ASME
SETTING THE STANDARD



**Transactions of the ASME, Journal of Applied Mechanics** (ISSN 0021-8936) is published bimonthly (Jan., Mar., May, July, Sept., Nov.) by The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300. **CHANGES OF ADDRESS** must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address. **STATEMENT from By-Laws.** The Society shall not be responsible for statements or opinions advanced in papers or printed in its publications (B7.1, Para. 3).


# Journal of Applied Mechanics

Published Bimonthly by ASME

VOLUME 76 • NUMBER 3 • MAY 2009

## RESEARCH PAPERS



(Contents continued on inside back cover)

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

(**Contents continued**)

**Journal of Applied Mechanics** **Volume 76, Number 3** **MAY 2009**

## TECHNICAL BRIEFS



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

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

# Simulation of Moment Lyapunov Exponents for Linear Homogeneous Stochastic Systems

**Wei-Chau Xie**

**Qinghua Huang**

Department of Civil and Environmental Engineering,
University of Waterloo,
Waterloo, ON, N2L 3G1, Canada

*Moment Lyapunov exponents are important characteristic numbers for describing the dynamic stability of a stochastic system. When the pth moment Lyapunov exponent is negative, the pth moment of the solution of the stochastic system is stable. Monte Carlo simulation approaches complement approximate analytical methods in the determination of moment Lyapunov exponents and provides criteria on assessing the accuracy of approximate analytical results. For stochastic dynamical systems described by Itô stochastic differential equations, the solutions are diffusion processes and their variances may increase with time. Due to the large variances of the solutions and round-off errors, bias errors in the simulation of moment Lyapunov exponents are significant in improper numerical algorithms. An improved algorithm for simulating the moment Lyapunov exponents of linear homogeneous stochastic systems is presented in this paper.*
[DOI: 10.1115/1.3063629]

## 1 Introduction

Consider a general $d$-dimensional linear homogeneous stochastic dynamical system:

$$\dot{\mathbf{X}}(t) = \mathbf{A}(\boldsymbol{\xi}(t))\mathbf{X}(t), \quad \mathbf{X}(0) = \mathbf{X}_0 \tag{1.1}$$

where $\mathbf{A}$ is analytic, and $\boldsymbol{\xi}(t)=\{\xi_1(t),\xi_2(t),\ldots,\xi_r(t)\}^T$ is an $r$-dimensional vector of stochastic processes. According to the works by Arnold et al. [1–3], under rather general conditions, the $p$th moment Lyapunov exponent of system (1.1), which is defined by

$$\Lambda(p) = \lim_{t\to\infty}\frac{1}{t}\log \mathbf{E}[\|\mathbf{X}(t)\|^p] \tag{1.2}$$

exists and characterizes the stability of the $p$th moment of the solution $\mathbf{X}(t)$, where $\mathbf{E}[\cdot]$ denotes the expected value and $\|\cdot\|$ denotes a suitable vector norm. The $p$th moment of the solution of system (1.1) $\mathbf{E}[\|\mathbf{X}(t)\|^p]$ is asymptotically stable if $\Lambda(p)<0$. The slope of the $p$th moment Lyapunov exponent $\Lambda(p)$ at $p=0$, i.e., $\Lambda'(0)$, is equal to the largest Lyapunov exponent $\lambda$, which is defined as

$$\lambda = \lim_{t\to\infty}\frac{1}{t}\log\|\mathbf{X}(t)\| \tag{1.3}$$

and describes the almost-sure or sample stability of the system. When the largest Lyapunov exponent $\lambda$ is negative, system (1.1) is asymptotically stable with probability 1.

Even if the solution of system (1.1) is almost surely stable with $\lambda<0$, i.e., $\|\mathbf{X}(t)\|\to 0$ as $t\to\infty$ with probability 1 at the exponential rate $\lambda$, it is still possible that the $p$th moment is unstable, i.e., $\Lambda(p)>0$, since generally almost-sure convergence cannot assure moment convergence. It is therefore important to obtain the moment Lyapunov exponents so that the complete dynamic stability behavior of stochastic system (1.1) can be described.

Although it is quite straightforward to set up the partial differential eigenvalue problem with the $p$th moment Lyapunov exponent $\Lambda(p)$ as the principal eigenvalue [1–5], the actual solution of the eigenvalue problem is very difficult. For certain simple two-dimensional or four-dimensional systems, approximate analytical methods, such as stochastic averaging or perturbation, have been applied to obtain approximate analytical results of the moment Lyapunov exponent $\Lambda(p)$ (see, e.g., Refs. [6–11]).

In general, numerical approaches, such as Monte Carlo simulations, have to be applied to determine the moment Lyapunov exponents. Furthermore, even when approximate analytical results are available, their accuracy have to be verified by numerical simulations.

When investigating the stability of a general stochastic dynamical system, it is usual to consider the corresponding linearized system near its stationary solution. The linearized system is homogeneous. This shows the importance of linear homogeneous systems in the research of stochastic dynamical systems.

There are some references discussing the numerical approximation of Lyapunov exponents, such as Refs. [12,13]. However, to the best knowledge of the authors, there is only one numerical algorithm for determining the moment Lyapunov exponents using Monte Carlo simulation published so far [14], which is described briefly in Sec. 2.

## 2 Numerical Algorithm Using Sample Norm

Consider the $d$-dimensional linear homogeneous stochastic dynamical system (1.1). In the cases that $\boldsymbol{\xi}(t)$ is described by Itô or Stratonovich stochastic differential equations, Eq. (1.1) is solved using an appropriate numerical discretization scheme with a time step $h$. To have an accurate estimation of the $p$th moment, $\mathbf{E}[\|\mathbf{X}(t)\|^p]$, a large number of sample realizations must be simulated since the evaluation of expectation is determined by the sample average.

When system (1.1) is stable, the solution decays exponentially in time, whereas, when it is unstable, the solution grows exponentially with time. To avoid float-point data overflow or underflow, it is essential to devise an appropriate scheme to normalize the solutions regularly during simulation. In practice, there is no need to

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 14, 2006; final manuscript received October 2, 2008; published online March 3, 2009. Review conducted by N. Sri Namachchivaya.

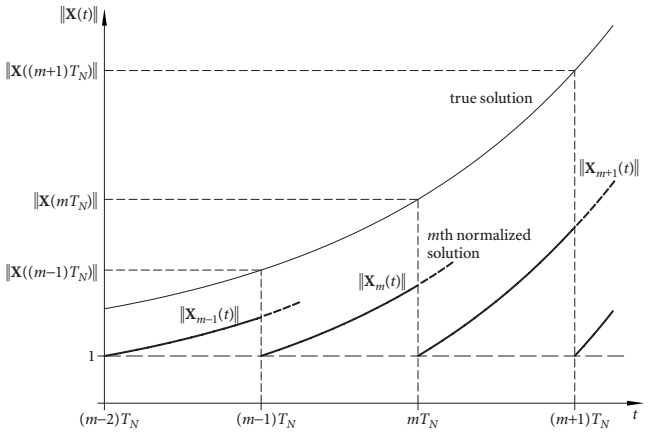



**Fig. 1 Growth of the solution and normalization**

normalize the solution at every iteration. Suppose the solution is normalized after every $K$ iterations or after every time period $T_N = Kh$.

Let $\mathcal{S}^{d-1}$ be the unit sphere in $d$-dimensional space $\mathcal{R}^d$. For a given initial condition $\mathbf{X}(0)=\mathbf{X}_0 \in \mathcal{S}^{d-1}$, i.e., $\|\mathbf{X}_0\|=1$, and simulation time $T=MT_N$, one has

$$\|\mathbf{X}(T)\| = \|\mathbf{X}(T,\mathbf{X}_0)\| = \prod_{m=1}^{M} \frac{\|\mathbf{X}(mT_N,\mathbf{X}_0)\|}{\|\mathbf{X}((m-1)T_N,\mathbf{X}_0)\|} \tag{2.1}$$

Since Eq. (1.1) is linear homogeneous, it can be seen that

$$\mathbf{X}_m(t) = \frac{\mathbf{X}((m-1)T_N+t,\mathbf{X}_0)}{\|\mathbf{X}((m-1)T_N,\mathbf{X}_0)\|}, \quad m=1,2,\ldots \tag{2.2}$$

solves Eq. (1.1) with the initial condition

$$\mathbf{X}_m(0) = \frac{\mathbf{X}((m-1)T_N,\mathbf{X}_0)}{\|\mathbf{X}((m-1)T_N,\mathbf{X}_0)\|} \in \mathcal{S}^{d-1} \tag{2.3}$$

Thus,

$$\|\mathbf{X}(T)\| = \prod_{m=1}^{M} \|\mathbf{X}_m(T_N)\| \tag{2.4}$$

Equations (2.2) and (2.3) indicate that normalization procedure can be performed for every time period $T_N$ such that the solution of the stochastic differentials always restarts from initial conditions with unit norm right after the normalization (see Fig. 1).

Let $\bar{\mathbf{X}}^h$ denote the solution from appropriate numerical discretization scheme with time step $h$. For a large simulation time $T$ with $M$ large, the $p$th moment Lyapunov exponent should be approximated as

$$\bar{\Lambda}^h(p) = \frac{1}{T}\log \mathbf{E}[\|\bar{\mathbf{X}}^h(T)\|^p] = \frac{1}{MT_N}\log \mathbf{E}\left[\prod_{m=1}^{M} \|\bar{\mathbf{X}}_m^h(T_N)\|^p\right] \tag{2.5}$$

In the algorithm in Ref. [14], the $p$th moment Lyapunov exponent is evaluated as

$$\bar{\Lambda}^h(p) = \frac{1}{MT_N}\log\prod_{m=1}^{M} \mathbf{E}[\|\bar{\mathbf{X}}_m^h(T_N)\|^p] \tag{2.6}$$

where the expectation is determined by the sample average

$$\mathbf{E}[\|\bar{\mathbf{X}}_m^h(T_N)\|^p] = \frac{1}{N}\sum_{s=1}^{N} \|\bar{\mathbf{X}}_m^{h,s}(T_N)\|^p \tag{2.7}$$

with $N$ being the sample size for simulation and $\bar{\mathbf{X}}_m^{h,s}$ being the $s$th sample path of $\bar{\mathbf{X}}_m^h$.

It is obvious that Eqs. (2.5) and (2.6) are different since all $\|\bar{\mathbf{X}}_m^h(T_N)\|$ are dependent and thus the expectation operation and the product operation cannot be interchanged. Notice that Eq. (2.6) can be rewritten as

$$\bar{\Lambda}^h(p) = \frac{1}{M}\sum_{m=1}^{M} \frac{1}{T_N}\log \mathbf{E}[\|\bar{\mathbf{X}}_m^h(T_N)\|^p] \tag{2.8}$$

It actually gives the average of $M$ moment Lyapunov exponents simulated for a time period of $T_N$ rather than the moment Lyapunov exponent simulated for a long time period of $T=MT_N$. Theoretically, the larger the value of $T_N$, the more accurate the approximation. Unfortunately, to avoid float-point data overflow and underflow, the value of $T_N$ cannot be very large. Although each simulation of the moment Lyapunov exponent for a relatively short time period of $T_N$ may not be accurate, for some systems the algorithm based on Eq. (2.6) yields satisfactory results because of the central limit theorem. However, there are systems for which Eq. (2.6) leads to erroneous results.

One possible revision to correct the insufficiency of algorithm (2.6) for linear homogeneous systems is to normalize the solutions by their expectations but not their norms, as shown in Eq. (2.2). With unit norm initial condition and the definition

$$\bar{\mathbf{Y}}_m^h(t) = \frac{\bar{\mathbf{X}}^h((m-1)T_N+t,\mathbf{X}_0)}{\mathbf{E}[\|\bar{\mathbf{X}}^h((m-1)T_N,\mathbf{X}_0)\|]}, \quad m=1,2,\ldots \tag{2.9}$$

the approximate moment Lyapunov exponents at time $T$ are given by

$$\begin{aligned}
\bar{\Lambda}^h(p) &= \frac{1}{MT_N} \times\log\left\{\frac{\mathbf{E}[\|\bar{\mathbf{X}}^h(MT_N)\|^p]}{\mathbf{E}[\|\bar{\mathbf{X}}^h((M-1)T_N)\|]^p}\prod_{m=1}^{M-1}\frac{\mathbf{E}[\|\bar{\mathbf{X}}^h(mT_N)\|]^p}{\mathbf{E}[\|\bar{\mathbf{X}}^h((m-1)T_N)\|]^p}\right\} \\
&= \frac{1}{MT_N}\left\{\log \mathbf{E}\left[\left\|\frac{\bar{\mathbf{X}}^h(MT_N)}{\mathbf{E}[\|\bar{\mathbf{X}}^h((M-1)T_N)\|]}\right\|^p\right] + \sum_{m=1}^{M-1} p\log \mathbf{E}\left[\left\|\frac{\bar{\mathbf{X}}^h(mT_N)}{\mathbf{E}[\|\bar{\mathbf{X}}^h((m-1)T_N)\|]}\right\|\right]\right\} \\
&= \frac{1}{MT_N}\left\{\log \mathbf{E}[\|\bar{\mathbf{Y}}_M^h(T_N)\|^p] + \sum_{m=1}^{M-1} p\log \mathbf{E}[\|\bar{\mathbf{Y}}_m^h(T_N)\|]\right\}
\end{aligned} \tag{2.10}$$

The solution of Eq. (1.1) may be a diffusion process and its variance may increase significantly with time. Although Eq. (2.10) is exact theoretically when $M$ is large enough, there are two main sources that will lead to significant numerical errors.

First, according to the central limit theorem, for independent and identically distributed (i.i.d.) random variables $x_1, x_2, \ldots$ with the same mean value $\mu$ and variance $\sigma^2$, the distribution of sample average $\bar{x}=(\Sigma_{s=1}^{N} x_s)/N$ will tend to the normal distribution $\mathcal{N}(\mu,\sigma^2/N)$. This means that Eq. (2.7) will not give acceptable results of the expected values when the variances of the solutions are so large that it is impossible to reduce the error of estimation to an acceptable level with a finite number of samples.

Second, due to the finite lengths of floating-point representations in computers, when two numbers are summed up, the smaller one will be neglected if the difference of their exponent bits exceeds the limit. If the system is unstable, its solution grows exponentially with time. Even when the system is stable and the chance that the solution takes extremely large values may be rare, once it happens, all the contributions from other samples will be eliminated. Thus this truncated error in estimating the expectations will be dominant in simulations with large variances.

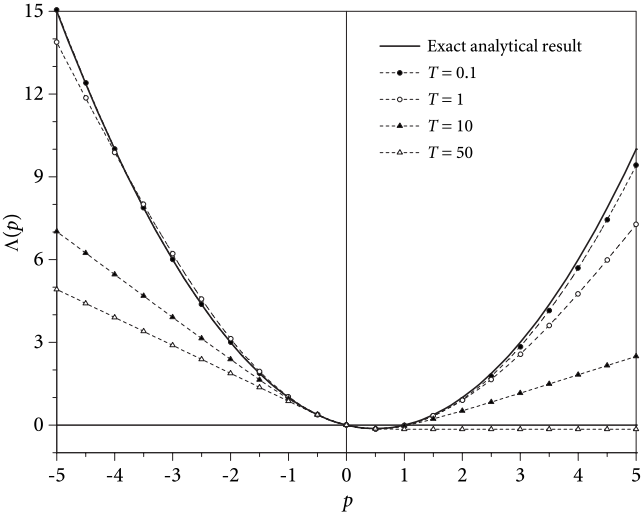



**Fig. 2 Simulation of moment Lyapunov exponents for a first-order linear system**

To illustrate, consider the first-order linear homogeneous stochastic system

$$dx(t) = ax(t)dt + \sigma x(t)dW(t) \tag{2.11}$$

where $a$ and $\sigma$ are real constants. The $p$th moment Lyapunov exponent can be easily determined as

$$\Lambda(p) = \lim_{t\to\infty}\frac{1}{t}\log \mathbf{E}[|x(t)|^p] = \frac{p}{2}[(p-1)\sigma^2 + 2a]$$

and the variance of norm is

$$\mathrm{var}[|x(t)|] = \mathbf{E}[|x(t)|^2] - \{\mathbf{E}[|x(t)|]\}^2 = e^{2at}(e^{\sigma^2 t} - 1)$$

Figure 2 shows the numerical results of the moment Lyapunov exponents for $a=0$, $\sigma=1$, and different values of the total time of simulation $T$, in which Eq. (2.11) is solved numerically using the explicit Euler scheme. The time step for iteration is $h=0.001$, the sample size is $N=5000$, and Eq. (2.10), i.e., the revised algorithm, is used to determine the approximate moment Lyapunov exponents. It is obvious that the longer the time $T$ for simulation, the worse the results. Because the variance of $|x(t)|$ increases exponentially with time, it is impossible to get an accurate estimate of the $p$th moment using sample average from finite sample sizes $N$ for $t$ large.

## 3 Estimation of the Expectation Through Logarithm of Norm

Because of the possible large errors in simulating moment Lyapunov exponents with increasing time of simulation, it is required to develop a new algorithm to overcome the difficulty in estimating the moments. Since the errors are caused by large variances of the solutions, it is clear that how to reduce the variances in order to obtain a good estimation of the moments using a finite number of samples is the key.

Notice that

$$\log \mathbf{E}[\|\mathbf{X}(T)\|^p] = \log \mathbf{E}[e^{p\log\|\mathbf{X}(T)\|}] = C(p) \tag{3.1}$$

where $C(p)$ is the cumulant generating function of $\log\|\mathbf{X}(T)\|$. In the special case when $\log\|\mathbf{X}(T)\|$ is normal, $C(p)$ takes the simple form

$$C(p) = p\mathbf{E}[\log\|\mathbf{X}(T)\|] + \tfrac{1}{2}p^2\,\mathrm{var}[\log\|\mathbf{X}(T)\|] \tag{3.2}$$

where $\mathrm{var}[\cdot]$ denotes the variance. Therefore, it may be possible to use the statistical properties of $\log\|\mathbf{X}(T)\|$ in estimating the moment Lyapunov exponents.

**3.1 Asymptotic Normality of Logarithm of Norm.** Since the simulation of stochastic dynamical systems is based on the theory of stochastic integrals, it is natural to start with the following $d$-dimensional Itô stochastic differential equation:

$$d\mathbf{X}(t) = \mathbf{B}_0(\mathbf{X},t)dt + \sum_{i=1}^{r}\mathbf{B}_i(\mathbf{X},t)dW_i \tag{3.3}$$

When the system is linear homogeneous with constant coefficients, Eq. (3.3) takes the form

$$d\mathbf{X}(t) = \mathbf{B}_0\mathbf{X}(t)dt + \sum_{i=1}^{r}\mathbf{B}_i\mathbf{X}(t)dW_i \tag{3.4}$$

where $\mathbf{B}_i$, $i=0,1,\ldots,r$, are the $d\times d$ constant matrices. Let $\rho(t) = \log\|\mathbf{X}(t)\|$. It has been shown (Ref. [15], p. 243) that the limit distribution of $(\rho(t)-\lambda t)/\sqrt{\mathrm{var}[\rho(t)]}$ is standard normal as $t\to\infty$ if $\mathrm{var}[\rho(t)]\to\infty$, and there exists a constant $\alpha$ such that, for any vector $\mathbf{Y}$,

$$\sum_{i=1}^{r}(\mathbf{Y}^T\mathbf{B}_i\mathbf{X})^2 \geq \alpha\|\mathbf{X}\|^2\|\mathbf{Y}\|^2 \tag{3.5}$$

is satisfied.

However, in applications, there are many cases that the nondegenerate condition (3.5) is not satisfied. An extended result by Arnold et al. [2] shows that, for any $\mathbf{X}(0)\neq\mathbf{0}$,

$$\lim_{t\to\infty}\frac{1}{t}\rho(t) = \lambda \quad \text{a.s.}, \quad \lim_{t\to\infty}\frac{1}{t}\log \mathbf{E}[\|\mathbf{X}(t)\|^p] = \Lambda(p) \tag{3.6}$$

and the normalized $\rho(t)$ converges weakly to Gaussian distribution with

$$\frac{\rho(t)-\lambda t}{\sqrt{t}} \overset{\text{distribution}}{\longrightarrow} \mathcal{N}(0,\Lambda''(0)), \quad t\to\infty \tag{3.7}$$

provided that

$$\dim \mathcal{LA}(\mathbf{g}_i: 0\leq i\leq r)(\mathbf{s}) = d-1 \quad \text{for all} \quad \mathbf{s}\in\mathcal{PJ}^{d-1} \tag{3.8}$$

where $\mathcal{LA}(\mathbf{g}_i)$ denotes the Lie algebra generated by the set of vector fields $\mathbf{g}_i$, dim denotes the dimension, $\mathcal{PJ}^{d-1}$ denotes the projective space obtained from $\mathcal{S}^{d-1}$ by identifying $\mathbf{s}=-\mathbf{s}$, and $\mathbf{g}_i$ are given by

$$\mathbf{g}_0(\mathbf{s}) = \mathbf{g}\left(\mathbf{B}_0 - \frac{1}{2}\sum_{i=1}^{r}\mathbf{B}_i^2,\mathbf{s}\right), \quad \mathbf{g}_i(\mathbf{s}) = \mathbf{g}(\mathbf{B}_i,\mathbf{s}), \quad i=1,\ldots,r \tag{3.9}$$

in which $\mathbf{g}$ is defined as, for any $d\times d$ matrix $\mathbf{B}$ and $\mathbf{s}\in\mathcal{S}^{d-1}$,

$$\mathbf{g}(\mathbf{B},\mathbf{s}) = (\mathbf{B} - (\mathbf{s}^T\mathbf{B}\mathbf{s})\mathbf{I})\mathbf{s} = \mathbf{B}\mathbf{s} - (\mathbf{s}^T\mathbf{B}\mathbf{s})\mathbf{s} \tag{3.10}$$

The linear homogeneous system with constant coefficients, i.e., Eq. (3.4), can be converted to the Stratonovich form. Thus it can describe the stochastic dynamical system

$$\dot{\mathbf{X}}(t) = \left(\mathbf{B}_0 - \frac{1}{2}\sum_{i=1}^{r}\mathbf{B}_i^2 + \sum_{i=1}^{r}\mathbf{B}_i\xi_i(t)\right)\mathbf{X}(t) = \mathbf{A}(\boldsymbol{\xi}(t))\mathbf{X}(t)$$

with $\xi_i(t), i=1,\ldots,r$ being the unit Gaussian white noises. It is very likely that noises $\xi_i(t)$ take other forms. The result obtained by Arnold et al. [1,3] solves this problem under some conditions.

Return to the general $d$-dimensional linear homogeneous system (1.1), with $\boldsymbol{\xi}(t)$ being a stationary ergodic diffusion vector process on a connected smooth manifold $\mathcal{M}$ described by Stratonovich stochastic differential equation:

$$d\boldsymbol{\xi}(t) = \mathbf{Q}_0(\boldsymbol{\xi}(t))dt + \sum_{i=1}^{r} \mathbf{Q}_i(\boldsymbol{\xi}(t)) \circ dW_i. \tag{3.11}$$

where $\mathbf{Q}_i$, $i=0,\ldots,r$, are smooth. If

$$\dim \mathcal{LA}(\mathbf{Q}_i: 1 \le i \le r)(\boldsymbol{\xi}) = \dim \mathcal{M} \quad \text{for all} \quad \boldsymbol{\xi} \in \mathcal{M} \tag{3.12}$$

$$\dim \mathcal{LA}(\mathbf{g}(\mathbf{A}(\boldsymbol{\xi}),\mathbf{s}): \boldsymbol{\xi} \in \mathcal{M})(\mathbf{s}) = d-1 \quad \text{for all} \quad \mathbf{s} \in \mathcal{PJ}^{d-1} \tag{3.13}$$

then for any $\mathbf{X}(0) \neq \mathbf{0}$, Eqs. (3.6) and (3.7) are still true. Fortunately, condition (3.13) is satisfied for most problems frequently considered in engineering applications [1]. Moreover, conditions (3.12) and (3.13) can be replaced by

$$\dim \mathcal{LA}\left(\mathbf{Q}_0 + \mathbf{g}(\mathbf{A},\mathbf{s}) + \frac{\partial}{\partial t}, \mathbf{Q}_1, \ldots, \mathbf{Q}_r\right)(\boldsymbol{\xi},\mathbf{s},t) = \dim \mathcal{M} + d \quad \text{for all} \quad (\boldsymbol{\xi},\mathbf{s},t) \in \mathcal{M} \times \mathcal{PJ}^{d-1} \times \mathcal{R} \tag{3.14}$$

As a result of Eqs. (3.6) and (3.7), one can write [6]

$$\Lambda(p) = \lambda p + \tfrac{1}{2}\Lambda''(0)p^2 + \mathcal{O}(p^2) \tag{3.15}$$

**3.2 Estimation Through Logarithm of Norm.** Suppose $\|\mathbf{X}(T)\|$ is obtained for linear homogeneous systems (3.4) or (1.1), and the corresponding system satisfies conditions (3.8), or (3.12) and (3.13). When $T$ is large enough and $\mathbf{X}(0) \neq \mathbf{0}$, $\rho(T) = \log\|\mathbf{X}(T)\|$ is near Gaussian according to the discussion in Sec. 3.1.

Let $\bar{\mathbf{X}}^h(T)$ still be the solution at time $T$ obtained from an appropriate numerical discretization scheme with time step $h$, and $\bar{\mathbf{X}}^{h,s}(T)$, $s=1,2,\ldots,N$, be the different samples of $\bar{\mathbf{X}}^h(T)$. Then $\bar{\mathbf{X}}^{h,s}(T)$ can be treated as i.i.d. random vectors with the same distribution as $\bar{\mathbf{X}}^h(T)$. Thus the $p$th moment Lyapunov exponent is approximated as

$$\bar{\Lambda}^h(p) = \frac{1}{T}\log \mathbf{E}[\|\bar{\mathbf{X}}^h(T)\|^p] = \frac{1}{NT}\log\left\{\prod_{s=1}^{N} \mathbf{E}[\|\bar{\mathbf{X}}^{h,s}(T)\|^p]\right\} = \frac{1}{NT}\log \mathbf{E}\left[\exp\left(p\sum_{s=1}^{N}\log\|\bar{\mathbf{X}}^{h,s}(T)\|\right)\right] \tag{3.16}$$

By defining

$$\bar{\rho}^h(T) = \log\|\bar{\mathbf{X}}^h(T)\|$$

$$\bar{\mu}_T^h = \mathbf{E}[\bar{\rho}^h(T)], \quad (\bar{\sigma}_T^h)^2 = \mathrm{var}[\bar{\rho}^h(T)]$$

$$\hat{\rho}^{h,s}(T) = \frac{\bar{\rho}^{h,s}(T) - \bar{\mu}_T^h}{\bar{\sigma}_T^h}, \quad R_N = \sum_{s=1}^{N} \hat{\rho}^{h,s}(T) \tag{3.17}$$

Eq. (3.16) becomes

$$\bar{\Lambda}^h(p) = \frac{1}{NT}\log \mathbf{E}[\exp(p\bar{\mu}_T^h N + p\bar{\sigma}_T^h R_N)] = p\frac{\bar{\mu}_T^h}{T} + \frac{1}{NT}\log \mathbf{E}[e^{p\bar{\sigma}_T^h R_N}] \tag{3.18}$$

With the notation

$$\zeta_N = \frac{R_N}{\sqrt{N}} \tag{3.19}$$

Eq. (3.18) is converted to

$$\bar{\Lambda}^h(p) = p\frac{\bar{\mu}_T^h}{T} + \frac{1}{NT}\log \mathbf{E}[e^{\sqrt{N}p\bar{\sigma}_T^h\zeta_N}] \tag{3.20}$$

Let $F_\zeta^{*N}(x)$ be the cumulative distribution function of $\zeta_N$, then $F_\zeta^{*N}(x)$ tends to the standard normal distribution of $\mathcal{N}(0,1)$ as $N \to \infty$ according to the central limit theorem, i.e., $F_\zeta^{*N}(x) \to \Phi(x)$, uniformly, where $\Phi(x)$ is the standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right)dy$$

Using the Edgeworth expansion theorem for distribution [16,17], one may have

$$F_\zeta^{*N}(x) = \Phi(x) + \sum_{k=3}^{\infty} c_k \Phi^{(k)}(x) \tag{3.21}$$

where the coefficients $c_k$ can be determined by the equality of moments on both sides of Eq. (3.21). Notice that $\zeta_N$ has zero mean and unit variance. If the $k$th central moment of $\bar{\rho}^h(T)$ is $\bar{\mu}_k^h$, for $k \ge 3$, then using integration by parts, it is easy to deduce from Eq. (3.21) that

$$c_3 = -\frac{1}{6}\frac{\bar{\mu}_3^h}{\sqrt{N}(\bar{\sigma}_T^h)^3}, \quad c_4 = \frac{1}{24}\frac{\bar{\mu}_4^h - 3(\bar{\sigma}_T^h)^4}{N(\bar{\sigma}_T^h)^4}, \cdots$$

Therefore, Eq. (3.20) yields

$$\begin{aligned}\bar{\Lambda}^h(p) &= p\frac{\bar{\mu}_T^h}{T} + \frac{1}{NT}\log\int_{-\infty}^{\infty} e^{\sqrt{N}p\bar{\sigma}_T^h x}\mathrm{d}F_\zeta^{*N}(x) \\ &= p\frac{\bar{\mu}_T^h}{T} + \frac{1}{NT}\log\{e^{1/2Np^2(\bar{\sigma}_T^h)^2}[1 - N^{3/2}p^3(\bar{\sigma}_T^h)^3 c_3 \\ &\quad + N^2p^4(\bar{\sigma}_T^h)^4 c_4 \\ &\quad + \cdots]\} = p\frac{\bar{\mu}_T^h}{T} + \frac{1}{2}p^2\frac{(\bar{\sigma}_T^h)^2}{T} \\ &\quad + \frac{1}{NT}\log\left\{1 + \frac{1}{6}Np^3\bar{\mu}_3^h \right. \\ &\quad \left. + \frac{1}{24}Np^4[\bar{\mu}_4^h - 3(\bar{\sigma}_T^h)^4] + \cdots\right\}\end{aligned} \tag{3.22}$$

The tail distribution of $F_\zeta^{*N}(x)$ is of paramount significance, since it is required to determine the expectation of $e^{\sqrt{N}p\bar{\sigma}_T^h\zeta_N}$ in Eq. (3.20) from a finite sample size $N$ in simulation. This means that accurate higher-order moments of $\bar{\rho}^h(T)$ are required in order to obtain a good approximation of $\mathbf{E}[e^{\sqrt{N}p\bar{\sigma}_T^h\zeta_N}]$. However, it is very difficult to do so in practice. If only lower-order moments are considered, the estimation error of moments may make the sum of a finite number of terms within the argument of the last logarithm in Eq. (3.22) be negative when $N$ becomes large, which will lead to invalid operation in simulation. Therefore, the distribution of $\bar{\rho}^h(T)$ has to be considered to find an appropriate estimation.

Noticing that Eq. (3.2) is true for normal distribution, one may attempt to see if the last logarithm term in Eq. (3.22) can be dropped in simulation since the distribution of the normalized $\bar{\rho}^h(T)$ approaches normal as $T$ goes to infinity.

From the definition of $\hat{\rho}^{h,s}(T)$, it can be seen that $\hat{\rho}^{h,s}(T)$, $s=1,2,\ldots,N$, are i.i.d. random variables with zero mean and unit variance. Moreover, using Eq. (3.1), it is obvious that the existence of moment Lyapunov exponents ensures that $\mathbf{E}[e^{\eta\hat{\rho}^{h,s}(T)}] < \infty$ for $|\eta| \le \eta_0$, where $\eta_0$ is some constant. Then according to the theorem proved by Komlós et al. [18,19], a sequence of standard normal random variables $z_s$, $s=1,2\ldots,N$, can be constructed such that, for every $N$ and all $x>0$, the partial sums $R_k = \Sigma_{s=1}^k \hat{\rho}^{h,s}(T)$ and $V_k = \Sigma_{s=1}^k z_s$ satisfy

$$\mathbf{P}\{\max_{k\le N}|R_k - V_k| > C_0 \log N + x\} < \delta_0 e^{-\delta x} \tag{3.23}$$

where $\mathbf{P}\{\cdot\}$ denotes the probability, $C_0$, $\delta$, and $\delta_0$ depend only on the distribution of $\hat{\rho}^{h,s}(T)$, and $\delta$ can be as large as possible by choosing $C_0$ large enough. Thus it follows $|R_N - V_N| = \mathcal{O}(\log N)$ almost surely for every $N$ [18,19].

Considering the near normality of $\hat{\rho}^{h,s}(T)$ and the finite sample size, events with zero probability are treated as not likely to happen in simulation. This means that $R_N$ is replaced by $V_N + \mathcal{O}(\log N)$ in the evaluation of expectation and thus Eq. (3.18) is approximated as

$$\begin{aligned}\bar{\Lambda}^h(p) &= p\frac{\bar{\mu}_T^h}{T} + \frac{1}{NT}\log \mathbf{E}[e^{p\bar{\sigma}_T^h\{V_N + \mathcal{O}(\log N)\}}] \\ &= p\frac{\bar{\mu}_T^h}{T} + \frac{1}{2}p^2\frac{(\bar{\sigma}_T^h)^2}{T} + \mathcal{O}\left(\frac{\log N}{N}\right)\end{aligned} \tag{3.24}$$

Hence, when the sample size $N$ is large enough, by neglecting the last term in Eq. (3.24) and estimating the mean and variance of logarithm of norm, the $p$th moment Lyapunov exponent is given by

$$\bar{\Lambda}^h(p) = \frac{1}{T}\left\{p\mathbf{E}[\log\|\overline{\mathbf{X}}^h(T)\|] + \frac{1}{2}p^2\,\mathrm{var}[\log\|\overline{\mathbf{X}}^h(T)\|]\right\} \tag{3.25}$$

It is obvious that the variance of $\log\|\overline{\mathbf{X}}^h(T)\|$ will be much smaller than the variance of $\|\overline{\mathbf{X}}^h(T)\|$ when $\|\overline{\mathbf{X}}^h(T)\|$ becomes large; therefore, obtaining a good estimation of the $p$th moment Lyapunov exponent through sample average is possible.

From Eq. (3.25), one also sees that the largest Lyapunov exponent can be approximated as

$$\bar{\lambda}^h = \frac{1}{T}\mathbf{E}[\log\|\overline{\mathbf{X}}^h(T)\|] \tag{3.26}$$

which is the same result as given by Talay [12].

## 4 Algorithm for Linear Homogeneous Stochastic Systems

Following Eq. (3.25), an algorithm for simulating the moment Lyapunov exponents of linear homogeneous stochastic dynamical system (3.4)

$$d\mathbf{X}(t) = \mathbf{B}_0\mathbf{X}(t)dt + \sum_{i=1}^{r}\mathbf{B}_i\mathbf{X}(t)\mathrm{d}W_i$$

and system (1.1)

$$\dot{\mathbf{X}}(t) = \mathbf{A}(\boldsymbol{\xi}(t))\mathbf{X}(t)$$

$$d\boldsymbol{\xi}(t) = \mathbf{Q}_0(\boldsymbol{\xi}(t))dt + \sum_{i=1}^{r}\mathbf{Q}_i(\boldsymbol{\xi}(t))\circ\mathrm{d}W_i$$

can be described as follows. Since the simulation requires the statistical properties of logarithm of the norm, the normalization operation described in Sec. 2 can be applied.

*Step 1*. Use an appropriate time discrete approximation, such as the Euler scheme, to discretize system (3.4) or (1.1) with time step $h$. Details of various numerical schemes for solving stochastic differential equations can be found in Ref. [20].

*Step 2*. Set the initial conditions of the state vector $\overline{\mathbf{X}}^h(t,\overline{\mathbf{X}}^h(0))$, by $\overline{\mathbf{X}}^h(0)\in\mathcal{S}^{d-1}$, i.e.,

$$\|\overline{\mathbf{X}}^{h,s}(0)\| = 1, \quad s = 1,2,\ldots,N$$

where $N$ is the sample size, and $\|\mathbf{X}\| = \sqrt{\mathbf{X}^T\mathbf{X}}$ is the Euclidean norm of vector $\mathbf{X}$.

*Step 3*. Solve the discretized system iteratively. Apply the normalization procedure as described in Sec. 2, i.e., normalization is performed after every $K$ iterations or after every time period $T_N = Kh$. At the $m$th normalization, or at $t = mT_N$, the $s$th sample is normalized using Eq. (2.3), i.e.,

$$\overline{\mathbf{X}}_{m+1}^{h,s}(0) = \frac{\overline{\mathbf{X}}^{h,s}(mT_N,\overline{\mathbf{X}}^{h,s}(0))}{\|\overline{\mathbf{X}}^{h,s}(mT_N,\overline{\mathbf{X}}^{h,s}(0))\|} = \frac{\overline{\mathbf{X}}_m^{h,s}(T_N,\overline{\mathbf{X}}_m^{h,s}(0))}{\|\overline{\mathbf{X}}_m^{h,s}(T_N,\overline{\mathbf{X}}_m^{h,s}(0))\|}, \quad m = 1,2,\ldots$$

Simulation is then continued with the initial condition $\overline{\mathbf{X}}_{m+1}^{h,s}(0)$ with $\|\overline{\mathbf{X}}_{m+1}^{h,s}(0)\| = 1$ for another $K$ iterations.

*Step 4*. Defining

$$\bar{\rho}^{h,s}(mT_N) = \log\|\overline{\mathbf{X}}^{h,s}(mT_N,\overline{\mathbf{X}}^{h,s}(0))\|,$$

$$\bar{\rho}_m^{h,s}(T_N) = \log\|\overline{\mathbf{X}}_m^{h,s}(T_N,\overline{\mathbf{X}}_m^{h,s}(0))\|$$

then, using Eq. (2.4),

$$\begin{aligned}\bar{\rho}^{h,s}(MT_N) &= \log\|\overline{\mathbf{X}}^{h,s}(MT_N,\overline{\mathbf{X}}^{h,s}(0))\| \\ &= \log\left[\prod_{m=1}^{M}\|\overline{\mathbf{X}}_m^{h,s}(T_N,\overline{\mathbf{X}}_m^{h,s}(0))\|\right] = \sum_{m=1}^{M}\bar{\rho}_m^{h,s}(T_N)\end{aligned}$$

*Step 5*. After $KM$ iterations, i.e., at time $T = MKh$, use the following equations:

$$\mathbf{E}[\log\|\overline{\mathbf{X}}^h(T)\|] = \mathbf{E}[\bar{\rho}^h(T)] = \frac{1}{N}\sum_{s=1}^{N}\bar{\rho}^{h,s}(T) = \bar{\mu}_T^h$$

$$\mathrm{var}[\log\|\overline{\mathbf{X}}^h(T)\|] = \mathrm{var}[\bar{\rho}^h(T)] = \frac{1}{N-1}\sum_{s=1}^{N}[\bar{\rho}^{h,s}(T)^2 - (\bar{\mu}_T^h)^2] = (\bar{\sigma}_T^h)^2 \tag{4.1}$$

to estimate the mean and variance of $\log\|\overline{\mathbf{X}}^h(T)\|$.

*Step 6*. Use Eq. (3.25), combining with Eqs. (4.1), i.e.,

$$\bar{\Lambda}^h(p) = \frac{1}{T}\left[p\bar{\mu}_T^h + \frac{1}{2}p^2(\bar{\sigma}_T^h)^2\right] \tag{4.2}$$

to calculate the moment Lyapunov exponents for all values of $p$ of interest.

To obtain a more accurate estimation of the mean and variance of $\log\|\overline{\mathbf{X}}^h(T)\|$, large sample size $N$ has to be used. Moreover, time for simulation $T$ has to be large enough to get a good approximation of moment Lyapunov exponents. This means iteration times in solving the system will be extremely large. Since the simulation of different samples can be implemented independently, parallel computation using Message Passing Interface[1] (MPI) or OPENMP application program interface[2] will be a great benefit to the simulation of moment Lyapunov exponents.

It should be mentioned that although higher-order schemes lead to more accurate results, it is easier to implement the Monte Carlo simulation using Euler scheme due to its simple form, especially when the system is high dimensional and complicated. Moreover, Romberg extrapolation may be applied to increase the precision of results to $\mathcal{O}(h^2)$ using Euler scheme, which is the order of $\mathcal{O}(h)$ [21].

## 5 Examples in Application

The Monte Carlo simulation algorithm presented in Sec. 4 is applicable to linear homogeneous stochastic dynamical systems, which have wide applications in engineering mechanics, such as oscillators under parametric excitations of noises. Before the algorithm is applied, conditions (3.8) (for system (3.4)), (3.12), and (3.13), or (3.14) (for system (1.1)) should be satisfied.

[1] http://www.mpiweb.org.
[2] http://www.openmp.org.

According to the definition of Lie algebra, $\mathcal{LA}(\mathbf{g}_i)$ is actually a vector space generated by the set of vectors $\mathbf{g}_i$, thus $\dim \mathcal{LA}(\mathbf{g}_i)(\mathbf{s})$ can be determined by the dimension of this vector space. It turns out that condition (3.8) is satisfied, provided that the space spanned by vectors $\{\mathbf{g}_i: 0 \le i \le r\}$ has dimension $d$ since $\mathbf{s} \in \mathcal{PT}^{d-1}$. For system (1.1), condition (3.13) is satisfied for the special case [3]

$$\mathbf{A}(\boldsymbol{\xi}) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_1(\boldsymbol{\xi}) & a_2(\boldsymbol{\xi}) & a_3(\boldsymbol{\xi}) & \cdots & a_d(\boldsymbol{\xi}) \end{bmatrix}$$

In particular, when $d=2$ and $a_2(\boldsymbol{\xi})$ is constant, i.e.,

$$\mathbf{A}(\boldsymbol{\xi}(t)) = \begin{bmatrix} 0 & 1 \\ -f(\boldsymbol{\xi}(t)) & -\beta \end{bmatrix}$$

system (1.1) becomes

$$\ddot{q}(t) + \beta\dot{q}(t) + f(\boldsymbol{\xi}(t))q(t) = 0 \tag{5.1}$$

which describes the motion of a damped oscillator under noise perturbation. When $f(\boldsymbol{\xi})$ is not a constant function, the vectors $\mathbf{g}(\mathbf{A}(\boldsymbol{\xi}), \mathbf{s})$ are not in the same direction for different values of $\boldsymbol{\xi}$. Thus it can be easily verified that condition (3.13) is true.

It is stated in Ref. [3] that most systems considered in physics and engineering satisfy the required conditions to ensure the existence of the moment Lyapunov exponents and the asymptotic normality of logarithm of norm. This means that the algorithm in Sec. 4 can be applied directly in most cases, and the asymptotic normality of the logarithm of norm may be verified through the histogram estimation obtained in simulation.

In this section, moment Lyapunov exponents of three single degree-of-freedom systems under white noise, real noise, and bounded noise excitations, respectively, are determined. The numerical results of simulation are compared with known approximate analytical results.

**5.1 An Oscillator Under Weak White Noise Excitation.** Consider the following two-dimensional oscillator under the excitation of white noises:

$$\ddot{q}(t) + [2\varepsilon\beta + \varepsilon^{1/2}\sigma_2\zeta_2(t)]\dot{q}(t) + \omega^2[1 + \varepsilon^{1/2}\sigma_1\zeta_1(t)]q(t) = 0 \tag{5.2}$$

where $\zeta_1(t)$ and $\zeta_2(t)$ are the unit Gaussian white noise processes, and $0 < \varepsilon \ll 1$ is a small parameter. Approximate moment Lyapunov exponents can be obtained by the method of perturbation [7,5], which is given by

$$\Lambda(p) = \varepsilon p\left(-\beta + \frac{p+2}{16}\omega^2\sigma_1^2 + \frac{3p+2}{16}\sigma_2^2\right) + \mathcal{O}(\varepsilon^3) \tag{5.3}$$

Equation (5.2) can be converted to the Itô differential equations

$$d\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -\varepsilon\left(2\beta - \frac{1}{2}\sigma_2^2\right) \end{bmatrix}\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}dt + \begin{bmatrix} 0 & 0 \\ -\varepsilon^{1/2}\omega^2\sigma_1 & 0 \end{bmatrix} \times \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}dW_1(t) + \begin{bmatrix} 0 & 0 \\ 0 & -\varepsilon^{1/2}\sigma_2 \end{bmatrix}\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}dW_2(t) \tag{5.4}$$

where $x_1(t)=q(t)$ and $x_2(t)=\dot{q}(t)$. It is obvious that system (5.4) is of the form (3.4) with $d=2$, $r=2$, and

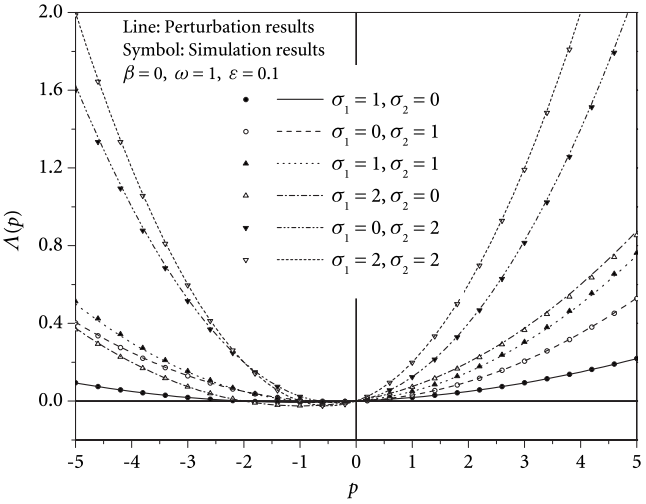



**Fig. 3 Moment Lyapunov exponents under white noise excitation ($\varepsilon = 0.1$)**

$$\mathbf{B}_0 = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -\varepsilon\left(2\beta - \frac{1}{2}\sigma_2^2\right) \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} 0 & 0 \\ -\varepsilon^{1/2}\omega^2\sigma_1 & 0 \end{bmatrix},$$
$$\mathbf{B}_2 = \begin{bmatrix} 0 & 0 \\ 0 & -\varepsilon^{1/2}\sigma_2 \end{bmatrix}$$

It can be verified that the set of vectors $\{\mathbf{g}_0, \mathbf{g}_1, \mathbf{g}_2\}$ spans a space with dimension 2 since all $\mathbf{g}_i$, $i=0, 1, 2$, are two-dimensional vectors. Thus condition (3.8) is satisfied and the algorithm presented in Sec. 4 can be applied to simulate the moment Lyapunov exponents.

The explicit Euler scheme is applied for the simulation

$$x_1^{k+1} = x_1^k + x_2^k \cdot h$$
$$x_2^{k+1} = x_2^k + \left[-\omega^2 x_1^k - \varepsilon\left(2\beta - \frac{1}{2}\sigma_2^2\right)x_2^k\right] \cdot h - \varepsilon^{1/2}\omega^2\sigma_1 x_1^k \cdot \Delta W_1^k - \varepsilon^{1/2}\sigma_2 \cdot \Delta W_2^k$$

The damping coefficient is set to $\beta=0$ and $\omega=1$. The sample size is $N=10{,}000$, time step $h=0.0001$, and the number of iterations is $MK=5\times10^7$, i.e., the total length of time of simulation is $T=5000$.

Figures 3 and 4 show the comparison of approximate analytical moment Lyapunov exponents given by Eq. (5.3) and the Monte Carlo simulation results for different values of $\sigma_1$, $\sigma_2$, and $\varepsilon$. It can be seen that the approximate analytical results fit rather well with the simulation results in most cases, implying that the algorithm in Sec. 4 works well as predicted.

To illustrate the asymptotic normality of logarithm of norm, the normalized histograms of $\log\|\mathbf{x}(T)\|$ for some typical values of $\sigma_1$ and $\sigma_2$ are plotted in Fig. 5. The corresponding normal density approximations given by

$$\bar{\varphi}(x) = \frac{1}{\sqrt{2\pi}\bar{\sigma}_T^h}\exp\left[-\frac{(x-\bar{\mu}_T^h)^2}{2(\bar{\sigma}_T^h)^2}\right] \tag{5.5}$$

are also shown in the same figure for comparison. It appears that, as the time of simulation $T$ is large enough, the distribution of $\log\|\mathbf{x}(T)\|$ does approach normal distribution.

**5.2 An Oscillator Under Real Noise Excitation.** Consider an oscillator under the excitation of real noise or Ornstein–Uhlenbeck process $\zeta(t)$,

$$\ddot{q}(t) + 2\varepsilon\beta\dot{q}(t) + \omega^2[1 - \varepsilon^{1/2}\zeta(t)]q(t) = 0$$

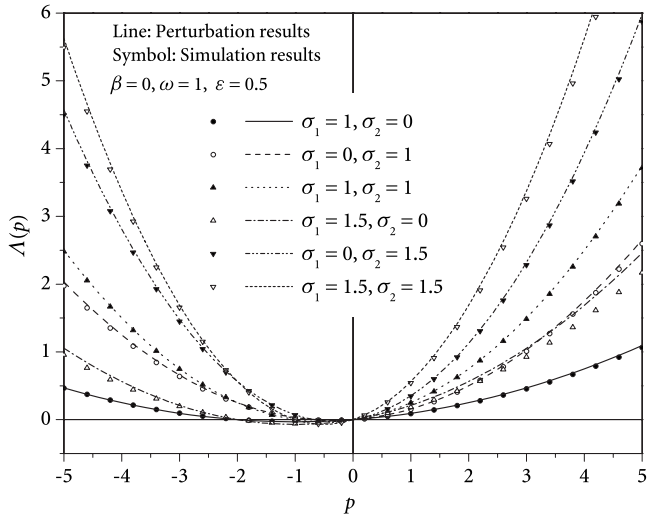



**Fig. 4 Moment Lyapunov exponents under white noise excitation ($\varepsilon = 0.5$)**

$$d\zeta(t) = -\alpha\zeta(t)dt + \sigma \mathrm{d}W(t) \tag{5.6}$$

Xie [10,5] determined the approximate moment Lyapunov exponents using the method of perturbation as

$$\Lambda(p) = -\varepsilon p\beta + p(p+2)\left[\varepsilon\frac{\sigma^2\omega^2}{16(\alpha^2+4\omega^2)} + \varepsilon^2\frac{(\alpha^4+22\alpha^2\omega^2+48\omega^4)\sigma^4\omega^4}{32\alpha(\alpha^2+\omega^2)(\alpha^2+4\omega^2)^3}\right] + \mathcal{O}(\varepsilon^3) \tag{5.7}$$

Obviously, the first equation in Eq. (5.6) takes the form of Eq. (5.1) with $f(x) = \omega^2 - \varepsilon^{1/2}\omega^2 x$. And the real noise $\zeta(t)$ satisfies condition (3.12). Therefore, the algorithm in Sec. 4 can be applied.

Letting

$$x_1(t) = q(t), \quad x_2(t) = \dot{q}(t), \quad x_3(t) = \zeta(t) \tag{5.8}$$

system (5.6) can be converted to the Itô differential equations

$$d\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} x_2 \\ -\omega^2 x_1 - 2\varepsilon\beta x_2 + \varepsilon^{1/2}\omega^2 x_3 x_1 \\ -\alpha x_3 \end{Bmatrix} dt + \begin{Bmatrix} 0 \\ 0 \\ \sigma \end{Bmatrix} dW(t)$$

The iteration equations using explicit Euler scheme are then given by

$$x_1^{k+1} = x_1^k + x_2^k \cdot h$$

$$x_2^{k+1} = x_2^k + (-\omega^2 x_1^k - 2\varepsilon\beta x_2^k + \varepsilon^{1/2}\omega^2 x_3^k x_1^k) \cdot h$$

$$x_3^{k+1} = x_3^k - \alpha x_3^k \cdot h + \sigma \cdot \Delta W^k$$

The norm for evaluating the moment Lyapunov exponents is $\|q(t)\| = (x_1^2 + x_2^2)^{1/2}$. The sample size for estimating the expected value is $N = 20{,}000$, time step $h = 0.0001$, and the total length of time of simulation is $T = 5000$, i.e., the number of iterations is $MK = 5 \times 10^7$.

Figures 6 and 7 show the typical results of the moment Lyapunov exponents for different values of $\alpha$ and $\sigma$, with the parameters taken as $\varepsilon = 0.1$, $\beta = 0.05$, and $\omega = 1$. It can be seen that when $\sigma$ is small, i.e., the noise intensity is weak, and for different values of $\alpha$, the approximate results from the perturbation method agree well with the simulation results. This is reasonable since the analytical approximations are obtained by weak noise expansion of eigenvalue problem governing the moment Lyapunov exponents.

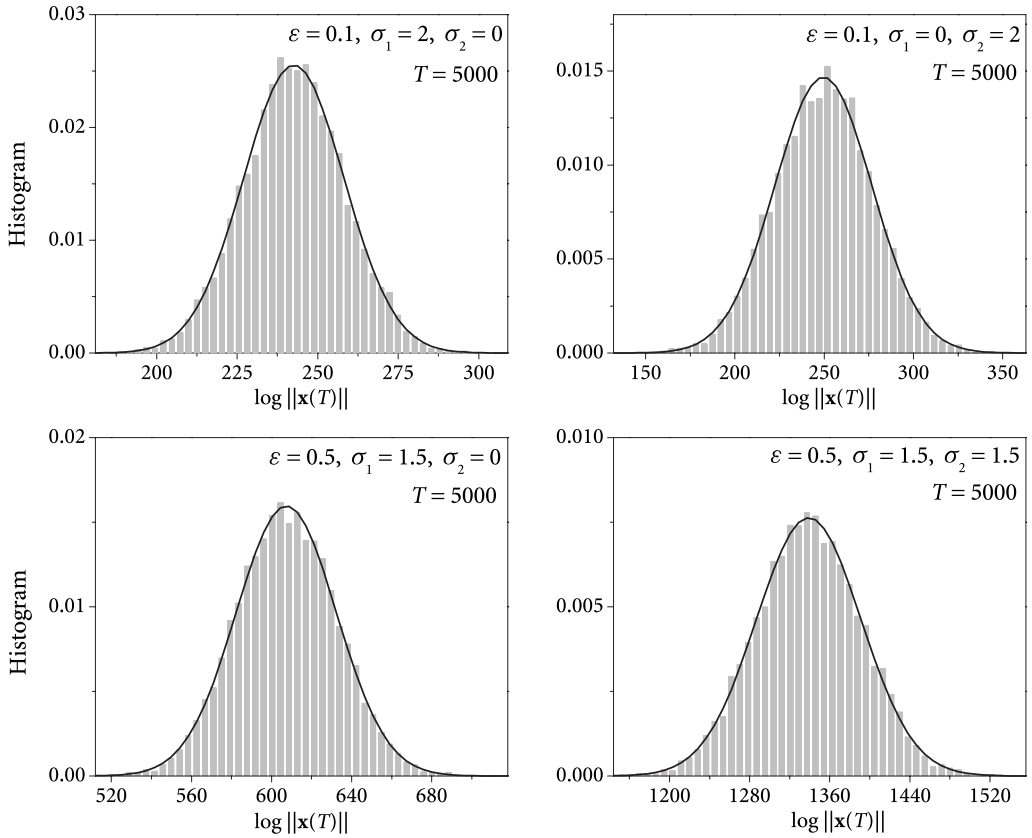



**Fig. 5 Histograms of logarithm of norm compared with normal density approximations**

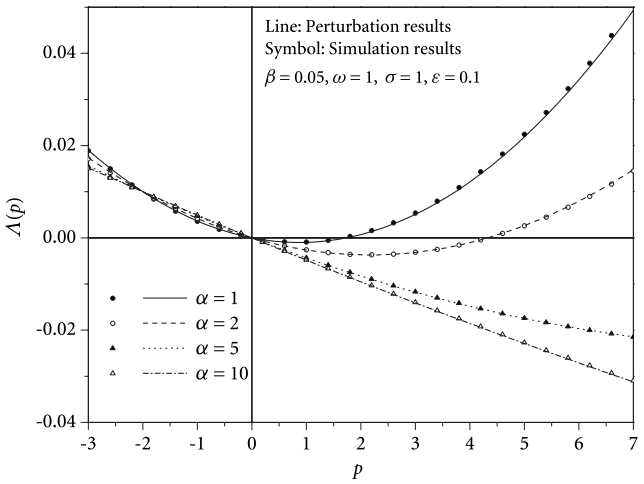



**Fig. 6 Moment Lyapunov exponents under real noise excitation for different $\alpha$**

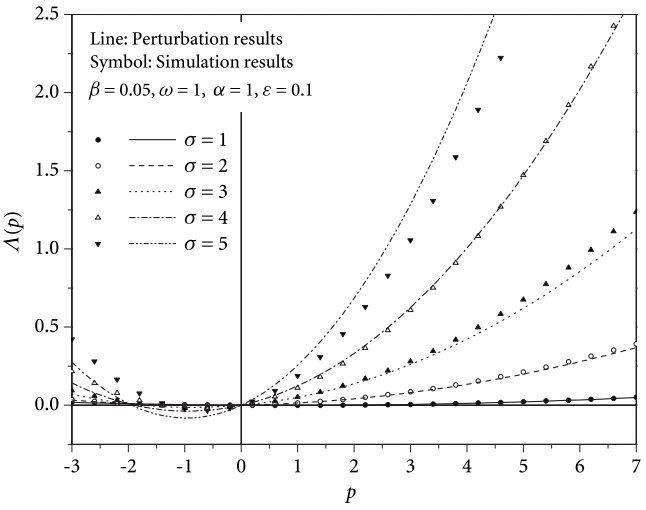



**Fig. 7 Moment Lyapunov exponents under real noise excitation for different $\sigma$**

**5.3 An Oscillator Under Bounded Noise Excitation.** Consider an oscillator under the excitation of bounded noise,

$$\ddot{q}(t) + 2\varepsilon\beta\dot{q}(t) + \omega^2[1 - \varepsilon\mu\cos\zeta(t)]q(t) = 0$$

$$d\zeta(t) = \nu dt + \varepsilon^{1/2}\sigma dW(t) \tag{5.9}$$

Xie [11,5] determined the approximate moment Lyapunov exponents using the method of perturbation as

$$\Lambda(p) = -\varepsilon p\beta + \frac{\varepsilon^3 p(p+2)\omega^4\mu^2\sigma^2\left[\nu^2 + 4(\omega^2 - \varepsilon^2\beta^2) + \frac{1}{4}\varepsilon^2\sigma^4\right]}{32(\omega^2 - \varepsilon^2\beta^2)\left[(2\sqrt{\omega^2 - \varepsilon^2\beta^2} + \nu)^2 + \frac{1}{4}\varepsilon^2\sigma^4\right]\left[(2\sqrt{\omega^2 - \varepsilon^2\beta^2} - \nu)^2 + \frac{1}{4}\varepsilon^2\sigma^4\right]} + o(\varepsilon^3) \tag{5.10}$$

Similar to the real noise case, condition (3.13) is satisfied with $f(x) = \omega^2 - \varepsilon\mu\omega^2 x$ in the form of Eq. (5.1). The bounded noise $\xi(t) = \cos\zeta(t)$ satisfies condition (3.12) since one has the Stratonovich stochastic differential equation

$$d\xi(t) = -\nu\sqrt{1-\xi^2}dt - \sigma\sqrt{1-\xi^2}\circ dW(t)$$

Using the same notation as Eq. (5.8), the Itô differential equations for system (5.9) become

$$d\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} x_2 \\ -\omega^2 x_1 - 2\varepsilon\beta x_2 + \varepsilon\omega^2\mu x_1\cos x_3 \\ \nu \end{Bmatrix} dt + \begin{Bmatrix} 0 \\ 0 \\ \varepsilon^{1/2}\sigma \end{Bmatrix} dW(t)$$

The iteration equations using explicit Euler scheme are then given by

$$x_1^{k+1} = x_1^k + x_2^k \cdot h$$

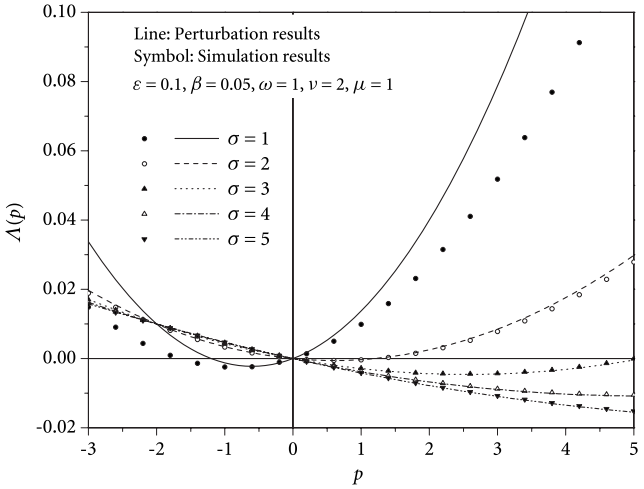



**Fig. 8 Moment Lyapunov exponents under bounded noise excitation for different $\sigma$**

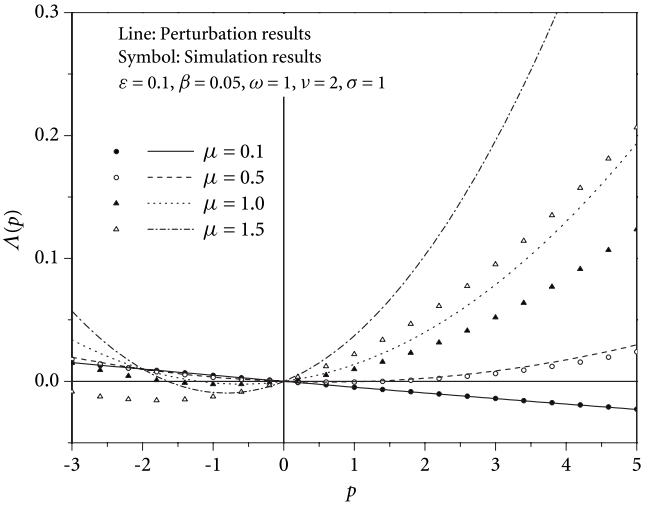



**Fig. 9 Moment Lyapunov exponents under bounded noise excitation for different $\mu$**

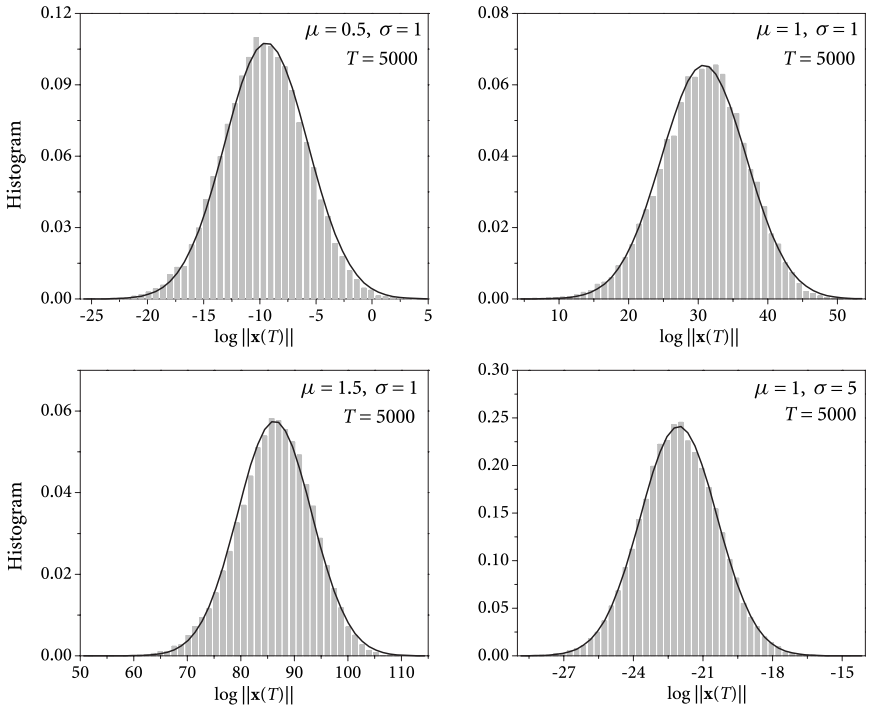



**Fig. 10 Histograms of logarithm of norm compared with normal density approximations**

$$x_2^{k+1} = x_2^k + (-\omega^2 x_1^k - 2\varepsilon\beta x_2^k + \varepsilon\omega^2\mu x_1^k \cos x_3^k) \cdot h$$

$$x_3^{k+1} = x_3^k + \nu \cdot h + \varepsilon^{1/2}\sigma \cdot \Delta W^k$$

The norm for evaluating the moment Lyapunov exponents is $\|q(t)\| = (x_1^2 + x_2^2)^{1/2}$. The sample size for estimating the expected value is $N = 20{,}000$, time step $h = 0.0001$, and the total length of time of simulation is $T = 5000$, i.e., the number of iterations is $MK = 5 \times 10^7$.

Typical results of the moment Lyapunov exponents for different values of $\sigma$ and $\mu$ are shown in Figs. 8 and 9, with the parameters taken as $\varepsilon = 0.1$, $\beta = 0.05$, $\omega = 1$, and $\nu = 2$. It can be seen that the approximate results from perturbation method agree well with the simulation results for small $\mu$ and large $\sigma$. This is the result that, in the eigenvalue problem governing the moment Lyapunov exponents, the approximate analysis using perturbation method requires $\varepsilon\mu$ be small enough and $\varepsilon\sigma^2$ be in the order $\mathcal{O}(1)$. The discrepancy between perturbation and simulation for small $\sigma$ and relatively large $\mu$ shows that some better approximation methods have to be considered.

The histograms of $\log\|q(T)\|$ for different values of $\mu$ and $\sigma$ are plotted in Fig. 10. Again, the figure shows the asymptotic normality of $\log\|q(T)\|$.

## 6 Conclusion

Contrary to intuition, Monte Carlo simulation of moment Lyapunov exponents of stochastic dynamical systems is a very difficult topic. Because the solution of a system grows exponentially when it is unstable and decays exponentially when it is stable, float-point overflow or underflow renders "brute-force" approaches inapplicable. Furthermore, because the variance of the solution may grow with time, it is very challenging to obtain an accurate estimation of the moments with finite sample size.

For linear homogeneous stochastic dynamical systems, a Monte Carlo simulation algorithm used to determine the moment Lyapunov exponents is presented in this paper. Since the limit distribution of the logarithm of norm of the solution is normal, the mean value and variance of the logarithm of norm, combined with normalization of the solution, are used to reduce the possible large variance of the solution so that the $p$th moment can be estimated. Numerical examples are presented to compare the Monte Carlo simulation results with approximate analytical results. This approach gives a better numerical approximation than the previous method [14], which uses the direct sample average of norm as the estimation of expectation.

## Acknowledgment

The research for this paper was supported, in part, by the Natural Sciences and Engineering Research Council of Canada.

Parts of the Monte Carlo simulation were made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca).

The authors are grateful to the referees for their suggestion, which helped to improve the paper.

## References

[1] Arnold, L., 1984, "A Formula Connecting Sample and Moment Stability of Linear Stochastic Systems," SIAM J. Appl. Math., **44**(4), pp. 793–802.

[2] Arnold, L., Oeljeklaus, E., and Pardoux, E., 1986, "Almost Sure and Moment Stability for Linear Itô Equations," *Lyapunov Exponents*, Vol. 1186 (Lecture Notes in Mathematics), L. Arnold and V. Wihstutz, eds., Springer-Verlag, Berlin, pp. 129–159.

[3] Arnold, L., Kliemann, W., and Oeljeklaus, E., 1986, "Lyapunov Exponents of Linear Stochastic Systems," *Lyapunov Exponents*, Vol. 1186 (Lecture Notes in Mathematics), L. Arnold and V. Wihstutz, eds., Springer-Verlag, Berlin, pp. 85–125.

[4] Wedig, W., 1988, "Lyapunov Exponent of Stochastic Systems and Related Bifurcation Problems," *Stochastic Structural Dynamics: Progress in Theory and Applications*, S. T. Ariaratnam, G. I. Schuëller, and I. Elishakoff, eds., Elsevier Applied Science, New York, pp. 315–327.

[5] Xie, W.-C., 2006, *Dynamic Stability of Structures*, Cambridge University Press, New York.

[6] Arnold, L., Doyle, M. M., and Sri Namachchivaya, N., 1997, "Small Noise Expansion of Moment Lyapunov Exponents for Two-Dimensional Systems," Dyn. Stab. Syst., **12**(3), pp. 187–211.

[7] Moshchuk, N., and Khasminskii, R. Z., 1998, "Moment Lyapunov Exponent and Stability Index for Linear Conservative System With Small Random Perturbation," SIAM J. Appl. Math., **58**(1), pp. 245–256.

[8] Sri Namachchivaya, N., and Van Roessel, H. J., 2001, "Moment Lyapunov Exponent and Stochastic Stability of Two Coupled Oscillators Driven by Real

Noise," Trans. ASME, J. Appl. Mech., **68**(6), pp. 903–914.
[9] Sri Namachchivaya, N., and Vedula, L., 2000, "Stabilization of Linear Systems by Noise: Application to Flow Induced Oscillations," Dyn. Stab. Syst., **15**(2), pp. 185–208.
[10] Xie, W.-C., 2001, "Moment Lyapunov Exponents of a Two-Dimensional System Under Real Noise Excitation," J. Sound Vib., **239**(1), pp. 139–155.
[11] Xie, W.-C., 2003, "Moment Lyapunov Exponents of a Two-Dimensional System Under Bounded Noise Parametric Excitation," J. Sound Vib., **263**(3), pp. 593–616.
[12] Talay, D., 1991, "Approximation of Upper Lyapunov Exponents of Bilinear Stochastic Differential Systems," SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal., **28**(4), pp. 1141–1164.
[13] Grorud, A., and Talay, D., 1996, "Approximation of Lyapunov Exponents of Nonlinear Stochastic Differential Equations," SIAM J. Appl. Math., **56**(2), pp. 627–650.
[14] Xie, W.-C., 2005, "Monte Carlo Simulation of Moment Lyapunov Exponents," ASME J. Appl. Mech., **72**(2), pp. 269–275.
[15] Khasminskii, R. Z., 1980, *Stochastic Stability of Differential Equations*, Kluwer Academic, Norwell, MA, English translation.
[16] Gnedenko, B. V., and Kolmogorov, A. N., 1954, *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, MA, translated from Russian.
[17] Feller, W., 1965, *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd ed., Wiley, New York.
[18] Komlós, J., Major, P., and Tusnády, G., 1975, "An Approximation of Partial Sums of Independent RV's and the Sample DF. I," Z. Wahrscheinlichkeitstheor. Verwandte Geb., **32**, pp. 111–131.
[19] Komlós, J., Major, P., and Tusnády, G., 1976, "An Approximation of Partial Sums of Independent RV's and the Sample DF. II," Z. Wahrscheinlichkeitstheor. Verwandte Geb., **34**, pp. 33–58.
[20] Kloeden, P. E., and Platen, E., 1992, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.
[21] Talay, D., and Tubaro, L., 1990, "Expansion of the Global Error for Numerical Schemes Solving Stochastic Differential Equations," Stoch. Anal. Appl., **8**(4), pp. 483–509.

# Marginal Instability and Intermittency in Stochastic Systems—Part II: Systems With Rapid Random Variations in Parameters

**M. F. Dimentberg**
Department of Mechanical Engineering,
Worcester Polytechnic Institute,
100 Institute Road,
Worcester, MA 01609

**A. Hera**
Information Technology Division,
Worcester Polytechnic Institute,
100 Institute Road,
Worcester, MA 01609

**A. Naess**
Centre for Ships and Ocean Structures,
and Department of Mathematical Sciences,
Norwegian University of Science and Technology,
NO-7491 Trondheim, Norway

*Dynamic systems with lumped parameters, which experience random temporal variations, are considered. The variations "smear" the boundary between the system's states, which are dynamically stable and unstable in the classical sense. The system's response within such a "twilight zone" of marginal instability is found to be of an intermittent nature, with alternating periods of zero (or almost-zero) response and rare short outbreaks. As long as it may be impractical to preclude completely such outbreaks for a designed system, subject to highly uncertain dynamic loads, the corresponding system's response should be analyzed. Results of such analyses are presented for cases of slow and rapid (broadband) parameter variations in Papers I and II, respectively. The former case has been studied in Paper I (2008, "Marginal Instability and Intermittency in Stochastic Systems—Part I: Systems With Slow Random Variations of Parameters," ASME J. Appl. Mech., 75(4), pp. 041002) for a linear model of the system using a parabolic approximation for the variations in the vicinity of their peaks (so-called Slepian model) together with Krylov–Bogoliubov averaging for the transient response. This resulted in a solution for the probability density function (PDF) of the response, which was of an intermittent nature indeed due to the specific algorithm of its generation. In the present paper (Paper II), rapid broadband parameter variations are considered, which can be described by the theory of Markov processes. The system is assumed to operate beyond its stochastic instability threshold—although only slightly—and its nonlinear model is used accordingly. The analysis is based on the solution of the Fokker–Planck–Kolmogorov partial differential equation for the relevant stationary PDF of the response. Several such PDFs are analyzed; they are found to have integrable singularities at the origin, indicating an intermittent nature of the response. Asymptotic analysis is performed for the first-passage problem for such response processes with highly singular PDFs, resulting in explicit formulas for an expected time interval between outbreaks in the intermittent response.*
[DOI: 10.1115/1.3086593]

## 1 Introduction

Classical definitions of stability and instability deal with long-term behavior of dynamic systems, that is, behavior at $t \to \infty$. These definitions are quite adequate for numerous engineering applications where the long-term operation of the systems is indeed required. However, they may become not perfectly appropriate for applications with limited service life, such as missiles and projectiles. In such applications a system may sometimes be qualified as acceptable in spite of being unstable in the classical sense as long as its deviations from the design state are small enough for a successful and safe operation. Thus design of such marginally unstable systems may be based on the analysis of their transient response within a limited service life.

The classical definitions of stability and instability may also prove to be not perfectly adequate for another class of dynamic systems—those that may be intended for long-term operation. Such systems are designed, as a rule, to operate within their stability domain in the classical sense as long as the "nominal" design parameters are considered. However, if the system's parameters may experience random temporal variations around their nominal or expected values, the system may become "temporarily unstable." Occasionally, this may happen whenever the ("classical") instability boundary is crossed. Whenever a complete elimination of such excursions of the system out of its stability domain may lead to impossible or impractical design, the corresponding short-time outbreaks in response should be analyzed to evaluate the system's reliability.

The resulting dynamic response of a system may be expected to be of an intermittent nature whereby relatively long periods of zero or almost-zero responses alternate with relatively rare spontaneous short-period high-level outbreaks. The name "intermittency" is used in fluid mechanics for transitional regimes between laminar and turbulent flows. The complete transition usually requires a *finite* change in the basic control parameter, such as the Reynolds number, and within the transitional range sporadic alternations, or on/off switching, between laminar and turbulent flow patterns can be observed. More recently the name was applied, by analogy, to dynamic systems with lumped parameters that exhibit the potential for such a "noninstantaneous" transition to chaos—with sporadic outbreaks in chaotic response within the transitional zone and zero or almost-zero response between the outbreaks; for an extensive survey of the topic, see Ref. [1].

In this paper analytical studies are presented for intermittency in systems with lumped parameters subject to *externally imposed*

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 8, 2007; final manuscript received October 8, 2008; published online March 5, 2009. Review conducted by Igor Mezic.

*random variations in parameters*. These variations may "smear" the stability boundary of a given system, that is, expand it into the transitional *zone*; the system may then become marginally unstable in the classical sense. Within such a zone, spontaneous high-level and relatively rare outbreaks alternate with periods of almost-zero response. This behavior of the response may also be qualified as being intermittent indeed. Its analysis may be of importance for engineering as long as it may be impractical for some applications to completely preclude the system's operation within the transitional zone; this may be the case, for example, with a structure subject to wind loads due to hurricanes and/or wave loads in severe storms. Thus, problems of first-passage failure and/or of low-cycle fatigue may be of concern for a system operating within the transitional state. These problems can be treated by existing methods for an analysis of random vibrations, which provide the possibility for estimating the system's reliability in the *design analysis*. Furthermore, relevant dynamic studies may also be of importance for the *interpretation* of test results for a machine or structure where intermittent behavior of the response is observed.

This paper presents results of such dynamic response analyses for two distinctly different mechanisms for generating response in systems with randomly varying parameters; in both cases, however, the response is manifestly intermittent. In Paper I [2] cases of slow parameter variations (compared with the system's lowest natural frequency) have been considered, with the nominal or expected state of the system being stable. Thus, spontaneous outbreaks in response were assumed to be essentially due to temporary excursions of the system's parameters into the instability domain. Response PDF analysis has been made using a parabolic approximation for parameter variations in the vicinities of peaks together with Krylov–Bogoliubov (KB)-averaging, which resulted in a solution for the transient response for linear models of the systems involved.

In the present Paper II, systems with "rapid" stationary random temporal variations in parameters are studied. The name is used for variations that are broadband with respect to the system or may even be approximated by white noises. These systems may be described by the theory of Markov processes [3–5], and they have clearly defined boundaries corresponding to various definitions of stochastic stability for the system's linear part [4]. The systems with such parameter variations operate *within the domain of stochastic instability* but very close to the corresponding instability threshold; this may happen in mechanical engineering because of high uncertainty of predictions for stochastic instability, whereas such a situation may be essentially natural in other applications, e.g., in population dynamics. The response is found to be of the intermittent nature indeed in such cases. It is clear that adequate modeling for this kind of steady-state response requires the system's *nonlinearity* to be accounted for. And the analysis provides the potential for predicting the response PDFs through the solution of the stationary Fokker–Planck–Kolmogorov (FPK) partial differential equation [3–5]. Several such stationary PDFs are analyzed; all of them are found to possess an integrable singularity at the origin, whereas the response itself does exhibit the intermittency indeed. Common characteristic features of these solutions are also certain other typical patterns of a stationary intermittent response; for example, if $A(t)$ with mean $\langle A\rangle$ and standard deviation $\sigma_A$ (where angular brackets denote probabilistic averaging) denotes the amplitude process, then typically a small relative stay time of $A(t)$ above $\langle A\rangle$ is observed, and also the inequality $\sigma_A \gg \langle A\rangle$. Asymptotic analysis is also performed for the first-passage problem for certain response processes with highly singular PDFs. This analysis provides explicit formulas for an expected time interval between outbreaks in the intermittent response.

## 2 Quasilinear Systems With Nonlinear Damping

The equation of motion for the single degree-of-freedom (SDOF) system to be studied may be written as

$$\ddot{X} + h(X,\dot{X}) + \Omega^2 X[1+\xi(t)] = 0 \tag{1}$$

where $\xi(t)$ is a stationary zero-mean broadband random process with power spectral density (PSD) $\Phi_{\xi\xi}(\omega)$. The latter is assumed to be proportional to a small parameter $\varepsilon$, and the same assumption is adopted for the nonlinear damping function $h$. Therefore the response $X(t)$ should be narrow-band, and its slowly varying amplitude $A(t)$ may be introduced as $X=A\sin\theta$, $\dot{X}=\Omega A\cos\theta$, $\theta=\Omega t+\phi$. This amplitude satisfies the following "shortened" first-order stochastic differential equation (SDE) as obtained by the asymptotic method of stochastic averaging [3–5]:

$$\begin{aligned}\dot{A} &= -\Omega^{-1}h_1(A) + 3B_\xi A - A\sqrt{2B_\xi}\bar{\xi}(t) \quad \text{where } B_\xi \\ &= (\pi\Omega^2/8)\cdot\Phi_{\xi\xi}(2\Omega) \text{ and } h_1(A) \\ &= (2\pi)^{-1}\int_0^{2\pi} h(A\sin\theta,\Omega A\cos\theta)\cos\theta d\theta\end{aligned} \tag{2}$$

Here $\bar{\xi}(t)$ is an "equivalent" standard zero-mean Gaussian white noise of unit intensity.

The FPK equation for the PDF $p(A)$ of the amplitude $A(t)$ as governed by the first-order SDE (Eq. (2)) is an ordinary differential equation (ODE) rather than a PDE, and it has the following steady-state solution:

$$p(A) = 2CA\exp\left[-\int_0^A (h_1(z)/\Omega B_\xi z^2)dz\right] \tag{3a}$$

where $C$ is a normalization constant. For the case of smooth $h_1(A)$, with $h_1(0)\neq 0$, this solution may have a singularity at its origin $A=0$. In particular, for the special case of a cubic damping nonlinearity where $h(X,\dot{X})=2\alpha\dot{X}+8\beta_1X^2\dot{X}$ and $\Omega^{-1}h_1(A)=\alpha A+\beta_1A^3$, the PDF (Eq. (3*a*)) is reduced, after imposing the relevant normalization condition, to the following form as obtained originally in Ref. [3]:

$$\begin{aligned}p(A) &= 2\beta^{1-\delta}A^{1-2\delta}\exp(-\beta A^2)/\Gamma(1-\delta) \quad \text{where } \delta=\alpha/2B_\xi, \quad \beta \\ &= \beta_1/2B_\xi\end{aligned} \tag{3b}$$

Here $\Gamma$ is a gamma-function. This PDF leads to the gamma-distribution for the scaled squared amplitude $v=\beta V=\beta A^2$. In particular, its PDF becomes

$$w(v) = v^{-\delta}\exp(-v)/\Gamma(1-\delta) \tag{4}$$

with mean value and standard deviation $m_v=1-\delta$ and $\sigma_v=(1-\delta)^{1/2}$, respectively. Both PDFs (3*b*) and (4) have singularities at their origins as long as the damping coefficient $\alpha$ is positive. They do exist only if these singularities are integrable, that is, if $\delta<1$; this condition for integrability of the PDFs is known to be the same as that for stochastic instability in the probability of the linear part of system (1): zero response is stable if $\delta>1$.

The PDF (3*b*) with an integrable singularity has been obtained as far back as in the late 1950s [3], with the name "undeveloped oscillations" being used to discriminate the corresponding response from that in the case of self-excited oscillations in a system with negative $\alpha$. However, a clear connection between the intermittent pattern of the response itself and the singularity in its PDF by direct numerical simulation was established only much later [6]. A relative stay time of a stationary gamma-distributed process $v(t)$ above its mean level has been suggested in Ref. [5] as a useful index of intermittency, which may be estimated in terms of its PDF as

$$\lambda_v = \frac{\text{Prob}(v > m_v)}{\text{Prob}(v < m_v)} = \frac{\int_{m_v}^{\infty} w(v)dv}{\int_0^{m_v} w(v)dv} = \frac{\Gamma(1-\delta,1-\delta)}{\Gamma(1-\delta)-\Gamma(1-\delta,1-\delta)} \tag{5}$$

Here the function $\Gamma(\cdot,\cdot)$, which depends on two arguments, is the incomplete gamma-function, and in view of its properties one may indeed regard a small value of the ratio $\lambda_v$ as an index for intermittency as long as $\lambda_v \to 0$ with $1-\delta \to 0$. Yet another index for intermittency is a large value of the ratio $\sigma_v/m_v$, which equals $(1-\delta)^{-1/2}$ for the process with PDF (4). Indeed, as long as $v(t)$ is ergodic, one may write

$$m_v = \lim_{T\to\infty} \frac{1}{T}\int_0^T v(t)dt \quad \text{and} \quad \sigma_v^2 = \lim_{T\to\infty} \frac{1}{T}\int_0^T [v(t)-m_v]^2 dt$$

For intermittent $v(t)$ a significant contribution to both integrals is provided by integration within the restricted time intervals of response outbreaks with their total duration being $T_{\text{out}}$, and as long as $v^2(t)$ and $v(t)$ are finite the ratio $\sigma_v/m_v$ should indeed be proportional to $\sqrt{T}$ for fixed $T_{\text{out}}$.

Besides the one-dimensional PDF, it may also be of interest to predict the mean or expected time between "outbreaks" of the intermittent process. The relevant analysis may be based on the solution to the first-passage problem for $T(V)$—the expected time until the first passage of a given barrier $V^*$, starting from an initial value $V<V^*$, where $V=A^2$ and $A(t)$ are governed by the SDE (2). The solution for $T(A)$, based on the theory of the backward Kolmogorov equation, is presented in the Solution Manual for the book [4]; it may be rewritten as

$$T(V) = (1/4B_\xi)\int_V^{V_*} z^{\delta-2}e^{\beta z} I(z)dz$$

$$I(z) = \int_0^z z'^{-\delta}e^{-\beta z'}dz' = \frac{1}{\beta^{1-\delta}}\int_0^{\tilde{z}} \tilde{z}'^{-\delta}e^{-\tilde{z}'}d\tilde{z}' = \frac{1}{\beta^{1-\delta}}\gamma(1-\delta,\tilde{z}) \tag{6}$$

where $\gamma(1-\delta,x)=\Gamma(1-\delta)-\Gamma(1-\delta,x)$.

Now, for small $1-\delta$ (case of intermittency) the approximation $\gamma(1-\delta,\tilde{z}) \cong \tilde{z}^{1-\delta}/(1-\delta)$ can be used in Eq. (6) [7,8]. Then

$$\alpha T(v) \cong \frac{1}{2}\cdot\frac{\delta}{1-\delta}\int_v^{v_*}\frac{e^z}{z}dz = \frac{1}{2}[\delta/(1-\delta)][Ei(v_*)-Ei(v)], \quad v = \beta V, \quad v_* = \beta V_* \tag{7}$$

where $Ei$ is the exponential integral function. This solution may be used to calculate the expected time interval between outbreaks as long as the definition of "outbreak" is introduced, say, as that of upcrossing a certain given level $V_*=\mu\langle V\rangle$ by $V(t)$ (choice of the nondimensional coefficient $\mu$ is application dependent).

The function $T(V)$ (Eq. (7)) may now be averaged over the stationary PDF (4) of $V(t)$, as suggested in Ref. [3], to estimate the mean or expected period between "clusters of level crossings." Using the adopted threshold for outbreaks in $V(t)$, that is, $\beta V_* = \mu\beta\langle V\rangle = \mu(1-\delta)$, and applying the approximation $Ei(x) \cong \gamma + \ln x + x$ (where $\gamma=0.5772\ldots$ is Euler's constant) for small values of the argument of $Ei$ [7,8] yields

$$\begin{aligned}\alpha\langle T\rangle_\mu &= \alpha\int_0^{\mu\langle v\rangle} T(v)w(v)dv \\ &= \frac{\delta}{2(1-\delta)\Gamma(1-\delta)}\int_0^{\mu(1-\delta)} v^{-\delta}e^{-v}[Ei(1-\delta)-Ei(v)]dv \\ &\simeq \frac{\delta}{2\Gamma(2-\delta)}\int_0^{\mu(1-\delta)} v^{-\delta}e^{-v}\cdot[\ln((1-\delta)/v)+1-\delta-v]dv \\ &\cong \frac{\mu^{1-\delta}\delta(1-\delta)^{1-\delta}}{2\Gamma(2-\delta)}\cdot\left[\frac{1}{2-\delta}+\frac{1}{(1-\delta)^2}\right] \\ &\cong \frac{\mu^{1-\delta}\delta}{2(1-\delta)^{1+\delta}\Gamma(2-\delta)} \cong \frac{\mu^{1-\delta}}{2(1-\delta)^{1+\delta}}\end{aligned} \tag{8}$$

The last row in Eq. (8) is based on the approximation $\exp(-v) \cong 1$ for $v \in [0, 1-\delta]$ and small $1-\delta$ in the integrand. This estimate for the expected time for response outbreak is seen to be independent of the system's nonlinearity.

An extensive Monte Carlo simulation study has been performed to verify the theoretical solutions (4) and (8) for the PDF $w(v)$ and the expected period between outbreaks, respectively. The response of system (1) has been simulated numerically for $\Omega=2$ and $\alpha=0.01$, whereas $\xi(t)$ was generated as a stationary zero-mean broadband random process with constant power spectral density $\Phi_{\xi\xi}(\omega)$ within a frequency range of $\omega<8\Omega$. Values of $\Phi_{\xi\xi}(\omega)$ were adjusted to assign various values for $\delta$ between 0.80 and 0.95. To test the insensitivity of the results to the system's nonlinearity, three different values of $\beta_1$ were considered in the numerical simulation, namely, 0.05, 0.02, and 0.005. Most of the simulations were conducted for a time interval equal to about 300,000 natural periods of the corresponding linear system, which contained about a thousand response outbreaks if $\delta<0.9$. A short sample of $x(t)$ that illustrates the response outbreaks is shown in Fig. 1(*a*).

Due to the limited accuracy of the numerical integration, it was expected that computational noise may affect the PDF of the scaled amplitude $v$ for very small values of $v$, thereby influencing the estimate of the expected period between outbreaks. This is illustrated in Fig. 1(*b*) where the theoretical solution (4) for the PDF $w(v)$ is compared with the computed PDF of the simulated response for $\delta=0.82$. The PDF of the simulated response seems to follow closely the theoretical curve down to a rather small value $v=4\times10^{-8}$. This level of accuracy seems to be adequate for some applications—such as for predicting fatigue life, especially with nonzero endurance limit. However, as will be seen from the following, the discrepancies between the theoretical and simulated $w(v)$ at small $v$'s may strongly influence the estimate of the period between outbreaks when used in the integrand in Eq. (8).

After extensive numerical simulations, it was found that for $\delta<0.9$, $w(v)$ can be approximated as uniform within the range of $v \in (v_0, v_1)$ and zero for $v<v_0$. In this paper the corrected formula for the PDF of the scaled squared amplitude is suggested as

$$w_C(v) = \begin{cases} Cv^{-\delta}\exp(-v)/\Gamma(1-\delta), & v > v_1 \\ Cv_1^{-\delta}\exp(-v_1)/\Gamma(1-\delta), & v \in [v_0, v_1] \\ 0, & v < v_0 \end{cases} \tag{4'}$$

Here $C$ is a normalization constant, whereas the two new parameters $v_0$ and $v_1$ depend on $\delta$ in general. Thus, for the "measured" PDF shown in Fig. 1(*b*), $v_0=1\times10^{-12}$ and $v_1=4\times10^{-8}$. To increase the accuracy of the predicted $\langle T\rangle$, expression (8) can be corrected using the measured PDF of $v$. In practical applications, this can be obtained from previous monitoring data if the environmental conditions are stationary, and anyway, numerical estimates for the response PDF may usually require less effort than those for the period between outbreaks.

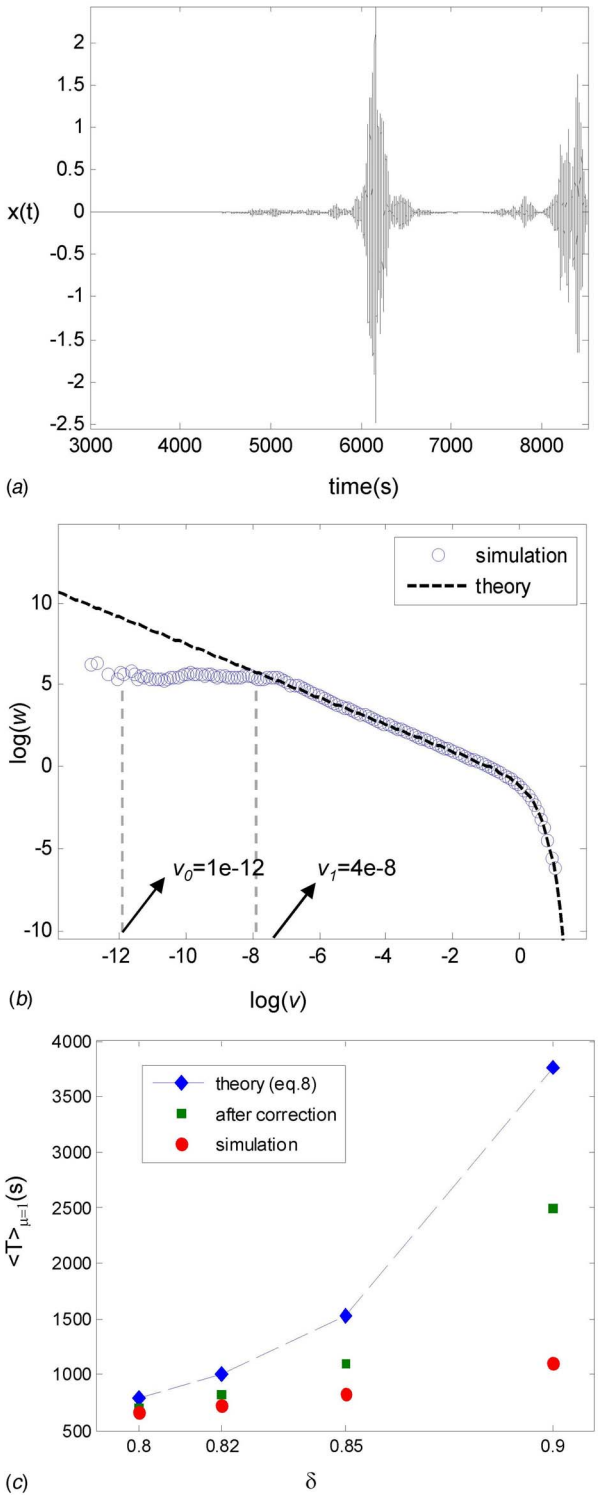



**Fig. 1** (*a*) **A short sample of response** $x(t)$ **illustrating the response outbreaks. System parameters are** $\Omega=2$, $\alpha=0.01$, $\beta_1=0.02$, **and** $\delta=0.82$. (*b*) **PDF of** $v$, $w(v)$, **predicted by theory (Eq. (4)) and calculated from simulation data** ($\delta=0.82, \beta_1=0.02$). **To emphasize the difference between the two plots, only the PDF for** $v\in[1\times10^{-4}, 1\times10^{2}]$ **has been used in the normalization of the results obtained from simulation data.** (*c*) **Expected period between outbreaks as shown by Eq. (8), Eq. (8) with correction for** $w(v)$, **and numerical simulation. The errors between numerical simulation and predicted results using the corrected PDF of** $v$ **are 5%** ($\delta=0.8$), **11%** ($\delta=0.82$), **23%** ($\delta=0.85$), **and 50%** ($\delta=0.90$). $\beta_1=0.02$.

Figure 1(*c*) compares three estimates of the expected period between outbreaks $\langle T\rangle\equiv\langle T\rangle_{\mu=1}$: theoretical Eq. (8) as based on solution (4) for $w(v)$, adjusted Eq. (8) with the integral calculated using the corrected $w(v)$ (Eq. (4′)), and numerical simulation. The correction of the PDF $w(v)$ is seen to reduce the expected time between outbreaks. The simulation results are close to those obtained by the corrected formula as long as the system is relatively deep within the instability domain although it still exhibits some intermittency, i.e., $\delta<0.9$. For $\delta\geq0.9$, the contribution of very small values of the response, of the order of numerical accuracy, into the system's behavior is significant, which explains the difference between simulation results and theory. Furthermore, additional simulations for different values of $\beta_1$ have confirmed that the expected time for response outbreak is independent of the system's nonlinearity, as implied by Eq. (8).

Concluding this section we may claim acceptable accuracy of the adjusted theoretical Eq. (8) for $\langle T\rangle$ as based on the use of the corrected PDF $w(v)$ in the integrand for $\delta<0.9$. Cases with $\delta\geq0.9$ would require longer samples of $x(t)$ with adequate accuracy to estimate both $w(v)$ and $\langle T\rangle$. Furthermore, some convention is needed to handle "irregular" cycles of $x(t)$—i.e., those with negative peaks and/or positive troughs—which may lead to ambiguity in evaluating $\langle T\rangle$.

## 3 Vibroimpact System With One-Sided Barrier

Consider now a system with a quite different type of nonlinear damping, which may effectively restrict response growth in case of stochastic instability. This is a single-barrier SDOF vibroimpact system with inelastic impacts as governed by the equation of motion for its displacement $Y(t)$,

$$\ddot{Y}+2\alpha\dot{Y}+\Omega^2Y[1+\xi(t)]=0 \quad \text{for } Y>-h \tag{9}$$

and the impact condition at the barrier at $Y=-h$,

$$\dot{Y}_+=-r\dot{Y}_- \quad \text{where } \dot{Y}_\pm=\dot{Y}(t_*\pm0),\ Y(t_*)=-h, \text{ and } 0<r\leq1 \tag{10}$$

Here $r$ is a restitution factor, whereas $\xi(t)$ is a stationary zero-mean Gaussian white noise with intensity $D_\xi=2\pi\Phi_{\xi\xi}=16B_\xi/\Omega^2$. The latter is assumed to be proportional to a small parameter, and so are the damping parameters $\alpha$ and $1-r$. The system (9), (10) has been analyzed using these assumptions in Ref. [9] (see also Ref. [5]) by a quasiconservative version of the stochastic averaging method [4,5]. The total response energy $H(t)=\frac{1}{2}(\dot{X}^2+\Omega^2X^2)$ was introduced accordingly, and the right hand side of the resulting first-order SDE for $H(t)$ was averaged over the (energy-dependent) natural period of system (9), (10) with $\varepsilon=0$. The stationary PDF of energy $p(H)$ was then derived as the solution to the corresponding FPK equation. We shall not consider this PDF in detail here since for the present purpose it is sufficient to note that it does exist (is normable) if and only if $\delta<1$ and $\Omega(1-r)/\pi\alpha>(1-\delta)/\delta$ [5]. The first of these conditions is the same as for system (1), whereas the second one guarantees convergence of the normalization integral of $p(H)$ at infinity so that the impact losses are sufficient for restricting response growth in the stochastically unstable system.

Both the SDE for $H(t)$ and the solution to the corresponding FPK equation for $p(H)$ are described by different analytical expressions for $H<H_h$ and for $H>H_h$, where $H_h=\Omega^2h^2/2$ is the system's potential energy at the barrier [5,9]. The first of the expressions for the PDF of energy is $p(H)=C_1H^{-\delta}$ for $H<H_h$, that is, for motions without impacts. It may be regarded as the conditional PDF $p_c(H)=p(H|H<H_h)$ if normalized within $[0, H_h]$ (the condition being $H<H_h$). Imposing this condition yields

$$p_c(H) = [(1-\delta)/H_h](H/H_h)^{-\delta} \tag{11}$$

The (conditional) mean and relative stay time above this mean of the response energy may be found from Eq. (11) as

$$\langle H\rangle = H_h\int_0^1 (1-\delta)\bar{H}^{-\delta}d\bar{H} = H_h[(1-\delta)/(2-\delta)], \quad \bar{H} = H/H_h$$

$$\lambda_H = \frac{\int_{\langle H\rangle}^{H_h} p_c(H)dH}{\int_0^{\langle H\rangle} p_c(H)dH} = \left(\frac{2-\delta}{1-\delta}\right)^{1-\delta} - 1 \tag{12}$$

The expression for $\lambda_H$ clearly indicates that the response should be intermittent indeed for small $1-\delta$.

Consider now the expected time $T(H)$ for reaching a given energy barrier $H^*$ "from below," which is from an initial state with energy $H<H^*$. It is obvious that formula (6) with $\beta=0$ (and the first expression for $I(z)$) should be directly applicable as long as $H=\Omega^2V/2$ provided $H_*<H_h$ so that the system remains linear. Thus

$$T(H) = \frac{1}{4B_\xi}\int_H^{H_*} z^{\delta-2}dz\int_0^z z'^{-\delta}dz' \quad \text{so that} \quad \alpha T(H)$$

$$= \frac{1}{2}\cdot\frac{\delta}{1-\delta}\cdot\ln\left(\frac{H_*}{H}\right) \tag{13}$$

And now, similar to the derivation of Eq. (8) we average this expression over $H$ using the conditional PDF (11) and substituting the selected fraction of the conditional mean $\langle H\rangle$ for the threshold $H^*$, that is, $H^*=\mu\langle H\rangle$. Thus

$$\alpha\langle T\rangle_\mu = \alpha\int_0^{\mu\langle H\rangle} T(H)p_c(H)dH = -\frac{1}{2}\delta\left(\frac{\mu\langle H\rangle}{H_h}\right)^{1-\delta}\int_0^1 y^{-\delta}\ln y dy$$

$$= (\delta/2)\mu^{1-\delta}(1-\delta)^{-(1+\delta)}(2-\delta)^{-(1-\delta)} \tag{14}$$

This estimate for the expected period between outbreaks is once again seen to be independent of the system's nonlinearity; moreover, up to higher-order terms in $1-\delta$, it coincides with the similar estimate (Eq. (8)) for the quasilinear system.

## 4 Vibroimpact System With Double-Sided Barrier

The next case of intermittency is considered for the system with *linear* damping but with a special type of nonlinearity in the restoring force that may provide a restriction of response growth for the case of stochastic instability. The equation of motion between impacts is

$$\ddot{Y} + 2\alpha\dot{Y}[1+\eta(t)] + \Omega^2 Y[1+\xi(t)] = 0 \quad \text{for} \quad -h < Y < h \tag{15}$$

where $\xi(t)$ and $\eta(t)$ are stationary zero-mean independent Gaussian white noises with intensities $D_\xi$ and $D_\eta=(2\alpha/\Omega)^2D_\xi$, respectively. Impact condition (10) with $r=1$ (elastic impact) is now imposed both at $Y=-h$ and at $Y=+h$. Under the above relation between intensities of excitations, the FPK equation for the joint PDF of response displacement and velocity has an exact stationary solution [10],

$$w(y,\dot{y}) = C/(y^2+\dot{y}^2/\Omega^2)^\delta \quad \text{where} \quad \delta = 2\alpha/D_\xi\Omega^2 + 1/2 \tag{16}$$

where $C$ is a normalization constant. It can be clearly seen that in the absence of barriers the stationary PDF (16) cannot exist since the normalization integral within the whole plane $(y,\dot{y})$ diverges at infinity in the present case of a linear system, which is stochastically unstable in probability if $\delta<1$. However, the nonlinearity due to the double-sided barrier may provide restriction for the response growth even in the absence of nonlinear damping—just through bounds on the apparent level of random parametric excitation as imposed by the barrier. Specifically, if $\frac{1}{2}<\delta<1$, then the PDF (16) is normalizable indeed and integration over $\dot{y}$ provides the PDF $p(y)$ of the displacement,

$$p(y) = \int_{-\infty}^{\infty}\frac{Cdv}{(y^2+v^2/\Omega^2)^\delta} = \frac{1-\delta}{h(y/h)^{2\delta-1}} \quad \text{for} \quad -h<y<h$$

$$C = \frac{(1-\delta)\Gamma(\delta)}{\Omega\sqrt{\pi}\cdot\Gamma(\delta-1/2)h^{2(1-\delta)}} \tag{17}$$

(Tables [8] were used, and the result is then normalized within $(-h/2,h/2)$.) The singularity at the origin in this PDF is seen to be integrable as long as $\delta>1/2$.

We may consider relative stay time $\lambda_{|Y|}$ of the magnitude of $Y(t)$ above its mean level as an index of intermittency. Thus

$$\langle|Y|\rangle = 2\int_0^h yp(y)dy = \frac{2(1-\delta)}{3-2\delta}\cdot h, \quad P = \text{Prob}\{Y<|Y|\}$$

$$= 2\int_0^{\langle|Y|\rangle} p(y)dy = [2(1-\delta)/(3-2\delta)]^{2(1-\delta)}$$

$$\text{and} \quad \lambda_{|Y|} = (1-P)/P \tag{18}$$

This index of intermittency is once again seen to approach zero with $1-\delta\to 0$.

Solutions (16) and (17) may be used to calculate the probability for staying in the nonimpact regime by integrating it in polar coordinates within a circle in the plane $(y,\dot{y})$. Thus

$$P_{\text{nonimp}} = C\Omega\int_0^{2\pi} d\theta\int_0^h \rho^{-2\delta}\rho d\rho = \frac{\sqrt{\pi}\Gamma(\delta)}{\Gamma(\delta-1/2)} \tag{19}$$

This probability is found to be close to unity for small $1-\delta$.

## 5 Lotka–Volterra System in Random Environment

Random oscillations are considered for the following stochastic Lotka–Volterra (LV) model [5,11,12]:

$$\dot{u} = -mu + k\beta uv, \quad \dot{v} = \alpha v[1+\xi(t)] - \beta uv - \gamma v^2, \quad \langle\xi(t)\rangle$$

$$= 0, \quad \langle\xi(t)\xi(t+\tau)\rangle = D\delta(\tau) \tag{20}$$

with one of the potential applications being a pair of interacting populations of the predator-prey or parasite-host type. In this case $u(t)$ and $v(t)$ are population sizes of predators (or parasites) and preys (or hosts), respectively, whereas $\xi(t)$ is a zero-mean Gaussian random white noise in the Stratonovich sense with intensity $D$ and $\delta$ is the Dirac delta-function; all coefficients in the SDE (20) are positive. The special case $\gamma=0$, $\xi(t)\equiv 0$ corresponds to the classical LV model with its singular point $u=\alpha/\beta$, $v=m/k\beta$ [12]. This classical model can describe periodic oscillations in population sizes, which are sometimes observed in nature; however, being conservative it cannot "withstand" random environmental variations [5,11,12]. Thus, to study its steady-state random response some kind of "damping" should be accounted for. Such a damping is provided by the interspecies competition term $\gamma v^2$ in the second SDE in Eq. (20). Upon addition of this term, the stable equilibrium point shifts to $u_0=(\alpha-\gamma m/k\beta)/\beta$, $v_0=m/k\beta$ and becomes asymptotically stable (either focus or node), whereas an additional equilibrium state appears at $u_*=0$, $v_*=\alpha/\gamma$. The latter is unstable if $v_o<v_*$ or $\gamma<\gamma_*=\alpha k\beta/m$. At the bifurcation point $\gamma=\gamma_*$, these two equilibrium states merge, with the state $u_*=0$, $v_*=\alpha/\gamma$ becoming stable for $\gamma>\gamma_*$. The physical meaning of these transformations is clear: beyond this transcritical bifurcation point growth of the prey, population is bounded by its interspecies

competition rather than by the activity of predators, and the latter become extinct because of food shortage.

The random variations in the preys' reproduction rate $\alpha$ simulate temporal variations in the environmental conditions. These variations lead to random oscillations in the sizes of both populations, which can be analyzed by using the transformations $x=\ln u$, $y=\ln v$. The SDE (20) is then reduced to the form

$$\dot{x} = -m + k\beta \exp y = \partial H/\partial y, \quad \dot{y} = \alpha - \beta \exp x - \gamma \exp y + \alpha\xi(t)$$
$$= -\partial H/\partial x - g \cdot (\partial H/\partial y) + \alpha\xi(t) \tag{21}$$

where $H(x,y)=k\beta \exp y - my + \beta \exp x - (\alpha - \gamma m/k\beta)x$ and $g=\gamma/k\beta$. The stationary (time-independent) joint PDF $p(x,y)$ of the transformed random state variables $x(t)$ and $y(t)$ as governed by the SDE (21), has been obtained in Ref. [5] as the solution to the corresponding FPK equation. Transformation to the original state variables then provides the joint stationary PDF $w(u,v)$, as defined within the first quadrant of the $(u,v)$-plane as a product of the individual one-dimensional PDFs of the population sizes of $u(t)$ and $v(t)$ [5,11],

$$w(u,v) = p(\Delta u/k) \cdot p(\Delta v) \quad \text{where } p(z) = z^{z_o-1}e^{-z}/\Gamma(z_o) \text{ and } \Delta = 2\gamma/D\alpha^2 \tag{22}$$

Here the universal expression for the PDF $p(z)$ of the gamma-distribution is used with $z=\Delta u/k$ and $z=\Delta v$ for the stationary PDFs of $u(t)$ and $v(t)$, respectively, and $z_o=\Delta u_o/k$ and $z_o=\Delta v_o$. Thus, both steady-state population sizes $u(t)$ and $v(t)$ are independent gamma-distributed random variables with the following mean values and standard deviations:

$$\langle u\rangle = u_0 = (\alpha - \gamma m/k\beta)/\beta = (\alpha/\beta)(1 - v_o/v_*), \quad \langle v\rangle = v_0 = m/k\beta$$
$$\sigma_u = \sqrt{\langle (u - \langle u\rangle))^2\rangle} = \sqrt{u_o k/\Delta}, \quad \sigma_v = \sqrt{\langle (v - \langle v\rangle))^2\rangle} = \sqrt{v_o/\Delta} \tag{23}$$

These PDFs both exist provided that $\gamma < \gamma_*$. If $\gamma > \gamma_*$, then the PDF of the predators' or the parasites' population degenerates into the Dirac delta-function, which implies extinction of predators due to food shortage. It can be seen from relations (22) and (23) and from previous analysis as presented in Sec. 2 that intermittent behavior should be expected of $v(t)$ $(u(t))$ whenever $\Delta v_0 \ll 1 (\Delta u_0/k \ll 1)$, so that $\sigma_v \gg v_o (\sigma_u \gg u_o)$, with rare and short pulselike intensive outbreaks in $v(t)$ $(u(t))$ and low-level oscillations between the pulses. Relevant indices of intermittency such as small relative stay time above the mean (relation (5)) and large value of the ratio $\sigma/m$ may also be used.

It may be convenient for applications to identify two different types of intermittency. The first type is observed only for a population of predators or parasites—when this population is close to extinction, so that $\Delta u_0/k \ll 1$ due to small $u_0$, whereas $\Delta$ is of the order of unity. This type I of intermittency in the proximity of the bifurcation point, i.e., for small $1 - v_o/v_*$ is similar to those considered in Secs. 2–4 (for systems close to their instability boundaries).

The other type of intermittency is observed in the case of high intensity $D$ of the environmental parameter variations, which lead to a small value of $\Delta$. It may be explained by the fact that the natural period $T(H)$ of the corresponding conservative (classical) LV system—i.e., one with $\gamma=0$, $\xi(t)\equiv 0$—increases indefinitely with $H$. If both $\gamma$ and $D$ are proportional to a small parameter, this case may also be analyzed by the quasiconservative version of the stochastic averaging method for a more general version of the SDEs (20)—if the product terms $uv$ are replaced by a more sophisticated interaction law $uV(v)$ [5].

It goes without saying that actually either types or mechanisms of intermittency may be involved for any given LV system. This is particularly clear from an analysis of yet another index for the on-off intermittency in the predator population—an expected number $n_+(u)$ of upcrossings per unit time of a given arbitrary level $u$ by $u(t)$ (crossings of the given level with positive derivative $\dot{u}$). Substituting Eq. (22), together with the first relation in Eq. (20), into the basic relation [5] for upcrossings yields [5,11]

$$n_+(u) = \int_0^\infty \dot{u} w(u,\dot{u}) d\dot{u} = \int_{v_o}^\infty k\beta u (v - v_0) w(u,v) dv = (k\beta/\Delta) \times (\Delta v_0)^{\Delta v_o} (\Delta u/k)^{\Delta u_o/k} \exp(-\Delta v_0 - \Delta u/k) \times [\Gamma(\Delta v_0)\Gamma(\Delta u_0/k)]^{-1} \tag{24}$$

The latter formula provides just the expected circular frequency of oscillations (in hertz) if upcrossings of the mean or expected level $(u=u_0)$ are considered, whereas its reciprocal is the expected time interval $\langle T\rangle$ between the upcrossings,

$$n_+(u_0) = n_\infty f(\Delta v_0) f(\Delta u_0/k), \quad f(z) = [(2\pi)^{1/2} z^{z-1/2} \exp(-z)]/\Gamma(z) \quad \text{with } \lim_{z\to\infty} f(z) = 1$$
$$n_\infty = \lim_{\Delta\to\infty} n_+(u_0) = \beta(k u_0 v_0)^{1/2}/2\pi = (\sqrt{\alpha m}/2\pi) \cdot \sqrt{1 - v_0/v_*} \quad \text{and} \quad \langle T\rangle = 1/(n_+(u_0)) \tag{25}$$

Here $n_\infty$ can be clearly identified as the system's "damped" natural frequency of oscillations with small deviations in $u(t)$ and $v(t)$ from their expected values. The first cofactor in expression (25) for $n_\infty$ is the natural frequency of the corresponding conservative or classical LV system [12], whereas the second cofactor represents a decrease in the natural frequency and an increase in the expected response period due to increasing interspecies competition (high $\gamma$). This second cofactor may be responsible for type I intermittency. On the other hand, two cofactors containing the function $f$ describe the effect of an increased response period with increasing intensity $D$ of the environmental variations. They may be responsible for type II intermittency with a high expected response period for small $\Delta$ as long as $f(z)$ decreases roughly as the square root of $z$ for small $z$.

Illustrations of both types of intermittency are shown in Fig. 2. Samples of $u(t)$ and $v(t)$ are obtained by numerical simulation for a system whose classical LV parameters are $m=1$, $\kappa=1$, $\beta=1$, and $\alpha=1$, whereas $\gamma=0.98$ and $D=1$ in Fig. 2(*a*) and $\gamma=0.5$ and $D=20$ in Fig. 2(*b*). Figure 2(*a*) illustrates intermittency of the first type $(u_0=0.01, \Delta=1.98)$ in the population size of predators, $u(t)$. The second type of intermittency ($\Delta=0.05$, $u_0=0.5$, and $v_0=1$) is illustrated in Fig. 2(*b*), where both population sizes, $u(t)$ and $v(t)$, display rear intensive outbreaks.

Finally, it may be of interest to note that formula (23) for the mean values $\langle u\rangle$ and $\langle v\rangle$ and the standard deviation of $v(t)$ may be derived without using solution (22) to the FPK equation but rather by direct application of the SDE calculus (method of moments) to the SDEs (20) and (21). As long as constant steady-state response moments are sought, one may impose conditions of zero expectations of the right hand sides of these SDEs. Indeed, from Eq. (21)

$$\langle e^y\rangle = \langle v\rangle = v_0 = m/k\beta \quad \text{and} \quad \beta\langle e^x\rangle = \beta\langle u\rangle = \alpha - \gamma\langle e^y\rangle = \alpha - \gamma\langle v\rangle = \alpha - \gamma m/k\beta = \beta u_0$$

Then, from Eq. (21) we obtain $\langle uv\rangle = m\langle u\rangle/k\beta = \langle u\rangle\langle v\rangle = u_0 v_0$ and finally

$$\langle v^2\rangle = [\alpha\langle v\rangle + \alpha\langle v\cdot\xi(t)\rangle - \beta\langle uv\rangle]/\gamma = v_0^2 + v_0/\Delta \quad \text{where } \alpha\langle v\cdot\xi(t)\rangle = \tfrac{1}{2}\alpha^2 D\langle v\rangle$$

is a Wong–Zakai correction [4,5] to the second Stratonovich SDE in Eq. (20).

Further attempts to directly predict response moments face the well-known closure problem for nonlinear systems [4,5]. On the other hand, this approach may be of some use for general systems

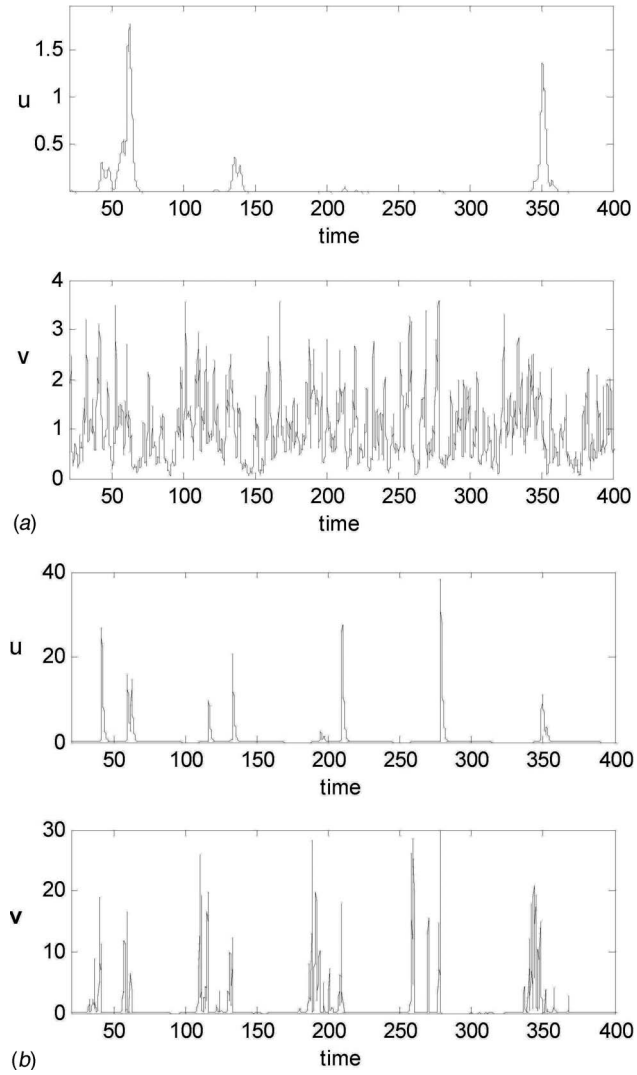



**Fig. 2** (*a*) **Samples of $u(t)$ and $v(t)$ as obtained from numerical simulation for the system with $m=1$, $\alpha=1$, $\beta=1$, $k=1$, $\gamma=0.98$, and $D=1$. It illustrates intermittency of type I ($u_0=0.01$ and $\Delta=1.98$). All quantities are nondimensional.** (*b*) **Samples of $u(t)$ and $v(t)$ as obtained from numerical simulation for the system with $m=1$, $\alpha=1$, $\beta=1$, $k=1$, $\gamma=0.5$, and $D=20$. It illustrates intermittency of type II ($u_0=0.5$, $v_0=1$, and $\Delta=0.05$). All quantities are nondimensional.**

of the LV type with more than two interacting species [12] where an analytical solution to the FPK equation may be unavailable. The present solution for $w(u,v)$ may be helpful for evaluating the accuracy of closure schemes for such cases. On the other hand, Ref. [13] may be referred to here where rigorous theorems were proved regarding moment hierarchy in a certain class of systems with distributed parameters that lead to intermittency.

## 6 Conclusions

The concept of marginal instability has been introduced for systems with random variations in parameters that may smear the system's boundary of dynamic instability. Operation of a system within the corresponding "twilight zone" between domains of stability and "real" instability may require analysis of the corresponding random response to evaluate the system's reliability. Application of relevant methods has been illustrated in this paper, which resulted in procedures for predicting PDFs and expected first-passage times of the response. Cases of slow and rapid random variations in parameters have been considered separately in Papers I [2] and II (this paper), respectively, with drastically different models of the basic systems being used for these cases.

The linear systems with slowly varying parameters, as considered in Paper I, were operating within the stability domain of the corresponding nominal system, i.e., one with mean or expected values of parameters. Thus, relatively rare short-time outbreaks in response were due to brief excursions into the instability domain. Analysis of the corresponding intermittent response was based on a parabolic approximation of parameter(s) in the vicinities of peaks together with the KB-averaging. It resulted in solutions for response PDFs, which may be important for predicting fatigue damage accumulation, whereas the first-passage problem for the response was reduced to that for the randomly varying parameter(s).

In the present paper (Paper II) the rapid random temporal variations in parameter(s) have been considered as described in the framework of the theory of Markov processes. The systems were assumed to operate within the domain of their stochastic instability, and nonlinear systems' models were used accordingly. Response PDFs were obtained as stationary solutions to the corresponding FPK equations, and analyses of these solutions for the case of operation in the vicinity of the stochastic instability boundary were presented. All of the solutions were found to contain an integrable singularity at the origin, indicating the intermittent nature of the response. Other indices for intermittency were also studied. Asymptotic analysis of the first-passage time has also been presented, which resulted in simple expressions for the expected time interval between response outbreaks. Yet another case of intermittency has also been studied for the LV system as used in population dynamics whereby strong "softening" nonlinearity may also become a source of intermittency at high intensity of parameter variations (together with closeness to the bifurcation point).

## Acknowledgment

The first author contributed to this work during his stay at the Centre for Ships and Ocean Structures (CeSOS) of the Norwegian University of Science and Technology (NTNU). The financial support provided by CeSOS is most highly appreciated.

## References

[1] Nayfeh, A. H., and Balachandran, B., 1995, *Applied Nonlinear Dynamics*, Wiley, New York.
[2] Dimentberg, M. F., Hera, A., and Naess, A., 2008, "Marginal Instability and Intermittency in Stochastic Systems—Part I: Systems With Slow Random Variations of Parameters," ASME J. Appl. Mech., **75**(4) p. 041002.
[3] Stratonovich, R. L., 1967, *Topics in the Theory of Random Noise*, Vol. II, Gordon and Breach, New York.
[4] Lin, Y. K., and Cai, G. Q., 1995, *Probabilistic Structural Dynamics: Advanced Theory and Applications*, McGraw-Hill, New York.
[5] Dimentberg, M. F., 1988, *Statistical Dynamics of Nonlinear and Time-Varying Systems*, Research Studies, Taunton, UK.
[6] Ibrahim, R., 1998, "Stabilization and Stochastic Bifurcation, With Application to Nonlinear Ocean Structures," *Stochastically Excited Nonlinear Ocean Structures*, M. F. Shlesinger and T. Swean, eds., World Scientific, Singapore, pp. 1–52.
[7] Abramowitz, M., and Stegun, I., 1972, *Handbook of Mathematical Functions*, Dover, New York.
[8] Gradstein, I. S., and Ryzhik, I. M., 1980, *Tables of Integrals, Series and Products*, Academic, New York.
[9] Dimentberg, M. F., and Menyailov, A., 1979, "Response of a Single-Mass Vibroimpact System to White-Noise Random Excitation," Z. Angew. Math. Mech., **59**, pp. 709–716.
[10] Dimentberg, M. F., 1982, "An Exact Solution to a Certain Nonlinear Random Vibration Problem," Int. J. Non-Linear Mech., **17**(4), pp. 231–236.
[11] Dimentberg, M. F., 2002, "Lotka-Volterra System in a Random Environment," Phys. Rev. E, **65**(3), p. 036204.
[12] Svirejev, Ju. M., and Logofet, D. O., 1978, *Stability of Biological Communities*, Nauka, Moscow, in Russian.
[13] Gärtner, J., and Molchanov, S. A., 1990, "Parabolic Problems for the Anderson Model. I. Intermittency and Related Topics," Commun. Math. Phys., **132**, pp. 613–655.

# A Screw Theory of Timoshenko Beams

**J. M. Selig**
Faculty of Business,
Computing and Information Management,
London South Bank University,
London SE1 0AA, UK
e-mail: seligjm@lsbu.ac.uk

**Xilun Ding**
Robotics Research Institute,
Beijing University of Aeronautics and Astronautics,
Beijing 100083, P. R. China

*In this work, the classic theory of Timoshenko beams is revisited using screw theory. The theory of screws is familiar from robotics and the theory of mechanisms. A key feature of the screw theory is that translations and rotations are treated on an equal footing and here this means that bending, torsion, and extensions can all be considered together in a particularly simple manner. By combining forces and torques into a six-dimensional vector called a wrench, Hooke's law for the Timoshenko beam can be written in a very simple form. From here simple expressions can be found for the kinetic and potential energy densities of the beam. Hence equations of motion for small vibrations of the beam can be easily derived. The screw theory also leads to a new understanding of the boundary conditions for beams. It is demonstrated that simple boundary conditions are closely related to mechanical joints. In order to set up the boundary conditions for a beam attached to a joint, a system of wrenches dual to the screws representing the freedoms of the joint must be found. Finally, a screw version of the Rayleigh–Ritz numerical method is introduced. An example is investigated in which the boundary conditions on the beam lead to vibrational modes of the beam involving bending, torsion, and extension at the same time.* [DOI: 10.1115/1.3063630]

## 1 Introduction

In this work, the well known theory of Timoshenko beams is re-examined from the point of view of screw theory. Screw theory is usually associated with spatial mechanisms and robots, although it dates back to the work of Ball at the end of the 19th century [1]. In robotics and mechanism theory, it has been found that considering translations and rotations together produces concise symbolic equations, which can be easily manipulated.

Euler–Bernoulli and Timoshenko beam theory treat the beam as a stack of elemental plates or sections. Under the effects of a stress, these plates undergo rigid transformations; this approximation is valid for small deflections in many materials. However, the usual analysis treats bending, torsion, and extension/compression separately and hence problems in which the beam experiences several of these deflections at once are hard to treat. Textbook problems are careful to study loading schemes and boundary conditions where only one type of deflection is relevant. However, more general types of loading and boundaries are common, for example, in the members of a spatial mechanism or robot. A natural way to treat these problems seems to be by using screw theory, since screw theory is based on the geometry of the group of rigid-body motions $SE(3)$ and its Lie algebra $se(3)$.

In this work, we show how the deflection of a beam can be modeled by a screw-valued function along the beam. A screw form of Hooke's law for Timoshenko beams is then presented. This introduces the compliance density of the beam. From this, a Lagrangian density for the beam can be derived and hence the equations of motion are easily found. The screw form of some possible boundary conditions is considered next. Then we turn to consideration of a numerical method, the Rayleigh–Ritz method, for computing the vibrational modes and frequencies of the beam. This also has an elegant derivation in terms of screw theory. Finally, we consider a particular example of a simply supported cylindrical beam where the joints supporting the beam at either end have perpendicular axes. Approximations to the first few modal frequencies and shapes are presented.

We begin by reviewing some elementary screw theory.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 18, 2006; final manuscript received October 2, 2008; published online March 5, 2009. Review conducted by Oliver M. O'Reilly.

## 2 Screw Theory

For rigid-body motions, a twist (or screw here) is an element of $se(3)$, the Lie algebra to the group of proper rigid-body motions. This can represent a small displacement of a rigid body or the velocity of the body.

A general screw can be written as $\mathbf{s}^T=(\boldsymbol{\omega}^T,\mathbf{v}^T)$, where $\boldsymbol{\omega}$ is the angular velocity and $\mathbf{v}$ is the linear velocity. So a screw is a combination of the angular velocity and linear velocity of the rigid body. Alternatively, $\mathbf{v}$ could represent a small translation and $\boldsymbol{\omega}$ a small rotation, that is, the components of $\boldsymbol{\omega}$ are small rotations about the $x$, $y$, and $z$ axes.

To each screw, one may associate a line in space, the screw axis, and a scalar $p$, called the pitch of the screw. A line in space is specified by its direction $\boldsymbol{\omega}$ and its moment $\mathbf{r}\times\boldsymbol{\omega}$, where $\mathbf{r}$ is any point on the line. The translational part of the screw $\mathbf{v}$ can be split into a part perpendicular to $\boldsymbol{\omega}$ and a part parallel to $\boldsymbol{\omega}$,

$$\mathbf{v}=\mathbf{r}\times\boldsymbol{\omega}+p\boldsymbol{\omega} \tag{1}$$

The axis of the screw is $\mathbf{l}^T=(\boldsymbol{\omega}^T,(\mathbf{r}\times\boldsymbol{\omega})^T)$, and the pitch of the screw is given by the quantity $(\boldsymbol{\omega}\cdot\mathbf{v})/(\boldsymbol{\omega}\cdot\boldsymbol{\omega})$ so long as $\boldsymbol{\omega}\neq 0$. When $\boldsymbol{\omega}=0$, the screw represents a pure translation and the pitch is said to be infinite. Zero pitch corresponds to pure rotations and positive pitch screws are said to be right handed while negative pitch screws are left handed.

An active transformation will move the screw according to the relation

$$\mathbf{s}'=H\mathbf{s}=\begin{pmatrix}R & 0\\ TR & R\end{pmatrix}\begin{pmatrix}\boldsymbol{\omega}\\ \mathbf{v}\end{pmatrix} \tag{2}$$

where $R$ is the rotation matrix of the transformation and $T$ is the antisymmetric matrix representing the translation ($T\mathbf{x}=\mathbf{t}\times\mathbf{x}$). The transformation matrix $H$ here is an element of the adjoint representation of the group $SE(3)$. The corresponding passive transformation is given by the inverse of the active transformation matrix $H$, as follows:

$$\bar{\mathbf{s}}=H^{-1}\mathbf{s}=\begin{pmatrix}R^T & 0\\ -R^TT & R^T\end{pmatrix}\begin{pmatrix}\boldsymbol{\omega}\\ \mathbf{v}\end{pmatrix} \tag{3}$$

A wrench is also a six-dimensional vector $\mathcal{W}^T=(\mathbf{M}^T,\mathbf{F}^T)$, but this is different from the twist. It is an element of the dual to the Lie algebra. The vector $\mathbf{F}$ is the force and $\mathbf{M}$ is the moment acting on the rigid body.

The generalized momentum of a rigid body, consisting of the linear and angular momentums, also forms an element of the dual to the Lie algebra. The key difference between Lie algebra elements and elements of the dual space is their transformation properties.

The pairing of a wrench and a screw $\mathcal{W}^T\mathbf{s}$ gives a quantity proportional to the energy or power, depending whether the screw represents a displacement or a velocity. These quantities are scalars; they are invariant with respect to rigid transformations. To ensure that the combination is invariant, wrenches must transform under the inverse transpose representation of $H$ for an active transformation and the transpose of $H$ for a coordinate change

$$\mathcal{W}'=H^{-T}\mathcal{W}=\begin{pmatrix}R & TR\\ 0 & R\end{pmatrix}\begin{pmatrix}\mathbf{M}\\ \mathbf{F}\end{pmatrix},\quad \overline{\mathcal{W}}=H^T\mathcal{W}=\begin{pmatrix}R^T & -R^TT\\ 0 & R^T\end{pmatrix}\begin{pmatrix}\mathbf{M}\\ \mathbf{F}\end{pmatrix} \tag{4}$$

The inverse transpose of $H$ is an element of the co-adjoint representation of $SE(3)$.

Wrenches also have axes and pitches: The pitch of a wrench $\mathcal{W}^T=(M^T,F^T)$ is given by $p=(\mathbf{F}\cdot\mathbf{M})/(\mathbf{F}\cdot\mathbf{F})$, and the axis of the wrench is the line with direction $\mathbf{F}$ and moment $\mathbf{M}-p\mathbf{F}$.

More details on the group $SE(3)$ and its applications to robotics and mechanism theory can be found in Ref. [2], for example. In Sec. 3, the application of this screw theory to the deflection of a beam will be developed.

## 3 Bending of Beams

**3.1 The Deflection Screw.** As mentioned above, in this approximation the beam is modeled by a stack of elemental cross sections or plates. Each of these plates may undergo a rigid-body motion. To keep track of them, we place a coordinate frame in each plate, its origin coincident with the centroid of the plate, and its $x$ and $y$-axes of this local frame will be aligned with the principle directions of the local cross section. These coordinate frames will be collectively referred to as the local moving frame. A globally fixed coordinate frame will also be required; for convenience, this will be located at the foot of the beam. In fact, it will be simplest to have the globally fixed frame and the moving frame coincide at the foot of the beam.

Again for simplicity, we will assume that the unstressed beam is straight so that the centerline of the beam lies along the $z$-axis. The passive transformation from the frame in the plate a distance $\mu$ from the foot of the beam to the global frame and is given by the matrix $H^{-1}$

$$H^{-1}=\begin{pmatrix}I_3 & 0\\ \mu T_{\mathbf{k}} & I_3\end{pmatrix} \tag{5}$$

with

$$T_{\mathbf{k}}=\begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \tag{6}$$

So, for example, the centerline of the beam will be given by

$$\overline{\mathbf{p}}_0=\begin{pmatrix}0\\ 0\\ \mu\end{pmatrix}=\mu\mathbf{k} \tag{7}$$

where $\mathbf{k}$ is the unit vector in the $z$ direction, but in the local moving frame this vector is simply

$$\mathbf{p}_0=\mathbf{0} \tag{8}$$

see Fig. 1. Notice that we are using a bar to denote vectors referred to the global frame.

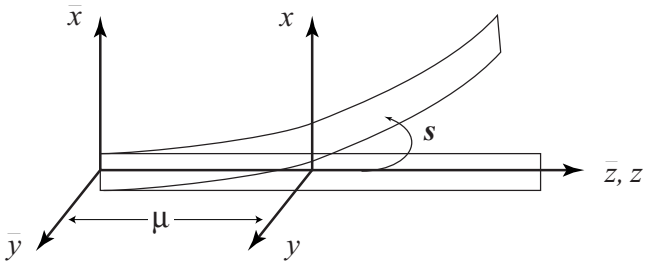



**Fig. 1 Coordinate frames along the beam**

When the beam is deflected, each plate will be subjected to a different rigid motion; the deflection might be modeled using a rigid-body motion at each point. This would lead to a theory that is sometimes referred to as "geometrically exact" beam theory or Cosserat theory; an attempt to describe such a theory in terms of screws was made in Ref. [3].

Here, however, we assume that the deflections are small and so can be described by elements of the Lie algebra. So the state of the beam will be described by a "deflection screw" $\mathbf{s}(\mu)$. The deflection screw specifies the position and orientation of each element in the beam relative to its position and orientation when the beam is undeformed, see Fig. 1. When we consider the dynamics of the beam in a moment, the deflection screw will also depend on time $\mathbf{s}(\mu,t)$. The small rotations of the axes in the stressed state will be written as a vector $\boldsymbol{\theta}$, and the translation vector will be written as $\mathbf{v}$. So the deflection screw, in the local moving frame, at arc length $\mu$ is given by

$$\mathbf{s}(\mu)=\begin{pmatrix}\boldsymbol{\theta}\\ \mathbf{v}\end{pmatrix}=\begin{pmatrix}\theta_x(\mu)\\ \theta_y(\mu)\\ \theta_z(\mu)\\ v_x(\mu)\\ v_y(\mu)\\ v_z(\mu)\end{pmatrix} \tag{9}$$

From screw theory, if a point $\mathbf{q}_0$ is subject to a screw $\mathbf{s}^T=(\boldsymbol{\theta}^T,\mathbf{v}^T)$, its new position will be given by $\mathbf{q}=\boldsymbol{\theta}\times\mathbf{q}_0+\mathbf{v}$. Applying this to the center of the beam in each element, we obtain the centerline curve $\mathbf{p}(\mu)$ given by

$$\mathbf{p}(\mu)=\boldsymbol{\theta}\times\mathbf{p}_0+\mathbf{v}=\mathbf{v} \tag{10}$$

In other words, we may identify the translational part of the deflection screw with the centerline curve of the beam.

Next, we turn to derivatives. The derivative of the deflection screw, expressed in the moving frame, will be denoted $\mathbf{s}'$ and the derivative in the global frame by $d\mathbf{s}/d\mu$. The two are related by the equation

$$\mathbf{s}'=\frac{d}{d\mu}\mathbf{s}+B\mathbf{s} \tag{11}$$

where the Lie algebra element $B$ is given by

$$B=H\frac{d}{d\mu}H^{-1}=-\frac{d}{d\mu}H\ H^{-1}=\begin{pmatrix}0 & 0\\ T_{\mathbf{k}} & 0\end{pmatrix} \tag{12}$$

with $H^{-1}$ as above. In particular, the derivative of the deflection screw is given by

$$\mathbf{s}' = \begin{pmatrix} \boldsymbol{\theta}' \\ \mathbf{v}' + \mathbf{k} \times \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \theta_x' \\ \theta_y' \\ \theta_z' \\ v_x' - \theta_y \\ v_y' + \theta_x \\ v_z' \end{pmatrix} \tag{13}$$

The second derivative can be found in the same way to be

$$\mathbf{s}'' = \begin{pmatrix} \boldsymbol{\theta}' \\ \mathbf{v}'' + 2\mathbf{k} \times \boldsymbol{\theta}' \end{pmatrix} \tag{14}$$

**3.2 Compliance.** We begin with the standard equations of beam theory,

$$\frac{d}{d\mu} p_z = \frac{F_z}{EA}, \quad \frac{d}{d\mu} \theta_z = \frac{M_z}{GJ}, \quad \frac{d}{d\mu} \theta_x = \frac{M_x}{EJ_x}, \quad \frac{d}{d\mu} \theta_y = \frac{M_y}{EJ_y} \tag{15}$$

where the symbols used have their usual meaning: $E$ is Young's modulus, $G$ is the shear modulus of the material, $A$ is the cross-sectional area of the beam, and the $J$s are the relevant area integrals. The vectors $\mathbf{F}^T = (F_x, F_y, F_z)$ and $\mathbf{M}^T = (M_x, M_y, M_z)$ represent the force and torque acting on the beam. The quantities $M_x$ and $M_y$ are usually called the bending moments, while $F_x$ and $F_y$ are the shear forces. The quantities can be combined into a single wrench acting on the beam

$$\mathcal{W} = \begin{pmatrix} \mathbf{M} \\ \mathbf{F} \end{pmatrix} \tag{16}$$

Now for the Euler–Bernoulli approximation, the gradient of the centerline is related to the bending angles by

$$p_x' = \theta_y \quad \text{and} \quad p_y' = -\theta_x \tag{17}$$

But in the Timoshenko approximation, we consider a loss of slope due to shearing, see Ref. [4], Sec. 39. The perpendicular to the face of the element is not tangent to the centerline curve here. For the Timoshenko beam, we have

$$p_x' - \theta_y = \frac{F_x}{a_x GA} \quad \text{and} \quad p_y' + \theta_x = \frac{F_y}{a_y GA} \tag{18}$$

where, as usual, $A$ is the cross sectional area of the beam and $G$ is the shear modulus. The constants $a_x$ and $a_y$ depend on the shape of the beam's cross section, for a circular cross section $a_x = a_y = 2/3$.

Comparing this with Eq. (13), we have the fourth and fifth components of $\mathbf{s}'$ are the right-hand sides of the above equations. This can be combined with Eq. (15) above to give the following version of Hooke's law:

$$\mathbf{s}' = c\mathcal{W} \tag{19}$$

The matrix $c$ is the compliance density of the beam and is easily seen to be given by

$$c = \operatorname{diag}\left( \frac{1}{EJ_x}, \frac{1}{EJ_y}, \frac{1}{GJ}, \frac{1}{a_x GA}, \frac{1}{a_y GA}, \frac{1}{EA} \right) \tag{20}$$

where $\operatorname{diag}(x_1, \ldots, x_n)$ denotes the diagonal matrix with entries $x_1, \ldots, x_n$. In the Euler–Bernoulli approximation, the fourth and fifth places here would be zero.

In Ref. [5], von Mises derived the compliance matrix of a straight beam using the standard equations of Euler–Bernoulli beam theory. We can find a similar compliance matrix for the Timoshenko beam by integrating the compliance density given above along the beam. To do this, we need to know the compliance density at all points along the beam but expressed in a common coordinate system. A convenient frame to choose is located at the center of the beam. In this frame, a beam element, a distance $\mu$ from the center, has compliance density $\bar{c}(\mu) = H^{-1} c H^{-T}$, where $H$ is the translation:

$$H = \begin{pmatrix} I_3 & 0 \\ \mu T_{\mathbf{k}} & I_3 \end{pmatrix} \tag{21}$$

So integrating the compliance density from $-l/2$ to $l/2$ gives the compliance matrix $C$, of a beam of length $l$,

$$C = \int_{-l/2}^{l/2} \bar{c}(\mu) d\mu = \operatorname{diag}\left( \frac{l}{EJ_x}, \frac{l}{EJ_y}, \frac{l}{GJ}, \frac{l}{a_x GA} + \frac{l^3}{12EJ_y}, \frac{l}{a_y GA} + \frac{l^3}{12EJ_x}, \frac{l}{EA} \right) \tag{22}$$

This can be seen to be a slight modification of the result given by von Mises. For comparison, consider a cylindrical steel rod of length 2 m and radius 5 cm; the terms in the compliance matrix are $l/(a_x GA) \approx 5 \times 10^{-9} \mathrm{N}^{-1}$ m and $l^3/(12EJ_y) \approx 6 \times 10^{-7} \mathrm{N}^{-1}$ m. That is, the Timoshenko correction is a hundred times smaller than the corresponding term in the Euler–Bernoulli theory. (The material properties used here are the same as those used in the example considered in Sec. 6 below.)

In this coarser approximation, Hooke's law is given by

$$\mathbf{s}(l/2) - \mathbf{s}(-l/2) = C\mathcal{W} \tag{23}$$

were $\mathbf{s}(l/2)$ and $\mathbf{s}(-l/2)$ are the deflection screws at either end of the beam and $\mathcal{W}$ is the wrench acting on the beam.

## 4 Variational Methods

**4.1 The Lagrangian.** In general, the potential energy of an elastic system is given by the work done in stressing the system. Here we have

$$E_p = \frac{1}{2} \int_0^l \mathcal{W}^T \mathbf{s}' d\mu \tag{24}$$

reverting to our original parametrization of the beam with $\mu$ running from 0 to $l$. Using the generalized version of Hooke's law given in Eq. (19) above, this can be written as

$$E_p = \frac{1}{2} \int_0^l \mathbf{s}'^T k \mathbf{s}' d\mu \tag{25}$$

where $k$ is the stiffness density given by $k = c^{-1} = \operatorname{diag}(EJ_x, EJ_y, GJ, a_x GA, a_y GA, EA)$.

The kinetic energy of the beam is given by integrating the kinetic energies of all the elements along the beam

$$E_k = \frac{1}{2} \int_0^l \dot{\mathbf{s}}^T n \dot{\mathbf{s}} d\mu \tag{26}$$

The quantity $n$ here represents an inertia density, and $\dot{\mathbf{s}} = d\mathbf{s}/dt$ represents the time derivative of the deflection screw.

The inertia density of the beam can be written as

$$n = \rho \operatorname{diag}(J_x, J_y, J, A, A, A) \tag{27}$$

where $\rho$ is the mass per unit length of the beam, $A$ is the cross-sectional area, and $J, J_x, J_y$ are the moments of area.

In Euler–Bernoulli theory, the rotational inertias $J_x$ and $J_y$ are assumed to be small and are neglected, but in the Timoshenko theory they are retained.

The Lagrangian density

$$\mathcal{L} = \tfrac{1}{2}(\dot{\mathbf{s}}^T n \dot{\mathbf{s}} - \mathbf{s}'^T k \mathbf{s}') \tag{28}$$

can be introduced. The Lagrangian function, the difference between the kinetic and potential energies, $L = E_k - E_p$, is given as an integral along the beam

$$L = \int_0^l \mathcal{L} d\mu = \frac{1}{2} \int_0^l \mathbf{s}^T n \dot{\mathbf{s}} - \mathbf{s}'^T k \mathbf{s}' d\mu \tag{29}$$

The Lagrangian density contains six dependent functions. In such a case, we expect six equations of motion, one for each dependent function. The general form of the equations of motion will be

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial \mu}\left(\frac{\partial \mathcal{L}}{\partial \phi'}\right) - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right) = 0 \tag{30}$$

see Ref. [6] or Ref. [7], for example. This is the Euler–Lagrange equation, which is necessary to make the Lagrangian stationary with respect to variations in the function $\theta_x, \theta_y, \theta_z, v_x, v_y$ or $v_z$. Since there are no external generalized forces, the right-hand side is zero.

**4.2 Equations of Motion.** The Euler–Lagrange equations are simple to evaluate.

We can summarize the six equations of motion into a single screw equation:

$$k\mathbf{s}'' + k'\mathbf{s}' - n\ddot{\mathbf{s}} = 0 \tag{31}$$

where the $6 \times 6$ matrix $k'$ here is given by

$$k' = \begin{pmatrix} 0 & 0 & 0 & 0 & -a_y GA & 0 \\ 0 & 0 & 0 & a_x GA & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_x GA & 0 & 0 & 0 & 0 \\ -a_y GA & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{32}$$

The matrix $k'$ is the derivative of $k$ in the moving frame and can be written as

$$k' = -B^T k - kB \tag{33}$$

where $B$ is as defined in Eq. (12) above.

The above gives a system of six coupled second order partial differential equations. In fact, it is only the bending and shearing that are coupled. The classic Timoshenko beam equation

$$\left(\frac{\partial^2}{\partial \mu^2} - \frac{\rho}{E}\frac{\partial^2}{\partial t^2}\right)\left(\frac{\partial^2}{\partial \mu^2} - \frac{\rho}{a_y G}\frac{\partial^2}{\partial t^2}\right)\theta_x + \frac{\rho A}{EJ_x}\frac{\partial^2 \theta_x}{\partial t^2} = 0 \tag{34}$$

is a single fourth order partial differential equation in one bending variable only. This classic equation can be recovered from the system given above by eliminating the shear variable $v_x$ or $v_y$ from the equations containing it. This, to some extent, validates our derivation.

**4.3 Boundary Conditions.** In order to solve the equations of motion given above, we need to specify the boundary conditions. In standard beam theory, we have three main types of boundary conditions. If the end of the beam is clamped, then the deflection screw must be zero, $\mathbf{s}(0,t) = \mathbf{0}$. Notice that for an Euler–Bernoulli beam this includes both $v_x = 0$ and $v_x' = \theta_y = 0$, and similar for bending in the $x$-direction.

If one end of the beam is free, say, the end at $\mu = l$, then there are no forces or torques acting at that end. Hence from Hooke's law, Eq. (19), we have that $\mathbf{s}'(l,t) = 0$.

The third type of boundary condition is a simply supported end. Essentially this can be thought of as attaching the beam to a revolute joint. In the "spatial" formalism presented here, this can be generalized to an arbitrary joint. Some directions will be free and some clamped. As an example, consider a beam attached to a revolute joint; the beam will be free to turn about the joint axis but will be clamped in all other directions. This means that the beam will not experience any torque about the joint axis. Suppose the joint is represented by a screw $\boldsymbol{\zeta}$; for a revolute joint this would be the axis of the joint. Now we can find six independent wrenches $\mathcal{W}_i$, which satisfy

$$\mathcal{W}_1^T \boldsymbol{\zeta} = 1 \quad \text{and} \quad \mathcal{W}_j^T \boldsymbol{\zeta} = 0, \quad j = 2,3,\ldots,6 \tag{35}$$

Notice here that $\mathcal{W}_1$ is the only wrench, which will do any work on a body moving with instantaneous velocity $\boldsymbol{\zeta}$. If $\boldsymbol{\zeta}$ is a line, corresponding to a revolute joint, then $\mathcal{W}_1$ would represent a pure torque about that joint. Now if the joint is located at the foot of the beam, $\mu = 0$, the boundary conditions can be written as

$$\mathcal{W}_1^T \mathbf{s}'(0,t) = 0 \quad \text{and} \quad \mathcal{W}_j^T \mathbf{s}(0,t) = 0, \quad j = 2,3,\ldots,6 \tag{36}$$

Finding the wrenches $\mathcal{W}_i$ is straightforward and can often be done by inspection. Although we have only referred to revolute joints in the above, the results apply without change to any one-degree-of-freedom joint. So beams terminated by prismatic or helical joints can also be treated. Moreover, the method given above easily extends to joints with several degrees of freedom, so cylindrical and spherical joint can also be studied. In fact, the clamped and free boundary conditions can be seen as particular cases of this general approach, for a clamped end there are no degrees of freedom while for a free end there are six degrees of freedom.

## 5 The Rayleigh–Ritz Method

The Rayleigh–Ritz method is a standard numerical method for approximating the frequencies and mode shapes for vibrating systems, see Ref. [8], Secs. 6–31, for example. Moreover this technique has often been used to study vibrations in Timoshenko beams in various situations, see Refs. [9–11] to cite just a few examples. The purpose of this section is to show how the Rayleigh–Ritz method can be easily combined with the screw theory. In particular, we show that the screw form of the boundary conditions developed above is simple to incorporate into the Rayleigh–Ritz method. In order to do this, we need to set up some notation and the simplest way of doing this seems to be to rehearse the familiar derivation of Rayleigh–Ritz method but using screw theory.

As is well known, the idea behind the Rayleigh–Ritz approximation is not to derive equations of motion from the Lagrangian but to approximate the solution to the variational problem directly. We seek a function that makes the Lagrangian given above stationary. The approximation comes from the fact that the functions we use only come from a small finite set of functions. The set of functions we choose is not too important, although a good choice can help with the accuracy or simplify the computations. Sometimes we can choose a set of functions that satisfy the boundary conditions; when this is not possible we must treat the boundary conditions as constraints.

**5.1 Basic Theory.** Assume that the displacement screw has the following approximate form:

$$\mathbf{s}(\mu,t) \approx (\mathbf{s}_1 f_1(\mu) + \mathbf{s}_2 f_2(\mu) + \cdots + \mathbf{s}_q f_q(\mu))\cos(\omega t + \phi) \tag{37}$$

that is, a simple harmonic vibration where the shape of the beam is determined by the functions $f_i(\mu)$. These functions are often referred to as shape functions and many different choices are possible.

The idea is to substitute the above into the expression for the Lagrangian and then minimize with respect to the coefficients. This is quite simple in this case since the coefficients are the constant screws $\mathbf{s}_1, \ldots, \mathbf{s}_q$.

It is convenient to introduce a "stacked" notation. Let $\mathbf{z}$ be the $6q \times 1$ vector consisting of the coefficient screws $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_q$ stacked on top of each other:

$$\mathbf{z} = \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_q \end{pmatrix} \tag{38}$$

Another way of putting this is to say that $\mathbf{z}$ is an element of the tensor product $se(3) \otimes \mathbb{R}^q$. With this notation, the approximate Lagrangian takes the form

$$L \approx \tfrac{1}{2}(\omega^2 \mathbf{z}^T N \mathbf{z} \sin^2(\omega t + \phi) - \mathbf{z}^T K \mathbf{z} \cos^2(\omega t + \phi)) \tag{39}$$

where the $6q \times 6q$ matrix $N$ is given by $N = n \otimes F$ and the matrix $K$ is given by $K = k \otimes F_2 + kB \otimes F_1 + B^T k \otimes F_1^T + B^T kB \otimes F$, where $B$ is as in Eq. (12) above. The $F$ matrices are $q \times q$ matrices of integrals:

$$F_{ij} = \int_0^l f_i f_j d\mu, \quad (F_1)_{ij} = \int_0^l f_i f_j' d\mu, \quad (F_2)_{ij} = \int_0^l f_i' f_j' d\mu,$$
$$i,j = 1,2, \ldots, q$$

Note that in substituting for $\mathbf{s}'$ it is necessary to make use of the relation (11), so that

$$\mathbf{s}'(\mu,t) \approx \left( \sum_{i=1}^{q} f_i'(\mu) \mathrm{s}_i + f_i(\mu) B \mathrm{s}_i \right) \cos(\omega t + \phi) \tag{40}$$

Minimizing the Lagrangian yields a linear eigenvalue problem

$$K\mathbf{z} - \omega^2 N \mathbf{z} = \mathbf{0} \tag{41}$$

**5.2 Boundary Conditions Again.** If the shape functions $f_i$ satisfy the boundary conditions, then the above argument needs no modification. However, it is more usual to choose shape functions that do not automatically satisfy the boundary conditions. To cope with this situation, we can treat the minimization of the Lagrangian as a constrained optimization problem.

Suppose we have a free end with a corresponding boundary condition $\mathbf{s}'(l,t) = \mathbf{0}$. Substituting Eq. (37) into this gives

$$\mathbf{s}'(l,t) = \left( \sum_{i=1}^{q} f_i'(l) \mathbf{s}_i + f_i(l) B \mathbf{s}_i \right) \cos(\omega t + \phi) = \mathbf{0} \tag{42}$$

see Eq. (40) above. Since this must be satisfied for all $t$, we can conclude that

$$\sum_{i=1}^{q} f_i'(l) \mathbf{s}_i + f_i(l) B \mathbf{s}_i = \mathbf{0} \tag{43}$$

Let us introduce a sequence of $6 \times 6$ matrices:

$$U_i^T = f_i'(l) I_6 + f_i(l) B, \quad i = 1,2, \ldots, q \tag{44}$$

where $I_6$ is the $6 \times 6$ identity matrix. The constraint equation can now be written in terms of our stacked variable $\mathbf{z}$ as

$$(U_1^T | U_2^T | \ldots | U_q^T) \mathbf{z} = 0 \tag{45}$$

Next we introduce six Lagrange multipliers $\boldsymbol{\lambda}^T = (\lambda_1, \lambda_2, \ldots, \lambda_6)$. The constrained optimal solution now satisfies the following linear equations:

$$K\mathbf{z} - \omega^2 N \mathbf{z} + \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_q \end{pmatrix} \boldsymbol{\lambda} = \mathbf{0} \tag{46}$$

Now we can combine these equations with the constraint Eq. (43) into a single set of linear homogeneous equations by extending our stacked notation. Let

$$\tilde{\mathbf{z}} = \begin{pmatrix} \mathbf{z} \\ \boldsymbol{\lambda} \end{pmatrix} \tag{47}$$

and set

$$\tilde{K} = \begin{pmatrix} & & U_1 \\ & K & \vdots \\ & & U_2 \\ U_1^T & \cdots & U_q^T \quad 0 \end{pmatrix} \quad \text{and} \quad \tilde{N} = \begin{pmatrix} & & 0 \\ & N & \vdots \\ & & 0 \\ 0 & \cdots & 0 \quad 0 \end{pmatrix} \tag{48}$$

All the equations can now be written as

$$(\tilde{K} - \omega^2 \tilde{N}) \tilde{\mathbf{z}} = \mathbf{0} \tag{49}$$

Hence the approximate vibrational frequencies will be given by the solution to the eigenvalue equation:

$$\det(\tilde{K} - \omega^2 \tilde{N}) = 0 \tag{50}$$

If the beam is connected to a joint, then from Sec. 4.3 above, we have the boundary conditions

$$\mathcal{W}_1^T \mathbf{s}'(0,t) = 0 \quad \text{and} \quad \mathcal{W}_j^T \mathbf{s}(0,t) = 0, \quad j = 2,3, \ldots, 6 \tag{51}$$

where the wrenches $\mathcal{W}_i$ are as described in Sec. 4.3. If we substitute our approximation for $\mathbf{s}(\mu,t)$ given in Eq. (37) above, we get

$$\sum_{i=1}^{q} \mathcal{W}_1^T (f_i'(0) I_6 + f_i(0) B) \mathbf{s}_i = 0 \quad \text{and} \quad \sum_{i=1}^{q} f_i(0) \mathcal{W}_j^T \mathbf{s}_i = 0,$$
$$j = 2,3, \ldots, 6 \tag{52}$$

Now let us construct a new sequence of $6 \times 6$ matrices:

$$V_i = (f_i'(0) \mathcal{W}_1 + f_i(0) B^T \mathcal{W}_1 | f_i(0) \mathcal{W}_2 | \cdots | f_i(0) \mathcal{W}_6), \quad i = 1,2, \ldots, q \tag{53}$$

That is, the columns of $V_i$ are the six wrenches multiplied by the appropriate boundary value of the shape function. The boundary condition can now be expressed as

$$\sum_{i=1}^{q} V_i^T \mathbf{s}_i = (V_1^T | V_2^T | \ldots | V_q^T) \mathbf{z} = \mathbf{0} \tag{54}$$

As above, we can now cast the constrained optimization problem into the same extended eigenproblem:

$$(\tilde{K} - \omega^2 \tilde{N}) \tilde{\mathbf{z}} = \mathbf{0} \tag{55}$$

where now $\tilde{K}$ is given by

$$\tilde{K} = \begin{pmatrix} & & V_1 \\ & K & \vdots \\ & & V_q \\ V_1^T & \cdots & V_q^T \quad 0 \end{pmatrix} \tag{56}$$

and $\tilde{N}$ as above.

Notice that the free end, considered first, is really a special case of the jointed end considered subsequently. In fact, it is easy to see how to treat an $n$ degree-of-freedom jointed end, where $0 \leq n \leq 6$. All that is needed is to find the wrenches $\mathcal{W}_i$ dual to the screws of freedom at the end of the beam and then define the columns matrices $V_i$ as $f_i'(0)\mathcal{W}_j + f_i(0) B^T \mathcal{W}_j$ or $f_i(0)\mathcal{W}_j$ depending on whether motion about $\mathbf{s}_j$ is clamped or free. This simplicity in treating different boundary conditions is one of the main utilities of the method.

## 6 Example: A Robot Link

In this example, the utility of the screw theory becomes even more clear. A standard case from classical beam theory is a beam hinged at one end. In the case of the robot link, the beam is hinged

Table 1 The first three frequencies (Hz) for different orders of approximating splines ($q$)

| $q$ | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ | $\omega_5$ | $\omega_6$ |
|---|---|---|---|---|---|---|
| 4 | 496.431 | 521.821 | 3492.494 | 4288.008 | 5030.611 | 8150.522 |
| 5 | 496.018 | 512.325 | 1603.640 | 1614.248 | 5030.611 | 6691.791 |
| 6 | 495.444 | 509.864 | 1594.628 | 1595.897 | 3347.973 | 3355.881 |
| 7 | 495.437 | 506.521 | 1581.400 | 1591.978 | 3292.574 | 3306.659 |
| 8 | 495.437 | 503.967 | 1581.171 | 1590.074 | 3231.632 | 3241.088 |
| 9 | 495.437 | 502.211 | 1581.086 | 1587.596 | 3229.490 | 3238.284 |
| 10 | 495.437 | 500.933 | 1581.085 | 1586.336 | 3228.191 | 3233.476 |

at both ends and moreover, these hinges may not be parallel. This means that in these cases the vibrational modes may not be pure bending, pure torsion, or pure extension but rather some mixtures of these. The tools developed above are particularly suited to this case.

As an example, we study here the vibrations of a 2 m cylindrical steel bar with a radius of 5 cm. We assume that the ends of the bar are constrained to rotate about two perpendicular joints. At one end $\mu=0$, we have a joint described by the screw $\mathbf{s}_a$ and at the other end $\mu=l$, the joint is $\mathbf{s}_b$ with

$$\mathbf{s}_a = (1,0,0,0,0,0)^T, \quad \mathbf{s}_b = (0,1,0,-l,0,0)^T \tag{57}$$

where $l=2$ for this example.

We propose to model the shape of the beam by a Bézier spline. That is, we chose the shape functions to be Bernstein polynomials, $f_i(\mu)=b_{q,i}(\mu)=\binom{q}{i}\mu^i(1-\mu)^{q-i}$. We need to redefine the numbering of the terms in the approximation here since the Bernstein polynomials in a Bézier curve are usually numbered from 0 to $q$, so the approximate shape of the beam will be given by

$$\mathbf{s}(\mu) \approx \mathbf{s}_0 b_{q,0}\left(\frac{\mu}{l}\right) + \mathbf{s}_1 b_{q,1}\left(\frac{\mu}{l}\right) + \mathbf{s}_2 b_{q,2}\left(\frac{\mu}{l}\right) + \cdots + \mathbf{s}_q b_{q,q}\left(\frac{\mu}{l}\right) \tag{58}$$

By judicially choosing the first and last screws, the boundary conditions can be satisfied automatically. Let $\mathbf{s}_0=\alpha\mathbf{s}_a$ and $\mathbf{s}_q=\beta\mathbf{s}_b$. The boundary conditions, $\mathcal{W}_j^T\mathbf{s}(m,t)=0, j=2,\ldots,6$ for $m=0$ and $m=l$, are then satisfied, (of course the constraint wrenches will be different at the two ends). Next we must deal with the other boundary conditions $\mathcal{W}_1^T\mathbf{s}'(m,t)=0$. The derivatives of the Bernstein polynomials at 0 are all zero, except for the first two for which we have

$$b'_{0,q}(0) = -q \quad \text{and} \quad b'_{1,q}(0) = q \tag{59}$$

Hence we have that

$$\mathcal{W}_1^T\mathbf{s}'(0,t) = 0 = -q\mathcal{W}_1^T\mathbf{s}_0 + q\mathcal{W}_1^T\mathbf{s}_1 + \mathcal{W}_1^T B\mathbf{s}_0 \tag{60}$$

In this example, $\mathcal{W}_1^T=(1,0,0,0,0,0)$, so that the boundary condition expresses the idea that there can be no torque about the $x$-axis at the end of the beam. The term $\mathcal{W}_1^T B\mathbf{s}_0$ is zero and so we can conclude that

$$\mathcal{W}_1^T\mathbf{s}_1 = \mathcal{W}_1^T\mathbf{s}_0 = \alpha \tag{61}$$

Now the boundary wrenches $\mathcal{W}_1,\ldots,\mathcal{W}_6$ are linearly independent so it will be possible to expand the identity matrix as

$$I_6 = \mathbf{s}_a\mathcal{W}_1^T + \mathbf{s}_{a2}\mathcal{W}_2^T + \cdots + \mathbf{s}_{a6}\mathcal{W}_6^T \tag{62}$$

That is, it will be possible to find the screws $\mathbf{s}_{ai}$ to satisfy this equation. However, we do not need to actually calculate them, all we need to do is to observe that

$$\mathbf{s}_0 = I_6\mathbf{s}_0 = \mathbf{s}_a\mathcal{W}_1^T\mathbf{s}_0 \tag{63}$$

since $\mathbf{s}_0$ already satisfies the boundary conditions $\mathcal{W}_i^T\mathbf{s}_0=0$, $i=2,\ldots,6$. Substituting from Eq. (61) above gives

$$\mathbf{s}_0 = \mathbf{s}_a\mathcal{W}_1^T\mathbf{s}_1 \tag{64}$$

This means that in the expression for the approximate Lagrangian $L\approx(1/4)(\omega^2\mathbf{z}^TN\mathbf{z}-\mathbf{z}^TK\mathbf{z})$, we can substitute for $\mathbf{s}_0$. The results are easier to see if we just concentrate on the kinetic energy term

$$\mathbf{z}^TN\mathbf{z} = \sum_{i=0}^{q}\sum_{j=0}^{q}\mathbf{s}_i^Tn\mathbf{s}_jF_{ij} = \mathbf{s}_0^Tn\mathbf{s}_0F_{00} + \sum_{i=1}^{q}\mathbf{s}_i^Tn\mathbf{s}_0F_{i0} + \sum_{j=1}^{q}\mathbf{s}_0^Tn\mathbf{s}_jF_{0j} + \sum_{i=1}^{q}\sum_{j=1}^{q}\mathbf{s}_i^Tn\mathbf{s}_jF_{ij} \tag{65}$$

here the first row, first column, and top-left-hand elements have been separated out. After the substitution, we can write

$$\mathbf{z}^TN\mathbf{z} = \mathbf{s}_1^T(\mathcal{W}_1\mathbf{s}_a^Tn\mathbf{s}_a\mathcal{W}_1^TF_{00} + n\mathbf{s}_a\mathcal{W}_1^TF_{10} + \mathcal{W}_1\mathbf{s}_a^TnF_{01} + nF_{11})\mathbf{s}_1 + \sum_{i=2}^{q}\mathbf{s}_i^T(n\mathbf{s}_a)\mathcal{W}_1F_{i0} + nF_{i1})\mathbf{s}_1 + \sum_{j=2}^{q}\mathbf{s}_1^T(\mathcal{W}_1\mathbf{s}_a^TnF_{0j} + nF_{1j})\mathbf{s}_j + \sum_{i=2}^{q}\sum_{j=2}^{q}\mathbf{s}_i^Tn\mathbf{s}_jF_{ij} \tag{66}$$

In terms of the large matrix $N$, this simply means that we must modify the first six rows and add them to the second six, modify the first six columns and add to the second six columns, and then modify the top-left $6\times6$ block and add to the next $6\times6$ diagonal block. After this the first six rows and columns, the ones associated with $\mathbf{s}_0$, can be removed reducing the overall size of the problem a little. The boundary condition at the other end of the beam can be treated in a similar way, eliminating the last six rows and columns. And of course the same procedure can be performed for the potential energy matrix $K$.

The computations for this example were performed using MATHEMATICA. The material properties of mild steel were used:

$$\rho = 7.8\times10^3\text{kg m}^{-3}, \quad E = 21.0\times10^{10}\text{N m}^{-2}, \quad G = 8.0\times10^{10}\text{N m}^{-2}$$

Results for first few frequencies are shown in Table 1. As can be seen, the results for the lowest frequency ($\omega_1$) seem to converge quickly, but the next two higher frequencies converge more slowly. Also the frequencies seem to be paired, the first two, third and fourth, and fifth and sixth are close together. This can be explained by the symmetry of the problem. If we rotate the beam through 180 deg about an axis halfway along the beam and set at 45 deg to the axes of the joints, then clearly nothing has changed. Hence we expect the solutions to be degenerate; that is, we expect the modes to occur in pairs with the same frequency and mode shapes related by the symmetry transformation described above. A better numerical approach would have taken account of this. However, it is a simple matter to find the shapes of the modes and these are shown in Fig. 2. Notice the large extensions near the beam ends.

The numerical approach used here was somewhat rudimentary, relying only on the standard routines supplied with the software to

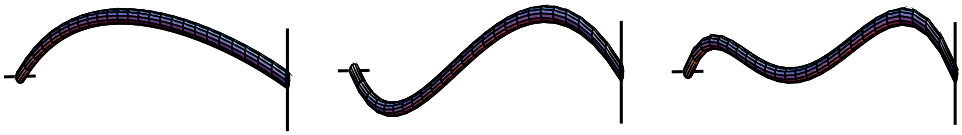


**Fig. 2 The first three modes shapes. The thick lines at the ends represent direction of the joint axes.**

find eigenvalues and eigenvectors. More important is the fact that these computations can be done quite simply and it should be expected that the methods described in this work can be used to produce accurate results if more attention is paid to numerical methods used. For example, it is well known that the success of the Rayleigh–Ritz method depends on the choice of shape functions. Here Bézier splines were used but perhaps some other choice would be better, B-splines, for instance (this would have the advantage of making the matrix $F$ tridiagonal).

## 7 Conclusions

In this work, traditional Timoshenko beam theory has been translated into screw theory. The advantage of this is that the different vibrational modes, bending, torsion, and extension/compression, can be treated on an equal footing and so problems involving all modes at the same time can be formulated and solved quite simply. The exposition given above only treats straight beams with uniform cross sections; however, there is no reason why this approach cannot be extended to curved rods with varying cross sections.

A variational method, based on the Lagrangian density along the beam, was used. This method provides an elegant and efficient way to derive the equations of motion. The equations of motion can be written as a screw equation, which neatly summarizes the equations for the different vibrational modes of the beam. Expressing the Lagrangian as the integral of a density along the beam also leads to a screw version of the familiar Rayleigh–Ritz method for computing the vibrational modes and frequencies of the beam.

The use of screw theory also leads to a new view of the boundary conditions. In standard beam theory, there are several different kinds of boundary conditions: clamped ends, free ends, and simply supported ends. By considering the freedom and constraints as screws and wrenches, all these different types of boundary conditions can be thought of as examples of a single type of boundary condition. In this scheme, the screws of freedom and wrenches of constraint must be identified at the ends of the beam. The number of freedoms and constraints will always total 6. With six constraints and no freedom, the end of the beam is clamped. On the other hand, if the beam has no constraints and has six degrees-of-freedom then it is free. In the case where the end of the beam has a single degree-of-freedom, we can think of the beam as a being attached to a mechanical joint. Usually, this will be a revolute joint corresponding to a pinned end or knife-edge. But now there is no essential difference between different kinds of one degree-of-freedom joint, so it would be a simple matter to study beams attached to helical or prismatic joints. A rolling node can be thought of as a two degree-of-freedom end, and a passive ball and socket joint is a three degree-of-freedom end.

The example given in Sec. 4 above is intended to be indicative only. By having simply supported ends but with the axes of freedom at right angles, the modal solutions will be combinations of bending in both directions perpendicular to the beam axis and also torsion and extension/compression along the beam axis. There is a large literature on coupled vibrations in Timoshenko beams, see Refs. [12,13] for just two examples. However, in these works coupled vibrations occur because of a nonsymmetric beam sections. In the example studied above, the beam section is symmetric, and the coupling arises because of the boundary conditions. We believe that this effect is less well studied. It was suggested above that this example might have some relevance for robot links. This is perhaps overstating things a little. Although a robot link has joints at either end, it cannot usually rotate freely about these joints since they are driven by motors. Often there is a lot of compliance at the joints of a robot and this compliance will usually be more significant than the compliance of the links for studying vibrations of the system. However, for space robots and especially designed flexible robots, this work may have some value. A better model might be a cantilever with mass at the free end. But this mass should not be a point mass, its inertia should be taken into account. Again the screw formalism used in this work will be useful here. In a real system, the ends of the link will have some structure, which is not the same as the body of the beam and these end-effects should be taken into account.

The present work only applies to small deformations of beams as the deflections are modeled by screws. A more accurate approach would be to model the deflection of the beam at each point along its length as a rigid motion; this leads to the well known Cosserat theory of beams, sometimes referred to as geometrically exact beam theory, see Refs. [14,15]. In future work, we hope to be able to study this problem from a screw theory perspective.

## Acknowledgment

The authors would like to thank the following for financial support. In China, this research was sponsored by the NSFC (Grant Nos. 50275002 and 50475001) and SRF for ROCS, SEM. From the UK, a visit by J.M.S. to China was made possible by EPSRC overseas travel grant EP/E034527/1.

## References

[1] Ball, R. S., 1900, *The Theory of Screws*, Cambridge University Press, Cambridge.
[2] Selig, J. M., 2005, *Geometrical Fundamentals of Robotics*, 2nd ed., Springer, New York.
[3] Selig, J. M., and Ding, X., 2001, "A Screw Theory of Static Beams," *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Maui, HI, pp. 2544–2550.
[4] Timoshenko, S., 1955, *Strength of Materials: Part 1*, 3rd ed., van Nostrand, New York.
[5] von Mises, R., 1924, "Motorrechnung, ein neues Hilfsmittel in der Mechanik," ZAMM, **4**(2), pp. 155–181 (English translation by E. J. Baker and K. Wohlhart, *Motor Calculus: A New Theoretical Device for Mechanics*, Institute for Mechanics, University of Technology Graz, Austria, 1996).
[6] Goldstein, H., Poole, C., and Safko, J., 2002, *Classical Mechanics*, 3rd ed., Addison-Wesley, New York.
[7] Pipes, L. A., and Harvill, L. R., 1970, *Applied Mathematics for Engineers and Physicists*, 3rd ed., McGraw-Hill, Tokyo.
[8] Meirovitch, L., 1967, *Analytical Methods in Vibrations*, Macmillan, New York.
[9] Aalami, B., and Atzori, B., 1974, "Flexural Vibrations and Timoshenko's Beam Theory," AIAA J., **12**(5), pp. 679–685.
[10] Oguamanam, D. C. D., and Heppler, G. R., 1996, "Effect of Rotating Speed on the Flexural Vibration of a Timoshenko Beam," *Proceedings of IEEE International Conference on Robotics and Automation*, Minneapolis, MN, pp. 2438–2443.
[11] Wang, S., 1997, "Unified Timoshenko Beam B-spline Rayleigh–Ritz Method for Vibration and Buckling Analysis of Thick and Thin Beams and Plates," Int. J. Numer. Methods Eng., **40**(3), pp. 473–491.
[12] Bercin, A. N., and Tanaka, M., 1997, "Coupled Flexural-Torsional Vibrations of Timoshenko Beams," J. Sound Vib., **207**(1), pp. 47–59.
[13] Bishop, R. E. D., and Price, W. G., 1977, "Coupled Bending and Twisting of a Timoshenko Beam," J. Sound Vib., **50**(4), pp. 469–477.
[14] Antman, S. S., 2005, *Nonlinear Problems of Elasticity*, 2nd ed., Springer, New York.
[15] Rubin, M. B., 2000, *Cosserat Theories: Shells, Rods and Points*, Kluwer, Dordrecht.

# Evaluation of Mixed-Mode Stress Intensity Factors for a Sharp Notch-tip With Curved and Stressed Edges

**J. H. Chang**
Professor

**J. F. Fan**
Graduate Research Assistant

Department of Civil Engineering,
National Central University,
Chungli, 320001 Taiwan

*For a sharp notch with curved edges and subjected to surface tractions along the edges, the fracture parameters (in particular, the stress intensity factors and the size of a singular-dominant zone) are significantly affected by the near-tip geometric and loading conditions. In this paper, a pair of contour integrals termed $J_{kR}$ is presented for calculating the mixed-mode stress intensity factors at such a sharp notch-tip. Furthermore, by proper use of the integrals, the extent of the singular-dominant zone can be effectively characterized. Since no a priori auxiliary (or, complementary) solutions are required in its formulation, the approach appears to be feasible for problems of arbitrary notch angles and curved shapes. Also, no special treatments are required for the modeling of the near-tip singular behavior so that the integration can be performed by direct use of numerical schemes such as finite element method.* [DOI: 10.1115/1.3002333]

*Keywords: sharp notch-tip, curved notch edges, edge tractions, $J_{kR}$-integrals, modified path-independence, singular-dominant zone*

## 1 Introduction

The edges for a sharp notch-tip may have arbitrary curved shapes. Also, the tractions along the notch edges—which may be due to pressurized fluids, contact pressure, interfacial friction of inclusions, etc.—are frequently of practical interest in many engineering applications. The associated fracture behavior in the near-tip area thus depends significantly on the notch geometry and loading condition. For problems containing complex geometric and loading conditions, direct evaluation of the near-tip stress field with numerical schemes such as finite element method appears to be difficult due to the complicated mechanical state around the singular point. Investigations on proper numerical approaches are therefore in need.

For a sharp notch with straight and traction-free edges, the order of stress singularity at the notch-tip is dependent on the notch angle and different from that of the crack case (i.e., $r^{-1/2}$). In particular, it is noted that the Mode I and Mode II asymptotic stresses are governed by different orders of singularity, with the strength of singularity for Mode II substantially weaker than that of Mode I [1–4]. The mixed-mode singular stress field for a notch-tip in linear elasticity can be effectively characterized by a pair of stress intensity factors (SIFs), similar to the concept customarily used in the crack case. Due to the difficulty in direct evaluation of SIFs, a number of indirect approaches have instead been proposed for the purpose. Most of the approaches are based on the concept of contour integrals. For example, a modified $H$-integral was presented to compute either pure- or mixed-mode SIFs in notched solids with various notch angles [5–8]. In order to perform the $H$-integral, a set of particularly complementary solutions needs to be developed. In addition to $H$, another pair of contour integrals termed $J_{kR}$ were presented by the authors [9] for the calculation. By comparing with the $H$-integral, $J_{kR}$ appears to be more straightforward in practice since no extra complementary solutions are required in their formulation.

On the other hand, for the special instance when the singular point corresponds to a crack tip, a variety of studies have been presented for problems containing curved traction-free cracks. Among them, theoretical studies were developed for a description of the mixed-mode asymptotic stress behavior for curved cracks with various shapes of curvature [10–12]. Also, a number of contour or area integrals have been proposed for the calculation of SIFs or other fracture parameters associated with curved cracks [13–15]. Nevertheless, more study is required in order to establish general formulation of problems containing notch-tips with curved and stressed edges.

The object of this paper is to evaluate the mixed-mode stress intensity factors for a sharp notch-tip with curved edges, which appears as a general case of the aforementioned curved crack problems. Also, the study is considered with the presence of surface tractions along the notch edges. A numerical procedure, based on the concept of the $J_{kR}$-integrals, is developed for this purpose. Since no extra auxiliary or complementary solutions are required in their formulation, $J_{kR}$ thus appear to be applicable for problems of arbitrary curved shapes. The approach is developed and incorporated with the finite element method for numerical calculation. No particular singular elements are used in the study.

## 2 The Asymptotic Stress Field

Consider an elastic body in a 2D field containing a sharp notch with its edges of arbitrary curved shape and of asymptotically tangential angle $\beta$ at its tip $O$ (Fig. 1). A local coordinate system originating at the notch-tip $O$ is introduced, with the notch angle being bisected by the (negative) $x_1$-axis. The body is then subjected to a system of loads, which includes particularly the traction vector $\mathbf{t}^n$ applied on the notch edges. In order to establish the expression for the asymptotic near-tip stress field, we need to consider the boundary value problem formulated in terms of the Airy stress function with the prescribed traction boundary conditions on the notch edges. This leads to a solution consisting of two parts, which are the particular solution $\boldsymbol{\sigma}^p$ (associated with the

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received October 11, 2006; final manuscript received August 31, 2007; published online March 5, 2009. Review conducted by Zhigang Suo.

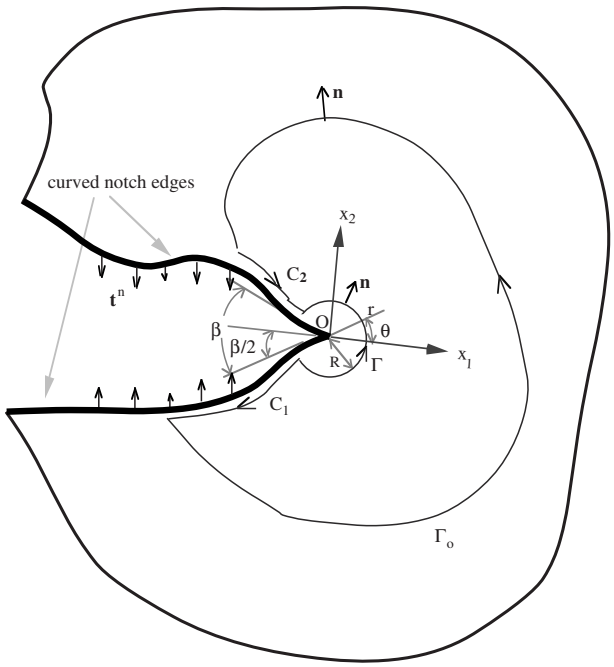



**Fig. 1 An elastic body in a 2D field, containing a notch of asymptotically tangential angle $\beta$ at tip *O***

notch edge traction $\mathbf{t}^n$) and the homogeneous solution $\boldsymbol{\sigma}^h$ (associated with the traction-free boundary condition). As a result, the asymptotic near-tip stresses $\boldsymbol{\sigma}$ can thus be expressed as

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^p + \boldsymbol{\sigma}^h \tag{1}$$

Note that, as shown in the Appendix, $\boldsymbol{\sigma}^p$ is of finite value and needs to be determined by solving the corresponding boundary value problem. As to $\boldsymbol{\sigma}^h$, by taking a specific Airy stress function employed by Williams [1] and taking the traction-free (homogeneous) boundary condition on the notch edges, we then have the following separable forms as functions of $\beta$:

$$\sigma_r^h(r,\theta;\beta) = r^{\lambda 1}\frac{(K_{\mathrm{I}})_\beta}{(2\pi-\beta)^{1/2}}[f_1(\beta)\cos\lambda_1\theta + g_1(\beta)\cos(\lambda_1+2)\theta] + r^{\lambda 2}\frac{(K_{\mathrm{II}})_\beta}{(2\pi-\beta)^{1/2}}[h_1(\beta)\sin\lambda_2\theta + p_1(\beta)\sin(\lambda_2+2)\theta] \tag{2}$$

$$\sigma_\theta^h(r,\theta;\beta) = r^{\lambda 1}\frac{(K_{\mathrm{I}})_\beta}{(2\pi-\beta)^{1/2}}[f_2(\beta)\cos\lambda_1\theta + g_2(\beta)\cos(\lambda_1+2)\theta] + r^{\lambda 2}\frac{(K_{\mathrm{II}})_\beta}{(2\pi-\beta)^{1/2}}[h_2(\beta)\sin\lambda_2\theta + p_2(\beta)\sin(\lambda_2+2)\theta] \tag{3}$$

$$\tau_{r\theta}^h(r,\theta;\beta) = r^{\lambda 1}\frac{(K_{\mathrm{I}})_\beta}{(2\pi-\beta)^{1/2}}[f_3(\beta)\sin\lambda_1\theta + g_3(\beta)\sin(\lambda_1+2)\theta] + r^{\lambda 2}\frac{(K_{\mathrm{II}})_\beta}{(2\pi-\beta)^{1/2}}[h_3(\beta)\cos\lambda_2\theta + p_3(\beta)\cos(\lambda_2+2)\theta] \tag{4}$$

where $(r,\theta)$ denote the polar components of the local coordinate and $-(\pi-\beta/2)\le\theta\le(\pi-\beta/2)$. The orders of singularity $\lambda_1$ and $\lambda_2$ are the smallest nontrivial real eigenvalues of the following pair of eigenequations:

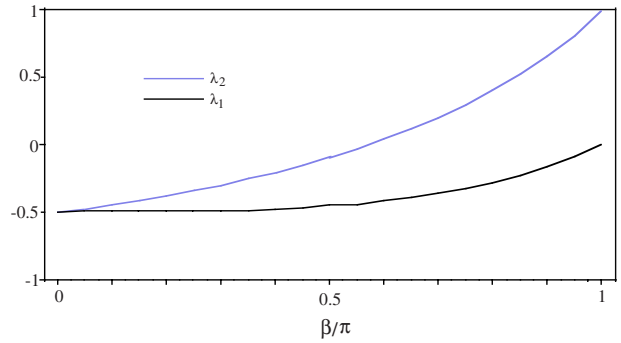



**Fig. 2 The values of $\lambda_1$ and $\lambda_2$, as functions of the notch angle $\beta$**

$$\sin[(\lambda+1)(2\pi-\beta)] \pm (\lambda+1)\sin\beta = 0 \tag{5}$$

The values of $\lambda_1$ and $\lambda_2$, as functions of $\beta$, are presented in Fig. 2. The figure shows that the Mode I (symmetric) and Mode II (antisymmetric) stresses are governed by different orders of singularity, except for the special case when $\beta=0$ (i.e., the crack problem, with $\lambda_1=\lambda_2=-0.5$). Since $\lambda_2\ge\lambda_1$, the strength of singularity for Mode II stresses is thus essentially weaker than those for Mode I. Also, for notch angles greater than $0.57\pi$ (approximately), $\lambda_2$ becomes positive and, consequently, Mode II stresses appear to be bounded. As an aside, the dimensionless coefficients $f_i$, $h_i$, $g_i$, and $p_i$ $(i=1,2,3)$ are also explicit functions of $\beta$. Detailed descriptions and discussions on variations of these coefficients with respect to $\beta$ can be found in the authors' previous work [9]. Note that Eqs. (2)–(5) are valid for problems containing both straight and curved notch edges in that they are applied in the near-tip region where the edges lie asymptotically along the tangential directions.

As shown in Eq. (1), the asymptotic near-tip stress field consists of the finite-valued $\boldsymbol{\sigma}^p$ and the singular-valued $\boldsymbol{\sigma}^h$. While the contribution from $\boldsymbol{\sigma}^p$ is negligible in the singular-dominant area, the size of the singular zone is inevitably and significantly affected by its appearance. Cautious investigation on such effects is therefore always necessary and will be presented in the following numerical examples.

With the above mentioned asymptotic stress field, the mixed-mode stress intensity factors for an arbitrary tangential notch angle $\beta$, denoted as $(K_{\mathrm{I}})_\beta$ and $(K_{\mathrm{II}})_\beta$, can be defined as

$$(K_{\mathrm{I}})_\beta \equiv \lim_{r\to 0}(2\pi-\beta)^{1/2}r^{-\lambda 1}\sigma_\theta(r,0;\beta) \approx \lim_{r\to 0}(2\pi-\beta)^{1/2}r^{-\lambda 1}\sigma_\theta^h(r,0;\beta) \tag{6}$$

$$(K_{\mathrm{II}})_\beta \equiv \lim_{r\to 0}(2\pi-\beta)^{1/2}r^{-\lambda 2}\tau_{r\theta}(r,0;\beta) \approx \lim_{r\to 0}(2\pi-\beta)^{1/2}r^{-\lambda 2}\tau_{r\theta}^h(r,0;\beta) \tag{7}$$

These SIFs are undetermined constants that account for the strength of stress singularity in the near-tip region. Note that the SIFs depend on both the far-field and near-tip loading conditions for a notch-tip with curved and stressed edges.

## 3 The $J_{kR}$-Integrals

According to the concept of the conventional $J_k$-integrals [16,17], the corresponding contour integrals for a sharp notch-tip can be defined as

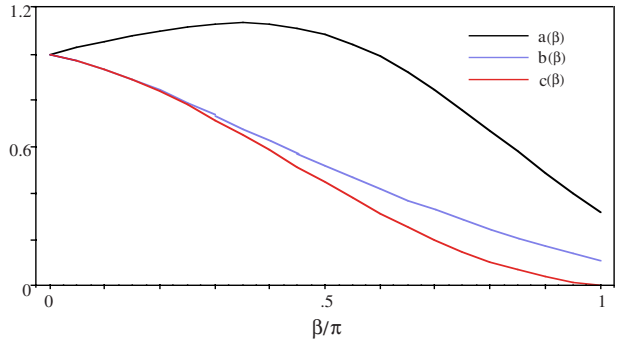



**Fig. 3 The variations of *a*, *b*, and *c* with respect to $\beta$**

$$J_k = \lim_{\Gamma \to 0} \int_{\Gamma} \left[ Wn_k - \sigma_{ij} n_j \left( \frac{\partial u_i}{\partial x_k} \right) \right] ds, \quad k = 1,2 \tag{8}$$

where $W$ is the strain energy density of the material, $\sigma_{ij}$ and $u_i$ are the Cartesian components of the stress tensor and the displacement vector, $n_j$ is the Cartesian components of the outward unit vector normal to $\Gamma$ (as shown in Fig. 1), and $s$ is the arc length along the contour. The path of integration, $\Gamma$ is defined as a counterclockwise contour encircling and shrinking onto the tip of the notch $O$ (this limiting case is not shown in Fig. 1). Still, Eq. (8) is valid for problems with both straight and curved notch edges in that the path of integration, $\Gamma$, is defined as a counterclockwise contour encircling and shrinking onto the tip $O$ (this limiting case is not shown in Fig. 1).

For the special case when $\beta=0$ (i.e., the crack problem), the integration in Eq. (8) results in a pair of finite-valued solutions for $J_1$ and $J_2$. It is well known that the results of $J_k$ can be used to evaluate the mixed-mode SIFs at the crack tip. However, for notch problems with generally $\beta>0$, the characteristic of finite values of $J_k$ is no longer valid. According to the authors' previous work [9], the relationship between $J_k$ and the SIFs (defined in Eqs. (6) and (7)) for a notch-tip can be written as

$$J_1 = \lim_{r \to 0} \frac{\Lambda}{E} [r^{2\lambda_1+1} a(\beta)(K_{\mathrm{I}})_\beta^2 + r^{2\lambda_2+1} b(\beta)(K_{\mathrm{II}})_\beta^2] \tag{9}$$

$$J_2 = \lim_{r \to 0} -\frac{2\Lambda}{E} r^{\lambda_1+\lambda_2+1} c(\beta)(K_{\mathrm{I}})_\beta (K_{\mathrm{II}})_\beta \tag{10}$$

where $\Lambda=1$ (for plane stress) or $1-\nu^2$ (for plane strain), $E$ is Young's modulus, and $\nu$ is Poisson's ratio.. Also, $a(\beta)$, $b(\beta)$, and $c(\beta)$ are dimensionless functions of $\beta$, and their variations with respect to $\beta$ are shown in Fig. 3. It is observed that, in addition to $(K_{\mathrm{I}})_\beta$ and $(K_{\mathrm{II}})_\beta$, three extra terms (i.e., $r^{2\lambda_1+1}$, $r^{2\lambda_2+1}$, and $r^{\lambda_1+\lambda_2+1}$) are contained in Eqs. (9) and (10). With the presence of these three terms, we anticipate that the values of $J_k$ vanish as the limiting condition $r \to 0$ by definition.

Due to their vanishing feature, $J_k$ are generally not suitable for direct use in determining the asymptotic stress field. Nevertheless, by choosing a small but finite cutoff radius $R$ for the circular integration path $\Gamma$ and denoting the "$J_{kR}$-integrals" as alternatives for $J_k$, we then rewrite Eqs. (9) and (10) as

$$J_{1R} = \frac{\Lambda}{E} [R^{2\lambda_1+1} a(\beta)(K_{\mathrm{I}})_\beta^2 + R^{2\lambda_2+1} b(\beta)(K_{\mathrm{II}})_\beta^2] \tag{11}$$

$$J_{2R} = -\frac{2\Lambda}{E} R^{\lambda_1+\lambda_2+1} c(\beta)(K_{\mathrm{I}})_\beta (K_{\mathrm{II}})_\beta \tag{12}$$

With the cutoff radius, the results of $J_{kR}$ appear to be dependent on the choice of $R$ and turn out to be of finite values. The SIFs and, consequently, the corresponding stress field can then be determined should the integrals be properly evaluated. However, in order to have appropriate solutions for $(K_{\mathrm{I}})_\beta$ and $(K_{\mathrm{II}})_\beta$, it is required that $R$ be taken in the region dominated by the asymptotic singular field $\boldsymbol{\sigma}^h$. Singular behavior is thus always involved in the calculation.

## 4 Modified Path-Independence

As described, the integration path $\Gamma$ for $J_{kR}$ is defined as a counterclockwise circular contour with center at the notch-tip $O$ and of small radius $R$. In finite element calculations, the discretized solutions will, in general, describe the behavior around the notch-tip $O$ more or less accurately, depending on the degree of local grid refinement and/or the adoption of special singular elements. Therefore, direct calculation of $J_{kR}$ along $\Gamma$ with numerical solutions appears to be difficult. Nevertheless, when there is no body force, we can alternatively have the $J_{kR}$-integrals rewritten as

$$J_{kR} = \int_{\Gamma_o} \left[ Wn_k - \sigma_{ij} n_j \left( \frac{\partial u_i}{\partial x_k} \right) \right] ds + \int_{C_1+C_2} Wn_k - t_i^n \left( \frac{\partial u_i}{\partial x_k} \right) ds \tag{13}$$

where $\Gamma_o$ is an arbitrary outer counterclockwise contour, and $C_1$ and $C_2$ are the portions of line segments along the curved notch edges, which are enclosed by $\Gamma_o$ and are terminated at a distance of $R$ away from the tip $O$, as shown in Fig. 1. Note that the integrand in the last term of Eq. (13) accounts for the contribution from the notch edge tractions.

For the first component of $J_{kR}$ (i.e., $J_{1R}$) under the special condition corresponding to a crack (i.e., $\beta=0$) with straight and traction-free crack surfaces, the last term on the right-hand side of Eq. (13) vanishes. The integration then reduces to the conventional $J_1$-integral for crack problems. In such a case, the value of integration remains unchanged along any arbitrarily chosen outer contour $\Gamma_o$ and this property is the well-known path-independence. However, for the general condition when $\beta>0$ with curved and stressed notch edges, the idea of path-independence for both $J_{1R}$ and $J_{2R}$ does not hold in the same manner. As a matter of fact, the concept needs to be modified by including the extra line integrals along the curved notch edges $C_1$ and $C_2$. Note that the outer contour $\Gamma_o$ can be arbitrarily chosen, except for the requirements to be inside the body, outside the cutoff radius $R$, and contain no other singularity in it. Such a concept, where $C_1$ and $C_2$ need to be included in the integration, has also been presented by Chen and Lu [18] for problems containing a V-shaped blunt notch.

Although the remote path $\Gamma_o$ can be chosen arbitrarily, the extra line segments $C_1$ and $C_2$ should both be terminated in the near-tip region. With these portions of line integrals, the asymptotic singular behavior is thus inevitably involved in the calculation. Cautious investigation for the numerical results is therefore necessary.

## 5 Numerical Examples

Three numerical example problems are presented in the following two subsections. In the first subsection, a notched specimen with straight and traction-free notch edges is considered. In the second subsection, we consider a notched specimen with curved and pressurized notch edges. These example problems are analyzed by using finite elements. Quadratic finite elements are used for displacement interpolation in the calculation. No particular singular element is used throughout the study.

**5.1 Straight and Traction-Free Notched Specimen.** The feasibility of $J_{kR}$ for a notch-tip with straight and traction-free edges has been demonstrated in the authors' previous work [9]. The aim of the following problem is thus to illustrate the numerical feature of $J_{kR}$ and, also, for later comparison with the results presented in the following subsection.

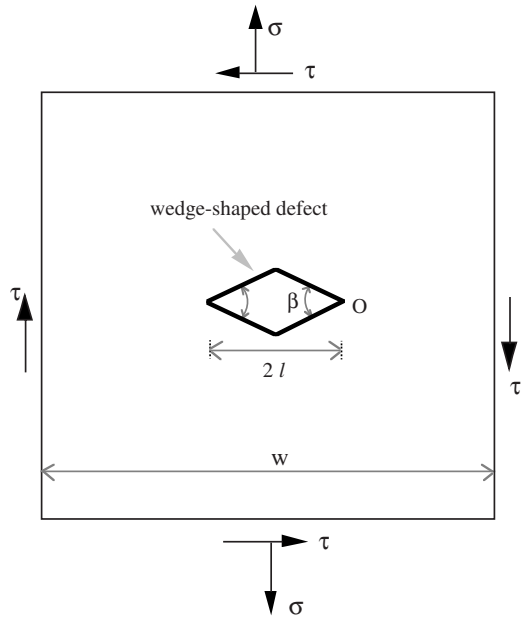



**Fig. 4 An elastic body containing a wedge-shaped defect**

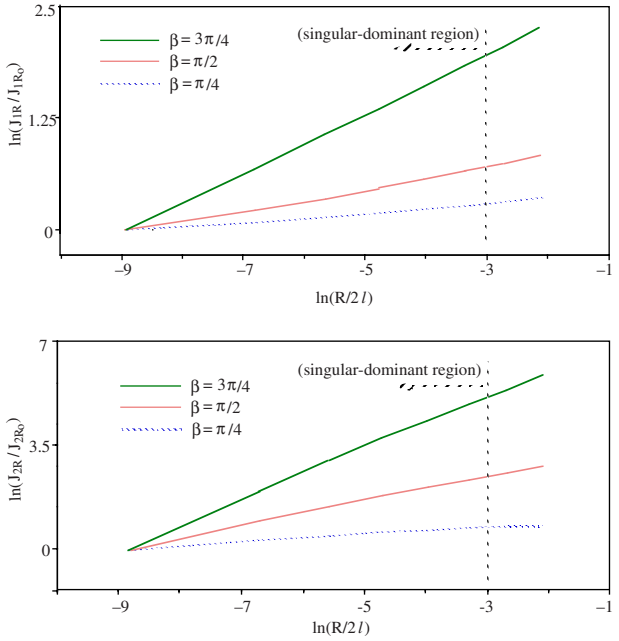



**Fig. 5 The variations of $J_{kR}/J_{kRo}$ with respect to $R/2l$ for $(\sigma,\tau)=(30,-20)$ kPa (loading case (i), mixed-mode, Problem 1)**

*5.1.1 Problem 1*. In this example problem, we consider a plane strain elastic specimen containing a central wedge-shaped defect and is subjected to combined loads $(\sigma,\tau)$, as shown in Fig. 4. The length of the defect $2l$ is relatively small compared with the size of the specimen so that the effect due to finite width can be neglected. Three instances of different notch angles (with $\beta$ equal to $\pi/4$, $\pi/2$, and $3\pi/4$, respectively) are taken. Two different loading conditions are considered, i.e., (Case (i)) a mixed-mode load with $(\sigma,\tau)=(30,-20)$ kPa (Case (ii)) and a Mode I load with $(\sigma,\tau)=(30,0)$ kPa.

The feature for the solutioins of $J_{kR}$ with respect to different selections of cutoff radius $R$ is examined. In order to properly characterize the near-tip behavior, it is necessary that $R$ be taken small enough to be inside the zone of dominance of the singular solution. Furthermore, to effectively illustrate the variation in $J_{kR}$ with respect to $R$, the solution at $R_o=0.25\times10^{-3}l$ is (arbitrarily) chosen as a reference. The normalized solutions of $J_{kR}/J_{kRo}$ for loading cases (i) and (ii) are depicted as functions of the scaled cutoff radius $R/2l$, as shown in Figs. 5 and 6, respectively. Note that the results of $J_{2R}$ for Case (ii) are not included in Fig. 6 in that they vanish under Mode I loading.

The asymptotic slope of each curve extracted from the numerical fields, along with the analytical solutions for the order of the leading terms $J_{1R}$ and $J_{2R}$ (as addressed in Eqs. (11) and (12)), are listed in Table 1 for loading case (i). We observe that the computed results for the slope of $\ln(J_{2R})$ appear to be well consistent with the values of $\lambda_1+\lambda_2+1$, with the errors remaining under 3%. On the other hand, substantial deviations between either $2\lambda_1+1$ or $2\lambda_2+1$ and the results of $\ln(J_{1R})$ are observed for $\beta=\pi/4$ and $\pi/2$ because the asymptotic solution of $J_{1R}$ is governed by the appearance of both $r^{2\lambda_1+1}$ and $r^{2\lambda_2+1}$ under the mixed-mode loads. Nevertheless, the influence of $r^{2\lambda_2+1}$ becomes less significant when $\beta$ increases so that the slope of $\ln(J_{1R})$ appears to be closer to the value of $2\lambda_1+1$, as is evident in the case of $\beta=3\pi/4$. As to loading case (ii), the asymptotic slope of each curve, along with the analytical solutions for the order of the leading term $J_{1R}$, are listed in Table 2. As shown, the numerical results are well compatible with the analytical values.

For problems with split singularities (i.e., the stronger and the weaker singularities), it is always necessary to investigate the proportions and the effects of both singular terms. To this end, a dimensionless parameter termed the "local mode mixity" $\eta$ was defined as

$$\eta=-\frac{K_{\mathrm{II}}}{K_{\mathrm{I}}}\left(\frac{r}{2l}\right)^{\lambda 2-\lambda 1} \tag{14}$$

The above parameter was originally proposed for the condition with a crack at a bimaterial interface [19]. Here, in order to characterize the proportion of the two modes for the notch problem, the associated local mode mixity for loading case (i) is evaluated and the variations of $|\eta|$ with respect to $r/2l$ is shown in Fig. 7. By observing the results from the three notch angles (i.e., $\beta=0$, $\pi/4$, and $\pi/2$) under the same loading condition, it is indicated

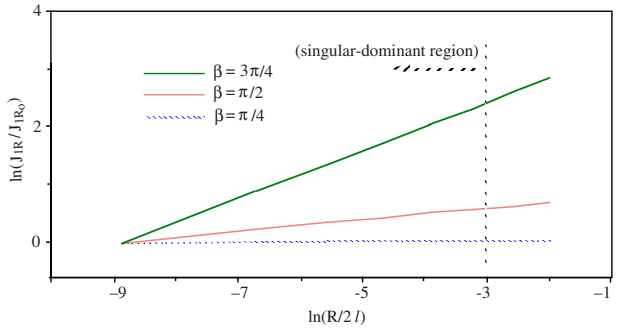



**Fig. 6 The variation of $J_{1R}/J_{1Ro}$ with respect to $R/2l$ for $(\sigma,\tau)=(30,0)$ kPa (loading case (ii), Mode I, Problem 1)**

**Table 1 The asymptotic slopes of $\ln(J_{kR})$ for Problem 1 (loading case (i), mixed-mode). Note that $w=30$ m, $l=1.5$ m, $E=207$ GPa, $\nu=0.3$, and $(\sigma,\tau)=(30,-20)$ kPa.**

| | $\ln(J_{1R})$ | $\ln(J_{2R})$ | $2\lambda_1+1$ | $2\lambda_2+1$ | $\lambda_1+\lambda_2+1$ |
|---|---|---|---|---|---|
| $\beta=\pi/4$ | 0.04 | 0.169 | 0.01 | 0.319 | 0.165 |
| $\beta=\pi/2$ | 0.102 | 0.465 | 0.089 | 0.817 | 0.453 |
| $\beta=3\pi/4$ | 0.349 | 0.998 | 0.347 | 1.604 | 0.976 |

**Table 2 The asymptotic slopes of $\ln(J_{1R})$ for Problem 1 (loading case (ii), Mode I). Note that $w=30$ m, $l=1.5$ m, $E=207$ GPa, $\nu=0.3$, $(\sigma,\tau)=(30,0)$ kPa.**

| | $\ln(J_{1R})$ | $2\lambda_1+1$ |
|---|---|---|
| $\beta=\pi/4$ | 0.009 | 0.01 |
| $\beta=\pi/2$ | 0.088 | 0.089 |
| $\beta=3\pi/4$ | 0.351 | 0.347 |

that the weaker singularity makes less contribution in the near-tip region for problems with larger notch angles. Also, such effect becomes more significant as one moves closer to the notch-tip.

**5.2 Curved and Pressurized Notched Specimen.** In this subsection, we consider a plane strain elastic specimen containing a central curved wedge-shaped defect. The defect is of asymptotic tangential angle $\beta$ at its two tips and consists of circular arc segments elsewhere (Fig. 8). Still, the span of the defect $2l$ is relatively small compared with the size of the specimen. The specimen is subjected to remote combined loads $(\sigma,\tau)$, as well as pressure $p$ along the edges of the wedge.

Two numerical problems are presented in this subsections. In the first problem, the $J_{kR}$-integrals are evaluated and the associated feature of the numerical results is investigated. Subsequently, the mixed-mode SIFs are evaluated in the second example problem.

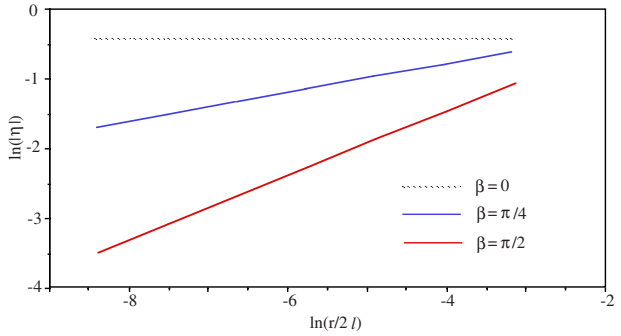



**Fig. 7 The variations in local mode mixity $\eta$ with respect to $r/2l$ for $(\sigma,\tau)=(30,-20)$ kPa (loading case (i), mixed-mode, Problem 1)**

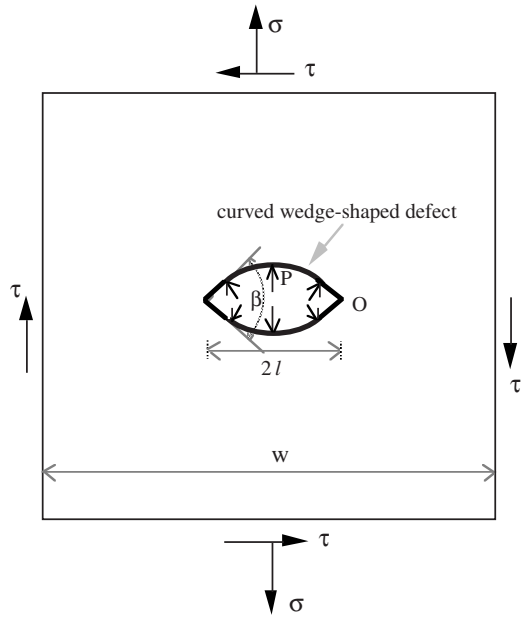
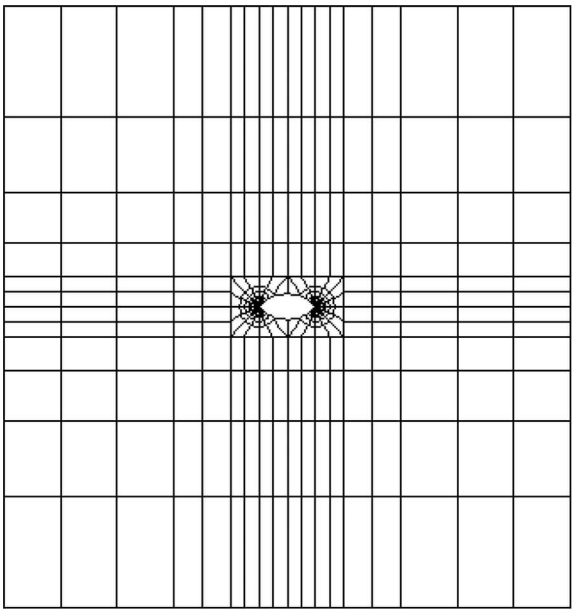



**Fig. 8 A homogeneous elastic body containing a curved wedge-shaped and pressurized defect**



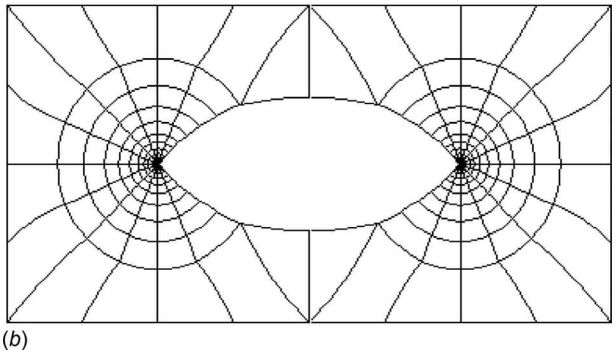




**Fig. 9 (*a*) The finite element model (of $\beta=\pi/2$) for the specimen in Fig. 8 and (*b*) the local finite element mesh in the near-tip area**

*5.2.1 Problem 2.1: The $J_{kR}$-integrals.* This problem is presented to illustrate the computation procedure and to study the behavior of the $J_{kR}$-integrals. The study is organized as follows. First, the property of the modified path-independence is examined. Next, the feature of $J_{kR}$ with respect to different selections of $R$ is investigated. Subsequently, the asymptotic behavior of $J_{kR}$ is observed. Finally, the size of the singular-dominant zone is examined. In the following calculations, three geometric instances of different notch angles (with $\beta$ equal to $\pi/4$, $\pi/2$, and $3\pi/4$, respectively) are considered. Also, two different loading conditions are considered, i.e., (Case (iii)) a mixed-mode load with $(\sigma,\tau,p)=(30,-20,10)$ kPa and (Case (iv)) a Mode I load with $(\sigma,\tau,p)=(30,0,10)$ kPa. The specimen is analyzed by using the finite element representation (e.g., of $\beta=\pi/2$) shown in Fig. 9(*a*). The discretized model is progressively refined as the elements approaching the notch-tip O. Details of the local mesh in the vicinity of the near-tip region is shown in Fig. 9(*b*).

To demonstrate the property of modified path-independence, three integration paths, each enclosing different region of the above finite element mesh, are defined in the calculation. The associated exterior contours $\Gamma_o$'s of these paths (for the instance

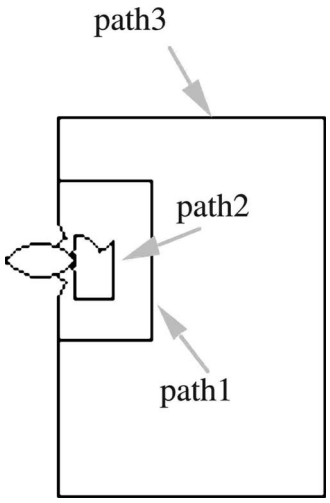



**Fig. 10 Three integration paths for the $J_{kR}$-integrals for the instance of $\beta=\pi/2$**

of, e.g., $\beta=\pi/2$) are depicted in Fig. 10. As illustrated in Eq. (13), the calculation consists of two parts, including the integrations along $\Gamma_o$ and the portions of the curved notch edges $C_1+C_2$. The results of each part for $R=(1.32\times10^{-2}l)$ m are shown in Table 3 (load case (iii)) and Table 4 (load case (iv)), respectively. We observe that the integration along $C_1+C_2$ makes rather significant contribution to the computation and thus accounts for the "modified" sense of path-independence. Note that, while both $\Gamma_o$ and $C_1+C_2$ yield finite-valued results in their separate integration under Mode I loading (Case (iv)), the solutions for $J_{2R}$ (i.e., the summation of both) are vanishingly small, as anticipated. As a summary, the finite element calculation from different paths, in general, yields very similar results of $J_{kR}$. Note that although the formulation is verified to be analytically path-independent in the modified sense, slight deviations among the results are observed because the equilibrium state is satisfied only weakly in the finite element computation.

Next, the behavior of $J_{kR}$ with respect to various values of $R$ is examined. Again, by arbitrarily choosing the solution with respect to $R_o(=0.25\times10^{-3}l)$ as the reference, the normalized results of $J_{kR}/J_{kRo}$ for loading cases (iii) and (iv) are depicted as functions of the scaled cutoff radius $R/2l$ and are shown in Figs. 11 and 12, respectively. Still, the results of $J_{2R}$ for Case (iv) are not included in Fig. 12 in that they vanish under Mode I loading.

The asymptotic slope for each curve in Fig. 11 (i.e., loading case (iii)), along with the analytical solutions for the order of the leading terms $J_{1R}$ and $J_{2R}$, are listed in Table 5. Still, the computed results for the slope of $\ln(J_{2R})$ appear to be well consistent with the values of $\lambda_1+\lambda_2+1$, with the errors remaining under 1.5%. On the other hand, the results for $\ln(J_{1R})$ are deviated from $2\lambda_1+1$ for $\beta=\pi/4$ and $\pi/2$ due to the appearance of both $r^{2\lambda_1+1}$ and $r^{2\lambda_2+1}$ in the asymptotic solution under mixed-mode loads. The influence of $r^{2\lambda_2+1}$ becomes less significant when $\beta$ increases so that the slope of $\ln(J_{1R})$ appears to be closer to the value of $2\lambda_1+1$, as is evident in the case of $\beta=3\pi/4$. As to loading case (iv), the asymptotic slope of each curve, along with the analytical solutions for the order of the leading term $J_{1R}$, are listed in Table 6. The numerical results are well compatible with the analytical values.

Finally, the size of the singular-dominant zone under different geometric and loading conditions is observed. For loading case (iii), the curves in Fig. 11 appear to remain linear in the region within 0.5% of the defect span and the values of the asymptotic

**Table 3 Path-independence for Problem 2.1 (unit: $10^{-2}$ Pa m) (Case (iii), mixed-mode). Note that $w=30$ m, $l=1.5$ m, $R/l=1.32\times10^{-2}$, $E=207$ GPa, and $\nu=0.3$.**

| | | $\int_{\Gamma_o}+\int_{C_1+C_2}=J_{1R}$ (Eq. (13)) | | | $\int_{\Gamma_o}+\int_{C_1+C_2}=J_{2R}$ (Eq. (13)) | | |
|---|---|---|---|---|---|---|---|
| $\beta=\pi/4$ | Path 1 | 3.472 | 0.157 | 3.629 | 1.605 | 0.004 | 1.609 |
| | Path 2 | 3.677 | −0.065 | 3.612 | 2.144 | −0.522 | 1.622 |
| | Path 3 | 3.465 | 0.150 | 3.615 | 1.611 | 0.003 | 1.614 |
| $\beta=\pi/2$ | Path 1 | 3.841 | −1.090 | 2.751 | 0.895 | −0.620 | 0.275 |
| | Path 2 | 3.236 | −0.445 | 2.791 | 1.184 | −0.912 | 0.272 |
| | Path 3 | 3.760 | −1.000 | 2.760 | 0.840 | −0.562 | 0.278 |
| $\beta=3\pi/4$ | Path 1 | 4.141 | −3.034 | 1.107 | 2.013 | −1.964 | 0.049 |
| | Path 2 | 2.723 | −1.628 | 1.095 | 0.612 | −0.567 | 0.045 |
| | Path 3 | 3.562 | −2.460 | 1.102 | 1.397 | −1.351 | 0.046 |

**Table 4 Path-independence for Problem 2.1 (unit: $10^{-2}$ Pa m) (Case (iv), Mode I). Note that $w=30$ m, $l=1.5$ m, $R/l=1.32\times10^{-2}$, $E=207$ GPa, and $\nu=0.3$.**

| | | $\int_{\Gamma_o}+\int_{C_1+C_2}=J_{1R}$ (Eq. (13)) | | | $\int_{\Gamma_o}+\int_{C_1+C_2}=J_{2R}$ (Eq. (13)) | | |
|---|---|---|---|---|---|---|---|
| $\beta=\pi/4$ | Path 1 | 2.475 | 0.803 | 3.278 | 0.004 | −0.004 | 0.000 |
| | Path 2 | 2.982 | 0.287 | 3.269 | 0.004 | −0.003 | 0.001 |
| | Path 3 | 2.520 | 0.752 | 3.272 | 0.004 | −0.004 | 0.000 |
| $\beta=\pi/2$ | Path 1 | 2.608 | 0.004 | 2.612 | 0.282 | −0.274 | 0.008 |
| | Path 2 | 2.907 | −0.235 | 2.672 | −0.068 | 0.080 | 0.012 |
| | Path 3 | 2.612 | 0.023 | 2.635 | 0.284 | −0.274 | 0.010 |
| $\beta=3\pi/4$ | Path 1 | 2.872 | −1.924 | 0.948 | 0.061 | −0.060 | 0.001 |
| | Path 2 | 2.460 | −1.519 | 0.941 | 0.026 | −0.025 | 0.001 |
| | Path 3 | 2.811 | −1.868 | 0.943 | 0.060 | −0.060 | 0.000 |

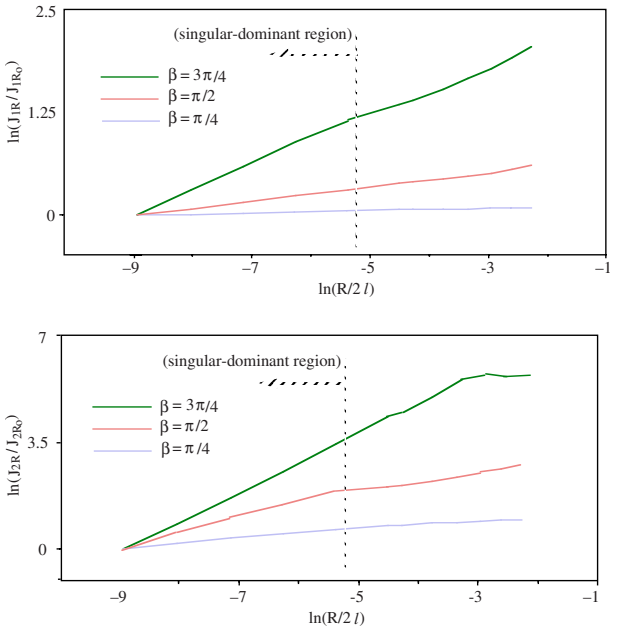



**Fig. 11 The variations of $J_{kR}/J_{kRo}$ with respect to $R/2l$ for $(\sigma,\tau,p)=(30,-20,10)$ kPa (loading case (iii), mixed-mode, Problem 2.1)**

slope are accurately calculated, as discussed previously. The numerical results for $J_{kR}$ in this region are thus well consistent with the asymptotic solutions shown in Eqs. (11) and (12). Such observed feature indicates that, in this case, the extent of the singular-dominant zone is within 0.5% of the defect span. Similarly, for loading case (iv), the linear portions of the curves in Fig. 12 indicate that the asymptotic singular region is within 0.8% of the defect span. On the other hand, for straight and traction-free notch edges (Problem 1), the singular-dominant region extends up to 5% of the defect span (Figs. 5 and 6). From the above results, it is observed that the size of the singular zone is significantly affected by the appearance of curved edges and the edge tractions. As a summary, it is demonstrated that, by cautionsly investigating

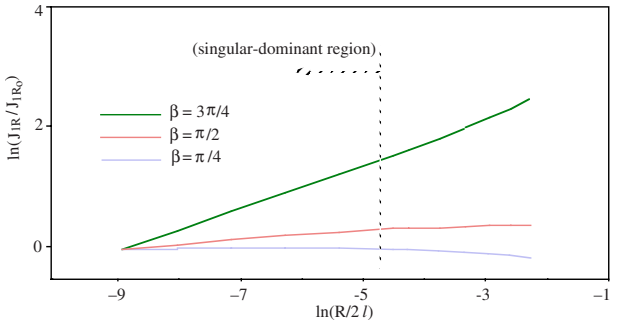



**Fig. 12 The variation of $J_{1R}/J_{1Ro}$ with respect to $R/2l$ for $(\sigma,\tau,p)=(30,0,10)$ kPa**

**Table 5 The asymptotic slopes of $\ln(J_{kR})$ for Problem 2.1 (loading case (iii), mixed-mode). Note that $w=30$ m, $l=1.5$ m, $E=207$ GPa, $\nu=0.3$, and $(\sigma,\tau,p)=(30,-20,10)$ kPa.**

| | $\ln(J_{1R})$ | $\ln(J_{2R})$ | $2\lambda_1+1$ | $2\lambda_2+1$ | $\lambda_1+\lambda_2+1$ |
|---|---|---|---|---|---|
| $\beta=\pi/4$ | 0.051 | 0.167 | 0.01 | 0.319 | 0.165 |
| $\beta=\pi/2$ | 0.147 | 0.459 | 0.089 | 0.817 | 0.453 |
| $\beta=3\pi/4$ | 0.343 | 0.966 | 0.347 | 1.604 | 0.976 |

**Table 6 The asymptotic slopes of $\ln(J_{1R})$ for Problem 2.1 (loading case (iv), mode-I). Note that $w=30$ m, $l=1.5$ m, $E=207$ GPa, $\nu=0.3$, and $(\sigma,\tau,p)=(30,0,10)$ kPa.**

| | $\ln(J_{1R})$ | $2\lambda_1+1$ |
|---|---|---|
| $\beta=\pi/4$ | 0.009 | 0.01 |
| $\beta=\pi/2$ | 0.085 | 0.089 |
| $\beta=3\pi/4$ | 0.343 | 0.347 |

the values of $J_{kR}$ resulting from different choices of $R$, the extent of notch-tip singularity for any particular geometry can be rationally and effectively characterized.

Since the concept of the $J_{kR}$-integrals for notches with curved and pressurized edges is originally presented in this paper, there is no analytical or numerical solution with which a direct test of the above computation scheme for the contour integrals can be carried out. Nevertheless, the numerical results show that the integrals are path-independent (in a modified sense). Also, the computed values of the order of their leading terms in the asymptotic near-tip region are well consistent with those of the analytical solutions. The feasibility of the proposed integrals can thus be appropriately demonstrated by the observation and comparison.

*5.2.2 Problem 2.2: The Mixed-Mode SIFs.* In this problem, we consider again the plane strain notched specimen with curved and pressurized notch edges and subjected to a combined load, as shown in Fig. 8. The mixed-mode SIFs, $(K_I)_\beta$ and $(K_{II})_\beta$ can be determined by the use of Eqs. (11) and (12) provided that the $J_{kR}$-integrals are properly evaluated.

Since the cutoff radius $R$ for $J_{kR}$ can be arbitrarily chosen, it is therefore anticipated that the resulting values of SIFs be independent of the selection of $R$. To demonstrate such a feature, we herein choose a number of different values of $R$, evaluate the corresponding $J_{kR}$-integrals, and then calculate the corresponding mixed-mode SIFs. Note that from the results in Problem 2.1 it is suggested that the scale of $R$ be within the range of $0.01l$ in order for accurate solutions of the SIFs. The numerical results for the SIFs associated with various values of $R$ for loading case (iii) are listed in Table 7. The results of $(K_I)_\beta$ and $(K_{II})_\beta$ appear to be almost invariant with respect to different selections of $R$, with maximum deviation under 3%, as expected.

It is interesting to note that, in Table 7, the value of $(K_{II})_{3\pi/4}$ is not available numerically due to the appearance of the regular leading term $r^{1.604}$. Such behavior has also been previously illustrated by the vanishingly small values of $J_{2R}$ listed in Table 3. As a matter of fact, since the asymptotic Mode II stress component is relatively negligible when the notch angle is greater than $0.57\pi$, the solution for $(K_{II})_\beta$ is therefore rather insignificant in such cases. Also listed in Table 7 are the corresponding (absolute values of) local mode mixity $|\eta|$. From the numerical results, it is observed that the weaker singularity makes less contribution as the notch angle increases from $\pi/4$ to $\pi/2$ and finally vanishes for $\beta=3\pi/4$, as anticipated. Still, such effect is more significant as one moving closer to the notch-tip.

As to loading case (iv), the values of the Mode I SIFs $(K_I)_\beta$ are shown in Table 8. The numerical results appear to be almost invariant with respect to different selections of $R$. Also, they are well consistent with the results of $(K_I)_\beta$, evaluated from the mixed-mode solutions shown in Table 7, as expected.

Although there is no analytical solution for direct verification of the above calculation, the validity of the formulation for the SIFs can nevertheless be alternatively addressed by observing the following two points. First, the numerical results of $(K_I)_\beta$ and $(K_{II})_\beta$ appear to be invariant with respect to different selections of the cutoff radius $R$, which is consistent with the concept indicated by Eqs. (11) and (12). Also, the computed values of $(K_I)_\beta$ under

**Table 7 The results of $(K_I)_\beta$ and $(K_{II})_\beta$ versus $R$ (Case (iii), mixed-mode) (unit: $10^4$ Pa $m^{-\lambda 1}$ for $(K_I)_\beta$, $10^4$ Pa $m^{-\lambda 2}$ for $(K_{II})_\beta$). Note that $w=30$ m, $l=1.5$ m, $E=207$ GPa, $\nu=0.3$, and $(\sigma,\tau,p)=(30,-20,10)$ kPa.**

| | $R/2l(10^{-3})$ | 0.294 | 0.713 | 1.722 | 4.15 |
|---|---|---|---|---|---|
| $\beta=\pi/4$ | $(K_I)_{\pi/4}$ | 8.557 | 8.564 | 8.536 | 8.501 |
| | $(K_{II})_{\pi/4}$ | −6.176 | −6.202 | −6.240 | −6.222 |
| | $\lvert\eta\rvert$ | 0.205 | 0.236 | 0.273 | 0.314 |
| $\beta=\pi/2$ | $(K_I)_{\pi/2}$ | 8.861 | 8.861 | 8.890 | 8.904 |
| | $(K_{II})_{\pi/2}$ | −8.508 | −8.586 | −8.409 | −8.458 |
| | $\lvert\eta\rvert$ | 0.050 | 0.069 | 0.093 | 0.129 |
| $\beta=3\pi/4$ | $(K_I)_{3\pi/4}$ | 11.10 | 11.20 | 10.93 | 11.10 |
| | $(K_{II})_{3\pi/4}$ | – | – | – | – |
| | $\lvert\eta\rvert$ | – | – | – | – |

**Table 8 The results of $(K_I)_\beta$ versus $R$ (Case (iv), Mode I) (unit: $10^4$ Pa $m^{-\lambda 1}$ for $(K_I)_\beta$). Note that $w=30$ m, $l=1.5$ m, $E=207$ GPa, $\nu=0.3$, and $(\sigma,\tau,p)=(30,0,10)$ kPa.**

| | $R/2l(10^{-3})$ | 0.294 | 0.713 | 1.722 | 4.15 |
|---|---|---|---|---|---|
| $\beta=\pi/4$ | $(K_I)_{\pi/4}$ | 8.543 | 8.494 | 8.550 | 8.522 |
| $\beta=\pi/2$ | $(K_I)_{\pi/2}$ | 8.826 | 8.812 | 8.918 | 8.854 |
| $\beta=3\pi/4$ | $(K_I)_{3\pi/4}$ | 10.83 | 11.07 | 10.98 | 10.87 |

loading case (iii) (mixed-mode) are well consistent with those under Case (iv) (Mode I). Such agreement is analytically supported by the mixed-mode feature described by Eqs. (2)–(4).

As an aside, it is important to note that the mixed-mode asymptotic singular stress field cannot be completely described by only using the SIFs in that the Mode I and Mode II stresses are governed by different orders of singularity and dependent on the tangential notch angle. A detailed structure of the singular stress field is therefore required in engineering fracture analysis. This can be achieved by substituting the SIFs into Eqs. (2)–(4), with which the asymptotic singular stress field can be appropriately calculated.

## 6 Conclusions

A numerical procedure is presented for the evaluation of the mixed-mode SIFs, i.e., $(K_I)_\beta$ and $(K_{II})_\beta$, for a sharp notch-tip with curved edges and subjected to surface tractions along the notch edges. This actually is a generalized extension of the earlier work conducted by the authors on the corresponding notched problems with straight and traction-free notch edges. In this approach, a pair of contour integrals $J_{kR}$ are defined by modifying the conventional $J_k$-integrals with the introduction of arbitrarily selected cutoff radius $R$ in the near-tip area. Furthermore, the relation between $J_{kR}$ and the SIFs is analytically established and expressed as functions of the asymptotic tangential notch angle $\beta$. Once the $J_{kR}$-integrals are evaluated, the SIFs and, consequently, the near-tip mixed-mode singular stress field can then be completely determined.

The $J_{kR}$-integrals are shown to be path-independent in a modified sense. That is, the integratioin can be performed along an arbitrarily chosen outer contour, along with the curve segments on the notch edges. With this property, accurate solutions can be achieved without using any particular numerical treatment. The feasibility of the presented approach is demonstrated in the numerical examples. As a matter of fact, it is verified both analytically and numerically that the results of $(K_I)_\beta$ and $(K_{II})_\beta$ are actually insensitive to different selections of cutoff radius $R$. Also, no a priori auxiliary solutions are required in the formulation. With the above superiorities, $J_{kR}$ thus appear to be efficient as a computational scheme in practice.

As an aside, it is important to note that the extent of the singular-dominant zone is significantly affected by the appearance of curved edges and the edge tractions. Such a feature can be effectively characterized by proper use of the $J_{kR}$-integrals through detailed investigations of the computed asymptotic behavior.


## Acknowledgment

This work has been partially supported by National Science Council Grant No. NSC 95-2211-E-008-096 to National Central University.


## Appendix

For the geometric configuration containing a sharp notch-tip (Fig. 1), the near-tip stress field can be formulated by adopting a specific airy stress function [1] and expressed in terms of expansions as follows

$$\sigma_r(r,\theta;\beta)=\sum_{i=1}^{\infty} r^{\lambda i-1}\left[\frac{d^2F_i(\theta)}{d\theta^2}+(\lambda_i+1)F_i(\theta)\right] \tag{A1}$$

$$\sigma_\theta(r,\theta;\beta)=\sum_{i=1}^{\infty}\lambda_i(\lambda_i+1)r^{\lambda i-1}F_i(\theta) \tag{A2}$$

$$\tau_{r\theta}(r,\theta);\beta=\sum_{i=1}^{\infty}-\lambda_i r^{\lambda i-1}\frac{dF_i(\theta)}{d\theta} \tag{A3}$$

where $\lambda_i$ is the order of the $i$th term remaining to be solved, and $F_i(\theta)$ is the $i$th angular function satifying the following compatibility equation:

$$\frac{d^4F_i(\theta)}{d\theta^4}+2(\lambda_i^2+1)\frac{d^2F_i(\theta)}{d\theta^2}+(\lambda_i^2-1)^2F_i(\theta)=0 \tag{A4}$$

Note that, in practice, truncated expressions of the above stress functions are usually utilized by taking, say, up to N terms of the expansion.

Due to the presence of the traction on the notch edges, the near-tip stresses thus appear to consist of the homogeneous and the particular parts, i.e., $\boldsymbol{\sigma}^h$ and $\boldsymbol{\sigma}^p$. Note that $\boldsymbol{\sigma}^h$ and $\boldsymbol{\sigma}^p$ can be expressed in terms of the same form of stress functions, yet with different combinations of expansions. For $\boldsymbol{\sigma}^h$, as shown in Eqs. (2)–(4), they were obtained by solving Eqs. (A1)–(A3) with the traction-free boundary condition on the notch edges [1]. As to $\boldsymbol{\sigma}^p$, on the other hand, we consider the traction vector $\mathbf{t}^n$ on the notch edges, i.e.,

$$\boldsymbol{\sigma}^p(r,\theta;\beta)\big|_{\theta=\pm(\pi-\beta/2)} \cdot \mathbf{n} = \mathbf{t}^n(r) \tag{A5}$$

By comparing Eqs. (A1)–(A3) with the prescribed boundary condition (A5), it can be seen that solutions of $\boldsymbol{\sigma}^p$ exist only if $\lambda_i \geq 1$ (for $i=1,2,\ldots,N$) in that $\mathbf{t}^n$ is of finite value. This, as a consequnce, leads to a finite-valued solution of $\boldsymbol{\sigma}^p$.

## References


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

[2] Bogy, D. B., 1972, "The Plane Solution for Anisotropic Elastic Wedges Under Normal and Shear Loading," ASME J. Appl. Mech., **39**, pp. 1103–1109.

[3] Awaji, H., Toshimitsu, Y. A., and Yokobori, T., 1986, "The Variation of the Stress Singularity at the Notch Tip as Notch Angle and Radius of Curvature," Compos. Struct., **22**(1), pp. 25–30.

[4] Peng, X., 1986, "On Stress Singularities at Tips of Plane Notches," Mech. Res. Commun., **13**, pp. 173–178.

[5] Carpenter, W. C., 1984, "Calculation of Fracture Parameters for a General Corner," Int. J. Fract., **24**, pp. 45–58.

[6] Sinclair, G. B., Okajima, M., and Griffin, J. H., 1984, "Path Independent Integrals for Computing Stress Intensity Factors at Sharp Corners in Elastic Plates," Int. J. Numer. Methods Eng., **20**, pp. 999–1008.

[7] Babuska, I., and Miller, A., 1984, "The Post-Processing Approach in the Finite Element Method, Part 2: The Calculation of Stress Intensity Factors," Int. J. Numer. Methods Eng., **20**, pp. 1111–1129.

[8] Labossiere, P. E. W., and Dunn, M. L., 1998, "Calculation of Stress Intensities at Sharp Notches in Anisotropic Media," Eng. Fract. Mech., **61**, pp. 635–654.

[9] Chang, J. H., and Kang, L. K., 2002, "Evaluation of the Stress Field Around a Notch Tip Using Contour Integrals," Int. J. Eng. Sci., **40**, pp. 569–586.

[10] Sih, G. C., Paris, P. C., and Erdogan, F., 1962, "Crack Tip, Stress Intensity Factors for Plane Extension and Plate Bending Problems," ASME J. Appl. Mech., **29**, pp. 306–312.

[11] Dreilich, L., and Gross, D., 1985, "The Curved Crack," Z. Angew. Math. Mech., **65**, pp. 132–134.

[12] Chen, Y. Z., 2004, "Singular Integral Equation Method for the Solution of Multiple Curved Crack Problems," Int. J. Solids Struct., **41**, pp. 3505–3519.

[13] Chang, J. H., 1994, "On the Calculation of Energy Release Rate for Curved Cracks of Rubbery Materials," Finite Elem. Anal. Design, **17**, pp. 323–338.

[14] Beom, H. G., Earmme, Y. Y., and Choi, S. Y., 1994, "Energy Release Rates for an Arbitrarily Curved Interface Crack," Mech. Mater., **18**, pp. 195–204.

[15] Lorentzon, M., and Eriksson, K., 2000, "A Path Independent Integral for the Crack Extension Force of the Circular Arc Crack," Eng. Fract. Mech., **66**, pp. 423–439.

[16] Rice, J. R., 1968, "Mathematical Analysis in the Mechanics of Fracuture," *Fracture*, Vol. 2, H. Liebowitz, ed., Academic Press, New York, pp. 191–213.

[17] Eshelby, J. D., 1974, "Calculation of Energy Release Rate," *Prospects of Fracture Mechanics*, G. C. Sih, H. C. Van Elst, and D. Broek, eds., Noordhoff, Groningen, pp. 69–84.

[18] Chen, Y. H., and Lu, T. J., 2004, "On the Path Dependence of the J-Integral in Notch Problems," Int. J. Solids Struct., **41**, pp. 607–618.

[19] Zhang, Z., and Suo, Z., 2007, "Split Singularities and the Competition Between Crack Penetration and Debond at a Bimaterial Interface," Int. J. Solids Struct., **44**, pp. 4559–4573.

**Jun Yu**
Engineer
LZA Technology,
The Thornton Tomasetti Group, Inc.,
24 Commerce Street,
Newark, NJ 07102
e-mail: jy232@columbia.edu

**Maura Imbimbo**
Associate Professor
Department of Mechanics, Structures and Environment,
University of Cassino,
Via G. Di Biasio 43,
Cassino 03043, Italy
e-mail: mimbimbo@unicas.it

**Raimondo Betti**
Professor
Department of Civil Engineering and Engineering Mechanics,
Columbia University,
640 S.W. Mudd Building,
New York, NY 10027
e-mail: betti@civil.columbia.edu

# Identification of Linear Structural Systems With a Limited Set of Input-Output Measurements

*In this paper, a methodology is presented for the identification of the complete mass, damping, and stiffness matrices of a dynamical system using a limited number of time histories of the input excitation and of the response output. Usually, in this type of inverse problems, the common assumption is that the excitation and the response are recorded at a sufficiently large number of locations so that the full-order mass, damping, and stiffness matrices can be estimated. However, in most applications, an incomplete set of recorded time histories is available and this impairs the possibility of a complete identification of a second-order model. In this proposed approach, all the complex eigenvectors are correctly identified at the instrumented locations (either at a sensor or at an actuator location). The remaining eigenvector components are instead obtained through a nonlinear least-squares optimization process that minimizes the output error between the measured and predicted responses at the instrumented locations. The effectiveness of this approach is shown through numerical examples and issues related to its robustness to noise polluted measurements and to uniqueness of the solution are addressed.*
[DOI: 10.1115/1.3002336]

*Keywords: identification, reduced order models, damage detection*

## 1 Introduction

In recent years, there has been an increasing interest in the civil and mechanical engineering communities in methodologies that are capable of detecting and quantifying structural damage in areas of a structure that are not easily accessible to engineers. A variety of nondestructive techniques has been successfully developed, thanks to technological advancements in computer technology, in signal processing, and in materials. However, the vast majority of these techniques focuses on the individual structural element, e.g., a beam or a column. On the contrary, damage identification techniques and methodologies that have a "global" perspective are still in the development phase, trying to overcome limitations imposed by computational and financial constraints.

In these global approaches to damage detection, one key factor is the reliability of the structural model used. Comparing two reliable models that correspond to the undamaged and damaged configurations of a structure can provide a wealth of useful information for accurately determining the location and amount of structural damage. Some of the noteworthy efforts in the identification of linear structural systems include the works by Agbabian et al. [1], Safak [2,3], Udwadia [4], Beck and Katafygiotis [5,6], Luş et al. [7], Alvin et al. [8,9], Farhat et al. [10], and Doebling et al. [11]. Although they have the common goal of identifying a mathematical/physical model of the system, all these works differ from the type of model they identify and from the methodology used in the identification process.

One way of determining mathematical (not necessarily physical) models of the structure is to use dynamic measurements of the structural input and/or output to identify a "black box" model that properly maps inputs and outputs. In mechanical and aerospace applications, it is quite common to use a "first-order" representation of a dynamic system. However, although it is quite popular in control-type applications, this type of representation of a system, being a black box type, is rather inconvenient for damage identification purposes.

For identifying areas of structural damage, it is more convenient to work with "second-order" models where the system's dynamics is represented by a set of second-order differential equations whose coefficients are physical parameters, such as mass, damping, and stiffness. However, the identification of such structural parameters is much more cumbersome than the identification of first-order models. A common trend among the most successful methodologies for the identification of the mass, damping, and stiffness characteristics of a structure is to start from an identified first-order representation of the system and, through proper transformations, to obtain the physical structural parameters [7–9,12]. The drawback of this type of approaches is that these methodologies require a full set of instrumentation (a sensor and/or an actuator on each degree of freedom) for identifying the mass, damping, and stiffness matrices of the structural model. However, in real life applications, even complex structures are usually instrumented with only few sensors and this impairs the complete identification of structural properties. In the case of an incomplete set of measurements, only limited information about the distribution of the structural stiffness, mass, and damping can be retrieved [13]. In addition, dealing with a limited set of input-output measurements raises issues related to the uniqueness of the identified solution [14–17].

In this paper, the results of a study focused on determining the mass, damping, and stiffness matrices for systems with few input/output measurements are presented. The success of the identification is closely related to the complexity of the adopted model. Considering different levels of model complexity and assuming that the identification problem has a unique solution, a satisfactory approximation of the second-order model can be obtained through a nonlinear optimization process that takes into account the geometrical constraints imposed by the different structural assumptions.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 29, 2007; final manuscript received June 6, 2008; published online March 5, 2009. Review conducted by N. Sri Namachchivaya.

## 2 First-Order Representation of a System

Let us consider an $N$ degree-of-freedom model of a structural system whose equations of motion can be written as

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}_c\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{L}\mathbf{u}(t) \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}_c$, and $\mathbf{K}$ are the $N\times N$ mass, damping, and stiffness matrices of the system, respectively, and $\mathbf{q}(t)$ is the $N\times 1$ vector of the generalized nodal displacements. The symbol $(\dot{})$ indicates the differentiation with respect to time. The vector $\mathbf{u}(t)$ indicates the $r\times 1$ input vector containing the $r$ time histories of the external excitations, while the matrix $\mathbf{L}$ represents the $N\times r$ input matrix.

By introducing the state vector $\mathbf{s}(t)=[\mathbf{q}(t)^T\dot{\mathbf{q}}(t)^T]^T$ and indicating with $\mathbf{y}(t)$ the $m\times 1$ output vector containing the $m$ time histories of the output measurements, Eq. (1) can be transformed to the first-order canonical form:

$$\dot{\mathbf{s}}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{B}\mathbf{u}(t)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{s}(t) + \mathbf{D}\mathbf{u}(t) \tag{2}$$

where the system matrix $\mathbf{A}$ and the input matrix $\mathbf{B}$ are given by

$$\mathbf{A} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C}_c \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ \mathbf{M}^{-1}\mathbf{L} \end{bmatrix} \tag{3}$$

The output matrix $\mathbf{C}$ and the direct transition matrix $\mathbf{D}$ are matrices of dimensions $m\times 2N$ and $m\times r$ and depend on the types of measurement available (displacement, velocities, or accelerations). By operating transformations on the state vector $\mathbf{s}$, it is possible to obtain an infinite number of sets of matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ that, although different from those in Eq. (3), still correctly represent the dynamics of the system.

Assuming that the system is controllable and observable, a first-order representation of the system (a set of matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$) can be obtained just using a set of input and corresponding output time histories, independent of the number of degrees-of-freedom of the structural model. Among the most popular techniques, eigensystem realization algorithm with data correlation (ERA/DC) and the observer Kalman filter identification (OKID) provide very reliable representation of the system's dynamics even with few input/output time histories.

## 3 Identification of Mass, Damping, and Stiffness Matrices

Identifying the full-order mass, damping, and stiffness matrices of a system, $\mathbf{M}$, $\mathbf{C}_c$, and $\mathbf{K}$, is quite challenging. One approach to tackle this problem is called "model updating" in which an initial model of the structure (usually a finite element model) is created and its describing parameters are updated so to match, for a given input, the output data as closely as possible. Model updating methodologies, although exempted from the limitation of one sensor/actuator per degree-of-freedom, are quite time consuming and face issues linked to the uniqueness of the identified model.

An alternative approach to determine the full-order mass, damping, and stiffness matrices is to start from the identified first-order state space model of the structure and derive, through a set of transformations, the corresponding mass, damping, and stiffness matrices. One of these methodologies [12] determines such matrices using the complex eigenvalues and eigenvectors of the associated damped eigenvalue problem:

$$(\lambda_i^2\mathbf{M} + \lambda_i\mathbf{C}_c + \mathbf{K})\psi_i = 0 \tag{4}$$

where $\lambda_i$ and $\psi_i (i=1,2,\ldots,2N)$ represent the $i$th complex eigenvalue and the associated complex eigenvector, respectively. The complex eigenvectors can be reordered in an eigenvector matrix $\psi=[\psi_1\psi_2\ldots\psi_{2N}]$ of dimension $N\times 2N$, while the corresponding complex eigenvalues $\lambda_i (i=1,2,\ldots,2N)$ are rearranged in a diagonal matrix $\mathbf{\Lambda}$ of dimension $2N\times 2N$. These eigenvalues and eigenvectors can be obtained by applying proper transformations to the identified state, and input and output matrices of the first-order model.

Using proper normalization for the complex eigenvalues [12], the mass, damping, and stiffness matrices of the structural system can be expressed as functions of its complex eigenvalues and eigenvectors:

$$\mathbf{M} = (\psi\mathbf{\Lambda}\psi^T)^{-1}, \quad \mathbf{K} = -(\psi\mathbf{\Lambda}^{-1}\psi^T)^{-1}, \quad \mathbf{C}_c = -\mathbf{M}\psi\mathbf{\Lambda}^2\psi^T\mathbf{M} \tag{5}$$

with the eigenvectors satisfying the following condition:

$$\psi\psi^T = 0_{N\times N} \tag{6}$$

If the structural system has a sufficient number of sensors and actuators ($m+r=N+1$), then the full eigenvector matrix $\psi$ of the system can determined by using the collocation requirement and the information at either sensor or actuator locations [12]. However, in real life applications, a structural system is usually insufficiently instrumented with some degrees-of-freedom deficient of both sensors and actuators ($m+r<N+1$). Using either information from sensor or actuator locations and the collocation requirement, only the components of the vibrational modes at the measured locations can be uniquely determined. Using the $m+r$ available input/output data with one colocated sensor-actuator pair, only $n=m+r-1$ rows of the complex eigenvector matrix $\psi$ can be determined, while the remaining $p=N-n$ rows remain unknown [12,13]. By rearranging the known ($\psi^{\text{kn}}$) and unknown ($\psi^{\text{un}}$) components, the eigenvector matrix $\psi$ can be partitioned into a known submatrix $\psi_1=\psi^{\text{kn}}$ of dimension $n\times 2N$ and an unknown submatrix $\psi_2=\psi^{\text{un}}$ of dimension $p\times 2N$:

$$\psi = \begin{bmatrix} \psi^{\text{kn}}_{n\times 2N} \\ \psi^{\text{un}}_{p\times 2N} \end{bmatrix} = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \tag{7}$$

The superscripts "kn" and "un" indicate whether, at that degree of freedom, a sensor or an actuator is present or not, implying that the corresponding modal quantity can be identified (known) or not (unknown).

Starting from the general expression of the mass, damping, and stiffness matrices in terms of the complex eigenvectors $\psi$, as shown in Eq. (5), we can express these matrices in partitioned forms as functions of the known and unknown partitions of the eigenvector matrix. For example, the full-order stiffness matrix $\mathbf{K}$ can be expressed as

$$\mathbf{K} = -(\psi\mathbf{\Lambda}^{-1}\psi^T)^{-1} = -\begin{bmatrix} \psi_1\mathbf{\Lambda}^{-1}\psi_1^T & \psi_1\mathbf{\Lambda}^{-1}\psi_2^T \\ \psi_2\mathbf{\Lambda}^{-1}\psi_1^T & \psi_2\mathbf{\Lambda}^{-1}\psi_2^T \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{bmatrix} \tag{8}$$

with each subpartition given by the following expressions:

$$\mathbf{K}_{11} = -[(\psi_1\mathbf{\Lambda}^{-1}\psi_1^T) - (\psi_1\mathbf{\Lambda}^{-1}\psi_2^T)(\psi_2\mathbf{\Lambda}^{-1}\psi_2^T)^{-1}(\psi_2\mathbf{\Lambda}^{-1}\psi_1^T)]^{-1}$$

$$\mathbf{K}_{22} = -[(\psi_2\mathbf{\Lambda}^{-1}\psi_2^T) - (\psi_2\mathbf{\Lambda}^{-1}\psi_1^T)(\psi_1\mathbf{\Lambda}^{-1}\psi_1^T)^{-1}(\psi_1\mathbf{\Lambda}^{-1}\psi_2^T)]^{-1}$$

$$\mathbf{K}_{12} = -\mathbf{K}_{11}^{-1}(\psi_1\mathbf{\Lambda}^{-1}\psi_2^T)(\psi_2\mathbf{\Lambda}^{-1}\psi_2^T)^{-1}$$

$$\mathbf{K}_{21} = -\mathbf{K}_{22}^{-1}(\psi_2\mathbf{\Lambda}^{-1}\psi_1^T)(\psi_1\mathbf{\Lambda}^{-1}\psi_1^T)^{-1} \tag{9}$$

In the case of limited sensor/actuator capabilities, only the known part ($\psi_1$) of the eigenvector matrix is available and this will impair the use of Eq. (9).

Using only the known partition of the complex eigenvector matrix, a "reduced-order" form for the mass, stiffness, and damping matrices of the structural system of dimension $n\times n$ can be expressed as

$$\mathbf{M}_r = (\psi_1\mathbf{\Lambda}\psi_1^T)^{-1} = \mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{-1}\mathbf{M}_{21} \tag{10}$$

$$\mathbf{K}_r = -(\psi_1\mathbf{\Lambda}^{-1}\psi_1^T)^{-1} = \mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21} \tag{11}$$

$$\mathbf{C}_{cr} = -\mathbf{M}_r \psi_1 \mathbf{\Lambda}^2 \psi_1^T \mathbf{M}_r \tag{12}$$

The use of such "reduced" models of the structural system is quite limited. In fact, they cannot be used to represent the dynamic characteristics of the structural system: The natural frequencies obtained from the reduced model do not match the ones from the real system. In addition, such models cannot be used to predict the structural response to future excitation since they represent smaller order systems that have been obtained by minimizing, at few locations, the output error for a given (past) excitation. However, they can provide indications on the locations where damage has occurred, although quantifying the damage with such reduced models is quite difficult, if not impossible.

To overcome these limitations imposed by the use of reduced-order models, an attempt is made to expand reduced second-order models of a system, obtained from a limited number of instrumentations, to "full" second-order models, which preserve most of the system information such as the input/output relationships at the instrumented locations and the normalized system eigenvalues. This is accomplished by determining the unknown components of the complex eigenvector matrix, $\psi_2$, through an optimization algorithm that minimizes the difference between the measured and simulated data at the sensor locations under a set of constraints imposed on the variables to update.

## 4 Optimization Algorithm

The aim of this optimization process is to determine the $2pN$ parameters of the vector $\psi_2$ ($\psi_{ij}^{un}$ $(i=1,2,\ldots,p; j=1,2,\ldots,2N)$) that minimize the objective function $J(\psi_2)$, representing the error between the measured and predicted outputs at the instrumented locations, under a set of linear and nonlinear constraints imposed by the type of structural models to be identified. Thus, this optimization can be expressed in a convenient mathematical form as

$$\min_{\psi_2 \in \Re^{2pN}} J(\psi_2) = \frac{1}{2} \sum_{h=t_i}^{t_f} [\hat{\mathbf{y}}(h,\psi_2) - \mathbf{y}(h)]^T [\hat{\mathbf{y}}(h,\psi_2) - \mathbf{y}(h)] \tag{13}$$

with the conditions

$$c_l(\psi_2) = 0, \quad l = 1,2,\ldots,m'$$

$$c_l(\psi_2) \geq 0, \quad l = m'+1,\ldots,s \tag{14}$$

being $\hat{\mathbf{y}}(h,\psi_2)$ and $\mathbf{y}(h)$ the predicted and the measured outputs at the time step $h$, with the time index $h$ varying from an initial time $t_i$ to a final time $t_f$. The conditions $c_l(l=1,2,\ldots,s)$ represent equality/inequality constraints of the nonlinear optimization related to either (1) the conditions required by the eigenvector normalization process, (2) special geometrical properties characterizing the system, or (3) to the physical meanings of the system parameters (e.g., floor stiffness cannot have negative values).

Any point $\psi_2$ that satisfies all the constraints $c_k$ is said to be *feasible*. The set of all the feasible points is termed the *feasible region* or *solution space*. Considering now that $\psi_2$, as defined in Eq. (7), contains $2pN$ undetermined components, then the initial dimension of its solution space, before any constraint is applied, is $2pN$. The introduction of any equality constraint will reduce this initial dimension. For details, see the work by Yu [17].

## 5 Expanding Reduced Order Models to "Full Order"

For a structural engineer, having a full order representation of the system, represented by the complete mass, damping, and stiffness matrices, is quite useful. In fact, such a representation can be used for predicting the structural response to future excitations at all possible locations in the structure as well as for locating and quantifying the possible structural damage (e.g., by comparing the structural stiffness). This is not possible using the reduced-order models of the structure: In fact, such models cannot be used for predicting the structural response to future excitation even at instrumented locations, while locations and amount of structural damage can only be vaguely characterized [13].

**Table 1 Number of total constraints and dimension of solution space for the different representative scenarios of reduced order systems**

| Constraints/cases | No. of Unknowns | No. of Constraints | No. of Solution Space |
|---|---|---|---|
| Case I | $2pN$ | $\frac{p}{2}(2N-p+1)$ | $\frac{p}{2}(2N+p-1)$ |
| Case II | $2pN$ | $\frac{p}{2}(4N-3p+1)$ | $\frac{p}{2}(3p-1)$ |
| Case III | $2pN$ | $\frac{p}{2}(4N-3p+1)$ | $\frac{p}{2}(3p-1)$ |
| Case IV | $2pN$ | $p(2N-1)$ | $p$ |
| Case V | $2pN$ | $2pN$ | 0 |

Reduced-order models of structural systems can be grouped into different categories on the basis of different structural constraints. According to the different forms of the mass, damping, and stiffness matrices, *five* representative scenarios have been defined and will be discussed separately. For all these scenarios, a common set of constraints can be defined expressing condition (6) in terms of $\psi_1$ and $\psi_2$ as

$$\psi\psi^T = \mathbf{0}_{N\times N} = \begin{bmatrix} \psi_1\psi_1^T & \psi_1\psi_2^T \\ \psi_2\psi_1^T & \psi_2\psi_2^T \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{0}_{n\times p} \\ \mathbf{0}_{p\times n} & \mathbf{0}_{p\times p} \end{bmatrix} \tag{15}$$

While the above constraint is automatically satisfied for the known part $\psi_1$, Eq. (15) provides two additional sets of conditions that can be used for the determination of the unknown part $\psi_2$. Such conditions are represented by the following sets of linear and nonlinear equations:

$$\psi_1\psi_2^T = \mathbf{0}_{n\times p}, \quad \psi_2\psi_2^T = \mathbf{0}_{p\times p} \tag{16}$$

**5.1 General M, $\mathbf{C}_c$, and K: Case I.** If the system can only be modeled with general (nondiagonal or block-diagonal) mass, damping, and stiffness matrices, then very little can be done in expanding the initial reduced-order model. In this case, the only constraint conditions that can be imposed on $\psi_2$ come from the above normalization equations, Eq. (15). Hence Eq. (16) will provide $n\times p$ and $p\times p$ equality constraints, respectively. However, due to symmetry, only the upper portion of $\psi_2\psi_2^T$ provides independent conditions and this reduces the total number of constraints to $(p/2)(2N-p+1)$. In this case, the dimension of the solution space ($SS_{DOF}$) becomes $(p/2)(N+p-1)$, as reported in Table 1. For this type of structural models, the missing components of the eigenvector matrix $\psi_2$ and, consequently, the real mass normalized eigenvectors and the associated natural frequencies cannot be exactly identified, with serious consequences on the accuracy of the identified model. The only subpartition matrices that can be determined are $\mathbf{M}_r$, $\mathbf{C}_{cr}$, and $\mathbf{K}_r$ (Eqs. (10)–(12)), which, however, are of limited use for identification and damage detection purposes. In addition, although they reproduce the recorded structural output at the sensor locations quite closely, these models cannot be used to predict the response of the system to future excitation, even at the instrumented locations. The reason is that the optimization identifies a model by minimizing the error between the measured output and the predicted output for a given input. However, since the identified model does not represent the dynamic characteristics of the real system, it cannot be expected to provide accurate prediction of the structural behavior to a different input excitation.

**5.2 Block Diagonal M With General $\mathbf{C}_c$ and K: Case II.** If we assume that the structural system can be represented by a block-diagonal mass matrix (an assumption widely used in mechanical and civil applications), then additional $n\times p$ constraints of the type $\psi_1\mathbf{\Lambda}\psi_2 = 0_{n\times p}$ can be imposed. Thus, the total number

of constraints becomes $(p/2)(4N-3p+1)$ and the dimension of the solution space reduces to $(p/2)(3p-1)$ (see Table 1). It is noteworthy that now the dimension of the solution space depends only on the number of degrees-of-freedom that are not instrumented ($p$) and not on the total number of the system's degrees-of-freedom ($N$).

In this case, the only subpartitions of the mass and damping matrices related to the instrumented degrees-of-freedom, $\mathbf{M}_{11}$ and $\mathbf{C}_{c11}$, are exactly identified as

$$\mathbf{M}_r = (\psi_1 \mathbf{\Lambda} \psi_1^T)^{-1} = \mathbf{M}_{11} \tag{17}$$

$$\mathbf{C}_{cr} = -\mathbf{M}_{11}\psi_1 \mathbf{\Lambda}^2 \psi_1^T \mathbf{M}_{11} = \mathbf{C}_{c11} \tag{18}$$

The remaining subpartitions of the full mass and damping matrices as well as the complete stiffness matrix cannot be determined uniquely. Similarly to Case I, the retrieved full-order model cannot be used for either response prediction or for damage assessment.

**5.3 Block-Diagonal M With the Assumption of Classical Damping $\mathbf{C}_c$: Case III.** Introducing the assumption that the structural damping can be expressed as a function of the mass and stiffness matrices leads to a further improvement in the identification process. In fact, using proper normalization conditions for the eigenvectors [18] (e.g., magnitudes of the real and the imaginary parts of the complex eigenvectors are equal ($|\Re(\psi)| = |\Im(\psi)|$)) will allow us to introduce additional constraints and, consequently, a substantial reduction in the number of unknown parameters. After some simple derivations, it can be shown that not only $\mathbf{M}_{11}$ and $\mathbf{C}_{c11}$ are exactly determined (as in Case II) but also that the subpartition of the stiffness matrix $\mathbf{K}_{11}$ becomes independent of the unknown $\psi_2$ and so can be exactly identified.

The fact that the assumption of classical damping allows the exact determination of subpartitions of the structural matrices related to the instrumented degrees-of-freedom is important for damage identification purposes. In fact, if damage occurs at the instrumented locations, the identified model will allow us to exactly detect and quantify the structural damage. However, in terms of structural response prediction, these identified models are still not capable of accurately predicting the response to future excitation.

**5.4 Diagonal M With Tridiagonal Damping $\mathbf{C}_c$ and Stiffness K: Case IV.** If the system can be modeled with a diagonal mass matrix and tridiagonal damping and stiffness matrices, the identified full-order system shows some interesting characteristics. This model represents systems that are composed of concentrated masses that are only connected to the adjacent ones with a spring and a dashpot, with both ends connected to supports.

In this case, the assumption of diagonal mass matrix imposes that the off-diagonal elements are equal to zero ($\psi_2 \mathbf{\Lambda} \psi_2^T = \text{diag}\{\}$). After accounting for the symmetry conditions, the total number of constraints changes to $p \times (2N-p)$ (Table 1) and thus, the solution space of $\psi_2$ reduces to $p^2$, again function only of the number of uninstrumented degrees-of-freedom.

The assumption of tridiagonal $\mathbf{K}$ and $\mathbf{C}_c$ matrices provides additional $(N-1)\times(N-2)$ independent equality constraints for the solution space of $\psi_2$. If the system has all but one unmeasured coordinates, i.e., $p=p_{\max}=N-1$, then all $(N-1)\times(N-2)$ "zero" elements of $\mathbf{K}$ and $\mathbf{C}_c$ represent independent conditions since the zero elements of the stiffness and damping matrices are independent of each other. Then,

$$\text{SS}_{\text{DOF}} = p_{\max}^2 - (N-1)\times(N-2) = (N-1) = p_{\max} \tag{19}$$

In this case, the dimension of the solution space of $\psi_2$ is reduced to $N-1$. The equivalence of these two numbers, $p_{\max}$ and $\text{SS}_{\text{DOF}}$, is not simply a coincidence but indicates a significant new result. Since the eigenvector matrix $\psi_2$ of dimension $(N-1)\times 2N$ needs to be full rank in rows (each row has at least one unknown factor) and the solution space is $N-1$ ($N-1$ unknown factors in total), it can be concluded that the unknown part of the identified complex eigenvector matrix contains, for each row, only one undetermined factor $\alpha_i$, with $i=1,2,\ldots,N-1$. This implies that the exact (unknown) and the identified (determined through optimization) components of the eigenvector matrices are related by the following relationships:

$$\hat{\psi}_1 = \psi_1(1,:), \quad \hat{\psi}_2 = \begin{bmatrix} \alpha_1\psi_2(1,:) \\ \alpha_2\psi_2(2,:) \\ \cdots \\ \alpha_{N-1}\psi_2(N-1,:) \end{bmatrix} \tag{20}$$

where $\psi_1$, $\psi_2$ and $\hat{\psi}_1$, $\hat{\psi}_2$ are the exact and the identified ($\hat{}$) parts of the eigenvector matrix while $\alpha_1, \alpha_2, \ldots, \alpha_{(N-1)}$ are the unknown proportionality constants. While all the components of the complex eigenvectors at the instrumented locations are correctly identified, the identified components of the eigenvectors at the noninstrumented locations are proportional to the exact values through some unknown constants. At each degree-of-freedom, the value of such a constant is the same for all the modes.

For the other cases where $p<p_{\max}$, not all the conditions imposing zero elements of $\mathbf{K}$ and $\mathbf{C}_c$ are independent. There are only $p^2-p$ independent new constraints because, among the total $p^2$ conditions, the first $p^2-p$ ones imply that there is an unknown factor $\alpha_i$ $(i=1,2,\ldots,p)$ at each unmeasured coordinate while the remaining $p$ conditions are $0\times\alpha_i=0$. Therefore,

$$\text{SS}_{\text{DOF}} = p^2 - [(N-1)\times(N-2) - p] = p \tag{21}$$

which implies that the identified $\hat{\psi}_1$ and $\hat{\psi}_2$ are similar to the corresponding ones in Eq. (20).

$$\hat{\psi}_1 = \begin{bmatrix} \psi_1(1,:) \\ \psi_1(2,:) \\ \cdots \\ \psi_1(N-p,:) \end{bmatrix}, \quad \hat{\psi}_2 = \begin{bmatrix} \alpha_1\psi_2(1,:) \\ \alpha_2\psi_2(2,:) \\ \cdots \\ \alpha_p\psi_2(p,:) \end{bmatrix} \tag{22}$$

Using these identified eigenvectors, $\hat{\psi}_1$ and $\hat{\psi}_2$, in the construction of the mass, damping, and stiffness matrices of the full-order models will lead to some interesting observations. The identified full-order mass, damping, and stiffness matrices show that the subpartitions corresponding to the instrumented degrees-of-freedom are exactly identified while the subpartitions corresponding to the unmeasured degrees-of-freedom are proportional to the exact ones through the $\alpha$ coefficients. For example, the identified full-order stiffness matrix, $\hat{\mathbf{K}}$, for an $N$ degrees-of-freedom system with only $n=N-p$ instrumented ones can be expressed as

$$\hat{\mathbf{K}} = -(\hat{\psi}\mathbf{\Lambda}^{-1}\hat{\psi}^T)^{-1} = \begin{bmatrix} \hat{\mathbf{K}}_{11} & \hat{\mathbf{K}}_{12} \\ \hat{\mathbf{K}}_{21} & \hat{\mathbf{K}}_{22} \end{bmatrix}$$

$$= \left[\begin{array}{c|ccc} \mathbf{K}_{11} & \alpha_1^{-1}k_{n+1,2} & \cdots & 0 \\ \hline \alpha_1^{-1}k_{2,n+1} & (\alpha_1\alpha_1)^{-1}k_{n+1,n+1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & (\alpha_p\alpha_p)^{-1}k_{NN} \end{array}\right] \tag{23}$$

where $\hat{\mathbf{K}}_{11}=\mathbf{K}_{11}$ is the $n\times n$ partition of the stiffness matrix related to the instrumented degrees-of-freedom (exactly identified) while $\hat{\mathbf{K}}_{12}=\hat{\mathbf{K}}_{21}^T$ and $\hat{\mathbf{K}}_{22}$ are the identified $n\times p$ and $p\times p$ subpartitions of the stiffness matrix related to the unmeasured DOFs. These two subpartitions differ from the correct ones because of the proportionality constants $\alpha$'s. Although the expanded mass and stiffness matrices are different from the exact ones, it can be proven that the system's natural frequencies obtained from the

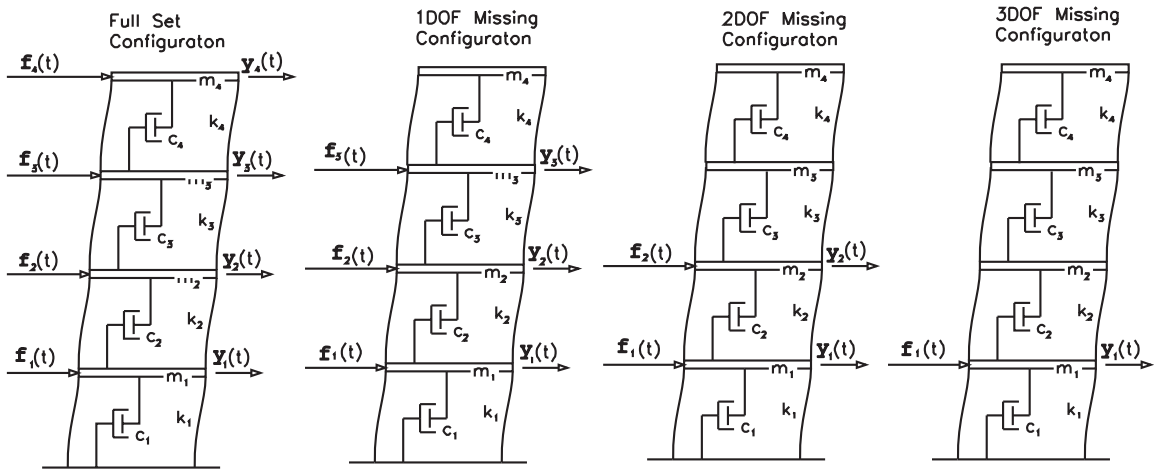



**Fig. 1 4DOF shear-type building system: reduced-order scenarios**

identified matrices are identical to the exact ones. For the proof, see Ref. [17].

The effects of the $\alpha$ coefficients can also be shown in the estimation of the system response. At measured locations, where all the contributions of the various modes are exactly estimated, the predicted response will be equal to the exact one while those predicted at locations with neither a sensor nor an actuator are proportional, through the same coefficients, to the exact ones:

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{\mathbf{y}}_1 \\ \hat{\mathbf{y}}_2 \end{bmatrix}, \quad \hat{\mathbf{y}}_1 = \mathbf{y}_1, \quad \hat{\mathbf{y}}_2 = \begin{bmatrix} \alpha_1 \mathbf{y}_2(1,:) \\ \alpha_2 \mathbf{y}_2(2,:) \\ \cdots \\ \alpha_p \mathbf{y}_2(p,:) \end{bmatrix} \tag{24}$$

Unfortunately, in real cases, not having the exact response available impairs the estimation of the $\alpha$ parameters. However, although these parameters remain unknown, it is still possible to use the identified structural stiffness and mass matrices to obtain correct assessment of the structural damage (see Sec. 9).

**5.5 Shear-Type Structural Models: Case V.** In many structural engineering applications, a common assumption is to consider shear-type models of structures represented by concentrated masses connected with the adjacent ones, with only one mass at the end connected to a support. This type of models, defined shear-type models, represents a special case of the previous tridiagonal systems (Fig. 1). In this case, additional constraints can be imposed by considering the following geometric relations for $\mathbf{K}$ and $\mathbf{C}_c$:

$$k_{(i-1)i} + k_{i(i+1)} = -k_{ii}, \quad c_{c(i-1)i} + c_{ci(i+1)} = -c_{cii}, \quad i = 2,3,\ldots,N-1$$

$$k_{(i-1)i} = -k_{ii}, \quad c_{c(i-1)i} = -c_{cii}, \quad i = N \tag{25}$$

From the previous discussion about Case IV, it is known that the solution space is less than or equal to $N-1$ and the number of elements of the stiffness matrix $k_1, k_2, \ldots, k_N$ is $N$. Considering the symmetry properties, the numbers of nonzero geometry/topology conditions, which can be used in Eqs. (13) and (14), are $2N-1$ ($N$ diagonal elements and $N-1$ off-diagonal elements) reducing the problem to $2N-1$ independent equations in less than $2N-1$ unknown factors. The first $p$ geometry relations in the corresponding undetermined parts of $\mathbf{K}$ ($\mathbf{K}_{12}$ and $\mathbf{K}_{22}$) are used to retrieve $p$ unknown factors $\alpha_1, \alpha_2, \ldots, \alpha_p$, and consequently, all the values of the elements of $\mathbf{K}$, $(k_1, k_2, \ldots, k_N)$, and $\mathbf{C}_c$, $(c_{c1}, c_{c2}, \ldots, c_{cN})$, can be obtained by the remaining topology conditions. Thus, the entire system is identified successfully even with the incomplete instrumentation setup.

Table 1 summarizes the number of unknowns, the available conditions, the number of constraints, and the dimension of the solution space for the five representative cases.

## 6 Uniqueness of the Solution for Reduced Order Model Systems

The crucial point in the identification of systems with a limited set of instrumentation is that only the part of complex eigenvectors, named $\psi_1$, can be exactly identified while the other component, $\psi_2$, can be estimated but without any guarantee to be exact. Only for structures that can be modeled as shear-type structures, the full eigenvector matrix can be correctly identified, provided that the instrumentation setup is such that guarantees the uniqueness of the solution. Consequently, one related question arises—Is there more than one subpartition $\psi_2$ that minimizes the cost functional (Eq. (13)) and satisfies all the appropriate constraints (Eq. (14))?

The answer to these questions can be found by looking into the problem of "uniqueness of the solution," which states that if an admissible solution is reached, then this is the exact one. This condition is quite simple to fulfill in the case of a complete set of instrumentation, leading directly to the identification of full-order models but rather difficult to satisfy with a limited set of instrumentation.

If the structural system has a complete set of instrumentation (at each degree of freedom, there is a sensor and/or an actuator), then there is only one system that can produce the given input-output time histories. In this case, the optimization problem is well posed and it will have a unique solution. However, if the instrumentation setup is not complete (sensors and/or actuators only at few degrees-of-freedom), then there could be more than one structural system that can produce, for the given input, the given output at the prescribed degrees-of-freedom, raising the issue of which system to identify. In the work by Franco et al. [16], the authors study the identification of a simple 3DOF shear-type structure subjected to a given input force in the central degree-of-freedom. As the output, the displacement response of the second degree-of-freedom is considered so that input and output time histories are measured at the same mass. The authors prove the lack of uniqueness of the solution by showing analytically that there are two different structural systems that have the same input-output response at the central mass. In this case, the optimization process will have no meaning since the uniqueness problem is ill posed. This multiplicity of the solution for the same 3DOF shear-type

**Table 2 Mass, damping, and stiffness matrices used for the system of Fig. 1**

| Exact system information of the test modal | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Mass $M$ | | | | Damping $C_c$ | | | | Stiffness $K$ | | | |
| 0.8 | 0.0 | 0.0 | 0.0 | 0.48 | −0.1 | 0.0 | 0.0 | 4.0 | −1.0 | 0.0 | 0.0 |
| 0.0 | 2.0 | 0.0 | 0.0 | −0.1 | 0.6 | −0.3 | 0.0 | −1.0 | 4.0 | −3.0 | 0.0 |
| 0.0 | 0.0 | 1.2 | 0.0 | 0.0 | −0.3 | 0.52 | −0.1 | 0.0 | −3.0 | 4.0 | −1.0 |
| 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | −0.1 | 0.2 | 0.0 | 0.0 | −1.0 | 1.0 |

system disappears when the input-output measurements are taken on the top or the bottom mass: In both these cases, there will be only one system that has a given input-output set of measurements. Only particular instrumentation setups (selection of sensor and actuator locations) and for particular structural models (e.g., shear-type model) can guarantee the uniqueness of the solution and hence validate the optimization process. For example, it can be shown [16,17] that the identification problem for shear-type structural model has a unique solution when measurement information is collected at (1) the bottom degree-of-freedom, (2) the top degree-of-freedom, or (3) any two contiguous degrees-of-freedom. For these three instrumentation setups, there is only one structural model that can map the input-output relationship and so guarantee the uniqueness of the solution. For these setups, the optimization process will lead, if well posed, to the unique structural system. However, for an arbitrary instrumentation setup, the uniqueness of the solution cannot be guaranteed ever for a shear-type structural model.

Among the five representative categories previously described, very little can be said about the uniqueness of the solution for the more general cases (Cases I, II, and III). For these cases, the subpartition matrices related to the instrumented locations are obtained more or less accurately, independent of whether the system's solution is unique. For example, for Case II, $\mathbf{M}_{11}$, $\mathbf{C}_{c11}$, and even $\mathbf{K}_{11}$ are the same for multiple solutions because of their independency of $\psi_2$ [17].

On the other hand, Cases IV and V provide interesting, although not complete, information about the uniqueness of the identified solution. In the work by Franco et al. [16] and by Yu [17], detailed discussions on the "uniqueness problem" have been presented. Starting from the pioneering work by Udwadia and Sharma [14] and Udwadia et al. [15] but using different approaches (Sylvester's dyalitic elimination [16] and Laplace transform [17]), they reached the conclusion that whether or not a system can be uniquely identified depends not only on the number of sensors and actuators available but also on their locations. For example, for structural models in Case V, the dimension of the solution space could be reduced to zero (Table 1). It can be shown that, given the known $\psi_1$ and starting from an arbitrary initial guess of $\psi_2$, the solution search will converge to one and only one point. However, for different initial $\psi_2$, it is still possible that the search will end at several different points, representing different solutions. In other words, the zero dimension solution space is not a sufficient condition for the unique identification, although it guarantees the determination of one of the possible solutions depending on the initial guess of $\psi_2$. As previously discussed, the identification problem for a shear-type structure using only few response measurements has a unique solution depending on whether the sensor and actuator placement satisfies certain geometric conditions (e.g., top mass, bottom mass, and two adjacent masses). If the sensors and actuators are placed at arbitrary locations, the uniqueness of the solution cannot be guaranteed impairing the entire identification problem.

## 7 Numerical Examples

To show the validity of the previous statements, a simple but general numerical example is presented. The system, shown in Fig. 1, is a shear/type four-story building system; the values of the mass, damping, and stiffness matrices used in this study are given in Table 2. Although it is a shear-type model (Case V) for which it is possible to obtain the exact solution, we will also consider the intermediate steps so to highlight the differences with an analogous Case IV model. For the purpose of showing the special property of the complex eigenvectors when the assumption of classical damping is introduced ($|\Re(\psi)|=|\Im(\psi)|$), a proportional damping matrix, $\mathbf{C}_c=0.1\mathbf{K}+0.1\mathbf{M}$ is considered. It is noteworthy that this approach is also valid for the case of general damping. To assure the unique solution of the identification process, the system is excited by only one actuator, located at the bottom floor, while the accelerations are measured at different floors for different scenarios (Table 3): (1) sensors at the first, second, and third floors, (2) sensors at the first and second floors, and (3) sensor only at the first floor (Fig. 1).

**Table 3 The locations of sensors and actuators for different reduced-order system scenarios**

| The locations of inputs and outputs | | |
|---|---|---|
| 1DOF unmeasured | 2DOF unmeasured | 3DOF unmeasured |
| 3 sensors: 1st, 2nd, 3rd floors<br>1 actuator: 1st floor | 2 sensors: 1st, 2nd floors<br>1 actuator: 1st floor | 1 sensor: 1st floor<br>1 actuator: 1st floor |

$$\psi^T = 10^{-1}\times \times \begin{bmatrix} 0.8824+0.8824j & 3.4188+3.4188j & 3.9066+3.9066j & 4.6339+4.6339j \\ 0.8824-0.8824j & 3.4188-3.4188j & 3.9066-3.9066j & 4.6339-4.6339j \\ -0.6395-0.6395j & -1.9375-1.9375j & -0.8032-0.8032j & 3.7689+3.7689j \\ -0.6395+0.6395j & -1.9375+1.9375j & -0.8032+0.8032j & 3.7689-3.7689j \\ -2.2889-2.2889j & -0.8508-0.8508j & 2.2010+2.2010j & -0.6226-0.6226j \\ -2.2889+2.2889j & -0.8508+0.8508j & 2.2010-2.2010j & -0.6226+0.6226j \\ -2.9021-2.9021j & 0.9934+0.9934j & -1.3027-1.3027j & 0.2942+0.2942j \\ -2.9021+2.9021j & 0.9934-0.9934j & -1.3027+1.3027j & 0.2942-0.2942j \end{bmatrix}$$

**Fig. 2 4DOF shear-type system: exact eigenvectors**

$$\hat{\psi}^T_{(1)} = \begin{bmatrix} \hat{\psi}^T_{1(1)} & \hat{\psi}^T_{2(1)} \end{bmatrix} = 10^{-1}\times$$
$$\times \left[\begin{array}{cccc} 0.8824+0.8824j & 3.4188+3.4188j & 3.9066+3.9066j & 8.0047+8.0047j \\ 0.8824-0.8824j & 3.4188-3.4188j & 3.9066-3.9066j & 8.0047-8.0047j \\ -0.6395-0.6395j & -1.9375-1.9375j & -0.8032-0.8032j & 6.5105+6.5105j \\ -0.6395+0.6395j & -1.9375+1.9375j & -0.8032+0.8032j & 6.5105-6.5105j \\ -2.2889-2.2889j & -0.8508-0.8508j & 2.2010+2.2010j & -1.0754-1.0754j \\ -2.2889+2.2889j & -0.8508+0.8508j & 2.2010-2.2010j & -1.0754+1.0754j \\ -2.9021-2.9021j & 0.9934+0.9934j & -1.3027-1.3027j & 0.5082+0.5082j \\ -2.9021+2.9021j & 0.9934-0.9934j & -1.3027+1.3027j & 0.5082-0.5082j \end{array}\right]$$
$$= \begin{bmatrix} \psi^T_{1(1)} & \alpha_{1(1)}\psi^T_{2(1)} \end{bmatrix}$$

**Fig. 3 4DOF shear-type system—identified eigenvectors: Case I**

In Figs. 2–4, the (transpose) matrices containing the exact complex eigenvectors and the identified complex eigenvectors for cases 1 and 3 are presented. Because of the assumption of classical damping and of the normalization scheme used, the real and the imaginary parts of the eigenvectors are identical. Since the structure is properly constrained and no rigid body motion is allowed, the eigenvectors appear in complex conjugate pairs. By comparing Fig. 2 with Fig. 3, it is clear that, for case 1, the first three rows of the eigenvector matrix (corresponding to the instrumented floors) are exactly identified, while the fourth identified one (the noninstrumented floor) is proportional to the corresponding exact one through the parameter $\alpha_1$. Similarly, for case 3 (Fig. 4), only the components of the eigenvectors corresponding to the first floor are identified correctly (first row), while the other three rows are only proportional to the exact ones. If the model representing the structure was ascribed to Case IV, the parameters $\alpha$ could not be determined and so the identified eigenvectors would be in their final stage (e.g., part of them would remain proportional to the exact ones). Instead, if the structure can be modeled as a Case V model (shear type), then the $\alpha$ parameters can be determined, as shown later, and the identified eigenvectors will perfectly match the exact ones.

Using the identified eigenvectors to reproduce the structural outputs, it is possible to highlight the effects of the $\alpha$ parameters on the structural response. If the structure is assumed to be represented with a model as in Case IV, the response at the instrumented degrees-of-freedom would match exactly the correct one, while at the uninstrumented degrees-of-freedom, it is only proportional to the correct one through the unknown $\alpha$ parameters. This is clearly shown in Fig. 5, where the exact and the identified outputs are plotted for the case of two sensors locations. The ratios between each predicted and exact output, $\hat{y}_i(t)$ and $y_i(t)$, respectively, are calculated and plotted for each floor. It can be observed that these ratios stay constant during the entire time histories: For the degrees-of-freedom with instrumentations, the values of these ratios is equal to 1, indicating that the identified time histories are identical to the exact ones, while for the unmeasured coordinates, these ratios are still constant but different from 1 (third floor $\alpha$=0.947 and fourth floor, $\alpha$=2.8368). It is interesting to see that these values differ from case to case (fourth floor case 1: $\alpha$=1.7274, case 2: $\alpha$=2.8368, while for case 3, $\alpha$ =12.663). If the structure is modeled as a shear-type structure (Case V), these $\alpha$ parameters could be uniquely determined and the predicted outputs match the exact ones.

The structural matrices (mass, damping, and stiffness) obtained from the identified eigenvectors will show identical pattern. The submatrices related to the instrumented degrees-of-freedom are exactly identified while those for the noninstrumented degrees-of-freedom are proportional to the exact ones. This is clearly seen in Fig. 6 where the stiffness matrices for case 1 (1DOF unmeasured), case 2 (2DOFs unmeasured), and case 3 (3DOFs unmeasured) are presented. For case 1, the upper $3\times3$ partition is correctly identified, while the last row and column are only proportional to the exact ones through the coefficient $\alpha_1$. Similarly, for case 3, the upper $1\times1$ submatrix is correct while the remaining submatrices are proportional to the correct ones through coefficients $\alpha_1$, $\alpha_2$, and $\alpha_3$.

These identified mass, damping, and stiffness matrices, although not entirely correct, can still be used to determine some characteristic parameters such as natural frequencies and mode shapes. In fact, the particular shape of the mass and stiffness matrices of the identified system for Cases IV and V is such that the $\alpha$ parameters cancel out in the characteristic equation of the associated eigenvalue problem as so the eigenvalues (natural frequencies) are not affected. On the contrary, the normalized real eigenvectors will present a similar pattern to the complex identified eigenvectors, keeping one part proportional to the exact one through the $\alpha$ parameters. Both these conclusions can also be seen in Fig. 5: The frequency content of the identified time histories is identical to that of the exact one while the amplitude is constantly different.

If the structure can be represented as a shear-type model, the proportional $\alpha$ factors can be obtained through the analysis of particular geometrical conditions (Eq. (25)). Consequently, the elements of the subpartitions of the stiffness and damping matrices related to uninstrumented DOFs can be exactly identified. For example, for the case where only 1DOF is not equipped with either a sensor or an actuator, the only unknown factor $\alpha_{1(1)}$ can be determined as follows:

$$\bar{k}_4 = k_{44} = [\alpha_{1(1)}]^2\hat{k}_{44}, \quad -\bar{k}_4 = k_{43} = \alpha_{1(1)}\hat{k}_{43}$$

$$\alpha_{1(1)} = -\hat{k}_{43}/\hat{k}_{44} = 0.5789/0.3351 = 1.7274 \tag{26}$$

so that the correct mass, damping, and stiffness matrices are identified. Consequently, the properties of the system can be retrieved as

$$\Rightarrow m_1 = 0.8, \quad m_2 = 2.0, \quad m_3 = 1.2, \quad m_4 = 1.0$$

$$k_1 = 3.0, \quad k_2 = 1.0, \quad k_3 = 3.0, \quad k_4 = 1.0 \tag{27}$$

The same reasoning applies to the cases when 2DOFS or 3DOFs are not instrumented. In those cases, the $\alpha$ factors are

$$\hat{\psi}^T_{(3)} = \begin{bmatrix} \hat{\psi}^T_{1(3)} & \hat{\psi}^T_{2(3)} \end{bmatrix} = 10^{-1}\times$$
$$\times \left[\begin{array}{cccc} 0.8824+0.8824j & 14.874+14.874j & 25.253+25.253j & 58.679+58.679j \\ 0.8824-0.8824j & 14.874-14.874j & 25.253-25.253j & 58.679-58.679j \\ -0.6395-0.6395j & -8.4294-8.4294j & -5.1925-5.1925j & 47.725+47.725j \\ -0.6395+0.6395j & -8.4294+8.4294j & -5.1925+5.1925j & 47.725-47.725j \\ -2.2889-2.2889j & -3.7015-3.7015j & 14.227+14.227j & -7.8832-7.8832j \\ -2.2889+2.2889j & -3.7015+3.7015j & 14.227-14.227j & -7.8832+7.8832j \\ -2.9021-2.9021j & 4.3218+4.3218j & -8.4210-8.4210j & 3.7255+3.7255j \\ -2.9021+2.9021j & 4.3218-4.3218j & -8.4210+8.4210j & 3.7255-3.7255j \end{array}\right]$$
$$= \begin{bmatrix} \psi^T_{1(3)} & \alpha_{1(3)}\psi^T_{21(3)} & \alpha_{2(3)}\psi^T_{22(3)} & \alpha_{3(3)}\psi^T_{23(3)} \end{bmatrix}$$

**Fig. 4 4DOF shear-type system—identified eigenvectors: Case III**

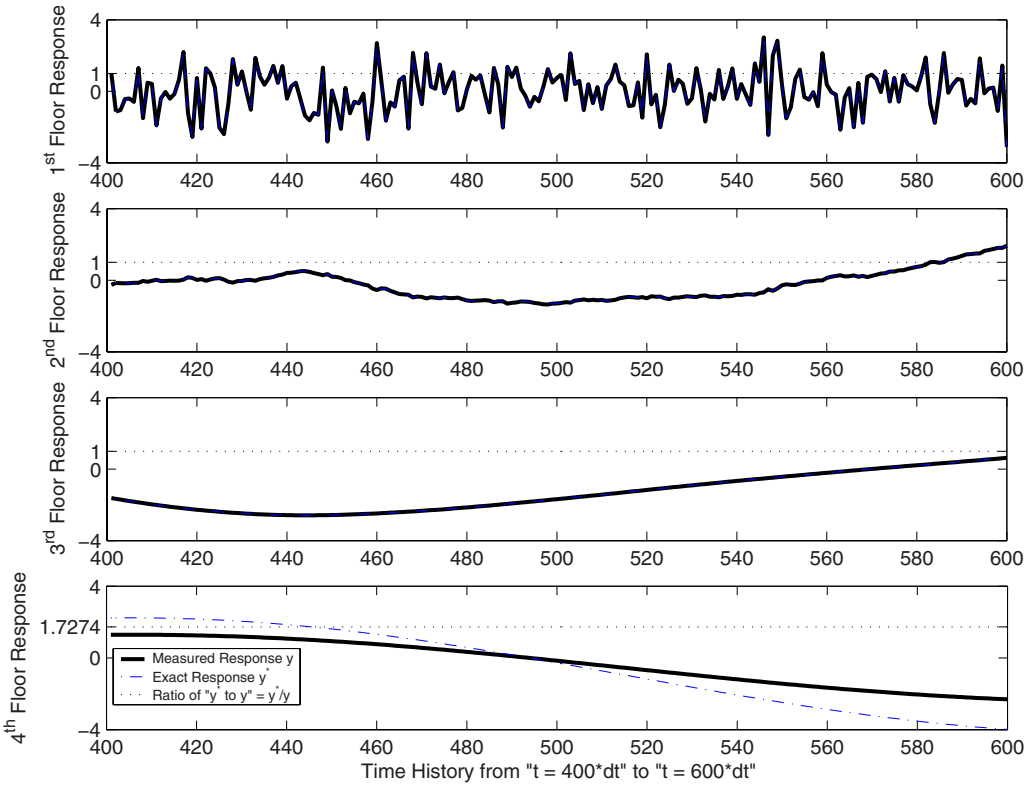



**Fig. 5 Comparison of the measured response y and the identified responses $\hat{y}=y^{*}$ at different floors of the 4DOF shear-type building for 1DOF unmeasured case**

$$\alpha_{1(2)} = 0.9470, \quad \alpha_{2(2)} = 2.8368$$

$$\alpha_{1(3)} = 4.3506, \quad \alpha_{2(3)} = 6.4642, \quad \alpha_{3(3)} = 12.663 \tag{28}$$

which lead to the determination of the correct mass, damping, and stiffness matrices.

$$\hat{\mathbf{K}}_{(1)} = \begin{bmatrix} \hat{\mathbf{K}}_{11(1)} & \hat{\mathbf{K}}_{12(1)} \\ \hat{\mathbf{K}}_{21(1)} & \hat{\mathbf{K}}_{22(1)} \end{bmatrix} = \left[\begin{array}{cccc} 4.0000 & -1.0000 & 0.0000 & 0.0000 \\ -1.0000 & 4.0000 & -3.0000 & 0.0000 \\ 0.0000 & -3.0000 & 4.0000 & -0.5789 \\ \hline 0.0000 & 0.0000 & -0.5789 & 0.3351 \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} k_{11} & k_{12} & 0 & 0 \\ k_{21} & k_{22} & k_{23} & 0 \\ 0 & k_{32} & k_{33} & \alpha_{1(1)}^{-1}k_{34} \\ \hline 0 & 0 & \alpha_{1(1)}^{-1}k_{43} & \alpha_{1(1)}^{-2}k_{44} \end{array}\right];$$

$$\hat{\mathbf{K}}_{(2)} = \begin{bmatrix} \hat{\mathbf{K}}_{11(2)} & \hat{\mathbf{K}}_{12(2)} \\ \hat{\mathbf{K}}_{21(2)} & \hat{\mathbf{K}}_{22(2)} \end{bmatrix} = \left[\begin{array}{cccc} 4.0000 & -1.0000 & 0.0000 & 0.0000 \\ -1.0000 & 4.0000 & -3.1680 & 0.0000 \\ \hline 0.0000 & -3.1680 & 4.4606 & -0.3723 \\ 0.0000 & 0.0000 & -0.3723 & 0.1243 \end{array}\right]$$

$$= \left[\begin{array}{cc|cc} k_{11} & k_{12} & 0 & 0 \\ k_{21} & k_{22} & \alpha_{1(2)}^{-1}k_{23} & 0 \\ \hline 0 & \alpha_{1(2)}^{-1}k_{32} & \alpha_{1(2)}^{-2}k_{33} & \alpha_{1(2)}^{-1}\alpha_{2(2)}^{-1}k_{34} \\ 0 & 0 & \alpha_{1(2)}^{-1}\alpha_{2(2)}^{-1}k_{43} & \alpha_{2(2)}^{-2}k_{44} \end{array}\right];$$

$$\hat{\mathbf{K}}_{(3)} = \begin{bmatrix} \hat{\mathbf{K}}_{11(3)} & \hat{\mathbf{K}}_{12(3)} \\ \hat{\mathbf{K}}_{21(3)} & \hat{\mathbf{K}}_{22(3)} \end{bmatrix} = \left[\begin{array}{c|ccc} 0.4800 & -0.2300 & 0.0000 & 0.0000 \\ \hline -0.2300 & 0.2113 & -0.1067 & 0.0000 \\ 0.0000 & -0.1067 & 0.0957 & -0.0122 \\ 0.0000 & 0.0000 & -0.0122 & 0.0062 \end{array}\right]$$

$$= \left[\begin{array}{c|ccc} k_{11} & \alpha_{1(3)}^{-1}k_{12} & 0 & 0 \\ \hline \alpha_{1(3)}^{-1}k_{21} & \alpha_{1(3)}^{-2}k_{22} & \alpha_{1(3)}^{-1}\alpha_{2(3)}^{-1}k_{23} & 0 \\ 0 & \alpha_{1(3)}^{-1}\alpha_{2(3)}^{-1}k_{32} & \alpha_{2(3)}^{-2}k_{33} & \alpha_{2(3)}^{-1}\alpha_{3(3)}^{-1}k_{34} \\ 0 & 0 & \alpha_{2(3)}^{-1}\alpha_{3(3)}^{-1}k_{43} & \alpha_{3(3)}^{-2}k_{44} \end{array}\right]$$

**Fig. 6 4DOF shear-type system: identified stiffness matrices**

## 8 Noise Effects on Identification

In the process of identifying reduced-order models and their expansion to full-order models, there are several factors, such as errors in the identification of $\psi_1$, $\psi_2$, and $\mathbf{\Lambda}$, that can affect the final outcome. As discussed in Ref. [17], it is reasonable to assume that all the eigenvalues $\mathbf{\Lambda}$ and the known components of the eigenvectors ($\psi_1$) can be identified with sufficient accuracy even in the presence of noise. Hence, the identification of reliable mass, damping, and stiffness matrices for the full-order system will depend on how the identification of the unknown part of complex eigenvector matrix, $\psi_2$, will be affected by noise.

To highlight the effects of measurement noise in the identification of the structural matrices, a simple 4DOF system with diagonal $\mathbf{M}$ and tridiagonal $\mathbf{C}_c$ and $\mathbf{K}$ matrices is considered. The system is subjected to external excitation at the first floor and the structural response is measured at the first and second floors. The input signal and structural responses are polluted with Gaussian, zero mean, white noise sequences (2%, 5%, and 10% rms). At each noise level, three noise patterns (output only, input only, and output/input noise polluted signals) are considered.

To highlight the effects of measurement noise on the identification, the coefficient of variation (CoV) of the identified components of $\hat{\psi}_2$ are computed and presented in Table 4 for all noise cases. Starting with the exact values of $\psi_1$ and $\mathbf{\Lambda}$ and using the noise polluted data in inputs or outputs or the combination of the two (first row in each noise level partition), the error is small enough to assume that $\psi_2$ is identified as accurately as in the noise free cases (CoV: $10^{-9}$–$10^{-11}$), independent of the noise level. Considering that only the values of $\psi_1$ are correct while the values in $\mathbf{\Lambda}$ are obtained by noise polluted data (second row in each noise level partition), satisfactory values of $\psi_2$ are still obtained for all noise levels (CoV: of the order $10^{-4}$). On the contrary, if the

**Table 4 Results for different noise polluted scenarios of reduced-order systems**

| Noise | STD/mean exact values | Output only | Noise patterns Input only | Input and output |
|---|---|---|---|---|
| | $\psi_1$ and $\Lambda$ | $8.8211\times10^{-11}$ | $6.3895\times10^{-11}$ | $1.2032\times10^{-11}$ |
| 2% | $\psi_1$ | $1.3010\times10^{-4}$ | $2.4420\times10^{-4}$ | $3.7365\times10^{-4}$ |
| | None | $5.5921\times10^{0}$ | $3.9407\times10^{0}$ | $1.0288\times10^{-1}$ |
| | $\psi_1$ and $\Lambda$ | $7.8525\times10^{-10}$ | $6.2229\times10^{-11}$ | $9.6750\times10^{-10}$ |
| 5% | $\psi_1$ | $3.5942\times10^{-4}$ | $5.0283\times10^{-4}$ | $5.9610\times10^{-4}$ |
| | None | $6.7392\times10^{-2}$ | $2.3771\times10^{0}$ | $5.8793\times10^{-2}$ |
| | $\psi_1$ and $\Lambda$ | $1.0310\times10^{-9}$ | $8.9614\times10^{-10}$ | $1.6556\times10^{-9}$ |
| 10% | $\psi_1$ | $7.4519\times10^{-4}$ | $1.7774\times10^{-3}$ | $2.3450\times10^{-3}$ |
| | None | $8.1390\times10^{-1}$ | $1.8915\times10^{1}$ | $4.4684\times10^{-2}$ |

values of $\psi_1$ are also obtained using noisy measurement data, this reflects in a poor identification of $\psi_2$ (CoV too large ($10^1$ even to claim convergence)).

From these results, it can be concluded that, for the proper identification of $\psi_2$, the eigenvector components at the instrumented locations, $\psi_1$, play a key role. If the components of $\psi_1$ are identified correctly [19], then for noise levels up to 10% rms, the identification of $\psi_2$ is quite accurate and the error in the full-order matrices is small.

## 9 Damage Detection Using the Identified Matrices

The identified full-order mass, damping, and stiffness matrices can be eventually used to locate and quantify structural damage. For example, by comparing the stiffness matrix identified before damage has occurred, $\mathbf{K}$, and after damage, $\mathbf{K}^{(d)}$, it is possible to identify the location where damage has occurred and to quantify what is the reduction of the element/floor stiffness. In order to do so, a reliable unique characterization of the stiffness matrix must be identified. As previously shown, this is possible only in Case V and with appropriate instrumentation setups. However, even for Case IV, where the $\alpha$ parameters remain unknown, it is possible to use the identified structural stiffness and mass matrices to obtain correct assessment of the structural damage. In fact, as shown in Ref. [17], by assuming that the structural mass does not change because of the damage, appropriately defined parameters can be defined as ratios of identified stiffnesses and masses. Because of the nature of the identified matrices, these newly defined damage parameters are such that the $\alpha$ parameters cancel out, providing a correct assessment of the amount and location of the structural damage.

In the work by Yu [17], one parameter that is introduced as a tool for damage detection is the so-called $\mathbf{K}/\mathbf{M}$ ratio. This parameter can be defined as the ratio between the square value of any $i$th-$j$th element of the stiffness matrix and the product of the corresponding $i$th and $j$th masses. For the measured DOFs ($i,j=1,2,\ldots,n$), such a ratio results in

$$\frac{\hat{k}_{ij}^2}{\hat{m}_{ii}\hat{m}_{jj}}=\frac{k_{ij}^2}{m_{ii}mjj} \tag{29}$$

where the ($\hat{}$) quantities indicate the identified ones. Equation (29) shows that, for the portion of the matrices related to instrumented locations, the identified mass and stiffness matrices provide the same value as the exact ones.

$$\widehat{\mathbf{M}}_{(1)}=\left[\begin{array}{c|cc} +1.0000 & +0.0000 & +0.0000 \\ \hline +0.0000 & +1.0582 & +0.0000 \\ +0.0000 & +0.0000 & +0.8923 \end{array}\right];\quad \widehat{\mathbf{K}}_{(1)}=\left[\begin{array}{c|cc} +3.7000 & -1.0287 & +0.0002 \\ \hline -1.0287 & +4.2327 & -2.9152 \\ +0.0000 & -2.9152 & +2.6770 \end{array}\right]$$

**Fig. 7 3DOF shear-type system—identified mass and stiffness matrices: 10% stiffness reduction at first floor, no noise**

For the DOFs where no instrumentation is present ($i,j=1,2,\ldots,p$), the $\mathbf{K}/\mathbf{M}$ ratio again provides identical values for both the identified matrices and the exact ones:

$$\frac{\hat{k}_{ij}^2}{\hat{m}_{ii}\hat{m}_{jj}}=\frac{k_{ij}^2(\alpha_{i-n}\alpha_{j-n})^2}{m_{ii}(\alpha_{i-n})^2m_{jj}(\alpha_{j-n})^2}=\frac{k_{ij}^2}{m_{ii}m_{jj}} \tag{30}$$

showing that the $\mathbf{K}/\mathbf{M}$ ratio is independent from the unknown factors embedded in the identified matrices $\hat{\mathbf{M}}$ and $\hat{\mathbf{K}}$. This is a valuable property that can be important in damage detection processes that use reduced-order models.

The same $\mathbf{K}/\mathbf{M}$ ratio can be obtained considering the identified mass and stiffness matrices for the damaged system ($\hat{\mathbf{M}}^d$ and $\hat{\mathbf{K}}^d$). Considering that the mass $m_{ii}(i=1,2,\ldots,N)$ does not change before and after damage, i.e., $m_{ii}=m_{ii}^{(d)}$, we can obtain information on damage location and amount in the entire system by a new damage index, the square root of the ratio between the two $\mathbf{K}/\mathbf{M}$ ratios corresponding to the damage and undamaged configurations, as shown:

$$\frac{k_{ij}^{(d)}}{k_{ij}}=\sqrt{\frac{(k_{ij}^{(d)})^2}{(k_{ij})^2}}=\sqrt{\frac{(k_{ij}^{(d)})^2/(m_{ii}^{(d)}m_{jj}^{(d)})}{(k_{ij})^2/(m_{ii}m_{jj})}}=\sqrt{\frac{(\hat{k}_{ij}^{(d)})^2/(\hat{m}_{ii}^{(d)}\hat{m}_{jj}^{(d)})}{(\hat{k}_{ij})^2/(\hat{m}_{ii}\hat{m}_{jj})}} \tag{31}$$

being $i,j=1,2,\ldots,N$. This approach allows us to explore the changes also in the off-diagonal terms of the stiffness matrix, giving the opportunity of refining the damage identification process [17].

As a numerical example, consider a three degree-of-freedom shear-type system, which can be modeled with a diagonal mass matrix $\mathbf{M}$ ($m_i=1$, with $i=1,2,3$), a tridiagonal stiffness matrix $\mathbf{K}$ ($k_1=k_3=3$, $k_2=1$), and a tridiagonal damping matrix $\mathbf{C}_c$ ($c_{c1}=c_{c3}=3$, $c_{c2}=1$). Structural damage occurs on the system in two different scenarios: (1) a stiffness reduction of 10% at the first floor, i.e., $k_1^{(d)}=2.7$, and (2) a stiffness reduction of 10% at the first and second floors $k_1^{(d)}=2.7$ and $k_2^{(d)}=0.9$. The system is identified using only one time history of the input and one of the output at the first floor. The identified mass and stiffness matrices for the two different damage scenarios are presented in Figs. 7 and 9 for the case with no measurement noise. In this case, it is evident that the comparison of the stiffness matrices from before and after damage does not provide any information on the damage amount and location, especially for the case of second damage scenario. Using the $\mathbf{K}/\mathbf{M}$ ratios clearly indicates the reduction of the stiffness on the first and second floors (Figs. 8 and 10). Even in the presence of 5% rms input and output measurement noise (Figs. 11 and 12),

$$\mathbf{d}^{(1)}=\begin{bmatrix} +0.9250 & +1.0000 & - \\ +1.0000 & +1.0000 & +1.0000 \\ - & +1.0000 & +1.0000 \end{bmatrix}$$

**Fig. 8 3DOF shear-type system—K/M ratios: 10% stiffness reduction at first floor, no noise**

$$\widehat{\mathbf{M}}_{(1\&2)}=\left[\begin{array}{c|cc} +1.0000 & +0.0000 & +0.0000 \\ \hline +0.0000 & +0.7668 & +0.0000 \\ +0.0000 & +0.0000 & +0.6713 \end{array}\right];\quad \widehat{\mathbf{K}}_{(1\&2)}=\left[\begin{array}{c|cc} +3.6000 & -0.7881 & +0.0000 \\ \hline -0.7881 & +2.9906 & -2.1524 \\ +0.0000 & -2.1524 & +2.0139 \end{array}\right]$$

**Fig. 9 3DOF shear-type system—identified mass and stiffness matrices: 10% stiffness reduction at first and second floors, no noise**

$$\mathbf{d}^{(1\&2)} = \begin{bmatrix} +0.9000 & +0.9000 & - \\ +0.9000 & +0.9750 & +1.0000 \\ - & +1.0000 & +1.0000 \end{bmatrix}$$

**Fig. 10 3DOF shear-type system—K/M ratios: 10% stiffness reduction at first and second floors, no noise**

$$\widehat{\mathbf{M}}_{(1)} = \begin{bmatrix} +1.0034 & -0.0006 & +0.0002 \\ -0.0006 & +1.0378 & -0.0531 \\ +0.0002 & -0.0531 & +1.0229 \end{bmatrix}; \quad \widehat{\mathbf{K}}_{(1)} = \begin{bmatrix} +3.7418 & -1.0459 & +0.0499 \\ -1.0459 & +4.2988 & -3.2733 \\ +0.0499 & -3.2733 & +3.2325 \end{bmatrix}$$

**Fig. 11 3DOF shear-type system—identified mass and stiffness matrices: 10% stiffness reduction at first, 5% rms noise in input and output measurements**

$$d_{5\%}^{(1)} = \begin{bmatrix} +0.9275 & +1.0009 & - \\ +1.0009 & +1.0004 & +0.9894 \\ - & +0.9894 & +0.9829 \end{bmatrix}$$

**Fig. 12 3DOF shear-type system—K/M ratios: 10% stiffness reduction at first floor, 5% rms noise in input and output measurements**

the $\mathbf{K}/\mathbf{M}$ ratios successfully indicate the 10% reduction in the stiffness of the first floor. For further details, the reader is referred to the work by Yu [17].

## 10 Conclusions

In this study, the identification of linear structural systems with a limited instrumentation setup has been addressed. In these cases, because the set of input/output measurement is not complete, the full mass, damping, and stiffness matrices cannot be generally retrieved. The lack of measurements at some degrees-of-freedom implies that only the components of the eigenvector matrix related to instrumented degrees-of-freedom can be successfully identified while nothing can be said about the components related to the unmeasured degrees-of-freedom.

In this paper, an attempt is made to expand the identified reduced-order models to full-order models by using an optimization algorithm based on a sensitivity analysis with respect to the unknown components of the complex eigenvector matrix $\psi_2$. To this purpose, five increasingly simplified structural models have been considered and their applicability to reduced-order problems has been analyzed.

If the structural model can be represented by a diagonal mass matrix and by tridiagonal damping and stiffness matrices, there is only one unknown multiplying factor for all modes at every unmeasured coordinate. This implies that the identified mass, damping, and stiffness matrices will still be different from the exact ones but will have a particular form that will allow the exact determination of the system's natural frequencies. When the structure can be modeled as a shear-type model, the geometric relations among some specific elements of the damping and stiffness matrices allow us to evaluate those unknown multiplying factors so that the full-order system can be retrieved even with an incomplete set of instrumentation.

The effectiveness of the approach has been illustrated by numerical examples. Noise effects on the identification of full-order models and on their effective use in damage detection have also been analyzed.

## References


[1] Agbabian, M. S., Masri, S. F., Miller, R. K., and Caughey, T. K., 1991, "System Identification Approach to Detection of Structural Changes," J. Eng. Mech., **117**(2), pp. 370–390.

[2] Safak, E., 1989, "Adaptive Modelling, Identification, and Control of Dynamic Structural Systems. I: Theory," J. Eng. Mech. **115**(11), pp. 2386–2405.

[3] Safak, E., 1989, "Adaptive Modelling, Identification, and Control of Dynamic Structural Systems. II: Applications," J. Eng. Mech. **115**(11), pp. 2406–2426.

[4] Udwadia, F. E., 1994, "Methodology for Optimum Sensor Locations for Parameter Identification in Dynamic Systems," J. Eng. Mech., **120**(2), pp. 368–390.

[5] 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.

[6] Beck, J. L., and Katafygiotis, L. S., 1998, "Updating Models and Their Uncertainties. II: Modal Identifiability," J. Eng. Mech., **124**(4), pp. 463–467.

[7] Luş, H., Betti, R., and Longman, R. W., 1999, "Identification of Linear Structural Systems Using Earthquake-Induced Vibration Data," Earthquake Eng. Struct. Dyn., **28**, pp. 1449–1467.

[8] Alvin, K. F., and Park, K. C., 1994, "Second-Order Structural Identification Procedure Via State-Space: Based System Identification," AIAA J., **32**(2), pp. 397–406.

[9] Alvin, K. F., Peterson, L. D., and Park, K. C., 1995, "Method for Determining Minimum-Order Mass and Stiffness Matrices From Modal Test Data," AIAA J., **33**(1), pp. 128–135.

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

[11] Doebling, S. W., Hemez, F. M., Barlow, M. S., Petersonorse, L. D., and Farhat, C., 1993, "Damage Detection in a Suspended Scale Model Truss Via Modal Update," *Proceedings of the 11th International Modal Analysis Conference*, Society for Experimental Mechanics, pp. 1083–1094.

[12] DeAngelis, M., Luş, H., Betti, R., and Longman, R. W., 2002, "Extracting Physical Parameters of Mechanical Models From Identified State Space Representations," ASME J. Appl. Mech., **69**(5), pp. 617–625.

[13] Luş, H., De Angelis, M., and Betti, R., 2003, "A New Approach for Reduced Order Modeling of Mechanical Systems Using Vibration Measurements," ASME J. Appl. Mech., **70**(5), pp. 715–723.

[14] Udwadia, F. E., and Sharma, D. K., 1978, "Some Uniqueness Results Related to Building Structural Identification," SIAM J. Appl. Math., **34**(1), pp. 104–108.

[15] Udwadia, F. E., Sharma, D. K., and Shah, P. C., 1978, "Uniqueness of Damping and Stiffness Distributions in the Identification of Soil and Structural Systems," SIAM J. Appl. Math., **45**(1), pp. 181–187.

[16] Franco, G., Betti, R., and Longman, R. W., 2006, "On the Uniqueness of Solutions for the Identification of Linear Structural Systems," ASME J. Appl. Mech., **73**(1), pp. 153–162.

[17] Yu., J., 2004, "Identification of Structural Systems With Limited Set of Instrumentation," Ph.D. thesis, Columbia University, New York.

[18] Sestieri, A., and Ibrahim, S. R., 1994, "Analysis of Errors and Approximations in the Use of Modal Co-Ordinates," J. Sound Vib., **177**(2), pp. 145–157.

[19] Luş, H., Betti, R., and Longman R. W. 2002, "Obtaining Refined First-Order Predictive Models of Linear Structural Systems," Earthquake Eng. Struct. Dyn. **31**, pp. 1413–1440.

# Suppressing Flutter Vibrations by Parametric Inertia Excitation

**Fadi Dohnal**
Institute of Sound and Vibration Research,
University of Southampton, Highfield,
Southampton SO17 1BJ, UK
e-mail: fd@isvr.soton.ac.uk

**Aleš Tondl**
Zborovská 41,
CZ-150 00 Prague 5, Czech Republic

*A theoretical study of a slender engineering structure with lateral and angular deflections is investigated under the action of flow-induced vibrations. This aero-elastic instability excites and couples the system's bending and torsion modes. Semiactive means due to open-loop parametric excitation are introduced to stabilize this self-excitation mechanism. The parametric excitation mechanism is modeled by time-harmonic variation in the concentrated mass and/or moment of inertia. The conditions for full suppression of the self-excited vibrations are determined analytically and compared with numerical results of an example system. For the first time, example systems are presented for which parametric antiresonance is established at the parametric combination frequency of the sum type.* [DOI: 10.1115/1.3063631]

*Keywords: aero-elasticity, flutter, vibration suppression, parametric excitation*

## 1 Introduction

For some structures exposed to the action of wind flow, this action can lead to a stability loss that initiates vibrations due to aero-elastic interaction, see, e.g., Ref. [1]. In the case where lateral and angular deflections have to be taken into account the simplest mathematical model is governed by a set of two linearly coupled differential equations of second order including aerodynamic forces. These forces can be modeled by introducing aeroelastic derivatives [2] or gyroscopic and nonconservative expression as widely discussed in literature, see, e.g., Refs. [3–5]. In the present study the second approach is chosen.

For the investigation of the stability of the equilibrium position only linear terms are considered. Such a model is expressed by the following equations of motion:

$$\ddot{y} + \omega_y^2 y - p\varphi + \beta\dot{y} - q\dot{\varphi} = 0 \tag{1a}$$

$$\ddot{\varphi} + \omega_\varphi^2 \varphi + py + \kappa\dot{\varphi} + q\dot{y} = 0 \tag{1b}$$

where $y$ represents the lateral deflection, $\varphi$ the angular deflection, $p$ and $q$ the coupling coefficients, and $\beta$ and $\kappa$ the damping coefficients. The damping coefficients can be negative due to the self-excitation mechanism, while the other coefficients are positive. When the equilibrium position becomes unstable, the vibration amplitudes resulting from the linearized equations in Eq. (1) are unlimited. However, the vibrations of the underlying nonlinear system are limited due to positive progressive damping. The validity of this mathematical model was proven experimentally in Refs. [3,4]. A stability analysis of systems as in Eq. (1) possessing gyroscopic and follower forces using different mathematical tools can be found in Refs. [6–9]. This paper examines a possibility to stabilize an otherwise unstable system by introducing time-periodic variations in the inertia coefficients, a parametric excitation.

The equations of motion in Eq. (1) describe the interaction of lateral and torsional vibrations, which may represent a serious danger and have to be either fully suppressed or, at least, limited. Tuned vibration absorbers (TVAs) or elastic mounting of the foundation connected with damping elements belong to the class of passive means, see, e.g., Refs. [10,11]. In recent years a special type of semiactive means using parametric excitation has been analyzed successfully of being capable to suppress self-excited vibrations. This semi-active means can by applied in the case when the passive means are not sufficiently effective. This phenomenon of damping the system's vibrations by introducing a nonresonant parametric excitation was found by Tondl [12] and thoroughly investigated in Refs. [13,14]. A specific parametric excitation that stabilizes an otherwise unstable system is called to be at parametric antiresonance. In these contributions general analytical stability conditions were derived showing that a parametric antiresonance may be expected if the parametric excitation frequency $\omega$ is close to a parametric resonance frequency of difference or sum type

$$\omega_0 = \frac{|\Omega_i \mp \Omega_j|}{N} \quad \text{for } i,j = 1,2,\ldots,n, \quad N = 1,2,\ldots \tag{2}$$

where $\Omega_k$ are the undamped natural frequencies of the system. To achieve this parametric antiresonance certain conditions have to be satisfied. Consequently, not every self-excited system can be stabilized by parametric excitation. First systems for which parametric antiresonances occur were found for time-periodic stiffness variations, e.g., chain systems [12,14–20], rotor systems [21,22], and cantilevers [23]. Parametric antiresonances of systems with periodic damping variation are investigated in Refs. [13,24] and systems with periodic mass or moment of inertia in Refs. [25,26]. Finally, a combination of stiffness, damping, and/or inertia variation is considered in Refs. [13,27].

The phenomenon of damping by parametric excitation has been confirmed at the parametric excitation frequencies of the difference type, $\omega \approx |\Omega_i - \Omega_j|/N$, for example systems having symmetric system matrices. Theoretical calculations in Refs. [13,26] give conditions for which the phenomenon of vibration suppression could occur at the parametric combination frequencies of the sum type, $\omega \approx (\Omega_i + \Omega_j)/N$. However, a physically feasible system configuration that meets these conditions could not be found so far. This contribution deals with the possibility of applying parametric excitation to suppress the selfexcited vibrations of coupled bending and torsion modes of the model in Eq. (1). For the first time, the present study describes example systems for which parametric antiresonance is established at frequencies of sum type in Eq. (2).

## 2 Formulation of Systems and Basic Considerations

Three systems with parametric excitation are considered, which equations of motion are listed below.

1. *System 1*. A parametric excitation in the lateral motion introduced by a periodic variation in the mass in Eq. (1). The equations of motion in its linearized form result in

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received August 6, 2007; final manuscript received October 2, 2008; published online March 9, 2009. Review conducted by N. Sri Namachchivaya.

$$(1+\varepsilon_1\cos\omega t)\ddot{y}+\omega_y^2 y-p\varphi-\varepsilon_1\omega\dot{y}\sin\omega t+\beta\dot{y}-q\dot{\varphi}=0 \tag{3a}$$

$$\ddot{\varphi}+\omega_\varphi^2\varphi+py+\kappa\dot{\varphi}+q\dot{y}=0 \tag{3b}$$

2. *System 2*. The parametric excitation is due to the periodic variation in the moment of inertia corresponding to the angular motion. For this configuration the linearized equations of motion yield

$$\ddot{y}+\omega_y^2 y-p\varphi+\beta\dot{y}-q\dot{\varphi}=0 \tag{4a}$$

$$(1+\varepsilon_2\cos\omega t)\ddot{\varphi}+\omega_\varphi^2\varphi+py-\varepsilon_2\omega\dot{\varphi}\sin\omega t+\kappa\dot{\varphi}+q\dot{y}=0 \tag{4b}$$

3. *System 3*. This system is a combination of the parametric excitations in Systems 1 and 2, respectively, and is governed by the following linearized equations of motion:

$$(1+\varepsilon_1\cos\omega t)\ddot{y}+\omega_y^2 y-p\varphi-\varepsilon_1\omega\dot{y}\sin\omega t+\beta\dot{y}-q\dot{\varphi}=0 \tag{5a}$$

$$(1+\varepsilon_2\cos\omega t)\ddot{\varphi}+\omega_\varphi^2\varphi+py-\varepsilon_2\omega\dot{\varphi}\sin\omega t+\kappa\dot{\varphi}+q\dot{y}=0 \tag{5b}$$

For $\varepsilon_2=0$ these equations simplify to Eq. (3) and for $\varepsilon_1=0$ to Eq. (4).

A first step toward the analysis is a transformation of the differential equations of motion into the quasinormal form assuming that the coefficients of the terms expressing damping ($\kappa$ and $\beta$) and parametric excitation ($\varepsilon_1$ and $\varepsilon_2$) are small. By introducing the transformation

$$y=x_1+x_2,\quad \varphi=a_1x_1+a_2x_2 \tag{6}$$

in Eqs. (3*a*), (3*b*), (4*a*), (4*b*), (5*a*), and (5*b*), the equations of motion with harmonic mass and/or inertia excitation can be transformed into the general quasinormal form

$$\ddot{x}_s+\Omega_s^2x_s+\sum_{k=1}^{n}\Theta_{sk}\dot{x}_k-\omega S_{sk}\dot{x}_k\sin\omega t+Q_{sk}x_k\cos\omega t=0 \tag{7}$$

For the systems considered here $n=2$ holds. The coefficients in Eq. (6) are determined by

$$a_k=-\frac{p}{\omega_\varphi^2-\Omega_k^2},\quad k=1,2 \tag{8}$$

where

$$(\Omega^2)_{1,2}=\tfrac{1}{2}(\omega_y^2+\omega_\varphi^2)\mp\tfrac{1}{2}\sqrt{(\omega_y^2-\omega_\varphi^2)^2-4p^2} \tag{9}$$

To obtain real values for $\Omega^2$ the condition

$$\omega_y^2+\omega_\varphi^2>4p^2 \tag{10}$$

must be met. The following relations can be derived for the system frequencies $\Omega_k$, assuming that $\Omega_1<\Omega_2$:

$$\omega_\varphi^2<\omega_y^2:\quad \Omega_1^2>\omega_\varphi^2\quad\text{and}\quad\Omega_2^2<\omega_y^2 \tag{11a}$$

$$\omega_\varphi^2>\omega_y^2:\quad \Omega_1^2>\omega_y^2\quad\text{and}\quad\Omega_2^2<\omega_\varphi^2 \tag{11b}$$

The diagonal coefficients $\Theta_{kk}$ in Eq. (7) correspond to the modal damping of the $k$th mode. When a coefficient $\Theta_{kk}$ is negative then the equilibrium position is unstable and self-excited vibrations in the $j$th vibration mode occur. According to the well-known theory of linear equations with periodic coefficients, e.g., Ref. [28], parametric instability intervals can be expected if the parametric excitation frequency $\omega$ is close to the parametric resonance frequencies $\omega_0$ in Eq. (2). A method for determining the instability intervals of the parametric excitation frequency $\omega$ is presented in Ref. [29]. In the case of harmonic stiffness excitation, the method proves that only one sign in Eq. (2) corresponds to a destabilizing parametric resonance while the other sign leads to a stabilizing parametric antiresonance. For the case of a system with harmonic mass excitation the same statement is found in Refs. [26,13]. However, for a system with harmonic damping excitation both signs in Eq. (2) lead to a destabilizing parametric resonance if the parametric excitation is symmetric (see Ref. [24]) and to a stabilizing parametric antiresonance if the parametric excitation is skew symmetric (see Ref. [13]). In all analyzed systems so far the parametric antiresonance can partly or even fully suppress self-excited vibrations.

## 3 Analytical Predictions

In Sec. 3.1, the equations of motion of the three example systems are transformed to their quasinormal form. In Sec. 3.2, the necessary and sufficient conditions for damping by parametric excitation are given explicitly for each system.

**3.1 Transformation to Quasinormal Form.** Applying a Taylor series for small values of $\varepsilon$ and respecting terms of up to first order,

$$\frac{1}{1+\varepsilon_k\cos\omega t}=1-\varepsilon_k\cos\omega t+\mathcal{O}(\varepsilon_k^2) \tag{12a}$$

$$\frac{\varepsilon_k\sin\omega t}{1+\varepsilon_k\cos\omega t}=\varepsilon_k\sin(\omega t)+\mathcal{O}(\varepsilon_k^2) \tag{12b}$$

the equations of motion in Eq. (5) can be rewritten as

$$\ddot{y}+\omega_y^2y-p\varphi-\varepsilon_1(\cos\omega t(\omega_y^2y-p\varphi)+\omega\sin\omega t\dot{y})+\beta\dot{y}-q\dot{\varphi}=0 \tag{13a}$$

$$\ddot{\varphi}+\omega_y^2\varphi+py-\varepsilon_2(\cos\omega t(\omega_\varphi^2\varphi+py)+\omega\sin\omega t\dot{\varphi})+\kappa\dot{\varphi}+q\dot{y}=0 \tag{13b}$$

Transforming Eq. (13) into its quasinormal form in Eq. (7) according to Eq. (6) yields

$$\begin{aligned}\ddot{x}_1+\Omega_1^2x_1=&-\frac{a_2}{a_2-a_1}(\beta(\dot{x}_1+\dot{x}_2)-q(a_1\dot{x}_1+a_2\dot{x}_2))\\&+\frac{\varepsilon_1a_2\cos\omega t}{a_2-a_1}(\omega_y^2(x_1+x_2)+p(a_1x_1+a_2x_2))\\&+\frac{\varepsilon_2}{a_2-a_1}(\cos\omega t(\omega_\varphi^2(a_1x_1+a_2x_2)+p(x_1+x_2))\\&-\omega\sin\omega t(\dot{x}_1+\dot{x}_2))+\frac{1}{a_2-a_1}(a_2\varepsilon_1\omega\sin\omega t(\dot{x}_1+\dot{x}_2)\\&+\kappa(a_1\dot{x}_1+a_2\dot{x}_2)+q(\dot{x}_1+\dot{x}_2))\end{aligned} \tag{14a}$$

$$\begin{aligned}\ddot{x}_2+\Omega_2^2x_2=&\frac{a_1}{a_2-a_1}(\beta(\dot{x}_1+\dot{x}_2)-q(a_1\dot{x}_1+a_2\dot{x}_2))\\&-\frac{a_1}{a_2-a_1}\varepsilon_2\cos\omega t(\omega_y^2(x_1+x_2)+p(a_1x_1+a_2x_2))\\&+\frac{\varepsilon_1}{a_2-a_1}(a_1\cos\omega t(\omega_\varphi^2(a_1x_1+a_2x_2)+p(x_1+x_2))\\&-\omega\sin\omega t(\dot{x}_1+\dot{x}_2))-\frac{1}{a_2-a_1}(a_1\varepsilon_2\omega\sin\omega t(\dot{x}_1+\dot{x}_2)\\&+\kappa(a_1\dot{x}_1+a_2\dot{x}_2)+q(\dot{x}_1+\dot{x}_2))\end{aligned} \tag{14b}$$

Thus, the parameters in Eq. (7) can be identified. The diagonal damping coefficients are

$$\Theta_{11}=\frac{1}{a_2-a_1}(a_2\beta-a_1\kappa-q(1+a_1a_2)) \tag{15a}$$

**Table 1 System and quasinormal form parameters**

| System parameter | $\omega_y^2$ | $\omega_\varphi^2$ | $p$ | $q$ | $\beta$ | $\kappa$ |
|---|---|---|---|---|---|---|
| Value | 1.0 | 0.55 | 0.1 | 0.03 | 0.05 | −0.015 |
| Quasinormal form parameter | $\Omega_1$ | $\Omega_2$ | $a_1$ | $a_2$ | $\Theta_{11}$ | $\Theta_{22}$ |
| Value | 0.76 | 0.99 | 4.27 | 0.23 | −0.004 | 0.039 |

$$\Theta_{22} = \frac{-1}{a_2 - a_1}(a_1\beta - a_2\kappa - q(1 + a_1 a_2)) \tag{15b}$$

-diagonal parametric excitation coefficients read

$$Q_{12} = \frac{-1}{a_2 - a_1}(a_2(\varepsilon_1\omega_y^2 - \varepsilon_2\omega_\varphi^2) - p(\varepsilon_2 + \varepsilon_1 a_2^2)),$$

$$S_{12} = \frac{a_2}{a_2 - a_1}(\varepsilon_2 - \varepsilon_1) \tag{16a}$$

$$Q_{21} = \frac{1}{a_2 - a_1}(a_1(\varepsilon_1\omega_y^2 - \varepsilon_2\omega_\varphi^2) - p(\varepsilon_2 + \varepsilon_1 a_1^2)),$$

$$S_{21} = \frac{-a_1}{a_2 - a_1}(\varepsilon_2 - \varepsilon_1) \tag{16b}$$

Setting $\varepsilon_2=0$ in Eq. (16) corresponds to System 1 and $\varepsilon_1=0$ to System 2.

**3.2 Stability Conditions.** In this section analytical conditions for parametric antiresonance are stated for the systems in Eqs. (3*a*), (3*b*), (4*a*), (4*b*), (5*a*), and (5*b*) as derived in Ref. [26] and later, using an approach based on Ref. [30], in Ref. [13]. It was proven that whether the parametric excitation in a system is due to a periodic stiffness, damping, and/or inertia variation, the necessary condition

$$\Theta_{11} + \Theta_{22} > 0 \tag{17}$$

must be met. To achieve full vibration suppression for the system in Eqs. (3*a*), (3*b*), (4*a*), (4*b*), (5*a*), and (5*b*) at the parametric combination frequencies of order $N=1$, $\omega=\omega_0=|\Omega_1 \mp \Omega_2|$ in Eq. (2), in addition to Eq. (17), the following condition needs to be satisfied:

$$\Theta_{11}\Theta_{22} \pm \frac{1}{4}\left(\frac{Q_{12}}{\Omega_2} - \omega_0 S_{12}\right)\left(\frac{Q_{21}}{\Omega_1} \pm \omega_0 S_{21}\right) > 0 \tag{18}$$

where either only the upper signs or only the lower signs are valid. In the special case of self-excitation the condition $\Theta_{11}\Theta_{22} < 0$ holds, and Eq. (18) demands the necessary but not sufficient condition

$$\left(\frac{Q_{12}}{\Omega_2} - \omega_0 S_{12}\right)\left(\frac{Q_{21}}{\Omega_1} \pm \omega_0 S_{21}\right) \gtrless 0 \tag{19}$$

For all three systems, the necessary stability condition in Eq. (17) simplifies to

$$\Theta_{11} + \Theta_{22} = \beta + \kappa > 0 \tag{20}$$

i.e., the sum of the damping coefficients must be positive. For the parameters in Table 1 this condition is met. Similar to Ref. [15], the second stability condition in Eq. (18) determines the minimum excitation amplitude to achieve a stabilizing parametric antiresonance and is investigated for each system separately.

1. *System 1*. For $\varepsilon_2=0$, the crucial coefficients of the quasinormal form become

$$Q_{12} = \frac{-\varepsilon_1 a_2}{a_2 - a_1}(\omega_y^2 - a_2 p), \quad S_{12} = \frac{-\varepsilon_1 a_2}{a_2 - a_1} \tag{21a}$$

$$Q_{21} = \frac{\varepsilon_1 a_1}{a_2 - a_1}(\omega_y^2 - a_1 p), \quad S_{21} = \frac{\varepsilon_1 a_1}{a_2 - a_1} \tag{21b}$$

The stability condition in Eq. (18) becomes for $\omega_0 = |\Omega_1 \mp \Omega_2|$, see Eq. (2),

$$\Theta_{11}\Theta_{22} \mp \frac{\varepsilon_1^2}{4}\frac{a_1 a_2}{(a_2 - a_1)^2}\left(\frac{\omega_y^2 - a_2 p}{\Omega_2} - \omega_0\right)\left(\frac{\omega_y^2 - a_1 p}{\Omega_1} \pm \omega_0\right) > 0 \tag{22}$$

2. *System 2*. Similar to System 1, for $\varepsilon_1=0$ the following crucial parameters for the stability condition in Eq. (18) can be identified:

$$Q_{12} = \frac{\varepsilon_2}{a_2 - a_1}(a_2\omega_\varphi^2 + p), \quad S_{12} = \frac{\varepsilon_2 a_2}{a_2 - a_1} \tag{23a}$$

$$Q_{21} = \frac{-\varepsilon_2}{a_2 - a_1}(a_1\omega_\varphi^2 + p), \quad S_{21} = \frac{-\varepsilon_2 a_1}{a_2 - a_1} \tag{23b}$$

Now, the stability condition in Eq. (18) reads

$$\Theta_{11}\Theta_{22} \mp \frac{\varepsilon_2^2}{4(a_2 - a_1)^2}\left(\frac{a_2\omega_\varphi^2 + p}{\Omega_2} + \omega_0 a_2\right)\left(\frac{a_1\omega_\varphi^2 + p}{\Omega_1} \mp \omega_0 a_1\right) > 0 \tag{24}$$

3. *System 3*. Finally, the stability condition in Eq. (18) for the combination of Systems 1 and 2 becomes

$$\Theta_{11}\Theta_{22} \mp \frac{1}{4(a_2 - a_1)^2} \times\left(\frac{a_2(\varepsilon_1\omega_y^2 - \varepsilon_2\omega_\varphi^2) - p(\varepsilon_2 + \varepsilon_1 a_2^2)}{\Omega_2} + \omega_0 a_2(\varepsilon_2 - \varepsilon_1)\right) \times\left(\frac{a_1(\varepsilon_1\omega_y^2 - \varepsilon_2\omega_\varphi^2) - p(\varepsilon_2 + \varepsilon_1 a_1^2)}{\Omega_1} \mp \omega_0 a_1(\varepsilon_2 - \varepsilon_1)\right) > 0 \tag{25}$$

Thus, there is an important influence of the phase shift between both parametric excitation components $\varepsilon_1$ and $\varepsilon_2$. The stability condition in Eq. (25) can be used to determine the necessary minimum excitation $\varepsilon_1$ for a given excitation $\varepsilon_2$ and vice versa.

In the case of synchronous excitation, $\varepsilon_1=\varepsilon_2=\varepsilon$, the quasinormal form coefficients $S_{12}$ and $S_{21}$ vanish according to Eq. (16). In this case the stability condition in Eq. (25) simplifies to

$$\Theta_{11}\Theta_{22} \pm \frac{Q_{12}Q_{21}}{4\Omega_1\Omega_2} > 0 \tag{26}$$

which is exactly the relation to achieve a parametric antiresonance for a pure stiffness variation, see Ref. [12]. For this synchronous excitation the parameters $Q_{12}$ and $Q_{21}$ reach their minimum value.

**Table 2 Resonance and antiresonance frequencies**

| Frequency | $2\Omega_2$ | $\Omega_1+\Omega_2$ | $2\Omega_1$ | $\Omega_2$ | $(\Omega_1+\Omega_2)/2$ | $\Omega_1$ | $2\Omega_2/3$ | $\Omega_2-\Omega_1$ |
|---|---|---|---|---|---|---|---|---|
| Value | 1.98 | 1.75 | 1.52 | 0.99 | 0.87 | 0.76 | 0.65 | 0.23 |

## 4 Numerical Calculations

The analytical predictions in Sec. 3 are compared with numerical calculations for specific values of the physical properties. The system parameters considered are listed in Table 1 determining the quasinormal form coefficients in Eqs. (16), (21), and (23). Some of the resulting parametric resonance and antiresonance frequencies in Eq. (2) are listed in Table 2.

First, the stability of System 1 in Eq. (3) is analyzed, $\varepsilon_2=0$. The numerical surface for largest eigenvalue in dependency of the parametric excitation frequency $\omega$ and the excitation amplitude $\varepsilon_1$ is plotted in Fig. 1(*a*). The system is stable for negative eigenvalues. Due to the self-excitation introduced by the negative damping coefficient $\kappa$, the system without parametric excitation, $\varepsilon_1=0$, possesses positive eigenvalues and is unstable. This is visualized by the horizontal level in Fig. 1(*a*). Activating the parametric excitation, $\varepsilon_1 \neq 0$, deforms this horizontal level. The parametric excitation is termed resonant in regions where this level is deformed toward negative eigenvalues and antiresonant in regions with deformation toward positive eigenvalues. The parametric resonances at the frequencies $2\Omega_1$, $2\Omega_2$, $\Omega_2$, and $\Omega_2-\Omega_1$ can be clearly identified. The most dominant destabilizing resonance is located at the principal parametric resonance frequency $2\Omega_1$, which corresponds to the amplitude $\varepsilon_1$. A parametric antiresonance of first order occurs at the frequency $\Omega_1+\Omega_2$. A further parametric antiresonance of second order is located at $(\Omega_1+\Omega_2)/2$ but is of negligible stabilizing effect.

Cutting the stability surface in Fig. 1(*a*) at the level 0 yields the stability boundary in Fig. 1(*b*). The shaded region denotes the stable antiresonance region. For the parameters as listed in Table 1, a parametric antiresonance at $|\Omega_1-\Omega_2|$ does not occur since the necessary condition in Eq. (19) is violated. However, this condition is satisfied at $\omega=\omega_0=\Omega_1+\Omega_2$. The analytical condition in Eq. (22) predicts the minimum value of $\varepsilon_1$ that is necessary to achieve parametric antiresonance to be 0.115. This value matches exactly the numerical one in Fig. 1(*b*). Summarizing, the self-excited System 1 can be effectively stabilized by the method of damping by parametric excitation.

The numerical stability boundaries of System 2 in Eq. (4), $\varepsilon_1=0$, are shown in Fig. 2. Similar to Fig. 1(*a*), parametric resonances occur at the frequencies $2\Omega_1$, $2\Omega_2$, $\Omega_1$, $\Omega_2$, and $\Omega_2-\Omega_1$ in Fig. 2(*a*). However, switching from System 1 ($\varepsilon_2=0$) to System 2 ($\varepsilon_1=0$) switches the dominant resonance frequency from $2\Omega_1$ to $\Omega_2$. The stability chart corresponding to Fig. 2(*a*) is plotted in Fig. 2(*b*). For the values in Table 1, the minimum value for the parametric excitation amplitude becomes $\varepsilon_2=0.115$ for a parametric antiresonance at $\omega=\omega_0=\Omega_1+\Omega_2$. Again, a parametric antiresonance at $|\Omega_1-\Omega_2|$ is not observed.

For both systems the stable antiresonance at $\Omega_1+\Omega_2$ is located in between the unstable principal parametric resonances $2\Omega_1$ and $2\Omega_2$. In regions where the parametric antiresonance is close to parametric resonance region, the stabilizing and the destabilizing effect interact. With increasing $\varepsilon_1$ in System 1, the parametric resonance frequency $2\Omega_1$ becomes more and more dominant and shifts the antiresonance region slightly toward the less dominant region $2\Omega_2$ and vice versa for increasing $\varepsilon_2$ in System 2. The antiresonance region is shifted toward the resonance region that is less dominant. Consequently, the skeleton line is shifted as well so that the minimum values of the largest eigenvalues are not found exactly at the analytically predicted frequency $\Omega_1+\Omega_2$ but slightly to the left in Fig. 1(*b*) and to the right in Fig. 2(*b*). For System 1 in Fig. 1(*b*) this shift is so large that the antiresonance region interacts strongly near $\omega \approx 1.5$ with the resonance at $2\Omega_2$. Here, the stabilizing effect is still present but becomes less effective than the destabilizing $2\Omega_2$ effect. The parametric antiresonance disappears and the antiresonance region is cut.

Finally, the stability of System 3 is determined. For a synchronous parametric excitation, $\varepsilon_1=\varepsilon_2$, the parameters $Q_{12}$ and $Q_{21}$ reach their minimum value so that the right expression in Eq. (26) becomes too small to compensate the negative left hand term. Hence, a parametric antiresonance cannot be established at the frequency of summation and difference type, respectively. The optimal case is when $\varepsilon_1=-\varepsilon_2=\varepsilon$, i.e., the excitations are in opposite phase, where the minimum value of $\varepsilon$ to achieve parametric antiresonance at the summation frequency becomes 0.057. Similar to Systems 1 and 2, an antiresonance cannot be achieved at the difference frequency.

Figure 3 shows the stability in dependency of the excitation amplitudes $\varepsilon_1$ and $\varepsilon_2$ according to Eq. (25). Only the part in the

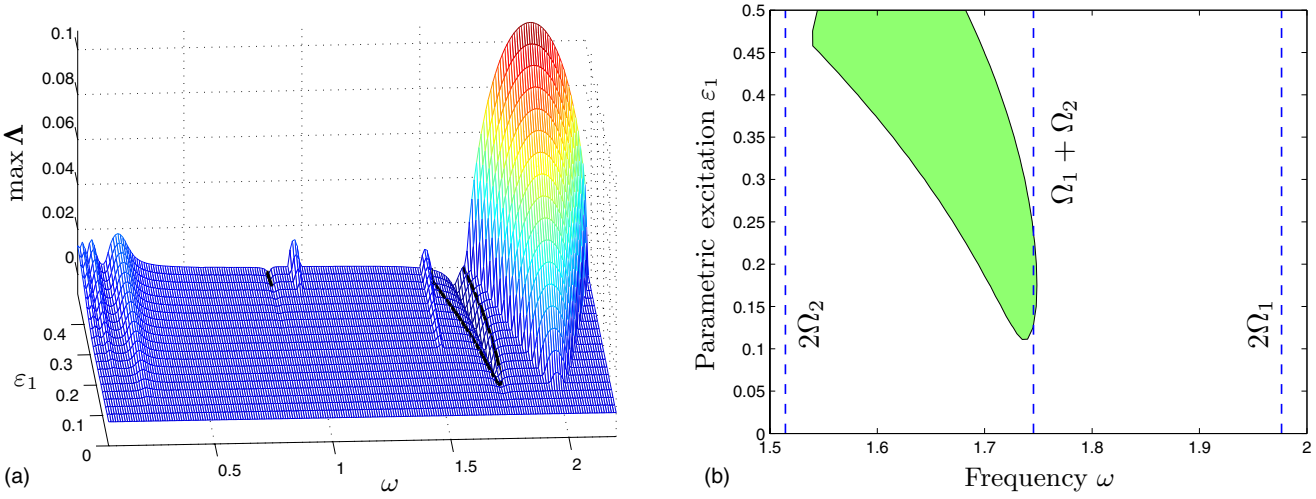



**Fig. 1 Numerical stability boundaries for System 1. (*a*) Numerical surface for largest eigenvalue. (*b*) Stability chart at max $\Lambda=0$. The shaded region denotes stable antiresonance region.**

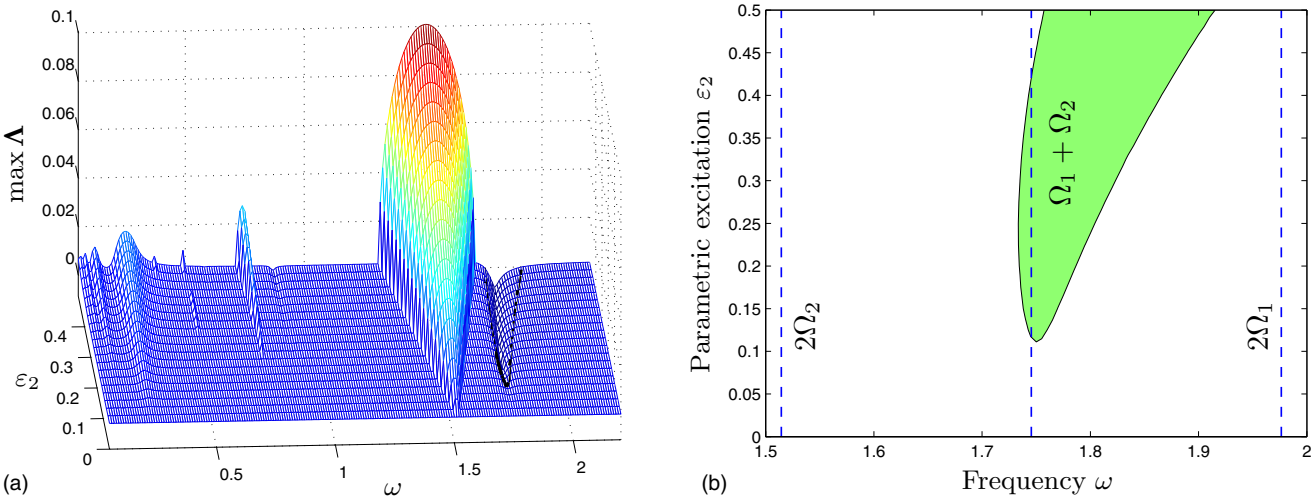



**Fig. 2 Numerical stability boundaries for System 2. (*a*) Numerical surface for largest eigenvalue. (*b*) Stability chart at max $\Lambda=0$. The shaded region denotes stable antiresonance region.**

vicinity of the main parametric antiresonance frequency $\Omega_1+\Omega_2$ is shown. For very high values of $\varepsilon_i$ tiny shapes occur also near $(\Omega_1+\Omega_2)/N$ for $N\geq 2$. Parameter combinations of $\varepsilon_1$ and $\varepsilon_2$ lying inside this stability body correspond to a stable system. A vertical slice at $\varepsilon_2=0$ corresponds to System 1 in Fig. 1(*b*). A horizontal slice at $\varepsilon_1=0$ corresponds to System 2 in Fig. 2(*b*). Similar to Figs. 1(*b*) and 2(*b*), the stability body is cut by the principal parametric resonance $2\Omega_1$ close to $\omega\approx 1.5$ and $2\Omega_2$ close to $\omega\approx 1.8$. In general, the wider the frequency interval becomes for fixed values of $\varepsilon_i$, the more the maximum eigenvalue is decreased, see Figs. 1(*a*) and 2(*a*). Hence, in general, increasing $\varepsilon_i$ increases the stabilizing effect, see Fig. 3. However, since the interaction in the example system is so large, at higher values of $|\varepsilon_i|$, the stabilizing effect is still present but becomes less effective than the destabilizing $2\Omega_2$ effect and the parametric antiresonance disappears.

The stability body in Fig. 3 is point symmetric around the $\varepsilon_1=\varepsilon_2=0$. As predicted by the analytical calculations, a synchronous parametric excitation, $\varepsilon_1=\varepsilon_2$, cannot establish a parametric antiresonance. An out of phase excitation, $\varepsilon_1=-\varepsilon_2$, corresponds to a slice at 45 deg. For this excitation the shortest distance of $|\varepsilon|=0.057$ to the stability body is achieved, which confirms the analytical prediction.

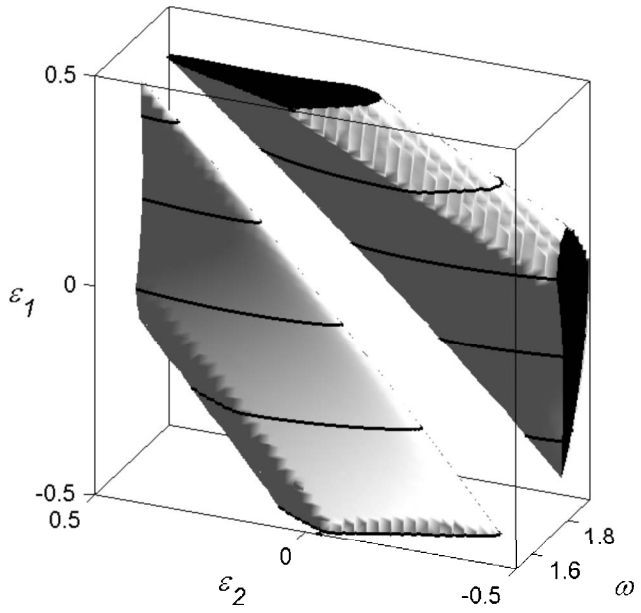



**Fig. 3 Numerical stability boundaries for System 3**

To emphasize the influence of the phase relation between the parametric excitation amplitudes $\varepsilon_i$, the stability body in Fig. 3 is cut for a fixed arbitrary value of $|\varepsilon_1|=0.1$. A slice for $\varepsilon_1=0.1$ is shown in Fig. 4(*b*) and for $\varepsilon_1=0.1$ in Fig. 5(*b*). Hence, the minimum value of $|\varepsilon_2|$ in Figs. 1(*b*) and 2(*b*) increases for $\varepsilon_1>0-0.211$ and decreases for $\varepsilon_1<0-1.010$. Similar results are obtained for a chosen fixed value of $\varepsilon_2$. Again, these values perfectly match the analytical predictions according to Eq. (25).

In the present study, the parametric mass and inertia excitation of the example system in Table 1 is analyzed for the case of synchronous variation, $\varepsilon_1=\varepsilon_2$, and the case of out of phase variation, $\varepsilon_1=-\varepsilon_2$. It was found that from these two cases the out of phase variation is optimal. For a parametric stiffness excitation of a two degrees of freedom system the optimal phase between stiffness excitations was proven to be at 0 deg (synchronous) or 180 deg (out of phase), depending on the system parameters [31]. Consequently, for the present study a more optimal configuration for an arbitrary phase relation between $\varepsilon_1$ and $\varepsilon_2$ is not expected.

Finally, it should be noted that it is necessary to avoid a tuning of the system close to its internal resonance, e.g., if the parametric combination resonance frequency of the sum type and first order, $\Omega_1+\Omega_2$, becomes close to a principal parametric resonance frequency of the same order, $2\Omega_1$ and $2\Omega_2$.

## 5 Conclusions

The suppression of selfexcited vibrations of an aero-elastic system with bending and torsion modes is investigated. The application of semi-active means is realized by a parametric excitation of mass and/or moment of inertia that may lead to full vibration suppression within certain limits of the system parameters if the amplitude and frequency are tuned properly. The frequency of the stiffness variation can be determined in advance, which leads to an open-loop control, without any feedback information necessary. A necessary condition for the application of this method is that the unstable system possesses at least one stable mode. As with previously investigated other self-excited systems, the parametric antiresonance effect was able to suppress the vibrations and to stabilize the system for frequencies close to the parametric combination frequency of the difference type. The outstanding feature of the physically feasible example systems presented is its

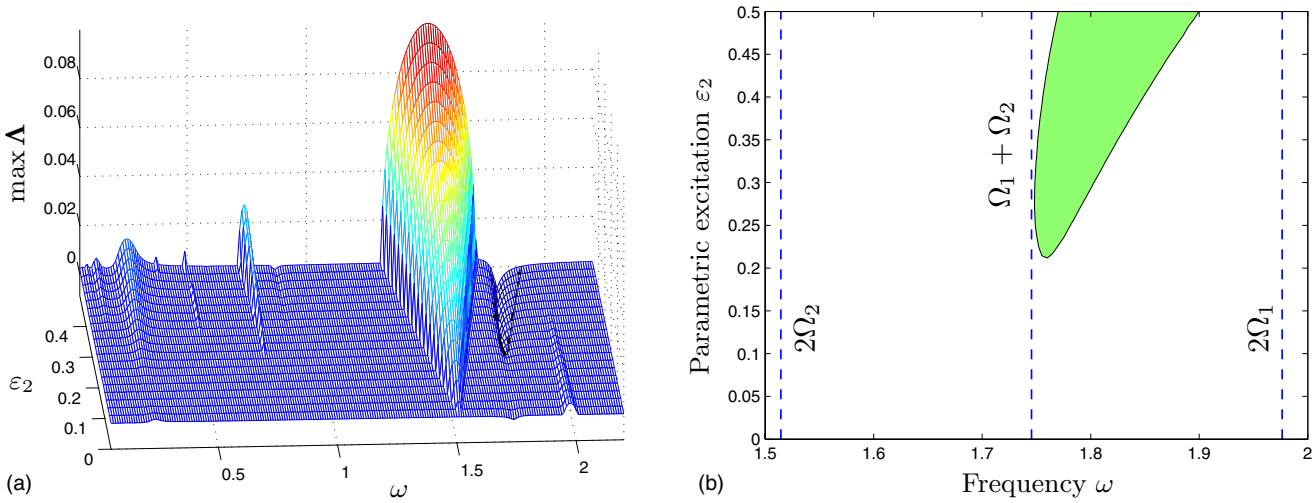



**Fig. 4 Numerical stability boundaries for System 3 for $\varepsilon_1 = +0.1$. (*a*) Numerical surface for largest eigenvalue. (*b*) Stability chart at max $\Lambda = 0$. The shaded region denotes stable antiresonance region.**

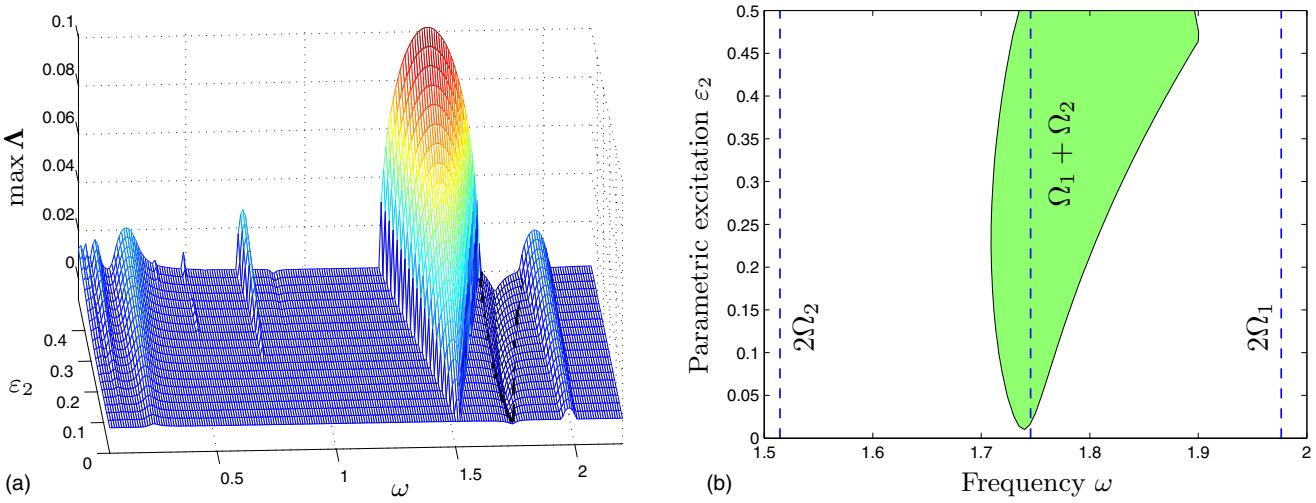



**Fig. 5 Numerical stability boundaries for System 3 for $\varepsilon_1 = -0.1$. (*a*) Numerical surface for largest eigenvalue. (*b*) Stability chart at max $\Lambda = 0$. The shaded region denotes stable antiresonance region.**

ability to achieve full vibration suppression, parametric antiresonance, close to the parametric combination frequency of the sum type.

## References


[1] Simiu, E., and Scanlan, R. H., 1984, *Wind Effects on Structures: An Introduction to Wind Engineering*, Wiley, New York.

[2] Scanlan, R. H., 1993, "Problematics in Formulation of Wind-Force Models for Bridge Decks," J. Eng. Mech., **119**(7), pp. 1353–1375.

[3] Pospíšil, S., Náprstek, J., and Hračov, S., 2006, "Stability Domains in Flow-Structure Interaction and Influence of Random Noises," J. Wind. Eng. Ind. Aerodyn., **94**, pp. 883–893.

[4] Náprstek, J., Pospíšil, S., and Hračov, S., 2007, "Analytical and Experimental Modelling of Non-Linear Aeroelastic Effects on Prismatic Bodies," J. Wind. Eng. Ind. Aerodyn., **95**, pp. 1315–1328.

[5] Ricciardelli, F., de Grenet, E. T., and Hangan, H., 2002, "Pressure Distribution, Aerodynamic Forces and Dynamic Response of Box Bridge Sections," J. Wind. Eng. Ind. Aerodyn., **90**(10), pp. 1135–1150.

[6] Bolotin, V. V., 1963, *Non-Conservative Problems of the Theory of Elastic Stability*, Pergamon, Oxford.

[7] Kirillov, O. N., 2007, "Destabilization Paradox Due to Breaking the Hamiltonian and Reversible Symmetry," Int. J. Non-Linear Mech., **42**, pp. 71–81.

[8] O'Reilly, O. M., Malhotra, N., and Sri Namachchivaya, N., 1995, "Reversible Dynamical Systems: Dissipation-Induced Destabilization and Follower Forces," Appl. Math. Comput., **70**, pp. 273–282.

[9] Náprstek, J., 2001, "Flutter Instability Due to Non-Conservative and Gyroscopic Forces (in Czech)," *Proceedings of the Colloquium Dynamics of Machines*, Institute of Thermomechanics ASCR, Prague, Feb. 6 and 7.

[10] Pourzeynali, S., and Datta, T. K., 2002, "Control of Flutter of Suspension Bridge Deck Using TMD," Wind Struct., **5**(5), pp. 407–422.

[11] Wilde, K., Fujino, Y., and Kawakami, T., 1999, "Analytical and Experimental Study on Passive Aerodynamic Control of Flutter of a Bridge Deck," J. Wind. Eng. Ind. Aerodyn., **80**(1-2), pp. 105–119.

[12] Tondl, A., 1978, "On the Interaction Between Self-Excited and Parametric Vibrations," Monographs and Memoranda 25, National Research Institute for Machine Design, Prague.

[13] Dohnal, F., 2005, "Damping of Mechanical Vibrations by Parametric Excitation," Ph.D. thesis, Vienna University of Technology, Vienna, Austria.

[14] Tondl, A., 1998, "To the Problem of Quenching Self-Excited Vibrations," Acta Tech. CSAV, **43**, pp. 109–116.

[15] Dohnal, F., Ecker, H., and Tondl, A., 2004, "Vibration Control of Self-Excited Oscillations by Parametric Stiffness Excitation," *Proceedings of the 11th International Congress of Sound and Vibration (ICSV11)*, St. Petersburg, Russia, Jul., pp. 339–346.

[16] Dohnal, F., 2008, "Damping by Parametric Stiffness Excitation—Parametric Resonance and Anti-Resonance," J. Vib. Control, **14**, pp. 669–688.

[17] Ecker, H., and Tondl, A., 2000, "Suppression of Flow-Induced Vibrations by a Dynamic Absorber With Parametric Excitation," *Proceedings of the Seventh International Conference on Flow-Induced Vibrations (FIV2000)*, Switzerland, Jun.

[18] Tondl, A., 1991, *Quenching of Self-Excited Vibrations*, Elsevier, Amsterdam.

[19] Tondl, A., and Ecker, H., 1999, "Cancelling of Self-Excited Vibrations by

Means of Parametric Excitation," *Proceedings of the 1999 ASME Design Engineering Technical Conferences (DETC)*, Las Vegas, NV, Sept. 12–15.
[20] Tondl, A., 2002, "Three-Mass Self-Excited Systems With Parametric Excitation," Acta Tech. CSAV, **47**, pp. 165–176.
[21] Ecker, H., Pumhössel, T., and Tondl, A., 2002, "A Study on Parametric Excitation for Suppressing Self-Excited Rotor Vibrations," *Proceedings of the Sixth International Conference on Rotor Dynamics (IFToMM)*, Sydney, Sept.
[22] Tondl, A., 2001, "Self-Excited Vibration Quenching in a Rotor System by Means of Parametric Excitation," Acta Tech. CSAV, **45**, pp. 199–211.
[23] Dohnal, F., Ecker, H., and Springer, H., 2008, "Amplification of Damping of an Elastic Beam by Parametric Excitation," Arch. Appl. Mech., **78**, pp. 935–947.
[24] Tondl, A., 2003, "Combination Resonances and Anti-Resonances in Systems Parametrically Excited by Harmonic Variation of Linear Damping Coefficients," Acta Tech. CSAV, **48**, pp. 239–248.
[25] Dohnal, F., 2005, "Vibration Suppression of Self-Excited Oscillations by Parametric Inertia Excitation," *Proceedings of the 76th International Conference of Gesellschaft für Angewandte Mathematik und Mechanik (GAMM)*, Luxembourg, Feb.
[26] Tondl, A., 2001, "Systems With Periodically Variable Masses," Acta Tech. CSAV, **46**, pp. 309–321.
[27] Dohnal, F., 2007, "Suppressing Self-Excited Vibrations by Synchronous and Time-Periodic Stiffness and Damping Variation," J. Sound Vib., **306**(1–2), pp. 136–152.
[28] Yakubovich, V. A., and Starzhinskii, V. M., 1975, *Linear Differential Equations With Periodic Coefficients*, Vols. 1 and 2, Wiley, London.
[29] Tondl, A., 1959, "The Method for the Determination of Instability Intervals of Quasi-Harmonic Vibration Systems (in Czech)," Applications of Mathematics, **4**(4), pp. 278–289.
[30] Abadi, A., 2003, "Nonlinear Dynamics of Self-Excitation in Autoparametric Systems," Ph.D. thesis, Utrecht University, Utrecht, The Netherlands.
[31] Dohnal, F., "Optimal Dynamic Stabilisation of a Linear System by Periodic Stiffness Excitation," J. Sound Vib., in press.

# Mechanism Maps for Frictional Attachment Between Fibrillar Surfaces

**Robert M. McMeeking**
Department of Mechanical Engineering,
and Department of Materials,
University of California,
Santa Barbara, CA 93106; and
Leibniz Institute for New Materials (INM),
Campus D2 2,
66123 Saarbruecken, Germany

**Lifeng Ma**
MOE Key Laboratory for Strength and Vibration,
Department of Engineering Mechanics,
Xi'an Jiaotong University,
Xi'an 710049, China

**Eduard Arzt**
Leibniz Institute for New Materials (INM),
Campus D2 2,
66123 Saarbruecken, Germany

*The mechanics of frictional attachment between surfaces with pillars, inspired by the head fixation system of dragonflies, is analyzed. The system consists of two surfaces of interdigitating pillars held together through friction, as by the densely packed bristles of two brushes when pressed together. The adhesive strength of the system is promoted by high elastic modulus, high friction coefficient, large aspect ratio, and dense packing of the fibers. However, the design is limited by the compressive buckling, the compressive indentation or cracking of the contacting pillars, yielding in shear or similar mechanisms that limit the achievable friction stress, and tensile failure of the pillars upon pull-out. Maps, which summarize the strength of the adhesive system and the failure limits and illustrate the trade-off among the design parameters, are presented. Case studies for steel, nylon, and ceramic pillars show that useful strength can be achieved in such attachments; when buckling during assembly and contact failure can be avoided, adhesive performance as high as 30% of the tensile strength of the pillar material may be possible.* [DOI: 10.1115/1.3002760]

## 1 Introduction

Adhesion by surface patterning is a common reversible fixation mechanism in biology. For example, the cordulegastrid dragonfly *Anotogaster sieboldii* can stabilize its head, e.g., [1], by moving it dorsally until two sets of opposed tapered pillars, one on the head and the other on the thorax, engage each other by sliding, similar to the scheme shown in Fig. 1. The contact between the sliding surfaces both stabilizes the lateral motion of the head relative to the thorax and inhibits its axial motion due to the friction that is generated if the pillars are required to slide axially relative to each other. It is of interest to investigate the mechanics of such an attachment to obtain insights into the physical phenomena that might control its strength and stability. Such features possibly include Hertzian contact, friction and adhesion among the pillars [2], buckling of the pillars during engagement [3], failure due to tensile overload of the pillars upon separation [4], yielding in shear upon pull-out that limits the friction stress [4], and damage of the contacts between the pillars due to excessive compression [2].

The system depicted in Fig. 1 inspires the possibility of artificial attachment devices based on the friction of interdigitating pillars. The possible design of such a system is depicted in Fig. 2. It should be noted that the proposed arrangement is commonly utilized by those who store their shoe brushes by pressing their bristle pads into each other. In such a configuration, the two brushes will adhere to each other due to the friction between the contacting bristles. In addition, the potential of the proposed system can be gauged from the effectiveness of nails, which, when driven into wood, are capable of very high attachment strengths through friction alone. In the absence of an interference fit between the wood and the nail, it will be pulled out easily. This example indicates clearly that the effectiveness of the proposed system depends on a design that can generate enough constraint so that sufficient friction can be created between the pillars when they are interdigitated.

The design of the fibrillar surfaces is envisaged to have a regular array of pillars on a hexagonal grid. Two identical surfaces can thus be interdigitated with each pillar surrounded by three from the opposite surface, in a pattern depicted in Fig. 3. The intention is for each pillar to remain stably constrained by the three with which it is in contact. Note that pillars on a square grid would appear to have the same property, but for geometric reasons associated with packing densities, a hexagonal array provides better adhesion/cohesion. The pillars are considered to be straight, without taper, so that a large length of overlap can be achieved without the attachment system being easily pulled apart. Interdigitating tapered fibrils as utilized by the dragonfly will be treated in future work. The spacing of the pillars must be such that an interference fit is obtained among the contacting fibrils. As a consequence, Hertzian contact will be achieved for elastic pillars, and a contact pressure will be generated between each pair of fibrils. The Hertzian pressure will generate friction by Amonton's law [2] and will thus constrain the pillar surfaces from being pulled apart unless sufficient tension is applied across the attachment system. Note, however, that this means that the system can only be deployed to a given strength level if the same degree of compression is applied when the attachment is assembled. Contrasts between static and dynamic friction may modify this in such a way that the tension to pull the system apart will exceed the compression to press the system together.

As noted above, the system will only deliver high adhesion strength if the required elastic constraint is not degraded. One obvious mechanism for loss of constraint is that the pillars may splay out and relax the contact compression. This will be most apparent on the perimeter of the pillar surfaces and can be neglected when the area is very large compared with the cross-section of the pillars. Another solution is to provide a stiff wall around the pillar surfaces. In the present paper, the unconstrained edge zone will be considered to be absent. The other obvious mechanism for loss of Hertzian pressure is if the compression among the pillars indents them plastically or cracks them. Thus a design constraint is presented in which the pressure at the contacts must be kept below a critical level. A high strength system necessarily generates a high tensile stress on the pillars when the surfaces are pulled apart. The performance of the attachment strategy is thus limited by the tensile strength of the pillars, another design constraint. Furthermore, damaging stress concentrators at the junction between the pillars and the system substrate must be

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received July 15, 2007; final manuscript received June 18, 2008; published online March 9, 2009. Review conducted by Kenneth M. Liechti.

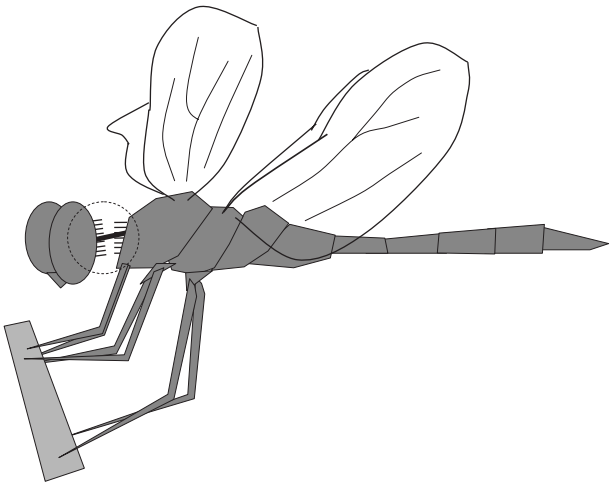


**Fig. 1 Schematic of the head stabilization mechanism in a dragonfly**

avoided. This design consideration will be considered to have been respected in what is presented below. The friction stress will also be limited by yielding in shear at the contact upon pull-out or by other shear failure mechanisms.

Problems can arise also when the system is pressed together. Most importantly, buckling of the pillars as they are axially compressed while engaging the attachment system is a fundamental limitation and will be addressed by analysis and as a design constraint in the present paper. We assume further that the tips are appropriately formed to avoid situations where pillars contact each other head-on or where two or more pillars become entrained in a gap meant to house only one.

Further complications can be anticipated. For example, if the diameter of the pillars is extremely small, JKR type contact cohesion or adhesion may be significant [2,5]. An obvious difficulty caused by this would be the clumping together of the pillars on a given surface due to adhesion. This would make the attachment system inoperative because regular interdigitation of the pillars will not occur. In addition, JKR adhesion would complicate the nature of friction at the contact, both inherently and due to the small scale of the contact [6,7], so that deviations from the anticipated performance will occur. For the purposes of the current paper, JKR type adhesion among the pillars will be neglected.

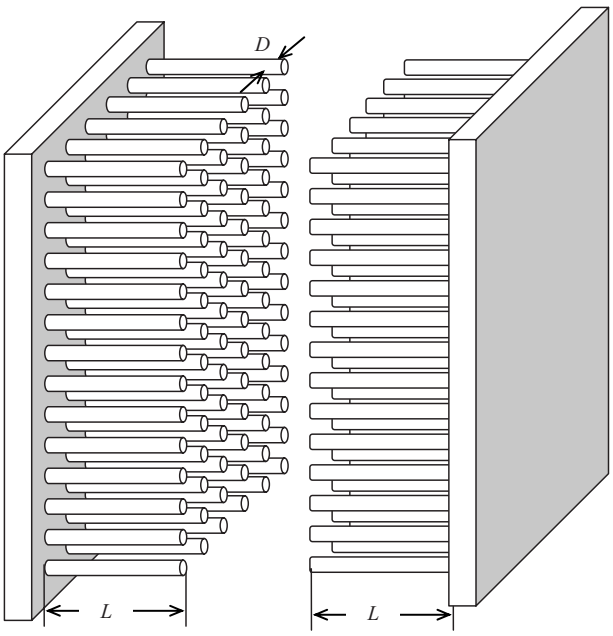



**Fig. 2 Attachment system consisting of two pillar arrays to be pressed together so that the pillars intersect and contact each other. The pull-off strength is due to the interpillar friction.**

In the present paper, we will explore the mechanics of a fibrillar attachment device based on friction, with special view to its limits due to compressive buckling, contact failure, and tensile failure. The approach will extend earlier models (see, e.g., Ref. [8]) and will provide a theoretical base for the practical design of frictional systems.

## 2 A Model for the Strength of the Attachment System

Consider an array of pillars laid out in a regular hexagonal pattern with spacing $s$ between centers (Fig. 3). The diameter of each pillar is $D$ and their length is $L$. Thus to ensure Hertzian contact between pillars from opposite surfaces, $s < \sqrt{3}D$. The difference

$$\delta = D - \frac{s}{\sqrt{3}} \tag{1}$$

is assumed to be small compared with $D$ so that infinitesimal strain Hertzian analysis is relevant.

Now consider two pillar surfaces as depicted in Fig. 4($a$) in which they are free from applied load and are just touching at the tips of the pillars. Next, compressive forces are applied to drive the surfaces together by a relative distance $\Delta$, as depicted in Fig. 4($b$). Thus, contact zones of length $\Delta$ and width $2a$ will be present between each pair of touching pillars, as indicated in Fig. 4($c$). When the attachment system is engaged, the elastic deformations, other than those due to frictional sliding, are tantamount to taking two parallel cylinders of length $\Delta$ and moving one cylinder toward the other by a distance $\delta$, starting from the line contact. When both pillar surfaces are made from the same linear elastic material, each experiences a deformation into the shape shown in Fig. 5.

Assuming that the contact strains remain small and the width of contact on the pillars is negligible compared with their circumference, Hertzian analysis [9] gives the average normal pressure as

$$p = \frac{\pi a E}{4D(1-\nu^2)} \tag{2}$$

when no length change of the cylinders is permitted and thus plane strain conditions prevail. This pressure will be accurate for the problem under consideration, except in a segment at each end of the contact approximately $2a$ in extent. It follows that when the contact length $\Delta$ is much greater than $4a$, the result in Eq. (2) is representative of the average pressure on the entire contact. In addition, the relationship between the parameters of Hertzian contact is given by [9]

$$\frac{\delta}{D} = \frac{4pa(1-\nu^2)}{\pi E D}\left[2\ln\frac{2D}{a} - 1\right] = \frac{a^2}{D^2}\left[2\ln\frac{2D}{a} - 1\right] \tag{3}$$

These expressions require numerical iteration to calculate $a$ for a given $\delta$. Thereafter, Eq. (2) can be used to obtain $p$.

Now consider an attachment system that has been interdigitated by pressing it together by the distance $\Delta$. To pull it apart, the frictional resistance of the sliding contacts must be overcome. Each pillar contacts three on the opposite surface, so the tensile force $P$ on one pillar (see Fig. 4($c$)) to pull the pillar surfaces apart is

$$P = 6\mu p a \Delta \tag{4}$$

where $\mu$ is the coefficient of friction. Note that if the pillars are made from a metal, polymer, or other materials that yield plastically, the frictional shear stress is limited to less than or equal to the yield stress in shear. Other shear failure mechanisms may be possible dependent on which material is being used. Such con-

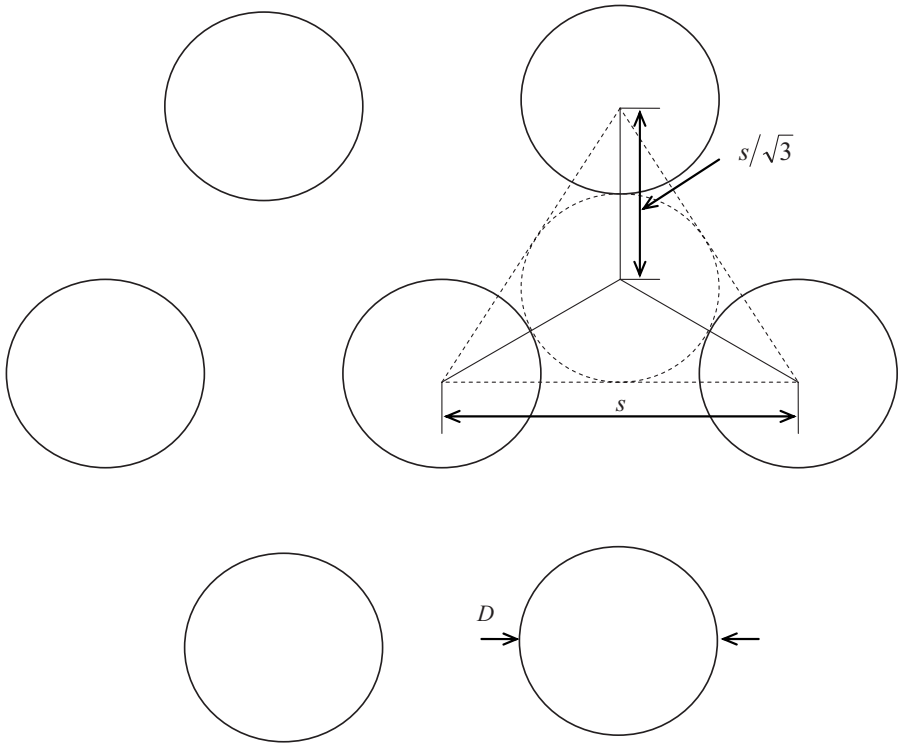



**Fig. 3 Plan view of the pillar array: the dashed circle corresponds to one interdigitating pillar**

straint limiting the frictional shear stress will be addressed below.

From Eq. (4), we can deduce that the tensile stress at the base of each pillar is

$$\sigma = \frac{24\mu pa\Delta}{\pi D^2} = \frac{6\mu a^2\Delta E}{D^3(1-\nu^2)} = \frac{96\mu p^2\Delta(1-\nu^2)}{\pi^2 DE} \tag{5}$$

which gives the result in various convenient forms. Given that $\delta \ll D$, the area fraction of the surface occupied by pillars is accurately represented as $\pi/6\sqrt{3}$. It follows that the average tensile stress required to separate the surfaces is given by

$$\Sigma = \frac{4\mu pa\Delta}{\sqrt{3}D^2} = \frac{\pi\mu a^2\Delta E}{\sqrt{3}D^3(1-\nu^2)} = \frac{16\mu p^2\Delta(1-\nu^2)}{\sqrt{3}\pi DE} \tag{6}$$

The maximum attachment strength, $\Sigma_c$, of the system occurs when the pillar surfaces are completely interdigitated so that $\Delta = L$. Thus

$$\Sigma_c = \frac{4\mu paL}{\sqrt{3}D^2} = \frac{\pi\mu a^2 LE}{\sqrt{3}D^3(1-\nu^2)} = \frac{16\mu p^2 L(1-\nu^2)}{\sqrt{3}\pi DE} \tag{7}$$

This result has been plotted in Fig. 6; the contours correspond to equal values of the normalized attachment strength $K = \Sigma_c(1-\nu^2)/\mu E$. This manner of plotting is preferred because of its universal form and its simplicity in terms of the most important variable parameters, i.e., $L/D$ and the contact strain measure, $\delta/D$, upon which $p(1-\nu^2)/E$ directly depends (Eqs. (2) and (3)). The plot, taken together with the second form on the right hand side of Eq. (7), shows that strong attachment is promoted by a high elastic modulus, high compressive strains and stresses at the contacts among the pillars, a high aspect ratio for the pillars, and a high coefficient of friction.

Now the three important failure mechanisms will be incorporated in Fig. 6. Because the maximum tensile stress in a pillar is $\sigma = 6\sqrt{3}\Sigma/\pi$, tensile failure will be avoided as long as

$$\Sigma_c < \pi\sigma_c/6\sqrt{3} \approx 0.302\sigma_c \tag{8}$$

where $\sigma_c$ is the tensile strength of the pillar material. Thus in favorable circumstances, the attachment system may deliver performance as high as 30% of the tensile strength of the pillars.

Another limiting mechanism is the compressive failure of the contacts among the pillars due to plastic indentation if the material is ductile or to compressive cracking if the material is brittle. The consequence of these phenomena will be the loss of contact pressure and thus a degradation of the strength of the attachment system. An alternative possibility is that fatigue will occur due to repeated engagements and separations of the system. Whatever the phenomenon involved, it can be avoided by limiting the contact pressure so that

$$p < p_c \tag{9}$$

where $p_c$ is the permissible pressure for the relevant failure mechanism.

As noted above, the frictional stress is limited by the yield stress in shear or any other shear criterion of failure. Let $\tau_c$ be the critical shear stress at failure or yielding. The attachment strength is thus limited to

$$\Sigma_c = \frac{4\tau_c aL}{\sqrt{3}D^2} = \frac{16\tau_c pL(1-\nu^2)}{\sqrt{3}\pi DE} \tag{10}$$

which replaces Eq. (7) when $p > \tau_c/\mu$. For clarity, this adjustment to the maps has not been made, and the contours of equal attachment strength in Fig. 6 represent the situation only when $p < \tau_c/\mu$. Violations of this restriction will be addressed in special cases below.

A more complex failure mechanism is buckling of the pillars when the attachment system is being engaged by the application of compressive force. The situation is illustrated in Fig. 7—showing the normal condition in Fig. 7(*a*) and the buckled state in Fig. 7(*b*)—in which one pillar surface moves laterally due

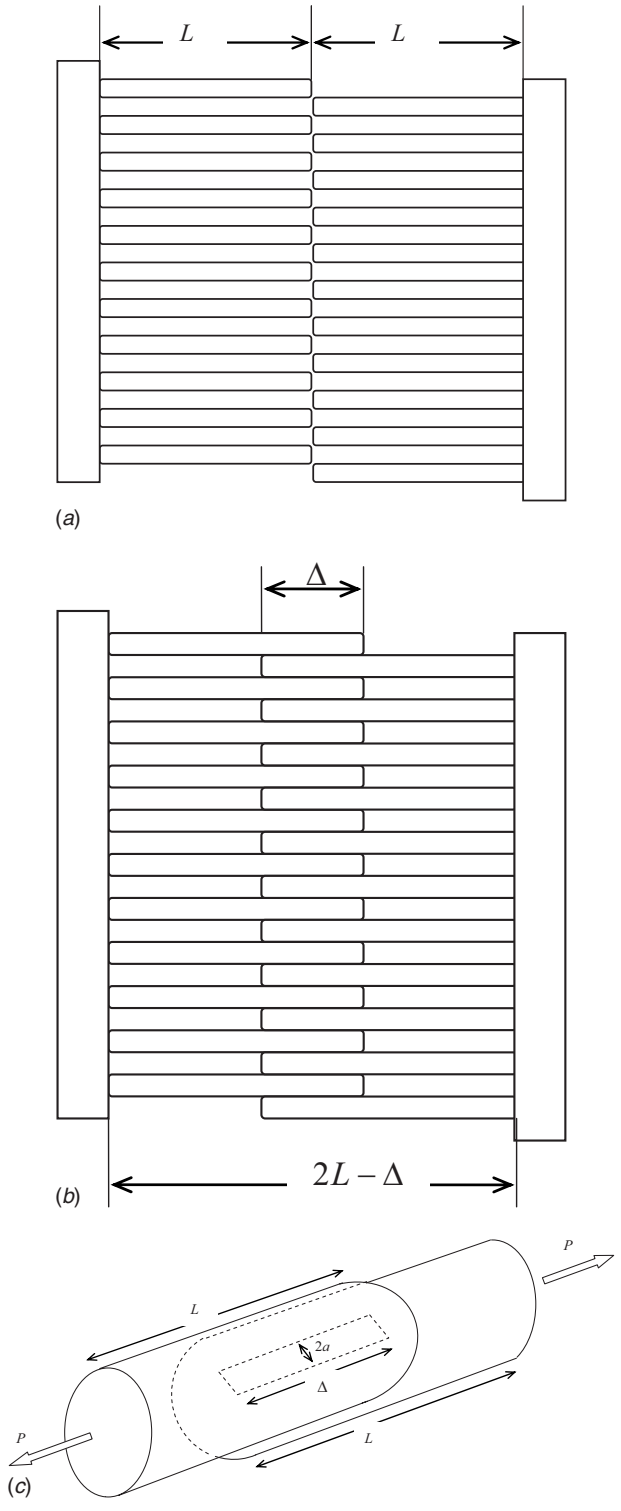



**Fig. 4 Schematic of the attachment device: (*a*) prior to being engaged to create the attachment, (*b*) after displacement by a distance Δ, and (*c*) the zone of contact between two touching pillars**

to a lack of constraint, in addition to its movement toward the other pillar surface. Note that if the pillar surfaces are constrained from such lateral motion, as when stiff shafts guide the pillar surfaces by sliding through holes, the buckling performance is improved.

The propensity to buckle will change with the degree of pillar overlap. When the overlap, Δ, is small, the column is at its longest

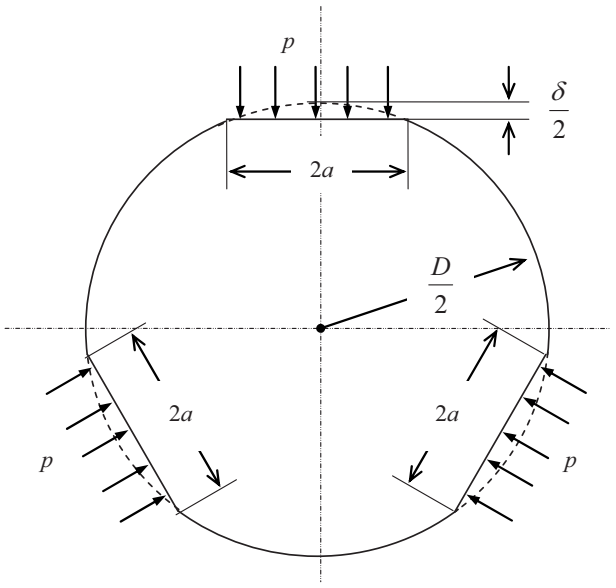



**Fig. 5 Approximate deformed cross-section of a pillar when constrained by three pillars from the opposite surface**

and most slender shape and is thus prone to buckling at smaller compressive loads. However, the frictional features of the sliding of one pillar relative to the other ensures that in this configuration, the applied compressive load, $F=6\mu pa\Delta$, is at its smallest level. Thus, buckling is not likely when Δ is small. On the other hand, when Δ is large and near its maximum level, $L$, the stockiness of the resulting column, consisting of two pillars that must buckle simultaneously, makes instability unlikely. (In fact, at this stage, the frictional constraints among the forest of contacting pillars in the system will inhibit buckling quite severely, so that buckling when $\Delta=L$ or close to $L$ is less likely than may be apparent from Fig. 7.) Thus, we conclude that the most dangerous stage for buckling during compression of the two surfaces toward each other to engage attachment is intermediate with $0<\Delta<L$.

An analysis of the problem (see the Appendix) shows that buckling becomes possible when

$$\beta=\cos\left[\sqrt{\frac{384\mu paL^2\Delta}{\pi ED^4}}\left(1-\frac{\Delta}{L}\right)\right]-\frac{\Delta}{2L}\sqrt{\frac{384\mu paL^2\Delta}{\pi ED^4}}\sin\left[\sqrt{\frac{384\mu paL^2\Delta}{\pi ED^4}}\left(1-\frac{\Delta}{L}\right)\right]=0 \tag{11}$$

Furthermore, $\beta>0$ precludes buckling. Thus, instability can be avoided if $\beta(\Delta/L)>0$ for all $\Delta/L$ within $0\leq\Delta/L\leq1$. Since $\beta(\Delta/L)$ is a smooth, continuous function of $\Delta/L$ in the given range, with $\beta(0)=\beta(1)=1$, it suffices to find the minima of $\beta(\Delta/L)$ in the range $0\leq\Delta/L\leq1$ and to set to zero the lowest minimum value of $\beta$ in the given range for $\Delta/L$. This provides the constraint that ensures that buckling is avoided upon engagement of the two surfaces by compression.

Numerical results (Appendix) show that the lowest minimum for $\beta$, simultaneous with it being zero, occurs at $\Delta/L=0.5305$ and that this takes place when $\mu(1-\nu^2)p^2L^3/(E^2D^3)=0.0594$. The interpretation for this result is that when the latter condition is satisfied, no buckling will occur until $\Delta/L=0.5305$, at which stage the lowest order mode becomes possible. If buckling is avoided at this stage by use of parametric values lower than those that trigger instability, then no buckling will take place later, even as the compressive load increases. It follows that the buckling condition can be avoided completely as long as

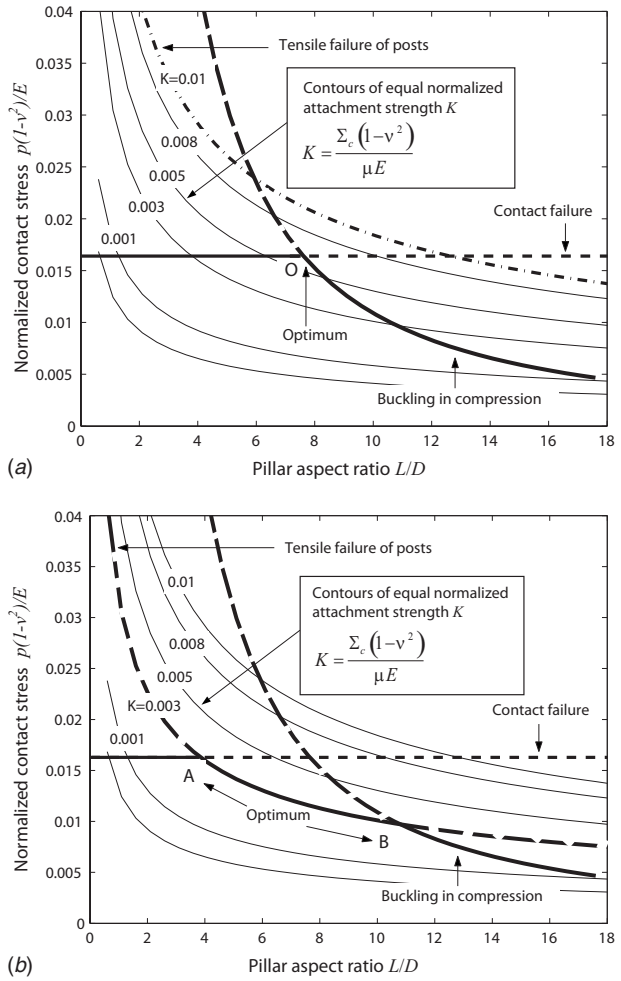



**Fig. 6 Map of attachment strength of the fibrillar system with failure limits: i.e., normalized contact stress $p(1-\nu^2)/E$ versus aspect ratio $L/D$ of the pillar. The curved contours are lines of equal normalized attachment strengths $K=\Sigma_c(1-\nu^2)/\mu E$. (*a*) The case of sufficient tensile strength (corresponding to $K=0.01$) where buckling upon engagement of the attachment system and compressive failure at the contacts together determine the optimum. (*b*) The case where the tensile strength of the pillars determines the optimum attachment strength ($K=0.003$).**

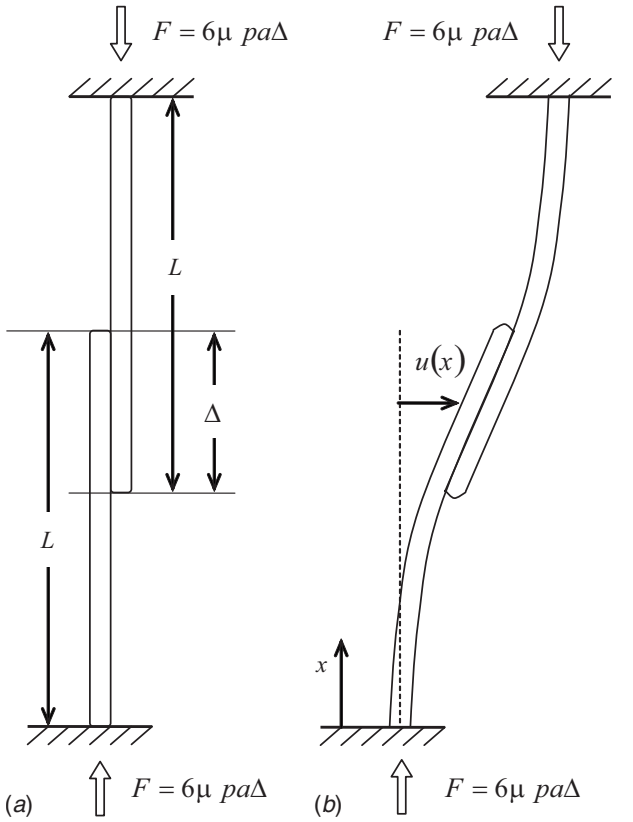



**Fig. 7 Two pillars sliding against each other to depict the unit problem for buckling during engagement of the attachment system by compression: (*a*) The pillars sliding past each other while straight. (*b*) A buckled pair of pillars.**

$$\frac{\mu(1-\nu^2)p^2L^3}{E^2D^3} < 0.0594 \tag{12}$$

Equation (12) provides another design constraint that must be fulfilled in a successful attachment system. Note that if failure in shear at the contact between pillars occurs so that the frictional stress is limited to $\tau_c$, then the product $\mu p^2$ in Eq. (12) should be replaced by $\tau_c p$.

The three failure constraints have been plotted in Fig. 6 to show the performance limits for the attachment system. This has been done for two different situations. In Fig. 6(*a*), a high value for the tensile strength (corresponding to $K=0.01$) has been assumed; the optimal design occurs at point O in the diagram, where the condition for contact failure and compressive buckling are simultaneously met. For a lower tensile strength ($K=0.003$, Fig. 6(*b*)), the tensile failure of the pillars will set in when the attachment system is pulled apart after having been compressed to the state represented by point O. Therefore, the optimal designs are now found along the locus of points AB. In both cases it has been assumed that failure in shear has not occurred at the contact, and the friction stress $\mu p$ can be sustained.

## 3 Optimum Attachment Strength: Limits and Examples

The maps shown in Fig. 6 suggest a straightforward method for identifying the optimal strength for a given material used to construct the fibrillar attachment system. In Figs. 6(*a*) and 6(*b*) it can be seen that the constraint involving compressive failure at the contact can be active in both cases, giving point O in Fig. 6(*a*) and point A in Fig. 6(*b*). Thus, to find the optimal performance, we can assume that $p=p_c$ and investigate whether tensile failure or buckling is the other active mechanism at the relevant optimal condition. The parameters of the active mechanisms of constraint can then be used to compute the optimal attachment strength. Note that this procedure is valid also when shear failure or yielding has occurred at the contact, but with the relevant formulas modified to account for this situation.

Consider first the case where buckling is the active constraint, and thus Fig. 6(*a*) characterizes the optimum. Buckling at the optimum point O (Eq. (12)) occurs at

$$\frac{\mu(1-\nu^2)p_c^2L^3}{E^2D^3} = 0.0594 \tag{13}$$

so that the aspect ratio of the pillars at the optimal condition is given by

$$\frac{L^3}{D^3} = 0.0594\frac{E^2}{\mu(1-\nu^2)p_c^2} \tag{14}$$

This result can be substituted into Eq. (7) to determine the optimum attachment strength if buckling and contact failure are the active constraints at the optimum. In the case of purely elastic response at the contact, so that the friction stress is equal to $\mu p$, the optimal attachment strength is then

$$\Sigma_c^{\text{buckling}} = 1.15\frac{\mu^{2/3}p_c^{4/3}(1-\nu^2)^{2/3}}{E^{1/3}} \tag{15}$$

If failure in shear has limited the friction stress to $\tau_c$, the group of terms $\mu p_c^2$ is replaced by $\tau_c p_c$ throughout Eqs. (13)–(15). The attachment strength in this situation is thus given by

$$\Sigma_c^{\text{buckling}} = 1.15\frac{\tau_c^{2/3}p_c^{2/3}(1-\nu^2)^{2/3}}{E^{1/3}} \tag{16}$$

If the tensile strength of the pillars is the controlling mechanism at the optimum point, then the optimal strength is given from Eq. (8) as

$$\Sigma_c^{\text{strength}} = 0.302\sigma_c \tag{17}$$

If $\Sigma_c^{\text{strength}} < \Sigma_c^{\text{buckling}}$, then the optimal design is associated with achieving a stress in the pillars equal to their tensile strength, and the optimum is given by $\Sigma_c = \Sigma_c^{\text{strength}}$. In contrast, if $\Sigma_c^{\text{buckling}} < \Sigma_c^{\text{strength}}$, the optimal design occurs in association with pillar buckling during compression and failure of the contacts due to excessive compressive stress on them. In this situation, $\Sigma_c = \Sigma_c^{\text{buckling}}$. Therefore, the ratio

$$R = \frac{\Sigma_c^{\text{buckling}}}{\Sigma_c^{\text{strength}}} \tag{18}$$

determines the optimal strength, with $R > 1$ showing that $\Sigma_c = \Sigma_c^{\text{strength}}$ and $R < 1$ indicating that $\Sigma_c = \Sigma_c^{\text{buckling}}$. If the contact response is purely elastic with the Coulomb friction,

$$R = 3.81\frac{\mu^{2/3}p_c^{4/3}(1-\nu^2)^{2/3}}{\sigma_c E^{1/3}} \tag{19}$$

whereas

$$R = 3.81\frac{\tau_c^{2/3}p_c^{2/3}(1-\nu^2)^{2/3}}{\sigma_c E^{1/3}} \tag{20}$$

when contact failure has occurred and the friction stress is limited to $\tau_c$.

The following examples show that useful attachment strength levels can be achieved by this frictional mechanism:

*(i) A mild steel system designed for repeated use.* In this case, the contact compression is limited to a value that keeps the stress magnitude under the contact to less than half the yield stress so that fatigue endurance can be ensured [4]. The tensile equivalent stress due to contact pressure alone has a maximum equal to $0.71p$ [2], whereas the maximum value due to frictional shear stress alone is $\sqrt{3}\mu p$. Thus, an upper bound to the total tensile equivalent stress is given by the root mean square of these two values. Assuming a coefficient of static friction for steel of 0.6 [4], the resulting bound on the total tensile equivalent stress is $1.26p$. To ensure that the tensile equivalent stress is kept below half the yield stress in tension, the contact stress is limited to $p_c = 0.4\sigma_y$, where $\sigma_y$ is the yield strength in tension. Thus, the resulting frictional shear stress when $p = p_c$ is $0.24\sigma_y$, below the yield strength in shear, equal to $\sigma_y/\sqrt{3}$ according to the von Mises criterion [4]. Therefore, Eq. (19) is used to obtain $R$, and to avoid fatigue, $\sigma_c$ is limited to $0.5\sigma_y$. With $E = 210$ GPa, $\nu = 0.3$, and $\sigma_y = 210$ MPa, we obtain a value for $R$ of 0.15. Thus buckling and the limitation on the contact compression together constrain the optimal design. As a consequence, the aspect ratio $L/D$ of the pillars at the optimal design as given by Eq. (14) is 88 and the optimal strength, from Eq. (15), is 4.8 MPa. A patch of pillars $14.4 \times 14.4$ mm$^2$ would thus be able to support 1 kN repeatedly without fatigue; such a load is representative of the weight of fully grown humans.

According to Eq. (2), with $p = p_c = 84$ MPa, a ratio $a/D = 0.00046$ is required to achieve the desired friction in the optimal design. Equation (3) then provides a ratio $\delta/D = 0.34 \times 10^{-5}$, so that the design is well within the limitations of infinitesimal strain Hertzian contact theory.

The estimate for attachment strength in this case may be conservative since an overestimate of the maximum tensile equivalent stress has been used in the analysis. A more accurate result could be obtained through finite element analysis. Countering this possibility is the fact that the value used for the coefficient of friction for steel against steel is at the high end of the data, valid for very smooth polished surfaces [2].

*(ii) A high strength steel system designed for a single use.* Since the system is to be used only once, for reliability the response should be elastic. Thus it suffices to keep the compression $p_c$ on the contact to a level that will ensure that all materials under it have a tensile equivalent stress below the yield strength, $\sigma_y$, including the combined effect of contact pressure and shear. Consequently, the limit on $p$ is taken to be $p_c = 0.8\sigma_y$. With the coefficient of friction equal to 0.6 as before, the friction shear stress during pullout will be below $\sigma_y/\sqrt{3}$. Therefore, Eq. (19) is used to obtain $R$. We will consider steel with high tensile yield strength of 1 GPa, with $\sigma_c$ equal to this value. Consequently, $R = 0.32$ so that the limit on the compressive stress and buckling together determine the optimal design, as is the case with the mild steel system considered previously. In this situation, the aspect ratio $L/D$ of the pillars in the optimal design is 19.6. The resulting strength, predicted by Eq. (15), is then 96 MPa. This means that 1 kN can be supported once only by a patch of pillars $3.2 \times 3.2$ mm$^2$. Repeated use will lead to low cycle fatigue [4].

According to Eq. (2), with $p = p_c = 0.8$ GPa, a ratio $a/D = 0.0044$ is required to achieve the desired friction for the optimal design. Equation (3) then provides a ratio $\delta/D = 0.00022$, and infinitesimal strain theory suffices as before. The small interference fits in both this design, and the mild steel one indicates that precision methods may be necessary to implement design and manufacturing of fibrillar attachment systems.

*(iii) A nylon system.* Consider a nylon 6/6 with an elastic modulus $E = 3$ GPa, Poisson's ratio of 0.25, and an ultimate tensile strength of 85 MPa [10]. To avoid excessive deformation under the contact, the design will be limited to an average compressive stress $p = p_c = 70$ MPa. The coefficient of friction will be taken to be 0.1 [10], and the tensile strength of 85 MPa will be taken to determine $\sigma_c$. It will also be assumed that the friction stress $\mu p_c = 7$ MPa is lower than the limiting shear strength $\tau_c$, so that Eq. (19) gives $R$. With such values, $R$ is found to be 0.18, so Eq. (14) determines $L/D$ and Eq. (15) gives the optimal strength. We find that the aspect ratio, $L/D$, for the optimal design is 10.5, and its attachment strength is 4.7 MPa.

The attachment strength level of 4.7 MPa for the case of nylon 6/6, comparable to that obtainable with mild steel, is encouraging and shows that 1 kN may be supported by such a system having dimensions of $14.5 \times 14.5$ mm$^2$. A more frictional polymer with $\mu = 0.5$, but otherwise with similar properties to nylon 6/6, would boost the attachment strength by almost a factor of 3. According to Eq. (2), with $p = p_c = 70$ MPa, a ratio $a/D = 0.028$ is required to achieve the necessary friction level for the optimal design. From Eq. (3), this corresponds to an effective contact strain $\delta/D = 0.0059$, acceptable for the purposes of the infinitesimal Hertzian contact analysis. The parameters in this case and the material involved suggest that economical devices can be readily produced, having quite high attachment strengths, especially when a polymer with a high coefficient of friction is utilized.

*(iv) A ceramic design.* Although it may be too fragile and unreliable to be practical, such a design is interesting to explore the limits of fibrillar attachment systems. Consider an attachment system made from materials largely having the characteristics of alumina, so that $E = 400$ GPa and $\nu = 0.25$. We will assume that a tensile strength of 1 GPa can be achieved reliably, so that $\sigma_c = 1$ GPa. In ceramics, there is typically a variation of strengths from component to component due to the random nature of the size of the inherent flaws that lead to tensile fracture. Excessive variance of the strength can be controlled by careful but expensive

processing of the ceramic during manufacture. Another strategy to enhance the uniformity, and thus the reliability, of the ceramic tensile strength properties is to use alternating material layers in tension and compression due to thermal strain mismatch [11]. We assume that some such strategy has been implemented and thus the tensile strength level of 1 GPa can be relied upon. Compressive strengths of ceramics are typically five to ten times the level of the tensile strength, so we will assume that a compression of 5 GPa can be sustained, and this will be used for the average compressive stress, $p=p_c$, on the contact. A friction coefficient of 0.4 [12] will be assumed, so that the friction stress is 2 GPa, and it will be assumed that this can be relied upon without damage or failure. Thus, Eq. (19) is used to evaluate $R$, which is found to be 2.3. The optimal design is this tensile strength limited, and Eq. (17) gives the attachment strength, which is 300 MPa. A device with dimensions $1.8\times1.8$ mm$^2$ will thus support a load of 1 kN.

Inspection to Fig. 6(*b*) indicates that the lowest pillar aspect ratio for an optimal design occurs when $p=p_c$. The aspect ratio for an optimal design can then be calculated from the third version of the right hand side of Eq. (7) and is found to be 4.3. According to Eq. (2), with $p=p_c=1$ GPa, a ratio $a/D=0.0030$ is required to achieve the necessary friction level for the optimal design. From Eq. (3), this corresponds to an effective contact strain $\delta/D=0.00011$, well within the valid regime for infinitesimal Hertzian contact analysis.

*(v) An example drawn from nature: keratin.* The head stabilization system of the cordulegastrid dragonfly *Anotogaster sieboldii* has tapered fibrils that are up to 50 $\mu$m long, are 10 $\mu$m at their base [1], and are composed of keratin, [13]. On the assumption that the attachment system operates by the mechanism described above, we will analyze it for optimality. (Note that evolution does not optimize but creates devices that are sufficiently good for fitness; nevertheless, it is interesting to explore how good the system could be.) Keratin has a Young's modulus in humid air of 2 GPa and a proportional limit of 40 MPa [14]. It is not clear whether this limit is a yield strength or simply a transition in the elastic stiffness of the material. The properties just given are measured on woolen and hair fibers, but Feughelman [14] stated that horn (also keratin) responds with the same stress-strain curve. Thus, we will assume that the keratin cuticle forming the fibrils of the *Anotogaster sieboldii* head arresting device has the properties just given. We take Poisson's ratio to be 0.4. The coefficient of friction of keratin on keratin is around 0.5 [1,8]. We assert that contact failure will occur at 120 MPa compression (i.e., three times the proportional limit—a reasonable guess since the stress-strain curve is almost horizontal beyond the proportional limit until the material stiffens before final failure [14]). Friction can therefore produce a shear stress of 60 MPa at the contact. The limiting frictional consideration will therefore be the proportional limit in shear, assumed to be $\tau_c=20$ MPa (i.e., half the proportional limit in tension), when $p$ will be equal to 40 MPa. When such values are substituted into Eq. (14) (with $\mu p_c^2$ replaced by $\tau_c p$), we obtain $L/D=7$, similar to the value observed in the natural device. (Note, however, that the natural fibrils are tapered, but we are treating the system as if it has nontapered ones.) Eq. (20) gives $R$ to be around 0.6, so that buckling is the controlling consideration at failure for an optimal device. Equation (16) then gives the strength to be 7 MPa, superior to mild steel designed for repeated use and better than nylon. On the other hand, the keratin case has strength inferior to mild steel when designed for single use and not as good as the ceramic system. A keratin attachment device of optimal design and with area of $1\times1$ mm$^2$ will then require 7 N to detach it. Note that these results were obtained with $p_c$ assumed to be 40 MPa. If, instead, $p_c$ is taken to be its assumed value of 120 MPa, $R$ is around 1.2, and tensile strength is the limiting condition. Eq. (17) then gives the attachment strength to be 12 MPa, corresponding to a detachment force of 12 N for a 1 mm$^2$ device. Gorb et al. [1] reported that Gorb and Popov [8] measured a detachment force of about 2 N for a synthetic device of such a net area simulating frictional fibrillar adhesion in insects. (The synthetic fibril material, a polymer, was not identified.) Thus our optimal estimates are consistent with experimental data, if somewhat high, an understandable result given that the synthetic system was not optimized.

## 4 Conclusion

The present analysis leads to practical conclusions for the optimum design of frictional fibrillar attachment systems. A trade-off between several governing parameters has been identified and conveniently displayed in a map. The main parameters are as follows:

- *Young's modulus.* The value of $E$ should be large to produce a high average pull-off stress between the two surfaces and to suppress buckling during compression. A possible limitation is given by the onset of brittleness in very stiff materials.
- *Friction coefficient.* The friction coefficient should be maximized in the interest of high pull-off forces. However, friction also enters in the compressive load and therefore facilitates buckling.
- *Pillar aspect ratio*. The pull-off force is increased by high aspect ratios which is limited by the onset of buckling.
- *Contact pressure*. For a given modulus, aspect ratio, and coefficient of friction, the contact pressure between the pillars enters quadratically in the attachment strength. It must, however, be limited by degradation of the material, e.g., by plasticity, fracture, or fatigue.
- *Pillar size.* Interestingly, within the Hertzian approach presented here, the adhesion performance does not depend on the absolute size of the pillars.

The case studies assuming mild steel, nylon, ceramic, and keratin pillars have shown that useful attachment strengths can be expected from a frictional mechanism. The maps can be helpful in guiding the optimum design of such attachment systems.

## Acknowledgment

R.McM. and L.M. acknowledge financial support for their research stays at the Max Planck Institute for Metals Research through a Humboldt Award and a Humboldt Fellowship, respectively.

## Appendix: Buckling of the Pillars During Compressive Engagement

The governing equation for the deflection of the pillars depicted in Fig. 7 is

$$EI\frac{d^2u(x)}{dx^2}+Fu(x)=Fu\left(L-\frac{\Delta}{2}\right) \tag{A1}$$

with the definitions of $F$, $u$, and $x$ as depicted in Fig. 7 and with $I$ being the second moment of area of the cross-section of the column. This equation applies to $0\le x\le L-\Delta$, so that the segment is omitted where the columns overlap and are sliding against each other by friction. This segment is considered to be straight because its bending resistance will be higher than elsewhere and the constraints among the contacting columns will inhibit bending. Note also that the moment contribution from the friction along the sides of the columns is omitted from Eq. (A1) due to its smallness. In addition, the second moment of area of a pillar with a circular cross-section is

$$I=\frac{\pi}{64}D^4 \tag{A2}$$

which is assumed to be the case in the results presented.

The solution to Eq. (A1) is

$$u = u\left(L - \frac{\Delta}{2}\right)\left[1 - \cos\left(\sqrt{\frac{F}{EI}}x\right)\right] \tag{A3}$$

in which $u(L-\Delta/2)$ is the kinematic variable representing the amplitude of the buckling. The compatibility condition for the straight segment of the column in $L-\Delta \leq x \leq L-\Delta/2$ is given by

$$u\left(L - \frac{\Delta}{2}\right) = u(L-\Delta) + \frac{\Delta}{2}\frac{du(L-\Delta)}{dx} \tag{A4}$$

which is valid at the inception of buckling and for small deflections from the straight configuration. The evaluation of Eq. (A3) at $x=L-\Delta$ and the insertion of Eq. (A4) into it give the eigenvalue equation for the buckling,

$$\cos\left[\sqrt{\frac{F}{EI}}(L-\Delta)\right] - \frac{\Delta}{2}\sqrt{\frac{F}{EI}}\sin\left[\sqrt{\frac{F}{EI}}(L-\Delta)\right] = 0 \tag{A5}$$

which normally would be solved for the critical values of $F$ leading to buckling. In our case, we replace $P$ in Eq. (4) with $F$ to account for the compressive situation under consideration. This, together with Eq. (A2), provides Eq. (11).

## References


[1] Gorb, S. N., Beutel, R. G., Gorb, E. V., Jiao, Y., Kastner, V., Niederegger, S., Popov, V. L., Scherge, M., Schwartz, U., and Voetsch, W., 2002, "Structural Design and Biomechanics of Friction-Based Releasable Attachment Devices in Insects," Integr. Comp. Biol., **42**, pp. 1127–1139.

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

[3] Timoshenko, S. P., and Gere, J. N., 1961, *Theory of Elastic Stability*, 2nd ed., McGraw-Hill, New York.

[4] McClintock, F. A., and Argon, A. S., 1966, *Mechanical Behavior of Materials*, Addison-Wesley, Reading, MA.

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

[6] Carpick, R. W., Agrait, N., Ogletree, D. F., and Salmeron, M., 1996, "Variation of the Interfacial Shear Strength and Adhesion of a Nanometer-Sized Contact," Langmuir, **12**, pp. 3334–3340.

[7] Kim, K. S., McMeeking, R. M., and Johnson, K. L., 1998, "Adhesion, Slip, Cohesive Zones and Energy Fluxes for Elastic Spheres in Contact," J. Mech. Phys. Solids, **46**, pp. 243–266.

[8] Gorb, S. N., and Popov, V. L., 2002, "Probabilistic Fasteners With Parabolic Elements: Biological Systems, Artificial Model and Theoretical Considerations," Philos. Trans. R. Soc. London, Ser. A, **360**, pp. 211–225.

[9] Chaudhury, M. K., Weaver, T., Hui, C. Y., and Kramer, E. J., 1996, "Adhesive Contact of a Cylindrical Lens and a Flat Sheet," J. Appl. Phys., **80**, pp. 30–37.

[10] www.matweb.com

[11] Rao, M. P., Sanchez-Herencia, A. J., Beltz, G. E., McMeeking, R. M., and Lange, F. F., 1999, "Laminar Ceramics That Exhibit a Threshold Strength," Science, **286**, pp. 102–105.

[12] www.ceramics.nist.gov

[13] Gorb, S. N., 2001, *Attachment Devices of Insect Cuticle*, Kluwer, Dordrecht.

[14] Feughelman, M., 1997, *Mechanical Properties and Structure of Alpha-Keratin Fibres: Wool, Human Hair and Related Fibres*, UNSW, Sydney.

# Accuracy and Convergence Using a Local Interaction Simulation Approach in One, Two, and Three Dimensions

**Shankar Sundararaman**

**Douglas E. Adams**[1]
e-mail: deadams@purdue.edu

Ray W. Herrick Laboratories,
Purdue University,
West Lafayette, IN 47907-2031

*Guided waves are utilized in structural health monitoring for identifying damage in material components. Simulations can be used to examine how elastic waves propagate in components to help in selecting measurement and data analysis techniques. In this work, the influence of grid size and the frequency sample rate on the amplitude accuracy and convergence of local interaction simulation approach/sharp interface model (LISA/SIM) numerical simulations are studied as they pertain to guided wave propagation in structural materials. These issues are studied in all three dimensions, and amplitude distortion with respect to the Courant–Friedrich–Lewy criterion is explored. The LISA/SIM enables accurate and fast modeling of localized and sharp changes in material properties across interfaces associated with heterogeneities and/ or boundaries. The validity of the simulation is demonstrated by comparing simulated responses with experimentally measured data. Additionally, Lamb wave dispersion curves are extracted through the course of the convergence study using a broadband pulse and the two-dimensional fast Fourier transform method.* [DOI: 10.1115/1.2871105]

## 1 Introduction

Wave-propagation-based methods are used for damage detection in structural materials. Elastic wave-propagation-based techniques are attractive because of their capability to propagate large distances, sensitivity to different types of flaws, and the ability to steer waves to reach hidden parts. There are many different techniques available in the literature that exploits elastic waves such as guided waves (including Lamb waves) for damage identification. Apart from structural inspection, elastic waves can be applied to a number of areas including seismology and material characterization. Lamb waves, also referred to as acoustoultrasonic waves, are present in platelike structures with free interfaces and have shown potential for damage identification in metallic and composite structures [1–5]. The structural inspection strategies using Lamb waves need to be carried out based on an understanding of the nature of particle displacements and dispersion characteristics of the Lamb mode of interest. Although experimental analysis is used for analyzing the propagation of waves in real structural materials, numerical simulations can provide useful information to help in selecting strategies for experimental measurement and data analysis.

The use of the local interaction simulation approach/sharp interface model (LISA/SIM) method for studying the propagation of waves was outlined in Ref. [6–9]. Agostini et al. [10] used the LISA model to study the propagation of Lamb waves and their interactions with damage in an orthotropic plate. The LISA model is capable of describing local mechanisms in the neighborhood of local heterogeneities in the form of holes or inclusions and allows freedom of interaction between the lattice nodes in surrounding cells. The SIM allows for exact treatment of the response parameters between interface cells and ensures that the LISA model can be smoothly applied to implement ultrasonic pulse propagation in heterogeneous structures [11].

While the LISA/SIM provides a computationally efficient, accurate, and simple method for developing a wave propagation model, there are issues associated with the stability and convergence of the signals especially in the case of finite dimensions and in the presence of heterogeneous interfaces. Some of these problems are associated with pulse distortion if the Courant–Friedrich–Lewy number (also referred to in the literature as the CFL number or simply the Courant number) is not satisfied. These issues are associated with the stability of the solution [12]. Apart from pulse distortion, amplitude distortion is a closely related error that is often not considered and instead simulation results are presented in the literature using normalized amplitudes. Although it is easier to implement the Courant criterion for one-dimensional problems, in two dimensions, this is often not possible. Iordache et al. [12] examined aspects of the pulse distortion problem associated with the inability to satisfy the criterion in the one-dimensional finite difference context. Delsanto et al. [13] studied the effects of convergence and stability on one-dimensional finite difference analysis while Ruffino and Delsanto [11] extended the studies to include the two-dimensional LISA model. Additional issues seen in the two-dimensional and three-dimensional models include artificial reflections from the boundaries of individual grid cells, inadequate spatial samples in any one direction, and pulse distortion due to the use of rectangular grids [11]. The issues of numerical stability and numerical dispersion associated with two-dimensional finite difference grids were also addressed by Cangellaris [14]. This problem is important especially in the context of the use of the LISA model as a test bed for studying guided wave behavior because of the multimodal nature of the signals. Any pulse distortions can lead to erroneous estimates of damage size and location based on remotely measured signatures and can severely limit the use of this simulation technique as a test bed for diagnostic studies of Lamb waves for damage identification.

Convergence and stability studies have been carried out for the bulk of the numerical methods such as finite element (FE), finite difference (FD), boundary element (BE), and to some extent to the one- and two-dimensional LISA models used for wave propagation. Further studies of these effects are needed to validate the LISA model through analytical and/or experimental comparisons

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 19, 2007; final manuscript received October 20, 2007; published online March 9, 2009. Review conducted by Prof. Sridhar Krishnaswamy.

so that the LISA model can be used as a base line for experimental damage identification techniques. Most of the stability and convergence studies are limited to pulse distortion in one and two dimensions with not many studies available for pulse distortion in three dimensions because of the increased complexity. The dispersive nature of some of the wave forms and the coupled nature of the elastic stiffnesses are additional sources of complexity. Amplitude distortion is studied here using numerical simulations instead of the analytical approaches (i.e., based on Taylor's series expansion) adopted by other researchers. Another aspect that necessitates this detailed numerical study is the adoption of rectangular cuboid grids while the original developments by Delsanto et al. [7–9] used square cube grids. Additionally, it is important to study the effects of the size of the specimen. One set of parameters for a specific geometry might not be relevant for another specimen with different thickness and geometrical characteristics.

In this work, numerical simulations in one, two, and three dimensions are used to show the stability and convergence of the LISA model especially in the context of Lamb wave propagation. Certain aspects of the accuracy and reliability of the LISA simulations were earlier studied by Ruffino and Delsanto [11] in the two-dimensional context for generic ultrasonic pulse propagation, and a more detailed discussion of convergence and stability problems in FD methods is provided in Ref. [15]. The work in this paper focuses on the impact of grid dimensions and sampling frequencies in all three dimensions for plates of different geometrical dimensions. The amplitude distortions seen in the signals are also discussed, and suitable remedies to modify the signatures are also suggested. Also, the numerical results are compared with experimental data. In addition to the study of accuracy and convergence of the LISA model, dispersion curves are obtained using the two-dimensional and three-dimensional LISA models and are compared to the analytical results. A two-dimensional Fourier transform, using the procedure illustrated by Alleyne and Cawley [16], is used to extract the wave-number-frequency curves. In this work, all the numerical results are carried out for a plate with free boundaries.

## 2 Guided Waves in Elastic Media

In an infinite homogeneous isotropic elastic medium, only pure modes are known to exist. The wave theory indicates primarily the presence of two wave forms—the longitudinal wave and the shear wave. The elastodynamic equation in 3D can be written as

$$\partial_l(\mathbf{S}_{klmn}w_{m,n}) + F_k = \rho\ddot{w}_k + \chi_k\dot{w}_k \quad k,l,m,n = 1,2,3 \tag{1}$$

where $\mathbf{S}(x_1,x_2,x_3)$ is the stiffness tensor, $w_k(x_1,x_2,x_3,t)$ is the $k$th displacement, $F_k(x_1,x_2,x_3)$ is the body force applied at a local point, $\rho(x_1,x_2,x_3)$ is the material density, and $\chi_k(x_1,x_2,x_3)$ is the material attenuation factor based on a proportional, viscoelastic damper.

In elastic media with finite thicknesses, the interaction of the two primary wave forms with the finite boundaries results in the presence of other wave forms including Rayleigh waves and several modes of Lamb and shear horizontal waves. The coupling of the longitudinal (P) and shear-vertical (SV) waves in plate-like structures under traction-free boundaries results in Lamb waves. The Rayleigh–Lamb equation is used to describe the phase velocity (or wave number) relationship for Lamb waves as a function of frequency for platelike structures and is given by

$$\frac{\tan\sqrt{1-\varsigma^2}\bar{d}}{\tan\sqrt{\xi^2-\varsigma^2}\bar{d}} + \left(\frac{4\varsigma^2\sqrt{1-\varsigma^2}\sqrt{\xi^2-\varsigma^2}}{(2\varsigma^2-1)^2}\right)^{\pm 1} = 0 \tag{2}$$

where $V_L$ is the Lamb wave velocity, $h$ is the thickness, $k_S$ is the $S$-wave wave number, $\varsigma = V_S/V_L$, $\xi = V_S/V_P$, $\bar{d} = k_S h/2$, $V_P$ is the $P$-wave velocity, $V_S$ is the $S$-wave velocity, and the exponent +1 corresponds to symmetric and −1 corresponds to antisymmetric modes. This equation is used to obtain the wave number (Figs 1(*a*) and 1(*b*)) dispersion curves [17]. The dispersion curves highlight the multimodal nature of the Lamb waves. Figure 1 highlights the dependence of this multimodal characteristic on the thickness of the material medium with higher thicknesses having more modes at any given frequency.

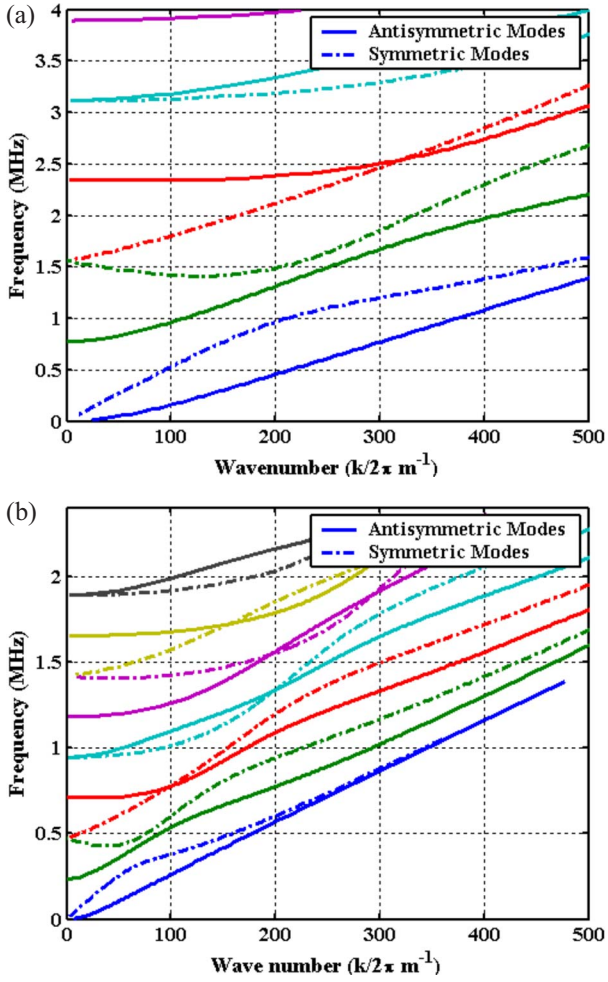



**Fig. 1 Rayleigh–Lamb wave-number-frequency dispersion relations for (*a*) 2 mm and (*b*) 6.6 mm aluminum plates with an elastic modulus of 70 GPa, density of 2700 kg/m$^3$, and Poisson's ratio of 0.334**

## 3 LISA for Modeling Ultrasonic Wave Propagation

The LISA/SIM in one, two, and three dimensions was implemented in MATLAB® and the simulations were carried out on an IBM® workstation. The elastodynamic differential equation (Eq. (1)) is discretized in either its complete form or in a dimensionally reduced form to yield the LISA difference relations. Most of these models are used to provide a qualitative idea of the effects of changes in stiffness and/or density within a structural medium; however, amplitude distortions need to be minimized in order to ensure that the signal features appear at the anticipated time instants without signal distortion corrupting these features. The methodology and the results of the discretization procedure for one-, two-, and three-dimensional models are presented in Refs. [7–9]. This work uses rectangular cuboid grids, details of which are presented in the Appendix and in Ref. [6]. The primary advantage of implementing these expressions is that they can be used for finer grid patterns where required (e.g., impedance changes associated with heterogeneous structures) and they can be used to model structures that cannot be discretized uniformly in all dimensions. The details of the expressions used in this work are

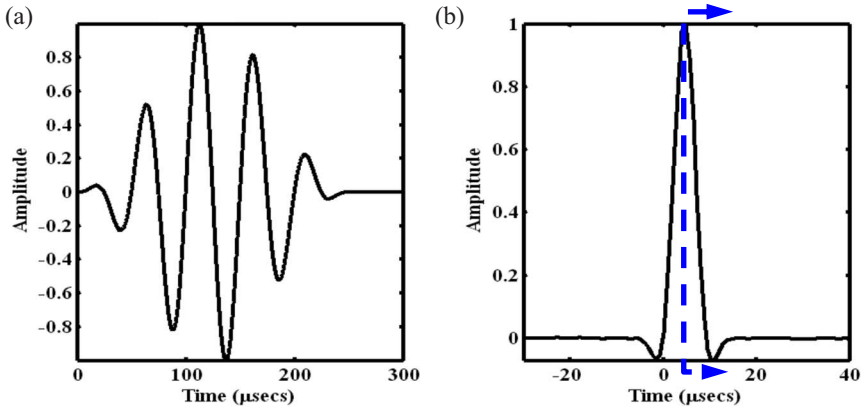



**Fig. 2 Simulation results input signal: (*a*) tone burst and (*b*) sinc pulse (with dashed line indicating typical start of signal used in the simulation)**

provided in the Appendix. The derivation is presented in Refs. [7–9] and is not repeated here.

## 4 Amplitude Accuracy of One-Dimensional and Three-Dimensional LISA Signatures

**4.1 Courant Stability Criterion.** The Courant stability criterion is often used as the minimum constraint criterion that explicit iterative difference methods for homogeneous media must meet to ensure the stability of the simulation associated with iterative difference equations. The LISA methodology, though different from the traditional FD method in its formulation for wave propagation in heterogeneous media, is also constrained by this criterion. Note that the iterative equations associated with the LISA are obtained by solving the discrete form of the differential equation, whereas the equations for the FD are obtained by solving the differential equation after discretization. The Courant criterion is essentially a time step refinement criterion and relates the grid spacing at a given time step and the maximum known wave speed present in the medium by

$$\mathrm{CFL} \equiv c_{\max}\Delta t\sqrt{\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}+\frac{1}{\Delta z^2}} \leq 1 \tag{3}$$

where $c_{\max}$ is the maximum wave speed (usually the longitudinal wave speed), $\Delta t$ is the time step, and $\Delta x$, $\Delta y$, and $\Delta z$ are the spatial steps in the three Cartesian coordinates $x$, $y$, and $z$ [18]. The criterion ensures that the proper choice of sampling frequency is made for a given grid spacing. The determination of the maximum wave speed is often a cause for concern because the wave speeds tend to decrease with smaller spatial steps. This decrease in speed occurs because larger grids tend to be stiffer, resulting in faster wave speeds than predicted by the material properties of the structure of interest. In one-dimensional models, this parameter tends to be somewhat straightforward; however, in higher dimensions, the CFL becomes much more complicated because of the presence of coupling terms (e.g., coupling stiffnesses).

Apart from the coarse spacing of grids, the other parameter that affects the wave speeds is the number of spatial samples per wavelength. To a large extent, grid dispersion can be mitigated by adequately sampling the shortest wavelength, thereby minimizing the error in the wave velocity. For instance, Harker [19] used the estimate of the theoretical percentage error of the FD method for a wave velocity $c$ propagating in the $x$ direction (i.e., the one-dimensional case),

$$c_\varepsilon = 100\left\{\frac{N_\lambda}{\pi}\frac{\Delta x}{c\Delta t}\sin^{-1}\left(\frac{c\Delta t}{\Delta x}\sin\left(\frac{\pi}{N_\lambda}\right)\right)-1\right\} \tag{4}$$

where $N_\lambda$ is the number of nodes at the excitation wavelength to provide a minimum constraint on the number of grid steps per wavelength [20]. Equation (4) is, however, seldom used. In fact, the general criterion for spatial grid spacing in the literature appears to be at least eight grids per the minimum wavelength present in the discretization [20,21]. Alleyne [21] also recommends a grid spacing upper limit of 20 per minimum wavelength to avoid computational issues (e.g., round-off error) associated with very fine grids. These computational issues are either in the form of excessively long run times or computational round-off errors, increasing errors present in the discretization medium. This result is relevant to second order iterative difference relations. For higher order iterative difference methods, this requirement can be relaxed.

The stability of the signatures was initially investigated for one- and three-dimensional models. Apart from this, the selection of Courant numbers much less than 1 was also investigated. In a majority of the numerical simulations, a narrowband Hanning window modulated sinusoidal tone-burst signal shown in Fig. 2(*a*) was used for studying the response patterns at a single frequency. In addition to the tone burst, pulse signatures such as the truncated sinc pulse (Fig. 2(*b*)) and rectangular pulse signatures were used to impart a broadband input pulse (in the frequency domain). A schematic setup of the plate structure used in the tests is shown in Fig 3.

**4.2 Amplitude Accuracy of One-Dimensional LISA Signatures.** A one-dimensional segment of length 150 mm was first investigated, and the signature variations associated with different choices of sampling frequency and spatial spacing were studied. The elastic modulus of the segment was assigned to be 104 GPa and the density 2800 $\mathrm{kg/m^3}$. Figure 4 indicates that the change in sampling frequency and spatial spacing affects the amplitude of the signature. The ratios of the test to a baseline wave speed for the different grid combinations (in order of sampling frequency and spatial spacing: (a) 5.24 MHz, 1 mm; (b) 10 MHz, 1 mm; (c) 10.375 MHz, 0.5 mm; and (d) 20.75 MHz, 0.5 mm) were found to be 1, 0.9875, 0.9825, and 0.9765 where the baseline wave speed corresponded to dataset (a). Additionally, the ratios of the time delay of arrival of the peak signatures with respect to the dataset (a) were found to be 1, 0.985, 0.9899, and 0.9838. The result shows that the wave speed slows down when the spatial and temporal spacing is decreased (i.e., if the sampling frequency is increased) and reinforces the argument presented in Sec. 3 that coarse grids are stiffer than fine grids. Furthermore, the results from the one-dimensional study indicate that the amplitude of the signal is affected by the CFL number. Previously, authors had noted that while there were significant differences between CFL numbers of 1 and 0.99, there was not as significant a difference between CFL numbers 0.99 and 0.1 [12]. However, while the shape of the signature is preserved and there is not a discernible

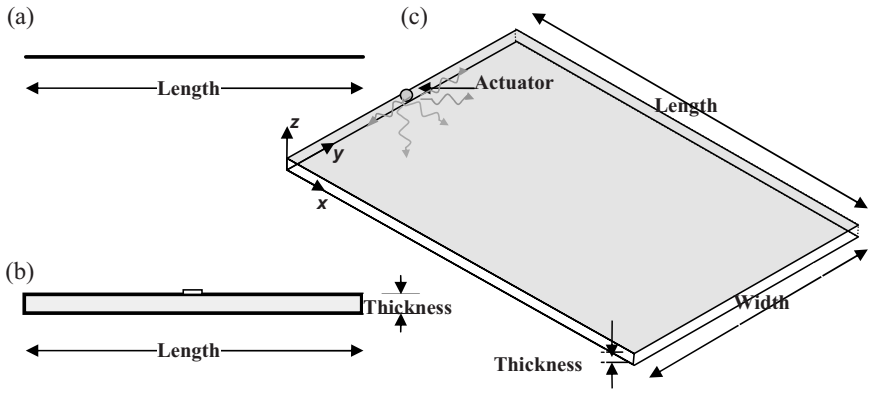



**Fig. 3 Schematic setup showing typical plate model with actuator for the (*a*) one-, (*b*) two-, and (*c*) three-dimensional models**

shape difference at least for the first wave packet studied in Fig. 4 when using CFL numbers smaller than 1, the result shows that the amplitude is influenced by the CFL number.

**4.3 Amplitude Accuracy of Three-Dimensional LISA Signatures.** A three-dimensional aluminum plate (material properties: elastic modulus of 70 GPa, density of 2700 $kg/m^3$, and Poisson's ratio of 0.33) of dimensions $300\times300\times2$ $mm^3$ was next investigated, and the signal variations due to choices of sampling frequency and spatial spacing were studied. The narrowband tone-burst signal illustrated in Fig. 2(*a*) was used as the input signal. The input was emitted from the bottom surface of the center of the plate (150,150,0) mm and the responses were measured at (150,200,0) mm. The change in grid spacing brought out noticeable changes in the shape of the wave form. This shape change as seen in Fig. 5 is independent of the sampling frequency and the wave pattern is seen to slow down with decreasing sampling frequency. The ratios of the amplitudes of the primary wave packet for all cases with respect to the base line wave packet (i.e., Case 1) and the relative delay in arrival of the peak signal amplitude in Fig. 5 with respect to the base line (Case 1) are given in Table 1. Again, this result reinforces the relationship between the CFL number, the amplitude and frequency distortions observed for the one-dimensional case, and the correlation between wave speed and signal amplitude. This result also indicates that the amplitude of the signal in three dimensions is also dependent on the spatial sampling, and the use of coarse grids result in artificially high maximum wave speeds.

**4.4 Trends in Amplitude Distortion.** In most studies, computational issues are centered on ensuring a minimum number of grids for the minimum wavelength in the simulation. Another ef-

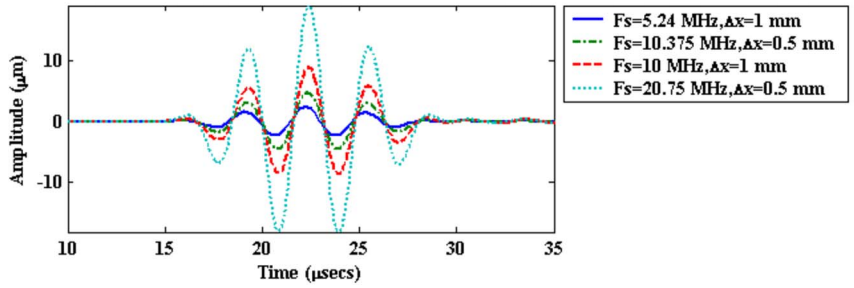



**Fig. 4 Amplitude variation as a function of sampling frequency and spatial sampling for the one-dimensional model**

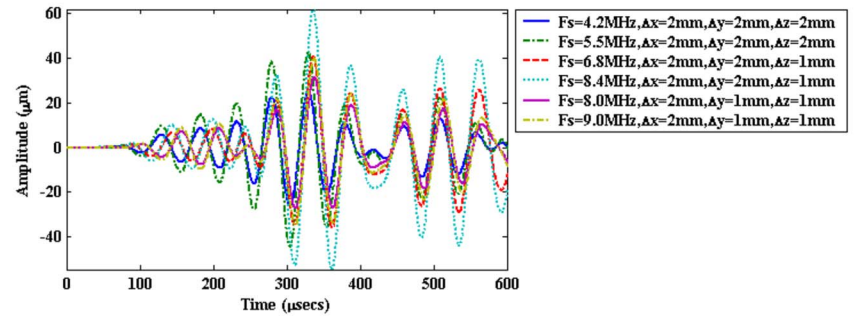



**Fig. 5 Amplitude variation as a function of sampling frequency and spatial sampling for a three-dimensional plate model**

**Table 1 Comparison of results obtained using the three-dimensional plate model**

| Case | $\Delta x$ (mm) | $\Delta y$ (mm) | $\Delta z$ (mm) | $F_S$ (MHz) | CFL | Amplitude ratio | Relative time of flight ($\mu$s) |
|---|---|---|---|---|---|---|---|
| (a) | 2 | 2 | 2 | 4.2 | 0.995 | 1 (base line) | 0 (base line) |
| (b) | 2 | 2 | 2 | 5.5 | 0.76 | 1.72 | 0.057 |
| (c) | 2 | 2 | 1 | 6.8 | 0.992 | 1.64 | 6.71 |
| (d) | 2 | 2 | 1 | 8.4 | 0.803 | 2.51 | 6.78 |
| (e) | 2 | 1 | 1 | 8.0 | 0.9935 | 1.27 | 7.99 |
| (f) | 2 | 1 | 1 | 9.0 | 0.8831 | 1.61 | 7.91 |

fect mentioned in the literature is the need for adequate spatial sampling in every dimension. In the thickness dimension reported in this analysis, the number of grid nodes did not exceed 3. According to Alleyne [21], at least five samples are needed in each dimension of the plate structure to accurately model the propagation of elastic waves. The modeling results presented for the two and three grid cases describe the primary and at most first two secondary wave packets. While the dispersion characteristic for each wave packet might not match the actual dispersion characteristic, adequate spatial sampling in the in-plane dimensions ensures that the reflections from the boundaries reach the sensor site at approximately the same instant as the expected times of arrival. With increasing number of sample points, the lack of adequate spatial sampling in the thickness direction grows and after some point significantly diverges from the real solution. As long as only the first two or three wave packets are considered for further analysis, the simulations would be valid even for lower numbers of grid samples in each dimension.

In the literature, researchers have primarily been concerned with the distortion of the input signal and refer to energy spreading because the numerical model does not meet the Courant limit. The results obtained in this section indicate the dependence of the amplitude on the same factors that influence the Courant criterion. A relationship that relates the new amplitude of the signal and the known base line amplitude estimated at a specified spatial and temporal sampling can be written as

$$A_{\text{new}} = \frac{A_{\text{old}}}{\left[\frac{F_S^{(\text{old})}}{F_S^{(\text{new})}}\right]^2 \frac{f(\Delta x, \Delta y, \Delta z)_{\text{old}}}{f(\Delta x, \Delta y, \Delta z)_{\text{new}}}} \tag{5}$$

where the subscript and superscript "old" is used to indicate the current dataset and "new" is used to indicate the new dataset, $A$ is the amplitude of the signal, $F_S$ is the sampling frequency of the signal, $f(\Delta x, \Delta y, \Delta z)$ is a function of the spatial sampling parameters (in one dimensions, $f(\Delta x, \Delta y, \Delta z) = (\Delta x) c_{\max}$, in two and three dimensions, this term is much more complicated because of the presence of coupling terms), and $c_{\max}$ is the maximum wave speed present in the medium at the wave-number-frequency combination of interest. The amplitude for fine grids and/or driving point displacements in two and three dimensions can be obtained using $f(\Delta x, \Delta y, \Delta z) = \Pi(\Delta x, \Delta y, \Delta z, c_{\max})$, where $\Pi$ indicates a product of the terms and $\Delta z = 1$ in the two-dimensional case.

The maximum wave speed as noted throughout this section is dependent on the material properties of the medium, wave form signal frequency, grid spacing, and to a much lesser extent on the sampling frequency. The unresolved issue is the exactness of the maximum wave speed present in the medium, and this aspect is further complicated by the presence of multiple highly dispersive Lamb modes at any given frequency. The exact amplitude of the displacements cannot be determined using the present numerical model without using known parameters regarding the displacement parameter as obtained from experimental/analytical developments and the minimum wavelength of the signal of interest. These assessments suggest that the two- and three-dimensional models can only provide a qualitative comparison of the signatures. A factor not considered in this study is the total number of time points simulated, which may significantly affect the results obtained. In summary, the expression in Eq. (5) can be used to compensate for the precise temporal and spatial parameters during or after the completion of the iterative simulation in order to mitigate signal shape distortion, signal amplitude distortion, and incorrect estimates of the wave speed of the propagating wave form. Furthermore, the expression is useful because it provides an alternate normalizing method for data obtained from FD or LISA simulations than unity normalization. Finally, the expression outlines the difficulty in using irregular grids to describe guided wave propagation in two- and three-dimensional grids because the different grid elements are known to have different associated maximum wave speeds, leading to mismatch at boundary points.

## 5 Convergence Study

In the previous section, it was seen that decreasing the spatial grid dimensions resulted in slowing of the wave form propagating through the specimen. The large discrepancies seen in the reported delays indicate that convergence was not achieved. Because it was not possible to implement rigorous convergence measures in the serial computing scenario, the convergence study was conducted in two dimensions to limit computation time. In all cases, an aluminum plate structure with an elastic modulus of 70 GPa, density of 2700 kg/m$^3$, and Poisson's ratio of 0.334 was modeled.

**5.1 Convergence Study at a Single Frequency.** First, a narrowband signal was used to excite the center of the length dimension in the plate section shown in Fig. 3(*b*) using the 20 kHz tone-burst signal described in Fig. 2(*a*). The objective of this study was to obtain information about the effect of grid size and study convergence of the LISA model. As part of this study, three different plate sections were considered: (i) $600 \times 2$ mm$^2$, (ii) $600 \times 4$ mm$^2$, and (iii) $150 \times 2$ mm$^2$. The input was primarily directed along the length of the plate ($u$ displacement), and one study was conducted with the input signal directed along the thickness dimension ($w$ displacement). Another study also attempted to quantify the effects of a different actuator location. The sampling frequency was varied throughout the study in order to ensure stable results. A common sampling frequency was not chosen so that total computational time could be reduced. It was noted in the previous section that the sampling frequency has little influence on the time of arrival. All signatures were obtained at a distance of 50 mm from the location of the excitation, and the time of flight was computed at the peak response of the first wave packet. The analysis was repeated for grid combinations of 1 mm, 0.5 mm, 0.25 mm, 0.125 mm, and 0.0625 mm in both directions, resulting in a maximum of 25 combinations of rectangular grid patterns. The sampling frequency in each case was varied according to the minimum grid dimension as (a) 10 MHz for a minimum grid dimension of 1 mm, (b) 20 MHz for the 0.5 mm grid, (c) 40 MHz for the 0.25 mm grid, (d) 80 MHz for the 0.125 mm grid, and (e) 160 MHz for the 0.0625 mm grid. The Courant number for all cases was chosen such that stable solutions were obtained.

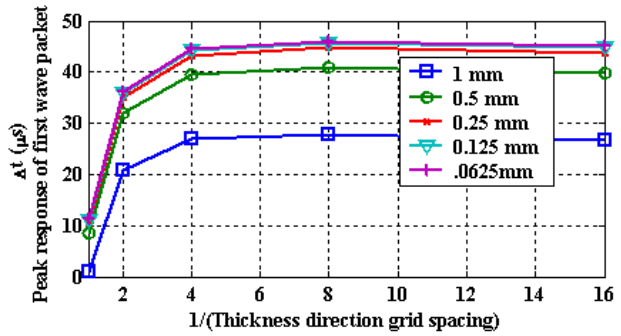



**Fig. 6 Convergence plot for varying grid sizes for a plate of dimensions 600×2 mm²; center frequency of input signal: 20 kHz; sensor location (0,100) mm**

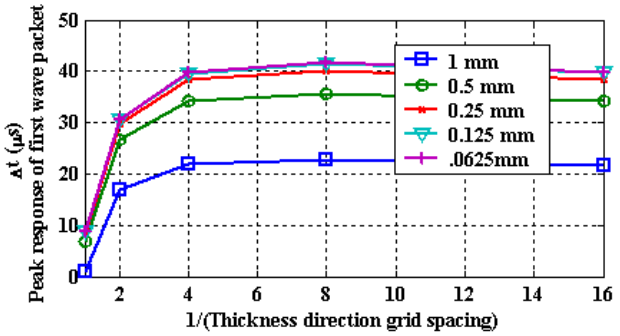



**Fig. 8 Convergence plot for varying grid sizes for a plate of dimensions 150×2 mm²; center frequency of input signal: 20 kHz; sensor location (0,100) mm**

The time of flight at each of these 25 combinations was then subtracted from the coarse grid of $1\times1$ mm$^2$, and the results were reported in Figs. 6–8 after the addition of 1 $\mu$s.

The results from the $600\times2$ mm$^2$ specimen (Fig. 6) indicated that the difference in the time of flight of arrival of the first wave packet for the $1\times1$ mm$^2$ grid and the $0.0625\times0.0625$ mm$^2$ grid was as much as 45 $\mu$s, approximately 9/10th of the time period of the excitation signal. A noticeable trend is that a reasonable match is obtained for grid sizes of $0.25\times0.25$ mm$^2$. However, the grids of dimension 1/16th mm displayed a slight trend reversal, indicating that the grid has certain size restrictions. Another aspect seen in Fig. 6 was that the reduction in the thickness dimension grids led to larger time of arrival differences than the reduction in the length dimension grids. Similar trends were also observed in Figs. 7 and 8, which were generated for a thicker specimen ($600\times4$ mm$^2$) and a shorter specimen ($150\times2$ mm$^2$). The difference in the time of flight of arrival of the first wave packet for the $1\times1$ mm$^2$ grid and the $0.0625\times0.0625$ mm$^2$ grid was about 12 $\mu$s for the $600\times4$ mm$^2$ specimen (Fig. 7) and was 42 $\mu$s for the $150\times2$ mm$^2$ specimen (Fig. 8). The results from Fig. 7 indicated that the higher number of spatial samples present in the thicker specimen ensured that the discrepancy between the coarse and the fine grid was not as high as in the case of the thinner specimen. The results from the shorter specimen (Fig. 8) are more surprising and indicate that the longer dimensions of that specimen have a negative impact on the wave speeds. This result indicates that the number of spatial grids also have lower and upper limits in order to get good convergence. The need for adequate spatial sampling is reinforced by the results of the comparison between the times of arrival of the first wave packet for the $0.5\times0.5$ mm$^2$ grid and the $0.0625\times0.0625$ mm$^2$ grid for the $600\times2$ mm$^2$ specimen. This difference is approximately 12 $\mu$s. Numerical results conducted on the plate structure also confirmed that the time of arrival was found to be independent of both input signal direction and input signal position.

As noted by Smith [22] among others, stability and convergence are different but related issues. An alternate form of convergence was examined by comparing the normalized peak amplitude of the first wave packet. The normalization was carried out based on the sampling frequency and spatial sampling as described in Eq. (3) and as noted in the $y$-axis caption of Fig. 9. The base line grid was chosen as the coarse grid of size $1\times1$ mm$^2$, and the sensor measurements were obtained 300 mm from the actuator. Convergence was examined by monitoring the amplitude distortions at the sensor location. The actuator was placed at the center of the specimen, as shown in Fig. 3(*b*). The result also suggests that the sensor amplitude is influenced by the number of spatial samples in each dimension.

**5.2 Convergence Study at Multiple Frequencies.** The convergence study was then extended to multiple frequencies by inducing a broadband excitation in the form of a narrow pulse. First, a rectangular pulse with a 2 MHz signal bandwidth was imparted at the center of a $609.6\times609.6\times2$ mm$^3$ plate (at (304.8,304.8,0) mm) with grid dimensions of approximately $1\times1\times1$ mm$^3$ using the three-dimensional LISA model for a homogeneous, isotropic aluminum plate. As before, an aluminum plate structure with an elastic modulus of 70 GPa, density of 2700 kg/m$^3$, and Poisson's ratio of 0.334 was modeled. The data were collected at 609 nodal locations approximately 1 mm apart along the center of the plate structure, and the first 608 nodal data were used for further analysis. The two-dimensional Fourier transform is given by [16]

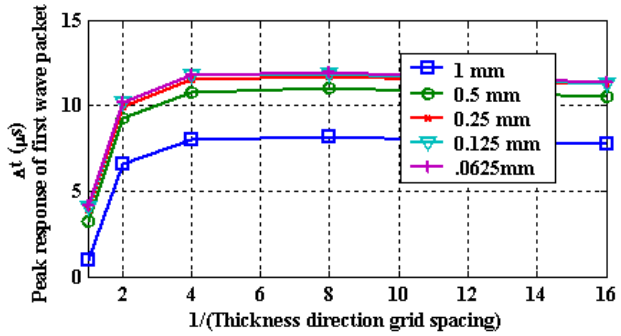



**Fig. 7 Convergence plot for varying grid sizes for a plate of dimensions 600×4 mm²; center frequency of input signal: 20 kHz; sensor location (0,100) mm**

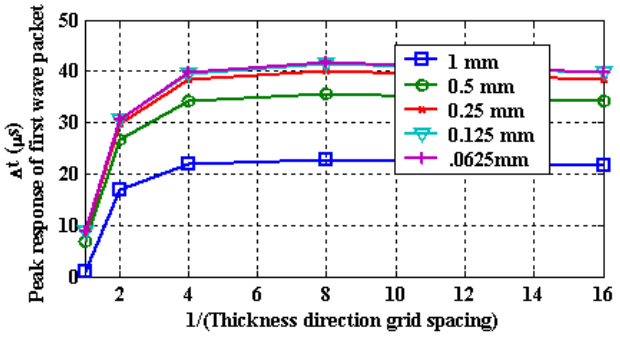
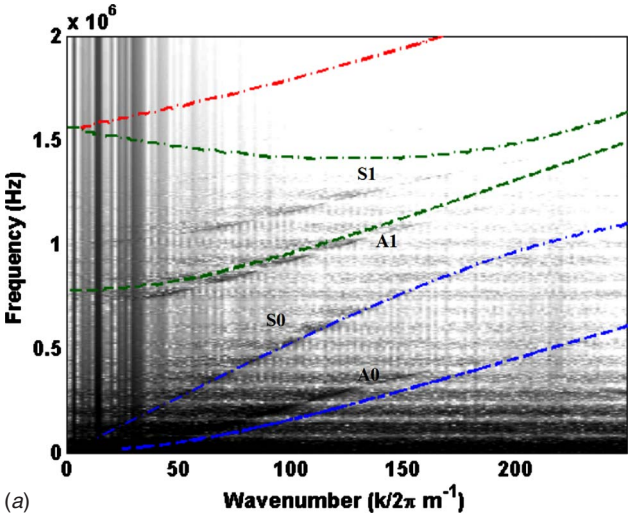
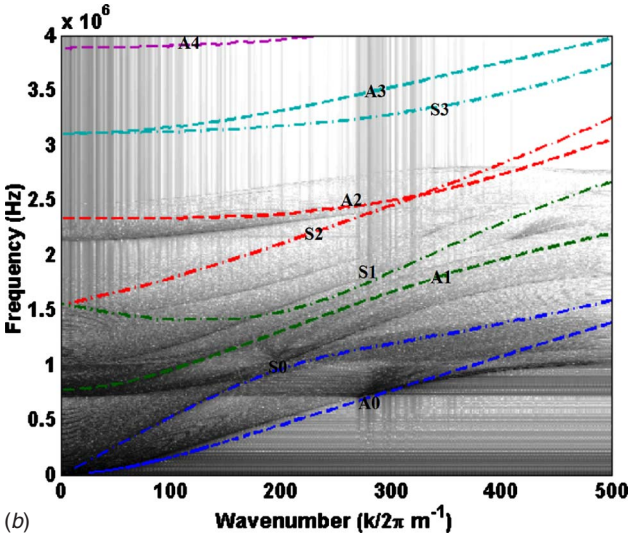



**Fig. 9 Convergence plot for varying grid sizes for a plate of dimensions 600×2 mm²; center frequency of input signal: 20 kHz; sensor location (0,0) mm, actuator location (0,300) mm**




(*a*)




(*b*)

**Fig. 10 Dispersion curve simulation results for a 609.6 ×609.6×2 mm³ aluminum plate with a grid spacing of 1 mm in all three dimensions extracted from excitation using (*a*) a rectangular pulse and (*b*) a half-sinc pulse**

$$U(k,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x,t) e^{-i(kx+\omega t)} dx dt \tag{6}$$

where $u(x,t)$ is the displacement at a spatial location $x$ and temporal point $t$, $\omega = 2\pi f$ is the angular frequency, and $k$ is the wave number. The implementation of the algorithm on the computer was carried out by first transforming to the frequency domain using the fast Fourier transform of the data rows in the spatial-time data matrix. Next, the spatial-frequency data thus obtained were transformed to the wave-number-frequency domain by computing the fast Fourier transform along the data columns. This double Fourier transform in the two different domains is referred to as a two-dimensional Fourier transform (or 2D FFT).

The results obtained numerically were then compared with the results (Figs. 10–13) obtained analytically from the Rayleigh–Lamb dispersion curves (Fig. 1). The results indicate a match between the fundamental symmetric ($S0$) and antisymmetric ($A0$) modes at the lower frequency-wave-number combinations and a progressive deviation from the analytical result with increasing frequency and wave number. The deviations indicate that the stiffness of the discretized elastic medium is higher than what it

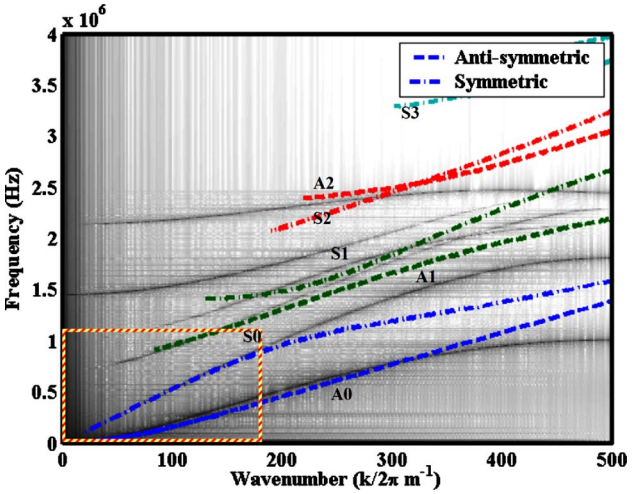



**Fig. 11 Dispersion curve simulation results for a 609.6 ×2 mm² aluminum section with a grid spacing of 1×1 mm². Boxed region indicates typical region in which experimental analysis is carried out using Lamb waves by health monitoring researchers.**

should be and suggests that the coarse geometry of the grid cell leads to stiffer results. This observation is similar to the results obtained for the narrowband excitation. The numerical results also match the analytical results for the second antisymmetric ($A1$) mode as well although the actual frequencies at which the waves appear are underpredicted.

The broadband excitation (rectangular pulse, Fig. 10($a$)) did not excite all the frequencies in the frequency range uniformly, and the frequency response of the rectangular pulse indicated that the input energy decreased with frequency. The sinc pulse shown in Fig. 2($b$) was instead used in a modified form (half the sinc pulse was used as indicated by the segment outlined in Fig. 2($b$)) to obtain a more uniform excitation in the frequency domain, and the result obtained using this excitation is shown in Fig. 10($b$). The dispersion curves obtained using the half-sinc pulse provided better results, both in clearly highlighting the actual curves and in the frequency bandwidth.

Additionally, a detailed convergence study was conducted for this broadband excitation. To facilitate the extraction of the dis-

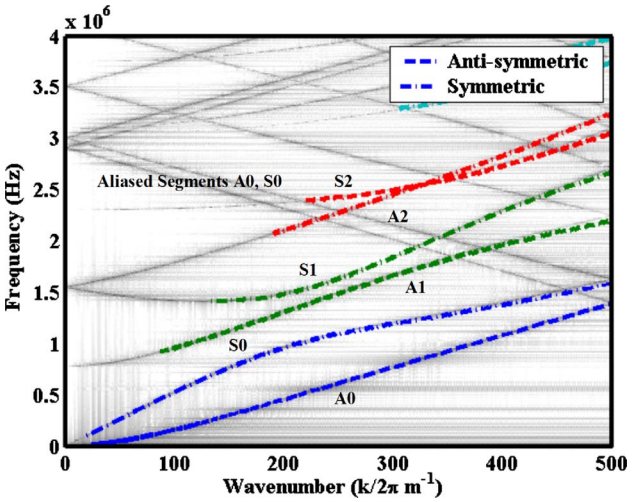



**Fig. 12 Dispersion curve simulation results for a 609.6 ×2 mm² aluminum section with a grid spacing of 1/8 ×1/8 mm²**

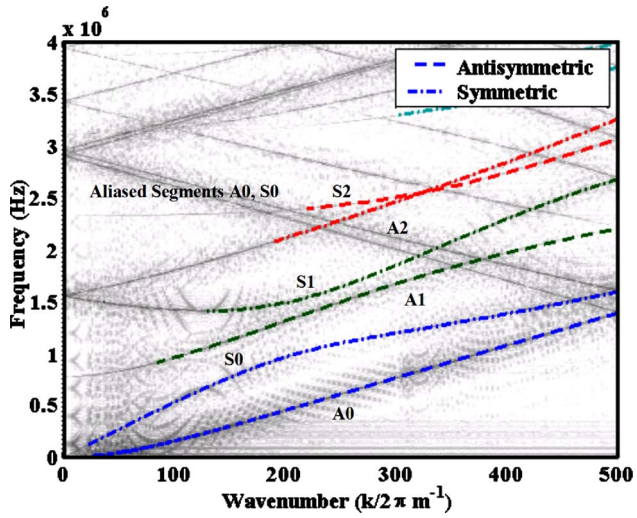



**Fig. 13 Dispersion curve simulation results for a 609.6 ×2 mm² aluminum section with a grid spacing of 1/16 ×1/16 mm²**

persion curves, finer grid spacing was adopted and because of the limitations of serial computing, the simulation study was carried out in two dimensions. A $609.6\times2$ mm$^2$ plate was considered with the same properties as the plate structure described earlier in this section. First, data were extracted from 608 points along a straight line in an approximate $1\times1$ mm$^2$ grid both to compare the results obtained in the three-dimensional model and to set up the comparisons for finer grid sizes. The results at this grid size after the 2D FFT process for extracting the wave number dispersion curves was carried out are shown in Fig. 11. The trends obtained here are similar to the trends observed in the 3D model and also indicate that the 2D model would be a satisfactory test bed for extracting the Lamb wave dispersion curves. The region marked by the striped rectangle in Fig. 11 especially shows a good match between the numerical and the analytical plots and indicates the typical region in which experimental data are acquired using Lamb waves by health monitoring researchers. The use of the sinc pulse also illuminates the response curves uniformly throughout the frequency range of interest.

The results from finer grid sizes are presented in Fig. 12 (grid size of $0.125\times0.125$ mm$^2$) and Fig. 13 (grid size of $0.0625\times0.0625$ mm$^2$). Both grid sizes were considered because the results observed in the narrowband study indicated a trend reversal (i.e., slowing of the wave packet) for finer grids ($0.0625\times0.0625$ mm$^2$). In both cases, the results indicate a very good match between numerical and analytical results, indicating the convergence of the grid at those sizes throughout the domain of interest.

In addition to the Lamb wave dispersion curves, there are lines seen in both figures that emanate from the right edge of the plot. The right edges of Figs. 12 and 13 highlight the limit of the $x$ axis illustrated by the Nyquist wave number (i.e., the spatial equivalent of the Nyquist frequency). The lines that emanate backward and forward from the right edge are spatially aliased segments of the different dispersion curves. It is also seen in Fig. 12 that close to the right extremity (i.e., at high wave numbers), the numerical results appear to diverge from the analytical result because of inadequate spatial sampling. This aspect is not seen in Fig. 13 and is indicative of convergence at this grid dimension over the entire domain.

## 6 Experimental Comparison of Data From Plate Structure

The results from the LISA/SIM numerical simulation at two different sets of grid spacing were then compared to experimental results obtained from an accelerometer to validate the model. A PCB Piezotronics Inc. manufactured 7–8 mm diameter 352C65 accelerometer (100 mV/g sensitivity, 0.0015 m/s$^2$ rms broadband resolution) was used to acquire wave forms excited by a 10 mm diameter PI P-010.00P (1000 V, 129 kHz, 4 g) piezostack actuator. The signal was created by a 33220A Agilent arbitrary wave form generator and the data were acquired using a Tektronix TDS5054B-NVT digital storage oscilloscope 500 MHz, 1 Gsample/s. The Data were sampled at a rate of 5 MHz and was averaged 64 times prior to storage using synchronous averaging at a burst rate of 71 ms. A five wave Hanning windowed sinusoidal signal centered at 20 kHz with a peak-to-peak amplitude of 10 V and described in Fig. 2($a$) was used for both the numerical and experimental investigations. Both transducers in the study were adhesively bonded to the structure through a layer of wax.

In the first set of experimental comparisons, an aluminum plate of dimension $609.6\times609.6\times2$ mm$^3$ was considered for the numerical and experimental investigations. The actuator was placed at (304.8,304.8,0) mm and the sensor was placed at (354.8,304.8,0) mm on the bottom surface of the plate structure. The numerical responses were time shifted (in order to synchronize with the actuation signal) and normalized. The displacements in the out-of-plane dimension were recorded from the numerical simulation, while the actuation was applied in the direction of the width dimension. The plots in Fig. 14 show a comparison of grids of approximate dimensions $2\times2\times2$ mm$^3$ and $2\times2\times1$ mm$^3$. The results indicate that the $2\times2\times1$ mm$^3$ grid provides a better match to the experimental result than the $2\times2\times2$ mm$^3$ grid. Furthermore, the $2\times2\times2$ mm$^3$ grid appears to have a faster wave velocity than the $2\times2\times1$ mm$^3$ grid. In a second set of results, experimental comparisons were made between $1\times1\times1$ mm$^3$ and $1\times1\times0.25$ mm$^3$ grids on an aluminum plate of dimensions

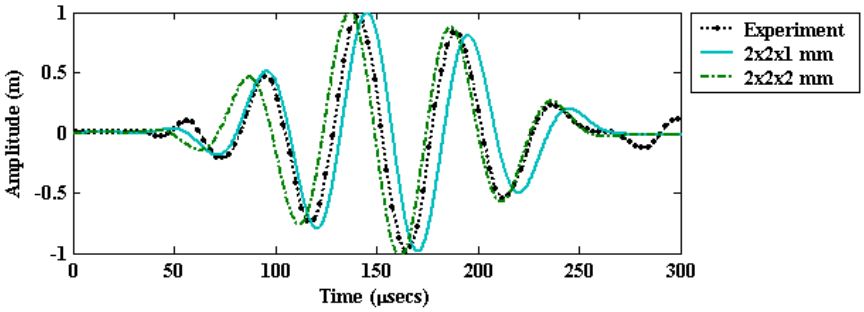



**Fig. 14 Comparison of experimental and numerical results obtained at two grid dimensions for a 609×609×2 mm³ aluminum plate**

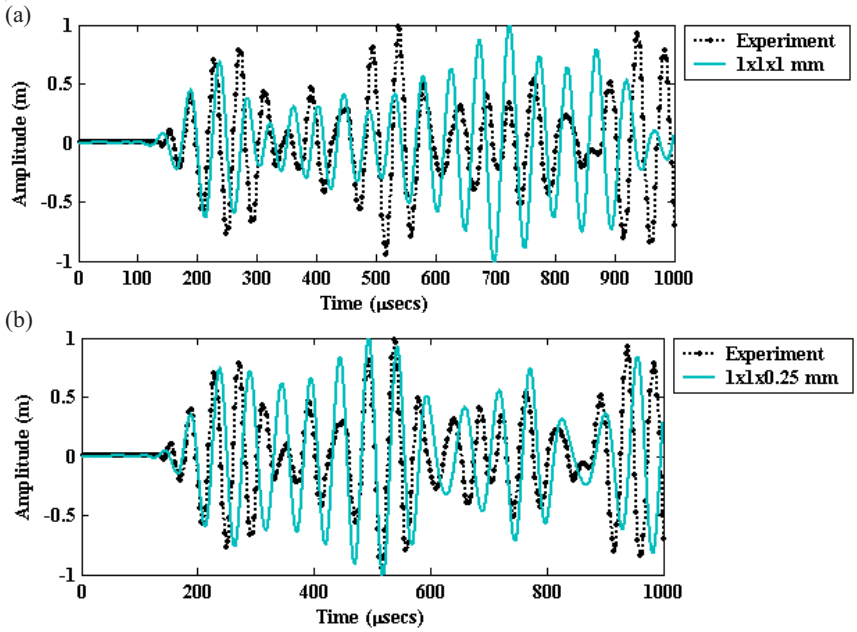



**Fig. 15 Comparison of experimental and numerical results obtained for a 1×1×1 mm³ grid and a 1×1×0.25 mm³ grid for a 300×300×2 mm³ plate at 20 kHz**

$300\times300\times2$ mm$^3$. The actuator in this case was placed at (150,150,0) mm and the sensor was placed at (200,150,0) mm. The results from Figs. 15(*a*) and 15(*b*) indicate that at a finer grid spacing, a better match of the response characteristics at other points in time can be more easily achieved. This match is especially evident in Fig. 15(*b*) where a number of the numerical and experimental wave packets match in the approximate shape and, unlike the results in Fig. 14, the match extends much beyond the first wave packet. However, for the results presented in Fig. 15, the signal was time shifted by 41 $\mu$s to compare the numerical results with the experimental results. This time of arrival mismatch and the match in shape indicates that the wave speed has slowed down uniformly along the grid and that coarser grids can still provide a qualitative match of experimental data. This result provides validation for the numerical model and further verification that the use of finer grids provides a convergence toward the experimental result.

## 7 Conclusions

In this work, results associated with the accuracy and convergence of the Lamb wave simulations using the LISA method are illustrated. The results obtained using this work confirm that the choice of sampling frequency and grid size determines the minimum requirements for a stable solution. A novelty of this work is the analysis of the use of rectangular cuboid grids for modeling Lamb waves in structures. In addition, an expression for the amplitude distortion observed from the numerical simulations has been obtained. Narrowband convergence studies were conducted on three different types of plates and it was found that longer and thinner plates required the use of finer grids than for shorter and/or thicker plates. The issues appear to be the need for adequate number of spatial samples in each dimension and grid instabilities associated with longer plate dimensions. These convergence studies were performed using changes in the peak amplitude of the signal in the form of delays in time of arrival and distortion in the peak amplitude. The overall trends appear to be similar in both cases with the changes in time of arrival, providing a more stable measure of convergence. Studies were also conducted using the 2D FFT to generate Lamb wave dispersion curves. This method, which uses a broadband sinc pulse for interrogation of the discretized medium, is an innovative feature of this work. The results were compared with the analytical forms obtained from the Rayleigh–Lamb equation. Additionally, broadband convergence was illustrated by matching the numerical dispersion curve with the analytical results. Finally, the numerical simulation results at two different grid levels for two different plate specimens were verified by comparing with the corresponding experimental results. The results indicate that a good match was obtained between numerical and experimental wave responses.

## Acknowledgment

The authors gratefully acknowledge Dr. Stuart J. Bolton, Purdue University, for several helpful suggestions and discussions.

## Appendix

### Two-Dimensional LISA Model

The 2D model is derived based on the assumption of translational invariance with respect to the third dimension. The discretization procedure reduces the 2D form of the differential equation to an appropriate iterative difference relationship. It is assumed that the model is discretized in small steps on a rectangular grid pattern with each grid point, in the 2D case, at the intersection of four different cells. There may be at most four different sets of material properties at any interface grid point. The material properties of each cell are denoted by the notation of the grid point at its lower left corner. These properties are averaged using the SIM. It is assumed that the displacements and material properties are continuous within a cell, while interface cells (e.g., cells at a boundary) are treated as discontinuous. The displacement components in two dimensions and in the absence of body forces can be obtained at each point $P(i,j)$ as

$$u^{i,j,t+1} = -\bar{q}_1^2 u^{i,j,t-1} + 2\bar{q}_1 u^{i,j,t} + \frac{\bar{q}_1 \Delta t^2}{4\bar{\rho}} \sum_{a,b=\pm} [-2u^{i,j,t}(\tilde{\eta}_1^2 \tilde{\sigma}_1 + \tilde{\eta}_2^2 \tilde{\mu}_{12}) + 2\tilde{\eta}_1^2 \tilde{\sigma}_1 u(1^a) + 2\tilde{\eta}_2^2 \tilde{\mu}_{12} u(2^b) + ab\tilde{\eta}_1 \tilde{\eta}_2 [(\tilde{\lambda}_{12} + \tilde{\mu}_{12})(v(3_b^a) - v) + (\tilde{\lambda}_{12} - \tilde{\mu}_{12})(v(2^b) - v(1^a))]] \quad \text{(A1)}$$

and

$$v^{i,j,t+1} = -\bar{q}_2^2 v^{i,j,t-1} + 2\bar{q}_2 v^{i,j,t} + \frac{\bar{q}_2 \Delta t^2}{4\bar{\rho}} \sum_{a,b=\pm} [-2v^{i,j,t}(\tilde{\eta}_2^2 \tilde{\sigma}_2 + \tilde{\eta}_1^2 \tilde{\mu}_{12}) + 2\tilde{\sigma}_2 \tilde{\eta}_2^2 v(2^b) + 2\tilde{\mu}_{12} \tilde{\eta}_1^2 v(1^a) + ab\,\tilde{\eta}_1 \tilde{\eta}_2 [(\tilde{\lambda}_{12} + \tilde{\mu}_{12})(u(3_b^a) - u) + (\tilde{\lambda}_{12} - \tilde{\mu}_{12})(u(1^a) - u(2^b))]] \quad \text{(A2)}$$

where $a,b=\pm 1$ and $\bar{a},\bar{b}=-1$ when $a,b=-1$ and $\bar{a},\bar{b}=0$ when $a,b=1$, $\bar{q}_k=(1+(\Delta t/2)\Sigma_{a,b}\tilde{\chi}_k/\bar{\rho})^{-1}$ is the transmission factor, and $\bar{\rho}=(1/4)\Sigma_{a,b}\tilde{\rho}(a,b)$, $(u,v)$ are the displacements, $\tilde{\sigma}_k=\sigma_k(a,b)$, $\tilde{\eta}_1=\eta_1(a,b)=1/\Delta x_{\bar{a}}$, $\tilde{\eta}_2=\eta_2(a,b)=1/\Delta y_{\bar{b}}$, and similar expressions hold for other quantities with a tilde. The expressions obtained here were found to match the expressions listed by Delsanto et al. by setting $\bar{q}_1=\bar{q}_2=1$, $1/\varepsilon=\tilde{\eta}_1=\tilde{\eta}_2$ (i.e., $\varepsilon=\Delta x_{\bar{a}}=\Delta y_{\bar{b}}$).

## Three-Dimensional LISA Model

In a similar fashion to the 2D case, the three-dimensional relations at each point $P(i,j,k)$ were obtained for an irregular rectangular cuboid grid (see Ref. [13] for details of the derivation) as

$$u^{i,j,k,t+1} = -\bar{q}_1^2 u^{i,j,k,t-1} + 2\bar{q}_1 u^{i,j,k,t} + \frac{\bar{q}_1 \Delta t^2}{8\bar{\rho}} \sum_{a,b,c=\pm} [-2u^{i,j,k,t}(\tilde{\eta}_1^2 \tilde{\sigma}_1 + \tilde{\eta}_2^2 \tilde{\mu}_{12} + \tilde{\eta}_3^2 \tilde{\mu}_{13}) + 2\tilde{\eta}_1^2 \tilde{\sigma}_1 u(1^a) + 2\tilde{\eta}_2^2 \tilde{\mu}_{12} u(2^b) + 2\tilde{\eta}_3^2 \tilde{\mu}_{13} u(3^c) + ab\,\tilde{\eta}_1 \tilde{\eta}_2 [(\tilde{\lambda}_{12} + \tilde{\mu}_{12})(v(6_b^a) - v) + (\tilde{\lambda}_{12} - \tilde{\mu}_{12})(v(2^b) - v(1^a))] + ac\,\tilde{\eta}_1 \tilde{\eta}_3 [(\tilde{\lambda}_{13} + \tilde{\mu}_{13})(w(5_c^a) - w) + (\tilde{\lambda}_{13} - \tilde{\mu}_{13})(w(3^c) - w(1^a))]] \quad \text{(A3)}$$

$$v^{i,j,k,t+1} = -\bar{q}_2^2 v^{i,j,k,t-1} + 2\bar{q}_2 v^{i,j,k,t} + \frac{\bar{q}_2 \Delta t^2}{8\bar{\rho}} \sum_{a,b,c=\pm} [-2v^{i,j,k,t}(\tilde{\eta}_2^2 \tilde{\sigma}_2 + \tilde{\eta}_3^2 \tilde{\mu}_{23} + \tilde{\eta}_1^2 \tilde{\mu}_{12}) + 2\tilde{\sigma}_2 \tilde{\eta}_2^2 v(2^b) + 2\tilde{\mu}_{23} \tilde{\eta}_3^2 v(3^c) + 2\tilde{\mu}_{12} \tilde{\eta}_1^2 v(1^a) + bc\,\tilde{\eta}_2 \tilde{\eta}_3 [(\tilde{\lambda}_{23} + \tilde{\mu}_{23})(w(4_c^b) - w) + (\tilde{\lambda}_{23} - \tilde{\mu}_{23})(w(3^c) - w(2^b))] + ab\,\tilde{\eta}_1 \tilde{\eta}_2 [(\tilde{\lambda}_{12} + \tilde{\mu}_{12})(u(6_b^a) - u) + (\tilde{\lambda}_{12} - \tilde{\mu}_{12})(u(1^a) - u(2^b))]] \quad \text{(A4)}$$

and

$$w^{i,j,k,t+1} = -\bar{q}_3^2 w^{i,j,k,t-1} + 2\bar{q}_3 w^{i,j,k,t} + \frac{\bar{q}_3 \Delta t^2}{8\bar{\rho}} \sum_{a,b,c=\pm} [-2w^{i,j,k,t}(\tilde{\eta}_3^2 \tilde{\sigma}_3 + \tilde{\eta}_1^2 \tilde{\mu}_{13} + \tilde{\eta}_2^2 \tilde{\mu}_{23}) + 2\tilde{\eta}_3^2 \tilde{\sigma}_3 w(3^c) + 2\tilde{\eta}_1^2 \tilde{\mu}_{13} w(1^a) + 2\tilde{\eta}_2^2 \tilde{\mu}_{23} w(2^b) + ac\,\tilde{\eta}_1 \tilde{\eta}_3 [(\tilde{\lambda}_{13} + \tilde{\mu}_{13})(u(5_c^a) - u) + (\tilde{\lambda}_{13} - \tilde{\mu}_{13})(u(1^a) - u(3^c))] + bc\,\tilde{\eta}_2 \tilde{\eta}_3 [(\tilde{\lambda}_{23} + \tilde{\mu}_{23})(v(4_c^b) - v) + (\tilde{\lambda}_{23} - \tilde{\mu}_{23})(v(2^b) - v(3^c))]] \quad \text{(A5)}$$

where $a,b,c=\pm 1$ and $\bar{a},\bar{b},\bar{c}=-1$ when $a,b,c=-1$ and $\bar{a},\bar{b},\bar{c}=0$ when $a,b,c=1$, $\bar{q}_k=(1+(\Delta t/2)\Sigma_{a,b,c}\tilde{\chi}_k/\bar{\rho})^{-1}$ is the transmission factor, $\bar{\rho}=1/8\Sigma_{a,b,c}\tilde{\rho}(a,b,c)$, $(u,v,w)$ are the displacements, $\tilde{\sigma}_k=\sigma_k(a,b,c)$, $\tilde{\eta}_1=\eta_1(a,b,c)=1/\Delta x_{\bar{a}}$, $\tilde{\eta}_2=\eta_2(a,b,c)=1/\Delta y_{\bar{b}}$, and $\tilde{\eta}_3=\eta_3(a,b,c)=1/\Delta z_{\bar{c}}$, and similar expressions hold for other quantities with a tilde. The expressions in Eqs. (A3)–(A5) hold only for orthotropic media. The expressions obtained here were found to match the expressions listed in Ref. [13] and were obtained by setting $\bar{q}_1=\bar{q}_2=\bar{q}_3=1$, $1/\varepsilon=\tilde{\eta}_1=\tilde{\eta}_2=\tilde{\eta}_3$ (i.e., $\varepsilon=\Delta x_{\bar{a}}=\Delta y_{\bar{b}}=\Delta z_{\bar{c}}$). In this paper, each grid point is at the intersection of 8 cells and 18 neighboring nodal points. Consequently, each grid point is surrounded by at most eight different physical and/or geometrical properties. The material properties of each cell are marked by the notation of the grid point at its lower left corner. These properties are averaged using the process described by Delsanto et al. with the SIM [9].

## References


[1] Alleyne, D. N., and Cawley, P., 1992, "The Interaction of Lamb Waves With Defects," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **39**(3), pp. 381–397.

[2] Bar-Cohen, Y., Mal, A., and Chang, Z., 1998, "Composite Material Defects Characterization Using Leaky Lamb Wave Dispersion Data," *Proceedings of SPIE, NDE Techniques for Aging Infrastructure and Manufacturing, Conference NDE of Materials and Composites II*, San Antonio, TX, Mar. 31–Apr. 2, Vol. 3396, Paper No. 3396-25.

[3] Tuzzeo, D., and di Scalea, F. L., 2001, "Noncontact Air-Coupled Guided Wave Ultrasonics for Detection of Thinning Defects in Aluminum Plates," Res. Nondestruct. Eval., **13**(2), pp. 61–77.

[4] Bartoli, I., di Scalea, F. L., Fateh, M., and Viola, E., 2005, "Modeling Guided Wave Propagation with Application to the Long-Range Defect Detection in Railroad Tracks," NDT & E Int., **38**, pp. 325–334.

[5] Rajagopalan, J., Balasubramanian, K., and Krishnamurthy, C. V., 2006, "A Phase Reconstruction Algorithm for Lamb Wave Based Structural Health Monitoring of Anisotropic Multilayered Composite Plates," J. Acoust. Soc. Am., **119**(2), pp. 872–878.

[6] Sundararaman, S., and Adams, D. E., 2007, "Simulation of Lamb Wave Propagation in a C458 Al-Li Friction Stir Welded Plate," *Proceedings of the SAMPE Conference*, Baltimore, MD, pp. 1–15.

[7] Delsanto, P. P., Whitcombe, T., Chaskelis, H. H., and Mignogna, R. B., 1992, "Connection Machine Simulation of Ultrasonic Wave Propagation in Materials I: The One-Dimensional Case," Wave Motion, **16**, pp. 65–80.

[8] Delsanto, P. P., Schechter, R. S., Chaskelis, H. H., Mignogna, R. B., and Kline, R. B., 1994, "Connection Machine Simulation of Ultrasonic Wave Propagation in Materials II: The Two-Dimensional Case," Wave Motion, **20**, pp. 295–314.

[9] Delsanto, P. P., Schechter, R. S., and Mignogna, R. B., 1997, "Connection Machine Simulation of Ultrasonic Wave Propagation in Materials III: The Three-Dimensional Case," Wave Motion, **26**, pp. 329–339.

[10] Agostini, V., Delsanto, P. P., Genesio, I., and Oliviero, D., 2003, "Simulation of Lamb Wave Propagation for the Characterization of Complex Structures," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **50**(4), pp. 441–448.

[11] Ruffino, E., and Delsanto, P. P., 1999, "Problems of Accuracy and Reliability in 2-D Lisa Simulations," Comput. Math. Appl., **38**, pp. 89–97.

[12] Iordache, D., Delsanto, P. P., and Scalerandi, M., 1997, "Pulse Distortions in the FD Simulations of Elastic Wave Propagation," Math. Comput. Modell., **25**(6), pp. 31–43.

[13] Delsanto, P. P., Iordache, D., Iordache, C., and Ruffino, E., 1997, "Analysis of Stability and Convergence in FD Simulations of the 1-D Ultrasonic Wave Propagation," Math. Comput. Modell., **25**(6), pp. 19–29.

[14] Cangellaris, A. C., 1993, "Numerical Stability and Numerical Dispersion of a Compact 2-D/FDTD Method Used for the Dispersion Analysis of Waveguides," IEEE Microw. Guid. Wave Lett., **3**(1), pp. 3–5.

[15] Strikwerda, J. C., 2004, *Finite Difference Schemes and Partial Difference Equations*, Society for Industrial and Applied Mathematics, Philadelphia.

[16] Alleyne, D. N., and Cawley, P., 1991, "A Two-Dimensional Fourier Transform Method for the Measurement of Propagating Multimode Signals," J. Acoust. Soc. Am., **89**, pp. 1159–1168.

[17] Graff, K. F., 1991, *Wave Motion in Elastic Solids*, Dover, New York.

[18] Virieux, J., 1986, "P-SV Wave Propagation in Heterogeneous Media: Velocity-Stress Finite Difference Method," Geophysics, **51**(4), pp. 889–901.

[19] Harker, A. H., 1988, *Elastic Waves in Solids With Applications to Non-Destructive Testing of Pipe Lines*, Hilger, Bristol.

[20] Balasubramanyam, R., Quinney, D., Challis, R. E., and Todd, C. P. D., 1996, "A Finite-Difference Simulation of Ultrasonic Lamb Waves in Metal Sheets with Experimental Verification," J. Phys. D, **29**(1), pp. 147–155.

[21] Alleyne, D. N., 1991, "The Nondestructive Testing of Plates Using Ultrasonic Lamb Waves," Ph.D. thesis, Imperial College of Science, Technology and Medicine, London.

[22] Smith, G. D., 1985, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, 3rd ed., Oxford University Press, Oxford.

# Scattering by a Cavity in an Exponentially Graded Half-Space

**P. A. Martin**
Department of Mathematical and Computer Sciences,
Colorado School of Mines,
Golden, CO 80401-1887
e-mail: pamartin@mines.edu

*An inhomogeneous half-space containing a cavity is bonded to a homogeneous half-space. Waves are incident on the interface and the problem is to calculate the scattered waves. For a circular cavity in an exponentially graded half-space, it is shown how to solve the problem by constructing an appropriate set of multipole functions. These functions are singular on the axis of the cavity, they satisfy the governing differential equation in each half-space, and they satisfy the continuity conditions across the interface between the two half-spaces. Seven recent publications are criticized: They do not take proper account of the interface between the two half-spaces.* [DOI: 10.1115/1.3086585]

## 1 Introduction

Consider two half-spaces, $x>0$ and $x<0$, welded together along the interface at $x=0$. The left half-space ($x<0$) is homogeneous. The right half-space is inhomogeneous. If a wave is incident from the left, it will be partly reflected and partly transmitted into the right half-space. We assume that these fields can be calculated.

Suppose now that the right half-space contains a cavity or some other defect (see Fig. 1). How are the basic fields described above modified by the presence of the cavity? In general, it is not easy to answer this question, as the associated mathematical problem is difficult, in general.

In some recent papers, Fang et al. claimed to solve a variety of such problems. All concern "exponential grading," meaning that the material parameters are proportional to $e^{-\beta x}$ for $x>0$, where $\beta$ is a given constant. The papers concern antiplane shear waves [1–4], thermal waves [5,6], and shear waves in a piezoelectric material [7]. All of these papers assume that the effect of the interface on the cavity can be found by introducing simple image terms, as if the interface were a mirror or a rigid wall. Unfortunately, this assumption is incorrect.

In this paper, we outline how the problems described above can be solved. We do this in the context of antiplane shear waves with exponential grading and a circular cavity. The main technical part concerns the derivation of suitable multipole potentials; these reveal the complicated image system.

The study of problems involving scatterers near boundaries or interfaces has a long history. For linear surface water waves interacting with a submerged circular cylinder, see the famous paper by Ursell [8]. For plane-strain elastic waves in a homogeneous half-space with a buried circular cavity, see Ref. [9]. There are also many papers on the scattering of electromagnetic waves by objects near plane boundaries; see, for example, Ref. [10].

Some problems involving objects near plane boundaries can be solved using images. However, determining the strength and location of the images may be difficult: Doing so will depend on the governing differential equations and on the conditions to be satisfied on the plane boundary. For two interesting examples where the location of the images is not obvious, we refer to Chap. 8 of Ting's book [11] (construction of static Green's functions in anisotropic elasticity) and a paper by Stevenson [12] (construction of Green's function for the anisotropic Helmholtz equation in a half-space).

The basic scattering problem is formulated in Sec. 2. The reflection-transmission problem (for which the cavity is absent) is solved in Sec. 3. The solution of this problem gives the "incident" field that will be scattered by the cavity. To solve the scattering problem, we construct an appropriate set of multipole functions (Sec. 4.1). Each of these satisfies the governing differential equations and the interface conditions, and is singular at the center of the circular cavity. Each multipole function is defined as a contour integral of Sommerfeld type; for a careful discussion of similar functions, see Refs. [13,9]. In Sec. 4.2, the multipole functions are combined so as to satisfy the boundary condition on the cavity, leading to an infinite linear system of algebraic equations. The far-field behavior of the multipole functions is deduced in Sec. 4.3. Closing remarks are made in Sec. 5.

## 2 Formulation

We consider the antiplane deformations of two elastic half-spaces, bonded together. In terms of Cartesian coordinates $(x,y)$, the half-space $x<0$ is homogeneous, the half-space $x>0$ is inhomogeneous ("graded"), and the interface is at $x=0$. The homogeneous region has shear modulus $\mu_0$ and density $\rho_0$ (both constants). The inhomogeneous region has shear modulus $\mu(x)$ and density $\rho(x)$ given by

$$\mu(x)=\mu_0 e^{2\beta x} \quad \text{and} \quad \rho(x)=\rho_0 e^{2\beta x} \tag{1}$$

where $\beta$ is a constant. Thus, the material parameters are continuous across the interface.

It is not our purpose here to discuss whether any real materials can be well represented by the functional forms given in Eq. (1). Certainly, the choices in Eq. (1) do lead to some mathematical simplifications and they have been used in the past; see, for example, Refs. [14,15].

For time-harmonic motions, with suppressed time-dependence $e^{-i\omega t}$, the governing equation is

$$\frac{\partial \tau_{xz}}{\partial x}+\frac{\partial \tau_{yz}}{\partial y}+u\omega^2\rho(x)=0$$

where $u(x,y)$ is the antiplane component of displacement and the stress components are given by

$$\tau_{xz}=\mu(x)\frac{\partial u}{\partial x} \quad \text{and} \quad \tau_{yz}=\mu(x)\frac{\partial u}{\partial y}$$

Thus, in the homogeneous region, where we write $u_0$ instead of $u$, we obtain the two-dimensional Helmholtz equation

$$(\nabla^2+k_0^2)u_0=0 \quad \text{with} \quad k_0^2=\rho_0\omega_2/\mu_0 \tag{2}$$

In the inhomogeneous region, we obtain

$$\nabla^2 u+2\beta\frac{\partial u}{\partial x}+k_0^2 u=0$$

This equation is satisfied by writing

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 12, 2007; final manuscript received March 10, 2008; published online March 10, 2009. Review conducted by Professor Sridhar Krishnaswamy.

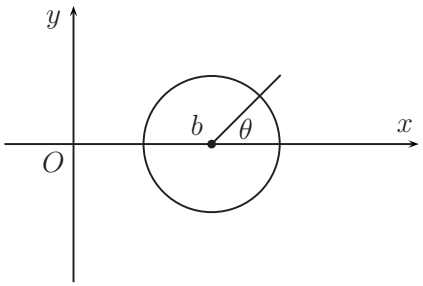



**Fig. 1 The scattering problem. The half-plane on the left of $x=0$ is homogeneous. The other half-plane is inhomogeneous. The circular cavity has radius $a$. A plane wave is incident from the left.**

$$u(x,y) = e^{-\beta x} w(x,y)$$

where $w$ satisfies a different two-dimensional Helmholtz equation

$$(\nabla^2 + k^2)w = 0 \quad \text{with} \quad k^2 = k_0^2 - \beta^2 \tag{3}$$

For simplicity, we assume that $k_0^2 > \beta^2$ and write $k = +\sqrt{k_0^2 - \beta^2}$.

The interface conditions require that the displacements and normal stresses be continuous, so that

$$u_0(0,y) = u(0,y) = w(0,y) \tag{4}$$

$$\left.\frac{\partial u_0}{\partial x}\right|_{x=0} = \left.\frac{\partial u}{\partial x}\right|_{x=0} = \left.\frac{\partial w}{\partial x}\right|_{x=0} - \beta w(0,y) \tag{5}$$

## 3 Incident Field

Suppose that a plane wave is incident on the interface from the homogeneous side. This wave is given by

$$u_{\text{in}}(x,y) = e^{ik_0(x\cos\alpha_0 + y\sin\alpha_0)}$$

where $\alpha_0$ is the angle of incidence, $|\alpha_0| < \pi/2$; $\alpha_0 = 0$ gives normal incidence. There will be a reflected wave $u_{\text{re}}$ and a transmitted wave $u_{\text{tr}}$. Evidently,

$$u_{\text{re}}(x,y) = \mathcal{R} e^{ik_0(-x\cos\alpha_0 + y\sin\alpha_0)}, \quad x < 0 \tag{6}$$

$$u_{\text{tr}}(x,y) = \mathcal{T} e^{-\beta x} e^{ik(x\cos\alpha + y\sin\alpha)}, \quad x > 0 \tag{7}$$

where $\mathcal{R}$, $\mathcal{T}$, and $\alpha$ are to be found. Writing $u_0 = u_{\text{in}} + u_{\text{re}}$ and $u = u_{\text{tr}}$, Eq. (4) gives

$$1 + \mathcal{R} = \mathcal{T} \quad \text{and} \quad k_0 \sin\alpha_0 = k \sin\alpha$$

Then, Eq. (5) gives

$$(1 - \mathcal{R}) i k_0 \cos\alpha_0 = \mathcal{T}(ik\cos\alpha - \beta)$$

Solving for $\mathcal{R}$ gives

$$\mathcal{R} = \frac{k_0\cos\alpha_0 - k\cos\alpha - i\beta}{k_0\cos\alpha_0 + k\cos\alpha + i\beta} = \frac{-i\beta}{k_0\cos\alpha_0 + k\cos\alpha}$$

and then $\mathcal{T} = 1 + \mathcal{R}$.

For a simple check, put $\beta = 0$; we obtain $k = k_0$, $\alpha = \alpha_0$, $\mathcal{R} = 0$, and $\mathcal{T} = 1$, as expected.

## 4 Scattering by a Buried Cavity

Next, we investigate how the wavefields of Sec. 3 are modified if there is a cavity in the inhomogeneous half-space, $x > 0$. See Fig. 1.

We suppose that the cavity's cross section is circular, with boundary

$$(x-b)^2 + y^2 = a^2 \quad \text{with} \quad 0 < a < b$$

We also introduce polar coordinates, $(r, \theta)$, so that

$$x = b + r\cos\theta \quad \text{and} \quad y = r\sin\theta \tag{8}$$

Thus, the cavity's boundary is given by $r = a$, and the boundary condition is

$$\frac{\partial u}{\partial r} = 0 \quad \text{on} \quad r = a \tag{9}$$

where $u$ is the total field in the inhomogeneous half-space.

To solve such a scattering problem, we write

$$u_0 = u_{\text{in}} + u_{\text{re}} + v_0, \quad x < 0$$

$$u = u_{\text{tr}} + v, \quad x > 0, \quad r > a$$

where $v_0$ solves Eq. (2), $v = we^{-\beta x}$, and $w$ solves Eq. (3). Also, $v_0$ must satisfy the Sommerfeld radiation condition and $v$ must decay with $x$.

**4.1 Multipole Functions.** To represent the scattered field, we introduce functions $\phi_n$ of the form

$$\phi_n = \begin{cases} e^{-\beta x}\{H_n^{(1)}(kr)e^{in\theta} + \Phi_n\}, & x > 0 \\ \Psi_n, & x < 0 \end{cases}$$

where $H_n^{(1)}$ is a Hankel function and $n$ is an arbitrary integer. We require that $\Phi_n$ solves Eq. (3) and $\Psi_n$ solves Eq. (2). In addition, $\Phi_n$ and $\Psi_n$ are to be chosen so that $\phi_n$ satisfies the interface conditions, Eqs. (4) and (5).

The use of polar coordinates is convenient for handling the circular cavity but it is inconvenient when trying to impose the conditions at $x = 0$. Therefore, we convert from polar coordinates to Cartesian coordinates using an integral representation; see the Appendix for details. In particular, if we insert Eq. (8) in Eq. (A4), we obtain the integral representation

$$H_n^{(1)}(kr)e^{in\theta} = \frac{(-1)^n}{\pi i}\int_{-\infty}^{\infty+\pi i} e^{k(b-x)\sinh\tau - iky\cosh\tau} e^{-n\tau}\,\mathrm{d}\tau$$

$$\text{for} \quad x < b, |y| < \infty \tag{10}$$

Notice that this formula is valid on the interface $x = 0$. The contour of integration in Eq. (10) is also described in the Appendix.

The form of Eq. (10) suggests using a similar integral representation for $\Phi_n$, and so we write

$$\Phi_n(x,y) = \frac{(-1)^n}{\pi i}\int_{-\infty}^{\infty+\pi i} A(\tau) e^{kx\sinh\tau} e^{-iky\cosh\tau + kb\sinh\tau - n\tau}\,\mathrm{d}\tau, \quad x > 0 \tag{11}$$

where $A(\tau)$ is to be found; $\Phi_n$ solves Eq. (3) automatically for any reasonable choice of $A$.

We shall also need a similar integral representation for $\Psi_n(x,y)$ in $x < 0$, where the wavenumber is $k_0$. However, in order to match solutions across the interface at $x = 0$, we shall require the same dependence on $y$ as in Eq. (11). Thus, we consider

$$\Psi_n(x,y) = \frac{(-1)^n}{\pi i}\int_{-\infty}^{\infty+\pi i} B(\tau) e^{x\Delta(\tau)} e^{-iky\cosh\tau + kb\sinh\tau - n\tau}\,\mathrm{d}\tau, \quad x < 0 \tag{12}$$

where $B(\tau)$ is to be found,

$$\Delta(\tau) = (k^2\cosh^2\tau - k_0^2)^{1/2} = (k^2\sinh^2\tau - \beta^2)^{1/2}$$

and the square root is taken so that $\mathrm{Re}\,\Delta \geq 0$ on the contour. Notice that Eq. (2) is satisfied automatically for any reasonable choice of $B$.

We are now ready to enforce the interface conditions. Continuity of $\phi_n$ across $x = 0$ gives $1 + A = B$ whereas continuity of $\partial\phi_n/\partial x$ gives

$$-k\sinh\tau-\beta+(k\sinh\tau-\beta)A(\tau)=\Delta(\tau)B(\tau)$$

Hence

$$A(\tau)=\frac{k\sinh\tau+\Delta+\beta}{k\sinh\tau-\Delta-\beta} \tag{13}$$

and

$$B(\tau)=\frac{2k\sinh\tau}{k\sinh\tau-\Delta-\beta} \tag{14}$$

These formulas complete the construction of the multipole functions $\phi_n$.

Note that when $\beta=0$, $k=k_0$, $\Delta(\tau)=-k\sinh\tau$, $A=0$, $B=1$, and $\Psi_n=H_n^{(1)}(kr)e^{in\theta}$, as expected.

**4.2 Imposing the Boundary Condition.** In the homogeneous half-space, we write

$$u_0=u_{\text{in}}+u_{\text{re}}+\sum_n c_n\phi_n$$

where $\Sigma_n$ denotes summation over all integers $n$. Similarly, in the graded half-space, we write

$$u(r,\theta)=u_{\text{tr}}+\sum_n c_n\phi_n$$

Then, by construction, the governing partial differential equations and the interface conditions along $x=0$ are all satisfied. It remains to determine the coefficients $c_n$ using the boundary condition on $r=a$, Eq. (9); this gives

$$\sum_n c_n\left.\frac{\partial\phi_n}{\partial r}\right|_{r=a}=-\left.\frac{\partial u_{\text{tr}}}{\partial r}\right|_{r=a} \tag{15}$$

To proceed, we write both sides of this equation as Fourier series in $\theta$. For the right-hand side, we have

$$u_{\text{tr}}=e^{-\beta x}\mathcal{T}_b e^{ikr\cos(\theta-\alpha)}=e^{-\beta x}\mathcal{T}_b\sum_m i^m J_m(kr)e^{im(\theta-\alpha)}$$

where $\mathcal{T}_b=\mathcal{T}\exp(ikb\cos\alpha)$ and $J_n$ is a Bessel function. Also, we have the expansion

$$e^{-\beta x}=e^{-\beta b}\sum_s(-1)^s I_s(\beta r)e^{is\theta}$$

where $I_n$ is a modified Bessel function. Hence,

$$u_{\text{tr}}(r,\theta)=e^{-\beta b}\sum_m(-1)^m U_m(r)e^{im\theta}$$

where

$$U_m(r)=\mathcal{T}_b\sum_s(-i)^s I_{m-s}(\beta r)J_s(kr)e^{-is\alpha}$$

In a similar way, we obtain

$$\Phi_n(r,\theta)=\sum_m(-1)^m f_m^n J_m(kr)e^{im\theta},\quad 0<r<b$$

with

$$f_m^n=\frac{(-1)^n}{\pi i}\int_{-\infty}^{\infty+\pi i}A(\tau)e^{2kb\sinh\tau}e^{-(m+n)\tau}\mathrm{d}\tau$$

Hence

$$\phi_n(r,\theta)=e^{-\beta b}\sum_m(-1)^m V_m^n(r)e^{im\theta}$$

where

$$V_m^n(r)=(-1)^n I_{m-n}(\beta r)H_n^{(1)}(kr)+\sum_s f_s^n I_{m-s}(\beta r)J_s(kr) \tag{16}$$

Thus, Eq. (15) and orthogonality of $\{e^{im\theta}\}$ give

$$\sum_n c_n V_m^{n\prime}(a)=-U_m'(a),\quad \text{all}\quad m \tag{17}$$

which is a linear system of algebraic equations for the coefficients $c_n$.

**4.3 Far-Field Behavior of $\Psi_n$.** We should expect cylindrical waves in the homogeneous half-space. These arise from the far-field behavior of $\Psi_n(x,y)$, for $x<0$. Thus, put

$$x=-R\cos\Theta,\quad y=R\sin\Theta,\quad |\Theta|<\pi/2$$

Then, making the substitution $k\cosh\tau=k_0\cosh s$ in Eq. (12) gives $\Delta=-k_0\sinh s$ and

$$\Psi_n=\frac{1}{\pi i}\int_{-\infty}^{\infty+\pi i}\mathcal{B}_n(k_0\cosh s;\beta)e^{k_0R\sinh(s-i\Theta)}\mathrm{d}s,\quad |\Theta|<\pi/2 \tag{18}$$

where

$$\mathcal{B}_n(\zeta;\beta)=\frac{2(\zeta^2-k_0^2)^{1/2}\exp\{-b(\zeta^2-k^2)^{1/2}\}}{(\zeta^2-k^2)^{1/2}+\beta+(\zeta^2-k_0^2)^{1/2}}\left(\frac{\zeta+(\zeta^2-k^2)^{1/2}}{(-k)}\right)^n$$

the square roots being defined to have non-negative real parts.

The formula for $\Psi_n$, Eq. (18), is convenient for estimating $\Psi_n$ when $k_0R\gg 1$, as we can use the saddle-point method ([16], Chap. 8). There is one relevant saddle point at $s=s_0$ where $s_0=i(\frac{1}{2}\pi+\Theta)$. As $\cosh s_0=-\sin\Theta$ and $\sinh(s_0+i\Theta)=i$, the standard argument gives

$$\Psi_n\sim\frac{1}{\pi i}\mathcal{B}_n(-k_0\sin\Theta;\beta)e^{ik_0R}\int\exp\left\{\tfrac{1}{2}ik_0R(s-s_0)^2\right\}\mathrm{d}s \tag{19}$$

$$\sim\sqrt{\frac{2}{\pi k_0R}}e^{i(k_0R-\pi/4)}\mathcal{B}_n(-k_0\sin\Theta;\beta)\quad\text{as}\quad R\to\infty \tag{20}$$

where the contour of integration in Eq. (19) passes through the saddle point.

When $\beta=0$, we obtain $\mathcal{B}_n(-k_0\sin\Theta;0)=i^n e^{ik_0b\cos\Theta}e^{-in\Theta}$. Then, Eq. (20) agrees with the known far-field expansion of $H_n^{(1)}\times(kr)e^{in\theta}$, when one takes into account that $\theta\sim\pi-\Theta$ and $r\sim R+b\cos\Theta$ as $R\to\infty$.

**4.4 Near-Field Behavior of $\Phi_n$.** As the expression for $\Phi_n$, Eq. (11), is similar to Eq. (10), it is reasonable to ask if $\Phi_n$ corresponds to a simple image term. To see that it does not, let us define polar coordinates centered at the mirror-image point, $(x,y)=(-b,0)$: $x=-b+r'\cos\theta'$, $y=r'\sin\theta'$. Then, calculations similar to those described in the Appendix show that

$$H_n^{(1)}(kr')e^{in(\pi-\theta')}=\frac{(-1)^n}{\pi i}\int_{-\infty}^{\infty+\pi i}e^{kx\sinh\tau}e^{-iky\cosh\tau+kb\sinh\tau-n\tau}\mathrm{d}\tau \tag{21}$$

for $|\theta'|<\pi/2$. The integral on the right-hand side of Eq. (21) should be compared with the integral defining $\Phi_n$, Eq. (11). For them to be equal, the function $A(\tau)$, defined by Eq. (13), would have to be constant: It is not, and it is not well approximated by a nonzero constant. Thus, it is not justified to replace $\Phi_n$ with a simple image term: We notice that Fang et al. [2] used image terms similar to those on the left-hand side of Eq. (21), with $\pi-\theta'$ replaced with $\theta'$.

## 5 Discussion

We have outlined how to solve the scattering problem for a cavity buried in a graded half-space; the result is the infinite linear algebraic system, Eq. (17). The system matrix is very complicated: One has to calculate $(\mathrm{d}/\mathrm{d}r)V_m^n(r)$ at $r=a$, where $V_m^n$ is defined by Eq. (16) as an infinite series of special functions with coefficients given as contour integrals. In principle, the system matrix could be computed but it is unclear whether this is a worthwhile exercise, given the limitations of the underlying model, with both shear modulus and density varying exponentially; see Eq. (1). However, it may be possible to extract asymptotic results from the exact system of equations for small cavities or for cavities that are far from the interface: This remains for future work.

## Appendix: Integral Representations

As explained in Sec. 4.1, we need to convert from polar coordinates to Cartesian coordinates in order to apply the interface conditions at $x=0$. This is done using certain integral formulas. Thus, from Ref. [17] (p. 178, Eq. (2)), we have the integral representation

$$H_n^{(1)}(kr) = \frac{1}{\pi i}\int_{-\infty}^{\infty+\pi i} e^{kr\sinh w - nw}\mathrm{d}w \tag{A1}$$

The integration is along any contour in the complex $w$-plane, starting at $w=-\infty$ and ending at $w=\pi i+\infty$. When $w=\xi+i\eta$, where $\xi$ and $\eta$ are real, $|e^{kr\sinh w}|=e^{kr\sinh\xi\cos\eta}$. Thus, we can generalize Eq. (A1) to

$$H_n^{(1)}(kr) = \frac{1}{\pi i}\int_{-\infty+i\eta_1}^{\infty+i\eta_2} e^{kr\sinh w - nw}\mathrm{d}w \tag{A2}$$

where the constants $\eta_1$ and $\eta_2$ must satisfy

$$-\tfrac{1}{2}\pi < \eta_1 < \tfrac{1}{2}\pi \quad \text{and} \quad \tfrac{1}{2}\pi < \eta_2 < \tfrac{3}{2}\pi$$

In other words, we have some flexibility in our choice of contour, flexibility that we shall exploit shortly.

Put $w=\tau+i(\theta-\pi)$. Then Eq. (A2) becomes

$$H_n^{(1)}(kr)e^{in\theta} = \frac{(-1)^n}{\pi i}\int_{-\infty+i\beta_1}^{\infty+i\beta_2} e^{-kr(\sinh\tau\cos\theta + i\cosh\tau\sin\theta)}e^{-n\tau}\mathrm{d}\tau \tag{A3}$$

where the constants $\beta_1$ and $\beta_2$ must satisfy

$$-\tfrac{1}{2}\pi < \beta_1+\theta-\pi < \tfrac{1}{2}\pi \quad \text{and} \quad \tfrac{1}{2}\pi < \beta_2+\theta-\pi < \tfrac{3}{2}\pi$$

In particular, the choices $\beta_1=0$ and $\beta_2=\pi$ show that

$$H_n^{(1)}(kr)e^{in\theta} = \frac{(-1)^n}{\pi i}\int_{-\infty}^{\infty+\pi i} e^{-kr(\sinh\tau\cos\theta + i\cosh\tau\sin\theta)}e^{-n\tau}\mathrm{d}\tau$$

$$\text{for} \quad \tfrac{1}{2}\pi < \theta < \tfrac{3}{2}\pi \tag{A4}$$

This is the integral representation that we use in Sec. 4.1.

## References


[1] Fang, X.-Q., Hu, C., and Du, S. Y., 2006, "Strain Energy Density of a Circular Cavity Buried in Semi-Infinite Functionally Graded Materials Subjected to Shear Waves," Theor. Appl. Fract. Mech., **46**, pp. 166–174.

[2] Fang, X.-Q., Hu, C., and Du, S.-Y., 2007, "Dynamic Stress of a Circular Cavity Buried in a Semi-Infinite Functionally Graded Material Subjected to Shear Waves," ASME J. Appl. Mech., **74**, pp. 916–922.

[3] Fang, X.-Q., Hu, C., and Huang, W.-H., 2007, "Strain Energy Density of a Circular Cavity Buried in a Semi-Infinite Slab of Functionally Graded Materials Subjected to Anti-Plane Shear Waves," Int. J. Solids Struct., **44**, pp. 6987–6998.

[4] Hu, C., Fang, X.-Q., and Huang, W.-H., 2008, "Multiple Scattering of Shear Waves and Dynamic Stress From a Circular Cavity Buried in a Semi-Infinite Slab of Functionally Graded Materials," Eng. Fract. Mech., **75**, pp. 1171–1183.

[5] Hu, C., Fang, X.-Q., and Du, S.-Y., 2007, "Multiple Scattering of Thermal Waves From a Subsurface Spheroid in Exponentially Graded Materials Based on Non-Fourier's Model," Infrared Phys. Technol., **50**, pp. 70–77.

[6] Fang, X. Q., and Hu, C., 2007, "Thermal Waves Scattering by a Subsurface Sphere in a Semi-Infinite Exponentially Graded Material Using Non-Fourier's Model," Thermochim. Acta, **453**, pp. 128–135.

[7] Fang, X.-Q., Hu, C., and Huang, W.-H., 2007, "Dynamic Stress of a Circular Cavity Buried in a Semi-Infinite Functionally Graded Piezoelectric Material Subjected to Shear Waves," Eur. J. Mech. A/Solids, **26**, pp. 1016–1028.

[8] Ursell, F., 1950, "Surface Waves on Deep Water in the Presence of a Submerged Circular Cylinder. I," Proc. Cambridge Philos. Soc., **46**, pp. 141–152.

[9] Gregory, R. D., 1970, "The Propagation of Waves in an Elastic Half-Space Containing a Cylindrical Cavity," Proc. Cambridge Philos. Soc., **67**, pp. 689–710.

[10] Borghi, R., Gori, F., Santarsiero, M., Frezza, F., and Schettini, G., 1996, "Plane-Wave Scattering by a Perfectly Conducting Circular Cylinder Near a Plane Surface: Cylindrical-Wave Approach," J. Opt. Soc. Am. A, **13**, pp. 483–493.

[11] Ting, T. C. T., 1996, *Anisotropic Elasticity*, Oxford University Press, Oxford.

[12] Stevenson, R. C., 1990, "Green's Function for the Helmholtz Equation of a Scalar Wave in an Anisotropic Halfspace," SIAM J. Appl. Math., **50**, pp. 199–215.

[13] Gregory, R. D., 1967, "An Expansion Theorem Applicable to Problems of Wave Propagation in an Elastic Half-Space Containing a Cavity," Proc. Cambridge Philos. Soc., **63**, pp. 1341–1367.

[14] Rangelov, T. V., Manolis, G. D., and Dinerva, P. S., 2005, "Elastodynamic Fundamental Solutions for Certain Families of 2d Inhomogeneous Anisotropic Domains: Basic Derivations," Eur. J. Mech. A/Solids, **24**, pp. 820–836.

[15] Destrade, M., 2007, "Seismic Rayleigh Waves on an Exponentially Graded, Orthotropic Half-Space," Proc. R. Soc. London, Ser. A, **463**, pp. 495–502.

[16] Copson, E. T., 1965, *Asymptotic Expansions*, Cambridge University Press, Cambridge.

[17] Watson, G. N., 1944, *Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge.

**Weiwei Yu**

**Dale G. Karr**
e-mail: dgkarr@umich.edu

Department of Naval Architecture and
Marine Engineering,
University of Michigan,
Ann Arbor, MI 48109

# Symmetrical Solutions for Edge-Loaded Annular Elastic Membranes

*The Föppl–Hencky nonlinear membrane theory is employed to study the axisymmetric deformation of annular elastic membranes. The general solutions for displacements and stresses are established for arbitrary edge boundary conditions. New exact solution results are developed for central loading and edge forcing conditions. Both positive and negative radial stress solutions are found. Comparisons are made for special cases to previously known solutions with excellent agreement.* [DOI: 10.1115/1.3005568]

## 1 Introduction

The development of nonlinear elastic membrane theory began a century ago with the work of Föppl [1]. Within this theory, strains are presumed to be sufficiently small such that linear elastic stress-strain relations are valid. The nonlinearity arises from the effects of small rotations of membrane elements. This nonlinearity is also present in plate theory attributed to Von Karman [2] and Way [3], coupling the effects of bending and in-plane stretching. An important particular application of this theory is that by Hencky [4]. Hencky provided power series solutions for a flat circular membrane under uniform vertical pressure in which the flexural rigidity is neglected.

A more general theory applicable for axisymmetric deformation including large rotations was developed by Reissner [5,6]. The range of applicability of the Föppl–Hencky theory and the Reissner theory for flat annular membrane problems was provided by Weinitschke [7]. Dickey [8] derived an exact theory for a plane circular membrane subjected to a vertical pressure, showing that the Föppl theory corresponded to the first term expansion for the exact theory.

In other closely related studies, Dickey [9] found axisymmetric solutions for a circular membrane under normal pressure, employing a numerical integration scheme. Employing an iterative method, detailed analyses of an annular membrane under surface and edge loads were provided by Weinitschke [10]. Existence and uniqueness of tensile and compressive conditions were examined. Callegari and Reiss [11] applied a shooting method to solve boundary value problems for the axisymmetric deformation of a circular membrane, also examining existence and uniqueness. Kelkar et al. [12] analyzed a circular membrane with fixed peripheral edges; displacement and stresses under three different loading conditions were found using a finite difference technique. Grabmüller and Novak [13] developed a refined integral equation solution technique, proving that tensile solutions may cease to exist when large radial displacements are prescribed at the inner radius. As discussed in Sec. 3, this is consistent with our findings.

In this paper we develop new exact axisymmetric solutions for annular membranes subjected to edge forcing, applying the Föppl theory. Some exact solutions for annular membrane problems for loading and edge conditions were first presented by Schwerin [14]. Of particular interest is Schwerin's analysis of the problem of an elastic membrane subjected to a vertical force acting at the interior of the membrane. Analytical solutions were given with Poisson's ratio less than or equal to one-third. We follow closely the approach of Schwerin, reducing the governing differential equations to a single nonlinear ordinary differential equation in terms of a single dependent variable. Both positive and negative radial stress solutions are found.

Such analyses are motivated by various applications. For example, applications of circular membrane theory for indentation of tires were pointed out by Yang and Hsu [15]. Chen and Cheng [16] analyzed ponding pressures acting on a circular membrane as a result of the weight of a liquid filling the space created by the deflection of a membrane. An outer circular region thus developed provided the membrane's self-weight was negligible relative to the liquid pressure. They developed an iterative technique to solve for the displacements in both the loaded and unloaded portions of the membrane.

Indentation problems also arise for marine, land, and air vehicles, when thin plates are subjected to impact loading. A common hypothesis is that membrane stresses predominate for large deformations [17]. Circular membrane analyses are a first-order approximation of the geometry for impact experiments and analyses [12,18].

Recent applications also include the analysis of circular holes in nonlinear circular membranes in an effort to analyze the effects of cavities and implanted rigid fixations introduced into thin tissue by clinical procedures [19,20]. Annular membrane models are also used in mechanical testing of soft materials such as polymers, elastomers, and biomaterials [21] and in the analysis of electrostatic deflection on elastic membranes for microelectromechanical actuators [22]. Theoretical and experimental results are presented by Begley and Mackin [23] for the spherical indentation of free-standing thin films; they also compared their results with the classical Schwerin-type analyses.

## 2 Analysis

As mentioned, Schwerin [14] analyzed the problem of an elastic membrane with displacements $u$ and $w$ equal to zero at the outer radius $r=r_a$. The central region of the membrane $0<r<r_i$ consisted of a rigid disk subjected to a vertical force resultant of magnitude $P$ acting at the center of the membrane. The annular region $r_i<r<r_a$ was subjected to a uniformly distributed vertical pressure $p$. In the following, in a manner similar to that of Schwerin, we develop the general solution to the governing differential equations of the elastic membrane with $p=0$, as shown in Fig. 1. Our analysis pertains to the annular region $r_i<r<r_a$ with loading provided by the stresses or displacements prescribed at the boundaries. Various boundary conditions on displacement or stress can thus be determined. Solutions for selected displacement boundary value problems are presented. For the particular case with the boundary conditions $u(r_i)=u(r_a)=w(r_a)=0$, our solutions are the same as those found by Schwerin provided that $\nu\leq\frac{1}{3}$. For these

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received August 14, 2007; final manuscript received September 30, 2008; published online March 10, 2009. Review conducted by Matthew R. Begley.

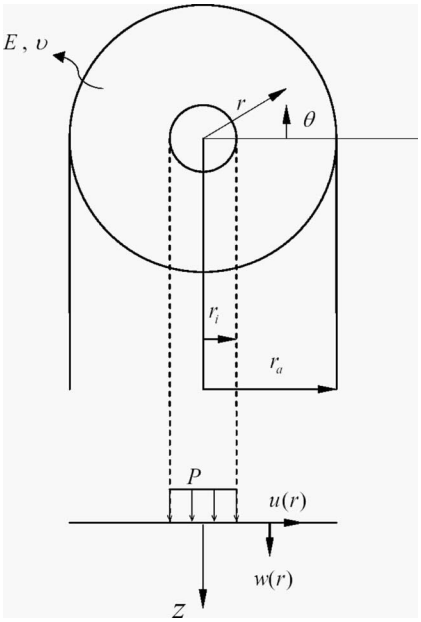



**Fig. 1 Geometry and notation for a center-loaded membrane. The inner portion provides a net vertical force *P*. The resulting edge-loaded annular membrane $r_i < r < r_a$ is analyzed.**

boundary conditions, we also find that solutions exist for $\nu > \frac{1}{3}$, in contrast to Schwerin's conclusion that axisymmetric solutions exist only for $\nu \leq \frac{1}{3}$.

For axisymmetric plane stress conditions, the circumferential strain and radial strain are related to the circumferential stress and radial stress, respectively, by

$$\varepsilon_\theta = \frac{1}{E}(\sigma_\theta - \nu\sigma_r) \tag{1}$$

$$\varepsilon_r = \frac{1}{E}(\sigma_r - \nu\sigma_\theta) \tag{2}$$

The strain-displacement relations are

$$\varepsilon_\theta = \frac{u}{r} \tag{3}$$

$$\varepsilon_r = \frac{du}{dr} + \frac{1}{2}\left(\frac{dw}{dr}\right)^2 \tag{4}$$

where $u$ is the displacement in the radial direction, and $w$ is the displacement in the $z$-direction. Substitution of the expressions for strain in Eq. (1) into Eq. (3) yields the following displacement-stress relation:

$$u = \frac{r}{E}(\sigma_\theta - \nu\sigma_r) \tag{5}$$

Differentiating Eq. (5) and substituting this result, along with Eq. (2) into Eq. (4), yields

$$\frac{1}{2}\left(\frac{dw}{dr}\right)^2 = \frac{1}{E}\left[(1+\nu)(\sigma_r - \sigma_\theta) - r\left(\frac{d\sigma_\theta}{dr} - \nu\frac{d\sigma_r}{dr}\right)\right] \tag{6}$$

The equilibrium of a differential element yields from the summation of forces in the radial direction

$$\frac{d\sigma_r}{dr} + \frac{1}{r}(\sigma_r - \sigma_\theta) = 0 \tag{7}$$

The summation of forces in the vertical direction yields

$$\sigma_r\frac{d^2w}{dr^2} + \frac{1}{r}\sigma_\theta\frac{dw}{dr} = 0 \tag{8}$$

Equations (6)–(8) involve three nonlinear differential equations of the first-order in $\sigma_r$ and $\sigma_\theta$, and of the second-order in displacement $w$. Accordingly, in the applications that follow, the solutions involve four boundary conditions for the axisymmetric deformation of the membrane.

The system of equations can be further reduced in several ways. The most popular approach is to introduce the stress function, from which we find

$$\sigma_r = \frac{1}{r}\frac{d\phi}{dr} \tag{9}$$

$$\sigma_\theta = \frac{d^2\phi}{dr^2} \tag{10}$$

Then, Eq. (7) is satisfied identically, and Eqs. (6) and (8) may be rewritten, respectively, as

$$\frac{E}{2}\left(\frac{dw}{dr}\right)^2 + r\frac{d}{dr}\left(\frac{d^2\phi}{dr^2} + \frac{1}{r}\frac{d\phi}{dr}\right) = 0 \tag{11}$$

and

$$\frac{d}{dr}\left(\frac{d\phi}{dr}\frac{dw}{dr}\right) = 0 \tag{12}$$

Equations (11) and (12) are special homogeneous cases of the nonlinear axisymmetric membrane theory.

An alternative system of equations can be found from Eqs. (6)–(8) by solving for $\sigma_\theta$ in terms of $\sigma_r$ using Eq. (7)

$$\sigma_\theta = \frac{d(r\sigma_r)}{dr} \tag{13}$$

Substituting this result into Eq. (6) yields

$$\frac{E}{2}\left(\frac{dw}{dr}\right)^2 + r\frac{d}{dr}\left[\sigma_r + \frac{d(r\sigma_r)}{dr}\right] = 0 \tag{14}$$

Also, substituting Eq. (13) into Eq. (8) yields

$$\frac{d}{dr}\left[r\sigma_r\frac{dw}{dr}\right] = 0 \tag{15}$$

Tuan [24] used equations in this form in an analysis of ponding of circular membranes using a fourth-order Runge–Kutta method. From Eq. (15) and from the equilibrium of the central disk, for the condition in which no pressure is applied to the annular region, we find

$$r\sigma_r\frac{dw}{dr} + \frac{P}{2\pi h} = 0 \tag{16}$$

Equations (14) and (16) now correspond, respectively, to Eqs. (5) and (4) of Schwerin's article [14].

Introducing the function

$$y \equiv r^2\frac{\sigma_r}{E} \tag{17}$$

and substituting the above equation into Eqs. (14) and (15) yields, respectively, two equations for $y(r)$ and $w(r)$

$$\frac{1}{2}\left(\frac{dw}{dr}\right)^2 + \frac{d^2y}{dr^2} - \frac{1}{r}\frac{dy}{dr} = 0 \tag{18}$$

$$\frac{d}{dr}\left(\frac{y}{r}\frac{dw}{dr}\right) = 0 \tag{19}$$

Integrating Eq. (19) yields the solution for the product

$$\frac{y}{r}\frac{dw}{dr} = 2a_1 \tag{20}$$

where $a_1$ is a constant of integration and the factor 2 is used to simplify expressions that follow. Further simplification can be achieved by introducing the new independent variable

$$x = r^2 \tag{21}$$

A single nonlinear ordinary differential equation for $y$ is then found by substituting the result (Eq. (20)) into Eq. (18)

$$y^2\frac{d^2y}{dx^2} = -\frac{1}{2}a_1^2 \tag{22}$$

The physical significance of the constant $a_1$ can be interpreted by noting that the net vertical force of any circular region of the membrane is

$$P = -2\pi r\frac{dw}{dr}\sigma_r h \tag{23}$$

Then from Eqs. (17) and (20) we find

$$a_1 = -\frac{P}{4\pi Eh} \tag{24}$$

Note that in the following we address only problems in which $P$, and hence $a_1$, are nonzero. This also implies, from Eq. (22), that $(d^2y/dx^2) < 0$. By introducing the nondimensional variable $V$,

$$V \equiv -K_1 y \tag{25}$$

where $K_1 = (1/r_a^2)\sqrt[3]{(2r_a^2/a_1^2)}$, and the nondimensional independent variable $\xi$,

$$\xi \equiv \frac{x}{r_a^2} \tag{26}$$

Equation (22) is now reduced to the equivalent form

$$\frac{d^2V}{d\xi^2} = \frac{1}{V^2} \tag{27}$$

By introducing

$$\rho \equiv \frac{dV}{d\xi} \tag{28}$$

we find

$$\rho d\rho = \frac{1}{V^2}dV \tag{29}$$

Integrating Eq. (29) yields

$$\frac{1}{2}\rho^2 = -\frac{1}{V} - \frac{a_2}{2} \tag{30}$$

where a second constant of integration $a_2$ has been introduced. From Eqs. (28) and (30) we also find

$$d\xi = \pm\frac{1}{\sqrt{2}}\frac{dV}{\sqrt{-\frac{1}{V} - \frac{a_2}{2}}} \tag{31}$$

Integration of Eq. (31) yields, with the introduction of a third integration constant $a_3$,

$$\xi + a_3 = \pm\frac{1}{\sqrt{2}}\int\frac{dV}{\sqrt{-\frac{1}{V} - \frac{a_2}{2}}} \tag{32}$$

The result of the integration of the right-hand side of Eq. (32) depends on whether $a_2$ is less than, greater than, or equal to zero. We introduce two parameters for conditions with $a_2$ nonzero

$$\varphi \equiv 2\cot^{-1}\left(\frac{1}{\sqrt{a_2}}\sqrt{-\frac{2}{V} - a_2}\right), \quad a_2 > 0 \tag{33}$$

$$\tau \equiv 2\tanh^{-1}\left(\frac{1}{\sqrt{-a_2}}\sqrt{-\frac{2}{V} - a_2}\right), \quad a_2 < 0 \quad V > 0 \tag{34}$$

$$\tau \equiv 2\coth^{-1}\left(\frac{1}{\sqrt{-a_2}}\sqrt{-\frac{2}{V} - a_2}\right), \quad a_2 < 0 \quad V < 0 \tag{35}$$

Equation (32) yields, using Eqs. (33)–(35) respectively,

$$\xi + a_3 = \pm\left(\frac{1}{a_2}\right)^{(3/2)}(\sin\varphi - \varphi), \quad a_2 > 0 \tag{36}$$

$$\xi + a_3 = \pm\left(\frac{1}{-a_2}\right)^{(3/2)}(\sinh\tau + \tau), \quad a_2 < 0 \quad V > 0 \tag{37}$$

$$\xi + a_3 = \mp\left(\frac{1}{-a_2}\right)^{(3/2)}(\sinh\tau - \tau), \quad a_2 < 0 \quad V < 0 \tag{38}$$

A fourth solution exists when $a_2 = 0$ in the form

$$\xi + a_3 = \mp\frac{\sqrt{2}}{3}(-V)^{(3/2)}, \quad a_2 = 0 \tag{39}$$

Expressions for $V$ in terms of $\varphi$ and $\tau$ are found from Eqs. (33)–(35), respectively,

$$V = -\frac{1}{a_2}(1 - \cos\varphi), \quad a_2 > 0 \tag{40}$$

$$V = -\frac{1}{a_2}(1 + \cosh\tau), \quad a_2 < 0 \quad V > 0 \tag{41}$$

$$V = \frac{1}{a_2}(\cosh\tau - 1), \quad a_2 < 0 \quad V < 0 \tag{42}$$

Solutions for the vertical displacement are now determined from Eq. (20), which can be rewritten as

$$\frac{dw}{d\xi} = -\frac{1}{V}\sqrt[3]{2a_1 r_a^2} \tag{43}$$

Integration of Eq. (43) upon substitution of either Eqs. (39)–(42)—as the case may be—yields

$$w + a_4 = \mp\sqrt[3]{2a_1 r_a^2}\frac{1}{\sqrt{a_2}}\varphi, \quad a_2 > 0 \tag{44}$$

$$w + a_4 = \mp\sqrt[3]{2a_1 r_a^2}\frac{1}{\sqrt{-a_2}}\tau, \quad a_2 < 0 \quad V > 0 \tag{45}$$

$$w + a_4 = \mp\sqrt[3]{2a_1 r_a^2}\frac{1}{\sqrt{-a_2}}\tau, \quad a_2 < 0 \quad V < 0 \tag{46}$$

$$w + a_4 = \mp\sqrt[3]{2a_1 r_a^2}\sqrt{-2V}, \quad a_2 = 0 \tag{47}$$

The constant of integration $a_4$ represents a vertical rigid body translation. Solutions for the horizontal displacement are now determined from Eq. (5), which can be rewritten as

$$u = \frac{1}{K_1 r_a}\sqrt{\xi}\left[\mp 2\sqrt{-\frac{2}{V} - a_2} + \frac{V}{\xi}(1 + \nu)\right] \tag{48}$$

Substitution of Eqs. (39)–(42) yields the radial displacement

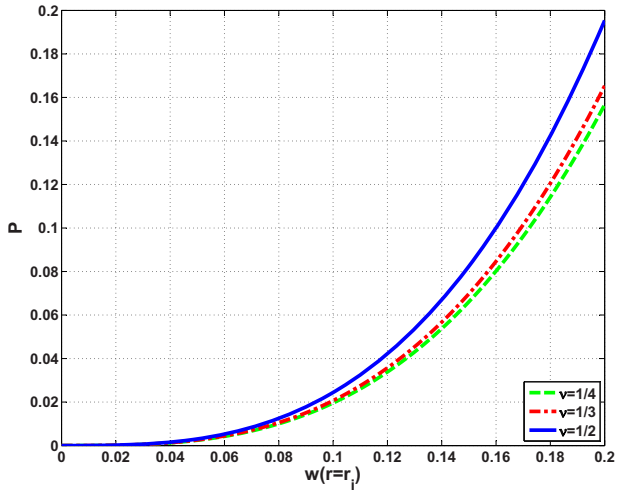



**Fig. 2 Vertical force *P*, applied at the inner boundary versus vertical displacement at inner boundary $w(r_i)$ for different Poisson's ratios with prescribed boundary conditions $w(r_a)=u(r_i)=u(r_a)=0$, and $(r_i/r_a)=0.5$**

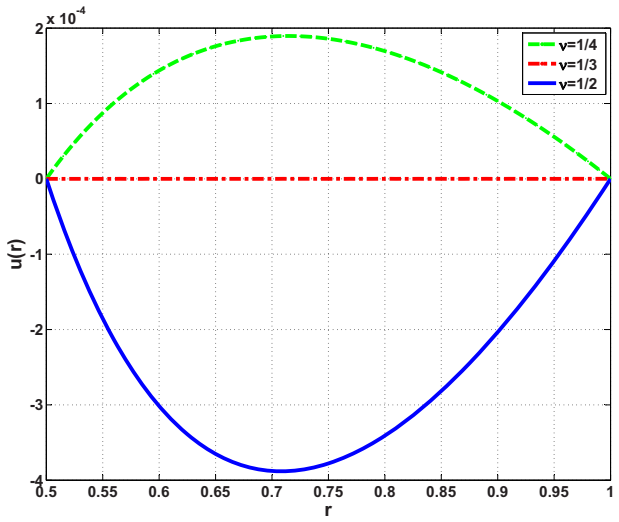



**Fig. 3 Horizontal displacement $u(r)$ for different Poisson's ratios with prescribed vertical displacements $w(r_i)=0.1$, $w(r_a)=u(r_i)=u(r_a)=0$, and $(r_i/r_a)=0.5$**

$$u=\frac{1}{K_1 r_a}\sqrt{\xi}\left[\mp 2\sqrt{a_2}\cot\frac{\varphi}{2}-\frac{(1-\cos\varphi)}{\xi a_2}(1+\nu)\right],\quad a_2>0 \tag{49}$$

$$u=\frac{1}{K_1 r_a}\sqrt{\xi}\left[\mp 2\sqrt{-a_2}\tanh\frac{\tau}{2}-\frac{(1+\cosh\tau)}{\xi a_2}(1+\nu)\right],\quad a_2<0\quad V>0 \tag{50}$$

$$u=\frac{1}{K_1 r_a}\sqrt{\xi}\left[\mp 2\sqrt{-a_2}\coth\frac{\tau}{2}+\frac{(\cosh\tau-1)}{\xi a_2}(1+\nu)\right],\quad a_2<0\quad V<0 \tag{51}$$

$$u=\frac{1}{K_1 r_a}\sqrt{\xi}\left[\mp 2\sqrt{-\frac{2}{V}}+\frac{V}{\xi}(1+\nu)\right],\quad a_2=0 \tag{52}$$

Note that from Eqs. (17) and (25), $V$ is of the opposite sign of the radial stress $\sigma_r$. Positive values of $V$ thus correspond to negative radial stress, changing the stability state of the membrane [25,26].

## 3 Results and Discussion

In this section, a membrane with uniformly distributed load around an inner circumference is studied. For all the calculations shown in the following figures, Young's modulus, the outer radius, and the thickness of the membrane of 1 are assumed. This is equivalent to nondimensionalizing the stresses $\bar{\sigma}=(\sigma/E)$, vertical force $\bar{P}=(P/Ehr_a)$, and displacements $\bar{u}=(u/r_a)$ and $\bar{w}=(w/r_a)$.

First, the relationships between the vertical displacement $w(r)$ and the net force $P$ applied uniformly around the inner circumference with various Poisson's ratios are examined. The boundary conditions are

- $w(r_i)$ is prescribed
- $w(r_a)=0$
- $u(r_a)=0$
- $u(r_i)=0$

Force-displacement relations are shown for $(r_i/r_a)=0.5$ in Fig. 2. A system of simultaneous equations must be solved to obtain these solutions. The system of equations depends on the signs of the constant $a_2$ and variable $V$. One particular method we found efficient is to first solve for $a_1$ in terms of a given vertical force $P$ using Eq. (24). This also provides a value for $K_1$. Evaluating $\xi$ at the inner and outer boundaries yields values of $\xi_1=0.25$ and $\xi_2=1$, respectively. Substituting these values into Eq. (36) yields two expressions for $\phi_1$ and $\phi_2$, as follows (assuming that $a_2>0$):

$$0.25+a_3=\pm\left(\frac{1}{a_2}\right)^{(3/2)}(\sin\varphi_1-\varphi_1),\quad a_2>0 \tag{53}$$

$$1+a_3=\pm\left(\frac{1}{a_2}\right)^{(3/2)}(\sin\varphi_2-\varphi_2),\quad a_2>0 \tag{54}$$

Two more equations from the boundary condition on the radial displacement at the inner and outer edges being zero yield from Eq. (49).

$$0=\frac{1}{K_1}\sqrt{0.25}\left[\mp 2\sqrt{a_2}\cot\frac{\varphi_1}{2}-\frac{(1-\cos\varphi_1)}{0.25a_2}(1+\nu)\right] \tag{55}$$

$$0=\frac{1}{K_1}\left[\mp 2\sqrt{a_2}\cot\frac{\varphi_2}{2}-\frac{(1-\cos\varphi_2)}{a_2}(1+\nu)\right] \tag{56}$$

This system involves four equations in four unknowns $\phi_1$, $\phi_2$, $a_2$, and $a_3$. These equations are solved using MATLAB package [27]. The boundary condition for $w(r_a)=0$ yields from Eq. (44)

$$a_4=\mp\sqrt[3]{2a_1}\frac{1}{\sqrt{a_2}}\varphi_2 \tag{57}$$

Equation (44) can then be applied to determine the displacement at the inner boundary for the given applied force, specifically.

$$w(r_i)=\mp\sqrt[3]{2a_1(0.25)^2}\frac{1}{\sqrt{a_2}}\varphi_1-a_4 \tag{58}$$

Note that our solutions for Poisson's ratios, $\upsilon\leq\frac{1}{3}$ are in full agreement with Schwerin's [14]. However, we find that axisymmetric solutions for $\upsilon>\frac{1}{3}$ exist as per Eqs. (46) and (51). As indicated, larger Poisson's ratios increase the stiffness of the membrane. Figure 3 shows the displacement $u(r)$ for the same boundary conditions with various Poisson's ratios with $w(r_i)=0.1$. Note that only for $\upsilon=\frac{1}{3}$ is $u(r)=0$. Maximum horizontal displacements are about three-orders of magnitude smaller than the maximum vertical displacement.

In order to establish the physical significance of the integration constant $a_2$, we studied the special case with $(r_i/r_a)=0.5$, the outer

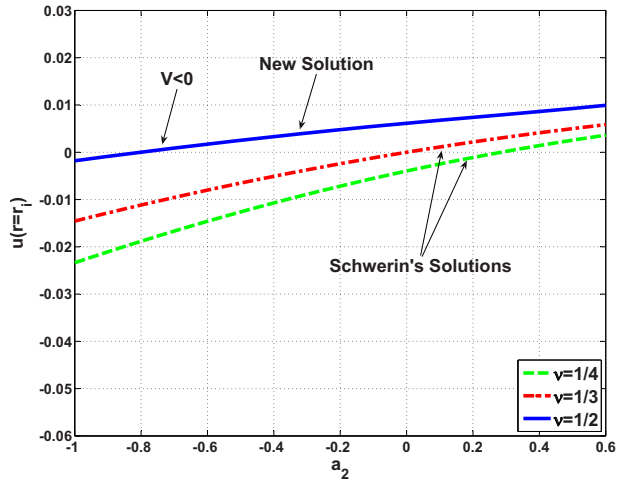



**Fig. 4 Integration constant $a_2$ versus horizontal displacement at inner boundary $u(r_i)$ with prescribed vertical edge displacement $w(r_i)=0.1$, $w(r_a)=u(r_a)=0$**

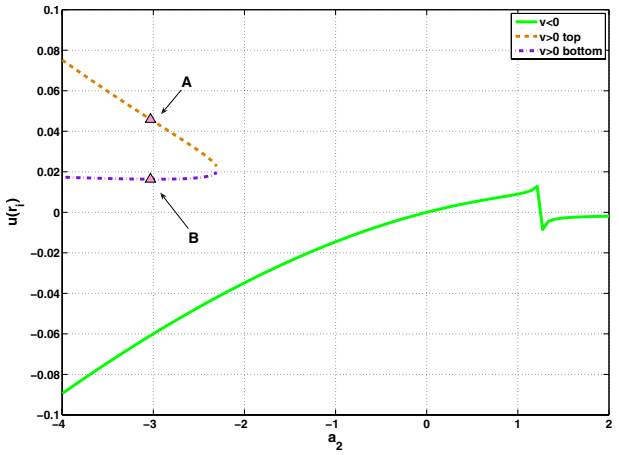



**Fig. 6 Integration constant $a_2$ versus horizontal displacement at inner boundary $u(r_i)$, with vertical displacement $w(r_i)=0.1$, Poisson's ratio $\nu=\frac{1}{3}$ and $(r_i/r_a)=0.5$**

edge displacement is zero, and the inner edge nondimensional vertical displacement fixed at 0.1. Figure 4 shows results as the inner radial displacement is varied for three different Poisson's ratios. The general trend is for the constant $a_2$ to increase as the inner radial displacement increases. This implies that as the radial membrane tensile strains increase, the constant $a_2$ increases for a given inner vertical displacement. This behavior is also reflected in Fig. 5 where Poisson's ratio is held at $\frac{1}{3}$ and the inner vertical displacement is varied.

In a wider region of $a_2$, we note that there is a discontinuity in solutions with $a_2>1$ as shown in Fig. 6. These solutions involve large slopes $dw/dr$, and are thus beyond the scope of the Föppl theory. Also as seen from Fig. 6, multiple solutions exist. In addition to the solution with $V<0$ (tensile radial stress), two solutions can be found for a single value of $a_2$. Two particular solutions for $V>0$, which corresponds to compressive radial stresses, are labeled A and B in Fig. 7. Point A corresponds to solutions from Eqs. (45) and (51) using the negative of the $\mp$ option, hence the "top" label in the figure. Point B corresponds to solutions using the positive option. The solutions for vertical displacement versus radial position are shown in Fig. 7. Note that the positive radial displacements are found at $r=r_i=0.5$. This results in compressive values of radial stress. The resulting radial and circumferential stresses are shown in Fig. 8.

Comparisons of our results are also made to previous results of other authors. As shown in Table 1, nondimensionalized radial and circumferential stresses and vertical displacement agree very well with the results of Schwerin [14] for $u(r_i)=u(r_a)=0$, $w(r_a)=0$ and $P=12.57$. We take the same nondimensionalized method that Schwerin used: $\overline{\sigma_r}=(\sigma_r/E)\cdot\sqrt[3]{(32\pi^2E^2h^2r_a^2/P^2)}$, $\overline{\sigma_\theta}=(\sigma_\theta/E)\cdot\sqrt[3]{(32\pi^2E^2h^2r_a^2/P^2)}$, and $\bar{w}=(w/r_a)\cdot\sqrt[3]{(2\pi r_a hE/P)}$. Also, the result of Tuan [24] with ponding radius equal to $b=90.53$ cm, radial and circumferential stresses at the outer boundary $a=185.70$ cm, are compared in Table 2. Differences of only 2.70% and 2.50% in radial and circumferential stresses, respectively, are found between the present results and those of Tuan's. This is presumably due to the calculation accuracy differences between the analytical and numerical methods.

We also examine frictionless indentation of a circular membrane. A rigid circular indentor of radius $r_i$ is presumed to have no frictional stress between it and that portion of the membrane in contact with the indentor. The flat membrane in contact with the indentor is thus subjected to in-plane stresses and for a given indentation state a point at the indentor edge, denoted by $r=r_c$, is defined by $u(r_c)+r_c=r_i$.

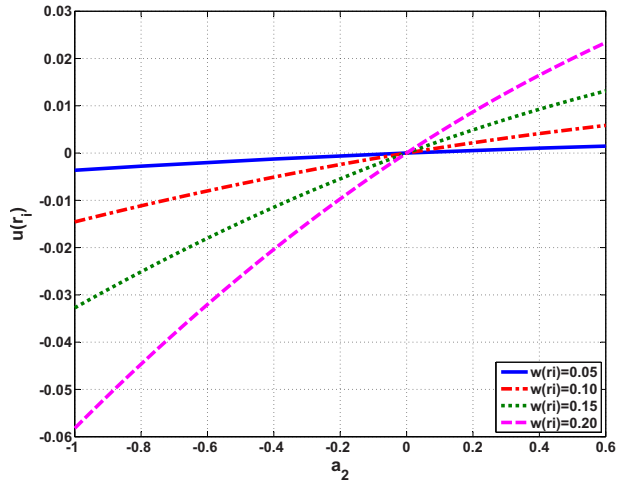



**Fig. 5 Integration constant $a_2$ versus horizontal displacement at inner boundary $u(r_i)$ for various vertical displacements $w(r_i)$ with Poisson's ratio $\nu=\frac{1}{3}$**

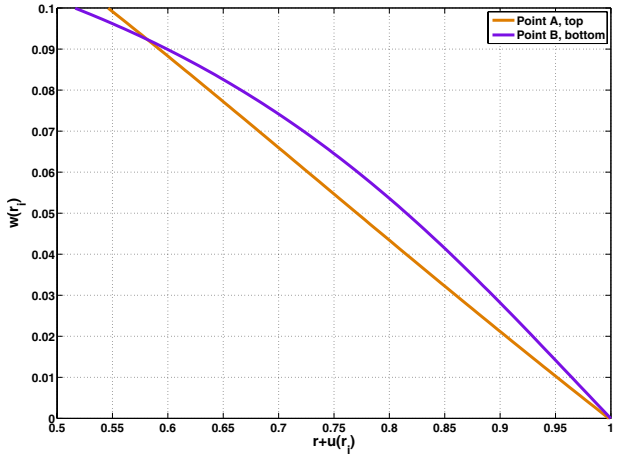



**Fig. 7 Vertical displacements $w(r)$ of points A and B in Fig. 6 versus radial position $r+u(r_i)$**

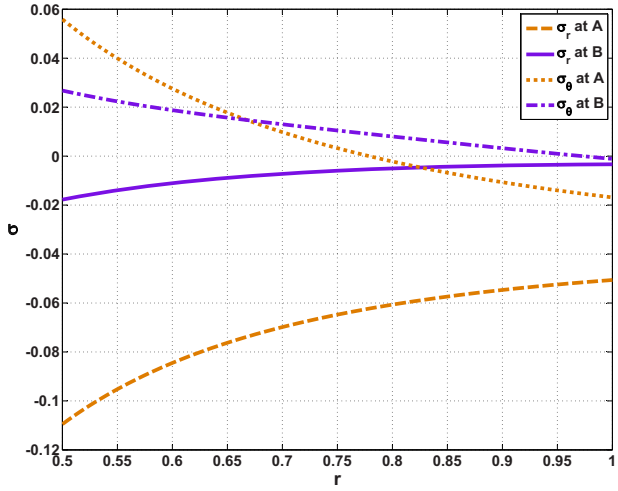



**Fig. 8 Nondimensional radial and circumferential stresses of points A and B of Fig. 6 versus radial position *r***

**Table 1 Comparison of results with Schwerin's (Schwerin's solutions are shown in italic)**

| $\xi=(r/r_a)^2$ | | 0.25 | 0.284 | 0.362 | 0.557 | 0.6755 | 0.8095 | 1.000 |
|---|---|---|---|---|---|---|---|---|
| Results | $\bar{\sigma}_r$ | 2.630 | 2.515 | 2.311 | 1.989 | 1.860 | 1.747 | 1.622 |
| Schwerin's Ref. | | *2.630* | *2.510* | *2.310* | *1.990* | *1.860* | *1.750* | *1.620* |
| Results | $\bar{\sigma}_\theta$ | 0.789 | 0.759 | 0.702 | 0.606 | 0.565 | 0.528 | 0.487 |
| Schwerin's Ref. | | *0.790* | *0.760* | *0.700* | *0.610* | *0.565* | *0.530* | *0.490* |
| Results | $\bar{w}$ | 0.678 | 0.628 | 0.527 | 0.326 | 0.226 | 0.126 | 0.000 |
| Schwerin's Ref. | | *0.680* | *0.630* | *0.530* | *0.330* | *0.225* | *0.125* | *0.000* |

**Table 2 Comparison of results with Tuan's**

| | $\sigma_r$ at $r=r_a$ (MPa) | $\sigma_\theta$ at $r=r_a$ (MPa) |
|---|---|---|
| Results | 16.46 | 5.43 |
| Tuan Ref. | 16.02 | 5.30 |

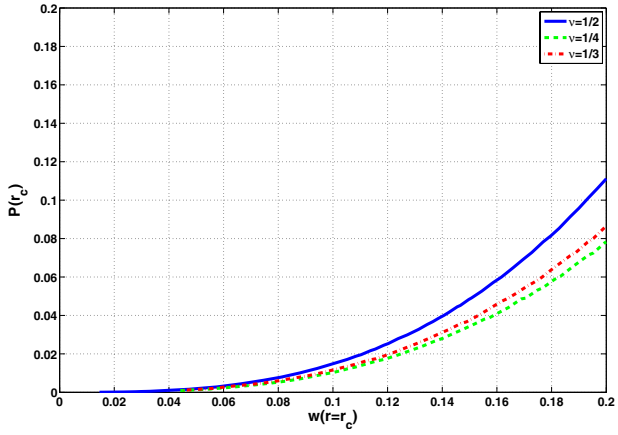



**Fig. 9 Vertical force *P*, applied with frictionless indentor at the inner boundary, versus vertical displacement at inner loaded boundary $w(r_c)$ for different Poisson's ratios with $u(r_c)+r_c=r_i$**

For a given $r_c$ and $r_i$, the radial stress in the contact portion of the radial membrane is determined. This stress is then matched with the inner portion of the unloaded annular membrane. By varying the transition point $r_c$, the resulting vertical force can be determined to provide the results shown in Fig. 9. Comparing Figs. 9 and 2, smaller pressure is needed for the same displacement in the frictionless case, as expected.

## 4 Conclusion

The nonlinear elastic annular membrane problem was solved for the prescribed edge loading. Analytical solutions were provided with particular emphasis on solutions for various values of the constant $a_2$. Vertical and horizontal displacements, and radial and circumferential stresses of the membrane were determined. Bifurcation of solutions occurred for some cases in which $a_2<0$. Also the small slope assumptions of the Föppl theory corresponded to a limited range of the constant $a_2$. Further comparisons were made to the results of Schwerin [14] and Tuan [24] with excellent agreement.


## Acknowledgment

The authors gratefully acknowledge that portions of this research were supported by the Department of Naval Architecture and Marine Engineering, University of Michigan and by the Office of Naval Research Grant No. N00014-03-1-0983, "Marine Research Consortium for Advanced Ship Design."


## References


[1] Föppl, A., 1907, *Vorlesungen uder Technische Mechanic*, Vol. 5, Teubner, Leipzig.
[2] Karman, V., 1910, "Festigkeitsprobleme in Maschinenbau," Encyklopadie Der Mathematischen Wissenschaften, Teubner, IV/4, pp. 311–385.
[3] Way, S., 1934, "Bending of Circular Plates With Large Deflection," Trans. ASME, **56**, pp. 627–633.
[4] Hencky, H., 1915, "Uber den Spannungszustand in Kreisrunden Platten Mit," Zeitschrift fur Mathematik und Physik, **63**, pp. 311–317.
[5] Reissner, E., 1948, "Note on Membrane Theory of Shells of Revolution," J. Math. Phys., **26**, pp. 290–293.
[6] Reissner, E., 1950, "On Axisymmetrical Deformation of Thin Shells of Revolution," Proc. Symp. Appl. Math., **3**, pp. 27–52.
[7] Weinitschke, H., 1980, "On Axisymmetric Deformations of Nonlinear Elastic Membranes," Mechanics Today, S. Nemat-Nasser, ed., Pergamon, Oxford, pp. 523–541.
[8] Dickey, R. W., 1983, "The Nonlinear Circular Membrane Under a Vertical Force," Q. Appl. Math., **41**, pp. 331–338.
[9] Dickey, R. W., 1967, "The Plane Circular Elastic Surface Under Normal Pressure," Arch. Ration. Mech. Anal., **26**, pp. 219–236.
[10] Weinitschke, H., 1987, "On Finite Displacements of Circular Elastic Membranes," Math. Methods Appl. Sci., **9**, pp. 76–98.
[11] Callegari, A. J., and Reiss, E. L., 1968, "Non-Linear Boundary Value Problems for the Circular Membrane," Arch. Ration. Mech. Anal., **31**, pp. 390–400.
[12] Kelkar, A., Elber, W., and Raju, I. S., 1985, "Large Deflections of Circular Isotropic Membranes Subjected to Arbitrary Axisymmetric Loading," Comput. Struct., **21**, pp. 413–421.
[13] Grabmüller, H., and Novak, E., 1987, "Nonlinear Boundary Value Problems for the Annular Membranes: A Note on Uniqueness of Positive Solutions," J. Elast., **17**, pp. 279–284.
[14] Schwerin, E., 1929, "Uber Spannungen und Formanderungen Kreisringformige Membranen," Z. Angew. Math. Mech., **12**, pp. 651–659.
[15] Yang, W., and Hsu, K. H., 1971, "Indentation of a Circular Membrane," ASME J. Appl. Mech., **38**, pp. 227–230.
[16] Chen, D., and Cheng, S., 1996, "Non-Linear Analysis of Prestretched Circular Membrane and a Modified Iteration Technique," Int. J. Solids Struct., **33**, pp. 545–553.
[17] Dienes, J. K., and Miles, J. W., 1977, "A Membrane Model for the Response of Thin Plates to Ballistic Impact," J. Mech. Phys. Solids, **25**, pp. 237–256.
[18] Wierzbicki, T., and Nurick, G. N., 1996, "Large Deflection of Thin Plates Under Localized Impulsive Loading," Int. J. Impact Eng., **18**, pp. 899–918.
[19] Mori, D., David, G., Humphrey, J. D., and Moores, J. E., 2005, "Stress Distribution in a Circular Membrane With a Central Fixation," ASME J. Biomech. Eng., **127**, pp. 549–553.
[20] David, G., and Humphrey, J., 2004, "Redistribution of Stress Due to a Circular Hole in a Nonlinear Anisotropic Membrane," J. Biomech., **37**, pp. 1197–1203.

[21] Scott, O. N., Begley, M. R., Komaragiri, U., and Mackin, T. J., 2004, "Indentation of Freestanding Circular Elastomer Films Using Spherical Indenters," Acta Mater., **52**, pp. 4877–4885.
[22] Pelesko, J., and Chen, X., 2003, "Electrostatic Deflections of Circular Elastic Membranes," J. Electrost., **57**, pp. 1–12.
[23] Begley, M. R., and Mackin, T. J., 2004, "Spherical Indentation of Freestanding Circular Thin Films in the Membrane Regime," J. Mech. Phys. Solids, **52**, pp. 2005–2023.
[24] Tuan, C. Y., 1998, "Ponding on Circular Membranes," Int. J. Solids Struct., **35**, pp. 269–283.
[25] Beck, A., and Grabmüller, H., 1992, "Wrinkle-Free Solutions of Circular Membrane Problems," Z. Angew. Math. Phys., **43**, pp. 481–504.
[26] Callegari, A. J., Reiss, E. L., and Keller, H. B., 1971, "Membrane Buckling: A Study of Solution Multiplicity," Commun. Pure Appl. Math., **24**, pp. 499–516.
[27] MATLAB, 2006, The MathWorks Inc.

# An Inviscid Solution for Modeling of Tornadolike Vortices

**Zhuyun Xu**
Vipac Engineers & Scientists Ltd.,
279 Normanby Road,
Port Melbourne VIC 3207, Austrailia
e-mail: zhuytunx@vipac.com.au

**Horia Hangan**
The Boundary Layer Wind Tunnel Laboratory,
The University of Western Ontario,
London, ON, N6A 5B9, Canada
e-mail: hmh@blwtl.uwo.ca

*An inviscid tornadolike vortex is analytically modeled using a free narrow jet solution combined with a modified Rankine vortex. An empirical and simplified solution to existing models is defined for flows similar to the ones simulated in Ward-type vortex chambers. Velocity profiles are calculated for a particular swirl ratio $S_r=0.28$. The model shows reasonable agreement with existing experimental measurements by Baker (1981, "Boundary Layers in Laminar Vortex Flows," Ph.D. thesis, Purdue University, West Lafayette, IN) and the numerical simulation by Wilson and Rotunno (1986, "Numerical Simulation of a Laminar End-Wall Vortex and Boundary Layer," Phys. Fluids, **29**(12), pp. 3993–4005).* [DOI: 10.1115/1.3063632]

*Keywords: tornadolike vortex, analytical model*

## 1 Introduction

Analytical models of tornadolike vortices are attractive alternatives to experimental and numerical simulations due to their relative simplicity and therefore use in desk-top engineering oriented solutions. There are several simple analytical vortex models available in literature [1]. The simplest analytical model cited by Ward [2] is a sink-vortex expressed by

$$(u+v)r = A + B = \text{const} \tag{1}$$

where $u$ and $v$ are the radial and swirl (or tangential) components of the velocity and $r$ is the radius of the vortex. This model can also be interpreted as $ur=A$ and $vr=B$. The product $ur$ is constant due to continuity in a two-dimensional flow and $vr$ is constant due to conservation of angular momentum. Clearly, the updraft motion is not included in this model. Another simple model, the modified Rankine vortex, reasonably approximates the swirl motion of the vortex:

$$v(r) = \frac{\Gamma}{\pi}\frac{r}{r_c^2 + r^2} \tag{2}$$

where $\Gamma$ is the circulation considered as a constant and $r_c$ is the radius of the vortex core.

Moreover, a three-dimensional model of the vortex was developed by Kuo [3]. The solution for the tangential velocity is similar to the laboratory model developed by Ying and Chang [4]. These solutions were further simplified and adopted to a wind engineering oriented study by Wen [5].

The Burgers–Rott model [6,7] is a more sophisticated model combining an inviscid stagnation flow (away from the wall, motion in the $r$-$w$ plane) and a decayed line vortex (swirl direction) to determine the three velocity components:

$$u = -Cr, \quad w = 2Cz, \quad v = \frac{\Gamma}{2\pi r}(1 - e^{-Cr^2/2\nu}) \tag{3}$$

where $C$ is a strength constant and $\nu$ is the kinematic viscosity. There are also some solutions combining vortex sinks (motion in the $x$-$y$ plane) and axial flow ($w$ direction). Shtern et al. [8] derived a class of solutions and applications including the vortex flow field modeling. Their streamline equation reads

$$z/r_0 = z_0/r_0 + a(r/r_0)^2 + b(r/r_0)^{\mathrm{Re}+2} + c(r/r_0)^4 \tag{4a}$$

$$\phi = \phi_0 + S\ln(r/r_0) \tag{4b}$$

where $a=W_c(2\ \mathrm{Re})^{-1}$, $b=W_r[\mathrm{Re}(\mathrm{Re}+2)]^{-1}$, $c=W_p(4\ \mathrm{Re})^{-1}$, and the swirl parameter $S=\Gamma/\mathrm{Re}$. There are five dimensionless parameters: Re (Reynolds number), $\Gamma$ (swirl Reynolds number), $W_c$ (a free parameter characterizing the uniform part of the axial flow), and $W_r$ and $W_p$ (characterizing the nonuniform shear of the axial velocity included by the axial advection and the radial pressure gradient, respectively). Lewellen and Lewellen [9,10] derived an idealized analytical model to explore the region of corner vortex flow and to explain the near surface intensification of the vortices. Moreover, a multitude of numerical simulations (e.g., Wilson and Rotunno [11] and Lewellen et al. [12,13]) and generic experiments of the vortices in tornado vortex chambers (TVCs) (e.g., Ward [2], Davies-Jones [14], Church et al. [15], and Baker [16]) were performed to determine the flow field associated with the vortices.

Herein, we combine a free narrow jet model and a modified Rankine vortex to provide simple algebraic expressions for the vortex model. Then, we use existing experimental measurements and numerical results to approximate the model's constants and to obtain physically fitted solutions. The obtained model is intended to provide more robust but compact solutions for engineering applications.

## 2 A Simple Inviscid Jet-Vortex Model

The vortex flow field is a complex three-dimensional flow field. A model of a jet away from a wall describes a two-dimensional motion (radial and axial), while a combined Rankine vortex describes essentially a one-dimensional swirl (tangential) motion. A three-dimensional vortex flow can therefore be constructed as a combination of a two-dimensional jet flow and a swirl flow modeled separately.

### 2.1 Motion in $r$-$z$ Plane

*2.1.1 Stokes' Stream Function.* At first, all length scales are nondimensionalized by the vortex radius at the influx boundary $R_{\mathrm{ref}}$ and all velocities are nondimensionalized by the radial influx velocity $u_{\mathrm{ref}}$ at the radius $R_{\mathrm{ref}}$. With $^*$ denoting the dimensional variables, we define the following:

$$r = \frac{r^*}{R_{\mathrm{ref}}}, \quad z = \frac{z^*}{R_{\mathrm{ref}}}, \quad u = \frac{u^*}{u_{\mathrm{ref}}}, \quad w = \frac{w^*}{u_{\mathrm{ref}}}, \quad v = \frac{v^*}{u_{\mathrm{ref}}} \tag{5}$$

According to Schlichting [17], Stokes' stream function expressed by

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received September 25, 2007; final manuscript received August 14, 2008; published online March 11, 2009. Review conducted by Nesreen Ghaddar.

$$\psi(z,r) = Dzg\left(\frac{r}{z}\right) \tag{6}$$

can be used to describe a flow similar to an upward narrow free jet. Stokes' stream function for a simple Gaussian flow can be expressed by

$$g(r) = \tfrac{1}{2}(1 - e^{-r^2}) \tag{7}$$

Using Eq. (7) as a function of the similarity variable $r/z$ in Eq. (6), we obtain

$$\psi(z,r) = Dz(1 - e^{-(r/z)^2}) \tag{8}$$

Equation (8) satisfies two boundary conditions: $\psi=0$ at $r=0$ and at $z=0$. With the influx condition

$$u = -1 \quad \text{at} \quad r = 1 \tag{9}$$

the constant $D$ in Eq. (8) is determined ($D=-1$). Then we obtain

$$\psi(z,r) = z(e^{-(r/z)^2} - 1) \tag{10}$$

*2.1.2 Velocities*. The radial and axial velocities can be directly derived from Eq. (10) and expressed by

$$u(z,r) = \frac{1}{r}\frac{\partial \psi}{\partial z} = \left(\frac{1}{r} + \frac{2r}{z^2}\right)e^{-(r/z)^2} - \frac{1}{r} \tag{11a}$$

$$w(z,r) = -\frac{1}{r}\frac{\partial \psi}{\partial r} = \frac{2}{z}e^{-(r/z)^2} \tag{11b}$$

**2.2 Swirl Motion.** A combined Rankine vortex expressed by Eq. (2) can be used to describe the swirl motion of the vortex. According to Ward [2], the radius of the vortex core, $r_c$, can be expressed by

$$r_c = r_0 \sin^2 \theta \tag{12}$$

where $r_0$ is the radius at the influx boundary of the control volume and $\theta$ is the influx angle expressed as

$$\theta = \arctan\left(\frac{v_0}{u_0}\right) \tag{13}$$

where $u_0$ and $v_0$ are the radial and tangential components of velocity at the influx boundary. The most important nondimensional parameter, the swirl ratio ($S_r$), is defined as

$$S_r = \frac{v_0}{2u_0} \tag{14}$$

Then, the radius of the vortex core becomes

$$r_c = r_0 \sin^2(\arctan(2S_r)) = \sin^2(\arctan(2S_r)) \tag{15}$$

where $r_0=1$ is used at the influx boundary. Substituting Eq. (15) into Eq. (2), and keeping in mind that for $r=1$, $u_0=1$, we obtain

$$v(r) = 2S_r(r_c^2 + 1)\frac{r}{r_c^2 + r^2} \tag{16}$$

Observations and numerical simulations [2,11] suggest that the radius of the vortex core, $r_c$, is not constant with height. Considering a funnel-shaped vortex, the core radius will become zero at the funnel tip. Assuming the funnel tip to be at the origin, a function of $z$, $f(z) \sim [0,1]$, is required to transform a cylindrical swirl vortex to a funnel-shaped vortex:

$$f(z) = \frac{z^m}{z^m + H} \tag{17}$$

where $m$ and $H$ are empirical constants. The height $z$ at which the full core is reached increases with $H$. Following a best fit of Baker's test data [11], the two empirical parameters in Eq. (17) were set at $m=0.05$ and $H=4$. Then, the core radius of the vortex funnel can be expressed by

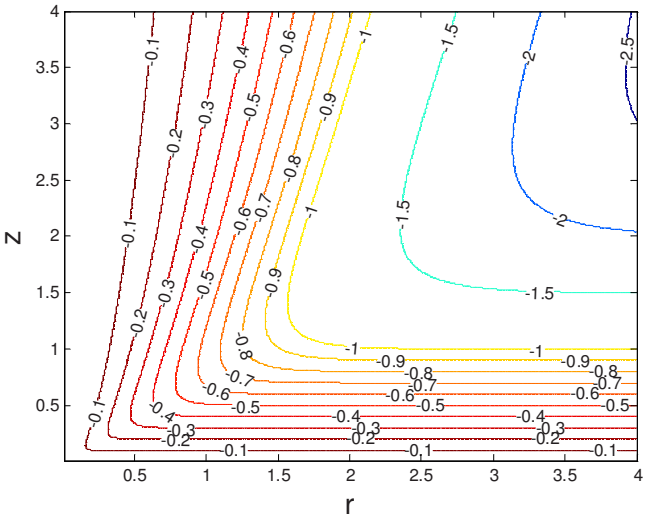



**Fig. 1 Contours of Stokes' stream function, which represent the modeled streamlines of the vortex (the values are −0.1, −0.2, −0.3, −0.4, −0.5, −0.6, −0.7, −0.8, −0.9, −1, −1.5, −2, and −2.5)**

$$r_c(z) = \frac{z^{0.05}}{z^{0.05} + 4}\sin^2(\arctan(2S_r)) \tag{18}$$

And Eq. (16) becomes

$$v(z,r) = 2S_r(r_c^2(z) + 1)\frac{r}{r_c^2(z) + r^2} \tag{19}$$

Equations (11) and (19) then describe the three-dimensional flow field of a vortex.

### 2.3 Flow Field Calculation Results

*2.3.1 Surface Flow*. As $z$ and $r_c$ tend to 0, a limit inviscid result can be obtained from Eqs. (11) and (19) as follows:

$$u(0,r) = -\frac{1}{r} \quad w(0,r) = 0 \quad v(0,r) = \frac{2S_r}{r} \tag{20}$$

Here the model recovers an exact sink-vortex in Eq. (1) with $a=-1$ and $b=2S_r$.

*2.3.2 Flow Field in a Finite Domain*. Stokes' stream function, vorticity, and the three components of velocity can be computed in the flow domain. In order to compare the model with available experimental data [11], a swirl ratio $S_r=0.28$ is chosen. The vorticity equation can be derived as

$$\Omega = \frac{1}{r}\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial r}\right) = \frac{2r}{z^3}e^{-(r/z)^2}\left(1 + \frac{2r^2}{z^2}\right) \tag{21}$$

A convergent layer is presented near the ground and a divergent upward flow above that are clearly revealed by Stokes' stream function (Fig. 1). The highest vorticity is concentrated near the origin in the proximity of the singularity (Fig. 2). Away from the axes, vorticity decreases rapidly. The axial velocity (Fig. 3) decreases with increasing height $z$. Radial velocities (Fig. 4) show two distinct regions: negative values at low levels, representing convergent inflow toward the vortex core, and positive values above associated with divergent outflow from the vortex. The radius of the free vortex in the center of Fig. 5 increases gradually and reaches a limit value corresponding to Eq. (16) as $z$ increases. As shown in Fig. 6, an approximate unity profile ($|u|=1$) is presented up to $z=0.4$. The influx velocity profile resulting from a modified model (see Sec. 3) is also shown.

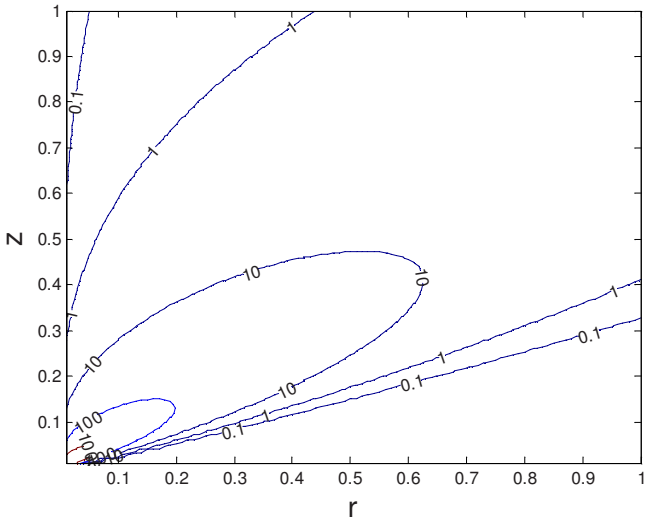



**Fig. 2 Vorticity contours (the values from the core to outside are 1000, 100, 10, 1, and 0.1)**

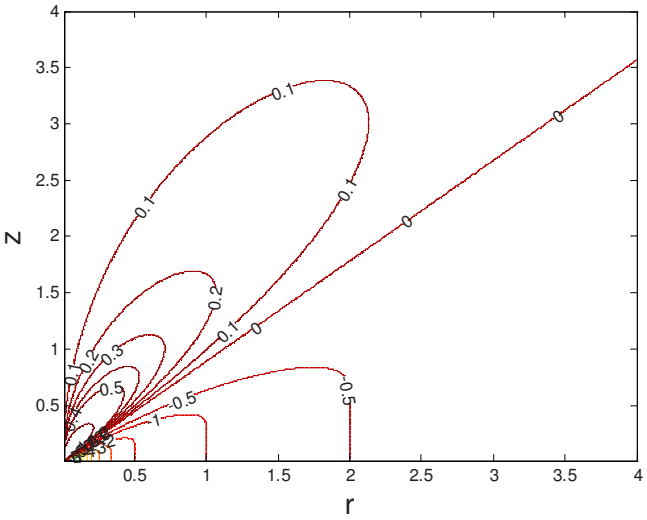



**Fig. 4 Contours of radial velocities (the values from the lower are −10, −5, −4, −3, −2, −1, −0.5, 0, 0.1, 0.2, 0.3, 0.4, 0.5, and 1)**

## 3 Flow Field Modeling of a Ward-Type Vortex Chamber

**3.1 Empirical Model.** In order to model the flow field in the inviscid region of a Ward-type vortex chamber, we consider a weighted combination of two similarity functions (the first and second powers of the similarity parameter $(r/z)$):

$$\psi(z,r) = WA_1 z^{B_1}(e^{-C_1(r/z)} - 1) + (1 - W)A_2 z^{B_2}(e^{-C_2(r/z)^2} - 1) \tag{22}$$

where $W$ is a weighting function. The constants $A_1$, $A_2$, $B_1$, $B_2$, $C_1$, and $C_2$ can be found by a least squared error fit with experimental or numerical measurements. Herein the fit is based on the test data by Baker and the numerical results by Wilson and Rotunno [11] providing three radial profiles at $z$=0.075, 0.25, and 0.625 and four vertical profiles at $r$=0.0475, 0.1025, 0.2125, and 0.75. This gives $[A_1, A_2, B_1, B_2, C_1\ C_2]$=[0.8, 4.6, 1.2, 2.2, 0.39, 0.62] leading to the following empirical model of the flow field:

$$\psi(z,r) = 0.8Wz^{1.2}(e^{-0.39(r/z)} - 1) + 4.6(1 - W)z^{2.2}(e^{-0.62(r/z)^2} - 1) \tag{23}$$

$$u(z,r) = \frac{1}{r}\frac{\partial \psi}{\partial z} = 0.8W\left(\frac{1.2z^{0.2}}{r}(e^{-0.39(r/z)} - 1) + \frac{0.39}{z^{0.8}}e^{-0.39(r/z)}\right) + 4.6(1 - W)\left(\frac{2.2z^{1.2}}{r}(e^{-0.62(r/z)^2} - 1) + \frac{1.24r}{z^{0.8}}e^{-0.62(r/z)^2}\right) \tag{24}$$

$$w(z,r) = -\frac{1}{r}\frac{\partial \psi}{\partial r} = 0.31W\frac{z^{0.2}}{r}e^{-0.39(r/z)} + 5.7(1 - W)z^{0.2}e^{-0.62(r/z)^2} \tag{25}$$

$$v(z,r) = 2S_r\frac{r(r_c^2(z) + 1)}{r_c^2(z) + r^2}\frac{z}{z + 0.05} \tag{26}$$

Equation (26) is obtained by multiplying Eq. (19) with an empirical fitted function of [0–1] based on the two data sets used.

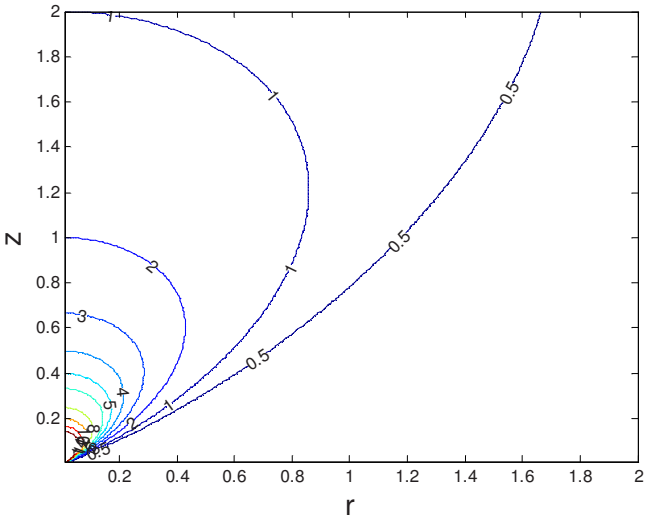



**Fig. 3 Contours of axial velocities (the values are 64, 32, 8, 6, 5, 4, 3, 2, 1, and 0.5)**

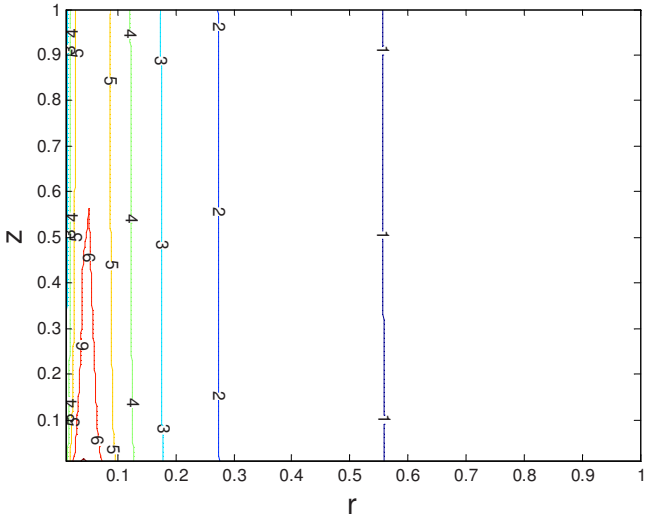



**Fig. 5 Contours of typical swirl velocities ($S_r$=0.28) (the values from the higher are 6, 5, 4, 3, 2, and 1)**

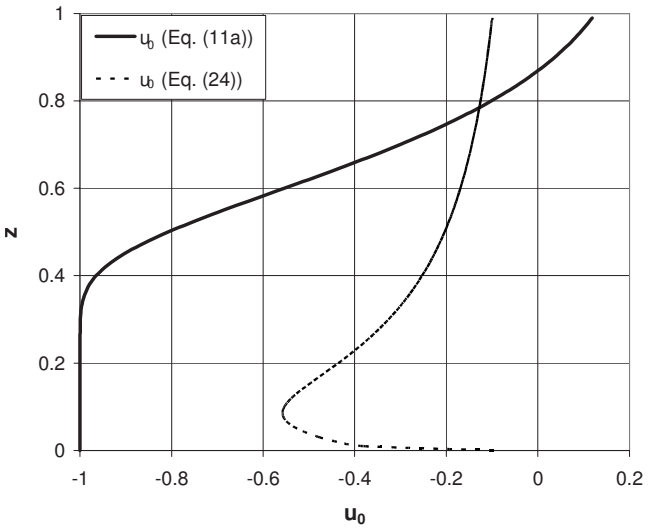



**Fig. 6 Influx radial velocity profiles at $r=1$**

Alternatively, a power or logarithmic law boundary layer function can be used to fit the data. The weighting function is chosen such that the second power term dominates in the core region and disappears for larger radius,

$$W = 1 - e^{-(Kr/r_c)} \tag{27}$$

$K=0.9$ is used in the above fitting.

**3.2 Flow Field Calculation Results and Comparison.** Comparisons of velocity profiles at various radial positions and heights are shown in Figs. 7 and 8, respectively. Figures 7(*a*)–7(*d*) compare the velocity variation versus $z$ at several radial positions. The thickness of the boundary layer in the experiment is about $z=0.025$ slightly increasing radially. The empirical model matches well with both test and numerical simulation. The comparison is better for small radii (Figs. 7(*a*)–7(*c*)) than for the larger radius (Fig. 7(*d*)), indicating a better fit for the core region than for the outer region. This can be overcome if more data sets for the outer region would be used to fit the model constants or by using a weighting function on the experimental input data. Figure 8 compares the velocity variation with radius for three heights ($z=0.075$, 0.25, and 0.625). Overall, the comparisons for all the three velocity components are reasonable both for the tendency and the magnitudes.

## 4 Concluding Remarks and Discussions

An inviscid solution of the vortex is obtained by combining a narrow free jet and a modified Rankine vortex. The two solutions are connected by the dominant flow characteristic, the swirl ratio $S_r$. The model's constants are modified using a least squared error fitting based on the experimental measurements by Baker and CFD simulation by Wilson and Rotunno [11]. The model-based velocity components are compared with the experimental data and CFD output, which indicates reasonable agreement. This model has inherent limitations related to its inviscid character and its dependency on available data sets. Experimental inputs (particularly for different swirl ratios) are expected to improve the robustness of the proposed model.

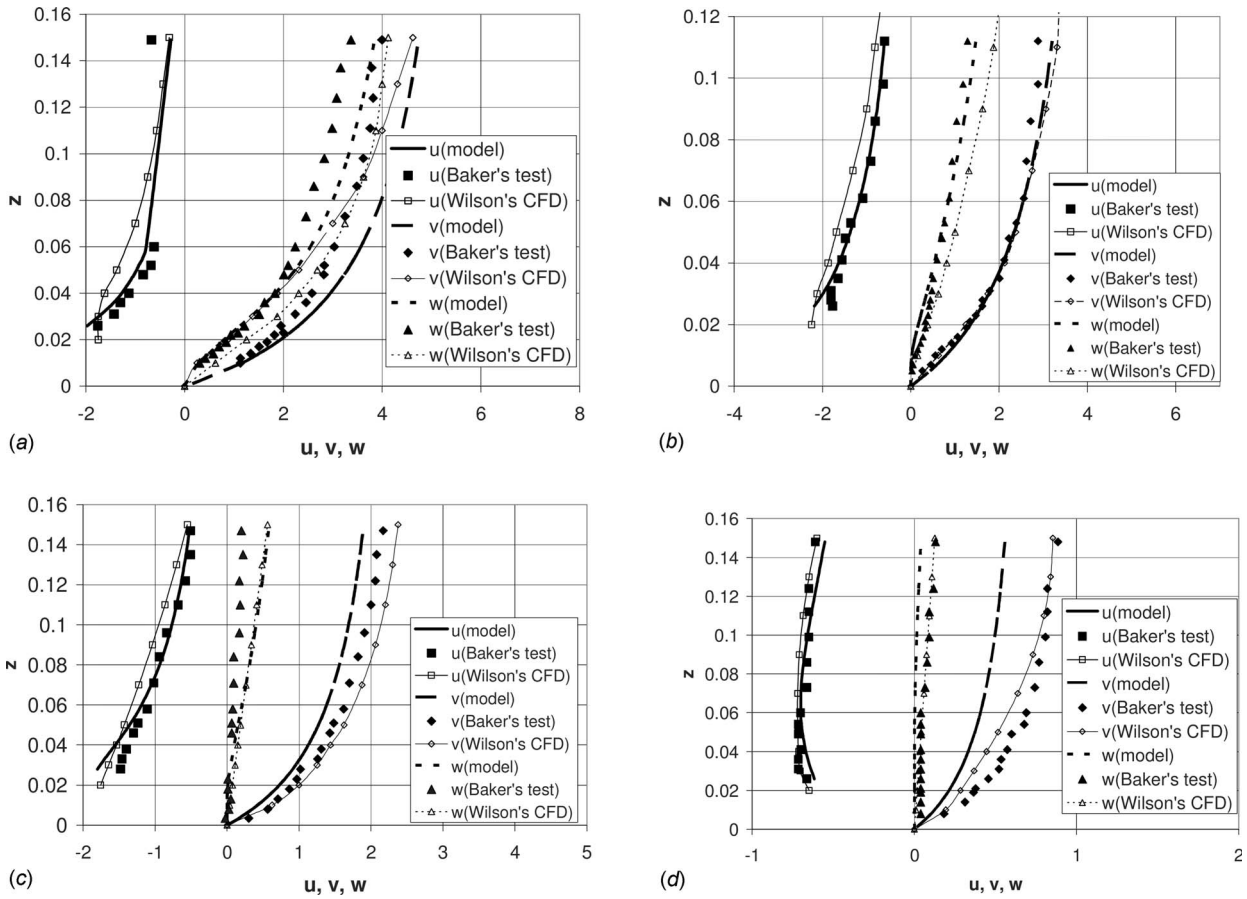



**Fig. 7 Comparisons of velocity profiles versus height at $r=0.75$ (modified model, Baker's test and Wilson and Rotunno's numerical simulation): (*a*) $r=0.0475$, (*b*) $r=0.1025$, (*c*) $r=0.2125$, and (*d*) $r=0.75$**

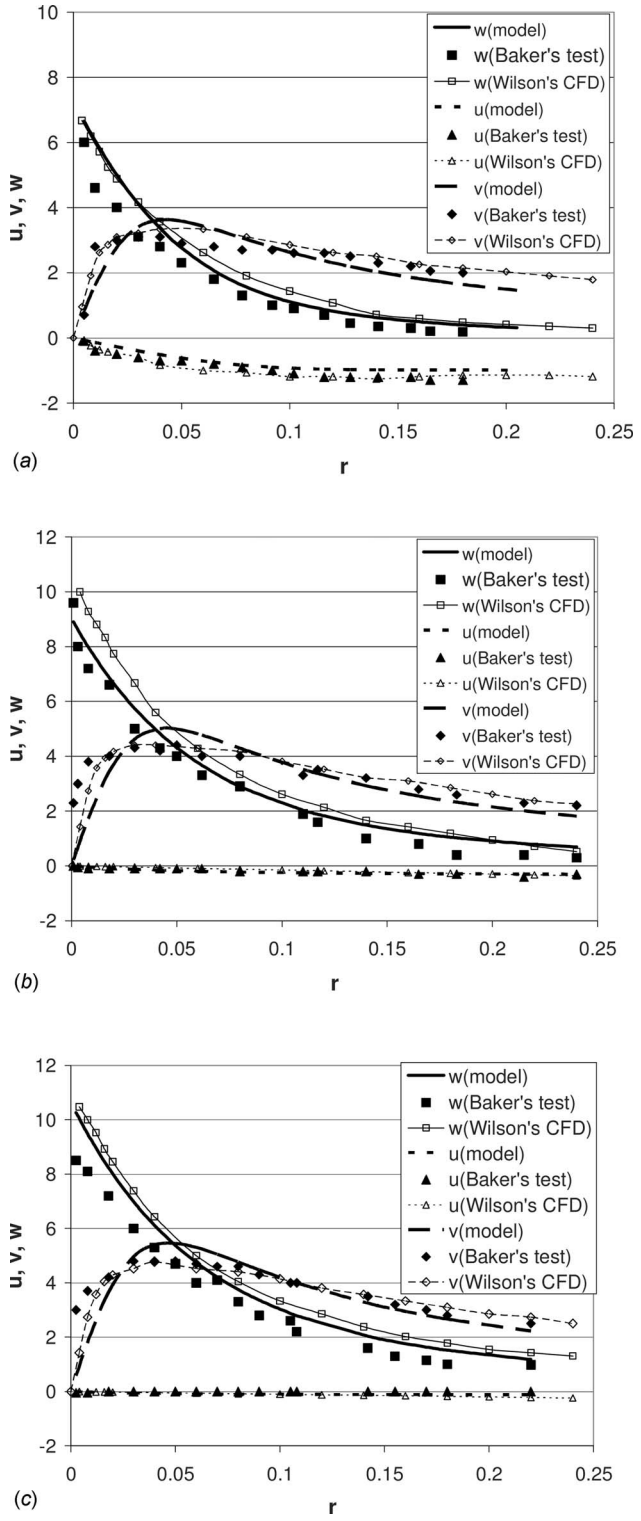



**Fig. 8 Comparison of velocity variation versus $r$ (modified model, Baker's test and Wilson and Rotunno's CFD): ($a$) $z=0.075$, ($b$) $z=0.25$, and ($c$) $z=0.625$**

Wen's simplified equations [5] set $u$ component to zero for the zone out of the viscous boundary layer, which is not suitable to describe the wind velocity profiles in a Ward-type vortex chamber. The Burgers–Rott model [6,7] described the $u$ and $w$ components with the simple stagnation flow equations (see Eq. (3)), without matching the experimental profiles for $u$ and $w$. The present model provides more flexibility to fit all the three velocity components to experimental or numerical wind fields.

## Acknowledgment

This work was made possible through funding provided by Natural Sciences and Engineering Research Council of Canada (NSERC) Grant No. 166732 and funding from Manitoba Hydro. Thanks are also due to Dr. David Surry for providing valuable suggestions.

## Nomenclature

$C$ = strength constant
$H$ = experimental fitted parameter
$m$ = experimental fitted parameter
$R_{ref}$ = vortex reference radius, m
$r^*$ and $z^*$ = radial and axial coordinates, m
$r$, $\theta$, and $z$ = dimensionless radial, azimuthal, and axial coordinates; $r=r^*/R_{ref}$, $z=z^*/R_{ref}$
$r_c$ = radius of the vortex core
$r_0$ = dimensionless radius at influx boundary, $r_0=1$
$S_r$ = swirl ratio
$u^*$, $v^*$, and $w^*$ = radial, azimuthal, and axial velocities, m/s
$u_{ref}$ = radial influx velocity at the inlet of the Ward-type vortex chamber, m/s
$u$, $v$, and $w$ = dimensionless radial, azimuthal, and axial velocities, $u=u^*/u_{ref}$, $w=w^*/u_{ref}$
$u_0$ = radial influx velocity at $r=1$

### Greek Symbols

$\nu$ = kinematical viscosity, $m^2/s$
$\Gamma$ = circulation of vortex
$\psi$ = Stokes' stream function
$\Omega$ = vorticity function

## References

[1] Lewellen, W. S., 1993, "Tornado Vortex Theory," *The Tornado: Its Structure, Dynamics, Prediction, and Hazards*, C. Church, D. Burgess, C. Doswell, and R. Davies-Jones, eds., American Geophysical Union, Washington, DC, pp. 19–39.
[2] Ward, B. N., 1972, "The Exploration of Certain Features of Tornado Dynamics Using a Laboratory Model," J. Atmos. Sci., **29**, pp. 1194–1204.
[3] Kuo, H. L., 1971, "Axisymmetric Flows in the Boundary Layer of a Maintained Vortex," J. Atmos. Sci., **28**(1), pp. 20–41.
[4] Ying, S. J., and Chang, C. C., 1970, "Exploratory Model Study of Tornado-Like Vortex Dynamics," J. Atmos. Sci., **27**(1), pp. 3–14.
[5] Wen, Y. K., 1975, "Dynamic Wind Loads on Tall Buildings," J. Struct. Div., **101**( *ST1*), pp. 169–185.
[6] Burgers, J. M., 1948, "A Mathematical Model Illustrating the Theory of Turbulence," Adv. Appl. Mech., **1**, pp. 197–199.
[7] Rott, N., 1958, "On the Viscous Core of a Line Vortex," Z. Angew. Math. Mech., **9**, pp. 543–553.
[8] Shtern, V., Borissov, A., and Hussain, F., 1997, "Vortex Sinks With Axial Flow: Solution and Applications," Phys. Fluids, **9**(10), pp. 2941–2959.
[9] Lewellen, D. S., and Lewellen, W. S., 2007, "Near-Surface Intensification of Tornado Vortices," J. Atmos. Sci., **64**, pp. 2176–2194.
[10] Lewellen, D. S., and Lewellen, W. S., 2007, "Near-Surface Vortex Intensification Through Corner Flow Collapse," J. Atmos. Sci., **64**, pp. 2195–2209.
[11] Wilson, T., and Rotunno, R., 1986, "Numerical Simulation of a Laminar End-Wall Vortex and Boundary Layer," Phys. Fluids, **29**(12), pp. 3993–4005.
[12] Lewellen, W. S., Lewellen, D. C., and Sykes, R. I., 1997, "Large-Eddy Simulation of a Tornado's Interaction With the Surface," J. Atmos. Sci., **54**, pp. 581–605.
[13] Lewellen, D. C., Lewellen, W. S., and Xia, J., 2000, "The Influence of a Local Swirl Ratio on Tornado Intensification Near the Surface," J. Atmos. Sci., **57**, pp. 527–544.
[14] Davies-Jones, R. P., 1973, "The Dependence of Core Radius on Swirl Ratio in a Tornado Simulator," J. Atmos. Sci., **30**, pp. 1427–1430.
[15] Church, C. R., Snow, J. T., and Agee, E. M., 1977, "Tornado Vortex Simulation at Purdue University," Bull. Am. Meteorol. Soc., **58**, pp. 900–908.
[16] Baker, G. L., 1981, "Boundary Layers in Laminar Vortex Flows," Ph.D. thesis, Purdue University, West Lafayette, IN.
[17] Schlichting, H., 1933, "Laminare Strahlenausbreitung," Z. Angew. Math. Mech., **13**, pp. 260–263.

**Craig A. Steeves**
Postdoctoral Scholar
Department of Materials,
University of California, Santa Barbara,
Santa Barbara, CA 93106

**Katherine H. Timpano**
Graduate Student
Department of Mechanical and
Aerospace Engineering,
Princeton University,
Princeton, NJ 08544

**Peter T. Maxwell**
Senior Development Engineer
Department of Materials,
University of California, Santa Barbara,
Santa Barbara, CA 93106

**Luigi Martinelli**
Associate Professor

**Richard B. Miles**
Professor

Department of Mechanical and
Aerospace Engineering,
Princeton University,
Princeton, NJ 08544

# Design and Manufacture of a Morphing Structure for a Shape-Adaptive Supersonic Wind Tunnel Nozzle

*Aerospace vehicles with fixed geometry are designed to operate at a predetermined flight condition. Variation of the aerodynamic environment, such as during acceleration, climbing, or turning, from the design condition reduces the efficiency of the vehicle. It would be advantageous to be able to adapt the vehicle geometry to maintain efficient flight over a range of aerodynamic conditions. Morphing sandwich structures offer sufficient strength and stiffness to serve as aerodynamic surfaces, while providing the shape-changing authority to attain a range of surface profiles without additional joints or seals. As a demonstration of the morphing concept in a supersonic environment, this paper describes the construction and testing of a morphing nozzle for a supersonic wind tunnel, which has been designed to operate isentropically over a Mach range from 2.5 to 3.8. The nozzle has been installed and operated in this Mach number range and the experimental results are presented.* [DOI: 10.1115/1.3005572]

## 1 Introduction

The performance of an air-breathing hypersonic engine, either a ramjet or a scramjet, is very sensitive to the aerodynamic conditions at the inlet, which involve multiple shocks, high heating, and large aerodynamic forces. Air is compressed and accelerated through a shock system forming on the vehicle forebody and at the entrance to the inlet. The shock structure must accelerate the air to the required velocity while maintaining the total pressure and minimizing the turbulence and boundary layer growth and separation. Due to this complexity, a fixed-geometry inlet cannot be optimally shaped for all flight conditions and hence will lose efficiency when the vehicle is not flying at the designed combination of Mach number, angle of attack, and altitude. Deleterious shock structures, such as a Mach reflection, may also form and choke the inlet. Boundary layer separation due to the inclement operating circumstances results in loss of total pressure at the inlet. These problems, and the resulting loss of efficiency, may be remedied by employing a shape-adaptive inlet, which can change its geometry to maintain an optimal shape for the changing flight conditions. However, present systems for shape changing involve complex joints, seals, and gaps, which are difficult to construct and have a negative impact on the local fluid dynamics.

Shape morphing concepts based on sandwich structures with actuated core members have continuous surfaces, which eliminate the problems associated with joints and gaps. Moreover, these structures can be morphed to achieve profiles with continuous curvature, which is important for shock control. Such morphing structures were proposed by [1] and demonstrated by [2]. The cellular core sandwich panels upon which these structures are based are described in detail in Refs. [3–5].

A shape-changing nozzle for a supersonic wind tunnel is a demonstration of the capabilities of morphing sandwich structures in an aerodynamic context. The ability of a morphing surface to conform to an aerodynamic shape, which is sensitive to small errors, will be shown. Furthermore, the morphing nozzle is a practical demonstration of a ramjet inlet operating in reverse: That is, the nozzle produces an isentropic acceleration from subsonic to supersonic velocity, while the ramjet inlet decelerates air from supersonic to subsonic velocity (in the reference frame of the vehicle). In order to apply shape morphing theories to aerodynamic systems, a morphing wind tunnel nozzle has been built; see Fig. 1.

This paper is organized as follows: First, the aerodynamic profiles needed for the wind tunnel will be explicated. A discussion of the analytical background necessary to fit the morphing structure to desired aerodynamic shapes follows. The third component describes the manufacture of the morphing structure and installation in the existing wind tunnel. Fourth, experimental results of the preliminary assessment of the wind tunnel performance will be reported. Finally, some limitations of these techniques are enumerated.

## 2 Determination of the Aerodynamic Shapes

Computational fluid dynamics (CFD) was used to determine ideal aerodynamic profiles for shock-free isentropic expansion at several Mach numbers. An inviscid code was used to determine the surface geometry, which was then modified for boundary layer thickness. The adjoint optimization methods used to calculate the optimal shapes are described in Refs. [6,7]. The nozzle shapes were designed with fixed exit dimensions, allowing the throat area to vary between Mach numbers. The reduction in effective exit

Contributed by the Applied Mechanics Division for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received September 10, 2007; final manuscript received April 29, 2008; published online March 11, 2009. Review conducted by Robert M. McMeeking.

**Fig. 1 A photograph of the morphing nozzle as installed in the in-draft wind tunnel. Visible are the curved morphing surface, the actuators, and the linear potentiometers used to control the system position.**

area due to boundary layer growth on the curved surface was accounted for in these curves by a slight enlargement of the exit area, deviating from the inviscid isentropic throat-to-exit ratio for each Mach number. The aerodynamic calculation was performed only in the supersonic region beyond the throat; as a consequence, the nozzle curvature was mirrored at the throat to produce the subsonic converging nozzle profile. In order to connect the nozzle to a pre-existing wind tunnel ejector system, a demonstration geometry with a 445 mm nozzle length, 50.8 mm width, and 48.5 mm exit height was chosen. All examples described herein utilize this geometry. For the purposes of the current design, nozzle curves were computed for Mach 2.5, Mach 3, and Mach 3.8 flow speeds; these shapes are presented in Fig. 2. The origin in the figure is taken to be at the curved wall at the connection between the nozzle and the diffuser.

## 3 Approximating Aerodynamic Surfaces With Morphing Structures

The structure to be used for the morphing wind tunnel nozzle is based on a corrugated sandwich beam, which has been demonstrated by [8]. As such, it changes shape in one dimension only;

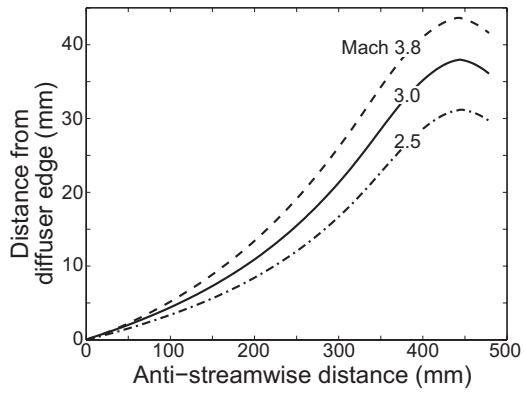



**Fig. 2 The desired shapes for the morphing wind tunnel nozzle. The nozzle is manufactured to conform to the Mach 3 shape and must be morphed to attain the Mach 2.5 and 3.8 shapes. The antistreamwise distance is the length from the connection between the diffuser and the nozzle in the direction opposite the flow. The throat of the nozzle is found at $x=445$ mm.**

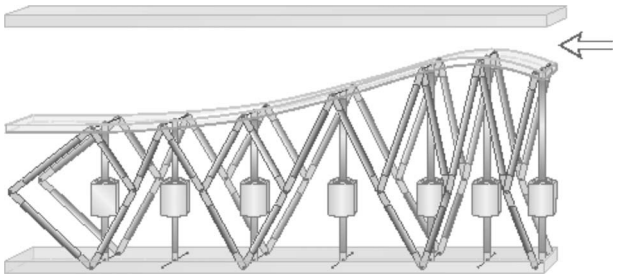


**Fig. 3 Isometric drawing of a morphing sandwich structure with actuators. The curved surface and the rigid back plate constitute the sandwich faces, while the oblique members and the vertical tie rods are the low-density sandwich core. The actuators are the rectangular elements of the vertical tie rods.**

other morphing structures based on Kagome core geometries can change shape in two directions [1,2]. In all cases, the shapes attainable through these morphing concepts involve only bending of the passive surface; that surface is not stretched or compressed. The structure described herein, and shown schematically in Fig. 3, has a core composed of vertical and oblique members, which provide actuation and shear stiffness, respectively, and are pin connected to transmit only axial forces. The aerodynamic surface is a continuous curved beam of streamwise length $\ell$, connected to the actuated and oblique members at unequally spaced nodes. In the neutral position, the surface conforms to the shape for a Mach 3 expansion and can be morphed into the correct profiles for isentropic expansions to other Mach numbers.

With this structural concept, a critical design challenge is to choose the best configuration of the actuators so that the initial structure can morph to approximate most closely the nozzle shapes for operation at other Mach numbers. It is emphasized that the morphing structure can attain only a limited family of shapes because only a finite number of locations are controlled through actuation; within this family, the shape that best achieves the desired function of the nozzle must be selected. This is accomplished by calculating the geometric changes, which can be achieved by small displacements of individual actuators, finding the best combination of actuator displacements for given actuator locations, and then shifting the actuator positions in a manner that improves the surface shape approximation. Because the required shape changes and slopes are small, it is assumed that it is possible to sum linearly the deformations of the morphing surface due to deflections from individual actuators. The desired shape of the aerodynamic surface, calculated by CFD, is defined as $\phi(x)$, where $x$ is distance from the diffuser in the direction opposite the flow (the antistreamwise direction). The ideal shape $\phi(x)$ will be approximated by $\omega(x)$ that belongs to the family of curves which can be obtained by deflections of the actuators in the morphing structure. In this work, an approximation is found in the least-squares sense, which means that the cost function $\psi$,

$$\psi = \int_0^l (\phi - \omega)^2 dx \tag{1}$$

is a minimum. The function $\omega(x)$ is expressed as $\omega(x)=a_i\omega_i(x)$ (utilizing the usual summation convention), where the $\omega_i$ are the linearly independent, and linearly summable, shape functions generated by displacements of the individual actuators. The $\omega_i$ are functions of the actuator configuration, the geometry of the morphing surface, and the boundary conditions, while the $a_i$ are constant coefficients. For the structure displayed in Fig. 3, the shape functions are linear sums of the deflected shapes of cantilever beams fixed at the leftmost end and subjected to a unit displacement at one actuator location (see Fig. 4), while displacements at the other actuator locations are prohibited. A full set of shape

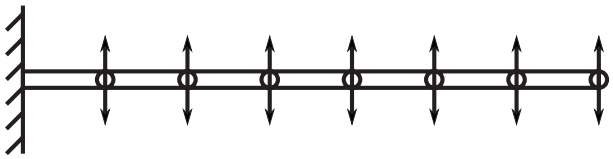


**Fig. 4 A flat cantilever beam with seven evenly spaced independently movable nodes. The shapes attainable by this configuration comprise the family of allowable shapes for the morphing wind tunnel surface.**

functions is shown in Fig. 5 for the demonstration surface with seven equispaced actuators.

Once the $\omega_i$ have been found, the Rayleigh–Ritz method is used to determine the constant coefficients $a_i$ by simultaneously solving the equations

$$\frac{\partial \psi}{\partial a_i} = -\int_0^l 2(\phi - a_j\omega_j)\omega_i dx = 0 \tag{2}$$

This is achieved through the solution of the linear system $b_{ij}a_j = r_i$, where

$$b_{ij} = \int_0^l \omega_i \omega_j dx \tag{3}$$

and

$$r_i = \int_0^l \phi \omega_i dx \tag{4}$$

The solution vector $a_i$ provides the mixture of the $\omega_i$ required to approximate most closely in the least-squares sense the desired surface profile $\phi$ for the given functions $\omega_i$, which depend on the actuator locations. The vector $a_i$ can therefore then be used to determine the optimal displacements of the individual actuators. This procedure assumes that the passive surface obeys elastic beam theory and that the surface is subject only to small deflections such that the surface displacements due to individual actuators can be superposed, and such that the longitudinal deflections of the joints at which the surface is connected to the actuators are small and can be neglected.

The approximation can be improved by varying the locations of the actuators. If the actuator locations are given by $\xi_i$, then the gradients of the cost function $\partial\psi/\partial\xi_i$ are calculated numerically given the present location of the actuators. The positions of the actuators are then adjusted using a path of steepest descent, the cost function is recalculated, and the process is iterated. While this

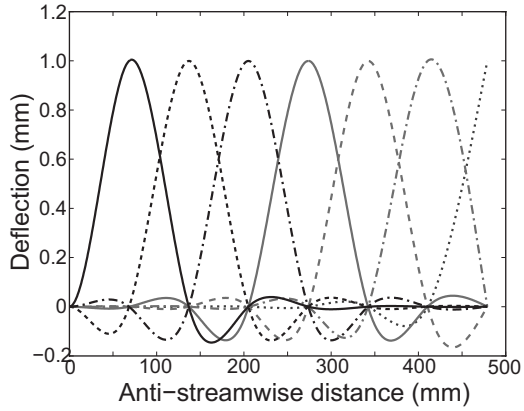



**Fig. 5 A set of shape functions for a morphing aerodynamic surface with seven equally spaced actuators**

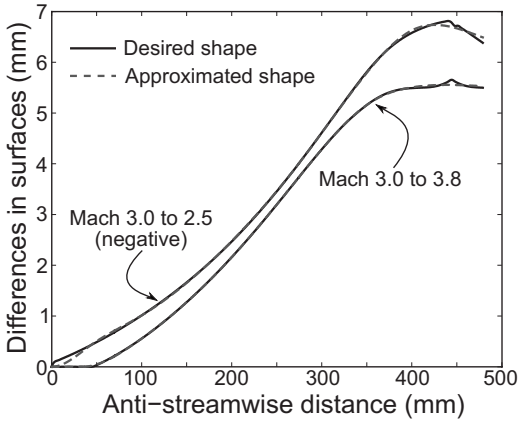



**Fig. 6 The required shape changes, in black, to move between the Mach 3.0 shape and the Mach 2.5 shape (shown in negative here) and the Mach 3.8 shape. The quality of the approximated fit after the actuator locations are improved is given in red.**

is guaranteed to improve the solution and eventually find a local minimum, finding the global minimum is not assured.

The morphing nozzle must change from the initial Mach 3 geometry to both the Mach 2.5 and the Mach 3.8 geometry. The three required shapes are shown in Fig. 2. Because the nozzle is constructed to conform to the Mach 3 shape, the actual morphing needs only to produce the deflections equivalent to the difference between the Mach 3 shape and either the Mach 2.5 or the Mach 3.8 shape. These difference curves are shown in Fig. 6. In both cases, a slight "cusp" is evident at the throat of the nozzle. This is a consequence of the technique used to calculate the curves and the necessity of extending the curves into the subsonic region before the throat. Utilizing the optimization procedure described above, the initial set of shape functions is modified and the quality of the shape change is improved. During the optimization procedure, the lengths of the intervals between actuators evolve from the initial equal spacing of 68 mm. Figure 7 shows the progress of these changes for the shape change from Mach 3.0 to Mach 3.8. The curve labeled "Interval 1" refers to the length of the space between the fixed end of the nozzle and the first actuator. Note that for manufacturing reasons, a minimum actuator spacing of 45 mm has been imposed. After approximately 1000 iterations, a steady state is reached. Comparing these interval lengths with the difference curves in Fig. 6 shows that the actuator spacing is reduced in regions where the curvature of the difference curves is

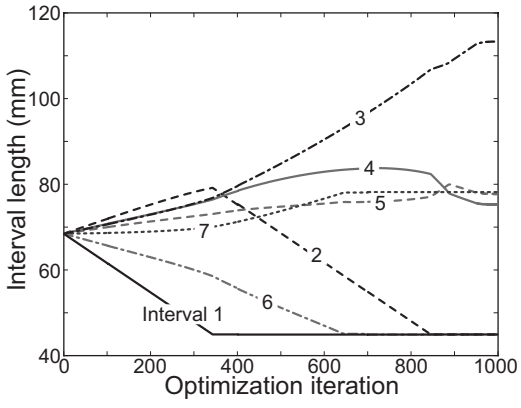



**Fig. 7 The progressive changes in the lengths of the intervals between the seven actuators during the optimization procedure for a change from the Mach 3.0 shape to the Mach 3.8 shape. Interval 1 is between the fixed end of the surface and the first actuator.**

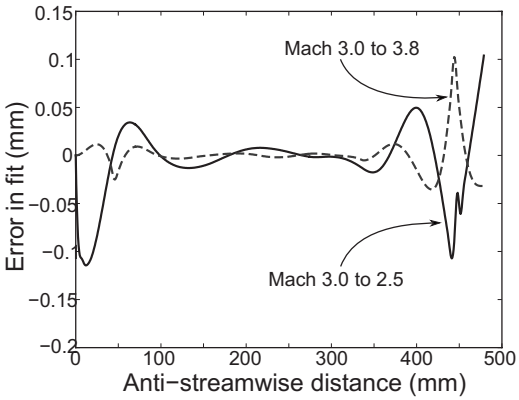



**Fig. 8 The differences between the desired shapes and the optimal shapes, in the least-squares sense, found by the shape fitting algorithm**

greatest. The resultant best fits are shown in Fig. 6, for comparison to the desired shapes. The differences between the desired shapes and the best shape attainable with this morphing structure are presented in Fig. 8. The maximum error in the fit is small, approximately 100 $\mu$m, but occurs at the throat of the tunnel.

Some comment on the differences between the desired shapes and the optimal shapes is required. Because of the CFD technique used to calculate the desired shapes, only the supersonic region beyond the throat of the nozzle was designed. The inlet region prior to the throat was taken to be a mirror image of the region immediately downstream of the throat. This leads to the cusps evident in the difference curves at the nozzle throat. These cusps are impossible to fit with the morphing system here (effectively beams in bending), and hence the error in the fit is worst at the nozzle throat. This is deleterious to the nozzle performance since the flow is highly sensitive to the conditions at the throat of the nozzle. This artifact of the CFD and the morphing system can be ameliorated by adding further constraints to the CFD calculations to limit the curvature of the difference curves and to specify the boundary conditions at the throat more precisely.

## 4 Construction of the Morphing Nozzle

The morphing structure is composed of a combination of stainless steel core members, a polycarbonate aerodynamic surface, and an aluminum base plate to which the actuators are fixed. This configuration is presented in Fig. 9. The core members are stainless steel rods, connected by pins at each end. The vertical tie rods are connected to the actuators using a set of universal joints to eliminate bending moments and are threaded in order to screw into the universal joints, which also act as turnbuckles for fine adjustment of the initial fit. Because the steel is several orders of magnitude stiffer than the polymer surface, the rods are assumed to be rigid in the analysis. The walls and flat roof of the tunnel, not shown in the photograph, are of 12.7 mm thick polycarbonate for optical access.

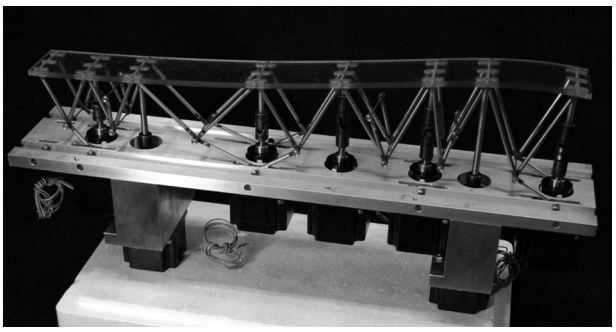


**Fig. 9 A photograph of the morphing structure prior to its installation in the wind tunnel. The tunnel walls and flat nonmorphing surface have been removed.**

A range of aerodynamic surfaces was manufactured, either by vacuum molding or by machining from block polycarbonate. The maximum thickness of the polycarbonate surface is limited by yielding of the material. The maximum imposed strain is determined from the maximum imposed curvature $\kappa_{max}$ due to the required shape change (the second spatial derivative of the approximated shapes shown in Fig. 6), such that $t_{max}=2\epsilon_Y\kappa_{max}^{-1}$, where $\epsilon_Y$ is the yield strain of the morphing face material. Polycarbonate was chosen as a surface material in part because its large yield strain permits very conservative designs. For this structure, $\epsilon_Y=0.02$ and $\kappa_{max}=1.5\times10^{-3}$ mm$^{-1}$, giving $t_{max}=27$ mm. A 10 mm thick sheet was chosen for the morphing surface to reduce the force necessary to morph the structure.

The actuators are custom-built Haydon Switch electric linear stepper motors with a 3.2 $\mu$m step size and a maximum force per motor of 880 N. Considerations of space limit the actuator arrangement; it would be preferable to install the actuators horizontally between adjoining nodes, as in the structure demonstrated by [8], but the small scale of this morphing structure prohibits this. AllMotion stepper controller chips were used to control and drive each actuator, and the chips were commanded through LabView software. The controllers were unable to sense an actuator's absolute position, so a 2 in. linear potentiometer was attached to the base of each actuator to measure actuator extension. These instruments gave precise measurements for the extension of the actuators and thus for the position of the surface at all times. The maximum actuator displacement during the transition from Mach 2.5 to Mach 3.8 is approximately 12 mm.

The morphing nozzle was integrated into a supersonic in-draft wind tunnel at the Princeton University Department of Mechanical and Aerospace Engineering. The nozzle itself is 445 mm in length with a fixed exit area of 2464 mm$^2$. The throat area varies from 877 mm$^2$ at Mach 2.5 to 252 mm$^2$ at Mach 3.8. The tunnel draws quiescent air from the laboratory through the nozzle, which gives constant and easily measurable stagnation conditions of room temperature and atmospheric pressure. The diverging end of the nozzle is attached to a diffuser connected to an air ejector system, which maintains the low back pressure that pulls in air from the room and creates supersonic flow through the test section. At Mach 3, the test section pressure is 2.7 kPa and the ejector back pressure is 20 kPa. A custom designed diffuser achieves the necessary pressure recovery after the test section. The diffuser slows the flow in a series of oblique shocks, decreasing the loss of total pressure produced by a normal shock diffuser.

## 5 Preliminary Assessment of the Nozzle Performance

Initial nozzle tests were designed to morph the nozzle through a full range of curve shapes and flow speeds, characterizing the nozzle performance. A 12 deg wedge was mounted to the flat surface of the nozzle test section to generate an oblique shock, which could be visualized with Schlieren imaging. Images were taken through two quartz windows in the sidewalls at the nozzle test section using a mercury lamp light source and a Qimaging Retiga 1300i Fast 1394 camera. Video and color stills were also taken with a Casio ExFilm EX-Z750 digital camera. Most nozzle runs began with the nozzle at the static Mach 2.5 contour before moving to Mach 3.0 or Mach 3.8 and then returning to Mach 2.5.

Figure 10 shows a series of images from one tunnel run in which the nozzle began at the Mach 2.5 contour and morphed to the Mach 3 contour. Taken in succession, these images show a clear change in shock angle, and thus Mach number, as the tunnel changes contour. Video footage of the nozzle showed a smooth and gradual shock angle change as the tunnel morphed, a visualization of the continuously changing Mach number in the nozzle

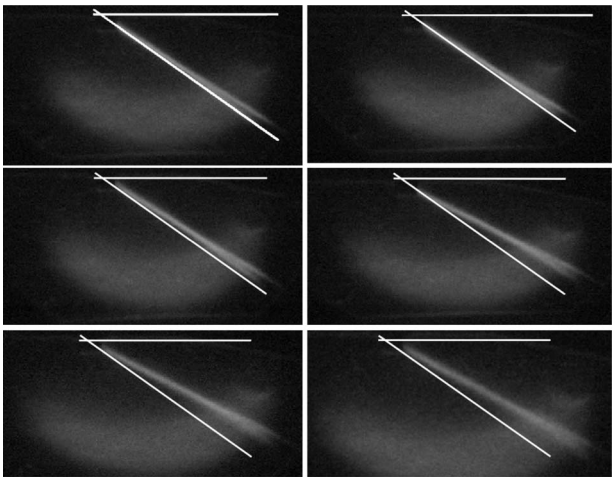


**Fig. 10 Gradual shock angle change from Mach 2.5 to Mach 3.0 read from left to right and top to bottom with the initial shock angle overlaid to show the change in shock angle as the surface deforms. Flow is left to right.**

test section. During motion of the morphing surface, no shocks were seen upstream of the wedge; this is evidence of an isentropic expansion through the nozzle to the test section.

An oblique shock was continuously visible during morphing from the Mach 2.5 contour all the way to the Mach 3.8 contour, across the morphing nozzle's entire flow speed range. The shock at the Mach 3.8 curve shape, however, appeared noticeably weaker in Schlieren images. Figure 11 is a series of four images showing the change from the Mach 3.8 contour to the Mach 3 contour. The shock was weak at the Mach 3.8 shape, but the same smooth, gradual shock angle change seen in the lower Mach number transitions was seen as the surface morphed to and from the Mach 3.8 curve shape.

One test was run without the wedge in the test section. This was meant to verify that the computed curves and the transitions between them were indeed isentropic, as they appeared to be when the wedge was in the test section. The three static contours, when properly formed, did not cause any shocks in the flow. They each created clean isentropic expansions at the expected static pressures. As the surface morphed, however, weak shocks were seen to appear and move swiftly downstream, as in Fig. 12. The shocks in these images appear to be coming from the flat wall, but they were more likely reflected shocks or expansions from the curved wall. As the aerodynamic surface morphed, it impinged into or

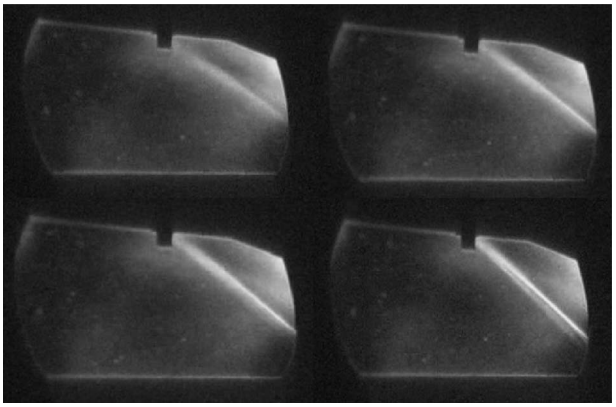


**Fig. 11 Oblique shock during the transition from the Mach 3.8 contour to the Mach 3 contour. The black tab in image shows the leading edge of the wedge.**

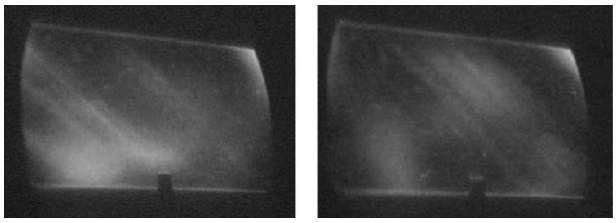


**Fig. 12 Oblique shocks move rapidly downstream during the transition between various curve shapes. Flow is left to right. The black tab marks the curved morphing side of the nozzle.**

retracted from the flow stream, causing disturbances to propagate downstream. These shocks were accompanied by jumps in static pressure followed by a pressure drop down to the expected values for each contour when the surface reached a static position.

Flow speed estimations were made with a static pressure reading immediately before the wedge. Based on the results from the nozzle test with no wedge, it is assumed, at least for static curve shapes, that the flow experiences an isentropic expansion through the nozzle. Flow speed can therefore be determined from the static pressure in the test section. An average pressure value was computed for each contour shape during each nozzle run, and then those values were averaged to get one mean value of static pressure for each of the three contours. For multiple runs over a period of two weeks, the pressure data were very consistent, as judged by very small standard deviations. To determine a characteristic flow speed at each contour, the average static pressure values were compared to the reservoir pressure using isentropic relations. Pressure values and the corresponding flow Mach number are shown in Table 1 for each of the three nozzle contours.

Clearly, there is a substantial disparity between the expected flow speeds and the observed values. There are two primary causes of this disparity: failure to correct fully for boundary layer effects and limitations of the wind tunnel ejector system. Initial curve design calculations corrected for the boundary layer thickness on the curved surface but did not account for boundary layer growth on the other nozzle walls. The presence of thick boundary layers on these walls significantly reduces the effective exit area, decreasing the Mach number for each contour and bringing the expected Mach number closer to those seen experimentally. The large error at the Mach 3.8 contour is most likely a result of the limitations of the wind tunnel ejector system. It was found that the ejector system was unable to draw the necessary mass flow to achieve Mach 3.8 flow in the nozzle, so the flow speed values observed must necessarily be lower than this upper limit, regardless of contour accuracy. This also explains the weakened appearance of the oblique shock at the Mach 3.8 contour.

## 6 Concluding Comments

To the best of the authors' knowledge, this is the first practical demonstration of a variable Mach number, continuously isentropic morphing supersonic wind tunnel. This nozzle is capable of continuously and smoothly changing output flow velocity through a range of Mach numbers. While weak oblique shocks were seen

**Table 1 Expected Mach numbers, measured pressures, and inferred Mach numbers for peration of the morphing wind tunnel nozzle**

| Design Mach number | Average of the average pressures (kPa) | Interpolated Mach number from isentropic tables |
|---|---|---|
| 2.5 | $8.2 \pm 1.3$ | 2.3 |
| 3.0 | $4.3 \pm 1.1$ | 2.8 |
| 3.8 | $1.75 \pm 0.3$ | 3.3 |

during shape transformations without the wedge installed, it is believed that these shocks can be eliminated by morphing the tunnel in smaller steps, rather than the large jumps (for example, between Mach 3 and Mach 3.8) used here. This would require the calculation of a larger series of ideal shapes but does not necessitate any new procedures. Ultimately, the nozzle is designed to demonstrate the concept of a morphing aerodynamic surface for applications in scramjet inlets. This research successfully validated the possibility of such a technology while addressing a number of engineering problems involved in such a design, including aerodynamic design, curve approximation, and positioning and control.

There remain a number of interesting theoretical issues surrounding the choice of the "best" approximate shape for the nozzle profile. In this work, the best shape has been identified with a least-squares fit. However, this provides no linkage between the surface profile and the aerodynamic performance of the system. For example, it is clear that the tunnel is more sensitive to positional errors at the throat than errors near the diffuser; the least-squares technique does not account for this. Alternate methods would be to use a weighted least-squares fit or to calculate the shapes directly through CFD, using the attainable nozzle shapes as further constraints. This latter approach would be possible for the morphing nozzle, which has few degrees of freedom, can be easily approximated as two-dimensional, and has a limited range of operating conditions. A morphing ramjet nozzle, which may need to respond quickly to changes in velocity, pressure, and attitude, and is fully three-dimensional, would overtax any CFD approach. Instead, a system with conjoined analytical and adaptive components might be preferable. We postulate a system wherein analytical (or highly simplified CFD) techniques are utilized to determine the large scale shape changes in the morphing structure, before an active feedback system, possibly based on a neural net, fine-tunes the surface to an ideal operating profile.

The installation and operation of the morphing nozzle revealed several other issues, which were only broadly considered in the preliminary design. A relatively thin morphing surface was chosen to reduce the loads induced on the actuators, balancing actuator performance, and structural rigidity. However, the aerodynamic loads caused noticeable deformations in the nozzle surface. Nevertheless, the flow appeared to be minimally sensitive to the magnitudes of deformation experienced. This may be a consequence of the relatively thick subsonic boundary layer suppressing shock formation. The morphing nozzle might therefore be less sensitive to small errors in surface approximation than initially expected. Ultimately the design of morphing systems is actuator dominated: The performance is governed by the space available for actuation and the actuation method that is selected. In the case described here, the maximum size and number of the actuators were governed by the length of the nozzle and the by the need to provide sufficient overall control of the morphing surface. This limited the power and speed of the actuators. It is believed that, as the system is scaled up, the actuator authority will increase more quickly than the aerodynamic loads, reducing the complexity of some of the design trades needed at the laboratory scale.

## Acknowledgment

The authors would like to thank Dan Ketchpel of Industrial Forming, Ventura, CA for his assistance with the manufacture of the aerodynamic surface and Sohail Zaidi and Phil Howard of Princeton University for their contributions to operation of the wind tunnel. The research reported in this paper was supported by the Office of Naval Research through a MURI program on Revolutionary Materials for Hypersonic Flight (Contract No. N00014-05-1-0439), principal investigator Anthony G. Evans.

## References

[1] Hutchinson, R. G., Wicks, N., Evans, A. G., Fleck, N. A., and Hutchinson, J. W., 2003, "Kagome Plate Structures for Actuation," Int. J. Solids Struct., **40**(25), pp. 6969–6980.
[2] Lucato, S. L., Wang, J., Maxwell, P. T., McMeeking, R. M., and Evans, A. G., 2004, "Design and Demonstration of a High Authority Shape Morphing Structure," Int. J. Solids Struct., **41**(13), pp. 3521–3543.
[3] Evans, A. G., Hutchinson, J. W., and Ashby, M. F., 1998, "Multifunctionality of Cellular Metal Systems," Prog. Mater. Sci., **43**(3), pp. 171–221.
[4] Evans, A. G., Hutchinson, J. W., Fleck, N. A., Ashby, M. F., and Wadley, H. N. G., 2001, "Topological Design of Multifunctional Cellular Metals," Prog. Mater. Sci., **46**(3–4), pp. 309–327.
[5] Evans, A. G., 2001, "Lightweight Materials and Structures," MRS Bull., **26**(10), pp. 790–797.
[6] Jameson, A. 1988, "Aerodynamic Design Via Control Theory," J. Sci. Comput., **30**(3), pp. 233–260.
[7] Jameson, A., 1990, "Control Theory for Optimum Design of Aerodynamic Shapes," *Proceedings of the IEEE Conference on Decision and Control*, Honolulu, Dec. 5–7, IEEE Control Systems Society, pp. 176–179.
[8] Zupan, M., Ashby, M. F., and Fleck, N. A., 2002, "Actuator Classification and Selection—The Development of a Database," Adv. Eng. Mater., **4**(12), pp. 933–939.

# Time-Derivative Preconditioning Methods for Multicomponent Flows—Part II: Two-Dimensional Applications

**Jeffrey A. Housman**
University of California Davis,
2132 Bainer Hall,
One Shields Avenue,
Davis, CA 95616

**Cetin C. Kiris**
NASA Advanced Supercomputing (NAS) Division,
NASA Ames Research Center,
Moffett Field, CA 94035

**Mohamed M. Hafez**
University of California Davis,
2132 Bainer Hall,
One Shields Avenue,
Davis, CA 95616

*A time-derivative preconditioned system of equations suitable for the numerical simulation of multicomponent/multiphase inviscid flows at all speeds was described in Part I of this paper. The system was shown to be hyperbolic in time and remain well conditioned in the incompressible limit, allowing time marching numerical methods to remain an efficient solution strategy. Application of conservative numerical methods to multicomponent flows containing sharp fluid interfaces was shown to generate nonphysical pressure and velocity oscillations across the contact surface, which separates the fluid components. It was demonstrated using the one-dimensional Riemann problem that these oscillations may lead to stability problems when the interface separates fluids with large density ratios, such as water and air. The effect of which leads to the requirement of small physical time steps and slow subiteration convergence for the implicit time marching numerical method. Alternatively, the nonconservative and hybrid formulations developed by the present authors were shown to eliminate this nonphysical behavior. While the nonconservative method did not converge to the correct weak solution for flow containing shocks, the hybrid method was able to capture the physically correct entropy solution and converge to the exact solution of the Riemann problem as the grid is refined. In Part II of this paper, the conservative, nonconservative, and hybrid formulations described in Part I are implemented within a two-dimensional structured body-fitted overset grid solver, and a study of two unsteady flow applications is reported. In the first application, a multiphase cavitating flow around a NACA0015 hydrofoil contained in a channel is solved, and sensitivity to the cavitation number and the spatial order of accuracy of the discretization are discussed. Next, the interaction of a shock moving in air with a cylindrical bubble of another fluid is analyzed. In the first case, the cylindrical bubble is filled with helium gas, and both the conservative and hybrid approaches perform similarly. In the second case, the bubble is filled with water and the conservative method fails to maintain numerical stability. The performance of the hybrid method is shown to be unchanged when the gas is replaced with a liquid, demonstrating the robustness and accuracy of the hybrid approach.* [DOI: 10.1115/1.3086592]

*Keywords: hybrid conservative/nonconservative method, split coefficient matrix (SCM) method, time-derivative preconditioning, dual time stepping*

## 1 Introduction

As described in Part I, many propulsion related flow applications require modeling of multicomponent and or multiphase flows over a wide range of Mach numbers. One example is the low speed flow of liquid propellants through the low pressure fuel turbopump (LPFTP) in the space shuttle main engine (SSME), see Ref. [1]. In this case, cavitation of the liquid propellant is likely to occur in the turbopump, which may cause blade damage and a reduction in the thrust produced by the SSME. Another example is the overpressure suppression system activated during the launch of a space vehicle. In this case, liquid water is injected into the exhaust plume and troughs of water baths are placed in the flame trench to reduce the ignition overpressure waves generated from the interaction of the exhaust plume and the flame trench at takeoff, see Ref. [2]. A time-derivative preconditioned numerical method appropriate for the simulation of compressible multicomponent and multiphase flows obeying arbitrary equations of state was described in Part I of this paper. The numerical method includes three formulations for the discretization of the convective flux derivatives: a conservative preconditioned Roe (PROE) method, a nonconservative preconditioned split coefficient matrix (PSCM) method, and a hybrid (HYBR) method, which combines the conservative and nonconservative methods. In order assess the accuracy and the robustness of the three formulations on multidimensional problems, each of the methods is applied to two separate flow applications. Performance and accuracy comparisons of each of the methods are discussed.

The present paper is organized as follows: First, steady and unsteady flows of liquid water through a channel containing a NACA0015 hydrofoil are computed. The inflow Mach number is set to induce cavitation at the hydrofoil upper surface near the leading edge. The solution is compared with the experimental results and a time accurate sequence of the cavitation phenomenon is presented. In the second application, the PROE and HYBR methods are used to predict the interaction of a shock wave traveling through air and hitting a cylindrical bubble of helium (or water). In the case of the helium bubble, both methods perform well and the HYBR method is shown to compute the correct shock speed and location. Alternatively, in the case of a water

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANISM. Manuscript received January 31, 2008; final manuscript received December 30, 2008; published online March 13, 2009. Review conducted by Tayfun E. Tezduyar.

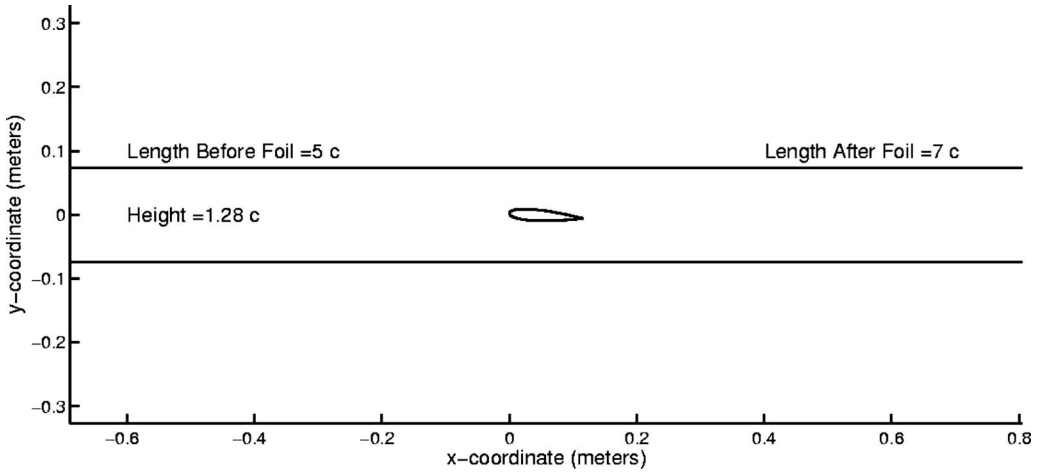



**Fig. 1 Geometry for NACA0015 inside a channel**

bubble, the PROE method fails to converge within the subiteration procedure for the fixed physical time step, while the HYBR method remains robust and efficient independent of the density ratio at the fluid interface.

## 2 Computed Results

The time-derivative preconditioned system of equations described in Part I of this paper is used to solve a series of two-dimensional steady and unsteady inviscid multicomponent flow problems. Three convective flux derivative discretizations are compared for the first application. These methods include the conservative PROE method, the nonconservative PSCM method, and the HYBR method, which combines the PROE and PSCM methods. In each of the test cases, either first-order upwind or third-order upwind biased differencing with total variation diminishing (TVD) minmod limiters are used in space, and first-order backward differencing in time (to ensure TVD boundedness). Note that the PROE method does not maintain bounded mass fractions ($0 \leq Y_i \leq 1$) even when first-order time and space discretizations are used (the proof of this is given in Ref. [3]).

To begin, the three methods are used to predict the cavitating liquid flow around a NACA0015 hydrofoil contained in a channel; this test case demonstrates the multiphase modeling capability of the methods. Then, to demonstrate the shock capturing capabilities of the preconditioned HYBR method, two shock bubble interaction test cases are solved. In the first case, the PROE and HYBR methods are used to predict the interaction of a $M_S=1.22$ shock wave traveling through air and hitting a cylindrical bubble of helium. Both methods perform well for this case, and the solutions predicted by the PROE and HYBR methods are indistinguishable. In the second case, the bubble of helium is replaced with a bubble of water and the PROE method becomes unstable as the shock wave hits the interface (for the moderately large time step used); this leads to subiteration divergence and failure of the method. Alternatively, the HYBR method remains efficient and stable by locally switching to the nonconservative form at the interface while predicting the correct shock-wave dynamics away from the sharp interface, demonstrating the superiority of the hybrid formulation for multicomponent flows at all speeds.

**2.1 Cavitating NACA0015 Hydrofoil.** The NACA0015 hydrofoil test case, studied in Ref. [4], is revisited here with modified inflow conditions inducing cavitation to occur and necessitating the inclusion of multiphase effects in the physical and numerical models. This case corresponds with case 2 of the Salvetti and Beux benchmark problem [5]. Here the multicomponent mixture model is used to predict cavitating flow through a channel containing a NACA0015 hydrofoil. In addition, the different convective term discretization strategies are analyzed and compared with each other in terms of accuracy and efficiency. The computed results are compared with the experimental results reported by Rapposelli et al. [6]. It is demonstrated that all three methods predict similar pressure distributions over the hydrofoil, but the nonconservative PSCM method is the most efficient for unsteady flows in terms of convergence rate and overall runtime. This case has also been solved by Li and Merkle [7] and Li et al. [8] using an equivalent mixture formulation and a numerical method similar to the present PROE method. There computation included viscous effects, which are neglected in this study. In their computations, the physical vapor pressure ($P_{\text{vap}}=3165$ Pa) was increased to 5372 Pa to simulate the effects of turbulence on cavitation. In the present study, computations using both the physical vapor pressure and the vapor pressure used by Li and Merkle [7] and Li et al. [8] are reported.

The profile of the NACA0015 hydrofoil is defined by the analytic function

$$\begin{matrix} y_{\text{upper}}(\bar{x}) = y(\bar{x}) \\ y_{\text{lower}}(\bar{x}) = -y(\bar{x}) \end{matrix} \quad \text{for} \quad \bar{x} = x/c \in [0,1]$$

where

$$y(\bar{x}) = 0.75(0.2969\sqrt{\bar{x}} - 0.12\bar{x} - 0.3537\bar{x}^2 + 0.2843\bar{x}^3 - 0.1015\bar{x}^4) \tag{1}$$

The chord of the hydrofoil is $c=0.115$ m and is contained in a channel 13 chord lengths long and 1.28 chord lengths high at an angle of attack of 4 deg. An outline of the geometry is displayed in Fig. 1. Discretization of the solution domain is carried out using a structured overset grid consisting of five zones and a total of 22,205 grid points. Figure 2 displays the overset region of the grid near the hydrofoil. The liquid water is assumed to obey a stiffened gas equation of state of the form

$$\rho = \frac{p + p_\infty}{RT} \quad \text{and} \quad h = C_p T \tag{2}$$

where the gas constant has its usual definition $R=C_p(\gamma-1)/\gamma$. The material properties are taken as

$$\gamma = 1.9276, \quad C_p = 8076.73 \text{ J/kg/K}, \quad p_\infty = 1.137279 \times 10^9 \text{ Pa}$$

These material constants are consistent with those reported in Ref. [9] to model water at standard atmospheric conditions. The boundary conditions include slip walls on the upper and lower surfaces of the channel, as well as the surface of the hydrofoil. At the inflow boundary, the velocities, temperature, and mass fraction are specified and the characteristic relation is used to update the

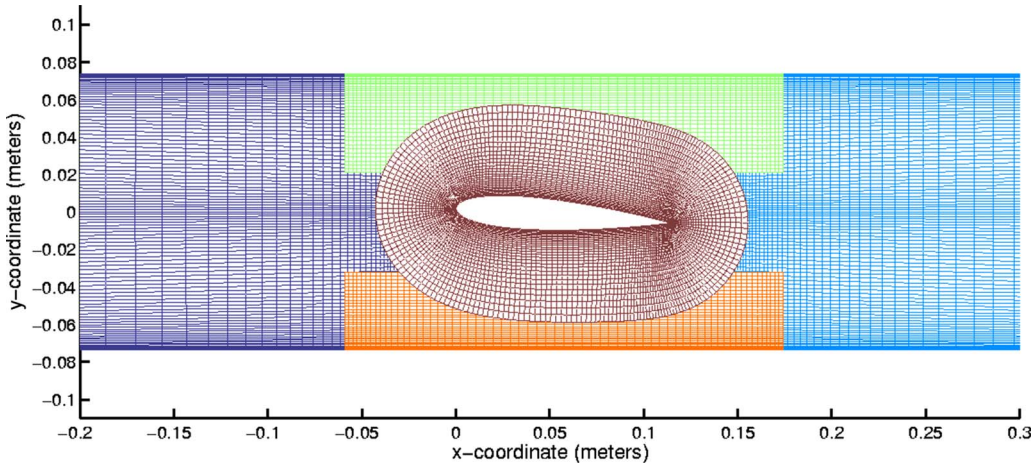



**Fig. 2 Structured overset grid for NACA0015 inside channel**

pressure. At the outflow boundary, a nonreflecting characteristic boundary condition is used. Points lying on the overset boundaries are updated using bilinear interpolation of the primitive variables $Q=(p,u,v,T,Y)$. The inlet conditions provided by the experimental study are

$$P_{\text{in}} = 0.12 \text{ bar}, \quad U_{\text{in}} = 3.41 \text{ m/s}, \quad T_{\text{in}} = 298 \text{ K}$$

To model the generation/condensation of vapor, a source term is added to the right-hand side of the governing equations of the form $(0,0,0,0,-\dot{m}^+-\dot{m}^-)^T$, where the mass fraction of the vapor phase is being solved for. Following Li et al. [8], the finite rate relations, as developed in Ref. [10], are used:

$$\dot{m}^+ = C_{\text{prod}}\left(\frac{\min(p-P_{\text{vap}},0)}{\frac{1}{2}\rho_{\text{ref}}U_{\text{ref}}^2 t_{\text{ref}}}\right)\rho(1-Y_v) \tag{3}$$

$$\dot{m}^- = C_{\text{dest}}\left(\frac{\max(p-P_{\text{vap}},0)}{\frac{1}{2}\rho_{\text{ref}}U_{\text{ref}}^2 t_{\text{ref}}}\right)\rho Y_v \tag{4}$$

The empirical constants $C_{\text{prod}}$ and $C_{\text{dest}}$ are set equal to the values used by Li et al. [8],

$$C_{\text{prod}} = 10^5 \quad \text{and} \quad C_{\text{dest}} = 1$$

To begin, the physical vapor pressure value of $P_{\text{vap}}=3165$ Pa is used in the source term evaluation, along with third-order convective flux derivatives with TVD limiters, and all the three methods are ran in steady-state mode. Cavitation occurs on the upper surface of the hydrofoil where a thin sheet of vapor/liquid mixture forms. The $C_p$ curve on the upper and lower surfaces of the hydrofoil are displayed in Fig. 3, along with the experimental data. The dashed line in the figure represents the value of

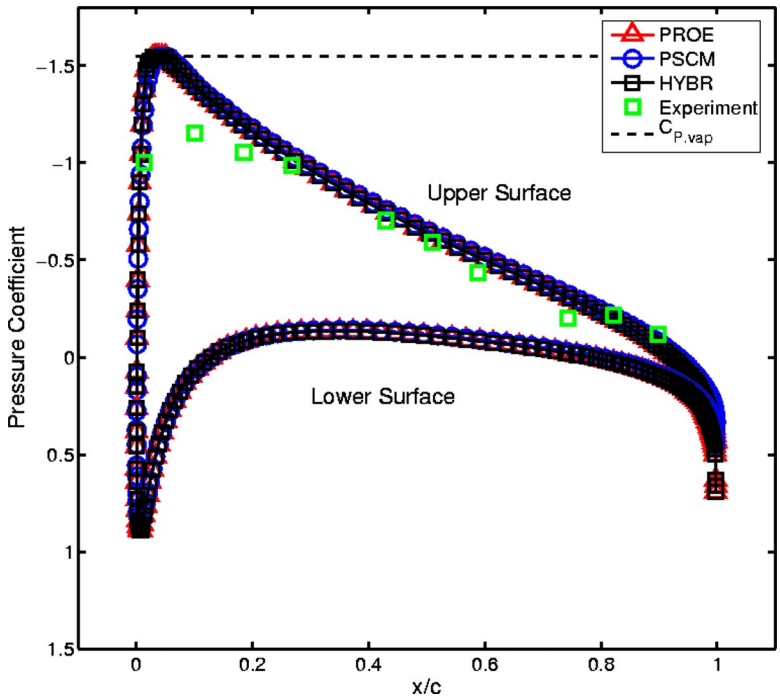



**Fig. 3 Plot of the $C_p$ curve on the upper and lower surfaces of the hydrofoil for $P_{\text{vap}}=3165$ Pa using third-order convective fluxes with TVD limiters for each of the methods**

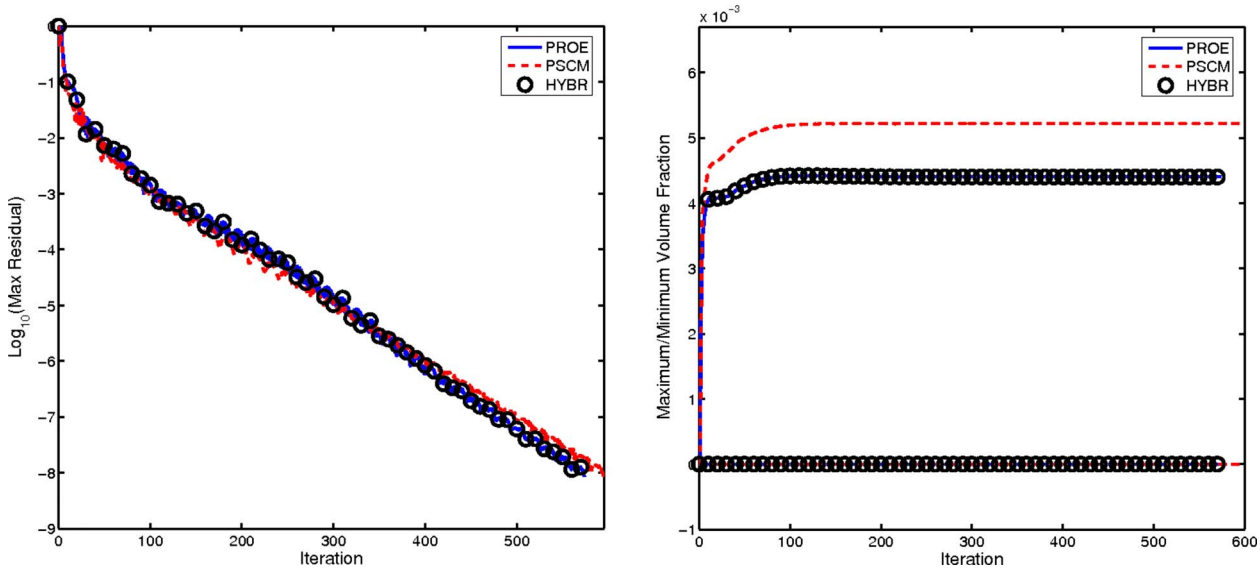



**Fig. 4 Convergence history of the maximum residual (top) and the computed maximum/minimum volume fraction for $P_{vap}$=3165 Pa using third-order convective fluxes with TVD limiters (bottom)**

$$\frac{P_{vap} - P_{in}}{\frac{1}{2}\rho_\infty U_\infty^2} \tag{5}$$

This is the pressure value at which the source term is activated and vapor is being generated. All three numerical methods converge to a similar solution, but none of them compares particularly well with the experimental data. All three methods reduce the maximum residual eight orders of magnitude in approximately the same number of iterations, as displayed in Fig. 4 along with the maximum and minimum volume fractions as a function of the iteration. Only small values of volume fraction were generated using the physical value of the vapor pressure. Since no sharp mass fractions are present in the solution, the results using the PROE and HYBR methods are identical.

In discussing these results with Venkateswaran, the second author of Ref. [8], he suggested to use the larger value of vapor pressure that they used, since they also experienced difficulties in matching their solution to the experimental data when using the physical value of the vapor pressure. The vapor pressure was then raised to $P_{vap}$=5372 Ps and the three methods were re-applied. To start, first-order convective fluxes were used and a steady solution was achieved. The $C_p$ curve on the upper and lower surfaces of the hydrofoil is displayed in Fig. 5. There is much better agreement with the experimental data near the leading edge, but the

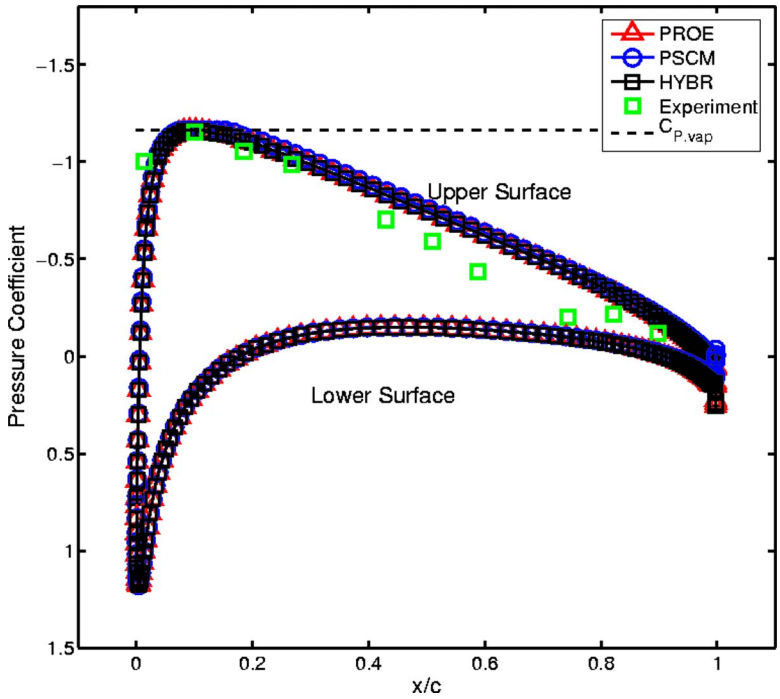



**Fig. 5 Plot of the $C_p$ curve on the upper and lower surfaces of the hydrofoil for $P_{vap}$=5372 Pa using first-order convective fluxes for each of the methods**

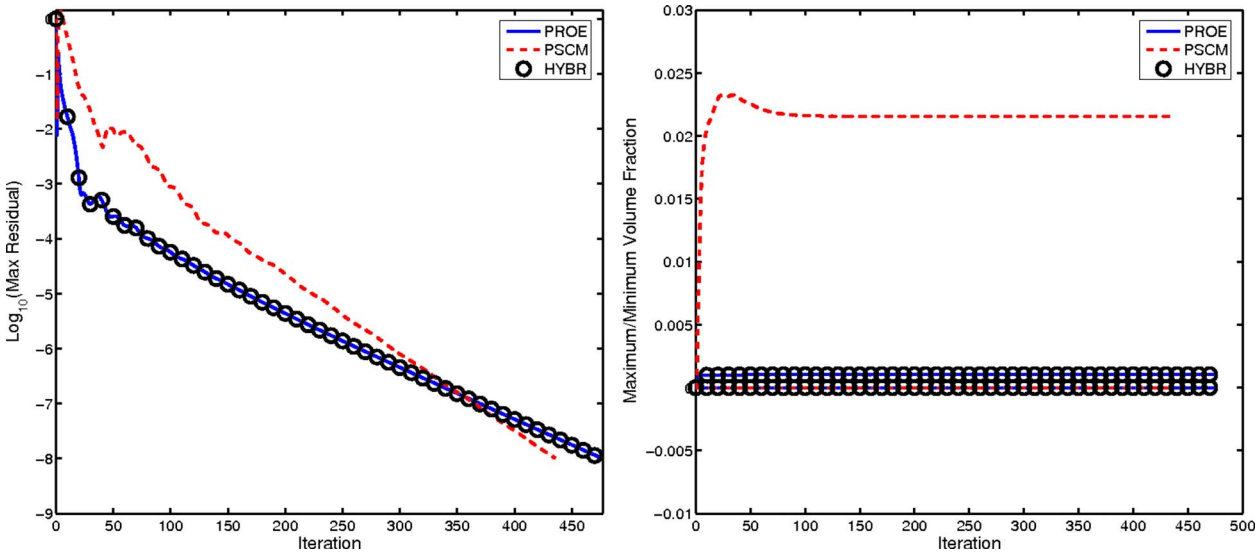



**Fig. 6 Convergence history of the maximum residual (top) and the computed maximum/minimum volume fraction for $P_{vap}=5372$ Pa using first-order convective fluxes (bottom)**

numerical diffusion appears to smooth the re-entrant jet such that the computed surface does not match the data near the trailing edge. Figure 6 plots the convergence history and maximum/minimum computed vapor volume fraction. Again the convergence history of all three methods are approximately the same, but a much larger vapor volume fraction is predicted than in the previous case. Additionally, no sharp interfaces are present in this calculation so the PROE and HYBR results are identical.

To reduce the numerical diffusion, third-order convective fluxes with TVD limiters were applied to the problem using the modified vapor pressure, $P_{vap}=5372$ Pa. During the time marching process, the steady-state residual began to oscillate and the pseudotime solution was analyzed to see whether the oscillations were caused by limiter rattling or some physical phenomenon. It was found that a vapor sheet generated on the upper surface of the hydrofoil was shedding. This implies that the flow is physically unsteady and the dual time stepping must be applied. This was not surprising since cavitation is an inherently unsteady phenomenon. Dual-time unsteady analysis was then performed at a physical time step of $\Delta t=1.0\times10^{-05}$ s, based on examining the shedding frequency observed in the pseudotime solution. Each method was ran with a pseudotime $CFL=1.0\times10^{+04}$, $U_\varepsilon=3.4$ m/s, and a convergence criterion of three orders of magnitude residual reduction or 50 subiterations (whichever is achieved first). The $C_p$ curve is highly unsteady due to the shedding vapor sheet, Fig. 7

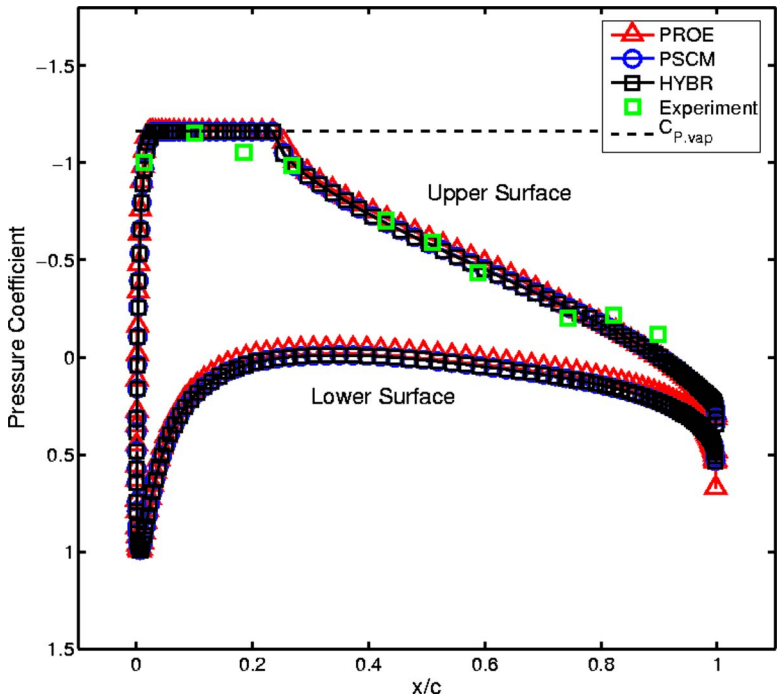



**Fig. 7 Plot of the $C_p$ curve on the upper and lower surfaces of the hydrofoil for $P_{vap}=5372$ Pa using third-order convective fluxes with TVD limiters for each of the methods**

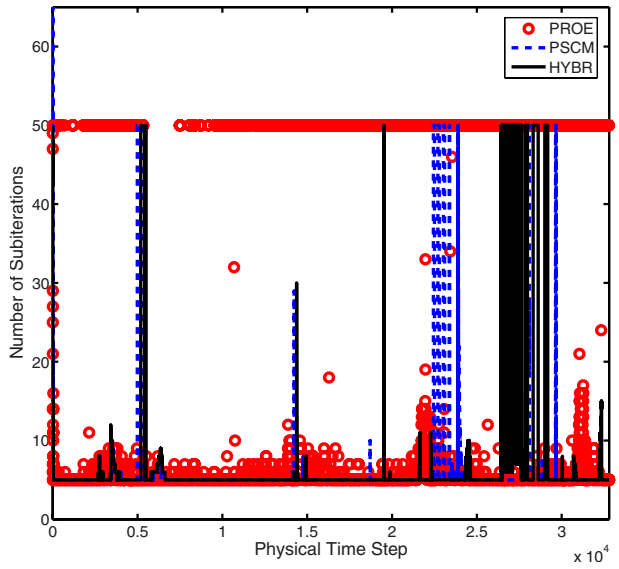



**Fig. 8 Number of subiterations required to reduce the residual three orders of magnitude as a function of the physical time step**

plots the $C_p$ curve during a physical time when the vapor sheet is attached to the hydrofoil. The result matches the experiment more closely than the first-order result, but a time-averaging procedure of the $C_p$ is necessary to assess the true accuracy of the solutions.

The convergence history for the three methods is displayed in Fig. 8. The PROE method appears to struggle during the subiteration procedure and performs the maximum 50 subiterations for almost every physical time step. Alternatively, the PSCM and HYBR methods are efficiently reducing the maximum residual and only perform the user specified minimum of 5 subiterations for most physical time steps. The unsteady flow is nearly incompressible and the results of the PSCM and HYBR methods are very similar. Since the PSCM method is the most efficient of the three methods, the computation was continued and Fig. 9 plots the maximum pressure as a function of time. From the figure, we observe the periodic nature of the flows. There appears to be a large pressure peak followed by a much smaller peak. Examining

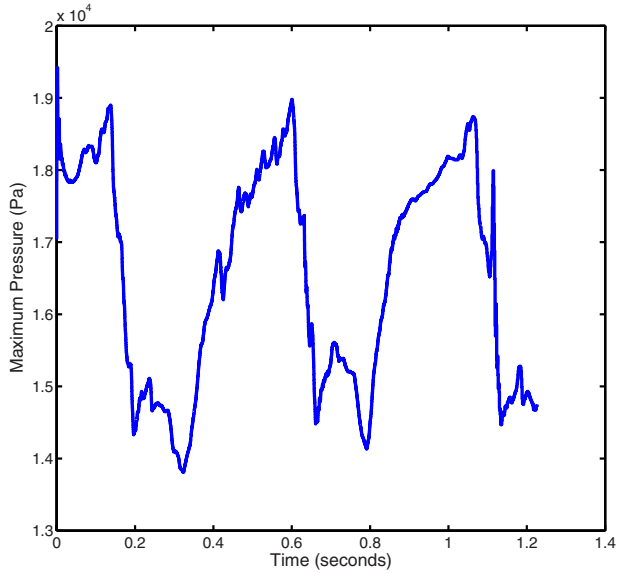



**Fig. 9 Maximum pressure as a function of time computed with the PSCM method**

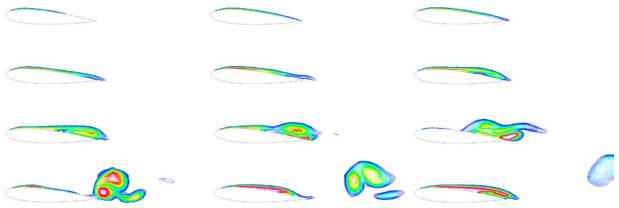


**Fig. 10 Volume fraction at $t$=0.02 to 0.24 s at intervals of 0.02 s**

the unsteady vapor volume fraction, the large peak is associated with a large sheet of vapor shedding from the hydrofoil, and the smaller peak corresponds to a smaller sheet following the larger one. To demonstrate the vapor shedding phenomenon, time sequences of contours of the vapor volume fraction are shown in Figs. 10–13. The contour levels are set from 0.1 (blue) to 0.9 (red). The time sequence starts with the initial forming of the vapor sheet on the upper surface of the hydrofoil. Next a large vapor cloud forms and a re-entrant jet pulls the vapor cloud from the foil and propagates it upstream. Next a small sheet of vapor rolls up and sheds of the trailing edge of the foil. The process is then repeated for another cycle, confirming the time periodic behavior of the flow.

**2.2 Shock Bubble Interaction.** This next application involves a multicomponent flow with sharp interfaces, compressibility effects, and shock waves. The simulation of a $M_S$=1.22 shock traveling through air and interacting with a cylindrical bubble of

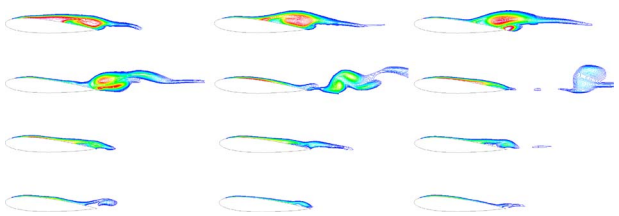


**Fig. 11 Volume fraction at $t$=0.26 to 0.48 s at intervals of 0.02 s**

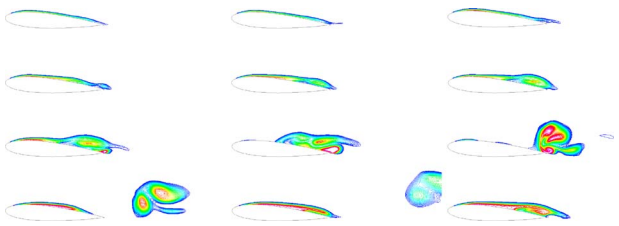


**Fig. 12 Volume fraction at $t$=0.50 to 0.72 s at intervals of 0.02 s**

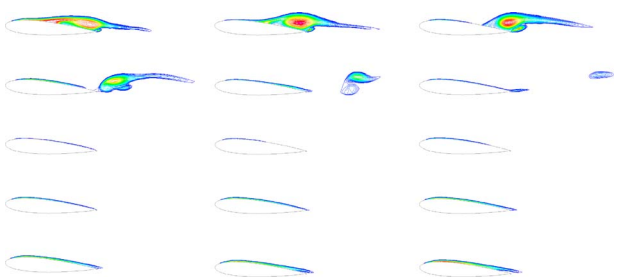


**Fig. 13 Volume fraction at $t$=0.74 to 1.02 s at intervals of 0.02 s**

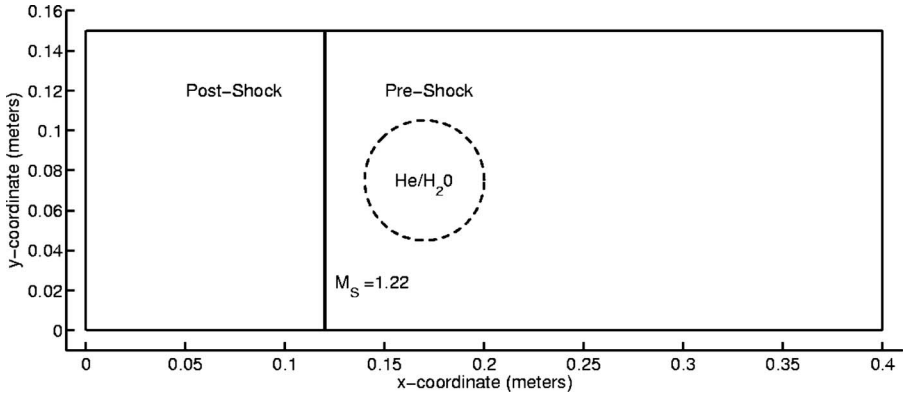



**Fig. 14 Configuration for shock/bubble interaction problem**

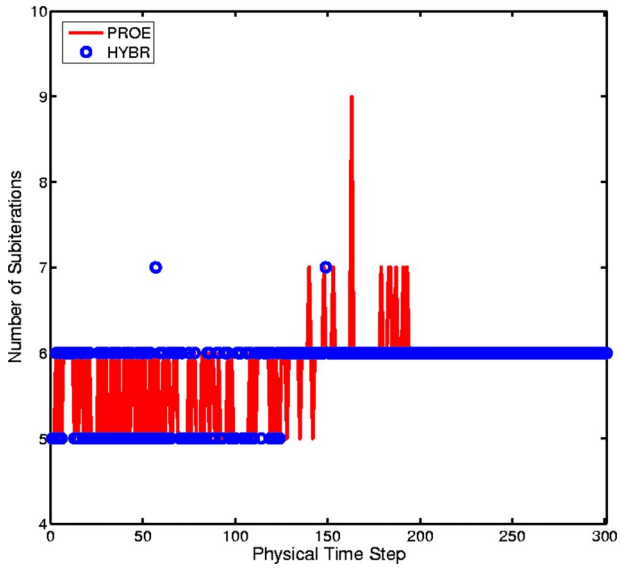



**Fig. 15 Subiteration convergence history of the PROE and HYBR methods for the $M_S$=1.22 shock air/helium interaction problem**

another fluid is performed. This case represents a preliminary step toward the modeling of fuel droplet vaporization and combustion in liquid propellant rocket engines. Here only a single droplet is considered and all chemistry and viscous forces are neglected. Since the test case requires the prediction of shock-wave dynamics, only the PROE and HYBR methods are applied. A schematic of the test case is displayed in Fig. 14. Initially a cylindrical bubble filled with helium gas is considered. The helium is assumed to obey an ideal gas equation of state where $\gamma_{\mathrm{He}}=1.67$ and $C_{p,\mathrm{He}}=519.37$ J/kg K. In the second case, the helium is replaced by water, which is assumed to obey the stiffened gas equation of state. The initial state of the preshock air is given by

$$P_{\mathrm{pre}} = 100{,}000 \text{ Pa}, \quad T_{\mathrm{pre}} = 300 \text{ K}, \quad u_{\mathrm{pre}} = 0 \text{ m/s}$$

and the mass fraction is set to zero inside the bubble and unity everywhere else. At these conditions, the slightly heavier helium gas creates a density ratio of 1.38 with the surrounding air, while the liquid water creates a much larger density ratio of 840. This is much closer to the density ratios present in LOX fueled rocket

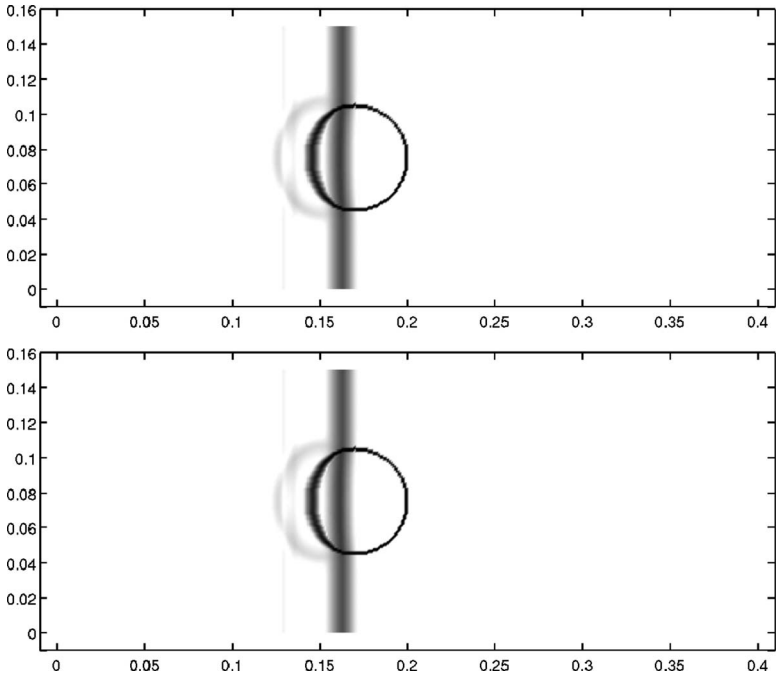


**Fig. 16 Numerical Schlieren for air/helium shock bubble interaction at time step 100: (Upper) PROE and (Lower) HYBR**

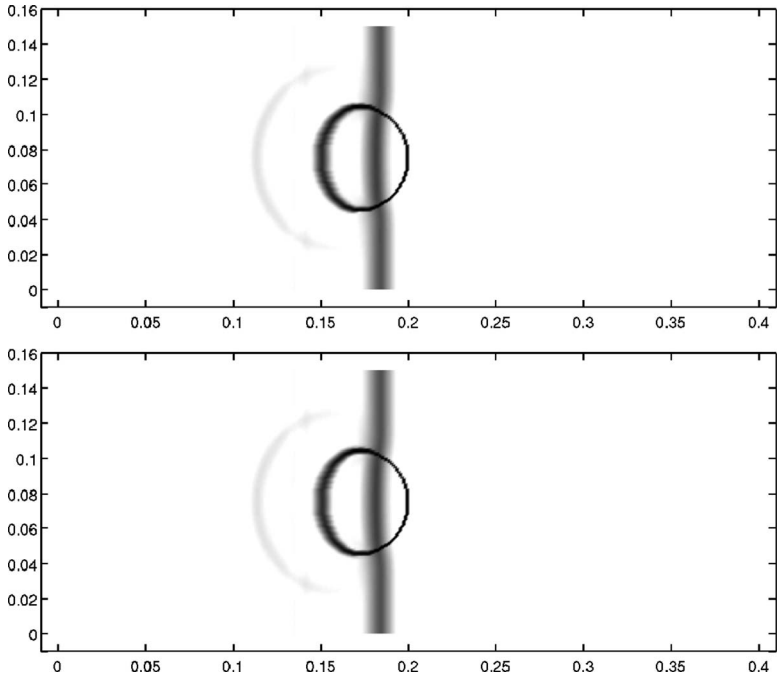



**Fig. 17 Numerical Schlieren for air/helium shock bubble interaction at time step 150: (Upper) PROE and (Lower) HYBR**

engines. The postshock state is derived using the one-dimensional shock relations, described in Ref. [11], and is given by

$$P_{\text{post}} = 156{,}980 \text{ Pa}, \quad T_{\text{post}} = 342.16 \text{ K}, \quad u_{\text{post}} = 115.85 \text{ m/s}$$

A uniform grid with dimensions $401 \times 151$ was generated for the present test case. This represents a grid resolution on the order of 1.0 mm, which is coarse compared with the resolution used by Quirk and Karni [12], where an adaptive grid method was used and an average mesh resolution of approximately 0.056 mm was considered. The purpose of this test is to compare the results predicted using the PROE and HYBR methods, not the accuracy of the solution compared with experimental data. The boundary conditions on the upper and lower walls of the channel are prescribed as slip walls, and extrapolation is used at the inflow and outflow boundaries. A physical time step $\Delta t = 1$ $\mu$s, based on the shock speed, is used along with a pseudotime $CFL = 10000$. A convergence criterion of three orders of magnitude reduction in the maximum residual or a maximum of 50 subiterations was used for

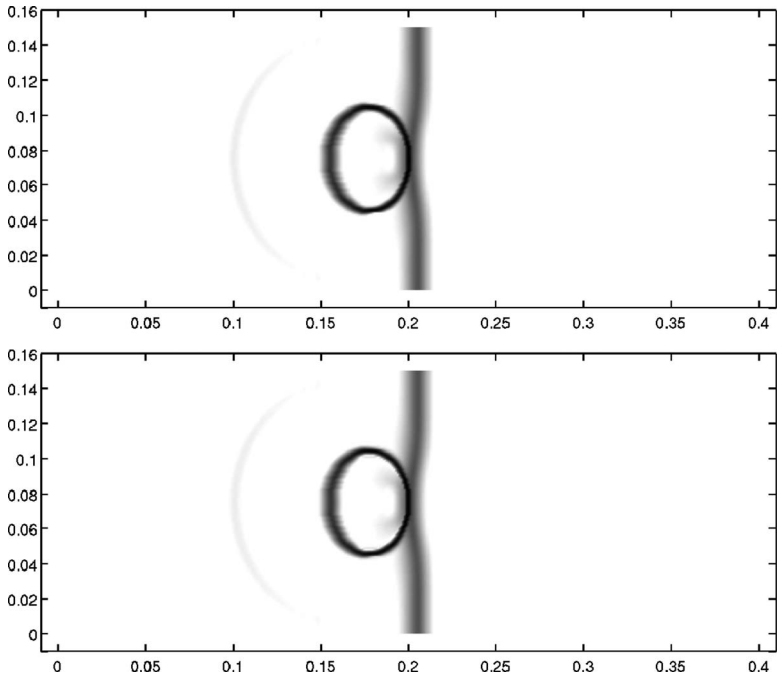



**Fig. 18 Numerical Schlieren for air/helium shock bubble interaction at time step 200: (Upper) PROE and (Lower) HYBR**

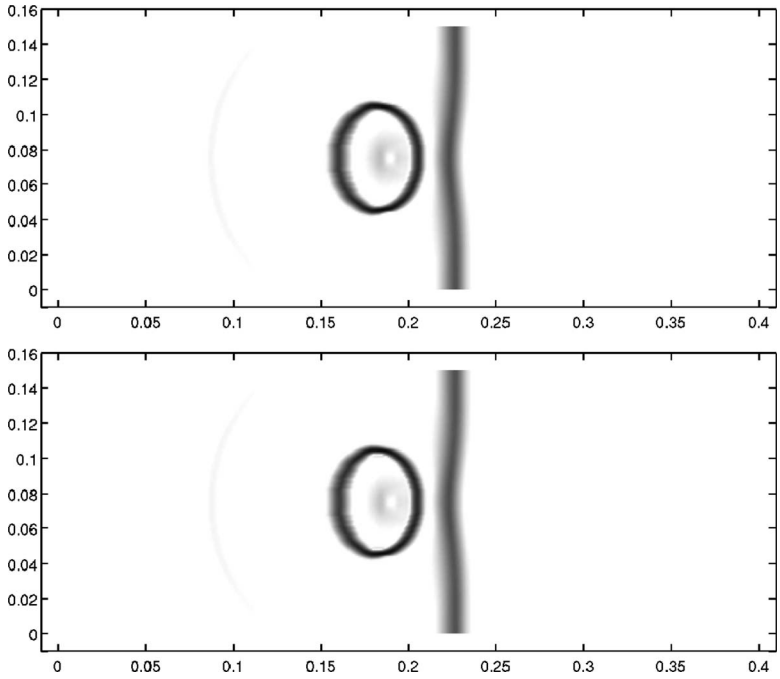



**Fig. 19 Numerical Schlieren for air/helium shock bubble interaction at time step 250: (Upper) PROE and (Lower) HYBR**

the calculations, which were performed for 300 physical time steps.

To begin, we discuss the results of the shock interacting with the helium bubble. In this case, both methods perform equivalently well in terms of convergence rate, as displayed in Fig. 15, which plots the number of subiterations required to meet the convergence criteria as a function of the physical time step. Additionally, both methods predict similar solution behavior, as displayed in Figs. 16–19, where numerical Schlieren images at particular time steps are plotted for each of the two methods. Here the numerical Schlieren is defined as $\log(|\nabla\rho|+1)$. Figure 16 displays the solution at the 100th time step corresponding to the time when the shock has impacted the bubble. In this figure, we see that most of the shock successfully transmitted through the bubble interface and only a small reflected shock appears in front of the bubble. Both solutions are practically identical and the HYBR method computes the correct shock location and strength. Figure 17 displays the solution at the 150th time step, at this time the shock is over halfway through the bubble and the reflected shock has traveled upstream. Figure 18 displays the solution at the 200th time step; now the shock is transmitting through the other side of the bubble, and another reflected shock appears inside the helium bubble. Finally, Fig. 19 displays the solution at the 250th time step, where the initial shock wave has transmitted through the bubble, and the reflected shocks are continuing to travel upstream of the bubble. In these results, we have demonstrated the capability of both methods to model multicomponent flows with shocks, and the shock capturing capabilities of the HYBR method have been confirmed in multidimensions.

The gas/gas shock bubble interaction case does not require low speed preconditioning and although the problem is multicomponent, the density ratios appearing in the solution are not large. Now the bubble is assumed to be filled with water, and since the sound speed of water is much larger than air, the shock in air becomes weak in water. Moreover, the density ratio is much larger and no smoothing of the initial interface is considered (as is often done when Eulerian methods are applied to the problem). This is a true test of the robustness of the methods. The solution contains shocks, multiple fluids, large density ratios, different equations of state, high speed in air, and low speed in water. In fact, with the present time step of 1 $\mu$s the PROE method diverged after 36 time steps. When the time step was reduced to one-tenth of a microsecond the solution diverged after 431 iterations, which is only 7.1 $\mu$s more of simulated physical time. Alternatively, the HYBR method experiences no problems in convergence and satisfies the convergence criteria throughout the computation, as shown in Fig. 20.

Examining the solution computed using the HYBR method, both pressure contours and numerical Schlieren images are displayed in Figs. 21–24. The max and min pressure contours are set to 100,000 Pa (blue) and 300,000 Pa (red). In Fig. 21, the shock

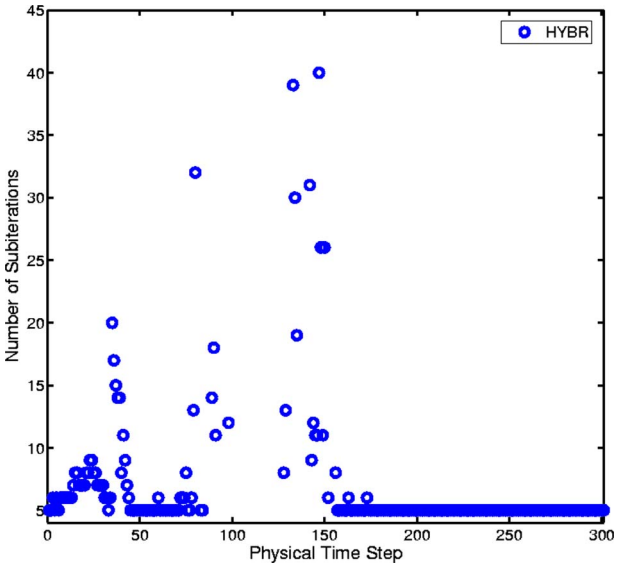



**Fig. 20 Subiteration convergence history of the PROE and HYBR methods for the $M_S$=1.22 shock air/water interaction problem**

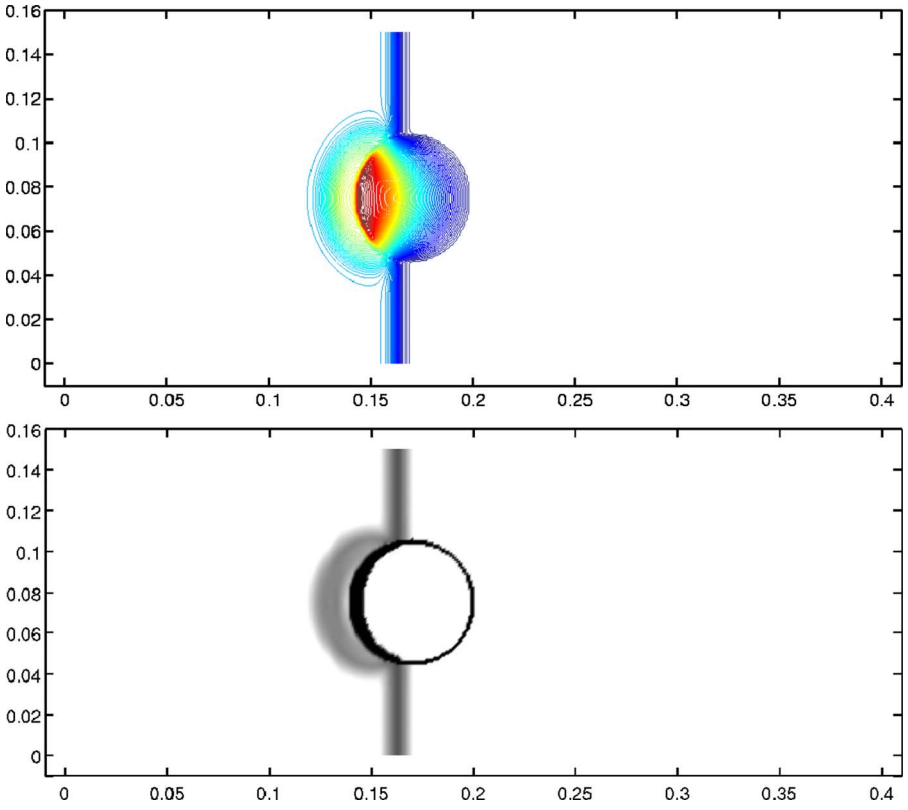



**Fig. 21 (Upper) pressure and (Lower) numerical Schlieren for air/water shock bubble interaction at time step 100 using the HYBR method**

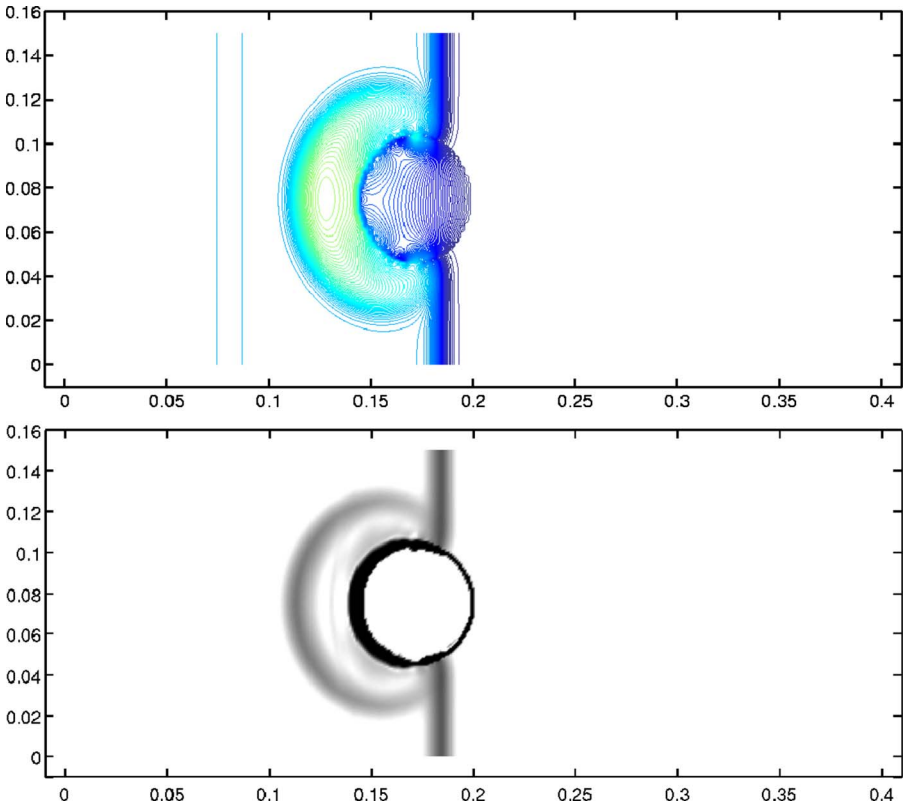



**Fig. 22 (Upper) pressure and (Lower) numerical Schlieren for air/water shock bubble interaction at time step 150 using the HYBR method**

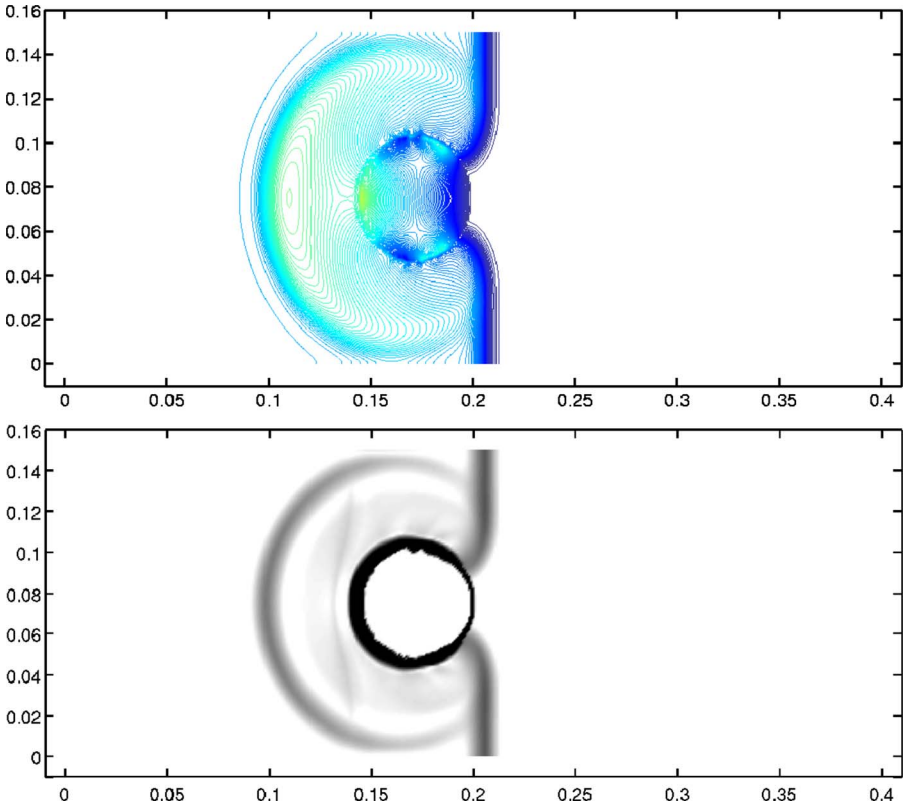



**Fig. 23 (Upper) pressure and (Lower) numerical Schlieren for air/water shock bubble interaction at time step 200 using the HYBR method**

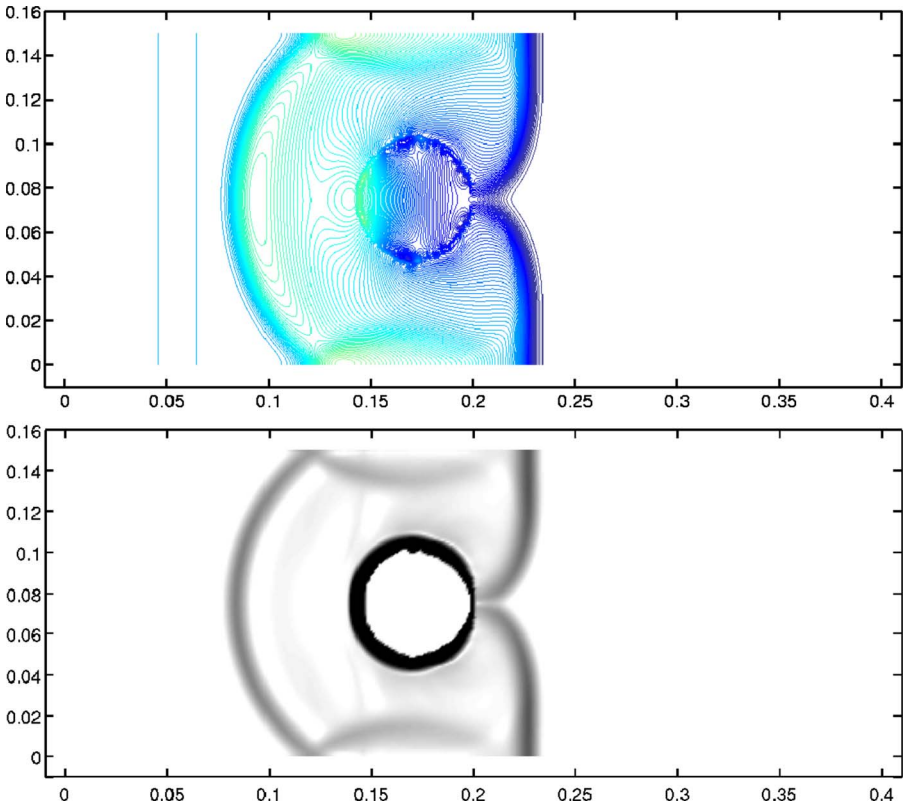



**Fig. 24 (Upper) pressure and (Lower) numerical Schlieren for air/water shock bubble interaction at time step 250 using the HYBR method**

has hit the bubble, but unlike the air/helium result, the shock can only weakly transmit into the water. Instead, a much larger reflected shock is observed near the front of the bubble. It is only the pressure contours that allow us to see the shock moving through the bubble. Figure 22 shows the shock wave after it has moved just over half the diameter of the bubble. The reflected shock is moving upstream and oscillations in pressure appear at the fluid interface. These oscillations are caused by the baroclinic generation of vorticity due to the misalignment of the pressure gradient and the density gradient. When the shock moves over the interface at different speeds inside and outside the bubble, the interface is perturbed leading the misalignment. If the numerical method produces nonphysical oscillations at the interface, as the PROE method has been shown to do, then there is no way to distinguish between the physical and nonphysical instabilities in the flow field. Figure 23 shows the initial shock as it leaves the bubble. The large change in fluid properties across the interface causes the shock wave in air to bend around the bubble as if it were a solid object. In Fig. 24, the shock has passed the bubble and the reflected shock has now reflected off the channel walls. In comparing the solutions of the air/water shock interaction with those of the air/helium results, it has been observed that a larger reflected shock is produced at the bubble interface, the strength of the transmitted shock is greatly reduced, and the outer shock bends around the fluid interface similar to the behavior observed when a shock passes over a solid cylinder. It has been demonstrated that the HYBR method is capable of predicting shock-wave dynamics correctly, deal with large density ratio fluid interfaces, and low speed flows.

## 3 Summary

A time-derivative preconditioned system of equations is applied to two multidimensional multicomponent and or multiphase flow applications. Three methods of spatial discretization are evaluated: a conservative preconditioned Roe method, a nonconservative preconditioned SCM method, and a hybrid method, which combines the conservative and nonconservative methods. It has been demonstrated that the conservative scheme lacks robustness when applied to multicomponent flows. The PSCM method remains robust when applied to multicomponent flows and is accurate for low speed flows, but converges to nonphysical solutions when shocks are present. The hybrid method, which combines the PROE and PSCM methods, retains the positive features of each of the methods resulting in a robust and efficient method for multicomponent flows at all speeds.

## References


[1] Kiris, C., Kwak, D., Chan, W., and Housman, J., 2008, "High Fidelity Simulations for Unsteady Flow Through Turbopumps and Flowliners," Comput. Fluids, **37**, pp. 536–546.

[2] Kiris, C., Chan, W., Kwak, D., and Housman, J., ""Time-Accurate Computational Analysis of the Flame Trench," ICCFD5, Seoul, South Korea, Jul. 7–11.

[3] Larrouturou, B., 1991, "How to Preserve the Mass Fractions Positivity When Computing Compressible Multi-Component Flows," J. Comput. Phys., **95**, pp. 59–84.

[4] Housman, J., Kiris, C., and Hafez, M., 2007, "Preconditioned Methods for Simulations of Low Speed Compressible Flows," Comput. Fluids, in print (available online).

[5] Salvetti, M.-V., and Beux, F., 2004, "Liquid Flow Around Non-Cavitating and Cavitating NACA0015 Hydrofoil," Numerical Workshop Problem From Mathematical and Numerical Aspects of Low Mach Number Flows Conference, Porquerolles, France.

[6] Rapposelli, E., Cervone, A., Bramanti, C., and d'Agostino, L., 2003, "Thermal Cavitation Experiments on a NACA0015 Hydrofoil," *Proceedings of FEDSM'03 4TH ASME/JSME Joint Fluids Engineering Conference*, Honolulu, HI.

[7] Li, D., and Merkle, C. L., 2004, "Liquid Flow Around Non-Cavitating and Cavitating NACA0015 Hydrofoil: Benchmark Problems for Gems Code," Numerical Workshop Problem from Mathematical and Numerical Aspects of Low Mach Number Flows Conference, Porquerolles, France, Jun. 21–25.

[8] Li, D., Venkateswaran, S., Lindau, J. W., and Merkle, C. L., 2005, "A Unified Computational Formulation for Multi-Component and Multi-Phase Flows," AIAA Paper No. AIAA-2005-1391.

[9] Chang, C.-H., and Liou, M.-S., 2005, "A Conservative Compressible Multifluid Model for Multiphase Flow: Shock-Interface Interaction Problems," AIAA Paper No. AIAA 2005-5344.

[10] Deshpande, M., Feng, J., and Merkle, C. L., 1993, "Navier–Stokes Analysis of 2-D Cavity Flows," ASME Cavitation and Multiphase Forum, Vol. FED-153, pp. 149–155.

[11] Toro, E. F., 1997, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer, New York.

[12] Quirk, J. J., and Karni, S., 1996, "On the Dynamics of a Shock-Bubble Interaction," J. Fluid Mech., **318**, pp. 129–163.

**Craig A. Steeves**
Postdoctoral Scholar

**Ming Y. He**
Project Scientist

Department of Materials,
University of California, Santa Barbara,
Santa Barbara, CA 93106

**Scott D. Kasen**
Graduate Student
Department of Materials Science and Engineering,
University of Virginia,
Charlottesville, VA 22904

**Lorenzo Valdevit**
Assistant Professor
Department of Mechanical and Aerospace Engineering,
University of California, Irvine,
Irvine, CA 92697

**Haydn N. G. Wadley**
Professor
Department of Materials Science and Engineering,
University of Virginia,
Charlottesville, VA 22904

**Anthony G. Evans**
Professor
Department of Materials,
University of California, Santa Barbara,
Santa Barbara, CA 93106

# Feasibility of Metallic Structural Heat Pipes as Sharp Leading Edges for Hypersonic Vehicles

*Hypersonic flight with hydrocarbon-fueled airbreathing propulsion requires sharp leading edges. This generates high temperatures at the leading edge surface, which cannot be sustained by most materials. By integrating a planar heat pipe into the structure of the leading edge, the heat can be conducted to large flat surfaces from which it can be radiated out to the environment, significantly reducing the temperatures at the leading edge and making metals feasible materials. This paper describes a method by which the leading edge thermal boundary conditions can be ascertained from standard hypersonic correlations, and then uses these boundary conditions along with a set of analytical approximations to predict the behavior of a planar leading edge heat pipe. The analytical predictions of the thermostructural performance are verified by finite element calculations. Given the results of the analysis, possible heat pipe fluid systems are assessed, and their applicability to the relevant conditions determined. The results indicate that the niobium alloy Cb-752, with lithium as the working fluid, is a feasible combination for Mach 6–8 flight with a 3 mm leading edge radius.* [DOI: 10.1115/1.3086440]

## 1 Background

For aerodynamic reasons hypersonic vehicles require sharp leading edges, with millimeter scale radius. When the edges are that sharp, the heat flux into the structure is intense. Specifically, at Mach 6–8 (the highest Mach number attainable with hydrocarbon fuels), the fluid stagnation temperature $T_{st} > 1400\,^{\circ}\mathrm{C}$, exceeding the realistic upper use temperature of most materials. Because the tip must remain sharp and have a stable shape, ablative solutions are very challenging. Instead, the heat must be rapidly redistributed through the solid to enable dissipation by radiation from the largest possible area of the vehicle surface. The three primary passive options to obtain such a solution are (i) carbon-fiber-based composites that retain load-bearing capability at $T \rightarrow T_{st}$, (ii) ultrahigh temperature ceramics (such as $HfB_2$) that combine refractoriness $T_M > T_{st}$ with large thermal conductivity at $T \rightarrow T_{st}$, and (iii) heat pipes, which enable allowable equilibrium temperatures by providing exceptional effective thermal conductivity, $k_{eff}$.

This assessment addresses the performance of heat pipes incorporated within metallic leading edge structures during steady-state hypersonic cruise. The preferred configuration has a curved surface fully defined by the thickness $t$, the leading edge radius $R_{le}$, and the angle $\phi_0$ (which is the complement of the wedge half-angle $\theta$) at which the curved front connects to the flat radiating surface (see Fig. 1). The design length of the radiating surface, $L$, must be chosen to ensure that the materials remain below their maximum use temperature.

Heat pipes have been pursued previously in the context of leading edges and other high temperature applications [1,2]. The present structural heat pipe differs in the sense that the design not only equilibrates the temperature but also supports transverse and shear loads (Fig. 2). For manufacturing facility and robustness, all-metallic designs are pursued, and their feasibility up to Mach 8 deduced.

Because of the crucial role of radiation, the temperatures induced are a strong function of the emissivity, $\epsilon$, of the radiating surface. While $\epsilon$ can be quite low for conventional alloys, the refractory alloys to be explored here can be designed to have larger values by pre-oxidizing to form either alumina or silica. The most well-documented are the nickel alloys used in turbines, which, when used with a bond coat, form a highly adherent, thin layer of $\alpha$-$Al_2O_3$ [3] with (at high temperature) $\epsilon \geq 0.9$.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 25, 2008; final manuscript received May 7, 2008; published online March 13, 2009. Review conducted by Martin Ostoja-Starzewski.

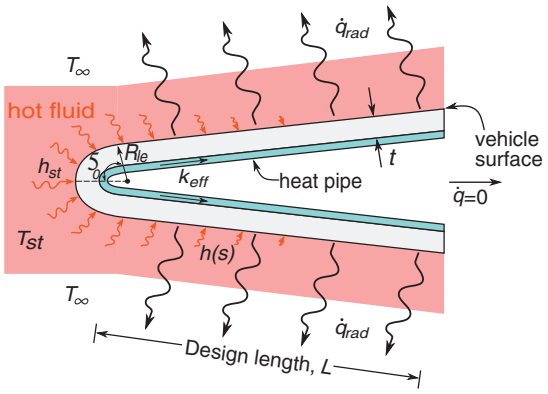



**Fig. 1 The local geometry and flow conditions near the leading edge**

To address this challenge, this paper is structured as follows: The aerothermodynamic environment that governs the heat flux into the leading edge is characterized using procedures based on the stagnation temperature and heat transfer coefficient. Steady-state temperatures and heat fluxes induced when the heat pipe is functioning are derived using analytic approximations with the fidelity assessed using selected finite element calculations. The designs to be explored are based on nickel (Inconel 625) and niobium (Cb-752) alloys having the properties in Table 1 and temperature-dependent yield strength characteristics presented in Fig. 3. Other refractory alloys, such as those based on molybdenum, tungsten, or rhenium, could be envisaged using the same basic protocol. The differing temperature and stress circumstances that arise during transients before the heat pipe begins to function are examined, as well as the influence of thin oxidation-protective coatings on the metallic surface. The operational requirements on the heat pipe are checked against models of heat pipe behavior to ensure the functionality of the system at the requisite heat flux and physical dimensions.

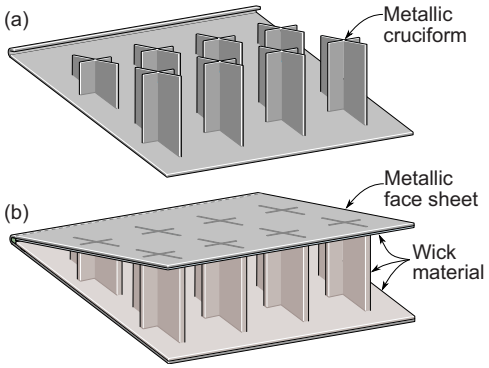



**Fig. 2 A structural heat pipe for the leading edge of a hypersonic vehicle showing (*a*) a cutaway view with the metallic wick material removed to show the cruciform structural members, and (*b*) the assembly with both metallic faces and the wick installed**

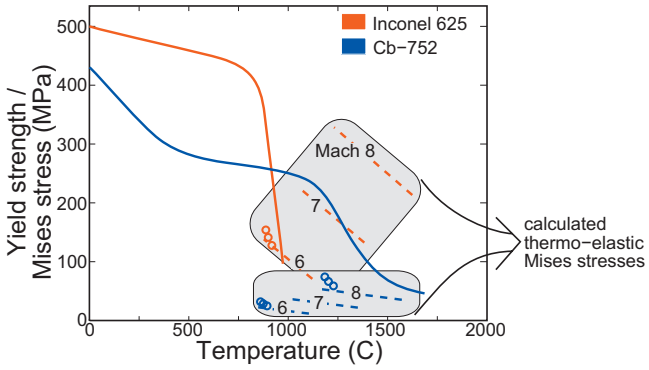



**Fig. 3 The temperature dependence of the yield strength for the two refractory alloys, Inconel 625 (red) and Cb-752 (blue), used in the analysis, shown as solid lines. Analytical cross-plots of maximum stress as a function of temperature are presented as dashed lines for Mach 6–8. The corresponding finite element results are plotted as circles.**

## 2 The Leading Edge Aerothermodynamic Environment

*Stagnation conditions.* For hypersonic vehicles powered by scramjet engines with hydrocarbon fuels, the Mach number at cruise ranges between 6 and 8. The relation between Mach number and altitude is determined by requiring that the vehicle flies at a constant dynamic pressure (typically 48 kPa [4]). For each altitude, the freestream temperature $T_\infty$, pressure $P_\infty$, and density $\rho_\infty$ are found in standard atmosphere tables [5], summarized in Table 2.

The total freestream enthalpy $H_\infty$ is related to the temperature and vehicle velocity, $u_\infty$, by

$$H_\infty = c_p T_\infty + u_\infty^2/2 \tag{1}$$

where $c_p = 1.04$ kJ/kg K is the specific heat of quiescent air. The pressure at the stagnation point is obtained by assuming that the kinetic energy of the fluid is converted completely to pressure, and that $P_{st} \gg P_\infty$ [6],

$$P_{st} \approx \rho_\infty u_\infty^2 \tag{2}$$

The stagnation temperature can be found using a Mollier diagram [7], which is implemented in the Hypersonic Airbreathing Propulsion (HAP) software [4]. The results are summarized in Table 2.

*Heat flux.* Herein is a strategy for determining the aerothermal loading. The premise is that the air close to the surface at the leading edge reaches the full gas stagnation temperature. Heat enters the leading edge at a rate governed by a spatially varying heat transfer coefficient. A heat pipe on the back surface of the leading edge transfers the heat to a flat region, from which the heat is then radiated into low temperature space (Fig. 1). The system is heat balanced, such that the total heat entering the vehicle through convection is equal to that leaving by radiation. The following describes how to calculate the heat transfer coefficient and heat flux.

**Table 1 Relevant material properties of nickel-based superalloy Inconel 625 and niobium alloy Cb-752. The temperature-dependent material properties are taken at 800°C.**

| | Density $\rho$ (kg/m$^3$) | Specific heat $c_p$ (J/kg K) | Thermal conductivity $k$ (W/m K) (at 800°C) | Thermal expansion $\alpha$ (ppm/K) (at 800°C) | Young's modulus $E$ (GPa) (at 800°C) |
|---|---|---|---|---|---|
| Inconel 625 | 8440 | 525 | 21 | 15 | 155 |
| Cb-752 | 9030 | 281 | 48 | 7.4 | 110 |

**Table 2 Atmospheric properties and calculated stagnation conditions for flight over the range of Mach 6 to 8 with constant dynamic pressure of 48 kPa**

| Mach No. | Altitude (km) | Temperature $T_\infty$ (K) | Pressure $P_\infty$ (Pa) | Density $\rho_\infty$ (kg/m$^3$) | Velocity $u_\infty$ (km/s) | Enthalpy $H_\infty$ (MJ/kg) | Stagnation temperature $T_{st}$ (K) |
|---|---|---|---|---|---|---|---|
| 6 | 26.93 | 223.5 | 1900 | 0.0296 | 1.80 | 1.85 | 1651 |
| 7 | 28.98 | 225.5 | 1396 | 0.0215 | 2.11 | 2.46 | 2122 |
| 8 | 30.76 | 227.3 | 1069 | 0.0160 | 2.42 | 3.16 | 2627 |

Fay and Riddell [8] used the results of Lees [9] to ascertain correlations for the heat flux at the stagnation point. Sutton and Graves [10] approximated these by

$$\dot{q} = (H_g - H_w)K\sqrt{\frac{P_{st}}{R_{le}}} \tag{3}$$

where $H_g$ is the local fluid enthalpy, $H_w$ is the local wall enthalpy, and $K$ is a heat transfer factor, which is a function of molecular weight, the mass fractions of the constituent gases, and a gas-dependent transport parameter. Sutton and Graves [10] used the results of Svehla [11] to calculate $K=3.6\times10^{-4}$ m$^{-1}$ kg$^{-1/2}$ for air. For the Mach numbers under investigation, $H_g \gg H_w$; consequently, by setting $H_w=0$ and $H_g \approx H_\infty$, Eq. (3) becomes the "cold-wall" heat flux. At the stagnation point,

$$\dot{q}_{cw,st} = H_\infty K\sqrt{\frac{P_{st}}{R_{le}}} \tag{4}$$

The heat transfer coefficient at this location $h_{st}$ can be approximated through

$$h_{st} = \dot{q}_{cw,st}/T_{st} \tag{5}$$

which varies around the curved leading edge as $h(\phi)=h_{st}\cos(\phi)$. Beyond the curved section, the heat transfer coefficient declines as $h(s)\sim s^{-1/2}$, calculated as shown in Fig. 4. The actual vehicle surface is given by the solid line. Equating the heat transfer at the transition from the curved to flat sections, the coefficient along the radiating surface becomes

$$h(s) = h_{st}\cos\phi_0\frac{\sqrt{R_{le}}}{\sqrt{s\tan\theta}} \tag{6}$$

These equations permit calculation of the fluid temperature and local heat transfer coefficient around the leading edge. Implicit in this approach is the assumption that $h$ is unaffected by the surface temperature $T_{surf}$ and, accordingly, can be used to determine the actual heat flux once $T_{surf}$ has been ascertained. The dependence of the stagnation point heat transfer coefficient on the leading edge radius for the Mach numbers of interest is displayed in Fig. 5; it is independent of the material properties of the leading edge.

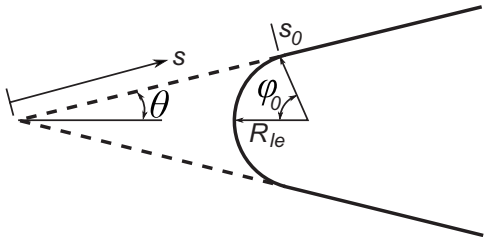



**Fig. 4 The assumed geometry for the calculation of the heat transfer coefficient along the flat radiating surface. The actual geometry is depicted with solid lines, while the dashed lines indicate the geometry for a perfectly sharp leading edge.**

## 3 Steady-State Temperatures and Thermal Stresses

**3.1 Analytical Estimates.** *Isothermal temperature*. When the heat pipe is functioning and the wall is thin, the solid approaches an isothermal temperature, $T_{iso}$, except very close to the leading edge. Thus, $T_{iso}$ becomes a useful reference temperature as well as the nominal temperature of the working fluid in the heat pipe. Under this thermal condition, the system is heat balanced. The spatial extent of the solid that attains $T_{iso}$ will be addressed later. The heat entering through the curved surface is given by

$$Q_{in}^c = \int_0^{\phi_0} R_{le}h_{st}\cos\phi(T_{st}-T_{iso})\mathrm{d}\phi \equiv R_{le}h_{st}\sin\phi_0(T_{st}-T_{iso}) \tag{7}$$

Through the flat surface the convective heat transfer is given by

$$Q_{in}^f = \int_{s_0}^{s_0+L} h(s)(T_{st}-T_{iso})\mathrm{d}s \equiv \frac{2h_\phi\sqrt{R_{le}}}{\tan\theta}(\sqrt{R_{le}+L\tan\theta}-\sqrt{R_{le}}) \times(T_{st}-T_{iso}) \tag{8}$$

with $h_\phi=h_{st}\cos(\phi_0)$ and $s$ the distance along the radiating surface as sketched in Fig. 4. The heat radiated out through the entire leading edge surface is

$$Q_{out} = \int_0^{L_{tot}} \epsilon\Sigma T_{iso}^4\mathrm{d}s \equiv \epsilon\Sigma L_{tot}T_{iso}^4 \tag{9}$$

where $L_{tot}$ is the total length of the heat pipe ($L_{tot}=R_{le}\phi_0+L$) and $\Sigma=5.67\times10^{-8}$ W/m$^2$ K$^4$ is the Stefan–Boltzmann constant. Heat balance requires that $Q_{in}=Q_{out}$, and hence

$$R_{le}h_{st}\sin\phi_0(T_{st}-T_{iso}) + \frac{2h_\phi\sqrt{R_{le}}}{\tan\theta}(\sqrt{R_{le}+L\tan\theta}-\sqrt{R_{le}})(T_{st}-T_{iso}) - \epsilon\Sigma L_{tot}T_{iso}^4 = 0 \tag{10}$$

which can be solved for $T_{iso}$. The results are plotted in Fig. 6(*a*), as a function of the length of the radiating surface, for $R_{le}$

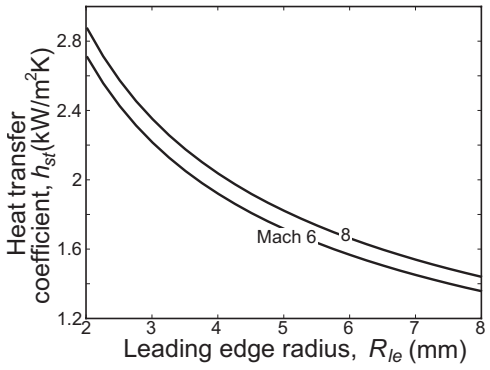



**Fig. 5 The stagnation point heat transfer coefficient as a function of the leading edge radius for a range of Mach numbers**

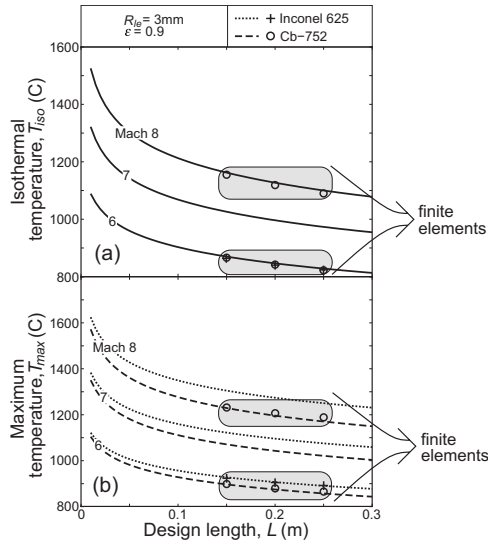



**Fig. 6 The leading edge temperatures calculated analytically (lines) over the relevant range of parameters, and by finite elements (circles and crosses) at Mach 6 and 8 for a vehicle with $R_{le}$=3 mm, $\epsilon$=0.9, and $\theta$=6 deg. (*a*) The isothermal temperature approximation as a function of the length of the radiating surface. Note that the analytical predictions are identical for the two materials, and that the finite element calculations at Mach 6 for Inconel 625 and Cb-752 fall on top of each other. (*b*) The maximum temperature at the stagnation point on the surface of the leading edge for Inconel 625 (dots and crosses) and Cb-752 (dashes and circles).**

=3 mm, $\epsilon$=0.9, and $\theta$=6 deg. Note that the isothermal temperature is sensitive to the design length $L$ but independent of the material chosen for the system.

*Maximum temperature.* The maximum temperature at the stagnation point can be estimated by assessing the heat flux across a small element, length $R_{le}d\phi$ (Fig. 7) by using a one-dimensional calculation for heat flow into a hollow cylinder. Implicit to this estimation is that the heat flux through the wall is much larger than that along the surface (justified below). The general equation for steady-state heat conduction is

$$k\left(\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}\right)=0 \tag{11}$$

with $k$ the thermal conductivity of the material and $r$ the radial distance from the center. The boundary conditions are set by the stagnation temperature and heat transfer coefficient of the gas, as well as by the premise that the internal surface is at temperature

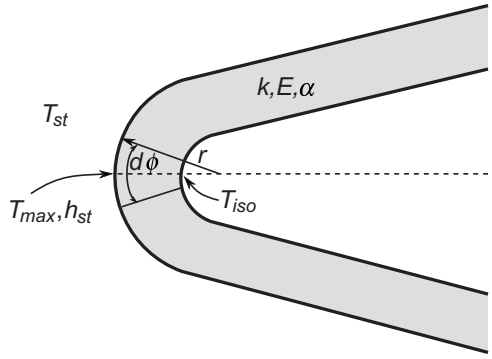



**Fig. 7 A sketch of the region along the stagnation line for which the temperatures are solved by a cylindrical finite difference scheme**

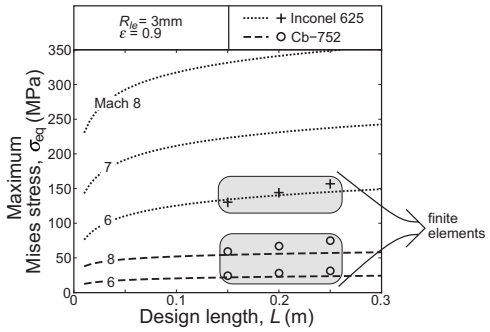



**Fig. 8 The maximum thermally induced Mises stress at the exterior surface along the stagnation plane for a vehicle with $R_{le}$=3 mm, $\epsilon$=0.9, and $\theta$=6 deg. The analytical results are lines, while the finite element results are circles and crosses.**

$T_{iso}$. The solution for the maximum material temperature at the outside surface is

$$T_{max}=\frac{T_{iso}+T_{st}\left(\frac{R_{le}h_{st}}{k}\ln\frac{R_{le}}{R_i}\right)}{1+\left(\frac{R_{le}h_{st}}{k}\ln\frac{R_{le}}{R_i}\right)} \tag{12}$$

where $R_i=R_{le}-t$ is the inner radius of the leading edge. The maximum temperatures $T_{max}$ for $R_{le}$=3 mm, $t$=1 mm, and $\epsilon$=0.9, for Inconel 625 ($k$=21 W/m K) and Cb-752 ($k$=48 W/m K) at Mach 6–8, are plotted in Fig. 6(*b*) for a range of design lengths. This temperature is a function of the material thermal conductivity with higher conductivity leading to lower $T_{max}$. The corresponding heat flux through the surface into the heat pipe is

$$\dot{q}=\frac{k(T_{max}-T_{iso})}{R_i\ln\frac{R_{le}}{R_i}} \tag{13}$$

*Stresses.* Because the structure is metallic, the Mises stress near the leading edge (where the temperature is largest), relative to the yield strength, governs the integrity. To estimate the thermoelastic stress (no yielding), the reference state is taken to be the long straight segment at $T_{iso}$. The thermal expansion misfit of the tip region, relative to this state, generates a stress. Neglecting bending (justified later), the stress is dominated by the out-of-plane strain misfit and varies through the thickness as

$$\sigma_{eq}(r)=\alpha E(T(r)-T_{iso}) \tag{14}$$

The maximum at the exterior surface is plotted in Fig. 8 over the relevant range of Mach numbers. The maximum increases with larger $L$ because, as the length increases, the total heat radiated away increases, *despite the decreased overall temperature*, requiring a larger thermal gradient at the leading edge to provide the greater heat flux. Hence, as the design length increases, the maximum temperature decreases while the maximum stress increases. Cross-plots of the relation between maximum stress and maximum temperature are shown in Fig. 3 for comparison with the material yield strength at that temperature. For longer design lengths (toward the left side of the graph) and lower Mach numbers, the operational conditions are clearly within the elastic envelope of the materials. However, above Mach 6, the stresses exceed the yield strength of Inconel 625 at the operating temperatures, thereby excluding this alloy from use. Conversely, Cb-752 retains its integrity even up to Mach 8, due to its combination of thermomechanical properties discussed below.

*Without a heat pipe.* In the absence of a heat pipe, heat is transferred only by longitudinal conduction along the curved region into the radiating surface, resulting in the temperature distributions plotted in Fig. 9 for Mach 7 flight for a Cb-752 leading

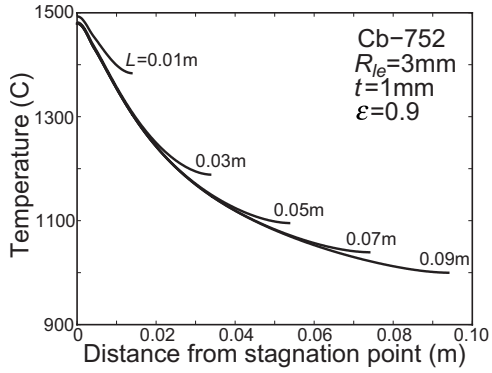



**Fig. 9 The behavior of the leading edge in the absence of a heat pipe: The temperature distribution along the length of an Cb-752 ($k$=48 W/m K) leading edge at Mach 7 for a range of design lengths $L$, assuming that the material is isothermal through the thickness and that $\epsilon$=0.9, $R_{le}$=3 mm, and $t$ =1 mm**

edge. A steep temperature gradient develops along the length and the maximum temperatures become insensitive to $L$. Comparison with Fig. 6 demonstrates the importance of having a functioning heat pipe.

**3.2 Finite Element Calculations.** *Method*. All calculations are conducted using the commercial finite element code ABAQUS. In the analysis, eight-node coupled temperature-displacement generalized plane strain elements are used. Symmetry conditions are imposed on $AD$ in the $y$-direction and on $CF$ in the $x$-direction (Fig. 10). The solids are incorporated as an elastic/plastic medium having temperature-dependent yield strength (Fig. 3) with nominal strain hardening.

The external thermal boundary conditions are the same as those described in Sec. 2. Outside the entire surface, the air is assigned the stagnation temperature, with a heat transfer coefficient that varies spatially over $\phi$ and $s$. This temperature difference with that computed for the solid surface, $T_{surf}$, then governs the inward heat flux. Radiation to ambient $T_\infty$ from the surface, at $T_{surf}$, determines the outward heat flux. ABAQUS simultaneously solves for the net flux and $T_{surf}$, as well as the Mises stresses and (when yielding occurs) equivalent plastic strains. Conduction through the solid is characterized by a (temperature invariant) thermal conductivity, $k$. The action of the internal heat pipe is simulated by imposing a thin (1 mm) compliant layer on the inside having exceptional effective thermal conductivity, $k_{eff}$ (Fig. 1). Effective thermal conductivities in excess of 100 times that of solid copper have been reported [12]. A conservative value of $k_{hp}$ =3 kW/m K was used in all simulations. The thermo-elastic properties of Inconel 625 and Cb-752 used in the calculations are those presented in Table 1 and Fig. 3. Results are presented for a

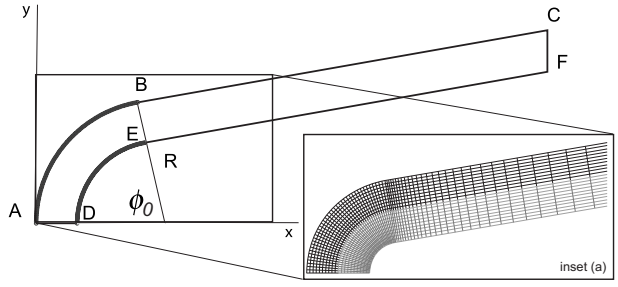



**Fig. 10 A typical finite element mesh used in the simulations. The solid metallic component is represented in black, while the heat pipe is in gray.**

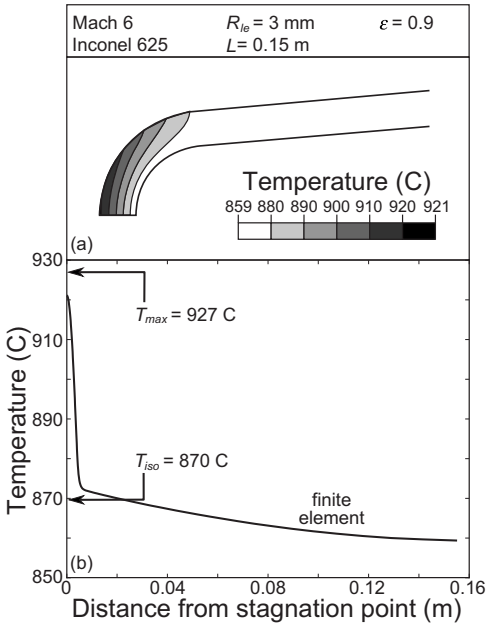



**Fig. 11 (*a*) Contours of the temperatures induced in the nickel alloy at Mach 6 when the heat pipe is functioning. (*b*) The temperatures along the external surface with analytic results for $T_{iso}$ and $T_{max}$ for $\epsilon$=0.9 are superimposed.**

leading edge with radius $R_{le}$=3 mm and wedge half-angle $\theta$ =6 deg.

*Temperatures and heat fluxes*. In order to compare with the analytic estimates, the temperatures have been obtained for the nickel alloy at Mach 6 and the niobium alloy at Mach 6 and 8 (as a function of design length). Representative temperature distributions are plotted in Figs. 11 and 12 (for $L$=0.15 m). The straight sections, as well as the inside of the curved tip, are at essentially uniform temperature, identified with $T_{iso}$. The outside temperature at the tip is appreciably hotter, with largest value, $T_{max}$, at the

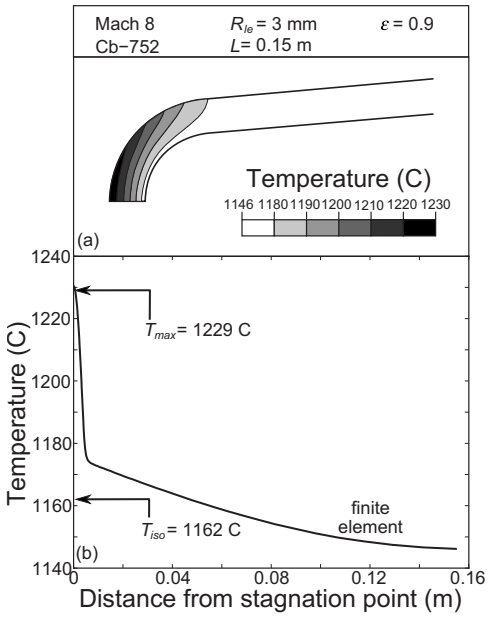



**Fig. 12 (*a*) Contours of the temperatures induced in the niobium alloy at Mach 8 when the heat pipe is functioning. (*b*) The temperatures along the external surface with analytic results for $T_{iso}$ and $T_{max}$ for $\epsilon$=0.9 are superimposed.**

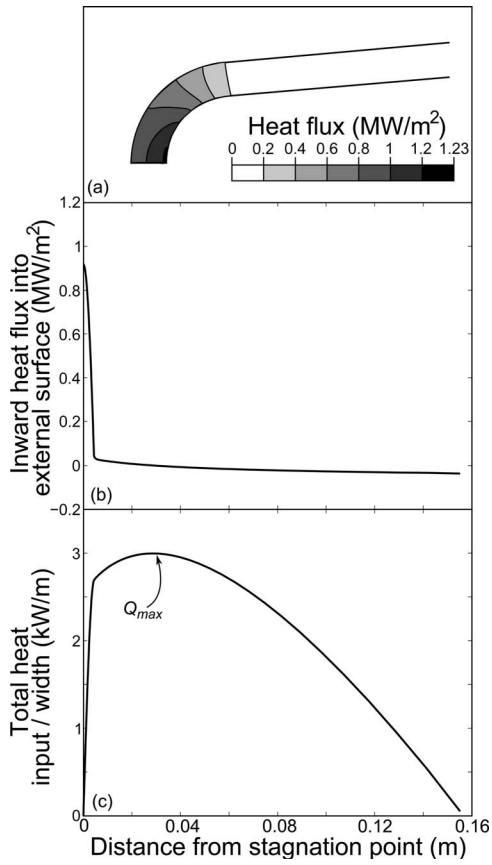



**Fig. 13 Heat flows at Mach 6 for an Inconel 625 leading edge. (*a*) Contours of heat flux at the design length $L=0.15$ m and $\epsilon=0.9$. (*b*) Local heat flux into the exterior surface. (*c*) Integrated heat input from the leading edge. At the design length the net total heat input is zero.**

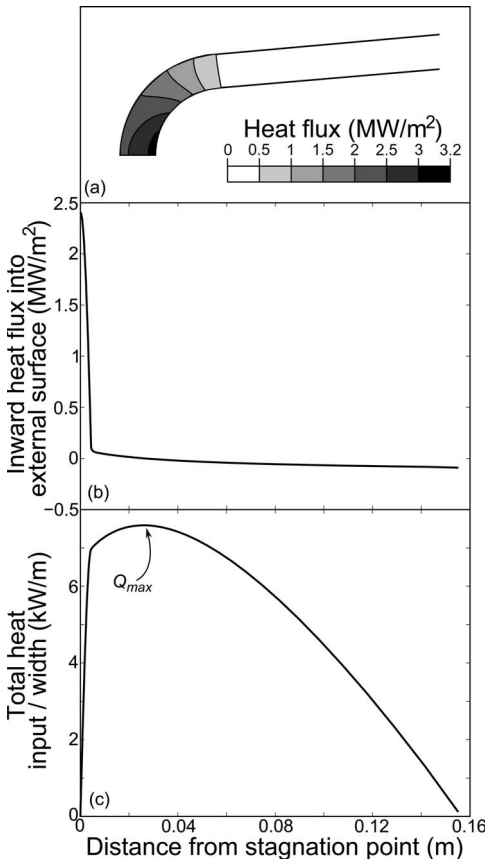



**Fig. 14 Heat flows at Mach 8 for a Cb-752 leading edge. (*a*) Contours of heat flux at the design length $L=0.15$ m and $\epsilon=0.9$. (*b*) Local heat flux into the exterior surface. (*c*) Integrated heat input from the leading edge. At the design length the net total heat input is zero.**

stagnation plane. Temperatures defined in this manner are compared with the analytic estimates on the figures. The close correspondence demonstrates the utility of the analytical approach for initial design purposes. Moreover, temperature comparisons along the flat section affirm that $T_{iso}$ is independent of the material choice. Recalling that $T_{iso}$ dictates the choice of the working fluid, then for a 3 mm radius, at Mach 6, either sodium or potassium should be applicable, while at Mach 8, lithium would be preferred, as elaborated later.

The corresponding heat fluxes are summarized in Figs. 13 and 14. Those at the external surface are directed inward over the curved section but outward over most of the flat segment. Consequently, the cumulative heat input reaches a maximum just beyond the transition and declines thereafter; by definition, the net heat input is zero at the design length. The maxima are the most pertinent since these are the fluxes that must be redistributed by the working fluid in the heat pipe. These are $Q_{max}=3$ kW/m at Mach 6 and $Q_{max}=7.6$ kW/m at Mach 8, with approximate scaling, $Q_{max} \sim u_\infty^3$. The contours over the curved region affirm that the flux entering the heat pipe near the stagnation plane is largest: $\dot{q}_{max} \approx 1.2$ MW/m$^2$ at Mach 6 and $\dot{q}_{max} \approx 3.2$ MW/m$^2$ at Mach 8. (Note that these are significantly lower than the cold-wall fluxes typically cited for leading edges at the same Mach number, altitude and radius.)

*Mises stresses*. Contours of the Mises stresses are plotted in Fig. 15. The maxima at the tip are about 10% larger than the analytic estimates (Fig. 8). The difference is attributed to a small bending effect not accounted for in the analytic formula. The results for Inconel 625 (Fig. 15(*a*)) reveal that, even at Mach 6, the stresses approach the yield strength at the associated temperature (see Fig. 3). Indeed, the most failure susceptible element at the tip on the exterior surface is sufficiently close to yield that the safety margin would be unacceptable. The corresponding Mises stress contours for the niobium alloy (Fig. 15(*b*)) reveal much larger margins of safety, even at Mach 8. The niobium alloy is preferable because of its higher conductivity, lower modulus, and lower thermal expansion, all of which decrease the maximum stress.

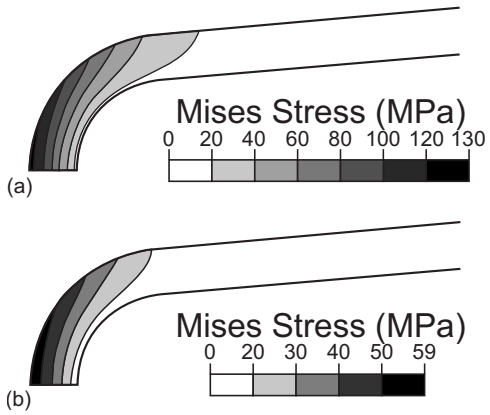



**Fig. 15 Mises stress contours when the heat pipe is functioning: (*a*) Mises stresses for Mach 6 with nickel-based superalloy Inconel 625; and (*b*) Mises stresses for Mach 8 with niobium alloy Cb-752**

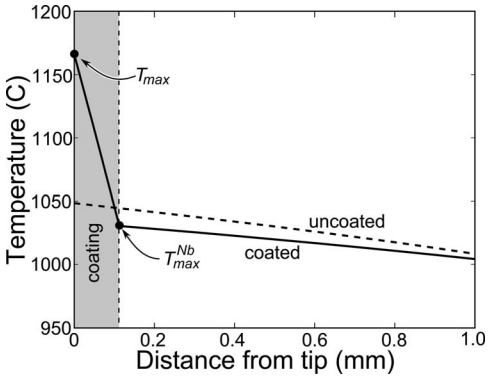



**Fig. 16 The temperature distribution through a niobium leading edge with a 100 $\mu$m anti-oxidation environmental barrier coating for a Mach 7 vehicle**

Coupled with greater yield strength at higher temperatures, niobium alloys are clearly a superior design choice.

## 4 Influence of Coatings and Transients

*Coatings*. The nickel and niobium alloys to be used may require oxidation protection coatings to assure multiflight capability. This is especially true for the Nb alloys being contemplated for use at the higher Mach numbers. To examine the possible effect of these coatings on temperatures in the system, a numerical calculation has been conducted for a Mach 7 design using a Cb-752 alloy with an oxidation protection coating having thickness $t_{coat}$ $=100$ $\mu$m. The coating is considered to be highly insulating, with thermal conductivity, $k=1$ W/m K. The temperature results are summarized in Fig. 16. Note that the maximum temperature developed in the niobium has been reduced relative to that without the coating, but the coating surface becomes extremely hot. Moreover, a large temperature gradient is induced in the coating, with high likelihood of delamination [13]. The impact of coatings on the viability of the envisioned leading edge systems thus requires careful assessment that balances their benefits in environmental protection with their adverse susceptibility to spalling.

*Transients.* Time-dependent calculations of heat diffusion through the leading edge elucidate the transient behavior of the system at startup. The leading edge at ambient temperature is instantaneously exposed to flight conditions. To activate the heat pipe, a minimum critical temperature, $T_{cr}$, must be attained at the back face; subsequently it provides heat transport dependent on the ensuing back face temperature. The heat pipe is assumed to obey a heat flux law

$$\dot{q} = \left( \frac{T_{bf} - T_{cr}}{T_{iso} - T_{cr}} \right)^m \dot{q}_{ss} \tag{15}$$

where $T_{bf}$ is the instantaneous temperature at the back face, $\dot{q}_{ss}$ is the heat flux at the steady state, as calculated by Eq. (13), and $m$ is an exponent governing the rate of heat pipe startup. Once the back face reaches $T_{iso}$, the heat flux in the system is in equilibrium. The time-dependent thermal behavior of the Mach 7 Cb-752 system is shown in Fig. 17(*a*) for $T_{cr}/T_{iso}=0.9$ and $m=1$.

Three transient phases are evident.

*Phase I*. The temperature of the outer surface rises quickly. The back face temperature rises more slowly, as heat diffuses through the thickness, causing an initial rapid stress elevation (see Fig. 17(*b*)); this phase concludes with a local maximum in stress.

*Phase II*. The system moves toward thermal equilibrium. The Mises stress gradually declines because the temperature of the back face increases more quickly than the temperature of the front face.

*Phase III*. The heat pipe begins to operate. This activation dramatically slows the rise in back face temperature, as the Mises stress increases to a global maximum. These calculations suggest that the maximum stress during the transient phase does not exceed that at steady state. However, the results are highly dependent on the choice of the startup temperature and the rate exponent. Experiments and detailed modeling of the heat pipe are needed to choose these accurately.

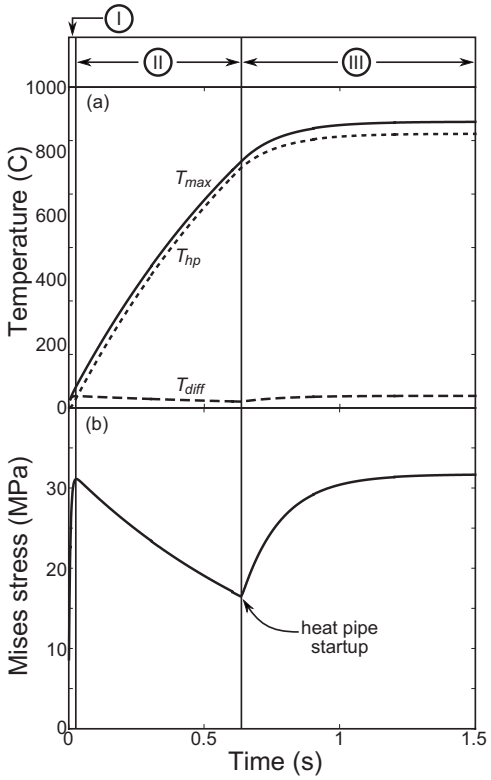



**Fig. 17 The idealized behavior of the leading edge during the startup phase for a Cb-752 leading edge at Mach 7. (*a*) The temperature at the external surface $T_{max}$, at the back surface $T_{hp}$, and the difference between the two temperatures $T_{diff}$. (*b*) The resulting Mises stress at the stagnation point.**

## 5 Heat Pipe Limitations

Heat pipes can fail when the heat transfer rate within the pipe is insufficient to transport the incident heat flux. Several operational phenomena affect this limit [12,14]. Only three are relevant to the leading edge environment: (i) the *sonic limit*, encountered when the mean vapor flow velocity approaches transonic values; (ii) the *capillary limit*, which arises when the drops in liquid and vapor pressures approach the capillary pumping pressure available within the wick; and (iii) the *boiling limit*, occurring when a critical superheating of the vapor is attained and bubbles stabilize in the wick of the evaporator zone. Detailed analyses of these phenomena, which relate these limits to the thermophysical properties of the working fluid, were derived for a hollow cylindrical heat pipe [12,14,15], which is not directly applicable here. For the leading edge, the corresponding analysis has been conducted for the geometry depicted in Fig. 18, and the upper portion of the leading edge is treated as a planar heat spreader. The leading edge is constructed either from Inconel 625 with thermochemically compatible sodium as the working fluid, or Cb-752 with lithium as the working fluid. The wick system is a woven wire mesh of the same material as the case.

**5.1 Sonic Limit.** The sonic limit is reached when the absorbed thermal flux per unit width is given by [14]

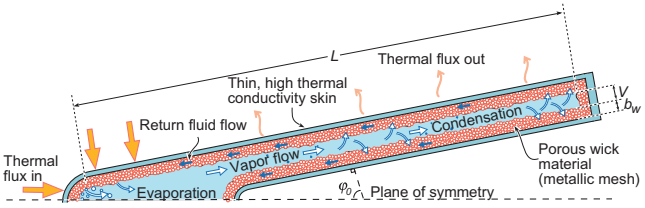



**Fig. 18 A cross-sectional schematic showing the operating principles and relevant geometry of a heat plate leading edge**

$$Q_{\max}^{\text{son}} = V\rho_v\lambda\left(\frac{\gamma R_g T_{\text{iso}}}{2(\gamma+1)}\right)^{1/2} \tag{16}$$

where $V$ is the vapor space height (see Fig. 18), $\rho_v$ is the vapor density, $\lambda$ is the latent heat of vaporization of the working fluid, $\gamma$ is the heat capacity ratio ($\gamma=1.67$ for monatomic sodium and lithium vapor), and $R_g$ is the gas constant for the vapor species within the heat pipe ($R_{\text{Na}}=361$ J/kg K and $R_{\text{Li}}=120$ J/kg K). The isothermal operating temperature of the heat pipe $T_{\text{iso}}$ (which equals the liquid-vapor saturation temperature) is limited by the onset of rapid softening and high creep rate of the case material. Note that only at low $T_{\text{iso}}$ is the sonic limit likely to restrict the heat transport rate.

**5.2 Capillary Limit.** The capillary limit can be derived by a pressure balance across the length of the heat pipe

$$\Delta P_{\max}^{\text{cap}} = \Delta P_l + \Delta P_v \tag{17}$$

where $\Delta P_{\max}^{\text{cap}}$ is the maximum capillary pumping pressure, with $\Delta P_l$ and $\Delta P_v$ the liquid and vapor pressure drops, respectively. The additional pressure drops attributed to inertial effects upon evaporation and condensation of the working fluid are deemed negligible [14]. The maximum capillary pumping pressure is given by the Young–Laplace equation [12,14,15] as

$$\Delta P_{\max}^{\text{cap}} = \frac{2\sigma}{R_{\text{eff}}} \tag{18}$$

where $\sigma$ is the liquid surface tension and $R_{\text{eff}}$ is the "effective pore radius" (the radius of curvature at the liquid-vapor interface in the evaporator zone), which is a function of the porous wick geometry. Its value has been experimentally determined for numerous wick structures and porosities. The assumed effective pore radius is $R_{\text{eff}}=1.27\times10^{-4}$ m, which corresponds to a woven mesh wick with ~4 cells/mm [14].

The maximum pressure drop in the liquid within the wick can be calculated from Darcy's model of laminar flow through a porous medium

$$\Delta P_l = \frac{\eta_l Q_{\max}^{\text{cap}}}{4\kappa b_w \rho_l \lambda} L_{\text{tot}} \tag{19}$$

where $\eta_l$ is the liquid dynamic viscosity, $\kappa$ is the wick permeability, $b_w$ is the wick thickness, and $\rho_l$ is the liquid density. The woven mesh wick with four unit cells per millimeter has permeability of $\kappa=1.93\times10^{-10}$ m$^2$ [14]. The absence of an adiabatic section implies that the evaporator zone length $L_e$ and the condenser zone length $L_c$ together occupy the entirety of the heat pipe such that $L_{\text{tot}}=L_e+L_c$.

When inertial forces dominate viscous forces, and the vapor is in the incompressible, laminar flow regime, as is the case here, the maximum vapor pressure drop can be derived from the governing equations describing the velocity distribution for fully developed fluid flow in a rectangular duct

$$\Delta P_v = \frac{6\eta_v Q_{\max}^{\text{cap}}}{V^3\rho_v\lambda} L_{\text{tot}} \tag{20}$$

where $\eta_v$ is the vapor dynamic viscosity.

Substituting Eqs. (18)–(20) into Eq. (17) and solving for $Q_{\max}^{\text{cap}}$ give an expression for the maximum absorbed thermal flux per unit width when constrained by capillary pumping limits

$$Q_{\max}^{\text{cap}} = \frac{4\sigma\lambda}{L_{\text{tot}}R_{\text{eff}}}\left(\frac{12\eta_n}{V^3\rho_v} + \frac{\eta_l}{2\kappa b_w\rho_l}\right)^{-1} \tag{21}$$

Liquid sodium has a heat of vaporization two orders of magnitude lower than liquid lithium at the same temperature, suggesting that the capillary limit is particularly relevant to the sodium-Inconel system.

**5.3 Boiling Limit.** The boiling limit is reached first where the thermal flux is largest, at the point where the stagnation line intercepts the inner surface of the heat pipe wall (see Figs. 13(*a*) and 14(*a*)). From direct application of Fourier's law, the boiling limit is given by

$$\dot{q}_{\max}^{\text{boi}} = \frac{k_w}{b_w}\Delta T_{\text{crit}} \tag{22}$$

where $k_w$ is the thermal conductivity of the saturated wick. The critical superheat, $\Delta T_{\text{crit}}$, is given by [14]

$$\Delta T_{\text{crit}} = \frac{2\sigma T_{\text{iso}}}{\lambda\rho_v}\left(\frac{1}{R_b} - \frac{1}{R_{\text{eff}}}\right) \tag{23}$$

where $R_b=10^{-7}$ m is an order-of-magnitude estimation of the bubble radius at nucleation. The saturated thermal conductivity for a sintered woven mesh wick is given by

$$k_w = k_l\left(\frac{k_s}{k_l}\right)^{(1-\beta)^{0.59}} \tag{24}$$

where $k_l$ is the liquid thermal conductivity, $k_s$ is that for the solid, and $\beta$ is the wick porosity ($\beta=0.63$). The maximum heat flux at the stagnation line, $\dot{q}_{\max}$, must be less than the allowable maximum local heat flux, $\dot{q}_{\max}^{\text{boi}}$.

For the geometry under consideration, the heat flux into the heat pipe is not uniform. We define the evaporator length as $L_e = Q_{\max}/\dot{q}_{\max}$, where $\dot{q}_{\max}$ is the maximum heat flux entering the vapor from the heat pipe wall at the stagnation point (see Figs. 13(*a*) and 14(*a*)). The total heat, which can be transferred from the wall to the fluid, is conservatively given by

$$Q_{\max}^{\text{boi}} = \dot{q}_{\max}^{\text{boi}} L_e \tag{25}$$

**5.4 Operational Limit Comparisons.** The sonic, capillary, and boiling limits are functions of temperature-dependent thermophysical working fluid properties, namely, the liquid and vapor densities, heat of vaporization, liquid surface tension, and the liquid and vapor viscosities. Values of these properties for sodium and lithium can be found in Ref. [16], resulting in the limits for the sodium/Inconel 625 and lithium/Cb-752 systems plotted in Figs. 19 and 20. The results in Figs. 13 and 14 give $L_e$ ~2.4 mm for both the Mach 6 and Mach 8 designs, and vapor space height $V=4.5$ mm and wick thickness of $b_w=1$ mm have been assumed. Both plots affirm that the heat pipes are functional under the proposed flight velocities and altitudes. While other limitations on the heat pipe operation exist, such as acoustic fluctuations due to aerodynamic loadings, they are currently beyond the purview of the standard models presented here.

## 6 Concluding Comments

This paper contains a systematic method for calculating heat fluxes, temperatures and thermal stresses in a sharp leading edge of a hypersonic vehicle which has an integrated planar heat pipe. The boundary conditions are ascertained through the use of standard expressions involving the flight velocity and altitude, the atmospheric properties, and the geometry of the vehicle. The temperatures and stresses in the leading edge are calculated using simple approximations, which are verified with finite element

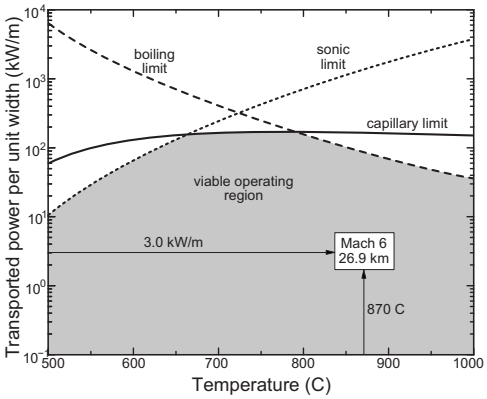



**Fig. 19 The predicted operational limits of a sodium-Inconel 625 heat plate leading edge with thermal flux leading edge input corresponding to Mach 6 (26.93 km, $\dot{q}_{max} \approx 1.2$ MW/m$^2$)**

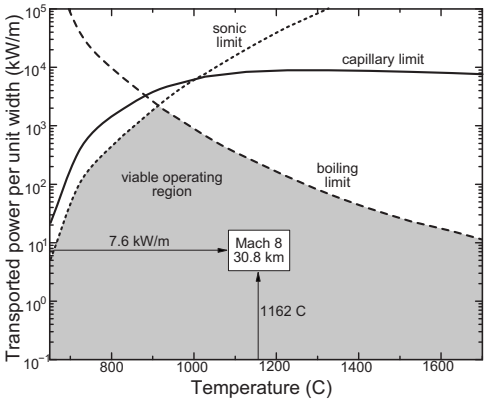



**Fig. 20 The predicted operating limits of a lithium-Cb-752 heat plate leading edge with thermal flux inputs corresponding to Mach 8 (30.76 km, $\dot{q}_{max} \approx 3.2$ MW/m$^2$)**

simulations. Conventional heat pipe estimates are then used to show the feasibility of planar metallic leading edge heat pipes. It is shown from these results that the niobium alloy Cb-752 is a better heat pipe material than the nickel alloy Inconel 625, and that a Cb-752 heat pipe with a lithium working fluid is a feasible choice for a 3 mm radius leading edge at Mach 8 or below.

The validity of the results presented here is highly dependent on choosing the correct thermal boundary conditions. The Fay–Riddell equations [8] are the most widely accepted technique for calculating leading edge heat flux, and the Sutton–Graves correlation [10] is a good approximation of these results. This paper has extrapolated from the cold-wall heat flux to determine a hot-wall heat transfer coefficient. It is believed that this is a conservative extrapolation; that is, the heat transfer coefficient should not increase with decreasing temperature gradient.

The isothermal approximation used in the analysis assumes the ideal functioning of the heat pipe. This assumption cannot be validated through the finite element simulations performed here. Instead, a full model of the internal workings of the heat pipe is required, including evaporation, condensation, and vapor and liquid transports. Such a model is beyond the scope of this paper. However, such validation can also be provided experimentally, so planar heat pipes are being built and tested in a laser facility, which delivers high heat flux to the leading edge. While these experiments cannot replicate the boundary conditions encountered during flight, the validity of the isothermal assumption and the expression for the maximum temperature can be ascertained by implementing the corresponding boundary conditions in the analytical models.

## Acknowledgment

The authors would like to thank Hossein Haj-Hariri of the Department of Mechanical and Aerospace Engineering, University of Virginia; David Marshall of Teledyne Scientific; Vince Cuda of Swales Aerospace; and George Jefferson of the Air Force Research Laboratory for discussions and very helpful assistance with this work. Funding has been provided by the Office of Naval Research through the MURI program Revolutionary Materials for Hypersonic Flight (Contract No. N00014-05-1–0439).

## References

[1] Glass, D. E., Camarda, C. J., Merrigan, M. A., and Sena, J. T., 1999, "Fabrication and Testing of Mo-Re Heat Pipes Embedded in Carbon/Carbon," J. Spacecr. Rockets, **36**(1), pp. 79–86.
[2] Glass, D. E., Camarda, C. J., Merrigan, M. A., Sena, J. T., and Reid, R. S., 1999, "Fabrication and Testing of a Leading-Edge-Shaped Heat Pipe," J. Spacecr. Rockets, **36**(6), pp. 921–923.
[3] Evans, A. G., Mumm, D. R., Hutchinson, J. W., Meier, G. H., and Pettit, F. S., 2001, "Mechanisms Controlling the Durability of Thermal Barrier Coatings," Prog. Mater. Sci., **46**, pp. 505–553.
[4] Heiser, W. H., and Pratt, D., 1993, *Hypersonic Airbreathing Propulsion*, American Institute of Aeronautics and Astronautics, Washington, DC.
[5] COESA, 1976, *U.S. Standard Atmosphere*, U.S. Government Printing Office, Washington, D.C.
[6] Bertin, J. J., 1994, *Hypersonic Aerothermodynamics*, American Institute of Aeronautics and Astronautics, Washington, DC.
[7] Moeckel, W. E., and Weston, K. G., 1958, "Composition and Thermodynamic Properties of Air in Chemical Equilibrium," Technical Note 4265, National Advisory Committee for Aeronautics, Lewis Flight Propulsion Laboratory, Cleveland, OH.
[8] Fay, J. A., and Riddell, F. R., 1958, "Theory of Stagnation Point Heat Transfer in Dissociated Air," J. Aeronaut. Sci., **25**(2), pp. 73–85.
[9] Lees, L., 1956, "Laminar Heat Transfer Over Blunt-Nosed Bodies at Hypersonic Flight Speeds," Jet Propul., **26**(4), pp. 259–269.
[10] Sutton, K., and Graves, R. A., 1971, "A General Stagnation-Point Convective-Heating Equation for Arbitrary Gas Mixtures," National Aeronautics and Space Administration, Langley Research Center, Technical Report No. R-376.
[11] Svehla, R. A., 1962, "Estimated Viscosities and Thermal Conductivities of Gases at High Temperatures," National Aeronautics and Space Administration, Lewis Research Center, Technical Report No. R-132.
[12] Peterson, G. P., 1994, *An Introduction to Heat Pipes: Modeling, Testing, and Applications*, Wiley, New York.
[13] Evans, A. G., and Hutchinson, J. W., 2007, "The Mechanics of Coating Delamination in Thermal Gradients," Surf. Coat. Technol., **201**(18), pp. 7905–7916.
[14] Faghri, A., 1995, *Heat Pipe Science and Technology*, Taylor & Francis, London.
[15] Chi, S. W., 1976, *Heat Pipe Theory and Practice*, Hemisphere, Washington, DC.
[16] Ohse, R. W., 1985, *Handbook of Thermodynamic and Transport Properties of Alkali Metals* (International Union of Pure and Applied Chemistry, Chemical Data Series No. 30), Blackwell Scientific, Oxford, UK.

# An Experimental Study of Contact Forces During Oblique Elastic Impact

**Philip P. Garland**[1]
e-mail: phil.garland@unb.ca

**Robert J. Rogers**
Professor

Department of Mechanical Engineering,
University of New Brunswick,
Fredericton, NB, E3B 5A3, Canada

*Low and high speed impacts frequently occur in many mechanical processes. Although widely studied, rarely are normal and tangential force time-waveforms measured, as generally these are very difficult measurements to do accurately. This paper presents, for the first time, a comprehensive set of experimentally obtained contact force waveforms during oblique elastic impact for a range of initial velocities and incidence angles. The experimental apparatus employed in this study was a simple pendulum consisting of a spherical steel striker suspended from a steel wire. The contact force time-waveforms were collected using a tri-axial piezoelectric force transducer sandwiched between a spherical target cap and a large block. The measured contact forces showed that loading was essentially limited to the normal and tangential directions in the horizontal plane. Analysis of the maximum normal and tangential forces for the near glancing angles of incidence indicated a friction coefficient that varies linearly with initial tangential velocity. The essential features of tangential force reversal during impact predicted by previous continuum models are confirmed by the experimental force results.*
[DOI: 10.1115/1.3063634]

*Keywords: oblique elastic impact, friction, experimental contact forces*

## 1 Introduction

Impact, or stereomechanics, has played a prominent role in the early studies of applied dynamics, and many famous figures such as Galileo, Mariotte, Descartes, Huygens, Newton, and Euler are known to have been interested in both the experimental and modeling aspects of impact [1]. Today, there are still many aspects of impact mechanics that are not well understood. The past 30 years have seen a significant amount of research activity on modeling oblique impact as it applies to several areas of engineering interest including tube/support interactions, granular flows, and robotic tasks.

Early approaches to impact modeling focused on rigid body approaches that applied the laws of conservation of linear momentum and an assumed coefficient of restitution to provide the solution of postimpact velocities. In the early models of oblique impact, it was assumed that no tangential force acted during impact so that the tangential velocity remained unchanged.

As the understanding of applied mechanics grew, the solution methods for the tangential force during oblique elastic impact began to be considered. Coulomb's friction law dictated that a friction force should act to decrease the tangential velocity for certain angles of incidence during impact. This implies that for high incidence angles (measured from normal), relative sliding of the two colliding bodies occurs throughout the impact duration. Further advances in the field of continuum mechanics indicated that even for near normal angles of incidence, tangential forces would be expected to develop due to the compliance of the colliding bodies. Unfortunately, these compliance relationships can be difficult to derive for all but the simplest contact zone geometries.

Maw et al. [2] provided the first continuum model solution for the problem of oblique elastic impact of spheres. Their analysis showed that oblique impact could be divided into three incidence angle regimes. For low incidence angles, the impact begins with full sticking of coincident points of the two bodies in the contact zone. The tangential force is less than the limiting friction value[2] and reverses direction at some point during the impact duration. Full sliding may begin at some point toward the end of impact. For intermediate angles, the impact begins with full relative sliding of coincident points in the contact zone, and the tangential force is equal to the limiting friction value. At some point during the impact, coincident points on some central portion of the contact zone begin to stick together, while coincident points on the outer annulus of the contact zone may still have some relative slipping. The tangential force leaves the limiting friction envelope and reverses direction. Reverse sliding of the coincident points occurs toward the end of impact, with the tangential force equal to the negative limiting friction value, and persists until contact is lost. For high incidence angles, full sliding persists throughout impact, and the tangential force is equal to the limiting friction value throughout. The division of low, intermediate, or high incidence angle is dependent on the material properties, friction, and geometry of the impacting bodies.

Many of the experimental studies on oblique impact that have been reported in the literature have focused on pre- and postimpact velocity results of the colliding bodies [2–11] rather than on contact force results. This information has been collected using high speed and/or stroboscopic photography. The results for several different material and geometric combinations have been presented. Those listed above are limited to cases of either spheres or disks impacting flat plates, or oblique impact of similar spheres. Also, the results for various initial velocities have been presented. In general, the results are consistent with the numerical model of Maw et al. [2] and show that the rebound angle of the contact point on the unconstrained body is negative (i.e., rebounds back toward the tangential approach direction) for near normal angles of incidence and positive for larger incidence angles.

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 11, 2008; final manuscript received September 18, 2008; published online March 13, 2009. Review conducted by Thomas W. Shield.

[2]Under an assumption of a constant Coulomb friction coefficient, the tangential force is limited to the product of the friction coefficient and the current normal force. For reversed sliding, the tangential force is equal to the negative of this product.

Experimental studies reporting the actual contact forces during oblique impact have been less common than those that report pre- and postimpact velocity values. Lewis and Rogers [12,13] collected oblique impact force data using a pendulum apparatus consisting of a steel ball attached to the bottom of an aluminum tube. The pendulum could be set so that the steel ball would strike a tri-axial piezoelectric force transducer at various incidence angles and initial velocities upon release. Calculation of the instantaneous ratio of tangential to normal force showed that the apparent coefficient of friction would start and end at zero and would rise to an approximate plateau at some point during the impact. For incidence angles less than about 40 deg from normal, the friction coefficient plateau value varied approximately linearly with impact angle; the friction coefficient plateau value was approximately constant for angles larger than 40 deg [12].

Experimental contact force waveform results for oblique elastic impact were presented by Osakue and Rogers [14]. In this study, a steel sphere suspended at the end of an aluminum tube pendulum was allowed, upon release, to obliquely impact a flat steel cap at various velocities and incidence angles. A single tri-axial piezoelectric force transducer, sandwiched between the cap and a steel mounting block, was used to obtain time-waveforms of the normal and tangential contact forces in the horizontal plane during impact; the vertical tangential force waveform was used to align the vertical position.

The results showed that the apparent coefficient of friction was weakly dependent on the velocity and could be described by a bilinear function of the incidence angle. Tangential force reversal occurred during the impact duration for several cases considered in this study. However, agreement between these results and those predicted by the previous models is difficult to judge as normalized tangential force waveforms were not included [14].

Cross [15] collected normal and tangential force waveforms for various balls bouncing against a flat piezoforce plate with various angles of incidence. The piezoforce plate, which provided the normal impact force results, was mounted on a wooden block that sat on cylindrical rollers. An inexpensive piezodisk was mounted on the side of the wooden block and responded to its acceleration. The tangential impact forces were found from the piezodisk's acceleration signal.

The balls used in this study included a tennis ball, a superball, a golf ball, a baseball, and a basketball. The results showed that tangential force reversal occurred for several of these balls at near normal angles of incidence. For higher angles of incidence, the tangential force waveforms were consistent with those expected for full sliding throughout impact. Cross also found that the postimpact spin of the ball was larger than that expected from the tangential force alone and surmised that the normal force may provide additional angular momentum to the ball by acting vertically through a point behind the ball's center.

The results given in the present paper were obtained using a similar pendulum apparatus as that used by Osakue and Rogers [14]. The aluminum tube has been replaced by a steel wire, the striking sphere has been replaced by a symmetrically shaped spherical striker, and a spherically shaped target cap is used instead of a flat cap. All three signals from the tri-axial force transducer were collected during the experiments, and the apparatus was very carefully aligned so that loading was effectively limited to the normal and tangential directions in the horizontal plane. The results present, for the first time, comprehensive sets of force waveforms that clearly show the phenomenon of tangential force reversal during oblique elastic impact. When properly normalized, these sets of force waveforms show wonderful consistency with one another and very reasonable agreement with previous theory.

## 2 Experimental Apparatus

A schematic of the experimental apparatus used in this study can be seen in Fig. 1. The apparatus is a simple pendulum setup with a spherical striker suspended on a steel wire. A magnetic gripper arm, which can be rotated to provide various incidence angles, is used to provide an initial offset of the pendulum. Upon release, the pendulum swings downward through a small distance and strikes a steel target cap having a spherical surface. A tri-axial piezoelectric force transducer is sandwiched between the target cap and a large steel mounting block. The mounting block is held in position using six bolts mounted on small angle-iron pieces located on the sides and back of the mounting block.

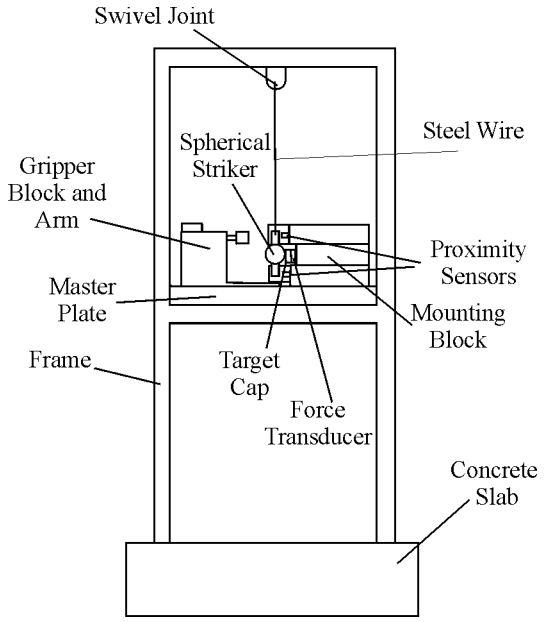



**Fig. 1 Schematic of experimental apparatus**

The actual initial velocity and incidence angle are precisely measured using two pairs of proximity sensors, which are housed on the mounting block. The proximity sensors measure the gap distance between themselves and two steel cubes that extend above and below the striking sphere. Each set of proximity sensors—one set above the force transducer and one below—consists of one sensor oriented along the normal impact direction and one sensor oriented to measure horizontal tangential motion.

**2.1 Apparatus Components.** The mounting block for the force transducer and proximity sensors can be seen in Fig. 2. The mounting block consists of three sections: the upper and lower sections hold the proximity sensors; the force transducer is attached to the middle section. In order to reduce mounting block

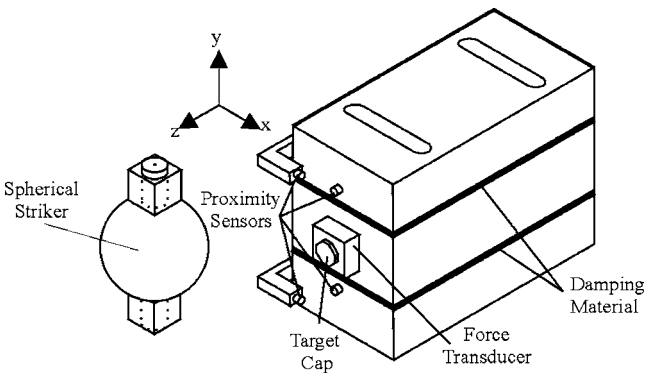



**Fig. 2 Spherical striker and mounting block**

natural frequency contamination of the force transducer's signal, a special damping material[3] was placed between the mounting block sections, which were held together using four bolts.

The radius of curvature of the steel ($E=207$ GPa; $\nu=0.3$; $\rho=7800$ kg/m$^3$) target cap's striking surface was 20 mm. A threaded rod, attached to the back of the cap, passes through the center of the transducer and preloads the force sensor by threading into the middle section of the mounting block.

The spherical striker and cubical proximity sensor targets are also shown in Fig. 2. The striker consists of a 40 mm diameter steel sphere with two 10 mm diameter, 20 mm high steel cylinders extending above and below. The proximity sensor targets are small steel cubes ($16\times16\times16$ mm$^3$) with cylindrical holes through their centers and fastened using set screws. This design allowed for interaction with the proximity sensors while avoiding significant rocking of the sphere on impact.

The composite mass moment of inertia of the striking sphere has a calculated value of $I=5.043\times10^{-5}$ kg m$^2$. A simple experimental exercise using a small accelerometer attached to the bottom cube gave an estimated moment of inertia within 7% of this value. The total mass of the spherical striker is 336.4 g.

**2.2 Sensors and Data Acquisition.** The contact forces were measured using a single Kistler™ (Winterhur, Switzerland) type 9251 piezoelectric force transducer [16]. This sensor is able to produce three charge signals proportional to the forces applied to the face of the transducer in three directions—one normal and two tangential. Three Kistler™ 5010B charge amplifiers were used to convert the charge signals into the appropriate voltage values.

The sensitivity of these sensors is dependent on installation [17], and therefore, calibration of each force sensor is required with each different application. In order to calibrate all three directions of the sensor and determine the appropriate sensitivities to be used for the impact study, a special calibration apparatus was constructed. This apparatus and the results of the calibration are the topics of another paper [18].

The proximity sensors used to collect initial velocity and incidence angle data were 5 mm diameter Bently Nevada™ (Minden, NV) eddy current sensors (model 20886-01 NFMP) with Proximitor™ signal conditioning units (model 20885-0). The return coil in these sensors measures the voltage drop as a metallic object passes through the magnetic field created by the source coil of the sensor. By measuring this voltage drop, the gap distance between the object and the sensor can be recorded. Calibration was performed in order to determine the correct voltage to distance conversion for these sensors using a micrometer equipped, spring loaded sled.

The data from all of the sensors were collected using two National Instruments™ (Austin, TX) PCI-6110 data acquisition boards, which were synchronized using a real time system integration (RTSI) cable. The resolution of these simultaneous sampling boards is 12 bits, and the data were sampled at 5 MHz/channel.

## 3 Experiment Preparation

The alignment procedure was an iterative process with data from several impacts being collected at nominal angles of +10 deg, 0 deg, and −10 deg. During this procedure, only the force transducer signal data were collected, as the tangential proximity sensors would interfere with the pendulum approach for the −10 deg tests.

Figures 3(*a*)–3(*c*) show the final alignment results of the normal (*z*-axis), horizontal (*x*-axis), and vertical (*y*-axis) force signals for five trials at the 0 deg position, respectively. The five normal force signals shown in Fig. 3 are virtually identical. The normal force appears as an approximate sine-squared wave with very little

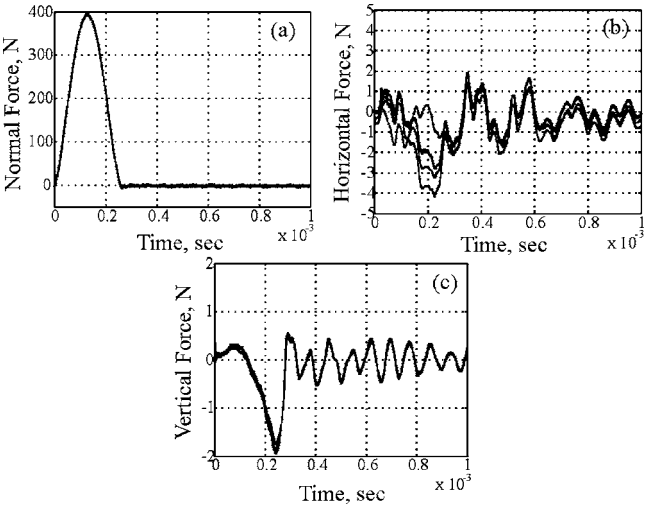



**Fig. 3 Alignment procedure results at 0 deg. (*a*) Normal force, (*b*) horizontal force, and (*c*) vertical force.**

postimpact ringing. The horizontal and vertical force signals show an almost random type waveform of less than 1% of the normal signals. As the crosstalk between the measuring directions for these sensors is listed as 1% [16], this level in the two tangential directions is accepted as essentially zero. The amount of postimpact ringing on the horizontal and vertical signals is on the same level of the actual forces generated during impact for this case.

The three force signals for five trials at the +10 deg and −10 deg positions are shown in Figs. 4(*a*)–4(*c*). As can be seen in Fig. 4(*a*), the normal force results of the five trials at +10 deg match those of the five trials at −10 deg very closely, so much so that it is difficult to distinguish between them.

The horizontal forces depicted in Fig. 4(*b*) show reasonable symmetry between the five trials at the +10 deg position and the five trials at the −10 deg position. Also, there is very little variability between the trials at similar angles. These waveforms show the characteristic sign reversal at about 60% of the impact duration. Somewhat interestingly, the postimpact ringing in the signal exhibits symmetry, but its level is small compared with the assumed actual horizontal force.

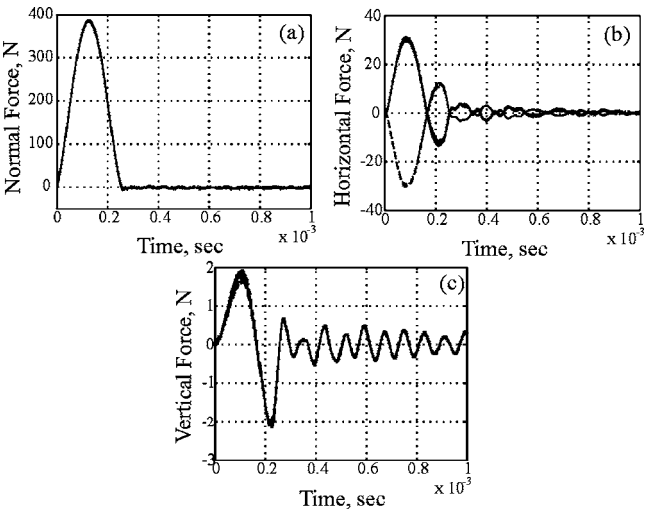



**Fig. 4 Alignment procedure results at +10 deg and −10 deg. (*a*) Normal force, (*b*) horizontal force, and (*c*) vertical force (+10 deg, solid line; −10 deg, dashed line).**

[3]Material No. DS-125 by H.L. Blachford Ltd. (Mississauga, Ontario, Canada).

Figure 4(*c*) shows that the vertical force waveforms for all ten trials are approximate sine waves, and the postimpact ringing portion of the signal is of a lower level. The characteristic shape of the signal during impact could be caused by rocking of the spherical striker due to either the imperfect symmetry of the striker about a horizontal plane or to a slight misalignment of the target cap and spherical striker. Another possible cause of the apparent rocking of the striker could be that the horizontal and vertical directions of the force sensor are not perfectly aligned with true horizontal and vertical. Since the level of force in this signal is within the allowable crosstalk between the sensor directions, this vertical force is taken as negligible.

Once the mounting block was assumed to be positioned correctly with respect to the master plate, experimental data were collected for 3 different nominal initial velocities—30 mm/s, 60 mm/s, and 90 mm/s—and 13 incidence angles—0–50 deg in steps of 5 deg, and 60 deg and 70 deg—for each of the nominal initial velocities. Six sets of data for each combination of nominal incidence angle and initial velocity were collected. The data were collected randomly within each nominal initial velocity, with the 90 mm/s data being collected first, followed by the 60 mm/s and then the 30 mm/s data. At the beginning of every nominal initial velocity and at every ten impacts thereafter, the surfaces of the target cap and spherical striker were cleaned using an ethanol solution to remove any possible surface buildup or debris.

## 4 Experimental Results

**4.1 Initial Conditions.** By calculating the rate of change in gap distance measured by the proximity sensors, one can find the approach velocity in both the normal and horizontal tangential directions and the incidence angle. Since the apparatus was equipped with four proximity sensors—two normal and two tangential—redundant information was collected. The decision to use all four sensors was driven by a desire to ensure proper approach motion of the striker.

Once the velocities at the proximity sensor locations had been found, they had to be transferred to equivalent velocities at the target cap. The velocity at the target cap whether measured from the top or bottom set of proximity sensors showed essentially identical results. The measured velocity values are taken as the mean of the top and bottom velocity measurements. The measured incidence angle is then calculated from

$$\phi = \arctan\left(\frac{V_{x,0}}{V_{z,0}}\right) \tag{1}$$

where the $x$-axis is the horizontal tangential direction and the $z$-axis is the normal direction, as shown in Fig. 2. The actual initial velocity of the striker is found from

$$V_0 = \sqrt{V_{z,0}^2 + V_{x,0}^2} \tag{2}$$

and had values of $V_0 = 27.78 \pm 0.98$ mm/s for nominal 30 mm/s, $V_0 = 58.11 \pm 0.67$ mm/s for nominal 60 mm/s, and $V_0 = 88.21 \pm 0.84$ mm/s for nominal 90 mm/s. These are the mean values $\pm 1$ standard deviation over all six trials at the 13 incidence angles.

**4.2 Impact Force Waveforms.** The presented force results are limited to the normal ($z$-axis) and horizontal ($x$-axis) tangential force waveforms. The vertical ($y$-axis) force waveforms for all tests were very similar to those seen in Figs. 3(*c*) and 4(*c*). In all cases, the vertical force signal during impact was within $\pm 2.5\%$ of the maximum normal force and therefore is considered negligible. Any discussion of tangential force in the remainder of this paper refers to the tangential force in the horizontal plane.

Although data were collected at 13 incidence angles for the three nominal initial velocities, in the interest of brevity Fig. 5 shows the normal (left side) and tangential (right side) force waveforms for five incidence angles—0 deg, 15 deg, 30 deg, 45 deg, and 70 deg—for the nominal initial velocity of 90 mm/s.

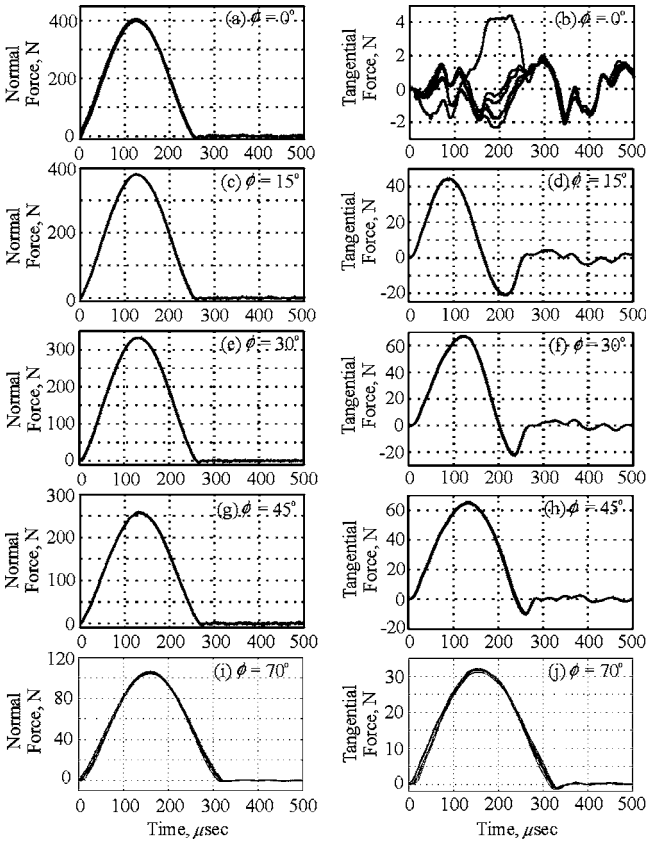



**Fig. 5 Experimental contact force waveforms during impact for initial velocity of 90 mm/s (normal force, left side; tangential force, right side)**

These graphs show all six trials for each incidence angle. Included in these graphs is the useful impact signal plus some postimpact signal. The postimpact ringing on the normal force signals is very small in all cases. The normal force waveforms show an approximate sine-squared type shape that is consistent with a Hertzian impact model. The normal force waveforms show decreasing maximum values and slightly increasing impact duration[4] as the incidence angle increases. In general, the variation in the force waveforms for any test case is quite small.

The tangential force waveforms show different characteristic curves depending on the actual incidence angle of the impact. For near zero incidence angles, the tangential force waveforms show a relatively small signal throughout the impact, with the postimpact ringing being of a similar level as the signal present during impact. For incidence angles of 15 deg, 30 deg, and 45 deg, the tangential force reverses direction within the impact duration.[5] The maximum and minimum tangential forces, as well as the time of tangential force reversal, increase with incidence angle. The postimpact ringing present in these signals is significant, and it would be reasonable to assume that the system natural frequencies slightly affect the force signals during impact. The variation in the tangential force waveforms is also quite small.

[4]The impact duration as judged from the normal force waveforms is taken as the time, after initial contact, at which the normal force crosses zero.

[5]The impact duration as judged from the tangential force waveform depends on whether tangential force reversal is expected and clearly present. For near normal incidence angles where tangential force reversal is clearly present, the impact duration is taken as the time at which the tangential force makes its second zero crossing (i.e., the first zero crossing after the reversed force direction phase). For glancing incidence angles, the impact duration as judged by the tangential force waveform is taken as the first zero crossing of tangential force after the initial contact.

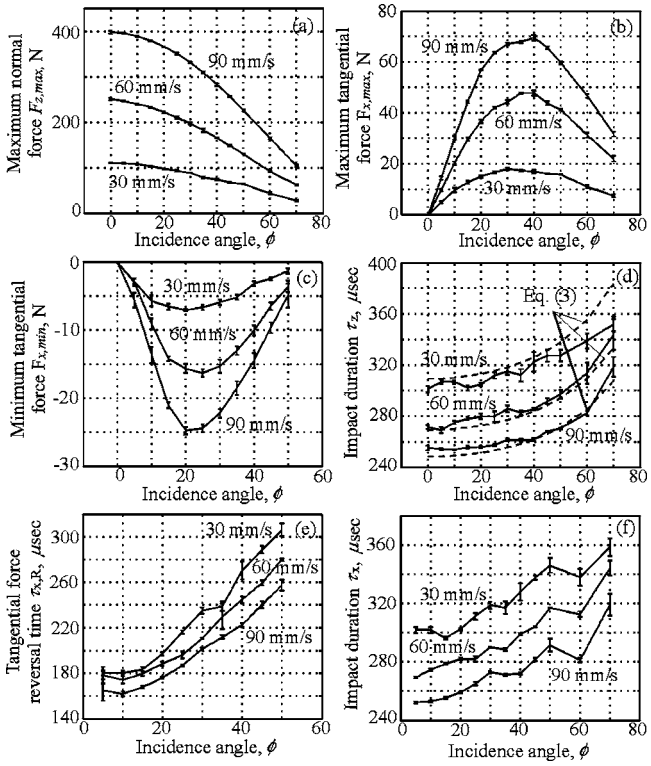



**Fig. 6 Variation of impact force parameters with incidence angle. (*a*) Maximum normal force, (*b*) maximum tangential force, (*c*) minimum tangential force, (*d*) impact duration, $\tau_z$, (*e*) tangential force reversal time, $\tau_{x,R}$, and (*f*) impact duration, $\tau_x$.**

At the nominal incidence angle of 70 deg, the tangential force waveforms do not appear to reverse direction during the impact. However, the postimpact ringing portion of these signals is significant and somewhat impairs the ability to judge whether the force reverses and to judge the exact time of the end of impact.

Very similar plots were obtained for nominal initial velocities of 30 mm/s and 60 mm/s. As expected, the maximum normal force decreased and the impact duration increased with decreasing initial velocity.

In some cases, the resolution of the graphs of experimental force waveforms does not give a good sense of the variation in these waveforms from trial to trial. As such, several parameters were isolated and plotted against the incidence angle in Fig. 6. These are the maximum normal force, $F_{z,\max}$; the impact duration as judged from the normal force waveform, $\tau_z$; the maximum tangential force, $F_{x,\max}$; the impact duration as judged by the tangential force waveform, $\tau_x$; and, if judged applicable, the minimum tangential force, $F_{x,\min}$, and the time of tangential force reversal, $\tau_{x,R}$.

As can be seen in Fig. 6(*a*), the maximum normal forces decrease with increasing incidence angle and decreasing initial velocity. The standard deviation (shown by error bars) of the maximum normal forces between trials at similar incidence angles is reasonably small. Figures 6(*b*) and 6(*c*) show the trends in maximum and minimum tangential forces, respectively. The complex nature of the relationship between friction, normal force, and incidence angle on which the tangential force depends makes it difficult to draw any conclusions as to the correctness of these trends; however, the standard deviation of the values between trials at similar incidence angles shows that the variation in the waveforms is reasonably small.

The impact time parameters shown in Figs. 6(*d*)–6(*f*) appear to be somewhat more variable and do not follow expected trends in some instances. For strictly normal impact, the impact duration is given by [19]

$$\tau = \frac{4\Gamma(2/5)\pi^{1/2}C^2V_{z,0}^{4/5}R}{5\Gamma(9/10)V_{z,0}} \approx \frac{2.9432C^2R}{V_{z,0}^{1/5}} \tag{3}$$

where $\Gamma$ indicates a gamma function, $R$ is the sphere's radius, and the parameter $C$ is given by

$$C = \left\{ \frac{15m(1-\nu)}{16GR^3} \right\}^{1/5} \tag{4}$$

where $m$ is the sphere's mass, $G$ is the shear modulus, and $\nu$ is Poisson's ratio. The results of Eq. (3) are included in Fig. 6(*d*) and show reasonable agreement with the experimental results for $\tau_z$.

Figures 6(*d*) and 6(*f*) show the trends in normal and tangential impact duration estimates, respectively. While the impact duration as judged from either the normal impact duration, $\tau_z$, or the tangential impact duration, $\tau_x$, does show increased time from near normal to the glancing incidence angles as expected, there are several cases where higher incidence angles show lower impact durations than those of lower incidence angles. With the exception of the lowest and highest incidence angles, the tangential impact duration, $\tau_x$, is larger than the normal impact duration, $\tau_z$, sometimes by nearly 20 $\mu$s. This difference in impact duration measurement is somewhat curious, but it seems reasonable that, given the significant postimpact ringing of the tangential force waveforms, this could be caused by system natural frequency contamination of the tangential force signals.

As can be seen in Fig. 6(*e*), the time of tangential force reversal, $\tau_{x,R}$, shows very small variation for incidence angles of 5 deg, 10 deg, and 15 deg for all three nominal initial velocities. Above these values, the time of force reversal increases with incidence angle and decreasing initial velocity.

**4.3 Friction Models.** Previous studies [2,20] indicate that trends in the expected tangential impact forces are difficult to identify outside of the framework of normalization of these forces. This normalization scheme is dependent on a properly identified friction coefficient or friction model. Although a full examination of the possible friction models and the factors that lead to differences in the friction data is beyond the scope of this study, an appropriate friction model is required for normalization of the experimental data.

At glancing incidence angles, from about 55 deg and higher, the tangential force should be equal to the product of the friction coefficient and normal force throughout the impact. This corresponds to full sliding of the coincidence points of the two bodies throughout the impact duration. The results of the previous simulations [2,20] indicate that use of the ratio of tangential to normal forces to determine the friction coefficient could also be valid, at least for the initial portion of the impact duration, for some of the higher incidence angles that exhibit tangential force reversal. Therefore, in an effort to increase the number of data points used to determine the coefficient of friction, the average value of the ratios of tangential to normal force up to one-half of the impact duration from the 45 deg, 50 deg, 60 deg, and 70 deg cases is used in determining the appropriate estimates of the friction coefficient for the three initial velocity data sets.

Using this method, the estimates of the coefficient of friction can be found from

$$\mu = \text{mean}\left(\frac{F_{x,t}}{F_{z,t}}\right), \quad t = 0, \ldots, \tau_z/2 \tag{5}$$

where $F_{x,t}$ and $F_{z,t}$ are the tangential and normal forces at time $t$, respectively.

Several researchers have found a dependence of friction on the relative tangential velocity between the contacting bodies

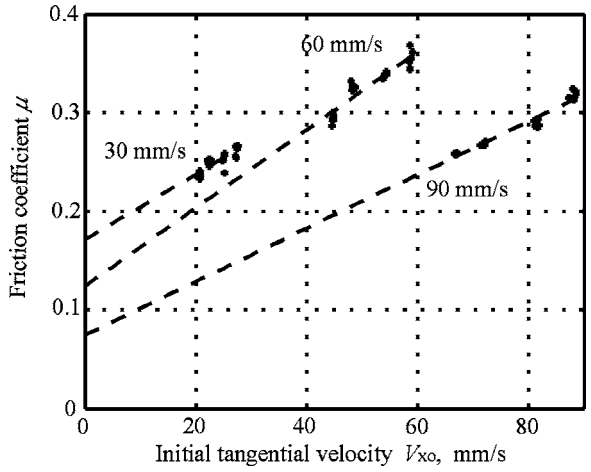



**Fig. 7 Friction coefficient versus initial tangential velocity**

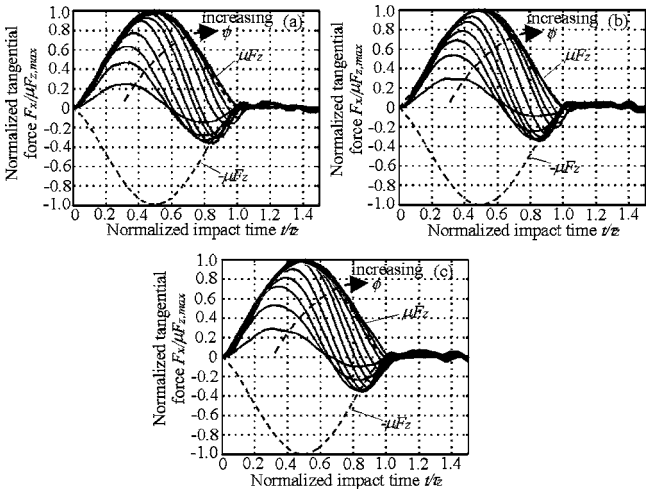



**Fig. 8 Normalized tangential force waveforms. (*a*) 30 mm/s data, (*b*) 60 mm/s data, and (*c*) 90 mm/s data (friction envelope, dashed line).**

[14,21,22], and initial analysis of the friction coefficient estimates indicated that a velocity dependence could be valid for the present data. Figure 7 shows linear trends for the friction coefficient values versus the initial tangential velocities for all three of the initial velocity cases of this study. As can be seen, the 60 mm/s friction data exhibit somewhat higher variability than the other two sets of data. It is clear from the figure that there is not a single friction curve that is valid for all data used. It is conceivable that the 60 mm/s and 30 mm/s friction data could belong to the same curve; however it is quite clear that the 90 mm/s friction data would not fit within the same model. The reason for the different friction models could be due to fretting wear of the surfaces of the striking sphere and target cap during the experiments, which was confirmed to have occurred by inspection under a microscope after the collection of the experimental data. In this case, the order in which the data were collected is significant.

The friction coefficient estimates suggest that a different friction model should be used for each different set of initial velocity experimental data. With this approach, each friction curve varies linearly with initial tangential velocity. In order to determine the appropriate friction coefficient for use with incidence angles less than 45 deg, the friction data for each curve were used to extrapolate to the appropriate tangential velocities. The collected friction data were then used to determine friction models given by

$$\mu_{30} = 0.1711 + 0.0033 \cdot V_{x,0}$$

$$\mu_{60} = 0.1260 + 0.0040 \cdot V_{x,0} \tag{6}$$

$$\mu_{90} = 0.0744 + 0.0028 \cdot V_{x,0}$$

where $V_{x,0}$ is the initial tangential velocity and $\mu_{30}$, $\mu_{60}$, and $\mu_{90}$ are the friction coefficient curves for the initial velocities of 30 mm/s, 60 mm/s, and 90 mm/s, respectively. These relationships are shown as the dashed lines in Fig. 7.

The incidence angle at which the impact commences in full sliding must now be determined under the framework of a linearly varying friction model. The normalized value of this incidence angle is given by [2]

$$\psi = \frac{2(1-\nu)}{\mu(2-\nu)} \frac{V_{x,0}}{V_{z,0}} \tag{7}$$

The general form of a friction model that varies linearly with tangential velocity is given by

$$\mu = \mu_0 + sV_{x,0} \tag{8}$$

where $\mu_0$ is the intercept value and $s$ is the slope of the friction line.

The incidence angle at which the impact commences in full sliding can be determined by setting $\psi=1$ and substituting Eq. (8) into Eq. (7). The values of Eq. (6) give 13.2 deg, 12.3 deg, and 7.4 deg for the critical incidence angle for the data sets of the nominal initial velocities of 30 mm/s, 60 mm/s, and 90 mm/s, respectively. Despite the rather different friction models of Eq. (6), these critical angles are not that dissimilar. For incidence angles greater than these critical values, one should expect the tangential force to lie on the limiting friction envelope at incidence.

**4.4 Normalized Tangential Force Waveforms.** The normalization of the tangential force waveforms contains two major parts: the normalization of the force values and the normalization of the time scale. Normalization of the force values is performed by dividing the tangential force by the product of the friction coefficient and the maximum value of the normal force waveform. Normalization of the time scale is performed by dividing the current time by the total impact duration. Since the tangential force waveforms seem more susceptible to system natural frequency contamination, impact duration estimates from the normal force waveforms are used in the normalization process.

Figures 8(*a*)–8(*c*) show the normalized tangential force waveforms for all the incidence angles for a single randomly chosen trial at the initial velocities of 30 mm/s, 60 mm/s, and 90 mm/s, respectively. The similarity of these graphs is quite remarkable. As the incidence angle is increased, the tangential force spends a higher percentage of the impact duration near the limiting friction envelope, and the time of tangential force reversal increases.

A few characteristics of the normalized tangential force waveforms do not agree with the generally accepted theory for partial slip models of oblique impact. The most striking is the fact that for all three cases, the normalized tangential force waveforms violate the friction envelope in the reversed slip regime toward the end of impact at all of the higher incidence angles. The reason for this violation is unclear but, given the significant amount of postimpact ringing, is most likely related to natural system response contamination, which is also believed to be the source of discrepancy between the apparent impact duration estimates.

At other points during the impact, the tangential force also appears to lie slightly outside the friction envelope. These do not appear to be systematic and could be due to a somewhat poorly defined friction model or, again, system natural frequency response contamination.

As can be seen, the 5 deg cases of the 90 mm/s and 60 mm/s data show a higher degree of waviness than the other waveforms.

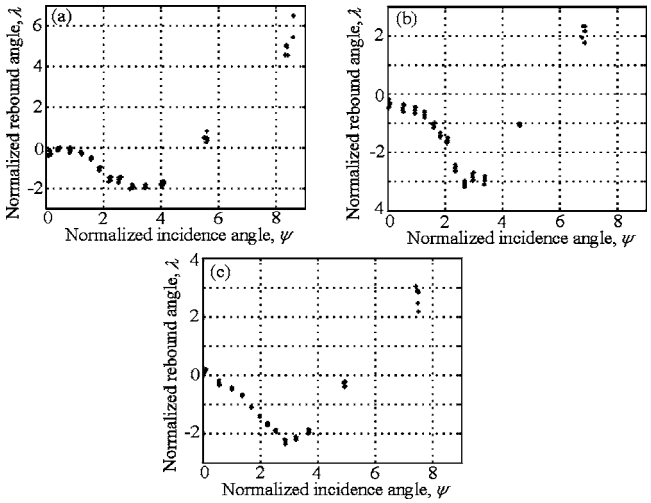



**Fig. 9 Normalized rebound versus normalized incidence angle. (*a*) 30 mm/s data, (*b*) 60 mm/s data, and (*c*) 90 mm/s data.**

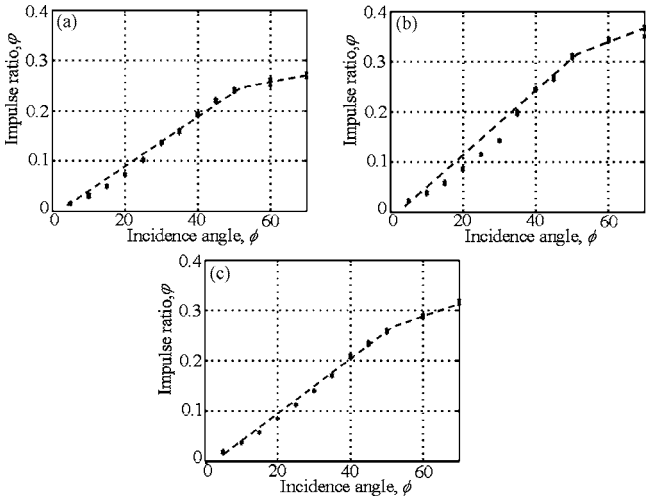



**Fig. 10 Impulse ratio versus incidence angle. (*a*) 30 mm/s data, (*b*) 60 mm/s data, and (*c*) 90 mm/s data.**

Also, the spacing between the waveforms for the 90 mm/s case is more uniform than that for the other two cases.

Those points aside, very reasonable agreement is seen in the force levels of all normalized waveforms of similar incidence angles between the three initial velocities. The limit of tangential force reversal is quite consistent, with the 50 deg case for all three data sets showing reversal prior to loss of contact and the 60 deg case of all three data sets showing full sliding throughout impact.

**4.5 Rebound Angles and Impulse Ratio.** As mentioned in Sec. 1, several researchers [3–5,9–11] have presented the results of surface rebound versus incidence angles of the contact zone obtained during oblique impact studies. Typically, these results are found from pre- and postimpact velocity measurements of the impacting body obtained from high speed photography. Even though postimpact velocity information was not collected during this study, the postimpact velocities can be calculated from the impulse values obtained from the contact force waveforms. The normal, tangential, and angular postimpact velocities can be found from

$$V_{z,N} = V_{z,0} - \frac{P_z}{m}, \quad V_{x,N} = V_{x,0} - \frac{P_x}{m}, \quad \omega_{y,N} = \omega_{y,0} - \frac{RP_x}{I} \tag{9}$$

where $\omega_{y,0}$ is the initial angular velocity about the $y$-axis, and $P_z$ and $P_x$ are the terminal values of the normal and tangential impulses, respectively. The effect of the normal force acting a small distance behind the center of mass of the spherical striker, due to elastic tangential deformation, is disregarded in the calculation of the angular velocity.[6]

The rebound angle of the contact zone is normalized, similar to the incidence angle in Eq. (7), by

$$\lambda = \frac{2(1-\nu)}{\mu(2-\nu)} \frac{(V_{x,N} + R\omega_{y,N})}{V_{z,N}} \tag{10}$$

where $\mu$ is the appropriate value of the friction coefficient found from Eq. (6).

Figures 9(*a*)–9(*c*) show the normalized rebound angle versus normalized incidence angle for the nominal initial velocities of 30 mm/s, 60 mm/s, and 90 mm/s, respectively. These results show that the normalized rebound angles are negative (i.e., the contact zone reflects toward the approach direction) over a wide range of the incidence angles considered. The rebound angles obtained are quite repeatable, with the 60 mm/s data set showing higher variability than the 30 mm/s or 90 mm/s data sets. Also, the 70 deg incidence angle data for all three data sets show much higher variation than the calculated rebound angles at other incidence angles. In general, the variations in rebound angle with incidence angle seen for all three data sets are quite similar and also agree quite well with the results of previous studies [3–5,9–11].

Another commonly presented result for oblique impact studies is the impulse ratio. This value is given by the equation

$$\varphi = \frac{P_x}{P_z} \tag{11}$$

Figures 10(*a*)–10(*c*) present the impulse ratio versus incidence angle for the nominal initial velocities of 30 mm/s, 60 mm/s, and 90 mm/s data sets, respectively. From these graphs, one can see that the impulse ratios of the experimental data show steadily increasing values with incidence angle. The results for all three data sets show an impulse ratio that does not reach a constant value at some critical incidence angle, as would be expected under an assumption of constant friction coefficient. Instead, the impulse ratio is an approximate bilinear function composed of two linear functions of incidence angle with different slopes. These approximately linear functions are superimposed on the impulse ratio plots. These results are strikingly similar to the experimental results presented by Calsamiglia et al. [3]. Those results, however, were derived from postimpact velocity information, not impact force information.

## 5 Conclusions

The impact force waveforms obtained during oblique impact of nonconforming elastic bodies with small circular contact zones have been presented. The results consist of force data from 13 incidence angles and 3 initial velocities. The apparatus used to obtain the impact forces consisted of a simple pendulum with a spherical striker attached to a steel wire. The normal force, as well as horizontal and vertical tangential forces, was collected during impact using a tri-axial piezoelectric force transducer. This transducer was located behind a spherical shaped target cap that the spherical striker impacted upon the release of the pendulum.

The normal force waveforms are consistent with a Hertzian impact model. As expected, the maximum normal force values decrease with increasing incidence angle and decreasing initial velocity. In general, the impact duration increases with incidence angle and decreasing initial velocity.

[6]The elastic tangential deformation is expected to be on the order of $10^{-5}$ times smaller than the radius of the spherical striker. Therefore, any moment caused by the normal force is negligible compared with that caused by the tangential force. This is different from the findings of the Cross [15] study in which sporting balls that experienced higher tangential deformation were used.

The experimental horizontal tangential force waveforms confirm the essential features of tangential force reversal predicted by previous continuum models for oblique impact. The vertical tangential force waveforms showed very little signal level, indicating essentially planar impact. The postimpact portion of the tangential force signals indicates a higher degree of postimpact ringing than the normal force waveform. This ringing is believed to be caused by the natural frequency response of the sensor's mounting block and could cause slight contamination of the horizontal tangential force signal during impact, as well as make it difficult to exactly judge the end of impact from the horizontal force waveform. Nevertheless, the results present, for the first time, comprehensive sets of force waveforms that clearly show the phenomena of tangential force reversal during oblique elastic impact.

Analysis of the ratio of horizontal tangential force to normal force at near glancing incidence angles indicates that the data are consistent with a friction coefficient that varies linearly with initial tangential velocity. Microscopic inspection of the surfaces of the target cap and spherical striker subsequent to experimental data collection indicated that slight fretting wear had occurred during the experiments, which meant that the value of the friction coefficient was dependent on the order in which the data were collected. As such, each set of initial velocity data had its own velocity-dependent friction model. Using these friction models, the horizontal tangential force waveforms were normalized, and the trends in the normalized tangential force waveforms with incidence angle are shown to be very consistent among the three data sets. These normalized force waveforms are also consistent with previous continuum model simulations, with the exception of the apparent friction envelope violation that occurs toward the end of impact. Also, the trends in both the rebound angle and the impulse ratio with incidence angle are consistent with previous results.

## Acknowledgment

The authors thank the Natural Sciences and Engineering Research Council of Canada for funding this research and the reviewers of the manuscript for the feedback provided.

## References

[1] Stronge, W. J., 2000, *Impact Mechanics*, Cambridge University Press, Cambridge, UK.
[2] Maw, N., Barber, J. R., and Fawcett, J. N., 1976, "The Oblique Impact of Elastic Spheres," Wear, **38**(1), pp. 101–114.
[3] Calsamiglia, J., Kennedy, S. W., Chatterjee, A., Ruina, A., and Jenkins, J. T., 1999, "Anomalous Frictional Behavior in Collisions of Thin Disks," ASME J. Appl. Mech., **66**(1), pp. 146–152.
[4] Foerster, 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.
[5] Kharaz, A. H., Gorham, D. A., and Salman, A. D., 2001, "An Experimental Study of Elastic Rebound of Spheres," Powder Technol., **120**(3), pp. 281–291.
[6] Labous, L., Rosato, A. D., and Dave, R. N., 1997, "Measurements of Collisional Properties of Spheres Using High-Speed Video Analysis," Phys. Rev. E, **56**(5), pp. 5717–5725.
[7] Lorenz, A., Tuozzolo, C., and Louge, M. Y., 1997, "Measurements of Impact Properties of Small, Nearly Spherical Particles," Exp. Mech., **37**(3), pp. 292–298.
[8] Louge, M. Y., and Adams, M. E., 2002, "Anomalous Behavior of Normal Kinematic Restitution in Oblique Impacts of a Hard Sphere on an Elastoplastic Plate," Phys. Rev. E, **65**, p. 021303.
[9] Maw, N., Barber, J. R., and Fawcett, J. N., 1981, "The Role of Elastic Tangential Compliance in Oblique Impact," ASME J. Lubr. Technol., **103**(1), pp. 74–80.
[10] Dong, H., and Moys, M. H., 2006, "Experimental Study of Oblique Impacts With Initial Spin," Powder Technol., **161**(1), pp. 22–31.
[11] Gorham, D. A., and Kharaz, A. H., 2000, "The Measurement of Particle Rebound Characteristics," Powder Technol., **112**(3), pp. 193–202.
[12] Lewis, A. D., and Rogers, R. J., 1988, "Experimental and Numerical Study of Forces During Oblique Impact," J. Sound Vib., **125**(3), pp. 403–412.
[13] Rogers, R. J., and Lewis, A. D., 1990, "Further Numerical Studies of Oblique Elastic Impact," J. Sound Vib., **141**(3), pp. 507–510.
[14] Osakue, E. E., and Rogers, R. J., 2001, "An Experimental Study of Friction During Planar Elastic Impact," ASME J. Pressure Vessel Technol., **123**(4), pp. 493–500.
[15] Cross, R., 2002, "Grip-Slip Behavior of a Bouncing Ball," Am. J. Phys., **70**(11), pp. 1093–1102.
[16] Kistler Instrumente AG, 1997, "Kistler 9251 Data Sheet," Kistler Instrumente AG, Report No. 000-145m-01.97, http://www.kistler.com/mediaaccess/9251A_000-145m-01.97.pdf.
[17] McConnell, K. G., 1995, *Vibration Testing: Theory and Practice*, Wiley, New York.
[18] Garland, P. P., and Rogers, R. J., 2008, "Dynamic Calibration of Tri-Axial Piezoelectric Force Transducers," Meas. Sci. Technol., **19**(9), p. 095202.
[19] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York.
[20] Garland, P. P., and Rogers, R. J., 2008, "An Analytical Solution for Shear Stress Distributions During Oblique Elastic Impact of Similar Spheres," ASME J. Comput. Nonlinear Dyn., **3**(1), p. 011002.
[21] Bowden, F. P., and Tabor, D., 1950, *Friction and Lubrication of Solids*, Oxford University Press, Oxford.
[22] Fandrich, M., 1998, "Modelling Nondestructive Impacts Macroscopically," Math. Comput. Modell., **28**(4–8), pp. 205–224.

# Shakedown Fatigue Limits for Materials With Minute Porosity

**Jehuda Tirosh**

**Sharon Peles**

Faculty of Mechanical Engineering,
Technion,
Haifa 32000, Israel

*The intention of this study is to predict the fatigue-safe long life behavior of elastoplastic porous materials subjected to zero-tension fluctuating load. It is assumed that the materials contain a dilute amount of voids (less than 5%) and obey Gurson's model of plastic yielding. The question to be answered is what would be the highest allowable stress amplitude that a porous material can endure (the "endurance limit") when undergoing an infinite number of loading/unloading cycles. To reach the answer we employ the two shakedown theorems: (a) Melan's static shakedown theorem ("elastic shakedown") for establishing the lower bound to fatigue limit and (b) Koiter's kinematic shakedown theorem ("plastic shakedown") for establishing its upper bound. The two bounds are formulated rigorously but solved with some numerical assistance, mainly due to the nonlinear pressure dependency of the material behavior and the complex description of the plastic flow near stress-free voids. Both bounds ("dual bounds") are adjusted to capture Gurson-like porous materials with noninteractive voids. General residual stresses (either real or virtual) are presented in the analysis. They are assumed to be time-independent as generated, say, by permanent temperature gradient between void surfaces and remote material boundaries. Such a situation is common, for instance, in ordinary porous sleeves (used in space industry and alike). A few experiments agree satisfactorily with the shakedown bounding concept.* [DOI: 10.1115/1.3005961]

*Keywords: endurance limit, cyclic loading, porous material, residual stress, shakedown theorems*

## 1 Introduction

The common scattered experimental data in measuring fatigue limit of most structural materials (such as steel, copper, aluminum alloys, etc.) are relatively large. It is reported to be ordinarily about 20% around an average value in (apparent) void-free materials. When the materials contain dilute amount of voids (about 5% or less, as in sintered materials), the scatter is even wider. Engineeringwise there is a clear need to assess *in advance* the long term fatigue behavior of such metals (in particular, the allowable stress amplitude), *at their various porosity levels*, in order to avoid premature failure. The shakedown bounds, illustrated in Fig. 1, are destined to handle this assessment by lower and upper bounds (applying Melan's [1] and Koiter's [2] theorems, respectively). The procedure of doing so is the subject of this paper.

Among numerous kinds of fatigue failures, the consideration given here faces only failures associated with overaccumulation of plastic strain ("ratcheting") in elastic–perfectly-plastic porous materials. Other failure modes are excluded.

Applications of the bounding theorems [1,2] are already described in several areas: in mechanical sciences by Polizzotto [3,4], Druyanov and Romman [5,6], Ponter and Engelhardt [7], and Tirosh and Peles [8,9]. In tribology sciences they are used by Kapoor and Williams [10] and Wong et al. [11] as a partial list. In fibrous composite they are used mainly by Dvorak and Tarn [12], Jansson and Leckie [13], and others (i.e., Ref. [14]). In pressure vessels and thermomechanical structures they are used by Ponter and Karadeniz [15,16], Xue et al. [17], and more. In damage mechanics they are used by Huang and Stein [18,19], Belouchrani and Weichert [20], and recently also in manufacturing sciences [21]. In practice, the above dual bounds to porous materials have not been studied hitherto, besides partially in a recent trial solution under the framework of two-phase materials [9].

One reason for this missing analysis could be the apparent lack of an admissible solution for residual stress field, which (according to Melan's proof [1]) defers plastic strain accumulation. The second reason is possibly the mathematical complexity arises from the nonunique kinematics of the plastic flow in nonhardening materials, particularly near existing voids. Under some idealizations (to be discussed) the application of fatigue shakedown theorems is described in due detail for elastic-perfectly-plastic materials, which contain dilute amount of porosity.

## 2 The Static Shakedown

The static shakedown theorem, proved by Melan [1], states that a given structure will shake down under prescribed fluctuating load if one can find an admissible, time-independent, residual stress field, $\sigma_{ij}^{(\mathrm{res})}(x_i)$, such that when it is superimposed on the unbounded elastic stress field, $\sigma_{ij}^{(e)}(x_i,t)$, the expression of their sum (1) never exceeds yielding Eq. (2).

$$\sigma_{ij}^{(\mathrm{sum})} \equiv [\sigma_{ij}^{(e)}(x_i,t) + \sigma_{ij}^{(\mathrm{res})}(x_i)] \tag{1}$$

It means that along a loading path with a given time-period $\hat{T}$ $(0 \le t \le \hat{T})$, the shakedown phenomena assure that the material will eventually respond elastically, namely,

$$\Phi\{\sigma_{ij}^{(\mathrm{sum})}(x_i,t)\} \le \sigma_0 \tag{2}$$

where the function $\Phi\{\ldots\}$ is a known yield function of the material and $\sigma_0$ is its yielding stress in unidirectional loading.

When considering $\Phi\{\ldots\}$ to present von Mises yield criterion, the so-called "shakedown condition" is generated by using Eq. (2) with its *equality sign*. It reads

$$\Phi = \tfrac{1}{2}[(\sigma_{rr}^{(\mathrm{sum})} - \sigma_{\theta\theta}^{(\mathrm{sum})})^2 + (\sigma_{\theta\theta}^{(\mathrm{sum})} - \sigma_{\varphi\varphi}^{(\mathrm{sum})})^2 + (\sigma_{\varphi\varphi}^{(\mathrm{sum})} - \sigma_{rr}^{(\mathrm{sum})})^2 + 3(\tau_{r\theta}^{(\mathrm{sum})})^2 = \sigma_0^2 \tag{3}$$

For porous materials, however, the yield function $\Phi\{\ldots\}$ is preferably considered by Gurson's yield criterion [22] akin to materials with dilute amount of pores.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 11, 2007; final manuscript received July 10, 2008; published online March 13, 2009. Review conducted by Krishna Garikipati.

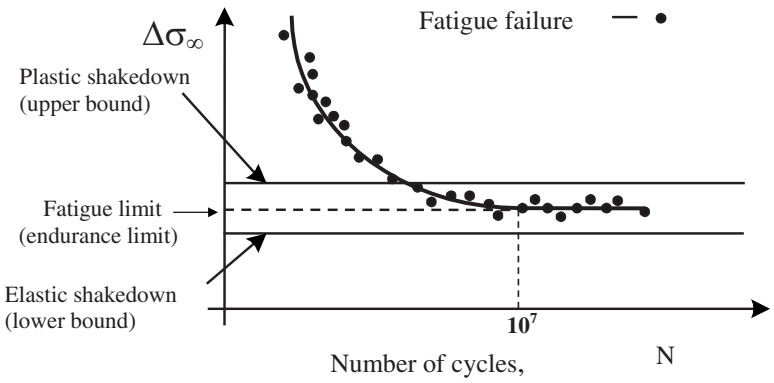



**Fig. 1 Schematic illustration of shakedown bounds to fatigue limit (endurance limit).**

Define $f$ as the volume fraction of the voids in the material. It is postulated that such a material can reach local yielding (shown later to originate along the stress-free boundary of the pores) by certain macrostresses $\sigma_{ij}^{(\mathrm{sum})}$ applied on the solid, or identically, on a representative unit-cell of the solid, having a single centered pore that is "smeared" across the cell to denote the macroproperty $f$ of the solid (and consequently implies that the associated stresses in the solid are macrostresses). Under such yielding, Gurson's *approximate yield function (*in an upper-bound sense [22]) reads

$$\frac{(3/2)s_{ij}^{(\mathrm{sum})}s_{ij}^{(\mathrm{sum})}}{\sigma_0}+2q_1 f\cosh\left(q_2\frac{\sigma_{kk}^{(\mathrm{sum})}}{\sigma_0}\right)=[1+q_3 f^2] \tag{4}$$

where $s_{ij}$ is defined by $s_{ij}=\sigma_{ij}-\frac{1}{3}\delta_{ij}\sigma_{kk}$. The summation convention for repeated subscripts has been employed.

The arbitrary scalars ($q_1$, $q_2$, and $q_3$) in Eq. (4) are chosen sometimes as 1.5, 1.1, and 2.25, in order to match more closely the experimental data suggested by Tvergaard [23,24]. In the limit, when the material porosity approaches zero, Eq. (4) coincides with von Mises criterion (3).

In the coming analysis, the material porosity is considered *permanently low*, so that the deviation between Eqs. (3) and (4) for, say, $f=0.03$, is relatively small (nearly 2%). Even though, in order to maintain generality, Eq. (4) rather than Eq. (3) is used in the oncoming analysis.

**2.1 Applications of Melan's Static Lower Bound.** The "beauty" of Melan's shakedown theorem [1] is that the actual elastoplastic solution, $\sigma_{ij}(x_i,t)$, does not have to be determined a priori. Instead, it is replaced with an *unbounded elastic solution*, $\sigma_{ij}^{(e)}(x_i,t)$, having the same geometry and subjected to the same boundary conditions as the actual problem. Such solutions are already given in classical papers (i.e., Southwell and Gough in 1926 [25]). They will be used in the sequel.

## 3 Elastic Solution

**3.1 The Stress Field Around a Spherical Void in a Material Subjected to Unidirectional Loading.** Consider a spherical void [25] with radius $r_0$ (shear modulus $\mu_0=0$ and Poisson ratio $\nu_0=0$) embedded in an infinite elastic solid (with properties $\mu$ and $\nu$, respectively). A remote alternating stress with amplitude $\sigma_\infty$ is prescribed on the porous solid as shown in Fig. 2($a$). The associated stress distribution (plotted in Fig. 2($b$)) is

$$\sigma_{rr}^{(e)}=\sigma_\infty\left[\sin^2\theta+\frac{1}{14-10\nu}\frac{1}{\rho^3}\left\{-38+10\nu+24\frac{1}{\rho^2}+\left(50-10\nu-36\frac{1}{\rho^2}\right)\cos^2\theta\right\}\right] \tag{5a}$$

$$\sigma_{\theta\theta}^{(e)}=\sigma_\infty\left[\cos^2\theta+\frac{1}{14-10\nu}\frac{1}{\rho^3}\left\{9-15\nu-12\frac{1}{\rho^2}-\left(5-10\nu-21\frac{1}{\rho^2}\right)\cos^2\theta\right\}\right] \tag{5b}$$

$$\sigma_{\varphi\varphi}^{(e)}=\frac{\sigma_\infty}{14-10\nu}\cdot\frac{1}{\rho^3}\left[9-15\nu-12\frac{1}{\rho^2}-15\left(1-2\nu-\frac{1}{\rho^2}\right)\cos^2\theta\right] \tag{5c}$$

$$\tau_{r\theta}^{(e)}=\frac{\sigma_\infty}{2}\sin 2\theta\left[1+\frac{1}{14-10\nu}\cdot\frac{1}{\rho^3}\left\{10(1+\nu)-24\frac{1}{\rho^2}\right\}\right] \tag{5d}$$

where $\rho\equiv(r/r_0)\geq 1$.

**3.2 Residual Stress Distribution.** The elastic shakedown theorem relies on a pre-existence of some residual stress field in the solid. There is no specific requirement that the field should be real, albeit it should be physically admissible (namely, to satisfy the elastic govern equations with homogenous boundary conditions). Consequently, one is allowed to choose entirely fictitious fields. Consider, for example, a permanent temperature drop, $\Delta T$, between the void-free surface and an ambient temperature at far distance, as occurred, for example, in sleeve bearings made of porous metals (usually copper) used frequently in space machineries. The temperature gradient is assumed to be sustained in a steady state condition and thus provides a source for an admissible thermal residual stress, the magnitude of which is $p=\alpha E\Delta T/(1-\nu)$ (where $E$ is the Young modulus and $\alpha$ is its coefficient of linear thermal expansion of the matrix material).

The residual stress for such a radially symmetric case near a void of radius $r_0$ is

$$\sigma_{rr}^{(\mathrm{res})}(r)=-p\left(\frac{1}{\rho}-\frac{1}{\rho^3}\right)$$

$$\sigma_{\theta\theta}^{(\mathrm{res})}(r)=-\frac{1}{2}p\left(\frac{1}{\rho}+\frac{1}{\rho^3}\right)\quad\text{for}\quad\rho\equiv(r/r_0)\geq 1$$

$$\sigma_{\phi\phi}^{(\mathrm{res})}(r)=-\frac{1}{2}p\left(\frac{1}{\rho}+\frac{1}{\rho^3}\right) \tag{6}$$

The above distribution (solved in Appendix A) is plotted in Fig. 3.

It is noted that if the magnitude of the residual stress, $p$, is sufficiently high, plastic yielding may occur even before prescribing the actual fatigue loading. By solving Eq. (6) (drawn in Fig. 3) it becomes clear that $p$ should be restricted to a certain maximum magnitude, say, $p_{\max}$, since Melan's shakedown theorem [1] is

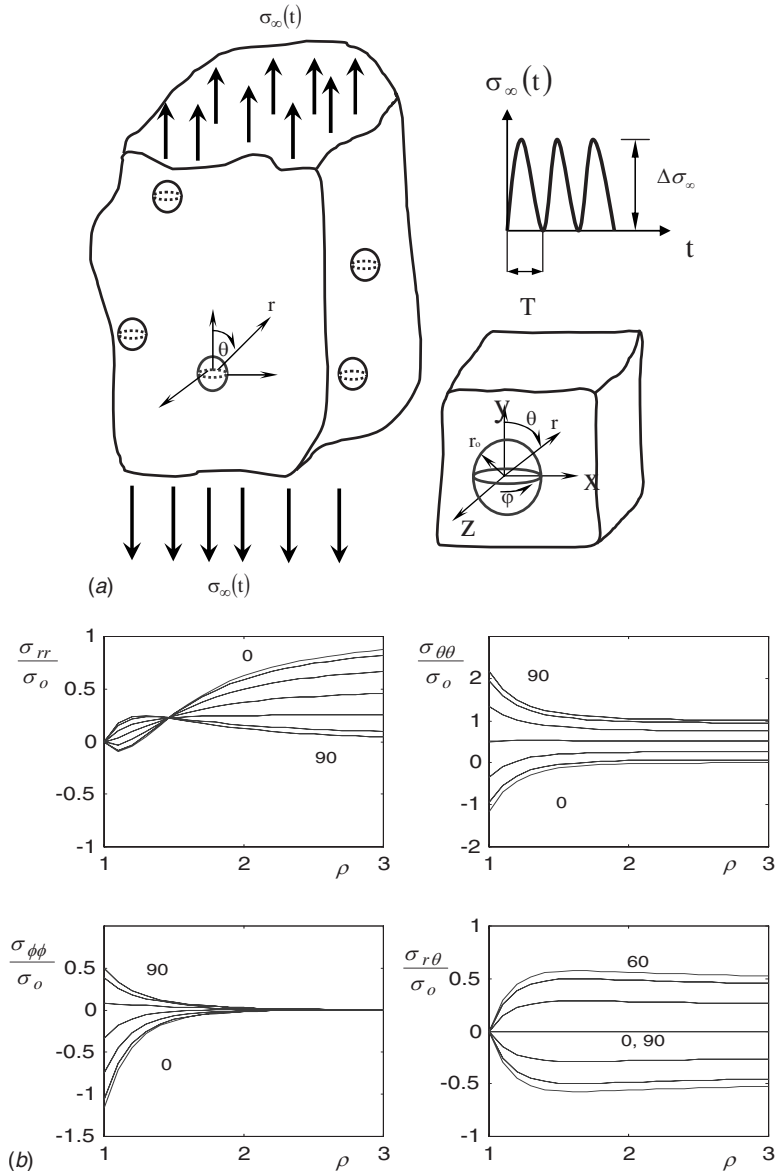



**Fig. 2 (*a*) A view of typical material with voids subjected to unidirectional cyclic loading. (*b*) Elastic stress distribution around a spherical void caused by remote unidirectional traction $\sigma_\infty$. The various curves are the radial distributions around the void at various angles (the zero angle is collinear with the load direction).**

applicable until the appearance of "the first yield point" in the otherwise elastic state of the solid. In order to limit the magnitude of the residual stress, $p_{\max}$, the search for the first yield point is done by using distribution (6) in yield function (4) while setting $\sigma_{ij}^{(e)}=0$. The end-result is

$$\frac{3p_{\max}^2}{\sigma_0^2}+2q_1 f\cosh\left(\frac{q_2 p_{\max}}{\sigma_0}\right)=1+q_3 f^2 \tag{7}$$

(plotted in Fig. 4) showing a decay of $p_{\max}$ with the increase in the material porosity.

Further assumptions are given as follows:

(a) The material properties are unaffected by changes (if any) in the material temperature.
(b) The porous material is idealized as a net of uniform cells. Each one (refers to as "a unit-cell," described in Appendix B) contains a single void whose volume fraction in the cell simulates the average volume fraction of the overall material. The conjugate variables (stress, strain, etc.) then implied their macro-expression.
(c) Each cell is presumed to behave under loading as the overall porous material would.
(d) The void volume of the cell remains constant (to stay consistent with Gurson's yielding model [22]).

## 4 The Static Shakedown Condition

By inserting Eqs. (5*a*)–(5*d*) and Eq. (6) into Eq. (4) one gets the yield function of the loaded material as

$$\frac{\sigma_\infty^2}{\sigma_0^2}\left\{Y\left(\theta,\rho,\frac{p}{\sigma_\infty},q_1,q_2,q_3,f\right)\right\}=1$$

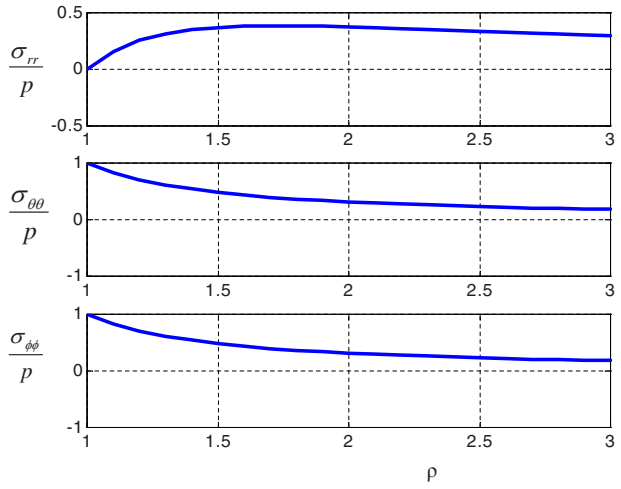



**Fig. 3 An admissible residual stress distribution along the radial distance $\rho = r/r_0$ (solved by Eq. (6)) normalized with its magnitude $p$ ($p = \alpha E \Delta T/(1-\nu)$)**

$$\text{for} \quad \rho \geq 10 \quad \text{and} \quad \leq \theta \leq 2\pi \tag{8a}$$

where the short-hand expression $Y$ in Eq. (8*a*) is

$$Y = \left[ \frac{(3/2)s_{ij}^{(\text{sum})} s_{ij}^{(\text{sum})}}{(1+q_3 f^2) - 2q_1 f \cosh(q_2 \sigma_{kk}^{(\text{sum})}/2\sigma_0)} \frac{1}{\sigma_\infty^2} \right] \tag{8b}$$

In order to find the location of the first yield point at which the material will yield under the prescribe fluctuating load, one need to maximize Eq. (8*b*) with respect to the void coordinates ($\rho$, $\theta$) in the unit-cell. This operation is described in a symbolic-convenient form in Eq. (9). The associated graphical results are the ellipselike curves shown in Fig. 5

$$\min \frac{\sigma_\infty}{\sigma_0} = \frac{1}{\max \left\{ Y\left( \theta, \rho, \frac{p}{\sigma_\infty}, q_1, q_2, q_3, f \right) \right\}^{1/2}} \tag{9}$$

The contours of the ellipses present the static shakedown condition (namely, Melan's [1] lower bound for *ratcheting-free* loading). The area enclosed by the ellipses represents the domain of fatigue-safe stress amplitudes at the corresponding residual stress at various porosity contents.

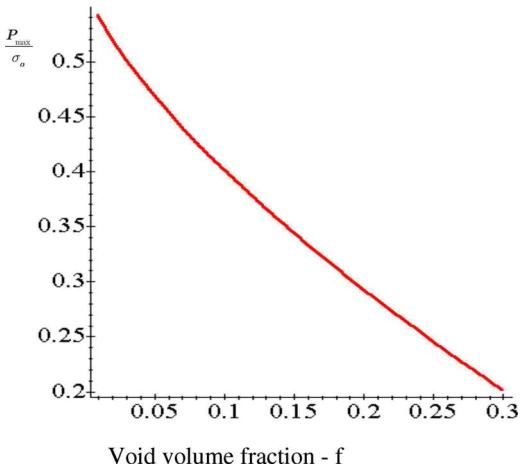



**Fig. 4 The restricted permissible magnitude of the residual stress, $p_{max}$, versus the porosity level of the solid as solved by Eq. (7)**

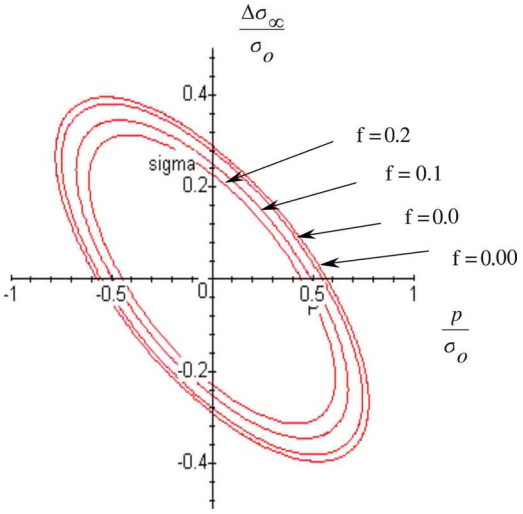



**Fig. 5 The loci of Melan's [1] elastic shakedown conditions (the limit load for ratcheting-free material response at a given void fraction *f*). The area inside each ellipse is the domain of fatigue-safe stress amplitudes.**

## 5 The Kinematic Shakedown

**5.1 The Theorem.** Koiter [2] proved in 1960 that a structure will shakedown under cyclic loading if the rate of plastic dissipation (and the associated frictional loss, if exists) exceeds the sum of the work rates of the applied traction and the body force $X_i$. This inequality reads (within a time period $\hat{T}$)

$$\int_0^{\hat{T}} dt \left\{ \int_V \sigma_{ij} \dot{\varepsilon}_{ij} dV \right\} \geq \int_0^{\hat{T}} dt \left\{ \int_V X_i u_i dV + \oint_{S_T} T_i u_i ds \right\} \tag{10}$$

The strain rate components $\dot{\varepsilon}_{ij}$ in Eq. (10) are distributed (on a microlevel) around a spherical pore located in the center of each unit-cell, and are derived from the trial plastic flow-field $u_i$ (to be devised later). The conjugate stress components $\sigma_{ij}$ are therefore microstresses inside the cell.

The energy-rate volume integral over the unit-cell (the left side of Eq. (10)) renders its overall dissipation rate, and thus turns the centered pore to an *average cell-porosity* $f$ (or equivalently, a "homogenized" property $f$ of the overall solid as well). By doing so, the emerged stresses are implied to be the *macrostress of the continuous porous solid*, akin to Koiter's theorem for continua [2].

In Eq. (10), $T_i(t)$ is the fluctuating traction on portion $S_T$ of the remote surface, $S$. The term on the left-hand side represents the plastic dissipation rate in the deformable volume $V$, considered as the sum of, say, $N$ similar cells (as in Appendix B). In the particular case of a nonhardening porous material, the constitutive relation based on the normality and convexity of Gurson's macrostress yield function [22] is

$$\dot{\varepsilon}_{ij} = \lambda \left[ s_{ij} + \delta_{ij} \frac{k_0 q_1 q_2 f}{\sqrt{3}} \sinh(q_2 \sigma_m) \right] \tag{11}$$

where the plastic coefficient $\lambda$ in Eq. (11) was shown to be (ibid.)

$$\lambda = \frac{\sqrt{I_2}}{k_0 \sqrt{1 + q_3 f^2 - 2q_1 f \cosh(q_2 \sigma_m) + (f^2 q_1^2 q_2^2/2) \sinh^2(q_2 \sigma_m)}} \tag{12}$$

$$s_{ij} = \sigma_{ij} - \frac{1}{3}\delta_{ij}\sigma_{kk}, \quad k_0 = \frac{\sigma_0}{\sqrt{3}}, \quad I_2 = \sqrt{\frac{1}{2}\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}} \tag{13}$$

**5.2 The Dissipative Energy Rates.** As mentioned, the plastic deformation work rate in porous materials is expressed as

$$\dot{W} = \int_V \sigma_{ij}\dot{\varepsilon}_{ij}dV = \int_V (s_{ij}\dot{\varepsilon}_{ij} + (1/3)\sigma_{kk}\dot{\varepsilon}_{ii})dV \tag{14}$$

The first term in the integral of Eq. (14) corresponds to the classical plastic dissipation rate while the second term denotes the dilatational work rate. The first term of Eq. (14) reads

$$\int_V s_{ij}\dot{\varepsilon}_{ij}dV = 2k_0\int_V \frac{(1 + q_3f^2 - 2q_1f\cosh(q_2\sigma_m))\sqrt{I_2}}{\sqrt{1 + q_3f^2 - 2q_1f\cosh(q_2\sigma_m) + (fq_1q_2/\sqrt{2})^2\ \sinh^2(q_2\sigma_m)}}dV \tag{15a}$$

and the second term of Eq. (14) leads to

$$\int_V (1/3)\sigma_{kk}\dot{\varepsilon}_{ii}dV = 2k_0\int_V \frac{\sigma_m q_1 q_2 f\sinh(q_2\sigma_m)\sqrt{I_2}}{\sqrt{1 + q_3f^2 - 2q_1f\cosh(q_2\sigma_m) + (fq_1q_2/\sqrt{2})^2\ \sinh^2(q_2\sigma_m)}}dV \tag{15b}$$

Equation (10) can now be rewritten in the following form:

$$\int_0^T dt\left\{\int_{V_m} X_i u_i dV + \int_{S_u+S_T} T_i(t)u_i ds\right\} \le \int_0^T dt\left\{\int_V (s_{ij}\dot{\varepsilon}_{ij} + 1/3\sigma_{kk}\dot{\varepsilon}_{ii})dV\right\} \tag{16}$$

where the "body force" $X_i$ is defined here as the force "density" (force/volume) generated by the residual stress field that surrounds each cavity. It renders the following:

$$X_i = \frac{\sigma_{ij}^{(\mathrm{res})}n_j \cdot 4\pi r^2}{\frac{4}{3}\pi r^3} = \frac{3\sigma_{ij}^{(\mathrm{res})}n_j}{r} \tag{17}$$

where $n_i$ is an outward unit vector on an arbitrary sphere of radius $r$, for $r \ge r_0$, and the force density is the overall tractions on the sphere surface per its volume.

Equation (17) is a reminder of the ordinary notion of body force in the sense that it manifests an internal source of energy-rate when forced to be displaced by $u_i$. After inserting solution (6) in Eq. (17) one gets its components as

$$X_r = -\frac{3p}{r_0}\left(\frac{1}{\rho^2} - \frac{1}{\rho^4}\right), \quad X_\theta = -\frac{3}{2}\frac{p}{r_0}\left(\frac{1}{\rho^2} + \frac{1}{\rho^4}\right)$$

$$X_\phi = -\frac{3}{2}\frac{p}{r_0}\left(\frac{1}{\rho^2} + \frac{1}{\rho^4}\right) \tag{18}$$

Since this artificial body force is highly concentrated near each cavity, the work-rate contribution for dilute porosity stays relatively small (about 5% less than otherwise). A complete harmonic cycle of the traction leads to *zero work-rate contribution* (due to internal work-rate cancellation) but a complete harmonic cycle is not always guaranteed during fluctuating fatigue loads. Consequently, we are attempting here to preserve the duality of the bounds by including the effect of the residual stresses in both bounds.

## 6 The Plastic Flow Field

The deformation field in nonhardening plasticity of metals is not unique, though the stress field is unique. Hence there are an infinite number of velocity fields that obey incompressibility, compatibility, and pertinent boundary conditions. A finite number of these fields can be combined to form a general approximate field, leaving several undetermined parameters to be determined by satisfying certain boundary conditions and (in certain cases) a minimal dissipative energy-rate. A procedure to do so was suggested by Rice and Tracey [26]. We will use herewith their results.

**6.1 A Smooth Plastic Flow (as Opposed to Flows With Rigid Zones).** The stream function $\psi(r,\theta,t)$, which describes a smooth admissible flow-field [22,26] around a spherical void due to a remote strain, $\dot{\varepsilon}$ (in vertical direction as seen in Fig. 6), was found to be

$$\psi = \frac{1}{8}\left(\rho^3 - 8D + 4E - 3\frac{E}{\rho^2}\right)r_0^3\dot{\varepsilon}\cos(\theta) + \frac{1}{8}\left(3\frac{E}{\rho^2} - 4E - \rho^3\right)r_0^3\dot{\varepsilon}\cos(3\theta) \tag{19}$$

where the unknown coefficients $D$ and $E$ are

$$D = 0.283e^{3\sigma_m^\infty/2\sigma_0}$$

$$E \cong -0.00087\cdot D^3 + 0.0066\cdot D^2 + 0.052\cdot D + 0.496 \tag{20}$$

The stream lines based on function (19) is plotted in Fig. 6.

Based on the definition relating velocity flow-field with a given stream function $\psi(r,\theta,t)$,

$$u_\theta = -\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial r}, \quad u_r = \frac{1}{r^2\sin\theta}\frac{\partial\psi}{\partial\theta} \tag{21}$$

one gets from Eq. (19) the following flow-field:

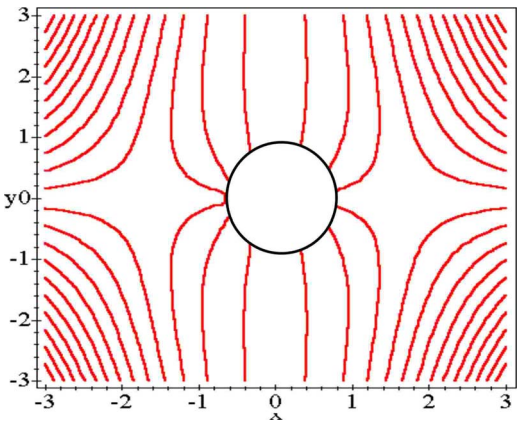



**Fig. 6 Stream lines of smooth plastic flow around a spherical void in a structure subjected to unidirectional remote load (in vertical direction)**

$$u_r = \left(\frac{3}{4}\rho + 3\frac{E}{\rho^2} - \frac{9}{4}\frac{E}{\rho^4}\right) r_0 \dot{\varepsilon} \cos(2\theta) + \left(\frac{E}{\rho^2} + \frac{\rho}{4} + \frac{D}{\rho^2} - \frac{3}{4}\frac{E}{\rho^4}\right) \dot{\varepsilon} r_0 \tag{22}$$

$$u_\theta = \frac{1}{4}\left(-3\rho - 6\frac{E}{\rho^4}\right) r_0 \dot{\varepsilon} \sin(2\theta) \tag{23}$$

The associated strain rate components, using

$$\dot{\varepsilon}_{ij} = 1/2(u_{i,j} + u_{j,i}) \tag{24}$$

give

$$\dot{\varepsilon}_{rr} = \frac{\partial u_r}{\partial r} = \left(9\frac{E}{\rho^5} - 6\frac{E}{\rho^3} + \frac{3}{4}\right)\dot{\varepsilon} \cos(2\theta) + \left(\frac{1}{4} - 2\frac{D}{\rho^3} + 3\frac{E}{\rho^5} - 2\frac{E}{\rho^3}\right)\dot{\varepsilon} \tag{25a}$$

$$\dot{\varepsilon}_{\theta\theta} = \frac{1}{r}\frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r} = \left(3\frac{E}{\rho^3} - \frac{21}{4}\frac{E}{\rho^5} - \frac{3}{4}\right)\dot{\varepsilon} \cos(2\theta) + \left(\frac{E}{\rho^3} + \frac{1}{4} + \frac{D}{\rho^3} - \frac{3}{4}\frac{E}{\rho^5}\right)\dot{\varepsilon} \tag{25b}$$

$$\dot{\varepsilon}_{\varphi\varphi} = \frac{u_r}{r} + u_\theta \frac{\cot(\theta)}{r} = \left(-\frac{3}{4} - \frac{6E}{4\rho^5}\right)\dot{\varepsilon} \cot(\theta)\sin(2\theta) + \left(\frac{3}{4} + 3\frac{E}{\rho^3} - \frac{9}{4}\frac{E}{\rho^5}\right)\dot{\varepsilon} \cos(2\theta) + \left(\frac{E}{\rho^3} + \frac{1}{4} + \frac{D}{\rho^3} - \frac{3}{4}\frac{E}{\rho^5}\right)\dot{\varepsilon} \tag{25c}$$

$$\dot{\varepsilon}_{r\theta} = \frac{1}{2}\left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} + \frac{1}{r}\frac{\partial u_r}{\partial \theta}\right) = \frac{3}{4}\dot{\varepsilon} \sin(2\theta)\left(8\frac{E}{\rho^5} - 4\frac{E}{\rho^3} - 1\right) \tag{25d}$$

where $\rho = r/r_0$.

**6.2 Calculating Energy Rates.** The traction work rate is given as

$$\int T_i u_i ds = \int_{S_T} \sigma_{ij} n_j u_i ds = \int_V (\sigma_{ij} u_i)_{,j} dv = \int_V \sigma_{ij} u_{i,j} dv = \int_V (\sigma_{xx}\dot{\varepsilon}_{xx} + \sigma_{yy}\dot{\varepsilon}_{yy} + \sigma_{zz}\dot{\varepsilon}_{zz} + 2(\sigma_{xy}\dot{\varepsilon}_{xy} + \sigma_{xz}\dot{\varepsilon}_{xz} + \sigma_{zy}\dot{\varepsilon}_{zy}))dv \tag{26a}$$

The work rate of the body force is calculated in the following manner:

$$\dot{W}_b = \int_V (X_r u_r + X_\theta u_\theta) dV = \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{r=r_0}^{r_0 n} (X_r u_r + X_\theta u_\theta) r^2 \sin\theta dr d\theta d\varphi \tag{26b}$$

The process of getting the upper-bound solution in terms of the variables $\sigma_\infty$ and $p$ is an algebraic evaluation of the terms in Eq. (10) given by Eqs. (11)–(14), (15*a*), (15*b*), (16)–(24), (25*a*)–(25*d*), (26*a*), and (26*b*). The result is plotted in Fig. 7.

The final description of the dual shakedown bounds (elastic and plastic) is plotted in Figs. 8(*a*) and 8(*b*) along with few experimental data taken from Sanderow et al. [27] and Katsushi et al. [28]. The data were generated by rotating-bending fatigue of steel structures made by sintered powder. The materials so produced have different amounts of porosities as measured by relative overall densities. The sintered materials could have included residual stress (from any reason) though not indicated in the data. Therefore, for demonstrating the possible effect of the residual stress on the shakedown bounds, we arbitrarily assumed extreme cases: The magnitude of the residual stress in the material is either zero (Fig. 8(*a*)) or $p = \pm 0.3\sigma_0$ (Fig. 8(*b*)).

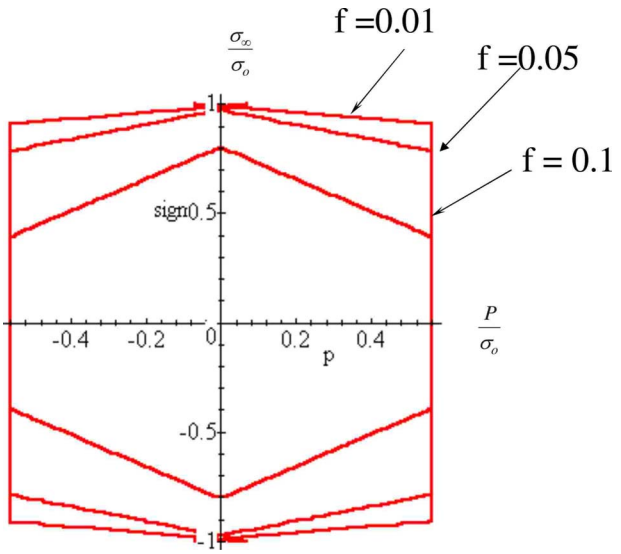



**Fig. 7 The plastic shakedown condition (the equality sign of Eq. (10)) for a safe fluctuating stress in materials with porosity fractions of $f=0.01$, $f=0.05$, and $f=0.1$**

The collection of the whole experimental data, which appear in Figs. 8(*a*) and 8(*b*) (as extracted from Sanderow et al. [27] and Katsushi et al. [28]), is summarized, for convenience, in Appendix C.

## 7 Discussion

The dual shakedown theorems provide a rough predictive tool for the highest allowable stress amplitude before ratcheting erupted. The beauty of the analysis stems from the fact that there is no need to postulate specific mechanisms by which the material fails, besides the common belief that it is associated with accumulation of plastic straining (ratcheting). Both shakedown bounds are destined to find the maximum stress amplitude, which holds more than $N=10^7$ cycles without signaling plastic accumulation.

Some specific conclusions are given as follows.

(1) Qualitatively, residual stress (whether real or imaginary) affects the computational assessment of the fatigue limit. In its presence, as seen in Fig. 5, the elastic shakedown indicates *higher fatigue limit* than in its absence for *negative* sign of $p$. In its presence at the plastic shakedown solution, as seen in Fig. 7, the residual stress *always* leads to a lower upper-bound of the fatigue limit (due to its symmetry effect with respect to the ordinary fatigue limit at $p=0$). Figures 8(*a*) and 8(*b*) intend to comprehend quantitatively both bounds with experiments [29].

(2) It is believed that initial void growth is retarded under compressive residual stress (and accelerated in tensile ones). This feature is conceptually "responsible" for the elastic shakedown of porous materials, as characterized in Fig. 5. More physical ground to such conclusion can be found in Chap. 13 of Ref. [30].

(3) The porosity content influences moderately the fatigue behavior. The endurance limit is always decreasing when the porosity content is increasing.

(4) In actual manufacturing processes it was evidenced that wrought, forged, and cast materials are commonly left with some amount of voids (usually less than 3%) to which the suggested dual bounds are applicable. Also, certain steel alloys contain "soft" particles and can be categorized as porous materials [31]. Therefore, the suggested assessment of ratchet-free loading seems to have a wide use.

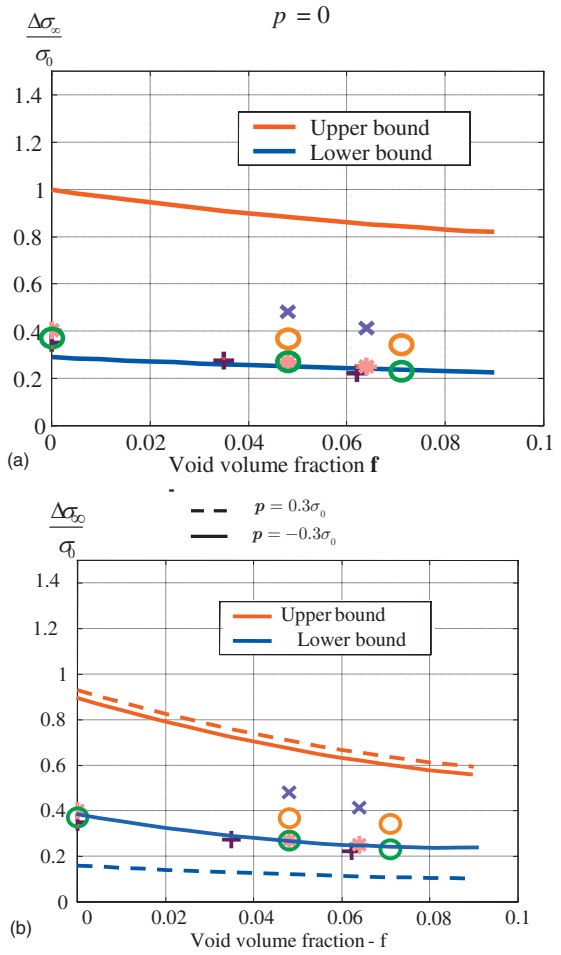



**Fig. 8 (*a*) Dual bounds for the ordinary porous materials (with zero residual stresses). The experimental data are from Sanderow et al. [27] and Katsushi et al. [28]. (*b*) Dual shakedown bounds for porous materials with a prescribed amount of residual stresses of magnitude $p=\pm 0.3\sigma_0$. The experimental data are taken from Sanderow et al. [27] and Katsushi et al. [28].**

(5) Admittedly, the dual bounds shown in this work are still relatively far from each other (e.g., Figs. 8(*a*) and 8(*b*)) for a reliable decision about the permissible fatigue stress amplitudes. Engineeringwise a smaller gap between the bounds is certainly desired, but presently no other experimental-free alternative is offered to assess such conditions in porous materials.


## Acknowledgment

We acknowledge the help of Dr. L. Ruvinski and Dr. M. Feldman from the Faculty of Mechanical Engineering. The encouraging environment and the generous funds for graduate students were all what we needed to produce this work.


## Appendix A: Residual Stress Distribution Around a Spherical Void

We visualize here a case of a temperature drop, $T$ (which is denoted as $\Delta T$ in the text), between the center of the cavity and its surrounding. Due to the radial symmetry there will be only three nonzero stress components: the radial component $\sigma_{rr}$, and two tangential components $\sigma_{\theta\theta}=\sigma_{\varphi\varphi}$. They are inter-related by a single equilibrium equation in the radial direction, namely,

$$\frac{d\sigma_{rr}}{dr}+\frac{2}{r}(\sigma_{rr}-\sigma_{\theta\theta})=0 \tag{A1}$$

The stress strain relations in the presence of a temperature drop are as follows:

$$\dot{\varepsilon}_{rr}-\alpha T=\frac{1}{E}(\sigma_{rr}-2\nu\sigma_{\theta\theta}) \tag{A2}$$

$$\dot{\varepsilon}_{\theta\theta}-\alpha T=\frac{1}{E}(\sigma_{\theta\theta}-\nu(\sigma_{rr}+\sigma_{\theta\theta})) \tag{A3}$$

If $u_r$ is the radial velocity, we get the strain rate components as

$$\dot{\varepsilon}_{rr}=\frac{du_r}{dr},\quad \dot{\varepsilon}_{\theta\theta}=\frac{u_r}{r} \tag{A4}$$

By substituting Eqs. (A2)–(A4) in Eq. (A1) we find the equilibrium equation in terms of the single variable $u$ as

$$\frac{d}{dr}\left[\frac{1}{r^2}\frac{d}{dr}(r^2u_r)\right]=\frac{1+\nu}{1-\nu}\alpha\frac{dT}{dr} \tag{A5}$$

whose solution is

$$u_r=\frac{1+\nu}{1-\nu}\alpha\frac{1}{r^2}\int_a^r \mathrm{Tr}^2\,dr+C_1r+\frac{C_2}{r^2} \tag{A6}$$

By conversion to stresses, Eq. (A6) becomes

$$\sigma_{rr}=-\frac{2\alpha E}{1-\nu}\frac{1}{r^3}\int_a^r \mathrm{Tr}^2\,dr+\frac{EC_1}{1-2\nu}-\frac{2EC_2}{1+\nu}\frac{1}{r^3} \tag{A7}$$

$$\sigma_\theta=\frac{\alpha E}{1-\nu}\frac{1}{r^3}\int_a^r \mathrm{Tr}^2\,dr+\frac{EC_1}{1-2\nu}-\frac{EC_2}{1+\nu}\frac{1}{r^3}-\frac{\alpha ET}{1-\nu} \tag{A8}$$

Denote $a$ and $b$ as the inner and outer radii of the stress-free void and remote boundary, namely,

$$\sigma_{rr}(r=b)=0,\quad \sigma_{rr}(r=a)=0 \tag{A9}$$

and assume that

$$T(r=a)=T_a,\quad T(r=b)=0$$

The integration constants $C_1$ and $C_2$ in Eqs. (A7) and (A8) are recovered, so that

$$\sigma_{rr}=\frac{2\alpha E}{1-\nu}\left[\frac{r^3-a^3}{(b^3-a^3)r^3}\int_a^b \mathrm{Tr}^2\,dr-\frac{1}{r^3}\int_a^r \mathrm{Tr}^2\,dr\right] \tag{A10}$$

$$\sigma_{\theta\theta}=\frac{2\alpha E}{1-\nu}\left[\frac{2r^3+a^3}{2(b^3-a^3)r^3}\int_a^b \mathrm{Tr}^2\,dr+\frac{1}{2r^3}+\int_a^r \mathrm{Tr}^2\,dr-\frac{1}{2}T\right] \tag{A11}$$

Consider, as an example, a cavitated material with steady heat flow. We denote the temperature at the inner surface by $T_a$ and the temperature at the outer surface as zero. Then the temperature at any radial distance from the void center is

$$T=\frac{T_a a}{b-a}\left(\frac{b}{r}-1\right) \tag{A12}$$

Substituting Eq. (A12) in Eqs. (A10) and (A11) we find

$$\sigma_{rr}=\frac{\alpha ET}{1-\nu}\frac{ab}{b^3-a^3}\left[a+b-\frac{1}{r}(b^2+ab+a^2)+\frac{a^2b^2}{r^3}\right] \tag{A13}$$

$$\sigma_{\theta\theta}=\frac{\alpha ET}{1-\nu}\frac{ab}{b^3-a^3}\left[a+b-\frac{1}{2r}(b^2+ab+a^2)-\frac{a^2b^2}{2r^3}\right] \tag{A14}$$

Define $p$ as the magnitude of the residual stress (see details in the text), then the residual stresses at infinitely large material ($b \to \infty$) are

$$\sigma_{rr}^{(\mathrm{res})} = p \cdot \lim_{b\to\infty}\left[\frac{(a/b)^2 + (a/b) - (1/r)\cdot(a + (a^2/b) + (a^3/b^2)) + (a^3/r^3)}{1-(a/b)^3}\right] = -\frac{pa}{r}\left[1 - \left(\frac{a}{r}\right)^2\right] \tag{A15}$$

$$\sigma_{\theta\theta}^{(\mathrm{res})} = p \cdot \lim_{b\to\infty}\left[\frac{(a/b)^2 + (a/b) - (1/2r)\cdot(a + (a^2/b) + (a^3/b^2)) - (a^3/2r^3)}{1-(a/b)^3}\right] = -\frac{pa}{2r}\left[1 + \left(\frac{a}{r}\right)^2\right] \tag{A16}$$

$$\sigma_{\varphi\varphi}^{(\mathrm{res})} = -\frac{pa}{2r}\left[1 + \left(\frac{a}{r}\right)^2\right] \tag{A17}$$

Let $a$ in Eqs. (A15)–(A17) be $r_0$ (as in the text) and let $r$ be normalized by $r_0$, then Eq. (6) in the text is obtained.

## Appendix B: The Porous Material Model

The porous material is modeled as a doubly periodic array of spherical (stress-free) voids. The half length of the unit-cell, $x_0$, is shown in terms of the porosity content $f$ (see Fig. 9).

## Appendix C: Collection of Fatigue Data

Table 1 was taken from Sanderow et al. [27] and Katsushi et al. [28].

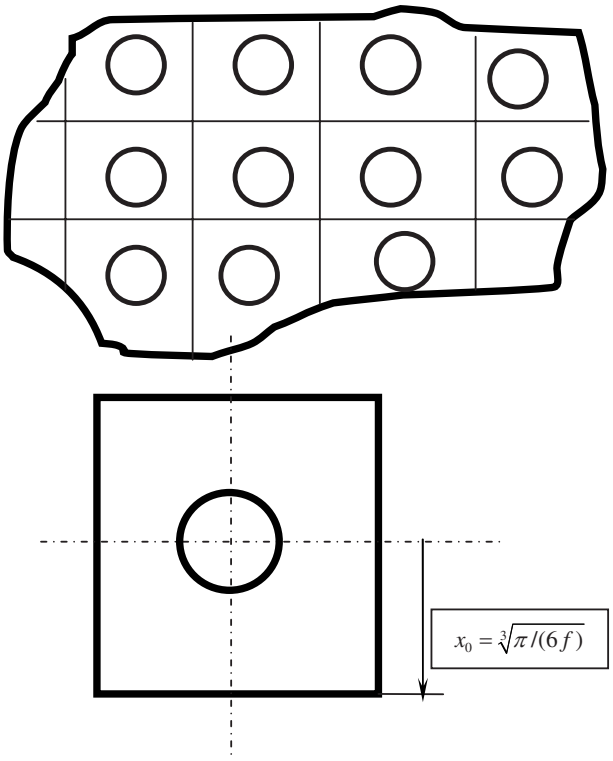



**Fig. 9 Porous material with uniform array of spherical voids**

**Table 1 Fatigue data of sintered powder steels, with fatigue limits and porosity *f***

| Material (steels) | Condition | Density ($g/cm^3$) | Yield stress, $\sigma_0$ (ksi) | Fatigue limit (ksi) | Porosity, $f$ (%) |
|---|---|---|---|---|---|
| FN-0205 | S | 6.91 | 57.4 | 21.8 | 6.2 |
| | HT 350 | 6.91 | 144 | 45.9 | 6.2 |
| | HT 500 | 6.91 | 132.8 | 37.4 | 6.2 |
| | HT 350 | 7.11 | 170.3 | 41.2 | 3.5 |
| | HT 500 | 7.11 | 154.5 | 40.1 | 3.5 |
| | S | 7.21 | 70.7 | 28.3 | 2.1 |
| | S | 7.37 | 82.7 | 28.6 | 0 |
| | HT 350 | 7.37 | 210.1 | 55.6 | 0 |
| | HT 500 | 7.37 | 189.2 | 55.85 | 0 |
| FN-0208 | HT 350 | 6.91 | 136.9 | 46.2 | 6.4 |
| | HT 500 | 6.91 | 140.6 | 46.4 | 6.4 |
| | HT 350 | 7.03 | 151.8 | 51 | 4.8 |
| | HT 500 | 7.39 | 205.6 | 61.3 | 0 |
| | HT 500 | 7.03 | 148.4 | 46.2 | 4.08 |
| | S | 6.91 | 65.1 | 21 | 6.4 |
| | S | 7.1 | 72.1 | 27.4 | 2.4 |
| | S | 7.28 | 93.8 | 33 | 0 |
| FN–0405 | S | 6.56 | 51.8 | 16.7 | 11.2 |
| | S | 7.03 | 74.4 | 27.3 | 3.4 |
| | S | 7.39 | 93.2 | 34.6 | 0 |
| | HT 350 | 6.86 | 147.8 | 44.05 | 7.17 |
| | HT 500 | 6.86 | 119.6 | 36.8 | 7.17 |
| | HT 500 | 7.03 | 140 | 46.3 | 4.87 |
| | HT 500 | 7.39 | 190.4 | 57.9 | 0 |

## References


[1] Melan, E., 1936, "Theorie Statisch Unbestimmter Systeme aus Idealplastischem Baustoff," Sitzungsber. Akad. Wiss. Wien, Math.-Naturwiss. Kl., Abt. 2A, **145**, pp. 195–218.

[2] Koiter, W. T., 1960, "General Theorems for Elastic-Plastic Solids," *Progress in Solid Mechanics I*, Vol. 6, I. N. Sneddon and R. Hill, eds., North-Holland, Amsterdam, pp. 203–313.

[3] Polizzoto, C., 1993a, "On the Condition to Prevent Plastic Shakedown of Structures: Part I Theory, Part II: The Plastic Shakedown Limit Load," ASME Trans. J. Appl. Mech., **60**, pp. 15–25.

[4] Polizzotto, C., 1993b, "A Study on Plastic Shakedown of Structures: Part I: Basic Properties Part II: Theorems," ASME Trans. J. Appl. Mech., **60**, pp. 318–330.

[5] Druyanov, B., and Roman, I., 1997, "Features of the Stress Path at the Post Adaptation Stage and Related Shakedown Conditions," Int. J. Solids Struct., **34**, pp. 3773–3780.

[6] Druyanov, B., and Roman, I., 1997, "Concept of the Limit Yield Condition in Shakedown Theory," Int. J. Solids Struct., **34**, pp. 1547–1556.

[7] Ponter, A. R. S., and Engelhardt, M., 2000, "Shakedown Limits for a General Yield Condition: Implementation and Application for a Von Mises Yield Condition," Eur. J. Mech. A/Solids, **19**, pp. 423–425.

[8] Tirosh, J., and Peles, S., 2001, "Bounds on the Fatigue Threshold in Metals," J. Mech. Phys. Solids, **49**, pp. 1301–1322.

[9] Tirosh, J., and Peles, S., 2003, "Shakedown Bounds for Fatigue Limit in Two Phase Materials," Int. J. Fract., **119**, pp. 65–81.

[10] Kapoor, A., and Williams, J. A., 1996, "Shakedown Limits in Rolling Sliding Point Contact on an Anisotropic Half Space," Wear, **191**, pp. 256–260.

[11] Wong, S. K., Kapoor, A., and Williams, J. A., 1997, "Shakedown Limits on

Coated and Engineered Surfaces," Wear, **203-204**, pp. 162–170.
[12] Dvorack, G. J., and Tarn, J. Q., 1975, "Fatigue and Shakedown in Metal Matrix Composite," *Fatigue of Composite Materials*, ASTM STP Vol. 569, American Society for Testing and Materials, Philadelphia, PA, pp. 145–168.
[13] Jansson, S., and Leckie, F. A., 1992, "Mechanical Behavior of a Continuous Fiber Reinforced Aluminum Matrix Composite Subjected to Transverse and Thermal Loading," J. Mech. Phys. Solids, **40**(3), pp. 593–612.
[14] Tirosh, J., 1998, "The Dual Shakedown Conditions for Dilute Fibrous Composites," J. Mech. Phys. Solids, **46**, pp. 167–185.
[15] Ponter, A. R. S., and Karadeniz, S., 1985, "An Extended Shakedown Theory for Structures That Suffer Cyclic Thermal Loading, Part I," ASME J. Appl. Mech., **52**, pp. 877–882.
[16] Ponter, A. R. S., and Karadeniz, S., 1985, "An Extended Shakedown Theory for Structures That Suffer Cyclic Thermal Loading, Part II, ASME J. Appl. Mech., **52**, pp. 883–889.
[17] Xue, M. D., Wang, X. F., Williams, F. W., and Xu, B. Y., 1997, "Lower Bound Shakedown Analysis of Axisymmetric Structure Subjected to Variable Mechanical and Thermal Loads," Int. J. Mech. Sci., **39**, pp. 965–976.
[18] Huang, Y. J., and Stein, E., 1995, "Prediction of the Fatigue Threshold for a Cracked Body Using Shakedown Theory," Fatigue Fract. Eng. Mater. Struct., **18**(3), pp. 363–370.
[19] Huang, Y. J., and Stein, E., 1996, "Shakedown of a Cracked Body Consisting of Kinematic Hardening Materials," Eng. Fract. Mech., **54**(1), pp. 107–112.
[20] Belouchrani, M. A., and Weichert, D., 1999, "An Extension of the Static Shakedown Theorem to Inelastic Cracked Structures," Int. J. Mech. Sci., **41**, pp. 163–177.
[21] Tirosh, J., 2008, "Extended Fatigue Life by Shot-Peening Process Via Shakedown Analysis," ASME J. Appl. Mech., **75**(1), p. 011005.
[22] Gurson, A. L., 1977, "Continuum Theory of Ductile Rupture by Void Nucleation and Growth—Part I: Yield Criteria and Flow Rules for Porous Ductile Media," ASME J. Eng. Mater. Technol., **114**, pp. 2–15.
[23] Tvergaard, V., 1981, "Influence of Voids on Shear Band Instabilities Under Plain Strain Conditions," Int. J. Fract., **17**, pp. 389–407.
[24] Tvergaard, V., 1981, "Ductile Fracture by Cavity Nucleation Between Larger Voids," Int. J. Fract., **18**, pp. 237–251.
[25] Southwell, R. V., and Gough, H. J., 1926, "On the Concentration of Stress in the Neighbourhood of a Small Spherical Flaw, and the Propagation of Fatigue Fracture in Statistically Isotropic Materials," Philos. Mag., **1**, pp. 71–96.
[26] Rice, J. R., and Tracey, D. M., 1969, "On the Ductile Enlargement of Voids in Triaxial Stress Field," J. Mech. Phys. Solids, **17**, pp. 201–217.
[27] Sanderow, S., Spirko, J. R., and Fredhoff, T. G., 1997, "Fatigue Properties of P/M Materials," *Advances in Powder Metallurgy and Particulated Materials*, Metal Powder Industries Federation, Chicago, IL.
[28] Katsushi, S., Shoji, H., Hironori, F., and Tohru, A., 1989, *Fatigue Strength of Steels With Thin Hard Coating*, Elsevier, Amsterdam, The Netherlands.
[29] Sonsino, C. M., and Ziese, J., 1993, "Fatigue Strength and Application of Cast Aluminum Alloys With Different Degrees of Porosity," Int. J. Fatigue, **15**, pp. 75–84.
[30] Suresh, S. 1999, *Fatigue of Materials* (Solid State Science Series), Cambridge University Press, Cambridge.
[31] Becker, R., Smelser, R. E., and Richmond, O., 1989, "The Effect of Void Shape on the Development of Damage and Fracture in Plain Strain Tension," J. Mech. Phys. Solids, **37**, pp. 111–129.

**Sergei Alexandrov**
Department of Mechanical Engineering,
Yung-Ta Institute of Technology and Commerce,
316 Chunshan Road,
Lin-Lo, Ping-Tung 909, Taiwan
e-mail: sergei_alexandrov@yahoo.com

**Yeau-Ren Jeng**
Department of Mechanical Engineering,
National Chung Cheng University,
168 University Road,
Ming-Hsiung, Chia-Yi 621, Taiwan

# Compression of Viscoplastic Material Between Rotating Plates

*An analysis is conducted of the two-dimensional flow of Bingham solids between two rotating plates. The maximum friction law is adopted at the plate surface. An asymptotic analysis of the solution is performed in the vicinity of the friction surface. Its results are used in a numerical procedure to obtain an accurate approximation of the solution near the friction surface. The through thickness distribution of velocities, the equivalent strain rate, and stresses is illustrated. Qualitative features of the solution are emphasized. The results are compared with the solution for rate-independent materials.*
[DOI: 10.1115/1.3005962]

## 1 Introduction

There are a number of rigid perfectly/plastic solutions obtained by inverse methods (see Refs. [1–3] among others). The monograph [1] also includes earlier solutions proposed by Prandtl and Nadai. In the case of problems with frictional interfaces, the friction law postulating that the friction stress at sliding is equal to a portion of the shear yield stress is usually adopted. A particular case of this friction law is the maximum friction law (the friction stress is equal to the shear yield stress). The maximum friction law is of special interest because the corresponding friction stress at sliding is higher or equal to the friction stress following from any other possible friction laws. Also, rigid perfectly/plastic solutions are singular in the vicinity of maximum friction surfaces [4]. Many of the aforementioned rigid perfectly/plastic solutions have been generalized on other rate-independent plasticity models, for instance Refs. [5–9], where the double-shearing model proposed in Ref. [10] was adopted. These solutions are singular in the vicinity of maximum friction surfaces as well. However, attempts to generalize the rigid perfectly/plastic solutions on rate-dependent materials are not completely successful. It is convenient to divide such materials into two groups: viscoplastic fluids (the yield stress vanishes at zero strain rate) and viscoplasic solids (the yield stress is of a finite value at zero strain rate) [11]. In the present paper, viscoplastic solids only are considered. It is possible to verify by inspection that the solution for the compression of a viscoplastic layer between parallel plates proposed in Ref. [12] is not valid in the case of the maximum friction law and predicts unrealistic results if the friction stress is less than the maximum friction stress but is still high enough. The solution for axisymmetric viscoplastic flow through infinite conical channels proposed in Ref. [13] does not satisfy the friction boundary conditions. Approximate solutions to the same and other similar problems found in Refs. [14,15] are valid at very small friction stresses only. A detailed discussion of plane-strain flows through wedge-shaped channels, including both converging and diverging flows, is provided in Ref. [16] where special attention is devoted to the qualitative behavior of the solutions, and difficulties similar to those appearing in the axisymmetric solutions are mentioned. One of the reasons for such difficulties with obtaining viscoplastic solutions of the class considered is that the velocity fields in the aforementioned solutions found by inverse methods require the regime of sliding at the friction surface whereas the general theory predicts the regime of sticking at maximum friction surfaces [17]. In the present paper, compression of a viscoplastic layer between rotating plates is studied, and it is shown that the problem permits a solution at the regime of sticking. As in the case of other plastic solutions obtained by inverse methods (for example Refs. [1–3,5,6,12,13]), the process is not feasible for practical realization. Therefore, it seems that the major significance of the solution and its potential applications lie in the area of numerical simulation of problems whose solutions cannot be found in a closed form. The qualitative behavior of the solution emphasized in the present paper shows that expressions like 0/0 can appear in the vicinity of maximum friction surfaces. Comparison made with the solution obtained in Ref. [8] for rate-independent materials demonstrates that the qualitative behavior of the solution depends on the material model chosen and that the solution may be singular. Asymptotic behavior of solutions and, in particular, the singularity enrichment functions are used in such finite element formulations as the extended finite element method (XFEM) [18]. Therefore, it is of importance to know the asymptotic behavior of solutions near certain lines or surfaces in advance.

## 2 Statement of the Problem

Consider the plane-strain deformation of a viscoplastic material between two plates rotating around a common axis with an angular velocity $\omega$ (Fig. 1). It is convenient to introduce a cylindrical coordinate system $r\theta z$ with its $z$-axis being perpendicular to the plane of flow. Also, the $z$-axis coincides with the axis of rotation of the plates. The flow is symmetric with respect to the axis $\theta=0$. It is therefore sufficient to obtain the solution in the region $0 \le \theta \le \alpha$, where $\alpha$ is the orientation of the plate surface (Fig. 1). Since the material is history-independent, it is sufficient to obtain the instantaneous solution, i.e., $\alpha$ is an arbitrary fixed constant. It is assumed that there is no material flux at $r=0$. A conventional system of equations for the quasistatic plane-strain deformation of a viscoplastic material consists of the following equilibrium equations:

$$\frac{\partial \sigma_{rr}}{\partial r}+\frac{1}{r}\frac{\partial \sigma_{r\theta}}{\partial \theta}+\frac{\sigma_{rr}-\sigma_{\theta\theta}}{r}=0, \quad \frac{1}{r}\frac{\partial \sigma_{\theta\theta}}{\partial \theta}+\frac{\partial \sigma_{r\theta}}{\partial r}+\frac{2\sigma_{r\theta}}{r}=0 \tag{1}$$

the yield criterion,

$$(\sigma_{rr}-\sigma_{\theta\theta})^2+4\sigma_{r\theta}^2=4k^2 \tag{2}$$

and its associated flow rule,

$$\xi_{rr}=\lambda(\sigma_{rr}-\sigma), \quad \xi_{\theta\theta}=\lambda(\sigma_{\theta\theta}-\sigma), \quad \xi_{r\theta}=\lambda\sigma_{r\theta} \tag{3}$$

Here $\sigma_{rr}$, $\sigma_{\theta\theta}$, and $\sigma_{r\theta}$ are the components of the stress tensor in the cylindrical coordinates, $\xi_{rr}$, $\xi_{\theta\theta}$, and $\xi_{r\theta}$ are the components of the strain rate tensor in the cylindrical coordinates, $\lambda$ is a nonnegative multiplier, $\sigma$ is the hydrostatic stress defined by

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received December 12, 2007; final manuscript received September 4, 2008; published online March 13, 2009. Review conducted by Antoinette Maniatty.

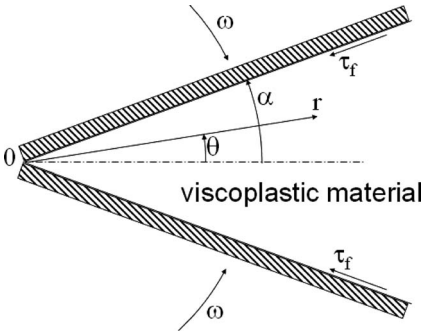



**Fig. 1 Geometry of the process**

$$\sigma = \frac{\sigma_{rr} + \sigma_{\theta\theta}}{2} \tag{4}$$

and $k$ is a prescribed function of the equivalent strain rate $\xi_{eq}$ defined by

$$\xi_{eq} = \sqrt{\frac{2}{3}}\sqrt{\xi_{rr}^2 + \xi_{\theta\theta}^2 + 2\xi_{r\theta}^2} \tag{5}$$

The physical meaning of $k$ is the shear yield stress at any given value of $\xi_{eq}$. By assumption, $dk/d\xi_{eq} > 0$ at any value of $\xi_{eq}$ and $k = k_0 > 0$ at $\xi_{eq} = 0$, where $k_0$ is the initial shear yield stress. The incompressibility equation in the form $\xi_{rr} + \xi_{\theta\theta} = 0$ immediately follows from Eqs. (3) and (4). The strain rate components are expressed in terms of the velocity components, $u_r$ and $u_\theta$, as

$$\xi_{rr} = \frac{\partial u_r}{\partial r}, \quad \xi_{\theta\theta} = \frac{1}{r}\frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r}, \quad \xi_{r\theta} = \frac{1}{2}\left(\frac{1}{r}\frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r}\right) \tag{6}$$

The yield criterion (2) is satisfied by the standard substitution

$$\sigma_{rr} = \sigma + k\cos 2\varphi, \quad \sigma_{\theta\theta} = \sigma - k\cos 2\varphi, \quad \sigma_{r\theta} = -k\sin 2\varphi \tag{7}$$

The geometry of the process assumes that $\xi_{\theta\theta} \le 0$ and, therefore, the incompressibility equation gives $\xi_{rr} \ge 0$. Hence, $\xi_{rr} - \xi_{\theta\theta} \ge 0$ and it follows from Eq. (3) that $\sigma_{rr} - \sigma_{\theta\theta} \ge 0$. Also, the direction of the friction stress (Fig. 1) requires that $\sigma_{r\theta} \le 0$. Using these inequalities it is possible to conclude from Eq. (7) that

$$0 \le \varphi \le \frac{\pi}{4} \tag{8}$$

Combining Eqs. (3) and (7) and excluding λ Eq. (3) can be transformed into

$$\xi_{rr} + \xi_{\theta\theta} = 0, \quad \xi_{\theta\theta} = \xi_{r\theta}\cot 2\varphi \tag{9}$$

Substituting Eq. (9) into Eq. (5) and taking into account Eq. (8) gives

$$\xi_{eq} = -\frac{2}{\sqrt{3}}\frac{\xi_{\theta\theta}}{\cos 2\varphi} \tag{10}$$

The stress boundary condition at the axis of symmetry is $\sigma_{r\theta} = 0$. Using Eqs. (7) and (8) this condition can be rewritten in the form

$$\varphi = 0 \tag{11}$$

at $\theta = 0$. The velocity boundary condition at the axis of symmetry is

$$u_\theta = 0 \tag{12}$$

at $\theta = 0$. One of the boundary conditions at the plate surface has the form

$$u_\theta = -\omega r \tag{13}$$

at $\theta = \alpha$. The other boundary condition at this surface depends on the friction law. Here the maximum friction law is adopted. In general, its mathematical formulation is

$$u_r = 0 \quad \text{at} \quad \tau_f \le k$$

$$\tau_f = k \quad \text{otherwise} \tag{14}$$

The first part of this law corresponds to the regime of sticking and the second part to the regime of sliding. However, it has been shown in Ref. [17] that the regime of sliding never occurs if $k \to \infty$ as $\xi_{eq} \to \infty$. In what follows such constitutive equations only are adopted. Therefore, Eq. (14) transforms into

$$u_r = 0 \tag{15}$$

at $\theta = \alpha$.

## 3 Velocity Equations

A typical assumption accepted in many classical rigid plastic solutions is that $\varphi$ is independent of one of the coordinates. In the case under consideration it is reasonable to suppose that $\varphi$ is independent of $r$. The velocity filed is assumed to be in the form

$$u_r = -\frac{\omega r}{2}\frac{df}{d\theta} \quad \text{and} \quad u_\theta = \omega r f(\theta) \tag{16}$$

where $f(\theta)$ is an arbitrary function of $\theta$. This velocity field predicts no material flux at $r = 0$. With the use of Eq. (6), it is possible to verify by inspection that the first equation of system (9) is satisfied at any choice of $f(\theta)$. The second equation of this system reduces to

$$\frac{dF}{d\theta} + 2F\tan 2\varphi = 0 \tag{17}$$

where

$$\frac{df}{d\theta} = F \tag{18}$$

Equation (17) is compatible with the assumptions made since $\varphi$ is independent of $r$. In terms of the new functions, boundary conditions (12), (13), and (15) become

$$f = 0 \tag{19}$$

at $\theta = 0$,

$$f = -1 \tag{20}$$

at $\theta = \alpha$, and

$$F = 0 \tag{21}$$

at $\theta = \alpha$. With the use of Eqs. (16) and (18) expression (10) for the equivalent strain rate transforms into

$$\xi_{eq} = -\frac{\omega}{\sqrt{3}}\frac{F}{\cos 2\varphi} \tag{22}$$

A Bingham solid is defined by

$$k = k_0 + k_0\frac{\xi_{eq}}{\xi_0}, \quad \xi_0 = \text{constant} \tag{23}$$

The quantity $\xi_0$ can be regarded as a characteristic strain rate. Substituting Eq. (22) into Eq. (23) gives

$$\frac{k}{k_0} = 1 - \frac{wF}{\cos 2\varphi}, \quad w = \frac{\omega}{\sqrt{3}\xi_0} \tag{24}$$

## 4 Stress Equations

Equation (24) and the assumptions made show that $k$ is independent of $r$. Therefore, as follows from Eq. (7), the shear stress $\sigma_{r\theta}$ is independent of $r$. Hence, using Eq. (7) equilibrium equations (1) can be transformed into

$$r\frac{\partial\sigma}{\partial r} - \frac{d(k\sin 2\varphi)}{d\theta} + 2k\cos 2\varphi = 0,$$

$$\frac{\partial\sigma}{\partial\theta} - \frac{d(k\cos 2\varphi)}{d\theta} - 2k\sin 2\varphi = 0 \tag{25}$$

The first equation of this system can be integrated to give

$$\sigma = \left[\frac{d(k\sin 2\varphi)}{d\theta} - 2k\cos 2\varphi\right]\ln\left(\frac{r}{R}\right) + k_0\sigma_0(\theta) \tag{26}$$

Here $R$ is an arbitrary constant and $\sigma_0(\theta)$ is an arbitrary function of $\theta$. It can be seen from the second equation of system (25) and (26) that the system is compatible if and only if

$$\frac{d(k\sin 2\varphi)}{d\theta} - 2k\cos 2\varphi = 2k_0A \tag{27}$$

where $A$ is constant. In this case Eq. (26) can be rewritten in the form

$$\frac{\sigma}{k_0} = 2A\ln\left(\frac{r}{R}\right) + \sigma_0(\theta) \tag{28}$$

Substituting Eq. (28) into the second equation of system (25) leads to

$$k_0\frac{d\sigma_0}{d\theta} = \frac{d(k\cos 2\varphi)}{d\theta} + 2k\sin 2\varphi \tag{29}$$

Since the arbitrary constant $R$ has already been introduced in Eq. (26), it is possible to assume with no loss of generality that the boundary condition to Eq. (29) is

$$\sigma_0 = 0 \tag{30}$$

at $\theta=\alpha$.

## 5 Solution

There are four equations to solve, Eqs. (17), (18), (27), and (29). Equation (29) can be solved independently after the solution to the other equations has been found. It is convenient to consider $\varphi$ as the independent variable. Then, $F$, $f$, and $\theta$ are functions of $\varphi$ to be determined from Eqs. (17), (18), and (27). In Eq. (27) the derivative $d/d\theta$ can be replaced with $(d/dF)(dF/d\theta)$ and $k$ can be excluded with the use of Eq. (24). Then, excluding $dF/d\theta$ by means of Eq. (17) gives

$$\frac{dF}{d\varphi} = \frac{2F\tan 2\varphi(\cos^3 2\varphi - wF)}{wF - \cos^3 2\varphi - A\cos^2 2\varphi} \tag{31}$$

Equations (17) and (18) can be transformed in a similar manner to give

$$\frac{d\theta}{d\varphi} = \frac{wF - \cos^3 2\varphi}{wF - \cos^3 2\varphi - A\cos^2 2\varphi} \tag{32}$$

and

$$\frac{df}{d\varphi} = \frac{(wF - \cos^3 2\varphi)F}{wF - \cos^3 2\varphi - A\cos^2 2\varphi} \tag{33}$$

Assume that $\varphi\neq\pi/4$ at $\theta=\alpha$. Since $F=0$ at $\theta=\alpha$, as follows from Eq. (21), Eq. (31) in the vicinity of $\theta=\alpha$ is represented in the form $dF/d\varphi=CF$ where $C$ is constant. Its general solution is $\ln F=C\varphi+C_0$ with $C_0$ being a constant of integration. Obviously, this solution cannot satisfy boundary condition (21) at any value of $C_0$. It is therefore necessary to assume that

$$\varphi = \frac{\pi}{4} \tag{34}$$

at $\theta=\alpha$. In this case the right hand side of Eqs. (31)–(33), reduces to 0/0 at $\theta=\alpha$ (or $\varphi=\pi/4$). It is possible to verify by substitution that the solution to Eq. (31) has the form

$$F = B\left(\frac{\pi}{4} - \varphi\right) + o\left(\frac{\pi}{4} - \varphi\right), \quad \varphi\to\frac{\pi}{4} \tag{35}$$

where $B$ is an arbitrary constant. As follows from Eq. (34), solution (35) satisfies boundary condition (21). Using Eq. (35) and expanding the numerator and denominator in Eqs. (32) and (33) in a series in the vicinity of $\varphi=\pi/4$ gives after integration

$$\theta = \alpha - \left(\frac{\pi}{4} - \varphi\right) + o\left(\frac{\pi}{4} - \varphi\right), \quad \varphi\to\frac{\pi}{4} \tag{36}$$

$$f = -1 - \frac{B}{2}\left(\frac{\pi}{4} - \varphi\right)^2 + o\left[\left(\frac{\pi}{4} - \varphi\right)^2\right], \quad \varphi\to\frac{\pi}{4} \tag{37}$$

The constants of integration were chosen to satisfy boundary conditions (20) and (34). Retaining terms up to $O(\pi/4-\varphi)$ in Eqs. (35) and (36) and up to $O[(\pi/4-\varphi)^2]$ in Eq. (37) it is possible to find the value of $F=F_c$, $\theta=\theta_c$, and $f=f_c$ at $\varphi=\varphi_c$ where $\varphi_c=\pi/4-\delta$ and $\delta\ll 1$. Obviously, the values of $F_c$ and $f_c$ depend on $B$.

In the interval $0\le\varphi\le\varphi_c$ the solution to Eqs. (31)–(33) should be found numerically with the use of $F_c$, $\theta_c$, and $f_c$ as the boundary conditions at $\varphi=\varphi_c$. The two constants, $A$ and $B$, should be determined by means of boundary conditions (11) and (19).

In Eq. (29), the derivatives $d/d\theta$ can be replaced with $(d/d\varphi)\times(d\varphi/d\theta)$ where $d\varphi/d\theta$ should be excluded by means of Eq. (32). Then, integrating with the use of boundary condition (30) and taking into account condition (34) gives

$$\sigma_0 = \cos 2\varphi - wF + 2\int_{\pi/4}^{\varphi}\frac{(\sin 2\varphi - wF\tan 2\varphi)(wF - \cos^3 2\varphi)}{(wF - \cos^3 2\varphi - A\cos^2 2\varphi)}d\varphi \tag{38}$$

Since the solution to Eqs. (31)–(33) determines $F(\varphi)$, $A$, and $B$, integration in Eq. (38) can be performed numerically with no difficulty, though the numerator and denominator of the integrand in Eq. (38) should be expanded in a series in the vicinity of $\varphi=\pi/4$ to find the asymptotic behavior of $\sigma_0$ near the friction surface. As a result,

$$\sigma_0 = -\frac{4A}{Bw}\left(\frac{\pi}{4} - \varphi\right)^2 + o\left[\left(\frac{\pi}{4} - \varphi\right)^2\right], \quad \varphi\to\frac{\pi}{4} \tag{39}$$

Then, the stress tensor is determined by Eq. (7) with the use of Eqs. (24) and (28).

Substituting Eq. (35) into Eq. (22) gives

$$\frac{\xi_{eq}}{\xi_0} = -\frac{wB}{2} + o(1), \quad \varphi\to\frac{\pi}{4} \tag{40}$$

It is also of interest to find the asymptotic behavior of the derivative of the equivalent strain rate with respect to $\theta$ in the vicinity of the friction surface. To this end, it is necessary to calculate the term of order $O[(\pi/4-\varphi)^2]$ in the asymptotic expansion of $F$. Assume that

$$F = B\left(\frac{\pi}{4} - \varphi\right) + B_1\left(\frac{\pi}{4} - \varphi\right)^2 + o\left[\left(\frac{\pi}{4} - \varphi\right)^2\right], \quad \varphi\to\frac{\pi}{4} \tag{41}$$

Substituting this expansion into Eq. (31) and equating coefficients of like powers of $(\pi/4-\varphi)$ to zero yields $B_1=4A/w$. Then, substituting Eq. (41) into Eq. (22) and differentiating with respect to $\varphi$ gives

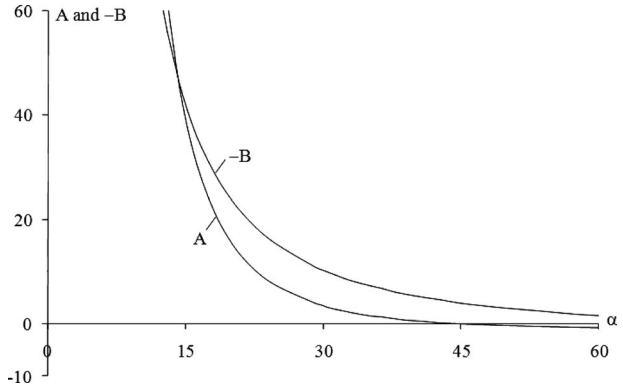



**Fig. 2 Variation of *A* and *B* with angle $\alpha$**

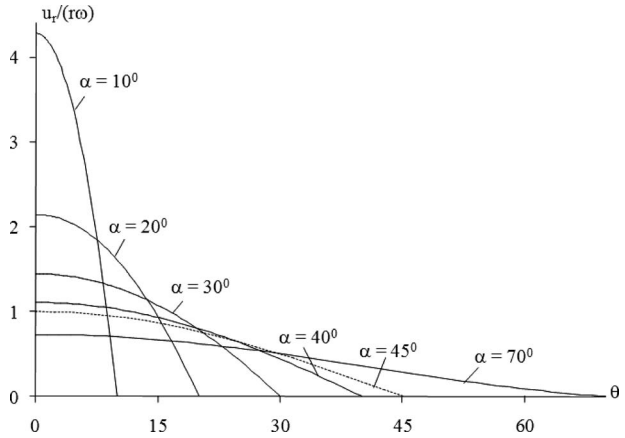



**Fig. 3 Variation of the radial velocity with the polar angle at different values of $\alpha$**

$$\frac{d(\xi_{eq}/\xi_0)}{d\varphi} = \frac{2A}{w} + o(1), \quad \varphi \to \frac{\pi}{4} \tag{42}$$

Taking into account Eq. (36), Eq. (42) can be rewritten in the form

$$\frac{d(\xi_{eq}/\xi_0)}{d\theta} = \frac{2A}{w} + o(1), \quad \theta \to \alpha \tag{43}$$

## 6 Special Case

A special case is obtained at $A=0$. In particular, Eqs. (31)–(33) reduce to

$$\frac{dF}{d\varphi} = -2F\tan 2\varphi, \quad \frac{d\theta}{d\varphi} = 1, \quad \frac{df}{d\varphi} = F \tag{44}$$

The solution to these equations satisfying boundary conditions (11), (19), and (20) is

$$F = -\frac{2\cos 2\varphi}{\sin 2\alpha}, \quad \theta = \varphi, \quad f = -\frac{\sin 2\varphi}{\sin 2\alpha} \tag{45}$$

As follows from this solution boundary condition (21) is satisfied if and only if $\alpha=\pi/4$. Then, Eq. (45) becomes

$$F = -2\cos 2\theta, \quad f = -\sin 2\theta \tag{46}$$

Also, substituting Eq. (45) at $\alpha=\pi/4$ into Eq. (22) it is possible to find

$$\xi_{eq} = 2w\xi_0 \tag{47}$$

Thus $\xi_{eq}$ and $k$ are constant in the entire plastic zone. Therefore, Eq. (29) can be immediately integrated with the use of boundary condition (30) to get $\sigma_0=0$ in the entire plastic zone. However, Eq. (28) should be replaced with $\sigma=$constant because $R$ cannot serve as a constant of integration in this special case. Then, the distribution of stress components is determined by Eq. (7), where $\varphi$ can be replaced with $\theta$ due to Eq. (45).

Since the condition $A=0$ corresponds to $\alpha=\pi/4$, it is reasonable to expect that the sense of $A$ is opposite at $\alpha>\pi/4$ and $\alpha<\pi/4$.

## 7 Numerical Results and Discussion

All calculations illustrated in Figs. 2–8 were performed at $w=1$ and $\delta=0.001$. The definition for $\delta$ is given after Eq. (37). The dependence of $A$ and $B$ introduced in Eqs. (27) and (35) on the value of $\alpha$ is depicted in Fig. 2. It is seen that $B$ is always negative whereas $A$ changes its sense at $\alpha=\pi/4$ such that it is negative at $\alpha>\pi/4$ and positive at $\alpha<\pi/4$. Then, it follows from Eq. (43) that the equivalent strain rate is an increasing function of $\theta$ in the vicinity of the friction surface at $\alpha<\pi/4$ and is a decreasing function of $\theta$ at $\alpha>\pi/4$. Also, Eq. (28) demonstrates that at any fixed value of $R$ the contact pressure becomes negative for sufficiently small $r$-values at $\alpha>\pi/4$ and for sufficiently large $r$-values at $\alpha<\pi/4$. It contradicts the friction law because one of its requirements is that the pressure is positive. However, it is a common defect of many solutions of this class. A typical example is the flow of a plastic material through infinite wedge-shaped and

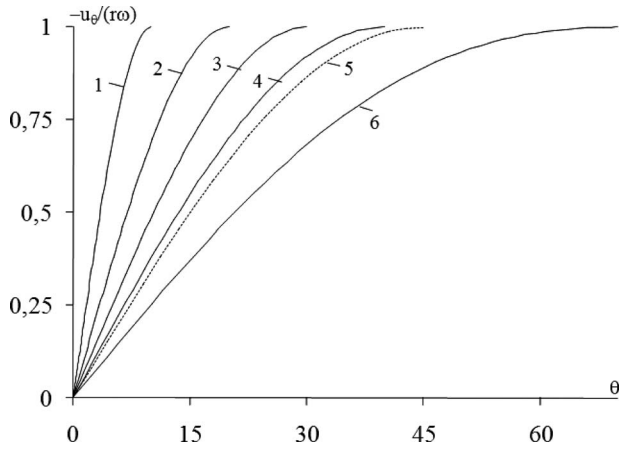



**Fig. 4 Variation of the circumferential velocity with the polar angle at different values of $\alpha$ (1–$\alpha$=10 deg, 2–$\alpha$=20 deg, 3–$\alpha$=30 deg, 4–$\alpha$=40 deg, 5–$\alpha$=45 deg, 6–$\alpha$=70 deg)**

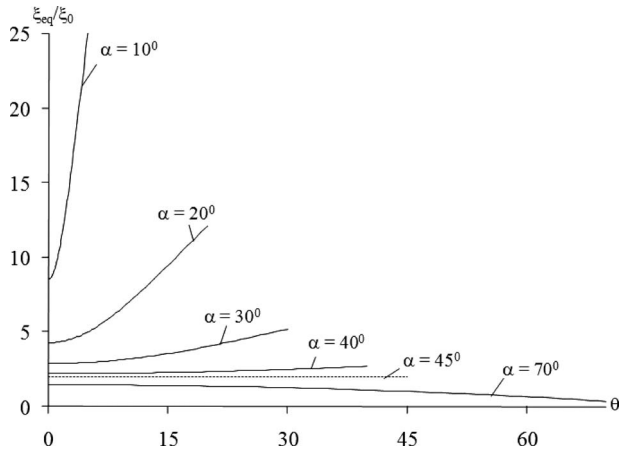



**Fig. 5 Variation of the equivalent strain rate with the polar angle at different values of $\alpha$**

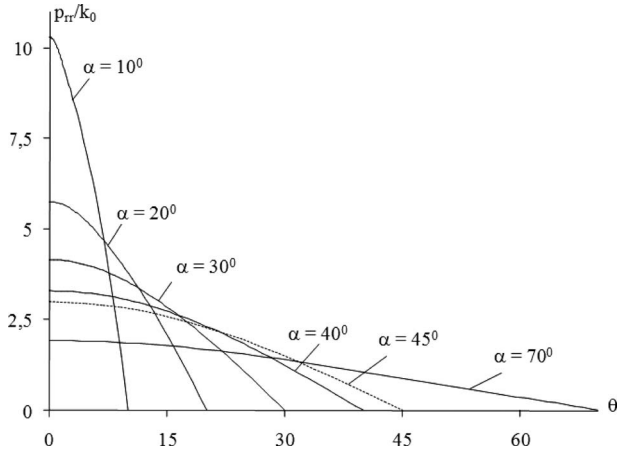



**Fig. 6 Variation of $p_{rr}$ value with the polar angle at different values of $\alpha$**

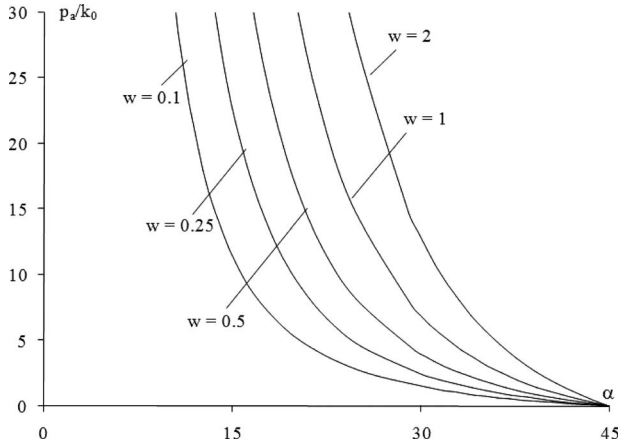



**Fig. 9 Variation of the average pressure with angle $\alpha$ at different values of $w$**

axisymmetric channels [1,2]. Also, the value of $R$ can be chosen such that the positive pressure occurs on any finite portion of the friction surface. The stress and velocity fields were calculated for several $\alpha$-values, including the special case corresponding to $\alpha = \pi/4$. The solution is illustrated in Figs. 3–8 where the special case is shown by dashed lines. The asymptotic expansions found in Eqs. (35)–(37), (39), and (42) were used to obtain an accurate representation of the solution in the vicinity of the friction surface. The variation of the velocity components with angle $\theta$ is depicted in Figs. 3 and 4 at different $\alpha$-values. The dependence of the equivalent strain rate on $\theta$ is shown in Fig. 5. It is seen that for $\alpha$-angles smaller than $\pi/4$ the equivalent strain rate is a monotonically increasing function of $\theta$ in the entire plastic zone, i.e., the equivalent strain rate is higher in the vicinity of the friction surface than at the center of flow. The tendency is opposite for $\alpha$-angles larger than $\pi/4$. It is interesting to note here that solutions of rate-independent plasticity predict a rigid zone in the vicinity of the friction surface at sufficiently large $\alpha$-angles (larger than $\pi/4$ for pressure-independent rigid perfectly plastic material) [8]. This result cannot be extended to the model under consideration. Assume that such a solution exists at $\alpha > \pi/4$ and that the rigid plastic boundary is given by the equation $\theta = \pi/4$. Then, Eq. (47) shows that $\xi_{eq} \neq 0$ and, therefore, $k > k_0$ on the plastic side of the rigid plastic boundary. On the other hand, $k = k_0$ in the rigid zone. Hence, the shear stress cannot have the same value of the two sides of the rigid plastic boundary, which contradicts the equilibrium equations.

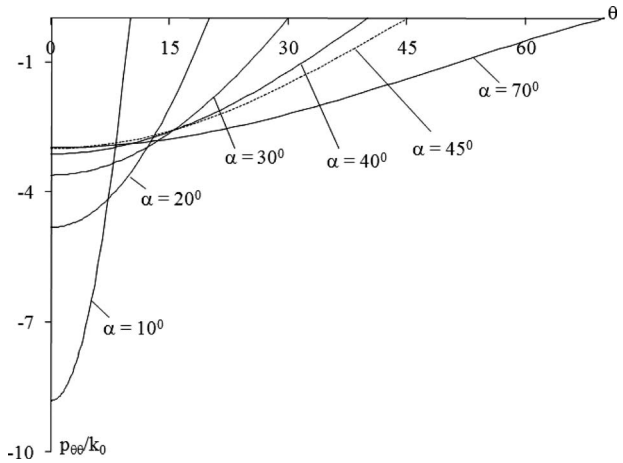



**Fig. 7 Variation of $p_{\theta\theta}$ value with the polar angle at different values of $\alpha$**

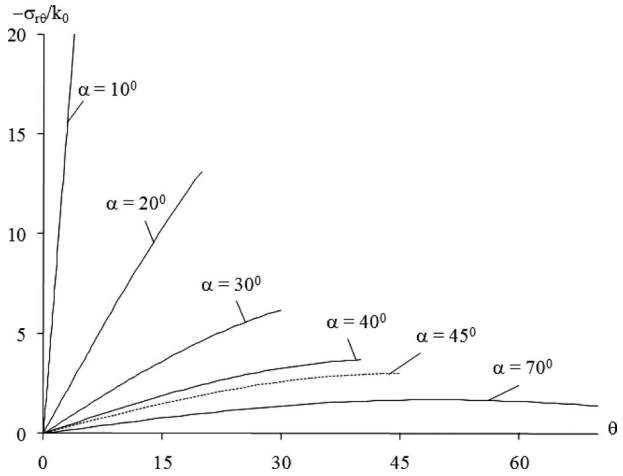



**Fig. 8 Variation of the shear stress with the polar angle at different values of $\alpha$**

The normal stresses cannot be completely calculated since $R$ remains undetermined. Therefore, for the illustration of the through thickness distribution of these stresses it is convenient to introduce the quantities $p_{rr} = \sigma_{rr} + p$ and $p_{\theta\theta} = \sigma_{\theta\theta} + p$, where $p$ is the pressure on the friction surface, i.e., $p$ is equal to $-\sigma_{\theta\theta}$ calculated at $\theta = \alpha$ (or $\varphi = \pi/4$). The variation of $p_{rr}$, $p_{\theta\theta}$, and the shear stress $\sigma_{r\theta}$ is shown in Figs. 6–8 at different $\alpha$-values. It is interesting to mention that the distribution of the shear stress has a maximum within the plastic zone at $\varphi > \pi/4$. The maximum coincides with the frictional interface at $\varphi = \pi/4$.

The average pressure at the friction surface defined as

$$\frac{p_a}{k_0} = -R^{-1}\int_0^R \frac{\sigma_{\theta\theta}}{k_0} dr \tag{48}$$

is independent of $R$. It follows from Eq. (7) that $\sigma_{\theta\theta} = \sigma$ at $\varphi = \pi/4$ (or $\theta = \alpha$). Therefore, substituting Eq. (28), with the use of Eq. (30), into Eq. (48) and integrating gives

$$p_a/k_0 = 2A \tag{49}$$

Since $A$ depends on $w$, this equation illustrates a strain rate effect. Based on the aforementioned numerical procedure, the variation of $A$ with $\alpha$ was calculated for several values of $w$ and, then, Eq. (49) was used to find the average pressure (Fig. 9). The material flux through the surface $r = R$ is determined from the incompressibility equation as $q = \omega R^2/2$. Since the domain $r < R$ was consid-

ered in Eq. (48), a physically meaningful result is obtained at $\alpha < \pi/4$ and, moreover, possible applications of this result are restricted to rather small angles of $\alpha$. It is somehow similar to the restriction that the thickness of the layer should be much smaller that its width in the famous problem of compression of the plastic layer between parallel plates (Prandtl problem) [1]. The dependence of $p_a$ on $w$ is in agreement with physical expectations, higher values of $w$ result in greater pressure.

A boundary layer theory for Bingham solids was proposed in Ref. [19]. In general, it predicts a very high velocity gradient in the vicinity of maximum friction surfaces. For the geometry considered it is equivalent to a high magnitude of the derivative of the radial velocity with respect to the polar angle. However, the solution found develops no boundary layer. It is seen in Fig. 3 and is obvious in the case of a very simple solution given in Sec. 6. A possible reason for this disagreement between the general theory and the exact analytic solution is that the latter has no characteristic length.

## 8 Conclusions

An analytic solution for viscoplastic flow between rotating plates was found. It was shown that an asymptotic analysis of the equations is necessary for obtaining an accurate representation of the solution in the vicinity of the plate surface where the maximum friction law is adopted. This finding may be of importance for developing numerical codes for problems of practical interest. An interesting feature of the solution is that the distribution of the equivalent strain rate is uniform at $\alpha = \pi/4$. This fact can be probably used in the experiment on determining parameters of the model. In particular, compression at $\alpha \approx \pi/4$ should provide more reasonable results because the inhomogeneity in the distribution of the equivalent strain rate is rather small. It is also of interest that no rigid zone appears in the material, though such a zone occurs in the solution to the same problem in the case of rate-independent materials [8], and rigid zones often occur in other viscoplastic solutions, for example, Refs. [20,21]. In the case of a rigid perfectly/plastic material the rigid zone appears at the friction surface at $\alpha > \pi/4$. In the present solution, the qualitative behavior of the equivalent strain rate changes at $\alpha = \pi/4$ (Fig. 5). In particular, the equivalent strain rate attains its minimum value at the friction surface at $\alpha > \pi/4$ and its maximum value at $\alpha < \pi/4$. Note that in the rigid perfectly/plastic solution $\xi_{eq} = 0$ in the rigid zone (i.e., in the vicinity of the friction surface at $\alpha > \pi/4$) and $\xi_{eq} \to \infty$ as $\theta \to \alpha$ at $\alpha < \pi/4$. As compared with squeeze flows studied in Refs. [20,21], the formulation of the present problem is quite different because there is no point on the friction surface where the friction stress vanishes, which, in fact, is the reason for the occurrence of rigid zones in the aforementioned studies.

The present study is intentionally restricted to the friction boundary condition in the form of the maximum friction law. It is motivated by the fact that most of available viscoplastic solutions obtained by inverse methods are not valid if the maximum friction law is adopted, for instance Refs. [12,13]. Therefore, the present solution contains a distinguished qualitative effect. It is interesting to note that the situation is quite different for viscoplastic fluids. In this case there are solutions to boundary value problems incorporating the maximum friction law [22]. Those were also found by an inverse method.

In spite of a possibility to adopt the present solution for the problem of compression of a finite block of material between rotating plates, by means of typical arguments used to apply exact and approximate plastic solutions found by inverse methods to material forming problems (see, for example Refs. (1) and (2)), the primary objective of this study is to provide the exact viscoplastic solution satisfying the maximum friction law. It is believed that the solution can serve as a benchmark problem for verifying numerical codes since accurate closed form solutions are necessary for this purpose [23].

## Acknowledgment

The authors are thankful to NSC (Taiwan) for financial support of this research through Grant No. 96-2221-E-132-007.

## References

[1] Hill, R., 1950, "The Mathematical Theory of Plasticity," Clarendon, Oxford, UK.
[2] Shield, R. T., 1955, "Plastic Flow in a Converging Conical Channel," J. Mech. Phys. Solids, **3**(4), pp. 246–258.
[3] Spencer, A. J. M., 1965, "A Theory of the Failure of Ductile Materials Reinforced by Elastic Fibres," Int. J. Mech. Sci., **7**, pp. 197–209.
[4] Alexandrov, S., and Richmond, O., 2001, "Singular Plastic Flow Fields Near Surfaces of Maximum Friction Stress," Int. J. Non-Linear Mech., **36**(1), pp. 1–11.
[5] Pemberton, C. S., 1965, "Flow of Imponderable Granular Materials in Wedge-Shaped Channels," J. Mech. Phys. Solids, **13**, pp. 351–360.
[6] Marshall, E. A., 1967, "The Compression of a Slab of Ideal Soil Between Rough Plates," Acta Mech., **3**, pp. 82–92.
[7] Alexandrov, S., and Lyamina, E., 2003, "Plane-Strain Compression of Material Obeying the Double-Shearing Model Between Rotating Plates," Int. J. Mech. Sci., **45**(9), pp. 1505–1517.
[8] Alexandrov, S., and Lyamina, E., 2003, "Compression of a Mean-Stress Sensitive Plastic Material by Rotating Plates," Mech. Solids, **38**(6), pp. 40–48.
[9] Spencer, A. J. M., 2005, "Compression and Shear of a Layer of Granular Material," J. Eng. Math., **52**, pp. 251–264.
[10] Spencer, A. J. M., 1964, "A Theory of Kinematics of Ideal Soils Under Plane Strain Conditions," J. Mech. Phys. Solids, **12**, pp. 337–351.
[11] Oldroyd, J. G., 1956, "Non-Newtonian Flow of Liquids and Solids," *Rheology: Theory and Applications*, Vol. 1, F. R. Eirich, ed., Academic, New York, pp. 653–682.
[12] Adams, M. J., Briscoe, B. J., Corfield, G. M., Lawrence, C. J., and Papathanasiou, T. D., 1997, "An Analysis of the Plane-Strain Compression of Viscous Materials," ASME J. Appl. Mech., **64**, pp. 420–424.
[13] Cristescu, N., 1975, "Plastic Flow Through Conical Converging Dies, Using a Viscoplastic Constitutive Equation," Int. J. Mech. Sci., **17**, pp. 425–433.
[14] Durban, D., 1984, "Rate Effects in Steady Forming Processes of Plastic Materials," Int. J. Mech. Sci., **26**(4), pp. 293–304.
[15] Durban, D., 1986, "On Generalized Radial Flow Patterns of Viscoplastic Solids With Some Applications," Int. J. Mech. Sci., **28**(2), pp. 97–110.
[16] Akulenko, L. D., Georgievskii, D. V., Klimov, D. M., Kumakshev, S. A., and Nesterov, S. V., 2006, "Deformation of a Bingham Viscoplastic Fluid in a Plane Confuser," Int. Appl. Mech., **42**(4), pp. 375–406.
[17] Alexandrov, S., and Alexandrova, N., 2000, "On the Maximum Friction Law in Viscoplasticity," Mech. Time-Depend. Mater., **4**(1), pp. 99–104.
[18] Rabinovich, D., Givoli, D., and Vigdergauz, S., 2007, "Xfem-Based Crack Detection Scheme Using a General Algorithm," Int. J. Numer. Methods Eng., **71**, pp. 1051–1080.
[19] Oldroyd, J. G., 1943, "Two-Dimensional Plastic Flow of a Bingham Solid: A Plastic Boundary-Layer Theory for Slow Motion," Proc. Cambridge Philos. Soc., **43**, pp. 383–395.
[20] Smyrnaios, D. N., and Tsamopoulos, J. A., 2001, "Squeeze Flow of Bingham Plastics," J. Non-Newtonian Fluid Mech., **100**, pp. 165–190.
[21] Matsoukas, A., and Mitsoulis, E., 2003, "Geometry Effects in Squeeze Flow of Bingham Plastics," J. Non-Newtonian Fluid Mech., **109**, pp. 231–240.
[22] deVries, G., Craig, D. B., and Haddow, J. B., 1971, "Pseudo-Plastic Converging Flow," Int. J. Mech. Sci., **13**, pp. 763–772.
[23] Helsing, J., and Jonsson, A., 2002, "On the Accuracy of Benchmark Tables and Graphical Results in the Applied Mechanics Literature," ASME J. Appl. Mech., **69**(1), pp. 88–90.

# An Argument Against Augmenting the Lagrangean for Nonholonomic Systems[1]

**Carlos M. Roithmayr**[2]
NASA Langley Research Center,
Hampton, VA 23681
e-mail: c.m.roithmayr@larc.nasa.gov

**Dewey H. Hodges**
Professor
School of Aerospace Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: dhodges@gatech.edu

*Although it is known that correct dynamical equations of motion for a nonholonomic system cannot be obtained from a Lagrangean that has been augmented with a sum of the nonholonomic constraint equations weighted with multipliers, previous publications suggest otherwise. One published example that was proposed in support of augmentation purportedly demonstrates that an accepted method fails to produce correct equations of motion whereas augmentation leads to correct equations. This present paper shows that, in fact, the opposite is true. The correct equations, previously discounted on the basis of a flawed application of the Newton–Euler method, are verified by using Kane's method together with a new approach for determining the directions of constraint forces.* [DOI: 10.1115/1.3086435]

*Keywords: holonomic constraint equations, nonholonomic constraint equations, Lagrangean, Kane's method, constraint forces*

## 1 Introduction

Dealing with nonholonomic constraint equations within the framework of variational methods is a controversial subject. For example, Ray [1] modified Hamilton's principle and augmented the Lagrangean by adjoining a sum of nonholonomic constraint equations weighted with multipliers. Later, Ray [2] reversed himself. In the erratum he compared the correct way of dealing with constraint equations that are linear in the time derivatives of the generalized coordinates to the incorrect approach of augmenting the Lagrangean that gives the wrong results, even when the constraint equations are linear. Saletan and Cromer [3] followed Ray and showed that the augmented Lagrangean gives correct equations of motion when the constraint equations are holonomic. They concluded that no such augmented Lagrangean exists in the nonholonomic case, in part, because they said that there is no way to determine initial conditions needed for the integration of differential equations governing the multipliers. Rosenberg [4] presented the same demonstration as Ray's erratum and concluded that although Hamilton's principle may be regarded as a variational principle for conservative holonomic systems, it cannot be so regarded for nonholonomic systems. In an effort to eliminate constraint violations, Rosen and Edelstein [5] made a proposal similar to that of Ray 30 years earlier; they account for nonlinear nonholonomic constraint equations in the same way that they do holonomic constraint equations. Hagedorn [6] pointed out that although their approach is justified in the holonomic case, it is incorrect for nonholonomic constraint equations, even when they are linear. He demonstrated this with an example and gave the well-known result for the correct way to handle linear equations, which does not come from modifying the Lagrangean. According to Hagedorn the mistake has been repeated many times over the past century and the pitfall has received attention in Refs. [7–9]. More recently, Flannery [10] examined the problems encountered by Ray and others and, after in-depth analysis, concluded "General [nonlinear] nonholonomic constraints are completely outside the scope of even the most fundamental principle of D'Alembert. The generalization of any principle based on [D'Alembert's] to general nonholonomic constraints is without foundation."

Rosen and Edelstein [6] offered a counterexample purportedly showing that Hagedorn's approach leads to incorrect results. Their conclusion, however, is based on a flawed application of the Newton–Euler method. The objectives of the present paper are to demonstrate the validity of Hagedorn's approach and bring to light mistakes made by Rosen and Edelstein in the development of their counterexample.

## 2 Identification of Constraint Forces

Suppose that a mechanical system $S$ is made up of $\nu$ particles $P_1, \ldots, P_\nu$. In the event $S$ is subject to configuration constraints, its configuration in a Newtonian reference frame $N$ can be described by generalized coordinates $q_1, \ldots, q_n$, and the motion of $S$ is characterized by motion variables $u_1, \ldots, u_n$ (also called generalized speeds [11]). If, in addition, $S$ is a simple nonholonomic system [11], the motion constraints imposed on $S$ are described with nonintegrable relationships that are linear in $u_1, \ldots, u_n$. Suppose further that the configuration constraints and motion constraints can be described at the velocity level by a total of $m$ independent equations of the form

$$\sum_{i=1}^{\nu} {}^N\mathbf{v}^{P_i} \cdot \mathbf{W}_{is} + Y_s = 0 \quad (s = 1, \ldots, m) \tag{1}$$

where ${}^N\mathbf{v}^{P_i}$ is the velocity of $P_i$ in $N$, $\mathbf{W}_{is}$ are vector functions in $N$ of $q_1, \ldots, q_n$ and the time $t$, and $Y_s$ are scalar functions of the same variables. According to Ref. [12], one can inspect these relationships and conclude that constraint forces are given by

$$\mathbf{C}_{is} = \mu_s \mathbf{W}_{is} \quad (i = 1, \ldots, \nu; \ \ s = 1, \ldots, m) \tag{2}$$

where $\mu_s$ are scalar multipliers. The constraint force $\mathbf{C}_{is}$ is evidently parallel to $\mathbf{W}_{is}$ and in general it must be applied to $P_i$ in order to satisfy the constraint equation having the form of Eqs. (1). The technique of inspecting Eqs. (1) and writing Eqs. (2) is a way of systematically establishing the direction and point of application of each constraint force.

## 3 Problems With the Counterexample

The planar system featured in the counterexample proposed by Rosen and Edelstein [6] is shown in Fig. 1. Two perpendicular unit vectors $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are fixed in an inertial reference frame $N$. A smooth rod $B$ whose axis is parallel to unit vector $\hat{\mathbf{b}}_1$ is inclined at a constant angle $\alpha$ to $\hat{\mathbf{n}}_1$; $B$ is permitted to translate along an axis parallel to $\hat{\mathbf{n}}_2$. A particle $P$ of mass $m$ moves along $B$, and the mass of $B$ is negligible in comparison to $m$. It is said that no forces are exerted on $P$ other than those necessary to prevent it from losing contact with $B$.

Analysis is facilitated by working with two generalized coordinates $q_1$ and $q_2$ shown in Fig. 1, where $q_1$ is the displacement in a prismatic joint connecting $B$ to $N$, and where $q_2$ is the displacement of $P$ along the rod. Two motion variables are introduced

[1]This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States. Approved for public release; distribution is unlimited.

[2]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received June 28, 2007; final manuscript received December 29, 2008; published online March 9, 2009. Review conducted by N. Sri Namachchivaya.

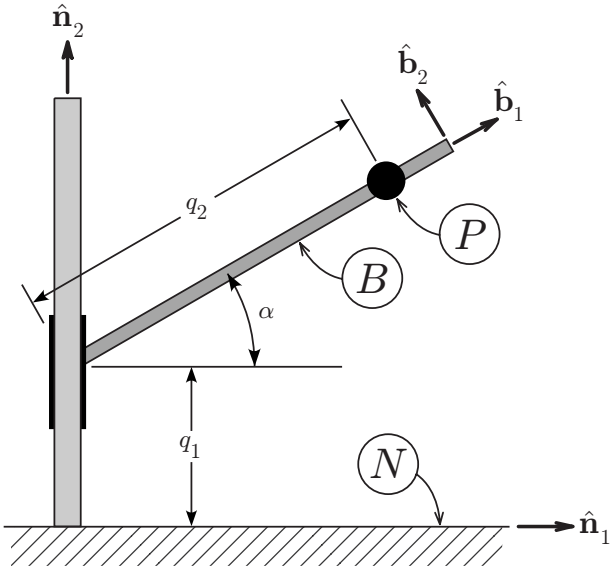



**Fig. 1 A particle moving on a sliding inclined rod**

simply as $u_r=\dot{q}_r$ $(r=1,2)$. A motion constraint is to be imposed upon the velocity of $P$ in $B$, expressed by the relationship

$$\varepsilon \cos \alpha u_2 - q_1 = 0 \tag{3}$$

where $\varepsilon$ is a positive constant. The equivalence of this expression and the constraint equation in Ref. [6] is demonstrated presently. Now, the velocity ${}^N\mathbf{v}^{\bar{B}}$ in $N$ of every point $\bar{B}$ fixed in $B$ is given by ${}^N\mathbf{v}^{\bar{B}}=u_1\hat{\mathbf{n}}_2$, and the velocity ${}^B\mathbf{v}^P$ of $P$ in $B$ is given by ${}^B\mathbf{v}^P=u_2\hat{\mathbf{b}}_1$. Henceforth, $\bar{B}$ is taken to be the point of $B$ that is coincident with $P$, and the velocity of $P$ in $N$ is simply ${}^N\mathbf{v}^P={}^N\mathbf{v}^{\bar{B}}+{}^B\mathbf{v}^P$. The nonholonomic constraint equation (3) can thus be written in vector form as

$$({}^N\mathbf{v}^P - {}^N\mathbf{v}^{\bar{B}}) \cdot \hat{\mathbf{b}}_1 - \frac{q_1}{\varepsilon \cos \alpha} = 0 \tag{4}$$

This relationship has the form of Eqs. (1). Let the particle $P$ play the part of $P_1$, and let the particle fixed in $B$ and coincident with $\bar{B}$ play the part of $P_2$. With $s=1$, the vector $\mathbf{W}_{11}$ is identified as $\hat{\mathbf{b}}_1$, whereas the vector $\mathbf{W}_{21}$ is $-\hat{\mathbf{b}}_1$. The scalar $Y_1$ is simply $-q_1/(\varepsilon \cos \alpha)$. Thus, with Eqs. (1) and (2) in mind, one may inspect Eq. (4) and conclude that $P$ must be subjected to a constraint force $\mathbf{C}_{11}$ that is parallel to $\hat{\mathbf{b}}_1$,

$$\mathbf{C}_{11} = \mu_1 \hat{\mathbf{b}}_1 \tag{5}$$

whereas a constraint force $\mathbf{C}_{21}=-\mathbf{C}_{11}$ is applied to $B$ at $\bar{B}$. The scalar multiplier $\mu_1$ is as yet unknown. In practice, the set of forces $\mathbf{C}_{11}$ and $-\mathbf{C}_{11}$ could be applied with a motorized gear attached to $P$ moving on a track of gear teeth fixed in $B$. Alternatively, friction could be exploited by using a capstan and pinch roller on opposite sides of $B$ in the way a similar mechanism is used to transport magnetic tape. Evidently the rod cannot be perfectly smooth as hypothesized in the problem statement, if the nonholonomic constraint equation (3) is to be satisfied.

There also exists a configuration constraint that prevents $P$ from moving in $B$ in the direction of $\hat{\mathbf{b}}_2$; at the velocity level, the holonomic constraint equation is expressed as

$$({}^N\mathbf{v}^P - {}^N\mathbf{v}^{\bar{B}}) \cdot \hat{\mathbf{b}}_2 = 0 \tag{6}$$

This relationship has the form of Eqs. (1). As before, $P$ and $\bar{B}$ play the roles of $P_1$ and $P_2$. With $s=2$, the vector $\mathbf{W}_{12}$ is identified as $\hat{\mathbf{b}}_2$, whereas the vector $\mathbf{W}_{22}$ is $-\hat{\mathbf{b}}_2$. The scalar $Y_2$ is 0. In view of Eqs. (1) and (2), satisfaction of the constraint requires the application of a constraint force

$$\mathbf{C}_{12} = \mu_2 \hat{\mathbf{b}}_2 \tag{7}$$

to $P$, and a constraint force $\mathbf{C}_{22}=-\mathbf{C}_{12}$ is applied to $B$ at $\bar{B}$. (In Ref. [6], the unknown multiplier $\mu_2$ is denoted by $N$.) This result is clearly in line with physical reasoning; in order to satisfy Eq. (6) and keep $P$ from losing contact with the rod, a constraint force of unknown magnitude and parallel to $\hat{\mathbf{b}}_2$ must be applied to $P$. In view of the law of action and reaction, a force of equal magnitude and opposite direction must be applied to $B$ at $\bar{B}$.

Now that the constraint forces $\mathbf{C}_{11}$ and $\mathbf{C}_{12}$ to be applied to $P$ have been identified, together with the constraint forces $\mathbf{C}_{21}$ and $\mathbf{C}_{22}$ applied to $B$ at $\bar{B}$, one is in a position to use Kane's method [11] to form two equations of motion $F_r+F_r^\star=0$ $(r=1,2)$ for the system $S$ composed of $P$ and $B$. The holonomic generalized active forces $F_r$ are given by

$$F_r = {}^N\mathbf{v}_r^P \cdot (\mathbf{C}_{11} + \mathbf{C}_{12}) + {}^N\mathbf{v}_r^{\bar{B}} \cdot (-\mathbf{C}_{11} - \mathbf{C}_{12}) \quad (r=1,2) \tag{8}$$

The holonomic generalized inertia forces $F_r^\star$ are constructed according to

$$F_r^\star = -{}^N\mathbf{v}_r^P \cdot m\, {}^N\mathbf{a}^P \quad (r=1,2) \tag{9}$$

One may inspect the relationships ${}^N\mathbf{v}^{\bar{B}}=u_1\hat{\mathbf{n}}_2$ and ${}^N\mathbf{v}^P=u_1\hat{\mathbf{n}}_2+u_2\hat{\mathbf{b}}_1$ to identify the holonomic partial velocities required by Eqs. (8) and (9).

$${}^N\mathbf{v}_1^P = \hat{\mathbf{n}}_2, \quad {}^N\mathbf{v}_2^P = \hat{\mathbf{b}}_1, \quad {}^N\mathbf{v}_1^{\bar{B}} = \hat{\mathbf{n}}_2, \quad {}^N\mathbf{v}_2^{\bar{B}} = \mathbf{0} \tag{10}$$

The acceleration of $P$ in $N$ is easily formed as ${}^N\mathbf{a}^P=\dot{u}_1\hat{\mathbf{n}}_2+\dot{u}_2\hat{\mathbf{b}}_1$. (The time derivative of $\hat{\mathbf{b}}_1$ in $N$ vanishes because $B$ has no angular velocity in $N$.) Hence, the dynamical equations of motion for $S$ in $N$ are found to be

$$m(\dot{u}_1 + \sin \alpha \dot{u}_2) = 0 \tag{11}$$

$$m(\sin \alpha \dot{u}_1 + \dot{u}_2) = \mu_1 \tag{12}$$

The nonholonomic constraint force $\mathbf{C}_{11}$ contributes to the holonomic generalized active forces, whereas the holonomic constraint force $\mathbf{C}_{12}$ does not. A third equation is needed to solve for the three unknowns $\dot{u}_1$, $\dot{u}_2$, and $\mu_1$; it is provided by the nonholonomic constraint equation (3) expressed at the acceleration level in scalar form as

$$\dot{u}_2 - \frac{u_1}{\varepsilon \cos \alpha} = 0 \tag{13}$$

Analytical solutions are then available:

$$\dot{u}_1 = -\frac{\tan \alpha}{\varepsilon} u_1 \tag{14}$$

$$\dot{u}_2 = \frac{u_1}{\varepsilon \cos \alpha} \tag{15}$$

$$\mu_1 = \frac{m \cos \alpha}{\varepsilon} u_1 \tag{16}$$

In the paragraph that precedes Eq. (16) [6], Rosen and Edelstein [6] stated that $P$ has zero acceleration along the path (the rod). It is worth noting that Eqs. (5), (15), and (16) together contradict this statement. If the proposed nonholonomic constraint equation is to be satisfied, $\dot{u}_2$ is in general nonzero; it vanishes only in the special case when the rod is stationary $(u_1=0)$.

The differential equations (14) and (15) yield closed-form solutions for $u_1$ and $u_2$; these can in turn be used to obtain closed-

form solutions for the generalized coordinates from the two kinematical differential equations $\dot{q}_r = u_r$ $(r=1,2)$. As shown in Ref. [12], these solutions verify the results attributed in Ref. [6] to Hagedorn's approach, namely, Eqs. (12) and (14a) in Sec. 3 of that work. Furthermore, the second-order differential equations (11a) and (11b) ascribed to Hagedorn's approach can be recovered from Eqs. (14)–(16) here. Such demonstrations are facilitated by establishing the following relationships between the Cartesian coordinates $x$ and $y$ and the generalized coordinates $q_1$ and $q_2$:

$$x = \cos \alpha q_2, \quad \dot{x} = \cos \alpha u_2, \quad \ddot{x} = \cos \alpha \dot{u}_2 \tag{17}$$

$$y = q_1 + \sin \alpha q_2, \quad \dot{y} = u_1 + \sin \alpha u_2, \quad \ddot{y} = \dot{u}_1 + \sin \alpha \dot{u}_2 \tag{18}$$

Moreover, appropriate substitution from these relationships shows that the original form of the nonholonomic constraint equation given in Ref. [6], $y - x \tan \alpha - \varepsilon \dot{x} = 0$, gives way to Eq. (3).

Rosen and Edelstein [6] rejected the aforementioned differential equations for $x$ and $y$, and the closed-form solutions, on the basis of their results obtained with the Newton–Euler method. With the analysis already performed here, it is evident that their application of the method is flawed, and the point in their development where the mistake was made can be identified immediately. In their Eqs. (2a) and (2b), they did not account for a constraint force $\mathbf{C}_{11}$ needed to ensure satisfaction of their nonholonomic constraint equation; they only considered $\mathbf{C}_{12}$ required to bring about the configuration constraint.

## 4 Conclusion

A straightforward application of Kane's method for simple nonholonomic systems, together with identification of the constraint forces needed to impose a motion constraint and a configuration constraint, are used to verify the results obtained with what was called the regular variational approach, brought to the reader's attention by Hagedorn in Ref. [6]. The conclusions reached by Hagedorn, Ray, Flannery, and others are thus affirmed; namely, the Lagrangean cannot be augmented by the sum of nonholonomic constraint equations weighted with multipliers, regardless of whether or not such equations are linear in the time derivatives of the generalized coordinates.

## References


[1] Ray, J. R., 1966, "Nonholonomic Constraints," Am. J. Phys., **34**(5), pp. 406–408.
[2] Ray, J. R., 1966, "Erratum: Nonholonomic Constraints," Am. J. Phys., **34**(12), pp. 1202–1203.
[3] Saletan, E. J., and Cromer, A. H., 1970, "A Variational Principle for Nonholonomic Systems," Am. J. Phys., **38**(7), pp. 892–897.
[4] Rosenberg, R. M., 1977, *Analytical Dynamics of Discrete Systems*, Plenum, New York, p. 220.
[5] Rosen, A., and Edelstein, E., 1997, "Investigation of a New Formulation of the Lagrange Method for Constrained Dynamic Systems," ASME J. Appl. Mech., **64**(1), pp. 116–122.
[6] Hagedorn, P., 1997, "Investigation of a New Formulation of the Lagrange Method for Constrained Dynamic Systems," ASME J. Appl. Mech., **64**(4), pp. 1024–1027.
[7] Neĭmark, J. I., and Fufaev, N. A., 1972, *Dynamics of Nonholonomic Systems*, American Mathematical Society, Providence, RI.
[8] Pars, L. A., 1979, *A Treatise on Analytical Dynamics*, Ox Bow, Connecticut.
[9] Whittaker, E. T., 1964, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed., Cambridge University Press, London.
[10] Flannery, M. R., 2005, "The Enigma of Nonholonomic Constraints," Am. J. Phys., **73**(3), pp. 265–272.
[11] Kane, T. R., and Levinson, D. A., 1985, *Dynamics: Theory and Applications*, McGraw-Hill, New York, Chaps. 2 and 4.
[12] Roithmayr, C. M., 2007, "Relating Constrained Motion to Force Through Newton's Second Law," Ph.D. thesis, Georgia Institute of Technology, Atlanta.

# Analysis of Non-Newtonian Reactive Flow in a Cylindrical Pipe

**Oluwole Daniel Makinde**
Faculty of Engineering,
Cape Peninsula University of Technology,
P.O. Box 652,
Cape Town 8000, South Africa
e-mail: makinded@cput.ac.za

*In this paper, a mathematical investigation on the effect of convective cooling on a reactive third-grade fluid flowing steadily through a cylindrical pipe is performed. It is assumed that the system exchange heat, with the ambient following Newton's cooling law and the reaction, is exothermic under Arrhenius kinetics, neglecting the consumption of the material. The simplified governing nonlinear equations of momentum and energy are obtained and solved using a special type of the Hermite–Padé approximation technique. The important properties of the overall flow structure including velocity field, temperature field, bifurcations, and thermal criticality conditions are discussed.*
[DOI: 10.1115/1.3086587]

*Keywords: cylindrical pipe, third-grade fluid, Arrhenius kinetics, thermal criticality, Hermite–Padé approximants, convective cooling*

## 1 Introduction

It is well known that the rheological properties of many fluids used in engineering and industries are not well modeled by the Navier–Stokes equations [1,2]. In recent years, there have been several studies [3–5] on the mechanics of non-Newtonian fluids, not only because of their technological significance but also in view of the interesting mathematical features presented by the nonlinear equations governing the flow. Furthermore, a large class of non-Newtonian fluids used in industries is chemically reactive, e.g., coal slurries, polymer solutions or melts, drilling mud, hydrocarbon oils, and grease. For instance, in modeling the non-Newtonian flow situations, such as coal-based slurries as retrofit fuels, the power-law model was used widely to characterize the rheological properties of the fluid [2]. Although the power-law model adequately fit the shear stress and shear rate measurements for many non-Newtonian fluids, it could not always be used to predict accurately the pressure loss data measured during the transport of a coal-liquid mixture in a fuel delivery system. Moreover, the power-law model could not predict correctly the normal stress effects that lead to phenomena such as road climbing, in which case the stresses are developed orthogonal to planes of shear. Consequently, a third-grade fluid model is fruitful and appropriate for this situation [6]. Meanwhile, the study of heat transfer and thermal criticality of reactive non-Newtonian fluids is extremely important in order to ensure safety of life and properties during handling and processing of such fluids [7–9]. Massoudi and Christie [10] studied the flow of a variable viscosity third-grade fluid and heat transfer in a pipe. They showed numerically that increasing non-Newtonian parameter lowered the temperature and velocity of the fluid in the pipe. Szeri and Rajagopal [11] studied the flow of non-Newtonian fluid between two heated horizontal parallel plates. They employed the third-grade fluid model and introduced temperature dependent viscosity.

The task of this present work is to study the effect of convective cooling on a non-Newtonian reactive flow in a cylindrical pipe using a special type of Hermite–Padé approximants. This is basically an extension of the recent work of Makinde [4] to include the effect of convective heat exchange with the ambient at the pipe surface. The mathematical formulation of the problem is established and solved in Secs. 2 and 3. In Sec. 4, we introduce and apply some rudiments of the Hermite–Padé approximation technique. Both numerical and graphical results are presented and discussed quantitatively with respect to various parameters embedded in the system in Sec. 5.

## 2 Analysis

The geometry of the problem is as shown in Fig. 1 below. The non-Newtonian reactive steady flow is induced by applied axial pressure gradient with convective cooling at the pipe surface. A cylindrical coordinate system $(r,z)$ is chosen such that $0z$ lies along the axis of pipe; $r$ is the distance measured radially with $r=a$ as the pipe radius. Following Refs. [3,5,10] and neglecting the reacting viscous fluid consumption, the governing equations for the momentum and heat balance can be written as

$$\frac{\mu}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right)+\frac{\beta_3}{r}\frac{d}{dr}\left(r\left(\frac{du}{dr}\right)^3\right)=\frac{dP}{dz} \tag{1}$$

$$\frac{k}{r}\frac{d}{dr}\left(r\frac{dT}{dr}\right)+\left(\frac{du}{dr}\right)^2\left(\mu+\beta_3\left(\frac{du}{dr}\right)^2\right)+QC_0Ae^{-E/RT}=0 \tag{2}$$

The appropriate boundary conditions are given as follows: the pipe surface is fixed, impermeable, and exchange heat with the ambient following Newton's cooling law,

$$u=0,\quad k\frac{dT}{dr}=-h(T-T_a)\quad \text{at } r=a \tag{3}$$

and the axisymmetric conditions along the centerline, i.e.,

$$\frac{du}{dr}=\frac{dT}{dr}=0\quad \text{at } r=0 \tag{4}$$

where the additional Arrhenius kinetics term in energy balance equation (2) is due to Refs. [7,12]. Here $T$ is the fluid temperature, $u$ is the fluid velocity, $U$ is the fluid characteristic velocity, $h$ is the heat transfer coefficient, $T_a$ is the ambient temperature, $k$ is the thermal conductivity of the material, $Q$ is the heat of reaction, $A$ is the rate constant, $E$ is the activation energy, $R$ is the universal gas constant, $C_0$ is the initial concentration of the reactant species, $\beta_3$ is the material coefficient, $P$ is the modified pressure, and $\mu$ is the fluid dynamic viscosity coefficient [7,8,11]. We introduce the following dimensionless variables into Eqs. (1)–(3):

$$\theta=\frac{E(T-T_a)}{RT_a^2},\quad \bar{r}=\frac{r}{a},\quad \lambda=\frac{QEAa^2C_0e^{-E/RT_a}}{T_a^2Rk},$$

$$W=\frac{u}{UG},\quad \mathrm{Bi}=\frac{ha}{k}$$

$$m=\frac{\mu G^2U^2e^{E/RT_a}}{QAa^2C_0},\quad \varepsilon=\frac{RT_a}{E},\quad G=-\frac{a^2}{\mu U}\frac{dP}{dz},\quad \gamma=\frac{\beta_3U^2G^2}{a^2\mu} \tag{5}$$

and obtain the dimensionless governing equations together with their corresponding boundary conditions as (neglecting the bar symbol for clarity)

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dW}{dr}\right)+\frac{\gamma}{r}\frac{d}{dr}\left(r\left(\frac{dW}{dr}\right)^3\right)=-1 \tag{6}$$

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 17, 2008; final manuscript received January 27, 2009; published online March 13, 2009. Review conducted by Nesreen Ghaddar.

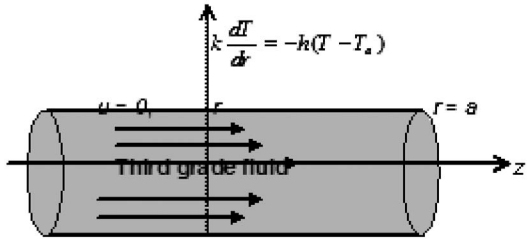



**Fig. 1 Geometry of the problem**

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\theta}{dr}\right) + \lambda\left[e^{(\theta/(1+\varepsilon\theta))} + m\left(\frac{dW}{dr}\right)^2\left(1 + \gamma\left(\frac{dW}{dr}\right)^2\right)\right] = 0 \tag{7}$$

with

$$W = 0, \quad \frac{d\theta}{dr} = -\mathrm{Bi}\,\theta \quad \text{at } r = 1 \tag{8}$$

$$\frac{dW}{dr} = \frac{d\theta}{dr} = 0 \quad \text{at } r = 0 \tag{9}$$

where $\lambda$, $\varepsilon$, $\gamma$, $m$, and Bi represent the Frank-Kamenetskii parameter, the activation energy parameter, the dimensionless non-Newtonian parameter, the viscous heating parameter, and the Biot number, respectively. In Sec. 3, Eqs. (6)–(9) are solved using both perturbation and multivariate series summation techniques [4,13,14].

## 3 Perturbation Method

Equations (6)–(9) are nonlinear, and it is convenient to form a power series expansion both in the dimensionless non-Newtonian parameter $\gamma$ and the Frank-Kamenetskii parameter $\lambda$, i.e.,

$$W = \sum_{i=0}^{\infty} W_i \gamma^j, \quad \theta = \sum_{i=0}^{\infty} \theta_i \lambda^i \tag{10}$$

Substituting the solution series in Eq. (10) into Eqs. (6)–(9) and collecting the coefficients of like powers of $\gamma$ and $\lambda$, we obtained and solved the equations for the coefficients of solution series iteratively. Using a computer symbolic algebra package (MAPLE), the first few terms of the above solution series in Eq. (10) can be easily obtained (see the Appendix). The physical quantities of interest in this problem are the skin-friction parameter ($C_f$) and the Nusselt number (Nu), which are defined by

$$C_f = \frac{a\iota_w}{\mu UG} = \frac{dW}{dr}(1) \tag{11a}$$

$$\mathrm{Nu} = \frac{aEq_w}{kRT_a^2} = -\frac{d\theta}{dr}(1) \tag{11b}$$

where $\iota_w = \mu du/dr$ and $q_w = -kdT/dr$ are the shear stress and the heat flux evaluated at the wall (i.e., $r=a$), respectively. We are aware that the power series solutions are valid for very small parameter values. However, using the Hermite–Padé approximation technique, we have extended the usability of the solution series beyond small parameter values, as illustrated in Sec. 4.

## 4 Thermal Criticality and Bifurcation Study

The determination of thermal criticality in a flow system is extremely important from the application point of view. Thermal criticality occurs when the rate of heat generation within the flow system exceeds the heat dissipation to the surroundings [4,11,13,14]. This condition is incipient thermal runaway or ignition in the flow system [3,12]. Thermal criticality is characterized by the onset of thermal instability in the system and the nonexistence of a steady-state solution to nonlinear problems for certain parameter values [6]. A primary objective of thermal criticality analysis is the prediction of the critical or unsafe flow conditions in order to avoid them [8]. The main tool employed in this section is a simple technique of series summation based on the generalization of the Hermite–Padé approximants and may be described as follows. Let us suppose that the partial sum

$$U_N(\lambda) = \sum_{n=0}^{N} a_n \lambda^n + O(\lambda^{N+1}) \quad \text{as } \lambda \to 0 \tag{12}$$

is given. It is important to note here that Eq. (12) can be used to approximate any output of the solution of the problem under investigation (e.g., the series for the wall heat flux parameter in terms of Nusselt number $\mathrm{Nu} = -d\theta/dr$ at $r=1$) since everything can be Taylor expanded in the given small parameter. Assuming that $U_N(\lambda)$ in Eq. (12) is a local representation of an algebraic function of $\lambda$ in the context of nonlinear problems, we seek a polynomial

$$F_d(\lambda, U_N) = \sum_{m=1}^{d}\sum_{k=0}^{m} f_{m-k,k}\lambda^{m-k}U_N^k \tag{13}$$

of degree $d \geq 2$, such that

$$\frac{\partial F_d}{\partial U_N}(0,0) = 1 \tag{14}$$

and

$$F_d(\lambda, U_N) = O(\lambda^{N+1}) \quad \text{as } \lambda \to 0 \tag{15}$$

Condition (13), which yields $f_{0,1} = 1$, ensures that the polynomial $F_d$ has only one root, which vanishes at $\lambda = 0$ and also normalizes $F_d$. The requirement (Eq. (14)) reduces the problem to a system of $N$ linear equations for the unknown coefficients of $F_d$. The entries of the underlying matrix depend on the $N$ given coefficients of $U_N$. Henceforth we shall take

$$N = \tfrac{1}{2}(d^2 + 3d + 2) \tag{16}$$

so that the number of equations equals the number of unknowns. The polynomial $F_d$ is a special type of Hermite–Padé approximant and is then investigated for bifurcation and criticality conditions ($\lambda_c$) numerically using the Newton diagram [15]. The critical value ($\lambda_c$) is obtained as the nearest singularity in the system and is tabulated in Sec. 5.

**Table 1 Computations showing the procedure rapid convergence and bifurcation point in the velocity field**

| $d$ | $N$ | $C_f$ | $\gamma_c$ |
|---|---|---|---|
| 1 | 9 | −0.750095 | −0.592617 |
| 2 | 12 | −0.750000 | −0.592592 |
| 3 | 15 | −0.750000 | −0.592592 |
| 4 | 18 | −0.750000 | −0.592592 |

## 5 Results and Discussion

The bifurcation procedure in Sec. 4 was applied to the first few terms of the solution series, and we obtained the results as shown in Tables 1 and 2 and Figs. 2–6. Table 1 demonstrates the rapid convergence of our procedure with respect to the non-Newtonian parameter ($\gamma_c$) dominant singularity and its critical exponent ($b_c$) for a third-grade fluid flowing steadily in a cylindrical pipe. A bifurcation point (i.e., a turning point) occurs in the flow field at $(\gamma_c, C_f) = (-16/27, -3/4)$, as shown in Fig. 6. It is noteworthy that the convergence of our procedure improves with gradual increase in the number of series coefficients utilized in the approximants.

**Table 2 Computations showing thermal criticality conditions for different parameter values**

| Bi | $m$ | $\gamma$ | Nu ($\varepsilon=0$) | $\lambda_c$ ($e=0$) | Nu ($\varepsilon=0.1$) | $\lambda_c$ ($\varepsilon=0.1$) |
|---|---|---|---|---|---|---|
| 0.1 | 1.0 | 0.1 | 0.101721745 | 0.06874448290 | 0.129969293 | 0.077293512 |
| 1.0 | 1.0 | 0.1 | 0.797386716 | 0.55307475391 | 1.021737269 | 0.622721247 |
| 10.0 | 1.0 | 0.1 | 1.892518128 | 1.60254039226 | 2.431123409 | 1.817184246 |
| 100 | 1.0 | 0.1 | 2.086597414 | 1.90674915118 | 2.659917357 | 2.163012837 |
| $\infty$ | 1.0 | 0.1 | 4.217362228 | 1.94607136100 | 5.369541143 | 2.20747552 |
| $\infty$ | 1.0 | 0.2 | 4.212964486 | 1.94675509035 | 5.364098368 | 2.20816105 |
| $\infty$ | 1.0 | 0.3 | 4.208225870 | 1.94748243186 | 5.358240483 | 2.20889033 |

Table 2 shows the thermal criticality conditions ($\lambda_c$) for a reactive third-grade liquid with respect to pipe flow. The magnitude of thermal criticality increases, while the rate of heat transfer across the wall decreases with an increase in the value of the Biot number and the non-Newtonian parameter ($\gamma$). This implies that increasing values of the non-Newtonian parameter and the intensity of convective cooling enhance the thermal stability of a reactive third-grade liquid. The results for the case of $\mathrm{Bi}\to\infty$ also agreed perfectly well with the one reported in Ref. [4]. The fluid velocity profile is parabolic, as shown in Fig. 2; however, a gradual decrease in the magnitude of fluid velocity is noticed with an increase in value of the non-Newtonian parameter. Figure 3 shows a transverse increase in the fluid temperature with the maximum temperature along the pipe centerline. An increase in the value of the Frank-Kamenetskii parameter ($\lambda$) due to Arrhenius kinetics causes a further increase in the fluid temperature, while a decrease in fluid temperature is observed with an increase in the convective cooling, as shown in Fig. 4. A slice of the bifurcation diagram for $\mathrm{Bi}>0$ in the ($\lambda$,Nu) plane is shown in Fig. 5. It represents the variation in wall heat flux (Nu) with the Frank-Kamenetskii parameter ($\lambda$). In particular, for every $0\le\varepsilon\le0.1$ there is a critical value $\lambda_c$ (a turning point) such that for $0\le\lambda<\lambda_c$ there are two solution branches (labeled I and II). The upper and lower solution branches occur due to Arrhenius kinetics in the governing thermal boundary layer equation (Eq. (2)). The system has no real solution for $\lambda_c<\lambda$ and displays a classical form, indicating thermal runaway. The magnitude of $\lambda_c$ increases with a decrease in the fluid activation energy ($\varepsilon=0.1$), hence preventing the early development of thermal runaway and enhancing thermal stability.

## 6 Conclusion

A novel hybrid numerical-analytical scheme based on a special type of Hermite–Padé approximants is utilized to investigate the effect of convective cooling on the thermal criticality of a reactive third-grade fluid flowing steadily through a cylindrical pipe. The

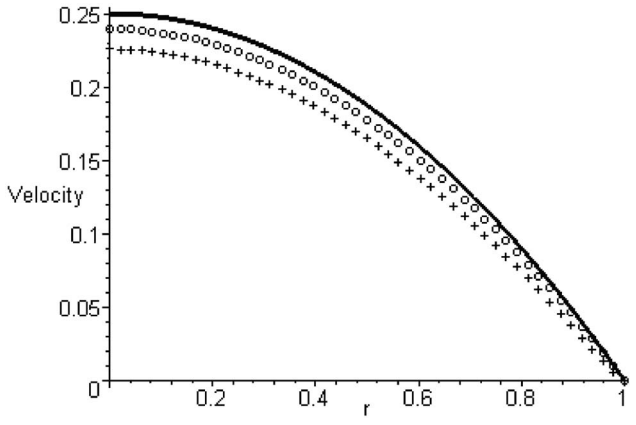



**Fig. 2 Velocity profile: _____, $\gamma=0$; oooooo, $\gamma=0.4$; +++++ +, $\gamma=0.8$**

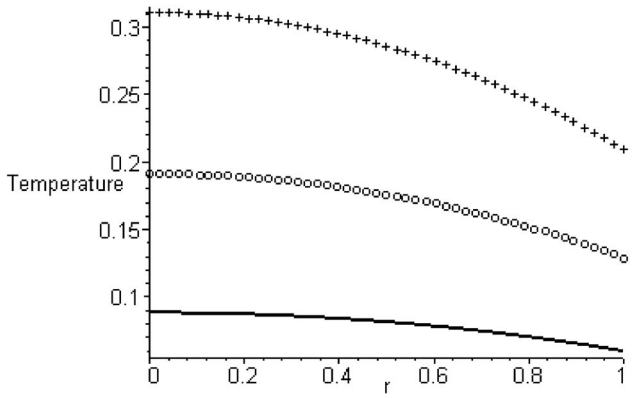



**Fig. 3 Temperature profile for Bi=1; $m$=1; $\varepsilon=0$; $\gamma=0.1$; _____, $\lambda=0.1$; ooooo, $\lambda=0.2$; +++++, $\lambda=0.3$**

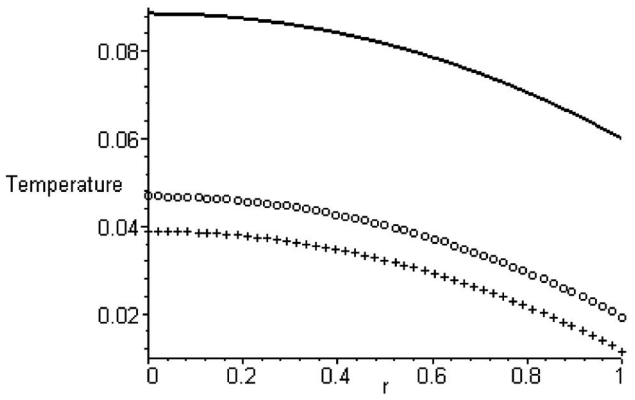



**Fig. 4 Temperature profile for $\lambda=0.1$; $m$=1; $\varepsilon=0$; $\gamma=0.1$; _____, Bi=1; ooooo, Bi=3; +++++, Bi=5**

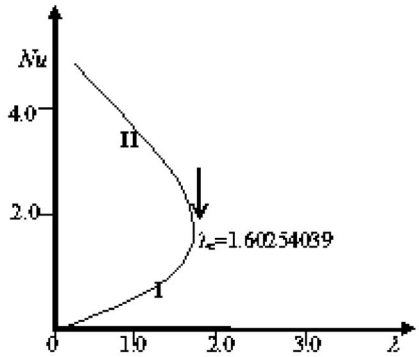



**Fig. 5 A slice of approximate bifurcation diagram in the ($\lambda$,Nu) plane**

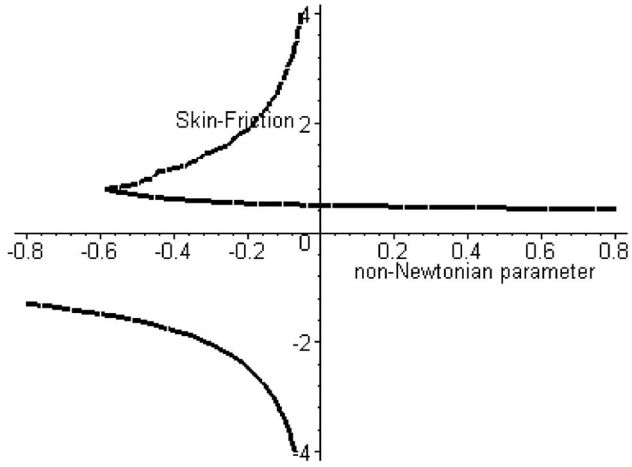



**Fig. 6 A bifurcation diagram in the $(\gamma, C_f)$ plane for a third-grade liquid**

procedure reveals accurately the thermal criticality conditions and various solution branches. It is observed that a combined increase in convective cooling parameter (Biot number) and non-Newtonian parameter enhances the thermal stability of a reactive third-grade fluid.


## Acknowledgment

The author would like to thank the National Research Foundation of South Africa Thuthuka program for financial support.


## Nomenclature

$T$ = fluid temperature
$U$ = characteristic velocity
$T_a$ = ambient temperature
$Q$ = heat of reaction
$E$ = activation energy
$C_0$ = reactant initial concentration
$r$ = radial coordinate
Bi = Biot number
$W$ = dimensionless velocity
$u$ = fluid velocity
$h$ = heat transfer coefficient
$k$ = thermal conductivity coefficient
$A$ = reaction rate constant
$R$ = universal gas constant
$P$ = fluid pressure
$z$ = axial coordinate
$m$ = viscous heating parameter
$G$ = pressure gradient parameter

### Greek Symbols

$\mu$ = dynamic viscosity coefficient
$\beta_3$ = material coefficient
$\lambda$ = Frank-Kamenetskii parameter
$\varepsilon$ = activation energy parameter
$\gamma$ = non-Newtonian parameter
$\theta$ = dimensionless temperature

## Appendix

This appendix gives the truncated series solutions to Eqs. (6)–(9) obtained using MAPLE,

$$W(r) = \frac{1}{4} - \frac{1}{4}r^2 + \left(\frac{1}{32}r^4 - \frac{1}{32}\right)\gamma + \left(-\frac{1}{64}r^6 + \frac{1}{64}\right)\gamma^2 + \left(\frac{3}{256}r^8 - \frac{3}{256}\right)\gamma^3 + \left(-\frac{11}{1024}r^{10} + \frac{11}{1024}\right)\gamma^4 + \left(\frac{91}{8192}r^{12} + \frac{91}{8192}\right)\gamma^5 + O(\gamma^6)$$

$$\begin{aligned}\theta(r) = &-\frac{1}{180,633,600}((225m\gamma^5 r^{14}\mathrm{Bi} - 4900m\gamma^4 r^{12}\mathrm{Bi} \\ &+ 42,336m\gamma^3 r^{10}\mathrm{Bi} - 132,300m\gamma^2 r^8\mathrm{Bi} - 313,600m\gamma r^6\mathrm{Bi} \\ &+ 2,822,400mr^4\mathrm{Bi} + 45,158,400r^2\mathrm{Bi} - 3150m\gamma^5 \\ &+ 58,800m\gamma^4 - 423,360m\gamma^3 + 1,058,400m\gamma^2 \\ &+ 1,881,600m\gamma - 11,289,600m - 90,316,800 - 225\mathrm{Bi}m\gamma^5 \\ &+ 4900\mathrm{Bi}m\gamma^4 - 42,336\mathrm{Bi}m\gamma^3 + 132,300\mathrm{Bi}m\gamma^2 \\ &+ 313,600\mathrm{Bi}m\gamma - 2,822,400\mathrm{Bi}m - 45,158,400\mathrm{Bi})\lambda)/\mathrm{Bi} \\ &+ \frac{1}{46,242,201,600}(-4,300,800\mathrm{Bi}m\gamma^4 + 8,670,412,800\mathrm{Bi} \\ &- 2,709,504r^2\mathrm{Bi}^2m\gamma^3 + 313,600r^2\mathrm{Bi}^2m\gamma^4 \\ &- 5,780,275,200r^2\mathrm{Bi} - 18,816,000\mathrm{Bi}^2m\gamma \\ &+ 8,467,200r^2\mathrm{Bi}^2m\gamma^2 + 225m\gamma^5 r^{16}\mathrm{Bi}^2 + 75,264m\gamma^3 r^{12}\mathrm{Bi}^2 \\ &- 6400m\gamma^4 r^{14}\mathrm{Bi}^2 - 338,688m\gamma^2 r^{10}\mathrm{Bi}^2 - 1,254,400m\gamma r^8\mathrm{Bi}^2 \\ &- 180,633,600r^2\mathrm{Bi}^2m + 31,610,880\mathrm{Bi}m\gamma^3 \\ &+ 963,379,200\mathrm{Bi}m - 240,844,800m\gamma - 135,475,200m\gamma^2 \\ &+ 54,190,080m\gamma^3 - 7,526,400m\gamma^4 + 403,200m\gamma^5 \\ &- 150,528,000\mathrm{Bi}m\gamma + 14,175\mathrm{Bi}^2m\gamma^5 - 307,200\mathrm{Bi}^2m\gamma^4 \\ &+ 20,070,400r^2\mathrm{Bi}^2m\gamma + 226,800\mathrm{Bi}m\gamma^5 + 11,560,550,400 \\ &+ 2,167,603,200\mathrm{Bi}^2 + 160,563,200\mathrm{Bi}^2m \\ &- 2,890,137,600r^2\mathrm{Bi}^2 + 722,534,400r^4\mathrm{Bi}^2 \\ &- 8,128,512\mathrm{Bi}^2m\gamma^2 + 2,634,240\mathrm{Bi}^2m\gamma^3 - 81,285,120\mathrm{Bi}m\gamma^2 \\ &+ 20,070,400mr^6\mathrm{Bi}^2 - 722,534,400r^2\mathrm{Bi}m \\ &+ 1,445,068,800m + 3,763,200r^2\mathrm{Bi}m\gamma^4 \\ &- 27,095,040r^2\mathrm{Bi}m\gamma^3 + 67,737,600r^2\mathrm{Bi}m\gamma^2 \\ &+ 120,422,400r^2\mathrm{Bi}m\gamma - 14,400r^2\mathrm{Bi}^2m\gamma^5 \\ &- 201,600r^2\mathrm{Bi}m\gamma^5)\lambda^2/\mathrm{Bi}^2 + O(\lambda)^3\end{aligned}$$

## References


[1] Beard, D. W., and Walters, K., 1964, "Elastico-Viscous Boundary-Layer Flows. I. Two-Dimensional Flow Near a Stagnation Point," Proc. Cambridge Philos. Soc., **60**, pp. 667–674.
[2] Schowalter, W. R., 1978, *Mechanics of Non-Newtonian Fluids*, Pergamon, Oxford.
[3] Hecht, A. M., 1973, "Theoretical Non-Newtonian Pipe-Flow Heat Transfer," AIChE J., **19**, pp. 197–199.
[4] Makinde, O. D., 2007, "Thermal Stability of a Reactive Third Grade Fluid in a Cylindrical Pipe: An Exploitation of Hermite–Padé Approximation Technique," Appl. Math. Comput., **189**, pp. 690–697.
[5] Yurusoy, M., and Pakdemirli, M., 2002, "Approximate Analytical Solutions for the Flow of a Third Grade Fluid in a Pipe," Int. J. Non-Linear Mech., **37**, pp. 187–195.
[6] Fosdick, R. L., Rajagopal, K. R., 1980, "Thermodynamics and Stability of Fluids of Third Grade," Proc. R. Soc. London, Ser. A, **369**(1738), pp. 351–377.
[7] Bebernes, J., and Eberly, D., 1989, *Mathematical Problems From Combustion Theory*, Springer-Verlag, New York.

[8] Bowes, P. C., 1984, *Self-Heating: Evaluating and Controlling the Hazard*, Elsevier, Amsterdam.
[9] Demirel, Y., and Kahraman, R., 2000, "Thermodynamic Analysis of Convective Heat Transfer in an Annular Packed Bed," Int. J. Heat Fluid Flow, **21**, pp. 442–448.
[10] Massoudi, M., and Christie, I., 1995, "Effects of Variable Viscosity and Viscous Dissipation on the Flow of a Third Grade Fluid in a Pipe," Int. J. Non-Linear Mech., **30**, pp. 687–699.
[11] Szeri, A. Z., and Rajagopal, K. R., 1985, "Flow of a Non-Newtonian Fluid Between Heated Parallel Plates," Int. J. Non-Linear Mech., **20**, pp. 91–101.
[12] Frank-Kamenetskii, D. A., 1969, *Diffusion and Heat Transfer in Chemical Kinetics*, Plenum, New York.
[13] Makinde, O. D., 2005, "Strong Exothermic Explosions in a Cylindrical Pipe: A Case Study of Series Summation Technique," Mech. Res. Commun., **32**, pp. 191–195.
[14] Makinde, O. D., 2006, "Thermal Ignition in a Reactive Viscous Flow Through a Channel Filled With a Porous Medium," ASME J. Heat Transfer, **128**, pp. 601–604.
[15] Vainberg, M. M., and Trenogin, V. A., 1974, *Theory of Branching of Solutions of Nonlinear Equations*, Noordoff, Leyden.

# Effective Properties of Carbon Nanotube and Piezoelectric Fiber Reinforced Hybrid Smart Composites

**M. C. Ray**

**R. C. Batra**

Department of Engineering Science and Mechanics,
Virginia Polytechnic Institute and State University,
MC 0219,
Blacksburg, VA 24061

*We propose a new hybrid piezoelectric composite comprised of armchair single-walled carbon nanotubes and piezoelectric fibers as reinforcements embedded in a conventional polymer matrix. Effective piezoelectric and elastic properties of this composite have been determined by a micromechanical analysis. Values of the effective piezoelectric coefficient $e_{31}$ of this composite that accounts for the in-plane actuation and of effective elastic properties are found to be significantly higher than those of the existing 1–3 piezoelectric composites without reinforced with carbon nanotubes.* [DOI: 10.1115/1.3063633]

## 1 Introduction

The discovery of carbon nanotubes (CNTs) [1] has stimulated extensive research devoted to the prediction of their elastic properties through experiments and theoretical modeling. Treacy et al. [2] experimentally determined that CNTs have Young's modulus in the terapascal range. Li and Chou [3] linked structural and molecular mechanics (MM) approaches to compute elastic properties of CNTs. Sears and Batra [4] used three MM potentials to simulate axial and torsional deformations of a CNT assuming that the tube can be regarded as a hollow cylinder of mean diameter equal to that of the CNT. They found the wall thickness, Young's modulus, and Poisson's ratio of the CNT. Shen and Li [5] assumed that a CNT should be modeled as a transversely isotropic material with the axis of transverse isotropy coincident with the centroidal axis of the tube. They determined values of the five elastic constants by using a MM potential and an energy equivalence principle. Batra and Sears [6] proposed that the axis of transverse isotropy of a CNT is a radial line rather than the centroidal axis of the tube and found that Young's modulus in the radial direction equals about 1/4 of that in the axial direction. Batra and Gupta [7,8] determined the wall thickness and material moduli of a CNT based on the frequencies of axial, torsional, and radial breathing modes. Wu et al. [9] developed an atomistic based finite deformation shell theory for single-walled CNT and found its stiffness in tension, bending, and torsion. A great deal of research has also been carried out on the prediction of effective elastic properties of CNT-reinforced composites [10–12].

Piezoelectric composites, often called piezocomposites, have been used as distributed actuators and sensors. Piezocomposites (PZCs), usually comprised of an epoxy reinforced with a monolithic piezoelectric material (PZT), provide a wide range of effective material properties not offered by existing PZTs, are anisotropic, and are characterized by good conformability and strength. One of the commercially available PZCs is the lamina of vertically reinforced 1–3 PZCs [13] and is being effectively used as underwater and high frequency ultrasonic transducers, and in medical imaging devices. In a 1–3 PZC lamina the poling direction of PZT fibers is along the laminate thickness, and the top and the bottom surfaces of the lamina are electroded. Smith and Auld [14] used the micromechanical isostrain/isostress technique to determine the effective moduli of a PZC and found that the magnitude of the effective piezoelectric coefficient $e_{33}$ is much larger than that of the effective piezoelectric coefficient $e_{31}$. Note that $e_{33}$ determines the magnitude of the induced actuating stress along the fiber direction due to a unit electric field applied across the thickness of the PZC lamina while $e_{31}$ gives the induced stress in the direction transverse to the fiber. Hence, the in-plane actuation of this PZC is negligible as compared with its out-of-plane actuation [15]. The control of bending deformations of a smart beam is generally attributed to the in-plane actuation induced by a PZT actuator. The in-plane actuation caused by the PZC can be enhanced by tailoring its effective piezoelectric coefficient $e_{31}$. Smith and Auld's [14] work also reveals that the magnitude of effective $e_{31}$ can be increased by improving upon elastic properties of the matrix. Since CNT reinforcements noticeably strengthen the polymer matrix, the matrix can also be reinforced with CNT and PZT fibers to form a new hybrid PZC with improved effective piezoelectric coefficient $e_{31}$. Here, we find values of effective moduli of a hybrid PZC that we call nanotube reinforced hybrid piezoelectric composite (NRHPC) by using a micromechanics approach proposed by Smith and Auld [14], and Benveniste and Dvorak [16].

## 2 Effective Moduli of a NRHPC

Figure 1 shows a schematic sketch of a lamina of NRHPC with CNT and PZT fibers aligned vertically. The cross section of CNT and PZT fibers is shown as square for simplicity since it does not enter into calculations. The analysis applies to straight prismatic fibers with parallel centroidal axes and fibers perpendicular to the lamina. The CNT fibers are assumed to be transversely isotropic with the axis of transverse isotropy along the centroidal axis, and the PZT fibers are poled in the thickness direction. The representative volume element considered for the micromechanics analysis is comprised of a CNT fiber and a PZT fiber surrounded by a polymer matrix of the same volume fraction as that in the composite. Using rectangular Cartesian coordinate axes exhibited in Fig. 1, constitutive equations for the PZT, the CNT, and the polymer matrix material are

$$\{\sigma^p\} = [C^p]\{\varepsilon^p\} - \{e^p\}E_z, \quad \{\sigma^n\} = [C^n]\{\varepsilon^n\}, \quad \text{and} \quad \{\sigma^m\} = [C^m]\{\varepsilon^m\} \tag{1}$$

where

$$\{\sigma^r\} = [\sigma^r_x \;\; \sigma^r_y \;\; \sigma^r_z \;\; \sigma^r_{yz} \;\; \sigma^r_{xz} \;\; \sigma^r_{xy}]^T,$$

$$\{\varepsilon^r\} = [\varepsilon^r_x \;\; \varepsilon^r_y \;\; \varepsilon^r_z \;\; \varepsilon^r_{yz} \;\; \varepsilon^r_{xz} \;\; \varepsilon^r_{xy}]^T,$$

$$[C^r] = \begin{bmatrix} C^r_{11} & C^r_{12} & C^r_{13} & 0 & 0 & 0 \\ C^r_{12} & C^r_{22} & C^r_{23} & 0 & 0 & 0 \\ C^r_{13} & C^r_{23} & C^r_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C^r_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C^r_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C^r_{66} \end{bmatrix},$$

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 29, 2008; final manuscript received July 24, 2008; published online March 13, 2009 Review conducted by Yonggang Huang.

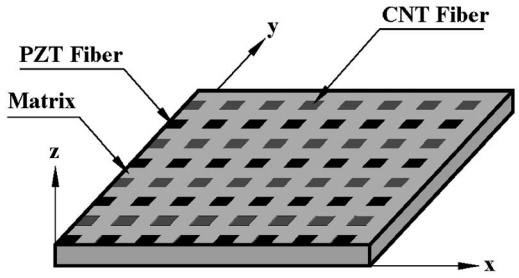



**Fig. 1 Schematic sketch of a NRHPC comprised of a polymer matrix reinforced with CNT and PZT fibers**

$$r = p,\ n,\ \text{and}\ m; \quad \{e^p\} = \begin{Bmatrix} e^p_{31} \\ e^p_{32} \\ e^p_{33} \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{2}$$

In Eq. (1), superscripts $p$, $n$, and $m$ denote, respectively, the PZT, the CNT, and the matrix. For the constituent phase $r$, $\sigma^r_x$, $\sigma^r_y$, and $\sigma^r_z$ represent normal stresses in the $x$, $y$, and $z$, directions, respectively; $\varepsilon^r_x$, $\varepsilon^r_y$, and $\varepsilon^r_z$ are the corresponding normal strains; $\sigma^r_{xy}$, $\sigma^r_{yz}$, and $\sigma^r_{zx}$ are the shear stresses; $\varepsilon^r_{xy}$, $\varepsilon^r_{yz}$, and $\varepsilon^r_{zx}$ are the shear strains; $C^r_{ij}$ $(i, j = 1,2,3,\ldots,6)$ are elastic constants; and $e^p_{31}$ and $e^p_{33}$ are the piezoelectric coefficients of the PZT. A field variable and a material property without a superscript represent quantities for the hybrid composite. We assume that all fibers are perfectly bonded to the matrix, and hence satisfy the following isofield conditions [14,16]:

$$\begin{Bmatrix} \varepsilon^p_z \\ \sigma^p_x \\ \sigma^p_y \\ \sigma^p_{yz} \\ \sigma^p_{xz} \\ \sigma^p_{xy} \end{Bmatrix} = \begin{Bmatrix} \varepsilon^n_z \\ \sigma^n_x \\ \sigma^n_y \\ \sigma^n_{yz} \\ \sigma^n_{xz} \\ \sigma^n_{xy} \end{Bmatrix} = \begin{Bmatrix} \varepsilon^m_z \\ \sigma^m_x \\ \sigma^m_y \\ \sigma^m_{yz} \\ \sigma^m_{xz} \\ \sigma^m_{xy} \end{Bmatrix} = \begin{Bmatrix} \varepsilon_z \\ \sigma_x \\ \sigma_y \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{Bmatrix} \tag{3}$$

As a limitation of the above assumptions, Smith and Auld [14] mentioned that for applications as transducers, this homogenization technique yields good results when the fiber sizes and spacings are sufficiently small as compared with the acoustic wavelengths. Since diameters of CNTs are very small they can be closely packed to make spacing between any two of them much smaller than the acoustic wavelength. Also, the assumption of uniform axial strain in the thickness direction in the three phases is not strictly valid unless the top and the bottom faces are bonded to rigid membranes and are uniformly pressed in the axial direction. However, for CNTs and PZTs distributed uniformly with very little space between them, the assumption gives reasonable results for applications as actuators of beams and plates.

Following the procedure outlined in Ref. [14], the normal stress $\sigma_z$, the normal strains $\varepsilon_x$ and $\varepsilon_y$, and the shear strains $\varepsilon_{xz}$, $\varepsilon_{yz}$, and $\varepsilon_{xy}$ in the homogenized composite can be expressed in terms of the corresponding stresses and strains in the constituent phases. Thus using Eqs. (1) and (3), we obtain

$$\{\sigma\} = [C_1]\{\varepsilon^p\} + [C_2]\{\varepsilon^n\} + [C_3]\{\varepsilon^m\} - \{e_1\}E_z \tag{4}$$

$$\{\varepsilon\} = [V_1]\{\varepsilon^p\} + [V_2]\{\varepsilon^n\} + [V_3]\{\varepsilon^m\} \tag{5}$$

$$[C_4]\{\varepsilon^p\} - [C_5]\{\varepsilon^n\} = \{e_2\}E_z \tag{6}$$

$$[C_5]\{\varepsilon^n\} - [C_6]\{\varepsilon^m\} = 0 \tag{7}$$

in which

$$[C_1] = \begin{bmatrix} C^p_{11} & C^p_{12} & C^p_{13} & 0 & 0 & 0 \\ C^p_{12} & C^p_{22} & C^p_{23} & 0 & 0 & 0 \\ v_p C^p_{13} & v_p C^p_{23} & v_p C^p_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C^p_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C^p_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C^p_{66} \end{bmatrix},$$

$$[C_2] = v_n \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ C^n_{13} & C^n_{23} & C^n_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[C_3] = v_m \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ C^m_{13} & C^m_{23} & C^m_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$[C_4] = \begin{bmatrix} C^p_{11} & C^p_{12} & C^p_{13} & 0 & 0 & 0 \\ C^p_{12} & C^p_{22} & C^p_{23} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & C^p_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C^p_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C^p_{66} \end{bmatrix}$$

$$[C_5] = \begin{bmatrix} C^n_{11} & C^n_{12} & C^n_{13} & 0 & 0 & 0 \\ C^n_{12} & C^n_{22} & C^n_{23} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & C^n_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C^n_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C^n_{66} \end{bmatrix},$$

$$[C_6] = \begin{bmatrix} C^m_{11} & C^m_{12} & C^m_{13} & 0 & 0 & 0 \\ C^m_{12} & C^m_{22} & C^m_{23} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & C^m_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C^m_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C^m_{66} \end{bmatrix}$$

$$[V_1] = \begin{bmatrix} v_p & 0 & 0 & 0 & 0 & 0 \\ 0 & v_p & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & v_p & 0 & 0 \\ 0 & 0 & 0 & 0 & v_p & 0 \\ 0 & 0 & 0 & 0 & 0 & v_p \end{bmatrix},$$

**Table 1 Material properties of the constituent phases**

| Material | Source | $C_{11}$ (GPa) | $C_{12}$ (GPa) | $C_{13}$ (GPa) | $C_{33}$ (GPa) | $C_{44}$ (GPa) | $e_{31}^p$ (C/m$^2$) | $e_{33}^p$ (C/m$^2$) |
|---|---|---|---|---|---|---|---|---|
| CNT (5,5) | [5] | 668 | 404 | 184 | 2153 | 791 | - | - |
| CNT (20,20) | [5] | 148 | 144 | 43.5 | 545 | 227 | - | - |
| CNT (50,50) | [5] | 55.1 | 54.9 | 17.5 | 218 | 92 | - | - |
| PZT5H | [14] | 151 | 98 | 96 | 124 | 23 | −5.1 | 27 |
| Spurr | [14] | 5.3 | 3.1 | 3.1 | 5.3 | 0.64 | - | - |

$$[V_2] = \begin{bmatrix} v_n & 0 & 0 & 0 & 0 & 0 \\ 0 & v_n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & v_n & 0 & 0 \\ 0 & 0 & 0 & 0 & v_n & 0 \\ 0 & 0 & 0 & 0 & 0 & v_n \end{bmatrix}$$

$$[V_3] = \begin{bmatrix} v_m & 0 & 0 & 0 & 0 & 0 \\ 0 & v_m & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & v_m & 0 & 0 \\ 0 & 0 & 0 & 0 & v_m & 0 \\ 0 & 0 & 0 & 0 & 0 & v_m \end{bmatrix},$$

$$\{e_1\} = \begin{Bmatrix} e_{31}^p \\ e_{32}^p \\ v_p e_{33}^p \\ 0 \\ 0 \\ 0 \end{Bmatrix}, \quad \text{and} \quad \{e_2\} = \begin{Bmatrix} e_{31}^p \\ e_{32}^p \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{8}$$

In Eq. (8) $v_p$, $v_n$, and $v_m$ represent volume fractions of PZTs, CNTs, and the matrix, respectively. The elimination of field variables of the constituent phases from Eq. (4) to Eq. (7) yields the following constitutive relation for the proposed hybrid NRHPC:

$$\{\sigma\} = [C]\{\varepsilon\} - \{e\}E_z \tag{9}$$

where

$$[C] = [C_1][V_5]^{-1} + [C_7][V_6]^{-1}$$

$$\{e\} = \{e_1\} - [C_1][V_5]^{-1}[V_4][C_6]^{-1}\{e_2\} + [C_7][V_6]^{-1}[V_1][C_4]^{-1}\{e_2\}$$

$$[V_4] = [V_3] + [V_2][C_5]^{-1}[C_6], \quad [V_5] = [V_1] + [V_4][C_6]^{-1}[C_4]$$

$$[V_6] = [V_4] + [V_1][C_4]^{-1}[C_6] \quad \text{and} \quad [C_7] = [C_3] + [C_2][C_5]^{-1}[C_6] \tag{10}$$

Comparing Eq. (9) with the constitutive relation $(1)_1$ for a PZT, the effective piezoelectric coefficients $e_{31}$, $e_{32}$, and $e_{33}$ of the NRHPC can be identified as $e_{31}=e(1)$, $e_{32}=e(2)$, and $e_{33}=e(3)$.

## 3 Results and Discussion

Material properties of CNTs, taken from Ref. [5], and of the PZT5H and the epoxy are listed in Table 1. Effective properties of the NRHPC, computed by simultaneously varying volume fractions of CNTs and PZT fibers, are compared with those given by Smith and Auld [14] for the PZT5H/epoxy composite.

Figure 2 depicts the variation of the effective piezoelectric coefficient $e_{31}$ of the NRHPC with the PZT fiber volume fraction, and for different volume fractions of CNTs. It is clear from these plots that the value of $e_{31}$ of the NRHPC is significantly enhanced by the addition of CNTs, and equals twice that of the 1–3 PZT5H/epoxy composites for $v_n=0.3$ and $v_p=0.4$. Furthermore, adding CNTs also improves values of elastic constants of the NRHPC over those of the existing 1–3 PZCs. As an example, we illustrate in Fig. 3 the variation of the effective elastic constant $C_{33}$ with respect to the PZT5H fiber volume fraction for different volume fractions of CNTs. Using Eq. (10), values of other effective elastic and piezoelectric constants can be easily computed for any combination of volume fractions of CNTs and PZT fibers. Also, it is evident from Figs. 2 and 3 that for a particular value of $v_p$, values of the effective piezoelectric coefficient $e_{31}$ and elastic properties increase with an increase in $v_n$. The significant improvement in

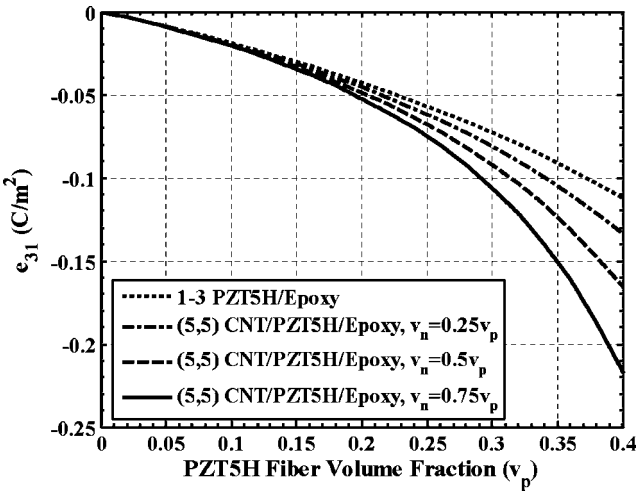



**Fig. 2 Effective piezoelectric coefficient $e_{31}$ of the NRHPC**

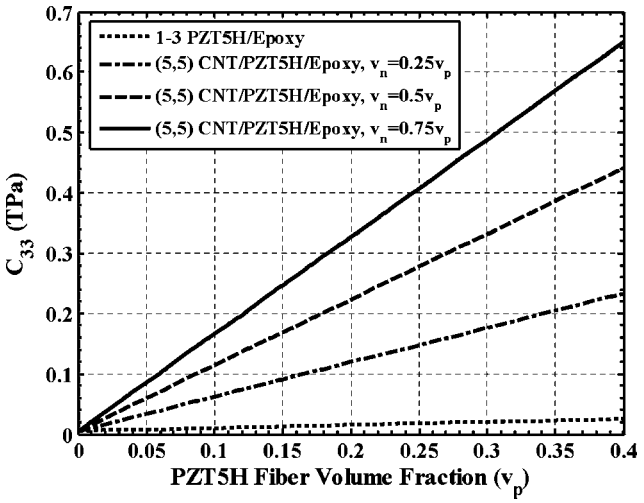



**Fig. 3 Effective elastic coefficient $C_{33}$ of the NRHPC**

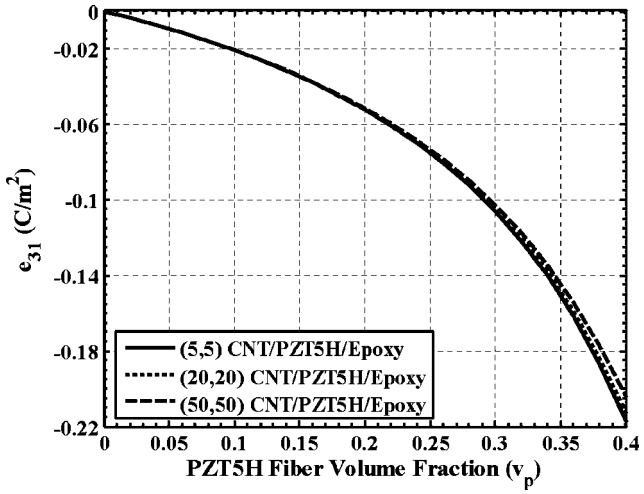



**Fig. 4 For $v_n=0.75v_p$, variation of the effective piezoelectric coefficient $e_{31}$ of the NRHPC with the diameter of CNTs**

effective properties of the NRHPC is attributed to the fact that CNT reinforcements enhance the elastic properties of the matrix. It is also found that the value of the other in-plane effective piezoelectric coefficient $e_{32}$ of the NRHPC equals that of the effective coefficient $e_{31}$. However, the addition of CNTs does not affect the value of the effective piezoelectric coefficient $e_{33}$. Results plotted in Figs. 4 and 5 reveal that as the diameter of the CNTs increases, magnitudes of both the $e_{31}$ and the elastic moduli decrease because elastic moduli of a CNT decrease with an increase in the diameter of the CNT.

## 4 Conclusions

We have proposed a hybrid piezoelectric composite comprised of a polymer matrix and single-walled CNTs and piezoceramic (PZT5H) fibers aligned parallel to each other along the thickness of the laminate. The PZT5H fibers are poled in the thickness direction. Values of the effective piezoelectric coefficients $e_{31}$ and $e_{33}$ are proportional to the in-plane and the out-of-plane actuations, respectively, due to a voltage applied across the thickness of the hybrid lamina. Effective moduli of the hybrid lamina have been determined by using the isostrain and the isostress assumptions. It is found that the value of $e_{31}$ of the proposed hybrid composite is significantly higher than that of the existing 1–3 piezocomposites [14] at the practically useful volume fraction of PZT fibers. For a fixed volume fraction of PZT fibers, the value of $e_{31}$ for the hybrid composite increases with an increase in the volume fraction of CNTs, and that of $e_{33}$ remains unaltered. Elastic moduli of the hybrid composite are also much larger than those of the existing 1–3 piezocomposites [15]. Because of increase in the value of $e_{31}$, the proposed hybrid composite can act as a distributed actuator for both in-plane and out-of-plane actuations while the in-plane actuation by the existing 1–3 piezocomposites is negligibly small [15].

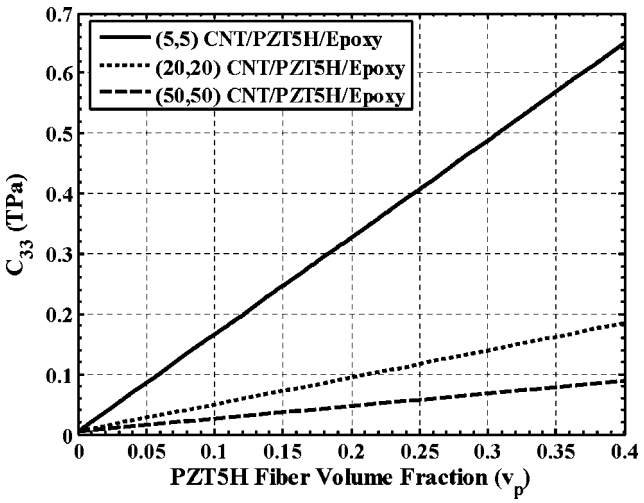



**Fig. 5 For $v_n=0.75v_p$, variation of the effective elastic constant $C_{33}$ of the NRHPC with the diameter of CNTs**

## Acknowledgment

This work was partially supported by the Office of Naval Research Grant No. N00014-98-06-0567 to Virginia Polytechnic Institute and State University with Dr. Y. D. S. Rajapakse as the program manager. Partial financial support from Virginia Tech's Institute of Critical Technologies and Sciences is gratefully acknowledged. Views expressed herein are those of authors, and neither of the funding agencies nor of VPI&SU.

## References

[1] Iijima, S., 1991, "Helical Microtubes of Graphitic Carbon," Nature (London), **354**, pp. 56–58.
[2] Treacy, M. M. J., Ebbessen, T. W., and Gibson, J. M., 1996, "Exceptionally High Young's Modulus Observed for Individual Carbon Nanotubes," Nature (London), **381**, pp. 678–680.
[3] Li, C., and Chou, T. W., 2003, "A Structural Mechanics Approach for the Analysis of Carbon Nanotubes," Int. J. Solids Struct., **40**, pp. 2487–2499.
[4] Sears, A., and Batra, R. C., 2004, "Macroscopic Properties of Carbon Nanotubes From Molecular Mechanics Simulations," Phys. Rev. B, **69**, p. 235406.
[5] Shen, L., and Li, J., 2004, "Transversely Isotropic Elastic Properties of Single-Walled Carbon Nanotubes," Phys. Rev. B, **69**, p. 045415.
[6] Batra, R. C., and Sears, A., 2007, "Uniform Radial Expansion/Contraction of Carbon Nanotubes and Their Transverse Elastic Moduli," Modell. Simul. Mater. Sci. Eng., **15**, pp. 835–844.
[7] Gupta, S. S., and Batra, R. C., 2008, "Continuum Structures Equivalent in Normal Mode Vibrations to Single-Walled Carbon Nanotubes," Comput. Mater. Sci., **43**, pp. 715–723.
[8] Batra, R. C., and Gupta, S. S., 2008, "Wall Thickness and Radial Breathing Modes of Single-walled Carbon Nanotubes," ASME J. Appl. Mech., **75**, p. 061010.
[9] Wu, J., Hwang, K. C., and Huang, Y., 2008, "An Atomistic-Based Finite-Deformation Shell Theory for Single-Wall Carbon Nanotubes," J. Mech. Phys. Solids, **56**, pp. 279–292.
[10] Thostenson, E. T., and Chou, T. W., 2003, "On the Elastic Properties of Carbon Nanotube Based Composites: Modeling and Characterization," J. Phys. D, **36**, pp. 573–582.
[11] Odegard, G. M., Gates, T. S., Wise, K. E., Park, C., and Siochi, E. J., 2003, "Constitutive Modeling of Nanotube-Reinforced Polymer Composites," Compos. Sci. Technol., **63**, pp. 1671–1687.
[12] Song, Y. S., and Youn, J. R., 2006, "Modeling of Effective Elastic Properties for Polymer Based Carbon Nanotube Composites," Polymer, **47**, pp. 1741–1748.
[13] Piezocomposites, Materials System Inc., 543 Great Road, Littleton, MA 01560.
[14] Smith, W. A., and Auld, B. A., 1991, "Modeling 1-3 Composite Piezoelectrics: Thickness Mode Oscillations," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **38**(1), pp. 40–47.
[15] Ray, M. C., and Pradhan, A. K., 2007, "On the Use of Vertically Reinforced 1-3 Piezoelectric Composites for Hybrid Damping of Laminated Composite Plates," Mechanics of Advanced Materials and Structures, **14**, pp. 245–261.
[16] Benveniste, Y., and Dvorak, G. J., 1992, "Uniform Fields and Universal Relations in Piezoelectric Composites," J. Mech. Phys. Solids, **40**, pp. 1295–1412.