



# A NUMERICAL COMPARISON OF ALTERNATIVE GALERKIN METHODS FOR EIGENVALUE ESTIMATION

K.-Y. LEE AND A. A. RENSHAW

*Department of Mechanical Engineering, Columbia University, MIC 4703, New York, NY 10027, USA*

*(Received 26 February 2001, and in final form 10 September 2001)*

This paper provides a concise and unified comparison of four distinct variations of Galerkin's method. The four Galerkin variations are the explicit (traditional), implicit, quadratic implicit, and diagonalized-implicit Galerkin methods. Results indicate that the explicit Galerkin method is superior to all implicit formulations. Among the implicit methods, the quadratic implicit and diagonalized-implicit perform equivalently and are significantly better than the standard implicit method. The explicit method is recommended in all the cases except those in which the number of trial functions is so large that numerical conditioning affects the eigenvalue estimate. In this case, the diagonalized-implicit method is recommended.

© 2002 Elsevier Science Ltd. All rights reserved.

## 1. INTRODUCTION

This paper describes and compares four distinct variations of Galerkin's method for obtaining eigenvalue estimates for continuous eigenvalue problems: the explicit, implicit quadratic-implicit, and diagonalized-implicit methods. The explicit method is the traditional method described in textbooks; reference [1] refers to this method as the configuration space method. The implicit method has been in the literature for a while and has been used by some researchers [1–4]; reference [1] refers to this method as the state-space method. Quadratic formulations of eigenvalues problems have been previously proposed and used to estimate eigenvalues [5, 6]. The diagonalized-implicit method has only recently been formally proposed [7].

The goal of this paper is to provide a side-by-side comparison of each Galerkin procedure using a consistent choice of trial function in order to determine whether one method is to be preferred over the others. In previous investigations [1, 2, 7], a distinction between different Galerkin procedures and alternative choices of trial function has not been maintained, and, as a result, the relative performance of each procedure is not always clear.

This paper provides a concise and unified comparison of the four Galerkin procedures for a series of test problems. The results indicate that the explicit Galerkin method is superior to all implicit formulations. This result was also obtained by Jha and Parker [1]. Among the implicit methods, the quadratic-implicit and diagonalized-implicit perform equally well and are significantly better than the standard implicit method.

## 2. PRELIMINARIES

## 2.1. NUMERICAL CONDITIONING

As motivation, we begin with two simple examples illustrating the effect of numerical conditioning on eigenvalue estimation. Related discussions can be found in textbooks such as in references [8–10].

Consider estimating the eigenvalues,  $\lambda$ , of the damped, axially moving string eigenvalue problem [7],

$$\lambda^2 M g + \lambda C g + D g = 0, \quad (1)$$

where the differential operators  $M$ ,  $C$ , and  $D$  are

$$M = 1, \quad C = 2c\pi + 2v \frac{\partial}{\partial x}, \quad D = (v^2 - 1) \frac{\partial^2}{\partial x^2}, \quad (2)$$

for parameters  $c$  and  $v$ , and  $g(x)$  is the eigenfunction defined on  $0 \leq x \leq 1$  with boundary conditions  $g(0) = g(1) = 0$ . Approximate  $g(x)$  with the linear expansion

$$g = \sum_{n=1}^N c_n u_n(x), \quad (3)$$

where the trial functions are  $u_n(x) = (1-x)x^n$  and the constants  $c_n$  are to be determined. Using Galerkin's method the eigenvalue estimates are given by the  $2N$ -dimensional, matrix eigenvalue problem

$$\lambda \mathbf{A} \mathbf{u} + \mathbf{B} \mathbf{u} = \mathbf{0}, \quad (4)$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{C} & \mathbf{D} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix}, \quad (5)$$

$$\mathbf{M}_{mn} = \langle M u_n, u_m \rangle, \quad \mathbf{C}_{mn} = \langle C u_n, u_m \rangle, \quad \mathbf{D}_{mn} = \langle D u_n, u_m \rangle, \quad (6)$$

where  $\mathbf{I}$  is the  $N \times N$  identity matrix,  $m, n = 1, \dots, N$ , and  $\langle \cdot, \cdot \rangle$  is the inner product associated with equation (1).

This eigenvalue problem is poorly conditioned. The condition number of  $\mathbf{M}$  exceeds  $10^6$  for  $N > 9$  (even sooner than when the inner product is done numerically), which means that inversion of  $\mathbf{M}$  is hazardous using double-precision arithmetic.  $\mathbf{D}$  is also poorly conditioned, so solving for  $\lambda^{-1}$  does not alleviate the problem. Any numerical errors arising from the inversion of  $\mathbf{M}$  will be carried over to any eigenvalue solver that inverts  $\mathbf{A}$  in order to solve equation (4). Example 7.7-2 in Golub and van Loan [8] gives a specific numerical example of this conditioning problem for the generalized eigenvalue problem in which a condition number of  $2 \times 10^6$  leads to a 22% error in the eigenvalue.

Numerical difficulties arising from poor numerical conditioning are normally circumvented by using a generalized eigenvalue routine rather than inverting  $\mathbf{M}$  (using, for instance, Gaussian elimination) and then applying a matrix eigenvalue solver. While the performance differences of these two solution algorithms are known to numerical specialists, they appear to be less well known in the dynamics community. For example, the standard graduate textbook in vibrations by Meirovitch [11] recommends the latter, less numerically stable approach.

Figure 1 compares the absolute error in the fundamental (smallest magnitude) eigenvalue estimate,  $\lambda_1$ , as a function of  $N$  for the damped, axially moving string for the two different solution algorithms: a generalized eigenvalue solver versus matrix inversion

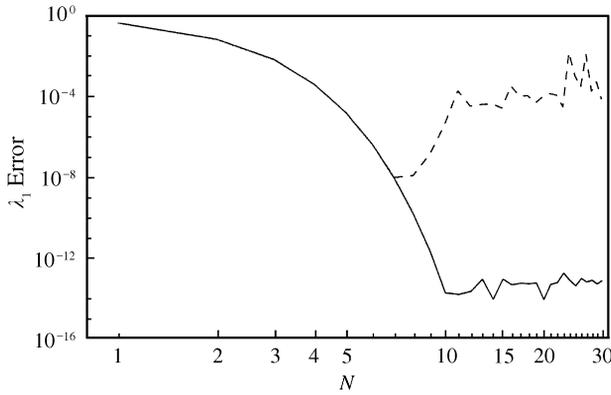


Figure 1. Absolute error of the fundamental eigenvalue estimate using polynomial trial functions for the damped, axially moving string as a function of  $N$ ,  $c = v = 0.5$ : —, generalized eigenvalue solver; ---, matrix inversion followed by a matrix eigenvalue solver.

followed by a matrix eigenvalue solver;  $c = v = 0.5$ . These results were calculated using Matlab. Initially, both estimates are indistinguishable and improve as  $N$  increases. At  $N = 7$ , however, the estimate using matrix inversion degenerates and quickly becomes worthless. For the generalized eigenvalue solver, the error levels off at about  $10^{-15}$  for  $N > 9$ .

As a second example, consider the same eigenvalue estimate using as trial functions the axially moving string eigenfunctions [12]

$$u_n = e^{in\pi vx} \sin(n\pi x) / n\pi \sqrt{1 - v^2}, \tag{7}$$

where now  $n = \pm 1, \pm 2, \pm 3, \dots \pm N$ . In contrast to the polynomial trial functions of the previous example, these trial functions actually are linearly dependent, at least in the limit of  $N \rightarrow \infty$  [13]. As a result, the same degeneracy occurs as in the previous example although at a slower rate (the condition number for  $\mathbf{M}$  is  $10^9$  for  $N = 15$ ). This degeneracy is visible in the results reported in reference [7]. Linearly dependent trial functions such as equation (7) can be used to accurately estimate eigenvalues but care must be taken to avoid numerical artifacts arising from poor numerical conditioning.

### 2.2. POSITIVE-DEFINITE, GYROSCOPIC EIGENSOLUTIONS

The trial functions for the implicit Galerkin method are normally eigensolutions. For simplicity, we restrict ourselves here to positive-definite, gyroscopic eigensolutions. For more general choices, see reference [7]. Consider the eigenvalue problem

$$\beta^2 M\phi + \beta G\phi + K\phi = 0, \tag{8}$$

where  $M$ ,  $G$ , and  $K$  are linear, real, differential operators in some Hilbert space  $H$  with complex inner product  $\langle \cdot, \cdot \rangle$ ,  $M > 0$ ,  $K > 0$ , and  $\phi \in H$ . The operators have the symmetries

$$M^* = M, \quad G^* = -G, \quad K^* = K, \tag{9}$$

where the  $*$  indicates the adjoint. The eigensolutions are  $\{\beta_n, \phi_n\}$  and are numbered such that  $n = \pm 1, \pm 2, \pm 3, \dots$  and  $\{\beta_{-n}, \phi_{-n}\} = \{\bar{\beta}_n, \bar{\phi}_n\}$ , where the overbar denotes complex conjugation.

The eigensolutions of this system have the following properties [12–14]. First, the eigenvalues  $\beta_n$  are all purely imaginary. Second, the operator  $L[\beta] = \beta^2 M + \beta G + K$  is self-adjoint. Third, the adjoint eigensolution is  $\{\beta_n^*, \phi_n^*\} = \{\bar{\beta}_n, \phi_n\} = \{-\beta_n, \phi_n\}$ . Fourth,

the eigensolutions can be normalized such that

$$\beta_n \beta_m \langle M \phi_n, \phi_m \rangle - \langle K \phi_n, \phi_m \rangle = -\delta_{nm}, \tag{10}$$

$$\beta_n (\beta_n + \beta_m) \langle M \phi_n, \phi_m \rangle + \beta_n \langle G \phi_n, \phi_m \rangle = -\delta_{nm}, \tag{11}$$

where  $\delta_{nm}$  is the Kronecker delta. Fifth, the eigensolutions are complete in the following sense. Let  $\{\lambda, g\}$  be an eigensolution. Then,  $g$  can be represented by the expansion

$$g = \sum_{n=\pm 1, \pm 2, \dots}^{\pm \infty} c_n \phi_n, \tag{12}$$

where the constants  $c_n$  satisfy the constraint

$$\lambda \sum_{n=\pm 1, \dots}^{\pm \infty} c_n \phi_n = \sum_{n=\pm 1, \dots}^{\pm \infty} c_n \beta_n \phi_n. \tag{13}$$

This last constraint (13) is the central fact used by the implicit Galerkin method.

### 3. EIGENVALUE PROBLEM FORMULATION

We wish to construct approximate solutions to the eigenvalue problem

$$\lambda^2 M g + \lambda C g + D g = 0 \tag{14}$$

for general, linear, differential operators  $M, C,$  and  $D$  defined on the Hilbert space  $H$  with  $g \in H$ . For clarity, we recast equation (14) in the first order form as

$$\lambda M f + C f + D g = 0, \quad \lambda g - f = 0, \tag{15, 16}$$

where now both  $f, g \in H$ .

Equations (15) and (16) can be written in the state-space form as

$$\lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} + \begin{bmatrix} C & D \\ -I & 0 \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{17}$$

Occasionally, it is convenient to replace the identity operator  $I$  in equation (17) by either  $C$  or  $D$  if either has special symmetry properties [12]; such reformulations are equivalent to equation (17) for both the explicit and implicit Galerkin method. For the explicit Galerkin method, all state-space formulations of equations (15) and (16) are equivalent to equation (17) since both equations (15) and (16) are strictly true. In the implicit Galerkin method, however, equation (16) is *not* explicitly enforced. Consequently, alternative state-space formulations of equations (15) and (16) are *not* necessarily equivalent to equation (17). Here, we consider three alternative formulations:

$$\lambda^2 \begin{bmatrix} 0 & M \\ M & C \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} + \begin{bmatrix} C & D \\ D & 0 \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{18}$$

$$\lambda \begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} + \begin{bmatrix} C & D \\ K & 0 \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{19}$$

$$\lambda \begin{bmatrix} 0 & M \\ M & G \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} + \begin{bmatrix} -M & 0 \\ (C - G) & D \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{20}$$

Formulation (18) recasts the eigenvalue problem in terms of  $\lambda^2$  instead of  $\lambda$ . Similar formulations have been considered by references [5, 6, 15, 16]. This formulation has the

advantage that, for a given number of trial function, twice as many eigenvalue estimates are obtained as are obtained in the linear case. This may be advantageous, but it has the fundamental drawback that it is not clear which of the two possible square roots of  $\lambda^2$  correspond to the original problem. Since the stabilities implied by the two different square roots are different when  $\lambda$  is not purely imaginary, the approximations obtained by equation (18) may be of limited value. The latter two formulations (19) and (20) are motivated by orthogonalities (10) and (11). Formulation (20) was used in reference [7].

#### 4. GALERKIN METHODS

##### 4.1. THE EXPLICIT GALERKIN METHOD

In the standard Galerkin method, herein referred to as the explicit method, both  $f$  and  $g$  are independently expressed as a linear combination of trial functions  $u_n \in H$ .

$$g = \sum_{n=1}^N c_n u_n, \quad f = \sum_{n=1}^N d_n u_n. \tag{21}$$

By inspection, we solve equation (16) by letting  $d_n = \lambda c_n$ . The inner products of equation (15) with  $u_m$  for  $m = 1, \dots, N$  then gives the matrix eigenvalue problem (5) and (6).

When the  $u_n$  include complex conjugate pairs as they do when  $u_n = \phi_n$ , expressions (21) can always be re-written in terms of real trial functions, i.e.,

$$c_1 u_1 + c_2 \bar{u}_1 = (c_1 + c_2) \text{Re}[u_1] + (c_1 - c_2) \text{Im}[u_1] \tag{22}$$

and similarly for  $d_n$ . It is therefore equivalent to use the real trial functions  $\text{Re}[u_1]$ ,  $\text{Im}[u_1]$ , etc. directly. Hence, the explicit Galerkin method always gives a  $2N \times 2N$ , real, generalized eigenvalue problem.

##### 4.2. THE IMPLICIT GALERKIN METHOD

The implicit Galerkin's method proceeds as follows. Use the eigenfunctions  $\phi_n$  of equation (8) as trial functions and let

$$g = \sum_{n=\pm 1}^{\pm N} c_n \phi_n \tag{23}$$

but simply assume

$$f = \sum_{n=\pm 1}^{\pm N} c_n \beta_n \phi_n. \tag{24}$$

In the limit  $N \rightarrow \infty$ , expansion theorem (13) can be used to show convergence. Equation (16) is satisfied only implicitly as  $N \rightarrow \infty$ . With equation (24) assumed, we then apply Galerkin's method to solve either equations (17), (18), (19) or (20). Equation (16) is ignored. (Note: we cannot apply Galerkin's method to solve equation (15); this formulation is degenerate.)

The principal advantage of the implicit method is that the  $d_n$  are eliminated from the problem, reducing the number of unknowns by half. Hence, we can use twice as many trial functions as in the explicit formulation to obtain the same size eigenvalue problem. However, no simple recombination of the trial functions produces expansions for both  $g$  and  $f$  in terms of purely real trial functions. Therefore, except in special cases, the implicit

procedure with expansion (23) leads to a  $2N \times 2N$  complex generalized eigenvalue problem.

We label and detail the matrix eigenvalue problem derived from each of the distinct formulations given above. In each case, the eigenvalue problem takes form (4), and we denote  $\mathbf{M}_{mn} = \langle M\phi_n, \phi_m \rangle$ ,  $\mathbf{C}_{mn} = \langle C\phi_n, \phi_m \rangle$ , etc.

(1) *Implicit formulation (I)*: Taking the inner product of equation (17) with  $[\bar{\beta}_m\phi_m \ \phi_m]^T$  gives

$$\mathbf{A}_{mn} = \beta_m\beta_n\mathbf{M}_{mn} + \mathbf{I}_{mn}, \quad \mathbf{B}_{mn} = \beta_m\beta_n\mathbf{C}_{mn} + \beta_m\mathbf{D}_{mn} - \beta_n\mathbf{I}_{mn}. \quad (25)$$

(2) *Quadratic-implicit formulation (IQ)*: Taking the inner product of equation (18) with  $[\bar{\beta}_m\phi_m \ \phi_m]^T$  gives

$$\mathbf{A}_{mn} = (\beta_m + \beta_n)\mathbf{M}_{mn} + \mathbf{C}_{mn}, \quad \mathbf{B}_{mn} = \beta_m\beta_n\mathbf{C}_{mn} + (\beta_m + \beta_n)\mathbf{D}_{mn}. \quad (26)$$

(3) *Diagonalized-implicit formulation (ID)*: Taking the inner product of equation (19) with  $[\bar{\beta}_m\phi_m \ \phi_m]^T$  and using orthogonality (10) gives

$$\mathbf{A}_{mn} = \beta_m\beta_n\mathbf{M}_{mn} - \mathbf{K}_{mn} = -\delta_{mn}, \quad \mathbf{B}_{mn} = \beta_m\beta_n\mathbf{C}_{mn} + \beta_m\mathbf{D}_{mn} + \beta_n\mathbf{K}_{mn}. \quad (27)$$

(4) *Alternative diagonalized-implicit formulation (ID)*: Taking the inner product of equation (20) with  $[\bar{\beta}_m\phi_m \ \phi_m]^T$  and using orthogonality (11) gives

$$\begin{aligned} \mathbf{A}_{mn} &= (\beta_m + \beta_n)\mathbf{M}_{mn} + \mathbf{G}_{mn} = -\delta_{mn}/\beta_n, \\ \mathbf{B}_{mn} &= \beta_m\beta_n\mathbf{M}_{mn} + \beta_n(\mathbf{C}_{mn} - \mathbf{G}_{mn}) + \mathbf{D}_{mn}. \end{aligned} \quad (28)$$

### 5. CONVERGENCE RESULTS

We examine the convergence of each Galerkin formulation for four test problems, the first three of which are positive-definite, gyroscopic eigenvalue problems. In each case, we plot the absolute error in  $\lambda_1$  and  $\lambda_2$ , the two smallest magnitude eigenvalues with positive imaginary parts, as functions of  $N$ . For the implicit formulation,  $N$  equal to half an odd integer corresponds to keeping one more positively indexed trial function than negative, i.e.,  $N = 1/2$  is a one-term expansion. In this manner, the convergence comparison is made on the basis of the size of the underlying matrix eigenvalue problem. This biases the results in favor of the implicit method since the implicit eigenvalue problem is complex, while the explicit eigenvalue problem is real; nevertheless, the explicit method still outperforms the implicit method in every case.

Since the results of the two implicit diagonalized methods are essentially the same, only the results for the first method are shown.

#### 5.1. EXAMPLE 1: ELASTICALLY SUPPORTED, AXIALLY MOVING STRING

The elastically supported, axially moving string eigenvalue problem is given by

$$M = 1, \quad C = 2v \frac{\partial}{\partial x}, \quad D = (v^2 - 1) \frac{\partial^2}{\partial x^2} + \mu, \quad (29)$$

where  $v$  is the axial velocity,  $\mu$  is the stiffness coefficient of the elastic foundation supporting the string, and  $g(0) = g(1) = 0$  [17]. The eigenvalues are

$$\lambda = in\pi(1 - v^2)\sqrt{1 + \mu/n^2\pi^2(1 - v^2)}. \quad (30)$$

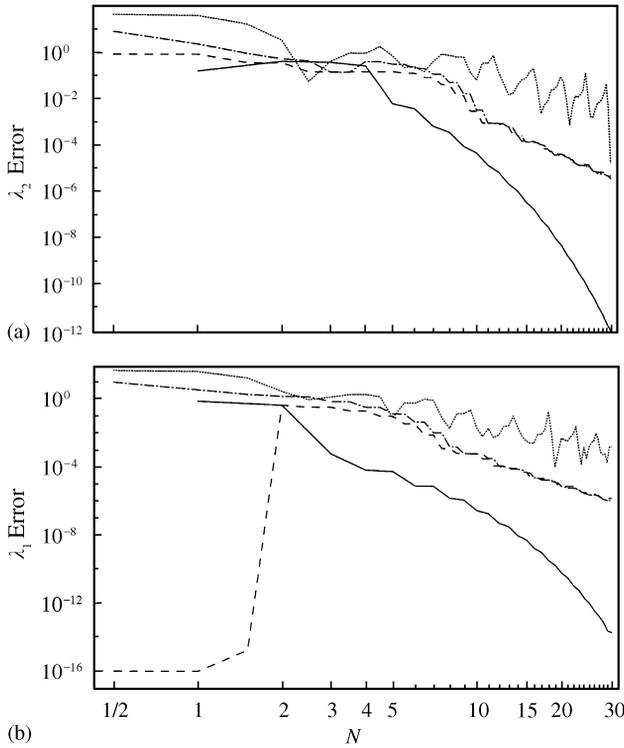


Figure 2. Absolute error in the first two eigenvalues of an axially moving string on an elastic foundation,  $v = 0.5, \mu = 100$ . Methods compared: —, E; ····, I; ----, IQ; -·-·, ID.

For trial functions, we use the axially moving string eigensolutions

$$\beta_n = i(1 - v^2)n\pi, \quad \phi_n(x) = e^{in\pi v x} \sin(n\pi x) / n\pi \sqrt{1 - v^2},$$

$$M = 1, \quad G = 2v \frac{\partial}{\partial x}, \quad K = (v^2 - 1) \frac{\partial^2}{\partial x^2}.$$

Figure 2 shows the absolute error in  $\lambda_1$  and  $\lambda_2$  for the different methods as functions of  $N$  for  $v = 0.5$  and  $\mu = 100$ . For  $N > 4$ , the explicit method is best for both eigenvalues; the next best is either the implicit diagonal or implicit quadratic which converge at the same rate; the slowest convergence is given by the standard implicit method.

This problem is unusual in that the first three estimates of the implicit quadratic method for  $\lambda_1$  are extraordinarily accurate. This artifact is interesting but of limited value since it does not happen in general and cannot be verified without knowing the eigenvalues. It does, however, rather dramatically indicate that convergence is not monotonic for the quadratic formulation.

### 5.2. EXAMPLE 2: AXIALLY MOVING CABLE

The operators of the axially moving cable eigenvalue problem are [18]

$$M = 1, \quad C = 2v \frac{\partial}{\partial x}, \quad D = \frac{\partial}{\partial x} \left( (v^2 - (m^2 + v^2)p(x)) \frac{\partial}{\partial x} \right), \quad (32)$$

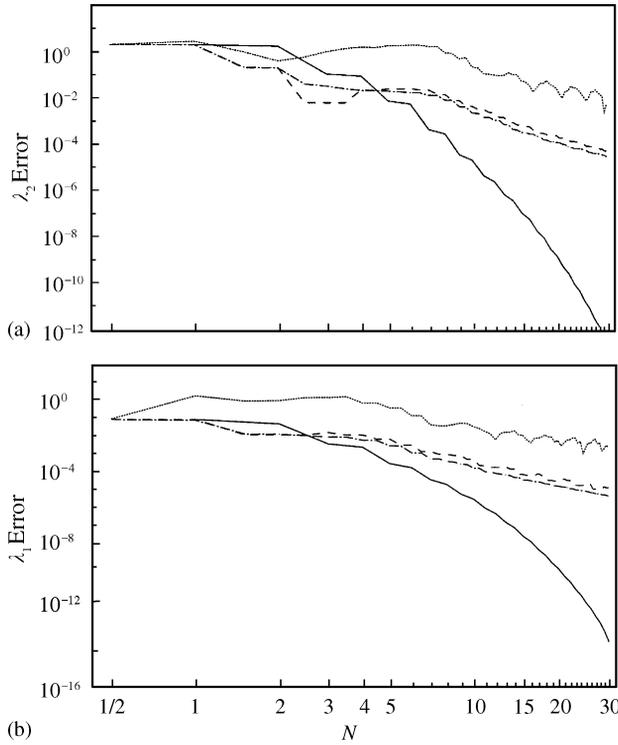


Figure 3. Absolute error in the first two eigenvalues of an axially moving cable,  $v = m^2 = 0.5$ : —, E; ····, I; ----, IQ; - · - ·, ID.

where  $v$  is the axial velocity,  $m^2$  is the dimensionless catenary parameter,  $p(x)$  is the tension distribution

$$p(x) = \left[ v^2 + \sqrt{m^4 + (x - 1)^2} \right] / (m^2 + v^2) \tag{33}$$

and  $g(0) = g(1) = 0$ . For this problem, there is no closed-form expression for the exact eigenvalues. In this case, the “exact” eigenvalues were computed using a shooting method (i.e., the ODE was integrated from  $x = 0$  to 1 adjusting the unknown boundary condition at  $x = 0$  until the boundary condition at  $x = 1$  was satisfied). For  $v = m^2 = 0.5$ , we take the first two eigenvalues to be

$$\lambda_1 = 2.278551952548856i, \quad \lambda_2 = 4.516201249694341i. \tag{34}$$

The axially moving string eigenfunctions (31) are used as trial functions.

Figure 3 shows the absolute error in  $\lambda_1$  and  $\lambda_2$  for the different methods as functions of  $N$  for  $v = m^2 = 0.5$ . As in the first example, for  $N > 4$ , the explicit method is best for both eigenvalues; the next best is either the implicit diagonal or implicit quadratic which converges at about the same rate; the slowest convergence is given by the standard implicit method.

### 5.3. EXAMPLE 3: AXIALLY MOVING BEAM WITH TAPERED STIFFNESS

The operators of the eigenvalue problem of the simply supported, axially moving beam with tapering bending stiffness are

$$M = 1, \quad C = 2v \frac{\partial}{\partial x}, \quad D = \frac{\partial^2}{\partial x^2} \left( (1 - \beta x) \frac{\partial^2}{\partial x^2} \right) + (v^2 - \mu^2) \frac{\partial^2}{\partial x^2}, \tag{35}$$

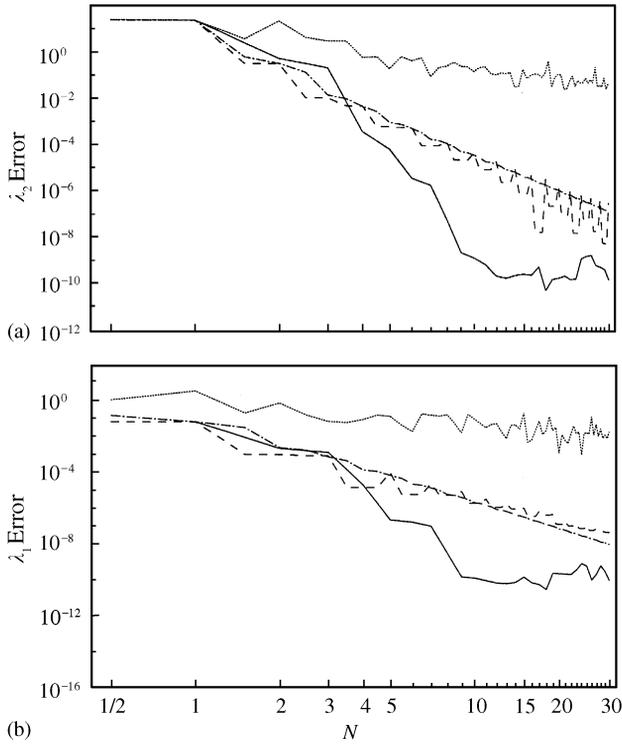


Figure 4. Absolute error in the first two eigenvalues of a tapered axially moving beam,  $v = \beta = 0.5$ ,  $\mu = 1$ : —, E; ···, I; ----, IQ; - · - ·, ID.

where  $v$  is the axial velocity,  $\mu$  is the beam tension, and  $\beta$  is the tapering factor. Simply supported boundary conditions are  $g(0) = g_{,xx}(0) = g(1) = g_{,xx}(1) = 0$ . These operators may be used to model travelling webs in which the bending stiffness changes due to heating and drying as the web crosses the span.

We use the axially moving beam eigenfunctions as trial functions for this problem:

$$G = 2v \frac{\partial}{\partial x}, \quad K = \frac{\partial^4}{\partial x^4} + (v^2 - \mu^2) \frac{\partial^2}{\partial x^2}. \tag{36}$$

The eigenfunctions themselves must be calculated numerically [12]. Appendix A gives a simplified, robust method for this calculation.

Figure 4 shows the absolute error in  $\lambda_1$  and  $\lambda_2$  for the different methods as functions of  $N$  for  $v = \beta = 0.5$  and  $\mu = 1$ . Once again the exact eigenvalues were determined using a shooting method and taken to be

$$\lambda_1 = 8.877457741110753i, \quad \lambda_2 = 34.296542960463400i. \tag{37}$$

As in the previous examples, the explicit method is best; the next best is either the implicit diagonal or implicit quadratic which converges at about the same rate; the slowest convergence is given by the standard implicit method. The final eigenvalue errors level off at about  $10^{-10}$ , which is about five orders of magnitude higher than in the previous examples. This appears to be caused by numerical errors derived from the lack of a closed form expression for the trial functions.

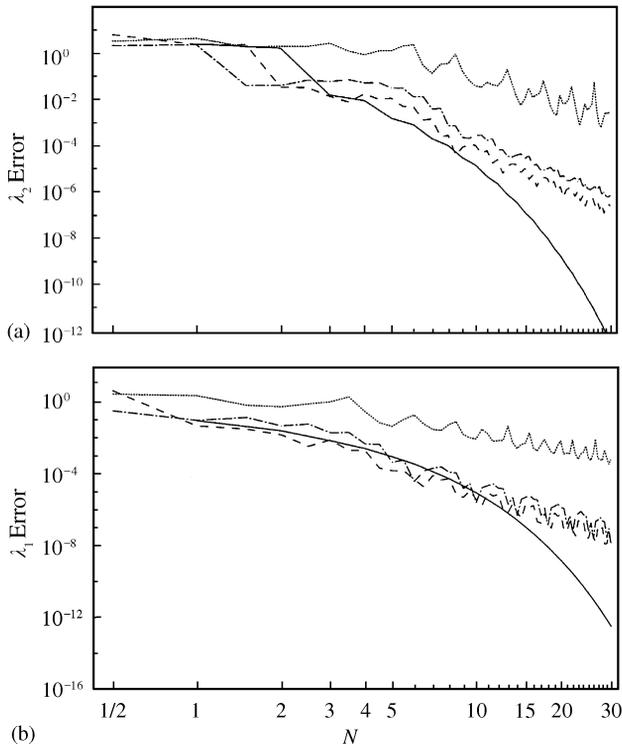


Figure 5. Absolute error in the first two eigenvalues of a damped, axially moving string,  $v = c = 0.5$ : —, E; ···, I; ----, IQ; -·-, ID.

5.4. EXAMPLE 4: DAMPED, AXIALLY MOVING STRING

The operators of the damped, axially moving string eigenvalue problem are

$$M = 1, \quad C = 2c\pi + 2v \frac{\partial}{\partial x}, \quad D = (v^2 - 1) \frac{\partial^2}{\partial x^2}, \tag{38}$$

where  $v$  is the axial velocity,  $c$  is the damping factor and  $g(0) = g(1) = 0$ . The exact eigenvalues are

$$\lambda_n = -(1 - v^2)\pi c + in\pi(1 - v^2)\sqrt{1 - c^2/n^2}. \tag{39}$$

The axially moving string eigenfunctions (31) are used as trial functions.

Figure 5 shows the absolute error in  $\lambda_1$  and  $\lambda_2$  for the different methods as functions of  $N$  for  $v = c = 0.5$ . Unlike the previous examples, in this example the quadratic-implicit and diagonalized-implicit method are competitive with the explicit method for  $N < 12$ . For  $N > 12$ , the explicit method is once again superior. The standard implicit method is significantly worse than the other methods.

6. DISCUSSION

The results demonstrate that the explicit Galerkin method is superior to *all* implicit methods. For very small  $N$  (say  $N < 4$  in the test problems), no method is superior for every problem; for moderate and large  $N$ , the explicit Galerkin method is superior.

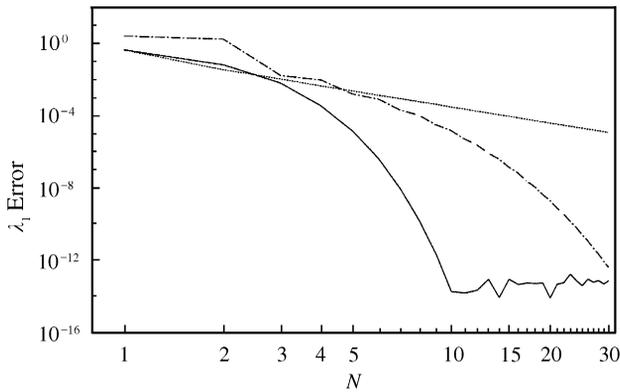


Figure 6. Absolute error in the fundamental eigenvalue of a damped, axially moving string using different trial functions,  $v = c = 0.5$ : —, polynomials; ···,  $\sin(n\pi x)$ ; - - -, axially moving string eigenfunctions. The explicit Galerkin method is used.

These conclusions are supported by the results of other investigations [1, 2, 7]. However, the results reported in these other investigations are not as clear-cut as those presented here because these other investigations have simultaneously examined two related, but independent issues: choice of trial function and trial function independence/numerical conditioning.

A common theme running throughout previous investigations has been the idea that the  $\phi_n$  may be a superior choice for Galerkin trial functions whenever the system to be examined is a “small” perturbation from the underlying positive-definite, gyroscopic system used to define  $\phi_n$ . While this idea is plausible, it is easy to show by counterexample that it is not true in general, at least for large  $N$ . Consider the damped axially moving string problem, solved using the explicit Galerkin method with the following three trial functions: the polynomials  $u_n(x) = (1-x)x^n$ ;  $\sin(n\pi x)$ ; and the axially moving string eigensolutions (31). If the perturbation idea were sound, then the polynomial trial functions would be the worst of these three sets of trial functions since the other two are derived from the perturbations  $c = v = 0$  and  $c = 0$  respectively. In fact, the polynomial trial functions are the best set, as shown in Figure 6 for  $c = v = 0.5$ . Despite the fact that the polynomials possess none of the gyroscopic character or spatial phase distribution of the axially moving string eigenfunctions, they constitute a preferred discretization basis for this particular problem.

For small  $N$  and, in particular,  $N = 1/2$  and 1, both the system perturbation idea and the implicit Galerkin method can sometimes be astonishingly accurate, as illustrated in Figure 2. It is unclear how general or how useful this might be.

The other issue that has arisen in previous investigations has been the numerical conditioning of the eigenvalue problem. We have been able to skirt around this issue here by restricting ourselves to simple problems with simple geometries for which we obtain excellent convergence with just a handful of trial functions. There are many real-world problems in which, due to numerical stiffness or complex geometry, hundreds or thousands of trial functions must be used in order to obtain consistently accurate predictions for the first few eigenvalues of the system. For example, Wickert [3] uses 240 degrees of freedom (80 axially moving string eigenfunctions) to solve the free vibration foil-bearing problem. For problems such as these, the numerical conditioning of the eigenvalue problem is at least as important as any dynamics encoded in the trial functions. (This, of course, is one reason why the finite element method is so successful.) Of the

methods considered here, only the diagonalized-implicit methods remain well conditioned for very large  $N$ . Consequently, the diagonalized-implicit method may be preferable to the explicit method in the case where  $N$  is sufficiently large that numerical conditioning problems hinder accurate, explicit eigenvalue estimation.

Finally, as one reviewer noted, all of our examples are derived from models of axially moving materials, and, as a result, our conclusions have only been tested for this narrow category of problems. We have limited our examples to one-dimensional, positive-definite, gyroscopic systems and compared only the first one or two eigenvalues in order to improve the clarity of the exposition; this naturally limits the scope of our conclusions. Similar conclusions may well apply to more general systems.

## 7. CONCLUSIONS

This paper provides a concise and unified comparison of four distinct variations of Galerkin's method for a series of test problems. The four Galerkin variations are the explicit, implicit, quadratic-implicit, and diagonalized-implicit Galerkin methods. Results indicate that the explicit Galerkin method is superior to all implicit formulations. Among the implicit methods, the quadratic-implicit and diagonalized-implicit perform equally well and are significantly better than the standard implicit method. The explicit method is recommended in all cases except those in which the number of trial functions is so large that numerical conditioning affects the eigenvalue estimation. In this case, the diagonalized-implicit method is recommended.

## REFERENCES

1. R. K. JHA and R. G. PARKER 2000 *Journal of Vibration and Acoustics* **122**, 290–294. Spatial discretization of axially moving media vibration problems.
2. J. A. WICKERT and C. D. MOTE JR 1991 *Applied Mechanics Review* **44**, S279–S284. Response and discretization methods for axially moving materials.
3. J. A. WICKERT 1993 *Journal of Vibration and Acoustics* **115**, 145–151. Free linear vibration of self-pressurized foil bearings.
4. J.-S. CHEN 1997 *Journal of Vibration and Acoustics* **119**, 152–157. Natural frequencies and stability of an axially-travelling string in contact with a stationary load system.
5. K. HUSEYIN 1976 *Journal of Sound and Vibration* **45**, 29–37. Standard forms of the eigenvalue problems associated with gyroscopic systems.
6. R. G. PARKER 1998 *Journal of Applied Mechanics* **65**, 134–140. On the eigenvalues and critical speed stability of gyroscopic continua.
7. K.-Y. LEE and A. A. RENSHAW 2000 *Journal of Vibration and Acoustics* **122**, 464–466. Eigenvalue estimation using linearly dependent eigenfunctions.
8. G. H. GOLUB and C. F. VAN LOAN 1983 *Matrix Computations*. Baltimore: The Johns Hopkins University Press.
9. L. N. TREFETHEN and D. BAU III 1997 *Numerical Linear Algebra*. Philadelphia, PA: SIAM.
10. W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING and B. P. FLANNERY 1992 *Numerical Recipes in C*. Cambridge: Cambridge University Press.
11. L. MEIROVITCH 1967 *Analytical Methods in Vibrations*. New York: Macmillan Publishing, Co.
12. J. A. WICKERT and C. D. MOTE JR 1990 *Journal of Applied Mechanics* **57**, 738–744. Classical vibration analysis of axially moving continua.
13. A. A. RENSHAW 1997 *Journal of Applied Mechanics* **64**, 238–240. Modal decoupling of systems described by three linear operators.
14. R. O. HRYNIV, P. LANCASTER and A. A. RENSHAW 1999 *Journal of Applied Mechanics* **66**, 660–664. A stability criterion for parameter dependent gyroscopic systems.
15. K. HUSEYIN and R. H. PLAUT 1974–1975 *Journal of Structural Mechanics* **3**, 163–177. Transverse vibration and stability of systems with gyroscopic forces.

16. R. H. PLAUT 1976 *American Institute of Aeronautics and Astronautics Journal* **14**, 431–435. Alternative formulations for discrete gyroscopic eigenvalue problems.
17. N. C. PERKINS 1990 *Journal of Vibration and Acoustics* **112**, 2–7. Linear dynamics of a translating string on an elastic foundation.
18. N. C. PERKINS and C. D. MOTE JR 1987 *Journal of Sound and Vibration* **114**, 325–340. Three-dimensional vibration of travelling elastic cables.

APPENDIX A: TENSIONED, AXIALLY MOVING BEAM EIGENFUNCTIONS

Since there is no known analytic formula, the eigensolutions for the tensioned, axially moving beam are computed numerically. Wickert and Mote [12] describe the standard calculation procedure in which the roots of the characteristic equation are found numerically. This method is straightforward but suffers from three defects: (1) it provides no good initial guesses for the roots; (2) once a root is found, additional calculations are required in order to determine its magnitude relative to other roots or the number of nodes in the eigensolution; and (3) the characteristic equation is generally written and evaluated as a complex-valued function even though it is purely imaginary (for the positive-definite case).

Here, we give an equivalent method for determining the simply supported eigensolutions in which each mode is found as a perturbation from the stationary beam eigensolutions. As a result, each of the above defects is corrected.

We parameterize the  $n$ th eigensolution of (36) ( $n = \pm 1, \pm 2, \dots$ ) in terms of two unknowns  $b$  and  $c$  as follows. Let

$$\bar{\mu} = \sqrt{\mu^2 + n^2\pi^2}, \quad a = (\mu^2 + c^2n^2\pi^2 + 2b^2 - v^2)^{1/2}/\bar{\mu}, \quad (\text{A.1, A.2})$$

$$\beta_n = i\omega = \text{isgn}(n)[(nc\pi - b)(nc\pi + b)(a^2\bar{\mu}^2 + b^2)]^{1/2}. \quad (\text{A.3})$$

A multi-dimensional root finder is used to find  $b$  and  $c$  such that both the dispersion relation and the characteristic equation are simultaneously satisfied. The dispersion relation is

$$b(a^2\bar{\mu}^2 + c^2n^2\pi^2) - \omega v = 0 \quad (\text{A.4})$$

and the characteristic equation is

$$\begin{aligned} &\sinh(a\bar{\mu}) \sin(nc\pi)[(a^2\bar{\mu}^2 + 2b^2 + n^2c^2\pi^2)^2 - 4b^2(b^2 + 2n^2c^2\pi^2)] \\ &+ 8ab^2cn\pi\bar{\mu}[\cosh(a\bar{\mu}) \cos(nc\pi) - \cos(2b)] = 0. \end{aligned} \quad (\text{A.5})$$

The unnormalized eigenfunction is

$$\phi_n = \sum_{k=1}^4 c_k e^{\alpha_k x}, \quad (\text{A.6})$$

where

$$\alpha_{1,2} = \pm inc\pi + ib, \quad \alpha_{3,4} = \pm a\bar{\mu} - ib \quad (\text{A.7})$$

and

$$\begin{aligned} c_1 &= 1, \quad c_2\Delta = -\alpha_1^2(e^{\alpha_4} - e^{\alpha_3}) + \alpha_3^2(e^{\alpha_4} - e^{\alpha_1}) - \alpha_4^2(e^{\alpha_3} - e^{\alpha_1}), \\ c_3\Delta &= \alpha_1^2(e^{\alpha_4} - e^{\alpha_2}) - \alpha_2^2(e^{\alpha_4} - e^{\alpha_1}) + \alpha_4^2(e^{\alpha_2} - e^{\alpha_1}), \\ c_4\Delta &= -\alpha_1^2(e^{\alpha_3} - e^{\alpha_2}) + \alpha_2^2(e^{\alpha_3} - e^{\alpha_1}) - \alpha_3^2(e^{\alpha_2} - e^{\alpha_1}), \\ \Delta &= \alpha_2^2(e^{\alpha_4} - e^{\alpha_3}) - \alpha_3^2(e^{\alpha_4} - e^{\alpha_2}) + \alpha_4^2(e^{\alpha_3} - e^{\alpha_2}). \end{aligned} \quad (\text{A.8})$$

The initial guess is based on the stationary beam, simply supported solution given by  $b = 0$  and  $c = 1$ . For the  $n$ th eigensolution, we start the non-linear root finding routine with the guesses

$$b = n\pi v / \sqrt{\mu^2 + 4n^2\pi^2}, \quad c = 1 - \left( \frac{0.2v^2}{n^2\pi^2 + 100\mu^2} \right)^{|n|} \sim 1. \quad (\text{A.9})$$

Numerical experiments suggest that this is an excellent initial guess. Over the range  $n = \pm 1, \dots, \pm 30$ ,  $0 \leq \mu \leq 100$  and  $0 \leq v \leq 0.8\sqrt{\mu^2 + \pi^2}$ , the maximum error in  $b$  and  $c$  is  $< 6\%$ . In all the cases, the multi-dimensional root finder converged quickly and no computational difficulties were found (such as having to take the square root of a negative number).