Convergence of the Classical Rayleigh-Ritz Method and the Finite Element Method

Leonard Meirovitch* and Moon K. Kwak†
Virginia Polytechnic Institute and State University, Blacksburg, Virginia

The Rayleigh-Ritz method is a technique for approximating the eigensolution associated with a distributed structure. The method amounts to approximating the solution of a differential eigenvalue problem having no known closed-form solution by a finite series of trial functions, thus replacing the differential eigenvalue problem by an algebraic eigenvalue problem. The finite element method can be regarded as a Rayleigh-Ritz method, at least for structures. The main difference between the finite element method and the classical Rayleigh-Ritz method lies in the nature of the admissible functions. An important question in both the classical Rayleigh-Ritz method and the finite element method is the speed of convergence. It is demonstrated in this paper that convergence of the classical Rayleigh-Ritz method can be vastly improved by introducing a new class of admissible functions, called quasi-comparison functions. Factors affecting the convergence of the finite element method are also discussed.

Introduction

THE eigenvalue problem for a distributed structure is defined by a differential equation and certain boundary conditions and is known as the differential eigenvalue problem.[1] More often than not, particularly for nonuniform structures characterized by coefficients depending on the spatial variable(s), a closed-form solution of the differential eigenvalue problem is not feasible. The Rayleigh-Ritz method is a technique for obtaining an approximate solution of the eigenvalue problem. The method consists of approximating the solution of a differential eigenvalue problem having no known closed-form solution by a finite series of trial functions multiplied by undetermined coefficients. There are two general approaches to the determination of the approximate solution. One approach is to work with the differential equation and to select the trial functions from the space of comparison functions, namely, functions that are $2p$ times differentiable and satisfy all of the boundary conditions, where $2p$ is the order of differential equation.[1] The other approach consists of defining and rendering a Rayleigh quotient for the system stationary. If the Rayleigh quotient is given in terms of expressions related to the potential and kinetic energy, then the trial functions can be selected from the space of admissible functions, namely, functions that are p times differentiable and satisfy the geometric boundary conditions alone.[1] Because admissible functions are considerably more abundant than comparison functions and are generally easier to produce, we choose the second approach.

The finite element method can be regarded as a Rayleigh-Ritz method, at least for structures. The main difference between the finite element method and the classical Rayleigh-Ritz method lies in the nature of the admissible functions.[2] Indeed, in the classical Rayleigh-Ritz method, the admissible functions are global functions, i.e., they are defined over the entire domain of the structure. Differentiability is seldom an issue, if ever. On the other hand, in the finite element method, the admissible functions are local functions, i.e., they are defined over smaller subdomains, such as two adjacent finite elements. The admissible functions are low-degree polynomials, satisfying the minimum differentiability requirements. They are often referred to as interpolation functions.

The finite element method has proved enormously successful in structural dynamics. One of the reasons for its success is that it can solve problems involving pronounced parameter nonuniformities and/or complicated geometries, where other methods cannot cope. Another reason is that its use has become mechanized, with the choice of interpolation functions for a given problem being very narrow. In fact, this is the reason that has inhibited the more frequent use of the classical Rayleigh-Ritz method. Indeed, in the classical Rayleigh-Ritz method, there is a wide choice of admissible functions for any given problem and there are no good criteria to help with the selection. Yet, in problems in which the parameter nonuniformities are not very pronounced, the classical Rayleigh-Ritz method is capable of yielding superior accuracy for the same number of degrees of freedom, or the same accuracy with fewer degrees of freedom.

One of the questions in both the classical Rayleigh-Ritz method and the finite element method is the speed of convergence. In the classical Rayleigh-Ritz method, and perhaps more indirectly in the finite element method, the question is related to the completeness of the set of admissible functions. The concept of completeness is more qualitative than quantitative in nature. Moreover, whether a set is complete or not can be answered only in connection with a given problem. Even then, there seems to exist a degree of completeness, as a set that is complete for a certain problem may not be as complete for a somewhat different problem. This question is important because in many cases it is possible to use as admissible functions for a given system the eigenfunctions of a related simpler system.

This paper is concerned with the convergence of the Rayleigh-Ritz method and, in particular, how completeness affects convergence. It is demonstrated that convergence of the classical Rayleigh-Ritz method can be improved vastly by introducing a new class of admissible functions, called quasi-comparison functions. The quasi-comparison functions are admissible functions such that, although individually they do not satisfy the natural boundary conditions, finite linear combinations thereof are capable of satisfying them to any desired degree of accuracy. Quasi-comparison functions can be pro-

Presented as AIAA Paper 88-2269 at the AIAA/ASME/ASCE/AHS 29th Structures, Structural Dynamics, and Materials Conference, Williamsburg, VA, April 18–20, 1988; received May 12, 1988.

*University Distinguished Professor, Department of Engineering Science and Mechanics. Fellow AIAA.

†Graduate Research Assistant; currently Assistant Professor, Department of Engineering Science and Mechanics. Member AIAA.

duced by enlarging the pool of admissible functions so as to enhance completeness. Convergence of the finite element method seems to be affected more by the ability of satisfying the differential equations than the boundary conditions.

A numerical example consisting of a rod in axial vibration fixed at one end and with a spring attached at the other end is considered. The eigenvalue problem is solved by both the classical Rayleigh-Ritz method and the finite element method. Results demonstrate the advantages of using quasi-comparison functions.

Differential Eigenvalue Problem

We are concerned with the differential eigenvalue problem

$$\mathcal{L}W(x) = \lambda m(x)W(x), \qquad 0<x<L \tag{1}$$

where $\mathscr{L}$ is a linear homogeneous differential operator of order $2p$, in which p is an integer, W is the displacement of a given point x of the elastic structure, λ is a parameter, $m(x)$ is the mass density at point x, and L is the length of the structure. The displacement is subject to the boundary conditions

$$B_i W = 0, \qquad i = 1,2,...,p, \qquad x = 0,L \tag{2}$$

where B_i are linear homogeneous differential operators of maximum order $2p-1$. The solution of the eigenvalue problem, Eqs. (1) and (2), consists of a denumerably infinite set of eigenvalues λ_r and associated eigenfunctions $W_r(x)$ $(r = 1,2,...)$.

To permit an expanded discussion of the eigenvalue problem and the various properties of its solution, we wish to review briefly certain pertinent definitions. To this end, we consider two functions f and g that are piecewise continuous in the domain $0 \le x \le L$ and define the inner product

$$(f,g) = \int_0^L fg\, \mathrm{d}x \tag{3}$$

If the inner product vanishes, functions f and g are said to be orthogonal over the domain $0 \le x \le L$. Next, let us consider a set of linearly independent functions $\phi_1,\phi_2,...$ and define

$$f_n = \sum_{r=1}^{n} c_r \phi_r \tag{4}$$

where c_r $(r = 1,2,...,n)$ are constant coefficients. Then, if by choosing n large enough the mean-square error

$$M = (f-f_n, f-f_n) = \int_0^L (f-f_n)^2\, \mathrm{d}x \tag{5}$$

can be made less than any arbitrarily small positive number ε, the set of functions $\phi_1,\phi_2,...$ is said to be complete.

There are three classes of functions of special interest: 1) eigenfunctions, 2) comparison functions, and 3) admissible functions. As indicated earlier, the eigenfunctions must satisfy both the differential equation, Eq. (1), and the boundary conditions, Eqs. (2). This is a relatively small class of functions, and more often than not it is not possible to solve the eigenvalue problem exactly and obtain the eigenfunctions. In such cases, we must be content with an approximate solution, which can be obtained in the form of a linear combination of members from the two other classes of functions. The class of comparison functions comprises functions that are $2p$ times differentiable and satisfy all the boundary conditions but not necessarily the differential equation. Clearly, the class of comparison functions is considerably larger than the class of eigenfunctions. To introduce the third class of functions, it is necessary to examine the boundary conditions closer. There are two types of boundary conditions: geometric boundary conditions and natural boundary conditions. As the name indicates, the first type involves the satisfaction of some geometric condition at the boundary, such as zero displacement or zero slope. In the case of a geometric boundary condition, the operator B_i is of maximum order $p-1$. On the other hand, the second type reflects force or bending moment balance at the boundary. In the case of a natural boundary condition, the operator B_i is of maximum order $2p-1$. The class of admissible functions comprises functions that are p times differentiable and satisfy the geometric boundary conditions only. It is obvious that the class of admissible functions is far larger than the class of comparison functions.

Next, let us consider two comparison functions u and v and define the inner product

$$(u,\mathscr{L}v) = \int_0^L u\mathscr{L}v\, \mathrm{d}x \tag{6}$$

Then, the differential operator $\mathscr{L}$ is said to be self-adjoint if

$$(u,\mathscr{L}v) = (v,\mathscr{L}u) \tag{7}$$

Whether Eq. (7) holds or not can be verified through integrations by parts, with due consideration of the boundary conditions. A mere function, such as the mass density m, is self-adjoint by definition. Hence, if $\mathscr{L}$ is self-adjoint, then the system is self-adjoint. In this case, integrations by parts yield

$$\int_0^L u\mathscr{L}v\, \mathrm{d}x = \int_0^L \sum_{k=0}^{p} a_k \frac{\mathrm{d}^k u}{\mathrm{d}x^k}\frac{\mathrm{d}^k v}{\mathrm{d}x^k}\mathrm{d}x + \sum_{\ell=0}^{p-1} b_\ell \frac{\mathrm{d}^\ell u}{\mathrm{d}x^\ell}\frac{\mathrm{d}^\ell v}{\mathrm{d}x^\ell}\bigg|_0^L \tag{8}$$

where a_k $(k = 0,1,...,p)$ and b_ℓ $(\ell = 0,1,...,p-1)$ are, in general, functions of x. Observing that the right side of Eq. (8) is symmetric in u and v, we conclude that self-adjointness implies symmetry of the eigenvalue problem in the sense indicated by Eq. (8).

It will prove convenient to introduce the energy inner product defined by

$$[u,v] = \int_0^L \sum_{k=0}^{p} a_k \frac{\mathrm{d}^k u}{\mathrm{d}x^k}\frac{\mathrm{d}^k v}{\mathrm{d}x^k}\mathrm{d}x + \sum_{\ell=0}^{p-1} b_\ell \frac{\mathrm{d}^\ell u}{\mathrm{d}x^\ell}\frac{\mathrm{d}^\ell v}{\mathrm{d}x^\ell}\bigg|_0^L \tag{9}$$

and we observe that the energy inner product is defined even when u and v are admissible functions rather than comparison functions. The term "energy inner product" can be explained by the fact that the product of u with itself, namely,

$$[u,u] = \int_0^L \sum_{k=0}^{p} a_k \left(\frac{\mathrm{d}^k u}{\mathrm{d}x^k}\right)^2 \mathrm{d}x + \sum_{\ell=0}^{p-1} b_\ell \left(\frac{\mathrm{d}^\ell u}{\mathrm{d}x^\ell}\right)^2\bigg|_0^L \tag{10}$$

can be interpreted as twice the maximum potential energy of the system whose eigenvalue problem is described by Eqs. (1) and (2). Now, let us consider the set of functions $\phi_1,\phi_2,...$ and define the approximation u^n to u

$$u^n = \sum_{r=1}^{n} d_r \phi_r \tag{11}$$

where d_r $(r = 1,2,...,n)$ are constant coefficients. Then, if by choosing a sufficiently large n the energy inner product $[u-u^n, u-u^n]$ can be made less than any arbitrarily small positive number ε, the set $\phi_1,\phi_2,...$ is said to be complete in energy.

As a consequence of the self-adjointness of the operator $\mathscr{L}$ and, hence, the self-adjointness of the system, the eigenvalues $\lambda_1,\lambda_2,...$ are real. As a corollary, the eigenfunctions $W_1,W_2,...$ are real. Moreover, if the system is self-adjoint, the eigenfunctions are orthogonal.[1] If the eigenfunctions are normalized so that they satisfy $\int_0^L mW_r^2\, \mathrm{d}x = 1$, $(r = 1,2,...)$ then they become orthonormal, where the orthonormality property is expressed by

$$\int_0^L mW_r W_s \,\mathrm{d}x = \delta_{rs}, \qquad r,s = 1,2,... \tag{12}$$

where δ_{rs} is the Kronecker delta. In view of Eq. (1), it follows that

$$\int_0^L W_r \mathscr{L} W_s \,\mathrm{d}x = \int_0^L W_s \mathscr{L} W_r \,\mathrm{d}x = \lambda_r \delta_{rs}, \qquad r,s = 1,2,... \tag{13}$$

Because the eigenfunctions are orthogonal over $0 \le x \le L$, they comprise a complete set. This fact permits the formulation of the following expansion theorem: Every function W with continuous $\mathscr{L}W$ and satisfying the boundary conditions of the system can be expanded in the absolutely and uniformly convergent series

$$W = \sum_{r=1}^{\infty} c_r W_r \tag{14}$$

where the coefficients c_r are given by

$$c_r = \int_0^L mWW_r \,\mathrm{d}x, \qquad r = 1,2,... \tag{15}$$

Rayleigh-Ritz Method

Let us consider once again the eigenvalue problem described by Eqs. (1) and (2), in which the operator $\mathscr{L}$ is self-adjoint, and define the Rayleigh quotient

$$R(W) = \frac{[W,W]}{(\sqrt{m}W,\sqrt{m}W)} \tag{16}$$

where W is a trial function, $[W, W]$ the energy inner product and $(\sqrt{m}W,\sqrt{m}W)$ the inner product of $\sqrt{m}W$ with itself.[1] Using the expansion theorem, Eq. (14), it is possible to show that Rayleigh's quotient has stationary values at the system eigenfunctions and that the stationary values are precisely the system eigenvalues.[1] Hence, as an alternative to solving the differential eigenvalue problem, Eqs. (1) and (2), one can consider rendering Rayleigh's quotient stationary. This variational approach is particularly attractive when the object is to obtain an approximate solution to the differential eigenvalue problem, because the solution using Rayleigh's quotient in the form given by Eq. (16) can be constructed from the space of admissible functions instead of the space of comparison functions.

Next, let us consider a trial function W from the space of admissible functions. For practical reasons, we cannot consider the entire space but only a finite-dimensional subspace, where the subspace is spanned by the functions $\phi_1,\phi_2,...,\phi_n,...$. The functions are such that 1) any n elements $\phi_1,\phi_2,...,\phi_n$ are linearly independent and 2) the sequence $\phi_1,\phi_2,...,\phi_n,...$ is complete in energy.

The Rayleigh-Ritz method is a procedure for deriving an approximate solution of the eigenvalue problem. For the preceding reasons, we are interested in solving the eigenvalure problem by rendering Rayleigh's quotient stationary. To this end, we consider a trial function in the form of the finite series

$$W^n = \sum_{i=1}^{n} u_i \phi_i \tag{17}$$

where u_i are coefficients to be determined and ϕ_i are known admissible functions ($i = 1,2,...,n$); the meaning of the superscript n is obvious. Inserting Eq. (17) into Eq. (16), we obtain

$$R(W^n) = \frac{[W^n,W^n]}{(\sqrt{m}W^n,\sqrt{m}W^n)} = R(u_1,u_2,...,u_n) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} k_{ij}u_i u_j}{\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}u_i u_j} \tag{18}$$

in which

$$k_{ij} = k_{ji} = [\phi_i,\phi_j] = \int_0^L \sum_{k=0}^{p} a_k \frac{\mathrm{d}^k\phi_i}{\mathrm{d}x^k}\frac{\mathrm{d}^k\phi_j}{\mathrm{d}x^k}\,\mathrm{d}x + \sum_{\ell=0}^{p-1} b_\ell \frac{\mathrm{d}^\ell\phi_i}{\mathrm{d}x^\ell}\frac{\mathrm{d}^\ell\phi_j}{\mathrm{d}x^\ell}\bigg|_0^L, \qquad i,j = 1,2,...,n \tag{19a}$$

$$m_{ij} = m_{ji} = (\sqrt{m}\phi_i,\sqrt{m}\phi_j) = \int_0^L m\phi_i\phi_j \,\mathrm{d}x, \qquad i,j = 1,2,...,n \tag{19b}$$

are symmetric stiffness and mass coefficients, respectively, where the symmetry is a direct consequence of the system being self-adjoint. The conditions for the stationarity of Rayleigh's quotient are

$$\frac{\partial R}{\partial u_r} = 0, \qquad r = 1,2,...,n \tag{20}$$

Equation (20) leads to n homogeneous algebraic equations,[1] which can be identified as representing the algebraic eigenvalue problem and expressed in the matrix form

$$\boldsymbol{Ku} = \lambda^n \boldsymbol{Mu} \tag{21}$$

where $K = [k_{ij}]$ and $M = [m_{ij}]$ are $n \times n$ real symmetric stiffness and mass matrices, respectively, $\boldsymbol{u}[u_1,u_2,...,u_n]^T$ is the n-dimensional vector of undetermined coefficients and λ^n is a parameter. Hence, spatial discretization has the effect of replacing a differential eigenvalue problem with an algebraic one. The solution of the algebraic eigenvalue problem yields the values $\lambda_1^n,\lambda_2^n,...,\lambda_n^n$ of the parameter; they are known as the estimated eigenvalues, or computed eigenvalues, as opposed to the actual eigenvalues $\lambda_1,\lambda_2,...\lambda_n$ of the distributed system. Clearly, we can only estimate the n lowest eigenvalues, and there are no estimates for $\lambda_{n+1},\lambda_{n+2},...$. To the eigenvalues $\lambda_1^n,\lambda_2^n,...,\lambda_n^n$ belong the eigenvectors $\boldsymbol{u}_1,\boldsymbol{u}_2,...,\boldsymbol{u}_n$, respectively. They can be used to produce estimates of the eigenfunctions or computed eigenfunctions. Indeed, introducing the n-vector $\phi = [\phi_1\phi_2\cdots\phi_n]^T$ of admissible functions and recalling Eq. (17), we can write the estimated eigenfunctions, or computed eigenfunctions,

$$W_r^n = \boldsymbol{u}_r^T\phi, \qquad r = 1,2,...,n \tag{22}$$

The question remains as to how the estimated eigenvalues and eigenfunctions compare with the actual ones. It can be shown[1] that

$$\lambda_r^n \ge \lambda_r, \qquad r = 1,2,...,n \tag{23}$$

or the estimated eigenvalues provide upper bonds for the actual eigenvalues. Unfortunately, the estimated eigenfunctions do not lend themselves to such meaningful comparison. However, it can be stated that, in general, the estimated eigenvalues are better approximations of the actual eigenvalues than the estimated eigenfunctions are of the actual eigenfunctions. The accuracy can be improved by increasing the number n of terms in Eq. (17). Indeed, because the admissible functions $\phi_1, \phi_2, \ldots, \phi_n, \ldots$ are from a complete set, we can write

$$\lim_{n\to\infty} \lambda_r^n = \lambda_r, \qquad r = 1,2,\ldots,n \tag{24}$$

Moreover, as n increases, the estimated eigenvalues approach the actual eigenvalues monotonically from above. Convergence is guaranteed if the admissible functions $\phi_1, \phi_2, \ldots$ are from a complete set, but the rate of convergence depends on the choice of admissible functions.

Although not stated above, one can infer that the admissible functions ϕ_r $(r = 1,2,\ldots,n)$ are global functions, in the sense that they extend over the entire length L of the structure. Moreover, the coefficients u_r $(r = 1,2,\ldots,n)$ do not represent actual displacements. Indeed, they are abstract quantities providing only a measure of the contribution of the individual admissible functions to the displacement. We refer to this version as the classical Rayleigh-Ritz method. The finite element method also can be regarded as a Rayleigh-Ritz method. In contrast, however, in the finite element method the admissible functions are local functions, in the sense that they extend over given segments of the structure only, where the segments are known as finite elements. In addition, the admissible functions tend to be very simple. In fact, they tend to be low-degree polynomials, satisfying the minimum differentiability conditions. Finally, the coefficients u_j represent actual displacements at points shared by two adjacent elements, where the points are known as nodes. Whereas in both the classical Rayleigh-Ritz method and the finite element method accuracy is improved by increasing the number of admissible functions, in the finite element method this is achieved by increasing the number of elements, which amounts to using a finer finite element mesh.

To examine how accuracy improves as the number of admissible functions increases, we propose to solve the eigenvalue problem for the system of Fig. 1 by both the classical Rayleigh-Ritz method and the finite element method. This example will provide the motivation for the real object of this investigation, namely, to develop the framework for an approach capable of achieving improved convergence characteristics.

The system of Fig. 1 represents a nonuniform rod in axial vibration with one end fixed and the other end restrained by a spring. The parameters are as follows:

$$EA(x) = (6EA/5)[1 - \tfrac{1}{2}(x/L)^2] \tag{25a}$$

$$m(x) = (6m/5)[1 - \tfrac{1}{2}(x/L)^2] \tag{25b}$$

$$k = EA/L \tag{25c}$$

where $EA(x)$ is the axial stiffness, in which E is the modulus of elasticity and $A(x)$ is the cross-sectional area, $m(x)$ is the mass per unit length, and k is the spring constant. The energy inner product, appearing at the numerator of Rayleigh's quotient, has the form[1]

$$[W,W] = \int_0^L EA(x)\left[\frac{dW(x)}{dx}\right]^2 dx + kW^2(L) \tag{26}$$

so that $p = 1$. Moreover, the weighted inner product, appearing at the denominator of Rayleigh's quotient, is

$$(\sqrt{m}W, \sqrt{m}W) = \int_0^L m(x)W^2(x)\, dx \tag{27}$$

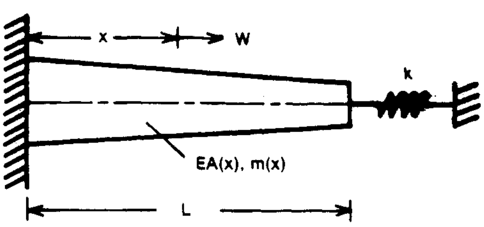

Fig. 1 Nonuniform rod in axial vibration.

For future reference, the boundary conditions are as follows[1]:

$$W(0) = 0 \tag{28a}$$

$$EA(x)\frac{dW(x)}{dx} + kW(x) = 0 \quad \text{at} \quad x = L \tag{28b}$$

where Eq. (28a) is recognized as a geometric boundary condition and Eq. (28b) as a natural boundary condition.

In the classical Rayleigh-Ritz method, we assume an approximate solution in the form

$$W(x) = \boldsymbol{a}^T\boldsymbol{\phi}(x) \tag{29}$$

where $\boldsymbol{a}$ is an n-vector of coefficients and $\boldsymbol{\phi}(x)$an n vector of admissible functions, and we note that the superscript n was omitted from W for simplicity of notation. Inserting Eq. (29) into Eq. (26), we obtain the energy inner product

$$[W,W] = \boldsymbol{a}^T[\boldsymbol{\phi},\boldsymbol{\phi}^T]\boldsymbol{a} = \boldsymbol{a}^T\boldsymbol{K}\boldsymbol{a} \tag{30}$$

where

$$K = [\boldsymbol{\phi},\boldsymbol{\phi}^T] = \int_0^L EA(x)\frac{d\boldsymbol{\phi}(x)}{dx}\frac{d\boldsymbol{\phi}^T(x)}{dx}\, dx + k\boldsymbol{\phi}(L)\boldsymbol{\phi}^T(L) \tag{31}$$

is recognized as the stiffness matrix. Moreover, introducing Eq. (29) into Eq. (27), we obtain the weighted inner product

$$(\sqrt{m}W,\sqrt{m}W) = \boldsymbol{a}^T(\sqrt{m}\boldsymbol{\phi},\sqrt{m}\boldsymbol{\phi}^T)\boldsymbol{a} = \boldsymbol{a}^T\boldsymbol{M}\boldsymbol{a} \tag{32}$$

where

$$M = (\sqrt{m}\boldsymbol{\phi},\sqrt{m}\boldsymbol{\phi}^T) = \int_0^L m(x)\boldsymbol{\phi}(x)\boldsymbol{\phi}^T(x)\, dx \tag{33}$$

is the mass matrix.

The admissible functions must satisfy only boundary condition (28a). We choose the admissible functions

$$\phi_i(x) = \sin\frac{(2i-1)\pi x}{2L}, \qquad i = 1,2,\ldots,n \tag{34}$$

which are recognized as the eigenfunctions of a related system, namely, a uniform rod in axial vibration with the left end fixed and the right end free. The stiffness and mass matrices are obtained by introducing Eqs. (34) into Eqs. (31) and (33), respectively.

In the case of the finite element method, we consider the approximate solution

$$W(x) = \boldsymbol{W}_j^T\boldsymbol{\phi}(x), \quad (j-1)h < x < jh, \qquad j = 1,2,\ldots,n \tag{35}$$

where $h = L/n$ is the width of the finite element and

$$\boldsymbol{W}_j = \begin{bmatrix} W_{j-1} \\ W_j \end{bmatrix} \tag{36a}$$

$$\boldsymbol{\phi}(x) = \begin{bmatrix} \phi_1(x) \\ \phi_2(x) \end{bmatrix} \tag{36b}$$

in which W_{j-1} and W_j are the values of the displacement $W(x)$ evaluated at $x = (j-1)h$ and $x = jh$, respectively, and

$$\phi_1(x) = j - (x/h) \tag{37a}$$

$$\phi_2(x) = 1 - [j - (x/h)] \tag{37b}$$

are admissible functions having the same shape for every element, namely, straight lines.[1] They are commonly known as linear interpolation functions. Following the same pattern as above, the energy inner product can be written in the form

$$[W,W] = \sum_{j=1}^{n} W_j^T K_j W_j = W^T K W \tag{38}$$

where K_j are 2×2 element stiffness matrices having the typical expression

$$K_j = \int_{(j-1)h}^{jh} EA(x) \frac{\mathrm{d}\phi(x)}{\mathrm{d}x} \frac{\mathrm{d}\phi^T(x)}{\mathrm{d}x} \mathrm{d}x + \delta_{jn} k\phi(L)\phi^T(L), \quad j = 1,2,...,n \tag{39}$$

Moreover, $W = [W_1 W_2 \cdots W_n]^T$ is an n-vector of nodal displacements and K is the overall stiffness matrix, obtained from the element stiffness matrices through an assembling process.[1] Similarly, the weighted energy inner product can be written as

$$(\sqrt{m}W, \sqrt{m}W) = \sum_{j=1}^{n} W_j^T M_j W_j = W^T M W \tag{40}$$

where M_j are the 2×2 element mass matrices

$$M_j = \int_{(j-1)h}^{jh} m(x)\phi(x)\phi^T(x)\, \mathrm{d}x, \qquad j = 1,2,...,n \tag{41}$$

and M is the overall mass matrix, obtained from the element mass matrices through the same assembling process mentioned earlier.

The eigenvalue problem was solved by both the classical Rayleigh-Ritz method and the finite element method. The lowest three natural frequencies are listed in Table 1 for $n = 1,2,...,30$ in the columns denoted by classical Rayleigh-Ritz (CRR) and finite element method (FEM), respectively. We observe that the FEM yields slightly better convergence than the CRR for the first frequency but poorer convergence for the second and third. Actually, the CRR gives better convergence for all frequencies higher than the first. Note that, because the computed natural frequencies provide upper bounds for the actual natural frequencies, lower estimates imply more accurate estimates. To explain the preceding results, we observe that the satisfaction of the natural boundary condition at $x = L$ tends to affect mostly the low modes and loses importance as the mode number increases, as the end $x = L$ acts more like a free end for the higher modes. This tends to give the edge to the FEM for the lower modes. This advantage disappears for higher modes, where the CRR has the edge, because the admissible functions used in the CRR are smoother than those used in the FEM and, hence, are better able to approximate the differential equation. As it turns out, convergence is relatively poor both for the CRR method using the admissible functions given by Eqs. (34) and for the FEM using linear interpolation functions. In the next section, we look into the reasons for these poor convergence characteristics and propose ways for accelerating the convergence.

Convergence Considerations

As can be verified from Table 1, convergence of the classical Rayleigh-Ritz method in conjunction with the admissible functions representing the eigenfunctions of a uniform fixed-free rod is relatively slow. Convergence of the finite element method using linear interpolation functions is slightly better for the first natural frequency, but it worsens progressively for higher natural frequencies. In this section, we seek an explanation for these poor results and propose ways for improving them.

In seeking to explain the results obtained in the preceding section, we wish to look first into the nature of the approximat solution. In the Rayleigh-Ritz method—both the classical and the finite element methods—we construct an approximate solution as a linear combination of admissible functions, i.e., functions that are complete in energy. Such functions satisfy the geometric boundary conditions by definition, but the question remains as to how well they approximate the differential equation and the natural boundary conditions. In this regard, the concept of completeness offers no help. Indeed, completeness is a mathematical concept that is qualitative in nature. It states what happens to the error caused by the approximation only as n becomes large. From a computational point of view, the concept of completeness is of dubious value, as the whole object of an approximate technique is to produce sufficient accuracy while the number of terms in the approximation is still relatively small. To bring this point into sharper focus, we consider Eqs. (17) and (34) and write

$$\frac{\mathrm{d}W^n}{\mathrm{d}x} = \sum_{i=1}^{n} u_i \frac{(2i-1)\pi}{2L} \cos\frac{(2i-1)\pi x}{2L} \tag{42}$$

and observe that every term in the series is zero at $x = L$, except when i approaches infinity. The implication is that the slope at $x = L$ is zero for finite n, so that the natural boundary condition, Eq. (28b), cannot be satisfied exactly for finite n.

As far as the approximate solution obtained in the preceding section by means of the finite element method in conjunction with linear interpolation functions, we conclude that the displacement is approximated by a curve consisting of straight lines, and hence displaying a corner at every nodal point. However, the actual displacement curve must be smooth.

Table 1 First, second, and third natural frequencies computed by the classical Rayleigh-Ritz method (CRR) and the finite element method (FEM)

n	ω_1^n		ω_2^n		ω_3^n	
	CRR	FEM	CRR	FEM	CRR	FEM
1	2.32965	2.67261	—	—	—	—
2	2.27291	2.32551	5.13905	6.27163	—	—
3	2.25352	2.26469	5.12823	5.68345	8.13148	9.88405
4	2.24369	2.24326	5.12158	5.43128	8.13028	9.41491
5	2.23781	2.23330	5.11727	5.31181	8.12835	8.98398
6	2.23391	2.22788	5.11430	5.24676	8.12667	8.72417
7	2.23115	2.22461	5.11215	5.20758	8.12533	8.56352
8	2.22910	2.22248	5.11053	5.18218	8.12426	8.45855
9	2.22751	2.22102	5.10926	5.16479	8.12340	8.38649
10	2.22625	2.21998	5.10826	5.15237	8.12270	8.33496
11	2.22523	2.21921	5.10743	5.14318	8.12212	8.29688
12	2.22439	2.21862	5.10675	5.13620	8.12163	8.26793
13	2.22367	2.21816	5.10617	5.13076	8.12121	8.24543
14	2.22307	2.21780	5.10568	5.12646	8.12086	8.22759
15	2.22254	2.21751	5.10525	5.12298	8.12055	8.21321
16	2.22209	2.21727	5.10488	5.12014	8.12028	8.20145
17	2.22169	2.21707	5.10456	5.11778	8.12004	8.19171
18	2.22133	2.21690	5.10427	5.11581	8.11983	8.18355
19	2.22101	2.21676	5.10401	5.11414	8.11964	8.17664
20	2.22073	2.21664	5.10378	5.11271	8.11947	8.17075
21	2.22047	2.21653	5.10357	5.11149	8.11932	8.16568
22	2.22024	2.21645	5.10338	5.11042	8.11918	8.16129
23	2.22003	2.21637	5.10321	5.10950	8.11905	8.15746
24	2.21983	2.21630	5.10305	5.10868	8.11893	8.15410
25	2.21966	2.21624	5.10290	5.10796	8.11882	8.15114
26	2.21949	2.21618	5.10277	5.10733	8.11873	8.14850
27	2.21934	2.21614	5.10264	5.10676	8.11863	8.14616
28	2.21920	2.21609	5.10253	5.10625	8.11855	8.14407
29	2.21907	2.21605	5.10242	5.10580	8.11847	8.14218
30	2.21895	2.21602	5.10232	5.10539	8.11840	8.14049

Therefore, the differential equation cannnot be satisfied exactly for finite n.

The preceding discussion points toward some ways of improving the approximations. In the classical Rayleigh-Ritz method, the need is for a better approximate satisfaction of the natural boundary condition, Eq. (28b). To this end, we consider first satisfying the natural boundary condition exactly by using the comparison functions

$$\phi_i = \sin\beta_i x, \qquad i = 1,2,...,n \tag{43}$$

instead of admissible functions, where β_i are such that Eq. (28b) is satisfied, or

$$EA(L)\beta_i \cos\beta_i L + k \sin\beta_i L = 0, \qquad i = 1,2,...,n \tag{44}$$

The first three natural frequencies computed by the classical Rayleigh-Ritz method using the comparison functions defined by Eqs. (43) and (44) are displayed in Table 2 for various values of n. As could be anticipated, the convergence characteristics are markedly superior to those based on the admissible functions $\sin(2i - 1)\pi x/2L$.

Improved finite element approximations can be obtained by using higher-degree interpolation functions,[1] which requires the introduction of internal nodes. Quadratic interpolation functions require one internal node per finite element, so that the nodal coordinate vector and the vector of interpolation functions have the form

$$\boldsymbol{W}_j = \begin{bmatrix} W_{j-1} \\ W_{j-\frac{1}{2}} \\ W_j \end{bmatrix} \tag{45a}$$

$$\boldsymbol{\phi}(x) = \begin{bmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{bmatrix} \tag{45b}$$

where the interpolation functions have the expression

$$\phi_1 = \left(j - \frac{x}{h}\right)\left[2\left(j - \frac{x}{h}\right) - 1\right] \tag{46a}$$

$$\phi_2 = 4\left(j - \frac{x}{h}\right)\left[1 - \left(j - \frac{x}{h}\right)\right] \tag{46b}$$

$$\phi_3 = 1 - 3\left(j - \frac{x}{h}\right) + 2\left(j - \frac{x}{h}\right)^2 \tag{46c}$$

Similarly, cubic interpolation functions require two internal nodes, so that

$$\boldsymbol{W}_j = \begin{bmatrix} W_{j-1} \\ W_{j-\frac{2}{3}} \\ W_{j-\frac{1}{3}} \\ W_j \end{bmatrix} \tag{47a}$$

$$\boldsymbol{\phi}(x) = \begin{bmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \\ \phi_4(x) \end{bmatrix} \tag{47b}$$

where

$$\phi_1 = \frac{1}{2}\left(j - \frac{x}{h}\right)\left[2 - 9\left(j - \frac{x}{h}\right) + 9\left(j - \frac{x}{h}\right)^2\right] \tag{48a}$$

$$\phi_2 = -\frac{9}{2}\left(j - \frac{x}{h}\right)\left[1 - 4\left(j - \frac{x}{h}\right) + 3\left(j - \frac{x}{h}\right)^2\right] \tag{48b}$$

$$\phi_3 = \frac{9}{2}\left(j = \frac{x}{h}\right)\left[2 - 5\left(j - \frac{x}{h}\right) + 3\left(j - \frac{x}{h}\right)^2\right] \tag{48c}$$

$$\phi_4 = 1 - \frac{11}{2}\left(j - \frac{x}{h}\right) + 9\left(j - \frac{x}{h}\right)^2 - \frac{9}{2}\left(j - \frac{x}{h}\right)^3 \tag{48d}$$

Table 2 First three natural frequencies computed by the classical Rayleigh-Ritz method using comparison functions (CRRC) and the finite element method using quadratic (FEMQ) and cubic (FEMC) interpolation functions

	ω_1^n			ω_2^n			ω_3^n		
n	CRRC	FEMQ	FEMC	CCRC	FEMQ	FEMC	CRRC	FEMQ	FEMC
1	2.22297	—	—	—	—	—	—	—	—
2	2.21647	2.23433	—	5.10630	6.27984	—	—	—	—
3	2.21573	—	2.21558	5.10070	—	5.25278	8.12426	—	10.76486
4	2.21559	2.21705	—	5.09984	5.18817	—	8.11790	9.13233	—
5	2.21555	—	—	5.09964	—	—	8.11680	—	—
6	2.21553	2.21584	2.21553	5.09957	5.12170	5.10352	8.11650	8.30344	8.21161
7	2.21553	—	—	5.09955	—	—	8.11640	—	—
8	2.21553	2.21563	—	5.09954	5.10695	—	8.11636	8.18985	—
9	2.21553	—	2.21553	5.09953	—	5.09988	8.11634	—	8.12705
10	2.21553	2.21557	—	5.09953	5.10265	—	8.11633	8.14851	—
11	2.21552	—	—	5.09953	—	—	8.11633	—	—
12		2.21555	2.21552	5.09953	5.10105	5.09959	8.11632	8.13242	8.11835
13		—		5.09953	—	—	8.11632	—	—
14		2.21554		5.09953	5.10036	—	8.11632	8.12520	—
15		—		5.09953	—	5.09954	8.11632	—	8.11688
16		2.21553		5.09953	5.10002	—	8.11632	8.12160	—
17		—		5.09953	—	—	8.11632	—	—
18		2.21553		5.09952	5.09983	5.09953	8.11632	8.11965	8.11651
19		—			—	—	8.11632	—	—
20		2.21553			5.09973	—	8.11632	8.11852	—
21		—			—	5.09953	8.11632	—	8.11640
22		2.21553			5.09966	—	8.11632	8.11783	—
23		—			—	—	8.11632	—	—
24		2.21553			5.09962	5.09953	8.11632	8.11739	8.11635
25		—			—	—	8.11632	—	—
26		2.21553			5.09960	—	8.11632	8.11710	—
27		—			—	5.09953	8.11632	—	8.11634
28		2.21553			5.09958	—	8.11632	8.11690	—
29		—			—	—	8.11632	—	—
30		2.21552			5.09956	5.09953	8.11632	8.11676	8.11633

The first three natural frequencies computed by means of the finite element method using quadratic and cubic interpolation functions are also shown in Table 2.

From Table 2, we conclude that the classical Rayleigh-Ritz method using comparison functions gives somewhat better results than those obtained by means of the finite element method using quadratic interpolation functions and comparable results to those obtained by using cubic interpolation functions. It is clear that the accuracy displayed in Table 2 was never achieved in Table 1.

Quasicomparison Functions: A New Class of Admissible Functions

According to the Rayleigh-Ritz theory, in formulating the algebraic eigenvalue problem by rendering the quotient stationary, where the quotient is in terms of the energy inner product $[W,W]$ instead of the inner product $(W,\mathscr{L}W)$, the approximation can be constructed from the space of admissible functions rather than comparison functions. The main difference between admissible functions and comparison functions lies in the fact that admissible functions need satisfy only the geometric boundary conditions and the comparison functions must satisfy all the boundary conditions. Of course, there is also the question of differentiability, but in the classical Rayleigh-Ritz method this question seldom arises, as the functions used tend to have sufficient smoothness to ensure the existence of derivatives of high order.

In using admissible functions in the case of the system of Fig. 1, the convergence was painfully slow. Consistent with this, satisfaction of the natural boundary condition at $x = L$ was not possible for a finite number of terms in series (17). On the other hand, the use of comparison functions yielded relatively fast convergence. However, the fact that each of the comparison functions must satisfy all the boundary conditions can be quite a burden. Although in the case of the system of Fig. 1 the burden was not particularly heavy, it was still necessary to solve a transcendental equation, Eq. (44). In other cases, such as in two- and three-dimensional structures, the satisfaction of natural boundary conditions can present a serious problem. Hence, the question is whether a way of circumventing this problem exists.

To improve convergence, we must attempt to produce a solution that satisfies the natural boundary conditions as closely as possible short of using comparison functions. To this end, we introduce a new class of admissible functions. This is the class of quasicomparison functions defined as admissible functions of such a nature that finite linear combinations thereof are capable of satisfying the natural boundary conditions as closely as desired. Hence, quasicomparison functions are functions that individually act like admissible functions but behave more like comparison functions when they appear in a finite group. In essence, quasicomparison functions can be regarded as being complete in boundary conditions, in addition to being complete in energy. Clearly, $\sin(2i-1)\pi x/2L$ do not qualify as quasicomparison functions, since no finite linear combination of these functions can produce a nonzero slope at $x = L$.

As a set of quasicomparison functions for the system of Fig. 1, we consider

$$\phi_i(x) = \sin(i\pi x/2L), \qquad i = 1,2,...,n \tag{49}$$

We note that none of the above functions satisfies the natural boundary condition at $x = L$ individually. However, the linear combination

$$W^2 = \sin(\pi x/2L) + c\,\sin(\pi x/L) \tag{50}$$

for example, can be made to satisfy the natural boundary conditions, Eq. (28b), by merely adjusting the constant c. The first three natural frequencies computed by the classical Rayleigh-Ritz method using the quasi-comparison functions given by Eqs. (49) are shown in Table 3. As can be concluded from Table 3, convergence is very rapid.

Table 3 First three natural frequencies computed by the classical Rayleigh-Ritz method using quasi-comparison functions

n	ω_1^n	ω_2^n	ω_3^n
1	2.32965	—	—
2	2.22359	5.98485	—
3	2.21615	5.10007	11.09264
4	2.21557	5.09957	8.15365
5	2.21553	5.09953	8.11632
6	2.21552	5.09952	8.11632
7	2.21552	5.09952	8.11632
8			8.11632
9			8.11632
10			8.11632
11			8.11632
12			8.11632
13			8.11631

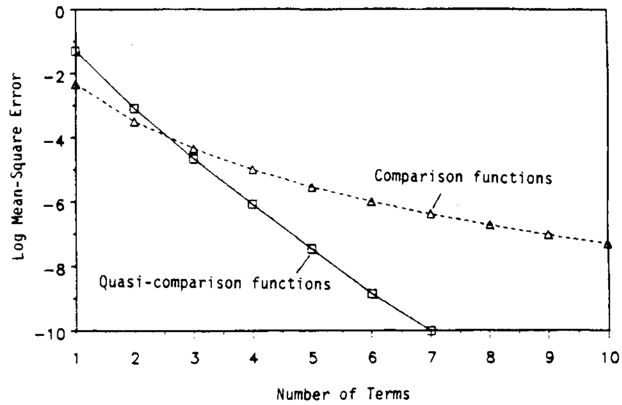

Fig. 2 Logarithm of the mean-square error of first computed eigenfunction.

The interesting thing about the results obtained by means of the quasicomparison functions is that, except for very low n, they are somewhat better than even those obtained by means of comparison functions, as can be concluded by comparing the results in Tables 2 and 3. The explanation is that the quasi-comparison functions are capable of approximating the solution throughout the domain $0 < x < L$ a little better than the comparison functions. As far as the poor results for small n are concerned, it should be pointed out that there is a minimum number of admissible functions required before the linear combinations of admissible functions become quasicomparison functions. Figure 2 shows a plot of the logarithm of the mean-square error of the first computed eigenfunction versus the number of terms for both quasicomparison and comparison functions. The plot shows smaller mean-square error for quasicomparison functions.

In examining Eqs. (34) and (49), we observe that the set of quasi-comparison functions $\sin(i\pi x/2L)$ contains the set of functions $\sin(i\pi x/L)$, in addition to the admissible functions $\sin(2i-1)\pi x/2L$. The functions $\sin(i\pi x/L)$ are zero at $x = L$, but their slope is not. Hence, the addition of the set of functions $\sin(i\pi x/L)$ provides the capability of satisfying the natural boundary conditions at $x = L$. Therefore, one can regard the set of quasi-comparison functions $\sin(i\pi x/2L)$ as being a "more complete set" of admissible functions than the set of admissible functions $\sin(2i-1)\pi x/2L$, in the sense that it can approximate the solution better not only throughout the closed-open domain $0 \leq x \leq L$, but also at the right boundary of the domain.

Summary and Conclusion

In producing an approximate solution of a differential eigenvalue problem by the Rayleigh-Ritz method, it is advantageous to use a variational approach, because it permits the use of admissible functions instead of comparison functions. The admissible functions need satisfy only the geometric boundary conditions and must be from a complete set in energy. In the classical Rayleigh-Ritz method, there is a wide choice of admissible functions. Although completeness in energy of the set of admissible functions guarantees convergence, the convergence rate can be slow. In some cases, the poor convergence can be traced to the inability of satisfying the natural boundary conditions with a finite number of admissible functions.

To accelerate convergence in the classical Rayleigh-Ritz method, a new class of admissible functions is introduced in this paper. This is the class of quasi-comparison functions defined as admissible functions such that, although individually do not satisfy the natural boundary conditions, linear combinations thereof are capable of satisfying them to any desired degree of accuracy. Convergence of the finite element method seems to be affected more by the ability of satisfying the differential equation than the boundary conditions.

A numerical example demonstrates the classical Rayleigh-Ritz method in conjunction with the quasi-comparison function is capable of producing superior results.

Acknowledgments

Sponsored in part by the U.S. Air Force/ASD and Air Force Office of Scientific Research Grant F33615-86-C-3233 monitored by V. B. Venkayya and A. K. Amos, whose support is greatly appreciated.

References

[1]Meirovitch. L., *Computational Methods in Structural Dynamics*, Sijhoff & Noordhoff, Alphen aan den Rijn, the Netherlands, 1980.

[2]Strang, G., and Fix, G. J., *An Analysis of the Finite Element Method,* Prentice-Hall, Englewood Cliffs, NJ, 1973.