, Skövde, Sweden

This paper gives a short review of two different methods for life prediction at high temperature; namely continuum damage mechanics (CDM) and the linear life-fraction rule (LFR). It is well known that the class of CDM theories with a separable evolution law gives a life prediction in accordance with the LFR. However, it appears to be an open question if this is a necessary condition. It is here shown that in order for a CDM theory to comply with the LFR it must have a separable evolution law. That is, if we can assume that a material follows the LFR, it is necessary to chose a separable evolution law for this material. The reverse is also true, to get a life-fraction different from unity, we must chose a non-separable evolution law. [DOI: 10.1115/1.2150502]

1 Introduction

At temperatures above about one-third of their homologous temperature, T_m, engineering materials tend to rupture when exposed to a constant stress, cf. e.g., Ref. [1]. This is known as creep rupture and has long been a subject of intense research due to its engineering importance. At relatively low temperatures and high stress, rupture is preceded by large strains while at high temperature and low stress, very small strain may precede rupture.

Experiments performed at constant temperature and stress often show an exponential relation between the time to rupture, t_R, and the applied uniaxial stress σ. Thus

$$t_R(\sigma) = A\sigma^a \tag{1}$$

where A and a are material constants that can be determined experimentally. In the sequel, the symbol $t_R(\sigma)$ is reserved for the time to creep rupture under *constant* stress. This function may be determined experimentally or theoretically.

In this paper, two alternative methods to predict life under *varying* stress are treated; the linear life-fraction rule (LFR) and the method of continuum damage mechanics (CDM). Specifically, we will derive the *necessary* and *sufficient* conditions for a CDM theory to yield the same predictions as the LFR. Our interest in this problem stems from recent results by Alfredsson and Stigh [2], who show that it is possible to develop thermodynamically consistent CDM theories which yield a linear life-fraction different from unity. In this introductory section, the two sets of methods/theories are first presented. In the following section, some known results on the *sufficient* conditions under which a CDM theory conform with the LFR are presented. In the second section, a simplified loading sequence is first presented and the *necessary* condition under which a CDM theory yields the same predictions as the LFR is derived. This result is shown to be valid under general loading. The paper is ended with a discussion on the implications of this result to other situations.

1.1 Robinson's Life-Fraction Rule. In the industry, the most used method for life prediction is Robinson's linear life-fraction rule [3]. According to this method, "the expenditure of each particular fraction of the life span at elevated temperature is independent of and without influence upon the expenditure of all other fractions of the life to rupture." With stepwise constant stress levels, σ_i, the life-fraction is defined by

$$L_f \equiv \sum_{i=1}^{n} \frac{t_i}{t_R(\sigma_i)} \tag{2}$$

where t_i is the time spent under stress σ_i. At rupture, $L_f = 1$, which is the rupture criteria. With a continuously varying stress, $\sigma(t)$, the life-fraction is generalized to

$$L_f(t) \equiv \int_0^t \frac{dt}{t_R(\sigma(t))} \tag{3}$$

The time to rupture, t^*, is solved from the rupture criteria $L_f(t^*) = 1$. Note that the symbol t^* is reserved for the time to rupture under general loading $\sigma(t)$ while the symbol t_R indicates the time to rupture under constant stress. Thus, $t_R(\sigma_0) = t^*(\sigma \equiv \sigma_0)$, where σ_0 is constant. It is immediately recognized that $L_f(t_R) = 1$ if $\sigma = \text{const}$. an important question is obviously how good these relations conform with experiments. Due to large scatter in experimental results, this appears to be an open question which might lack a materially independent answer. An attempt to study the problem is reported by Jansson [4]. In the study, experiments are performed on an austenitic stainless steel (AISI 316) at 700°C. Both experiments with a constant stress, yielding the relation $t_R(\sigma)$, and step-loading are performed. In these, the stress is held constant during one time period and then changed to another constant level which is kept until rupture, cf. Fig. 1. Two types of experiments are conducted, "step-up" where the second stress level is higher than the first and "step-down" where it is lower. The life-fraction varies according to Fig. 1. Thus, step-up appears to give a somewhat lower life-fraction at rupture than step-down. The difference is perhaps too small to be of any engineering significance. This is indeed the basic assumption in much engineering work.

It should be emphasized that the LFR do not couple the gradual degradation of the material with other constitutive properties. This can lead to large errors when analyzing a hyperstatic structure. If, for instance, the applied load on a hyperstatic structure is kept constant, the degradation leads to a redistribution of the stress with time. Generally, the larger stresses decrease and the lower increase. This means that an ignorant application of a design rule according to Eq. (2) gives an overly conservative estimate of the failure time.

1.2 Continuum Damage Mechanics. A method to consider the degradation of constitutive properties, integrated with a theory for life prediction, is provided by the method of continuum damage mechanics (CDM). The original ideas of Kachanov [5] is inspired by the observation that, at temperatures above about $T_m/2$, grain boundary cavities form with time. The cavities nucleate, grow, and eventually coalesce to form a major crack which leads to final creep rupture. In order to consider the influence of the grain boundary cavities, a "net stress," $\tilde{\sigma}$, and a damage variable, ω, are introduced; for a virgin material, $\omega = 0$ and for rupture $\omega = 1$. The net stress, which might be visualized as the load on the remaining "load bearing cross section" of the material, is related to the stress by

$$\tilde{\sigma} = \frac{\sigma}{1 - \omega}. \tag{4}$$

Thus, for a virgin material without cavities, the net stress equals the acting stress. At rupture no load bearing area remains and the net stress tends to infinity when $\omega \to 1$. Kachanov assumes the damage growth law

---

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 12, 2005; final manuscript received October 24, 2005. Review conducted by K. Ravi-Chandar.

Copyright © 2006 by ASME

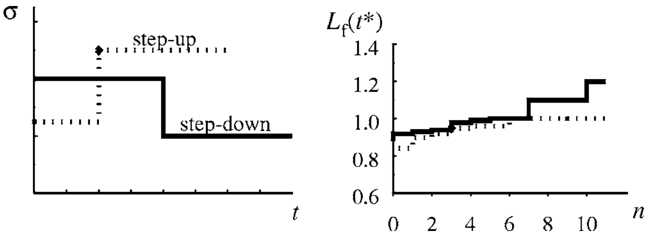


Fig. 1 Left panel: Step-up and step-down experiments. Right panel: Life-fraction at rupture during step-up and step-down experiments performed with an austenitic stainless steel at 700°C. The ordinate shows the life-fraction at rupture and the abscissa shows the number *n* of experiments yielding a life-fraction less than the value on the ordinate. Experimental results from Ref. [4].

$$\dot{\omega} = C\left(\frac{\sigma}{1-\omega}\right)^{\nu} \tag{5}$$

where C and ν are material constants. With a known $\sigma(t)$, Eq. (5) provides a separable differential equation for $\omega(t)$. With the initial condition $\omega(0)=0$ and the rupture criteria $\omega(t^*)=1$, integration yields

$$C(1+\nu)\int_0^{t^*} \sigma(t)^{\nu}\mathrm{d}t = 1 \tag{6}$$

This equation provides an implicit expression for the time to rupture t^* under varying stress. If the stress is constant, the remaining integral is readily evaluated and the time to rupture is given by

$$t_{\mathrm{R}} = \frac{1}{C(\nu+1)\sigma^{\nu}} \tag{7}$$

Thus, the theory predicts an exponential relation between t_R and σ, which is indeed often observed, cf. Eq. (1). From this equation, the material constants C and ν can be determined from experiments performed with constant stress.

1.3 Conditions Under Which CDM Implies the LFR. If we assume that a material follows Kachanov's equation, Eq. (5), it also follows the LFR. To show this, form the life-fraction at rupture, Eq. (3), using the time to rupture under constant stress from Kachanov's theory, cf. Eq. (7). The result is

$$L_{\mathrm{f}} = C(\nu+1)\int_0^{t^*} \sigma(t)^{\nu}\mathrm{d}t = 1 \tag{8}$$

where Eq. (6) is used in the last equality. Thus, Kachanov's version of CDM implies the LFR. This result was shown in Refs. [6,7]. It is also known that any separable evolution law implies the LFR. That is, if

$$\dot{\omega} = g(\sigma)h(\omega) \tag{9}$$

then $L_{\mathrm{f}}(t^*)=1$. This is shown by integrating the separable differential equation between $\omega(0)=\omega_1$ and $\omega(t^*)=\omega_2$, where we have introduced generalized initial and rupture conditions $\omega(0)=\omega_1$ and $\omega(t^*)=\omega_2$, respectively. The result is

$$\int_0^{t^*} g(\sigma(t))\mathrm{d}t = \int_{\omega_1}^{\omega_2} \frac{\mathrm{d}\omega}{h(\omega)} \equiv \Psi, \tag{10}$$

where the symbol Ψ is introduced for the integral on the right-hand side. Note that with a specific CDM model, Ψ is a constant. Now, if σ is a constant, the left-hand side is readily integrated and the time to rupture is given by

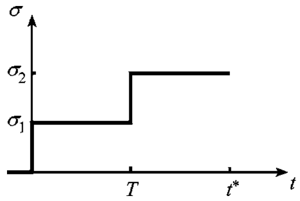


Fig. 2 Two-step load history. During the time $0<t<T$, $\sigma=\sigma_1$. From $t=T$, $\sigma=\sigma_2$ which is held constant until rupture at $t=t^*$.

$$t_{\mathrm{R}}(\sigma) = \frac{\Psi}{g(\sigma)} \tag{11}$$

With a varying stress, the life-fraction at rupture is given by Eq. (3) using Eq. (11), viz.

$$L_{\mathrm{f}}(t^*) = \int_0^{t^*} \frac{g(\sigma(t))}{\Psi}\mathrm{d}t = 1 \tag{12}$$

where Eq. (10) is used in the last equality.

Thus, any CDM theory with a separable evolution law yields a life-fraction equal to unity at rupture. However, in a hyperstatic structure, any CDM theory will generally predict another lifetime than the LFR since the LFR does not provide a modeling of the load redistribution due to the degradation of the material. This phenomenon is usually modeled with a CDM theory.

2 Necessary Conditions

In the introduction we showed that any evolution law

$$\dot{\omega} = f(\sigma,\omega) \tag{13}$$

that is separable according to Eq. (9) yields $L_{\mathrm{f}}(t^*)=1$ at rupture. We will now show that this is a *necessary condition* for a CDM theory to conform with the LFR. To show this, first study a simplified two-step load history, cf. Fig. 2. For this load history, we will show that Eq. (13) yields $L_{\mathrm{f}}(t^*)=1$ at rupture only if the function f is separable according to Eq. (9). Thus, if a CDM theory shall conform with the LFR for *any load history*, the only possibility is for f to be separable according to Eq. (9).

For greater generality, chose $\omega(0)=\omega_1$ as the initial value of damage and $\omega(t^*)=\omega_2$ as the rupture criteria. During the time period $0<t<T$, $\sigma=\sigma_1$ and the evolution law, Eq. (13), is separable. Integration yields

$$\int_{\omega_1}^{\omega_{\mathrm{T}}} \frac{\mathrm{d}\hat{\omega}}{f(\sigma_1,\hat{\omega})} = T \tag{14}$$

where ω_{T} is the damage at $t=\mathrm{T}$. This equation is now conveniently rewritten according to

$$T = \int_{\omega_1}^{\omega_{\mathrm{T}}} \ldots = \int_{\omega_1}^{\omega_2} \ldots - \int_{\omega_{\mathrm{T}}}^{\omega_2} \ldots = t_{\mathrm{R}}(\sigma_1) - \int_{\omega_{\mathrm{T}}}^{\omega_2} \frac{\mathrm{d}\hat{\omega}}{f(\sigma_1,\hat{\omega})} \tag{15}$$

Similarly, in the remaining life, i.e., in the time interval $T<t\leq t^*$, $\sigma=\sigma_2$ and Eq. (13) is again a separable differential equation. Integration yields

$$\int_{\omega_{\mathrm{T}}}^{\omega_2} \frac{\mathrm{d}\hat{\omega}}{f(\sigma_2,\hat{\omega})} = t^* - T \tag{16}$$

We can now form the life-fraction at rupture, cf. Eq. (2)

$$L_{\mathrm{f}}(t^{*}) = \frac{T}{t_{\mathrm{R}}(\sigma_1)} + \frac{t^{*}-T}{t_{\mathrm{R}}(\sigma_2)} = \frac{t_{\mathrm{R}}(\sigma_1) - \displaystyle\int_{\omega_{\mathrm{T}}}^{\omega_2} \frac{\mathrm{d}\hat{\omega}}{f(\sigma_1,\hat{\omega})}}{t_{\mathrm{R}}(\sigma_1)} + \frac{\displaystyle\int_{\omega_{\mathrm{T}}}^{\omega_2} \frac{\mathrm{d}\hat{\omega}}{f(\sigma_2,\hat{\omega})}}{t_{\mathrm{R}}(\sigma_2)} \tag{17}$$

where Eqs. (15) and (16) are used in the last equality. Finally, we evaluate the rupture times for the load levels σ_1 and σ_2 using Eqs. (14) and (16). In Eq. (14) we replace ω_{T} with ω_2 and T with $t_{\mathrm{R}}(\sigma_1)$ and in Eq. (16) we replace ω_{T} with ω_1 and $t^{*}-T$ with $t_{\mathrm{R}}(\sigma_2)$. Inserting in Eq. (17) and simplifying the result yields

$$L_{\mathrm{f}}(t^{*}) = 1 - \frac{F(\omega_{\mathrm{T}},\sigma_1)}{F(\omega_1,\sigma_1)} + \frac{F(\omega_{\mathrm{T}},\sigma_2)}{F(\omega_1,\sigma_2)} \tag{18}$$

where we have defined

$$F(\omega,\sigma) \equiv \int_{\omega}^{\omega_2} \frac{\mathrm{d}\hat{\omega}}{f(\sigma,\hat{\omega})} \tag{19}$$

Thus, a necessary condition for $L_{\mathrm{f}}(t^{*})=1$ is that the relation

$$\frac{F(\omega,\sigma)}{F(\omega_1,\sigma)} \equiv C \tag{20}$$

is independent of the parameter σ. As evident from the equation above, the function C is, at most, a function of ω and σ. If C is independent of σ we can rewrite Eq. (20) as

$$F(\omega,\sigma) = C(\omega)F(\omega_1,\sigma) \tag{21}$$

Thus

$$C(\omega)F(\omega_1,\sigma) = \int_{\omega}^{\omega_2} \frac{\mathrm{d}\hat{\omega}}{f(\sigma,\hat{\omega})}, \tag{22}$$

where we have used Eq. (19) in the last equality. Differentiation with respect to ω yields

$$C'(\omega)F(\omega_1,\sigma) = -\frac{1}{f(\sigma,\omega)} \tag{23}$$

Thus

$$f(\sigma,\omega) = g(\sigma)h(\omega) \tag{24}$$

is a necessary condition for a CDM theory to predict the same life as the LFR for a step-loading. Since a separable evolution law always implies the LFR, we have shown that this is the *only* possibility for general loading.

3 Discussion

CDM provides methods to develop constitutive equations within the concept of continuum mechanics to deal with the gradual degradation of materials during loading. In recent years, the implications of the second law of thermodynamics on CDM theories have been scrutinized to give guidance on the structure of a CDM theory. A recent example is given in Ref. [2] that gives a framework for a class of CDM theories. The basic idea is that, if the damage variable is constant, the material responds as the undamaged material with modified stress-like variables. The modifications are, in principle, given by the structure $\tilde{\sigma}=\sigma/N(\omega)$. By identifying the "damage stress," Ω, as the work-conjugated quantity to ω, the requirements of the evolution law for damage to conform with the dissipation inequality are identified. It is shown that this stress is partitioned in two parts; one is the elastoplastic damage stress, Ω^{ep}, which essentially measures the release of elastic and plastic free energy during growth of damage. The second part is identified as a "cohesive damage stress," Ω^{c}, which measures the increase of cohesive energy during growth of damage. This second term facilitates a possibility to model healing of damage without violating the second law of thermodynamics. In Ref. [2], the theory is exemplified with the following evolution law

$$\dot{\omega} = \frac{\sigma^2}{2E^2(1-\omega)^2} - \frac{E_{\mathrm{c}}}{E}\omega \tag{25}$$

which is both thermodynamically consistent and nonseparable. In the equation, E and E_{c} are material parameters. As expected for a nonseparable evolution law, $L_{\mathrm{f}}(t^{*}) \neq 1$. Moreover, as intuitively expected the life-fraction evolves quicker than the damage, i.e., $L_{\mathrm{f}}(t^{*}) > 1$. In order to develop a model yielding $L_{\mathrm{f}}(t^{*}) < 1$ one may couple a time-independent damage growth to the time-dependent part. This would result in an additional increase of damage for each load cycle that yields a quicker damage growth than expected from the LFR.

The result of this paper is readily extended to fatigue if the evolution law of *fatigue damage* can be written

$$\frac{\mathrm{d}\omega}{\mathrm{d}N} = f\left(\frac{\Delta\sigma}{N(\omega)}\right) \tag{26}$$

Here $\Delta\sigma$ is the stress range and $\mathrm{d}\omega/\mathrm{d}N$ the rate of damage growth per unit load cycle N. The corresponding life-fraction rule is the well-known Palmgren-Miner rule, cf. Refs. [8,9]

$$\sum_{0}^{N_{\mathrm{P}}} \frac{n_i}{N_{\mathrm{f}}(\Delta\sigma_i)} = 1 \tag{27}$$

Here, n_i is the number of load cycles at the constant stress range $\Delta\sigma_i$, $N_{\mathrm{f}}(\Delta\sigma_i)$ is the number of load cycles to fracture at the same stress range, and N_{P} is the number of load blocks to fracture. A simple change of interpretation of the symbols leads to the same conclusion as for creep rupture.

We conclude this paper by noting that the result is valid for multiaxial stress states as well. For instance, introducing an effective stress according to

$$\sigma = \alpha\sigma_{\mathrm{eff}} + 3(1-\alpha)\sigma_{\mathrm{m}} \tag{28}$$

where $\sigma_{\mathrm{eff}} \equiv (3/2)s_{ij}s_{ij}$, $\sigma_{\mathrm{m}} \equiv \sigma_{kk}/3$, and $s_{ij} \equiv \sigma_{ij} - \delta_{ij}\sigma_{\mathrm{m}}$ are the von Mises effective stress, the mean stress, and the deviatoric stress, respectively. With $0 \leqslant \alpha \leqslant 1$, the damage evolution can be modeled to vary from one being driven by the hydrostatic stress ($\alpha=0$) to one being driven by the deviatoric stress ($\alpha=1$). With this effective stress, uniaxial tests can be used to determine the evolution law $f(\sigma,\omega)$. The value of α has to be determined from multiaxial tests. With this effective stress, the results of this paper are immediately applicable. With more elaborate evolution laws, the ability to conform with the LFR has to be checked. If possible, a suitable method would be to identify the effective stress associated with the specific model.

References

[1] Ashby, M. F., and Jones, D. R. H., 1980, *Engineering Materials 1—An Introduction to Their Properties and Applications*, Pergamon, Oxford.

[2] Alfredsson, K. S., and Stigh, U., 2004, "Continuum Damage Mechanics Revised—A Principle for Mechanical and Thermal Equivalence," Int. J. Solids Struct., **41**, pp. 4025–4045.

[3] Robinson, E. L., 1952, "Effect of Temperature Variation on the Long-Time Rupture Strength of Steels," Trans. ASME, **74**, pp. 777–780.

[4] Jansson, S., 1985, "Damage, Crack Growth and Rupture in Creep," dissertation, Division of Solid Mechanics, Chalmers University of Technology, Göteborg, Sweden.

[5] Kachanov, L. M., 1958, "On the Time to Failure Under Creep Conditions," Izv. Akad. Nauk SSR, Otd. Tekhn. Nauk, **8**, pp. 26–31.

[6] Odqvist, F. K. G., and Hult, J., 1961, "Some Aspects of Creep Rupture," Ark. Fys., **19**, pp. 379–382.

[7] Odqvist, F. K. G., and Hult, J., 1962, *Kriechfestigkeit Metallischer Werkstoffe*, Springer-Verlag, Berlin.

[8] Palmgren, A., 1924, "Die Lebensdauer von Kugellagern," Z. Verein Deutscher Ingenieure, **68**, pp. 339–347.

[9] Miner, M. A., 1945, "Cumulative Damage in Fatigue," ASME J. Appl. Mech., **67**, pp. A159–A164.

Circumferential Waves of Infinite Hollow Poroelastic Cylinders

M. Tajuddin

S. Ahmed Shah

Department of Mathematics,
Osmania University,
Hyderabad 500007 (A.P) India

Employing Biot's theory of wave propagation in liquid saturated porous media, the frequency equation of circumferential waves for a permeable and an impermeable surface of an infinite hollow poroelastic cylinder is derived in the presence of dissipation and then discussed. Phase velocity and attenuation are determined for different dissipations and then discussed. By ignoring liquid effects, the results of purely elastic solid are obtained as a special case. [DOI: 10.1115/1.2164513]

1 Introduction

Waves propagating around the circumference of a poroelastic circular cylinder are known as circumferential waves. The free vibrations of a solid cylinder of an elastic material are given by Love [1]. Gazis [2] studied the free vibrations of an infinite thick-walled hollow elastic cylinder. Employing Biot's theory [3] of wave propagation, Tajuddin [4] studied free vibrations of an infinite isotropic poroelastic cylinder, taking the general displacement components of a vibratory system following the analysis of Zamanaek [5]. A review of the work based on Biot's theory was given by Paria [6]. A historical formulation of porous media theories is given by Boer and Ehlers [7]. Tajuddin [8–13] studied some problems of wave phenomena characterizing Biot's theory. Wisse et al. [14] presented the experimental results of guided wave modes in porous cylinders. Chao et al. [15] studied the shock-induced borehole waves in porous formations.

In the present analysis, the frequency equation of circumferential waves of an infinite hollow poroelastic cylinder for a permeable and an impermeable surface is derived in the presence of dissipation. Plots of nondimensional frequency versus ratio of thickness to inner radius of the hollow cylinder are presented, for two different types of poroelastic materials, namely, sandstone saturated with kerosene and sandstone saturated with water. The experimental values of physical parameters are presented by Fatt [16] and Yew and Jogi [17], respectively. These material parameters are employed to study the nondimensional frequency, phase velocity, and attenuation The results are presented graphically and then discussed. The considered problem is of importance in civil engineering, ceramic industry where the frequency play an important role. The investigation is also applicable to Bio-Mechanics, wherein osseous tissue, bony elements saturated with fluid are approximated by hollow poroelastic cylinder.

2 Solution of the Problem

Consider a thick-walled homogeneous, isotropic infinite poroelastic hollow cylinder of inner and outer radii r_1 and r_2, respectively. Let (r,θ,z) be the cylindrical polar coordinates such that z-axis coincides with the axis of the cylinder. The solid displacement functions $\mathbf{u}(u,v,0)$ which can readily be evaluated from field equations of Biot [3] representing steady-state harmonic vibrations, are

$$\begin{aligned}u(r,\theta,t)=[&C_1\xi_1J_n'(\xi_1r)+C_2\xi_1Y_n'(\xi_1r)+C_3\xi_2J_n'(\xi_2r)+C_4\xi_2Y_n'(\xi_2r)\\&-nr^{-1}\{A_1J_n(\xi_3r)+B_1Y_n(\xi_3r)\}]^{\sin}_{\cos}(n\theta)e^{i\omega t}\\v(r,\theta,t)=[&nr^{-1}\{C_1J_n(\xi_1r)+C_2Y_n(\xi_1r)+C_3J_n(\xi_2r)+C_4Y_n(\xi_2r)\}\\&-\xi_3\{A_1J_n'(\xi_3r)+B_1Y_n'(\xi_3r)\}]^{\cos}_{\sin}(n\theta)e^{i\omega t},\end{aligned}\tag{1}$$

where ω is the wave frequency, n is the integer number of waves around the circumference and/or also known as angular wave number $(=kh)$, where k is the complex wave number and $h\ [=(r_2-r_1)]>0$, is the thickness of the hollow poroelastic cylinder. Also J_n and Y_n are Bessel functions of first and second kind of order n, respectively. A "prime" over a quantity denotes differentiation with respect to r. In (1) C_1, C_2, C_3, C_4, A_1, B_1 are all constants, and

$$\xi_i=\omega v_i^{-1}\ .\ (i=1,2,3)\tag{2}$$

In Eq. (2), v_1, v_2, v_3 are the dilatational wave velocities of the first and second kind and a shear wave velocity, which are all complex following Gardner [18].

By substituting the displacement functions u, v from (1) into stress displacement relations [3], the relevant solid stresses and liquid pressure are

$$\begin{aligned}\sigma_{rr}+s=[&C_1M_{11}(r)+C_2M_{12}(r)+C_3M_{13}(r)+C_4M_{14}(r)\\&+A_1M_{15}(r)+B_1M_{16}(r)]^{\sin}_{\cos}(n\theta)e^{i\omega t},\\\sigma_{r\theta}=[&C_1M_{21}(r)+C_2M_{22}(r)+C_3M_{23}(r)+C_4M_{24}(r)+A_1M_{25}(r)\\&+B_1M_{26}(r)]^{\cos}_{\sin}(n\theta)e^{i\omega t},\\s=[&C_1M_{31}(r)+C_2M_{32}(r)+C_3M_{33}(r)+C_4M_{34}(r)]^{\sin}_{\cos}(n\theta)e^{i\omega t}.\\\frac{\partial s}{\partial r}=[&C_1N_{31}(r)+C_2N_{32}(r)+C_3N_{33}(r)+C_4N_{34}(r)]^{\sin}_{\cos}(n\theta)e^{i\omega t}.\end{aligned}\tag{3}$$

The coefficients $M_{ij}(r)$ and $N_{ij}(r)$ are

$$\begin{aligned}M_{11}(r)&=-2N\xi_1r^{-1}J_n'(\xi_1r)+[2N(n^2r^{-2}-\xi_1^2)+(Q+R)\delta_1^2\xi_1^2-(A\\&\quad+Q)\xi_1^2]J_n(\xi_1r),\\M_{15}(r)&=-2Nn\xi_3r^{-1}J_n'(\xi_3r)+2Nnr^{-2}J_n(\xi_3r),\\M_{21}(r)&=-2Nn\xi_1r^{-1}J_n'(\xi_1r)-Nnr^{-2}J_n(\xi_1r),\\M_{25}(r)&=r^{-1}\xi_3J_n'(\xi_3r)+(\xi_3^2-2n^2r^{-2})J_n(\xi_3r),\\M_{31}(r)&=(R\delta_1^2-Q)\xi_1^2J_n(\xi_1r),\quad N_{31}(r)=(R\delta_1^2-Q)\xi_1^3J_n'(\xi_1r),\end{aligned}\tag{4}$$

and $M_{12}(r)$, $M_{16}(r)$, $M_{22}(r)$, $M_{26}(r)$ are similar to $M_{11}(r)$, $M_{15}(r)$, $M_{21}(r)$, $M_{25}(r)$ with J_n, J_n' replaced by Y_n, Y_n', respectively, $M_{14}(r)$, $M_{24}(r)$ are similar to $M_{11}(r)$, $M_{21}(r)$ with ξ_1, δ_1, J_n, J_n' replaced by ξ_2, δ_2, Y_n, Y_n', respectively, and $M_{13}(r)$, $M_{23}(r)$, $M_{33}(r)$, $N_{33}(r)$ are similar to $M_{11}(r)$, $M_{21}(r)$, $M_{31}(r)$, $N_{31}(r)$ with ξ_1, δ_1 replaced by ξ_2, δ_2, respectively. Also, $M_{32}(r)$ is similar to $M_{31}(r)$ with J_n replaced by Y_n, $N_{32}(r)$ is similar to $N_{31}(r)$ with J_n' replaced by Y_n', $M_{34}(r)$ is similar to $M_{31}(r)$ with ξ_1, δ_1, and J_n replaced by ξ_2, δ_2, and Y_n, respectively, and $N_{34}(r)$ is similar to $N_{31}(r)$ with ξ_1, δ_1, J_n' replaced by ξ_2, δ_2, Y_n' respectively.

In Eqs. (4)

$$\delta_i^2=(RK_{12}-QK_{22})^{-1}[RK_{11}-QK_{12}-v_i^{-2}(PR-Q)^2],\quad(i=1,2),$$

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 24, 2004; final manuscript received October 26, 2005. Review conducted by E. M. Arruda.

$$P = A + 2N, \quad K_{11} = \rho_{11} - ib\omega^{-1}, \quad K_{12} = \rho_{12} + ib\omega^{-1},$$

$$K_{22} = \rho_{22} - ib\omega^{-1}. \tag{5}$$

The recursion relations of Bessel functions have been employed using Abramowitz and Stegun [19].

3 Frequency Equation

The boundary conditions for free vibrations of a permeable surface are

$$\sigma_{rr} + s = 0, \quad \sigma_{r\theta} = 0, \quad s = 0 \text{ at } r = r_1 \text{ and } r_2. \tag{6}$$

Equations (3) and (6) together give six homogeneous equations for the six constants C_1, C_2, C_3, C_4, A_1, and B_1. By eliminating these constants, the frequency equation for a permeable surface is

$$D = 0, \tag{7}$$

where the determinant D is

$$D = \begin{vmatrix} M_{11}(r_1) & M_{12}(r_1) & M_{13}(r_1) & M_{14}(r_1) & M_{15}(r_1) & M_{16}(r_1) \\ M_{21}(r_1) & M_{22}(r_1) & M_{23}(r_1) & M_{24}(r_1) & M_{25}(r_1) & M_{26}(r_1) \\ M_{31}(r_1) & M_{32}(r_1) & M_{33}(r_1) & M_{34}(r_1) & 0 & 0 \\ M_{11}(r_2) & M_{12}(r_2) & M_{13}(r_2) & M_{14}(r_2) & M_{15}(r_2) & M_{16}(r_2) \\ M_{21}(r_2) & M_{22}(r_2) & M_{23}(r_2) & M_{24}(r_2) & M_{25}(r_2) & M_{26}(r_2) \\ M_{31}(r_2) & M_{32}(r_2) & M_{33}(r_2) & M_{34}(r_2) & 0 & 0 \end{vmatrix}.$$

In case of an impermeable surface, the boundary conditions for free vibrations are

$$\sigma_{rr} + s = 0, \quad \sigma_{r\theta} = 0, \quad \frac{\partial s}{\partial r} = 0 \text{ at } r = r_1 \text{ and } r_2.$$

Arguing on similar lines of a permeable surface, the frequency equation of an impermeable surface is similar to Eq. (7) with the third and sixth rows replaced by $M_{3j}(r) = N_{3j}(r)$, $(j = 1, 2, 3, 4)$. These are the frequency equations of a thick-walled hollow poroelastic cylinder, examined by Malla Reddy and Tajuddin [12], and these were not studied for small values of the ratio of the thickness of the hollow poroelastic cylinder to that of inner radius of the cylinder. Accordingly this study is also warranted.

For small values of hr_1^{-1}, the waves exist along the circumference of inner cylinder and are known as circumferential waves. Considering the determinant D given in (7) as a function of $\xi_i r_1$ and hr_1^{-1} $(i = 1, 2, 3)$, and following the analysis of Gazis [2], one obtains the frequency equation of circumferential waves of a permeable surface to be identical to (7) with the fourth, fifth and sixth rows replaced by $M_{ij}(r_2) = M'_{ij}(r_1)$ $(i = 1, 2,\ J = 1, 2, 3, \ldots, 6,\ i = 3,\ j = 1, 2, 3, 4)$. The frequency equation of circumferential waves of an impermeable surface is similar to Eq. (7) with the third and sixth row replaced by $M_{ij}(r_1) = N_{ij}(r_1)$, $M_{ij}(r_2) = N'_{ij}(r_1)$ $(i = 3,\ j = 1, 2, 3, 4)$, fourth and fifth rows by $M_{ij}(r_2) = M'_{ij}(r_1)$ $(i = 1, 2,\ j = 1, 2, \ldots, 6)$ respectively.

Employing Eqs. (4) into the frequency equation of circumferential waves of a permeable surface, after a lengthy calculation, gives

$$\xi_3 r_1 = 2\{(n^2 + 1)[1 - RN(PR - Q^2)^{-1}]\}^{1/2}. \tag{8}$$

Arguing on similar lines, the frequency equation of an impermeable surface reduces to

$$\xi_3 r_1 = 2\{(n^2 + 1)(1 + N(PR - Q^2)[PR(Q + R)K_{12}K_{22}^{-1} - P(PR - Q^2) - QR(P + Q)]^{-1})\}^{1/2}, \tag{9}$$

where ξ_3 is defined in (2). Thus the frequency Eqs. (8) and (9) are complex valued velocity equations.

By ignoring the liquid effects in Eq. (8), that is, when $\rho_{12} \to 0$, $\rho_{22} \to 0$, $N \to \mu$, and $(A\text{-}Q^2R^{-1}) \to \lambda$ gives

$$\xi_3 r_1 = 2[(n^2 + 1)(1 - d)]^{1/2},$$

where $d = \mu(\lambda + 2\mu)^{-1}$; λ and μ are the familiar elastic Lame' constants. This is the result for a purely elastic solid derived by Gazis (Eq. (63) [2]). Equation (9) has no counterpart in a purely elastic solid.

To analyze the parameters of physical interest, it is convenient to introduce the following nondimensional variables,

$$a_1 = PH^{-1}, \quad a_2 = QH^{-1}, \quad a_3 = RH^{-1}, \quad a_4 = NH^{-1},$$

$$m_{11} = \rho_{11}\rho^{-1}, \quad m_{12} = \rho_{12}\rho^{-1}, \quad m_{22} = \rho_{22}\rho^{-1}, \quad b_1 = bh(c_0\rho)^{-1},$$

where $\rho = \rho_{11} + 2\rho_{12} + \rho_{22}$, $H = P + 2Q + R$, $c_0^2 = N\rho^{-1}$. Let $g = r_2 r_1^{-1}$, so that $hr_1^{-1} = (g - 1)$, and $\Omega = \omega h c_0^{-1}$. (10)

In Eq. (10), Ω, b_1, m_{11}, m_{12}, m_{22} are nondimensional frequency, dissipation, and mass coefficients, while a_1, a_2, a_3, a_4 are nondimensional poroelastic constants.

Introducing nondimensional quantities given in (10) into Eq. (8), the frequency equation of circumferential waves for a permeable surface gives

$$\frac{\Omega^2}{(g - 1)^2}\left[\frac{\Omega^2 m_{22}(m_{11}m_{22} - m_{12}^2) + b_1^2}{\Omega^2 m_{22}^2 + b_1^2}\right] = 4(n^2 + 1)[1 - a_3 a_4(a_1 a_3 - a_2^2)^{-1}]. \tag{11}$$

Similarly, the frequency equation of circumferential waves for an impermeable surface is

$$\frac{\Omega^2}{(g - 1)^2}\left[\frac{\Omega^2 m_{22}(m_{11}m_{22} - m_{12}^2) + b_1^2}{\Omega^2 m_{22}^2 + b_1^2}\right] = 4(n^2 + 1)\left\{1 + \frac{[T_1T_2(\Omega^2 m_{11}m_{22} - b_1^2) + T_1T_3(\Omega^2 m_{22}^2 + b_1^2)][\Omega^2 m_{22}^2 + b_1^2]}{[T_2(\Omega^2 m_{11}m_{22} - b_1^2) + T_3(\Omega^2 m_{22}^2 + b_1^2)]^2 + [T_2 b_1 \Omega(m_{12} + m_{22})]^2}\right\}. \tag{12}$$

In Eq. (12), T_1, T_2, and T_3 are defined as

$$T_1 = a_4(a_1 a_3 - a_2^2), \quad T_2 = a_1 a_2 a_3 + a_1 a_3^2,$$

$$T_3 = a_1 a_2^2 - a_1^2 a_3 - a_1 a_2 a_3 - a_2^2 a_3.$$

4 Phase Velocity and Attenuation

Let $n = kh$, where k is the wave number and h is the thickness. Due to the dissipative nature of the medium, the wave number is complex, that is, $k = k_r + ik_i$. The waves generated obey the diffusion process, and thereby become attenuated. Setting $n = kh$ in the frequency equation (11) the phase velocity is $C_p = \mathrm{Re}(\omega k^{-1})$ and the attenuation is $x_h = |k_i|^{-1}$. Thus the nondimensional phase velocity and attenuation of circumferential waves of a permeable surface respectively, give

$$c_p c_0^{-1} = 2\sqrt{2}\Omega(B_1 + B_2)^{-1/2}, \quad x_h h^{-1} = 2\sqrt{2}(B_1 - B_2)^{-1/2}, \tag{13}$$

where the terms B_1 and B_2 are defined by

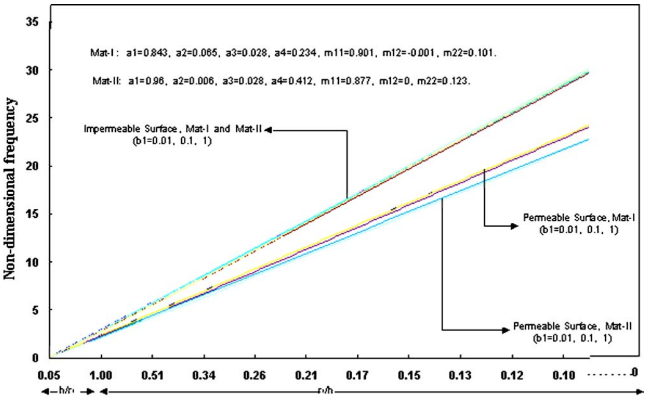


Fig. 1 Frequency as a function of h/r_1 for $n=1$

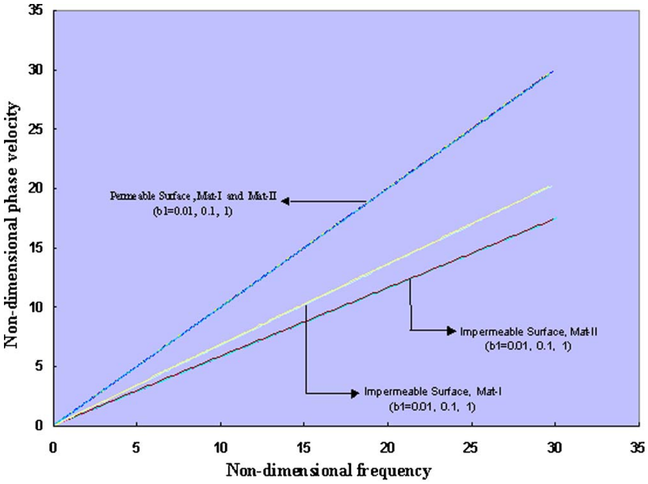


Fig. 2 Phase velocity as a function of frequency

$$B_1=\left[\frac{\Omega^4(E_r^2+E_i^2)(a_1a_3-a_2^2)^2}{(g-1)^4(a_1a_3-a_2^2-a_3a_4)^2}-\frac{8\Omega^2E_r(a_1a_3-a_2^2)}{(g-1)^2(a_1a_3-a_2^2-a_3a_4)}+16\right]^{1/2},$$

$$B_2=\left[\frac{\Omega^2E_r(a_1a_3-a_2^2)}{(g-1)^2(a_1a_3-a_2^2-a_3a_4)}-4\right]. \qquad (14)$$

In Eqs. (14), E_r and E_i are

$$E_r=\frac{\Omega^2m_{22}(m_{11}m_{22}-m_{12}^2)+b_1^2}{\Omega^2m_{22}^2+b_1^2},\quad E_i=\frac{b_1\Omega(m_{12}+m_{22})^2}{\Omega^2m_{22}^2+b_1^2}. \qquad (15)$$

Arguing on similar lines, the nondimensional phase velocity and attenuation of circumferential waves of an impermeable surface respectively, give

$$c_pc_0^{-1}=2\sqrt{2}\Omega(B_3+B_4)^{-1/2},\quad x_hh^{-1}=2\sqrt{2}(B_3-B_4)^{-1/2}, \qquad (16)$$

where B_3 and B_4 are

$$B_3=\left\{\frac{\Omega^4(E_r^2+E_i^2)[(G_3^2+G_1G_3G_4)^2+(G_2G_3G_4)^2]}{(g-1)^4[(G_3+G_1G_4)^2+(G_2G_4)^2]^2}-\frac{8\Omega^2[(G_3^2+G_1G_3G_4)E_r+(G_2G_3G_4)E_i]}{(g-1)^2[(G_3+G_1G_4)^2+(G_2G_4)^2]}+16\right\}^{1/2},$$

$$B_4=\left\{\frac{\Omega^2[(G_3^2+G_1G_3G_4)E_r+(G_2G_3G_4)E_i]}{(g-1)^2[(G_3+G_1G_4)^2+(G_2G_4)^2]}-4\right\}. \qquad (17)$$

In Eqs. (17), G_i $(i=1,2,3,4)$ are

$$G_1=T_1T_2(\Omega^2m_{12}m_{22}-b_1^2)+T_1T_3(\Omega^2m_{22}^2+b_1^2),$$

$$G_2=T_1T_2\Omega b_1(m_{12}+m_{22}),$$

$$G_3=[T_2(\Omega^2m_{12}m_{22}-b_1^2)+T_3(\Omega^2m_{22}^2+b_1^2)]^2+[T_2b_1\Omega(m_{12}+m_{22})]^2,\quad G_4=\Omega^2m_{22}^2+b_1^2. \qquad (18)$$

5 Results and Discussion

The physical parameters of sandstone saturated with kerosene (material-I) and with water (material-II) are respectively, presented by Fatt [16] and Yew and Jogi [17]. For a given material, the frequency Eqs. (11) and (12) of permeable and impermeable surfaces with different dissipations (b_1) give a relation between non-dimensional frequency (Ω) and the ratio of thickness to inner radius (hr_1^{-1}). The dissipation parameters chosen are 0.01, 0.1, and 1, respectively. Frequency Ω vs hr_1^{-1} is computed for both the materials and presented in Fig. 1 taking a wide spectrum of values of hr_1^{-1} for $n=1$. $n=1$ represents the first flexural mode. Figure 1 shows the frequency Ω versus hr_1^{-1} is linear and that the frequency increases as n increases for the referred materials. The frequency is therefore higher for higher flexural modes. The frequency of the permeable surface for material-I is higher than that of material-II, while the frequency of an impermeable surface is the same for both the materials. Moreover, the frequency for an impermeable surface is higher than that of a permeable surface.

Phase velocity and attenuation versus frequency is computed each for permeable and impermeable surfaces and exhibited in Figs. 2 and 3, respectively. Figure 2 shows that phase velocity is again linear irrespective of dissipations. The phase velocity is higher for a permeable surface than for an impermeable surface. Phase velocity of an impermeable surface in material-I is higher than that of material-II while for a permeable surface, it is same for both the materials. The attenuation exhibited in Fig. 3 is similar in phenomenon to phase velocity but nonlinear for a permeable surface and different for different dissipations. Attenuation in material-I is less than that in material-II for a permeable surface but linear and similar for both the materials for an impermeable surface.

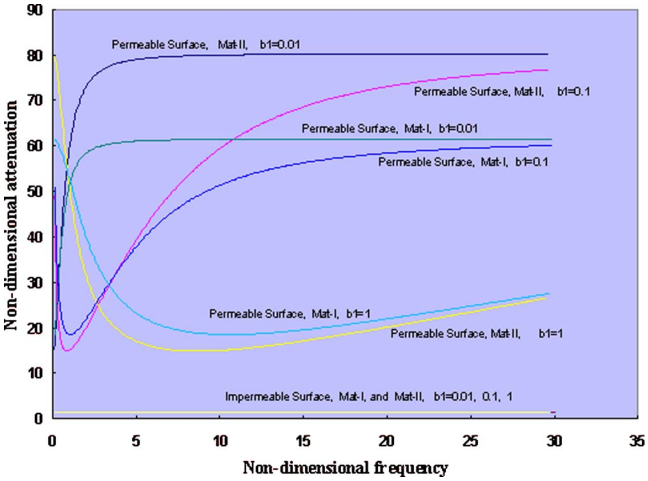


Fig. 3 Attenuation as a function of frequency

Acknowledgment

The authors are thankful to Dr. Ellen M. Arruda, Associate Editor, and reviewers for stimulating suggestions.

References

[1] Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, Dover, New York.

[2] Gazis, D. C., 1958, “Exact Analysis of Plane-Strain Vibrations of Thick-Walled Hollow Cylinders,” J. Acoust. Soc. Am., **30**, pp. 786–794.

[3] Biot, M. A., 1956, “Theory of Propagation of Elastic Waves in Fluid-Saturated Porous Solid,” J. Acoust. Soc. Am., **28**, pp. 168–178.

[4] Tajuddin, M., 1978, “Vibrations in Fluid-Saturated Porous Elastic Cylinders,” Rev. Roum. Sci. Tech., Mec. Appl., **23**, pp. 371–379.

[5] Zamanaek, J., 1971, “Experimental and Theoretical Investigation of Elastic Wave Propagation in a Cylinder,” J. Acoust. Soc. Am., **51**, pp. 265–283.

[6] Paria, G., 1963, “Flow of Fluid Through Porous Deformable Solids,” Appl. Mech. Rev., **16**, pp. 421–423.

[7] Boer, R. de, and Ehlers, W., 1988, “A Historical Review of the Formulation of Porous Media Theories,” Acta Mech., **79**, pp. 1–8.

[8] Tajuddin, M., and Sarma, K. S., 1980, “Torsional Vibrations of Poroelastic Cylinders,” ASME J. Appl. Mech., **47**, pp. 214–216.

[9] Tajuddin, M., 1982, “Rayleigh Waves on a Concave Cylindrical Poroelastic Surface,” Indian J. Pure Appl. Math., **13**, pp. 1278–1282.

[10] Tajuddin, M., 1984, “Rayleigh Waves in a Poroelastic Half-Space,” J. Acoust. Soc. Am., **75**, pp. 682–684.

[11] Tajuddin, M., and Moiz, A. A., 1984, “Rayleigh Waves on a Convex Cylindrical Poroelastic Surface,” J. Acoust. Soc. Am., **76**, pp. 1252–1254.

[12] Malla Reddy, P., and Tajuddin, M., 2000, “Exact Analysis of the Plane-Strain Vibrations of Thick Walled Hollow Poroelastic Cylinders,” Int. J. Solids Struct., **37**, pp. 3439–3456.

[13] Malla Reddy, P., and Tajuddin, M., 2003, “Edge Waves in Poroelastic Plate Under Plane-Stress Conditions,” J. Acoust. Soc. Am., **114**, pp. 185–193.

[14] Wisse, C. J., Smeulders, D. M. J., van Dongen, M. E. H., and Chao, G., 2002, “Guided Wave Modes in Porous Cylinders: Experimental Results,” J. Acoust. Soc. Am., **112**, pp. 890–895.

[15] Chao, G., Smeulders, D. M. J., and van Dongen, M. E. H., 2004, “Shock-Induced Borehole Waves in Porous Formations: Theory and Experiments,” J. Acoust. Soc. Am., **116**, pp. 693–702.

[16] Fatt, I., 1959, “The Biot-Willis Elastic Coefficients for a Sandstone,” J. Appl. Mech., **26**, pp. 296–297.

[17] Yew, C. H., and Jogi, P. N., 1976, “Study of Wave Motions in Fluid-Saturated Porous Rocks,” J. Acoust. Soc. Am., **60**, pp. 2–8.

[18] Gardner, G. H. F., 1962, “Extensional Waves in Fluid Saturated Porous Cylinders,” J. Acoust. Soc. Am., **34**, pp. 36–40.

[19] Abramowitz, A., and Stegun, I. A., 1965, *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, DC.

Extending Koiter's Simplified Equations to Shells of Arbitrary Closed Cross Section

James G. Simmonds
Fellow ASME
Department of Civil Engineering,
University of Virginia,
Charlottesville, VA 22904-4742

The techniques used by Koiter in 1968 to derive a simplified set of linear equilibrium equations for an elastically isotropic circular cylindrical shell in terms of displacements and the associated pointwise error estimate engendered in Love's uncoupled strain-energy density are here extended to derive analogous simplified equilibrium equations and an error estimate for elastically isotropic cylindrical shells of arbitrary closed cross section.
[DOI: 10.1115/1.2165236]

To Pierre Ladevèze, mécanicien extraordinaire, with respect and affection, on the occasion of his 60th birthday.

1 Introduction

In a report [1] of 1968, unpublished in the archival literature, Koiter presented what he called the "modified, simplest possible accurate linear theory of thin circular cylindrical shells." Derived from the Sanders-Koiter theory [2,3], Koiter's equations of equilibrium with body forces (which, by d'Alembert's Principle, can immediately be converted into equations of motion) are remarkable in several ways. First, they are much simpler than the analogous equations of Love [4] or Flügge [5]—see [6] for a comparison—but of equal accuracy. In particular, the two equations of tangential equilibrium, expressed in terms u and v, the midsurface axial and circumferential displacements, respectively, are the *same* as those of the simplified theory of Donnell [7]. Moreover, if the tangential equilibrium equations are used to differentially eliminate u and v from the normal equilibrium equation, then the resulting eighth-order partial differential equation for the outward midsurface normal displacement w is identical to that derived in a less rigorous manner by Morley [8]. Notably, this equation is *free of Poisson's ratio* ν, except as it appears in the small dimensionless parameter $h^2/12(1-\nu^2)R^2$, where h is the constant thickness of the shell and R is the midsurface radius. Published derivations of Koiter's simplified equations may be found in a paper by Mangelsdorf [9] and in a book by Niordson [10].

In the present Note we extend Koiter's simplifications to elastically isotropic cylindrical shells of arbitrary closed cross section. The main impediment is that the dimensionless cross-sectional curvature κ is no longer a constant. Following Koiter's approach, we add small terms to Love's uncoupled quadratic strain-energy density V to put as many terms as possible in the bending strain-energy density into divergence form. Such terms have no influence on the associated Euler-Lagrange equations. If V^* is the modified (coupled) strain-energy density resulting from the addition of these small terms to V,then, using basic inequalities, we show that

$$|V_* - V| = (h/R)\{C_1|\kappa|[1 + O(|\kappa|h/R)]V + C_2|\kappa^\bullet|\sqrt{\lambda V\int_0^\lambda [1 + O(|\kappa|h/R)]Vdx}\} \tag{1}$$

where C_1 and C_2 are $O(1)$ constants independent of h/R, $2\pi R$ is the midsurface circumference, $R\lambda$ is the length of the cylindrical midsurface, and $()^\bullet \equiv \partial()/\partial y$, where Ry is circumferential distance. Later, we shall also use the abbreviation $()' \equiv \partial()/\partial x$, where Rx is distance along a generator of the midsurface.

For a *circular* cylindrical shell (κ=1) all terms in the bending strain-energy density involving u and v can be cast into divergence form, but for nonconstant κ, some terms involving the circumferential displacement remain. Nevertheless, the simplifications of the unmodified Sanders-Koiter first-approximation shell equations are substantial.

Finally, we note that our final equilibrium equations apply without change to nonclosed sections. Only the boundary terms in the associated energy functional and one integral inequality have to be modified.

2 The Governing Equations

With $\{u,v\}=R\{\bar{u},\bar{v}\}$ and $R^{-1}\kappa$ denoting the curvature of the cross section of the midsurface, the dimensionless extensional and bending strains of the Sanders-Koiter theory can be expressed in terms of the dimensionless midsurface displacements and their derivatives as

$$\{\gamma_x, \gamma_y, 2\gamma\} = \{\bar{u}', \bar{v}^\bullet + \kappa\bar{w}, \bar{u}^\bullet + \bar{v}'\}$$

$$\{k_x, k_y, k\} = \left\{\varphi_x', \varphi_y^\bullet, \varphi_y' + \tfrac{1}{2}\kappa\gamma\right\}, \tag{2}$$

where the (dimensionless) rotations are given by

$$\{\varphi_x, \varphi_y\} = \{\bar{w}', \bar{w}^\bullet - \kappa\bar{v}\}. \tag{3}$$

Love's uncoupled strain-energy density can be written as $V=(1-\nu^2)^{-1}Eh\bar{V}$, where E is Young's modulus and

$$\begin{aligned}\bar{V} &= \tfrac{1}{2}\{(\gamma_x + \gamma_y)^2 + 2(1-\nu)(\gamma^2 - \gamma_x\gamma_y) \\ &\quad + \varepsilon^2[(k_x + k_y)^2 + 2(1-\nu)(k^2 - k_x k_y)]\} \\ &\equiv \bar{V}_e + \varepsilon^2\bar{V}_b.\end{aligned} \tag{4}$$

Here, $\varepsilon^2=h^2/12R^2$ and "e" and "b" are mnemonics for "extension" and "bending."

3 Modifications of the Bending Invariants

Noting (2)–(4), we define the *modified first bending invariant* as

$$I_* \equiv I + \kappa\gamma_y = (\Delta + \kappa^2)\bar{w} - \kappa^\bullet\bar{v}, \tag{5}$$

where $I \equiv k_x + k_y$ and Δ is the surface Laplacian. Thus,

$$I_*^2 - I^2 = \kappa(2\gamma_y I + \kappa\gamma_y^2). \tag{6}$$

To define the *second modified bending invariant* II_*, first note that

$$(k - \tfrac{1}{2}\kappa\gamma)^2 - k_x k_y = II - \kappa\gamma k + \tfrac{1}{4}\kappa^2\gamma^2, \tag{7}$$

where $II \equiv k^2 - k_x k_y$. Furthermore, because

$$k - \tfrac{1}{2}\kappa\gamma = \varphi_y' = \varphi_x^\bullet - \kappa\bar{v}', \tag{8}$$

by (2) and (3), we may write

Contributed by the Applied Mechanics Division of ASME for publication in the ASME Journal of Applied Mechanics. Manuscript received July 6, 2005; final manuscript received November 6, 2005. Review conducted by N. Triantafyllidis. Discussion the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5070 and will be accepted until four months after final publication in the paper itself in the ASME Journal of Applied Mechanics.

$$
\begin{aligned}
(k-\tfrac{1}{2}\kappa\gamma)^2-k_xk_y &= (\varphi_x^{\bullet}-\kappa\bar{v}')\varphi_y'-\varphi_x'\varphi_y^{\bullet} \\
&= [(\varphi_x^{\bullet}-\kappa\bar{v}')\varphi_y]'-(\varphi_x'\varphi_y)^{\bullet}+\kappa\bar{v}''\varphi_y. \qquad (9)
\end{aligned}
$$

By the divergence theorem, the integral over the midsurface of the first two terms in the second line can be replaced by line integrals around the boundary of the cylindrical midsurface.

To deal with the last term in the second line of (9), use (2) and note that $\bar{v}''=2\gamma'-\gamma_x^{\bullet}$. Thus,

$$
\kappa\bar{v}''\varphi_y=(2\kappa\gamma\varphi_y)'-(\kappa\gamma_x\varphi_y)^{\bullet}-2\kappa\gamma k+\kappa^2\gamma^2+\kappa\gamma_xk_y+\kappa^{\bullet}\gamma_x\varphi_y. \qquad (10)
$$

To further reduce the last four terms in (10), note by (2) that

$$
\begin{aligned}
\varphi_y(x,y)-\varphi_y(0,y) &= \int_0^x \varphi_y'(\xi,y)d\xi \\
&= \int_0^x \left(k-\frac{1}{2}\kappa\gamma\right)d\xi. \qquad (11)
\end{aligned}
$$

Thus, because $\gamma_x=\bar{u}'$,

$$
\kappa^{\bullet}\gamma_x\varphi_y=[\kappa^{\bullet}\bar{u}\varphi_y(0,y)]'+\kappa^{\bullet}\gamma_x\int_0^x\left(k-\frac{1}{2}\kappa\gamma\right)d\xi \qquad (12)
$$

or, more symmetrically,

$$
\kappa^{\bullet}\gamma_x\varphi_y=[\kappa^{\bullet}\bar{u}f(y)]'+\mathcal{E}, \qquad (13)
$$

where

$$
f\equiv\tfrac{1}{2}[\varphi_y(0,y)+\varphi_y(\lambda,y)] \qquad (14)
$$

and

$$
\mathcal{E}\equiv\frac{1}{2}\kappa^{\bullet}\gamma_x\left[\left(\int_0^x-\int_x^\lambda\right)\left(k-\frac{1}{2}\kappa\gamma\right)d\xi\right] \qquad (15)
$$

is an additional error term produced by nonconstant curvature.

Substituting (13) and (10) into (9) and the resulting equation into (7), we are led to define the second modified bending invariant as

$$
\begin{aligned}
II_* &\equiv II+\kappa\left(\gamma k-\gamma_xk_y-\tfrac{3}{4}\kappa\gamma^2\right)-\mathcal{E} \\
&= [(\varphi_x^{\bullet}+\kappa\bar{u}^{\bullet})\varphi_y+\kappa^{\bullet}\bar{u}f(y)]'-[(\varphi_x'+\kappa\gamma_x)\varphi_y]^{\bullet}. \qquad (16)
\end{aligned}
$$

4 Equations of Equilibrium in Terms of Displacements

With

$$
\bar{V}_b^*\equiv\tfrac{1}{2}[I_*^2+2(1-\nu)II_*] \qquad (17)
$$

and $(1-\nu^2)^{-1}E(h/R)\{\bar{p}_x,\bar{p}_y,\bar{p}\}$ denoting the axial, circumferential, and outward normal components of the external surface load, consider the dimensionless energy functional

$$
\begin{aligned}
\bar{\Pi}_*[\bar{u},\bar{v},\bar{w}] \equiv{}& \int_0^{2\pi}\int_0^\lambda\left[\bar{V}_e+\frac{1}{2}\varepsilon^2I_*^2-(\bar{p}_x\bar{u}+\bar{p}_y\bar{v}+\bar{p}\bar{w})\right]dxdy \\
&+\int_0^{2\pi}\{\varepsilon^2(1-\nu)[(\varphi_x^{\bullet}+\kappa\bar{u}^{\bullet})\varphi_y+\kappa^{\bullet}\bar{u}f(y)] \\
&-(\hat{N}_x\bar{u}+\hat{S}_x\bar{v}+\hat{Q}_x\bar{w}+\hat{M}_x\varphi_x)\}_0^\lambda dy, \qquad (18)
\end{aligned}
$$

written here for prescribed end loads $(1-\nu^2)^{-1}Eh\{\hat{N}_x, \hat{S}_x,\hat{Q}_x,\hat{M}_x\}_0^\lambda$. The associated Euler-Lagrange equations are

$$
\bar{u}''+\tfrac{1}{2}(1-\nu)\bar{u}^{\bullet\bullet}+\tfrac{1}{2}(1+\nu)\bar{v}'^{\bullet}+\nu\kappa\bar{w}'+\bar{p}_x=0 \qquad (19)
$$

$$
\tfrac{1}{2}(1+\nu)\bar{u}'^{\bullet}+\tfrac{1}{2}(1-\nu)\bar{v}''+\bar{v}^{\bullet\bullet}+(\kappa\bar{w})^{\bullet}
$$

$$
\varepsilon^2\kappa^{\bullet}[(\Delta+\kappa^2)\bar{w}-\kappa^{\bullet}\bar{v}]+\bar{p}_y=0 \qquad (20)
$$

$$
\kappa(\nu\bar{u}'+\bar{v}^{\bullet}+\kappa\bar{w})+\varepsilon^2(\Delta+\kappa^2)[(\Delta+\kappa^2)\bar{w}-\kappa^{\bullet}\bar{v}]-\bar{p}=0. \qquad (21)
$$

These equations reduce to Koiter's [1] for a *circular* cylindrical shell ($\kappa=1$).

5 Error Estimates

From (4), (6), and (16),

$$
\begin{aligned}
|\bar{V}_*-\bar{V}|=\varepsilon^2\left|\tfrac{1}{2}(I_*^2-I^2)+(1-\nu)(II_*-II)\right| &\leqslant |\kappa|\varepsilon Q \\
&+\varepsilon^2\{\kappa^2[\tfrac{1}{2}\gamma_y^2+\tfrac{3}{4}(1-\nu)\gamma^2]+(1-\nu)|\mathcal{E}|\}, \qquad (22)
\end{aligned}
$$

where

$$
Q\equiv\varepsilon[(k_x+k_y)\gamma_y+(1-\nu)(\gamma k-\gamma_xk_y)] \qquad (23)
$$

is a quadratic form. Let

$$
\mu_{\max}=\max(Q/\bar{V}),\quad \bar{V}\neq 0, \qquad (24)
$$

taken over all strains. The associated eigenvalue problem involves a 6×6 determinant for computing $\mu_{\max}$ that, because γ and k uncouple from the other four strains, reduces to evaluating a 4×4 and a 2×2 determinant. The six eigenvalues are readily computed by hand, the largest being

$$
\mu_{\max}=\sqrt{\frac{3-2\nu+\sqrt{5-4\nu}}{2(1-\nu^2)}}\leqslant 1.618\quad \text{if } 0\leqslant\nu\leqslant\tfrac{1}{2}. \qquad (25)
$$

(This agrees with Koiter's upper bound [1].) Furthermore, because

$$
\bar{V}_e=\tfrac{1}{2}[(1-\nu^2)\gamma_y^2+(\gamma_x+\nu\gamma_y)^2+2(1-\nu)\gamma^2], \qquad (26)
$$

the term in brackets in the second line of (22) is bounded by $(1-\nu^2)^{-1}\bar{V}$.

Finally, we turn to the last term in (22). From (15),

$$
(1-\nu)\varepsilon^2|\mathcal{E}|\leqslant\tfrac{1}{2}(1-\nu)\varepsilon^2|\kappa^{\bullet}||\gamma_x|\int_0^\lambda|k-\tfrac{1}{2}\kappa\gamma|dx. \qquad (27)
$$

It is not difficult to show that

$$
(1-\nu)\varepsilon^2\left(k^2-\kappa\gamma k+\tfrac{1}{4}\kappa^2\gamma^2\right)\leqslant\left(1+\tfrac{1}{2}\varepsilon|\kappa|\right)^2\bar{V}. \qquad (28)
$$

Hence, reversing the roles of γ_x and γ_y in (26) and using the Schwarz inequality, we obtain

$$
(1-\nu)\varepsilon^2|\mathcal{E}|\leqslant\varepsilon|\kappa^{\bullet}|[2(1-\nu^2)]^{-1/2}\left[\lambda\bar{V}\int_0^\lambda\left(1+\frac{1}{2}\varepsilon|\kappa|\right)^2\bar{V}dx\right]^{1/2} \qquad (29)
$$

Substituting (24) and (29) into (22) and setting $\varepsilon=2h/\sqrt{3}R$, $V=(1-\nu^2)^{-1}Eh\bar{V}$, we obtain the inequality (1) with

$$
C_1=\mu_{\max}/2\sqrt{3}\ \text{and}\ C_2=1/\sqrt{3(1-\nu^2)}. \qquad (30)
$$

6 Conclusion

If $\kappa=1$, the inequality (1) reduces to Koiter's pointwise result [1]. If $\kappa\neq 1$, our error estimate is only partially pointwise and the dimensionless energy functional (18) associated with the simplifications proposed here contains the somewhat unusual boundary term $\frac{1}{2}(1-\nu)\varepsilon^2\int_0^{2\pi}\kappa^{\bullet}\bar{u}[\varphi_y(0,y)+\varphi_y(\lambda,y)]_0^\lambda dy$. On the other hand, if the term $\kappa^{\bullet}\gamma_x\varphi_y$ in (10) is *not* replaced by a divergence term plus the error term $\mathcal{E}$, as in (13), then the error estimate (1) is identical to Koiter's, the above-mentioned boundary term does not appear in the dimensionless energy functional, and the complicating terms $(1-\nu)\varepsilon^2\left\{\kappa^{\bullet}(\bar{w}^{\bullet}-\kappa\bar{v})',\tfrac{1}{2}[(\kappa^2)^{\bullet}\bar{u}]',-(\kappa^{\bullet}\bar{u}')^{\bullet}\right\}$must be added to the equilibrium Eqs. (19)–(21), respectively.

References

[1] Koiter, W. T., 1968, "Summary of Equations for Modified, Simplest Possible Accurate Linear Theory of Thin Circular Cylindrical Shells," Report No. 442, Lab. Tech. Mech., T. H. Delft.

[2] Sanders, J. L., Jr., 1959, "An Improved First-Approximation Theory for Thin Shells," NASA Report No. 24.

[3] Koiter, W. T., 1960, "A Consistent First Approximation in the General Theory of Thin Elastic Shells," The Theory of Thin Elastic Shells, *Proceedings of the IUTAM Symposium*, Delft, 1959, W. T. Koiter, ed., North-Holland, Amsterdam.

[4] Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., Dover, New York, p. 573.

[5] Flügge, W., 1973, *Stresses in Shells*, 2nd ed., Springer-Verlag, New York, p. 215.

[6] Simmonds, J. G., 1966, "A Set of Simple, Accurate Equations for Circular Cylindrical Elastic Shells," Int. J. Solids Struct. **2**, pp. 525–541.

[7] Donnell, L. H., 1933, "Stability of Thin-Walled Tubes Under Torsion," NACA TR 479.

[8] Morley, L. S. D., 1959, "An Improvement on Donnell's Approximation for Thin-Walled Circular Cylinders," Q. J. Mech. Appl. Math. **12**, pp. 89–99.

[9] Mangelsdorf, C. P., 1973, "Morley-Koiter Equations for Thin-Walled Circular Cylindrical Shells. Part 1. General Solutions for Symmetrical Shells of Uniform Thickness," J. Appl. Mech. **40**, pp. 961–965.

[10] Niordson, F. I., 1985, *Shell Theory*, North-Holland, Amsterdam, Chap. 11.