SIAM J. NUMER. ANAL.
Vol. 9, No. 1, March 1972

HIGHER ORDER CONVERGENCE RESULTS FOR THE RAYLEIGH–RITZ METHOD APPLIED TO EIGENVALUE PROBLEMS. I: ESTIMATES RELATING RAYLEIGH–RITZ AND GALERKIN APPROXIMATIONS TO EIGENFUNCTIONS*

J. G. PIERCE† AND R. S. VARGA‡

Abstract. The application of the Rayleigh–Ritz method for approximating the solutions of linear eigenvalue problems in several dimensions is investigated. The object is to improve upon known error estimates for the approximate eigenfunctions. The L^2 and uniform norm error estimates for the approximate eigenfunctions are related to the corresponding error estimates for the Galerkin approximation of the eigenfunctions. A basic result, independent of dimension, is obtained, which shows that these two approximations are quite close in the relevant energy norm. Then, known results for Galerkin approximation can be directly applied to obtain error estimates for Rayleigh–Ritz approximation.

1. Introduction. The Rayleigh–Ritz method for approximating the eigenvalues and eigenfunctions of the linear eigenvalue problem (cf. (2.1))

$$\mathfrak{N}u(x) = \lambda \mathfrak{M}u(x), \qquad x \in \Omega \subset R^N, \tag{1.1}$$

subject to the homogeneous boundary conditions (cf. (2.2))

$$\mathfrak{B}u(x) = 0, \qquad x \in \partial\Omega, \tag{1.2}$$

has been described extensively in the literature (cf. Collatz [5], Courant and Hilbert [6], Gould [7], Kamke [10], [11] and Mikhlin [12]). Attention has recently been focused on applications of this method to subspaces of piecewise-polynomial functions, such as spline functions. Such techniques provide numerical schemes which are easily adapted for use on high-speed digital computers. Moreover, known error estimates for the approximation of smooth functions in such subspaces can then be directly applied to obtain error estimates for the approximate eigenvalues and eigenfunctions.

The one-dimensional problem has been considered by several authors. In 1965, Wendroff [17] considered the approximation of the eigenvalues and eigenfunctions of Sturm–Liouville problems using subspaces of piecewise linear functions. Birkhoff, de Boor, Swartz and Wendroff [2] were able to improve these results using subspaces of cubic splines, and once-differentiable piecewise cubic polynomials. For computational aspects, see also Johnson [9]. These techniques were then extended to higher order one-dimensional problems and quite general finite-dimensional subspaces by Ciarlet, Schultz and Varga [4]. Specializing to L-spline subspaces, they were then able to obtain high order error estimates for the approximate eigenvalues and eigenfunctions. The results of Birkhoff and de Boor [1] demonstrated that the error bounds of [4] for the approximate *eigenvalues* are best possible in a certain sense. However, as we shall show, the corresponding results for the approximate *eigenfunctions* in the uniform norm are not best possible.

* Received by the editors September 15, 1970.

† Department of Mathematics, University of Southern California, Los Angeles, California 90007. The work of this author was supported in part by the Office of Naval Research under Contract N0014-67-A-0467-0011.

‡ Department of Mathematics, Kent State University, Kent, Ohio 44240. The work of this author was supported in part by the Atomic Energy Commission under Grant (11-1)-2075.

Moreover, the assumption made in [2] and [4] that the eigenvalues are simple will be shown to be unnecessary.

More recently, eigenvalue problems in more than one dimension have been considered by Schultz [15], Strang and Fix [16], and Hald and Widlund [8].

The object of this paper is to relate the L^2 and uniform norm error estimates for the approximate eigenfunctions to the corresponding error estimates for the Galerkin approximation of the eigenfunctions. In § 2, we discuss the relevant theory for the problem (1.1)–(1.2), and in § 3, we briefly describe the Rayleigh–Ritz method. Section 4 contains the proof of a basic result (Theorem 4.2), independent of dimension, demonstrating that the Rayleigh–Ritz approximate eigenfunction and the Galerkin approximation to the eigenfunction are quite close when the corresponding eigenvalue is simple. In the case of multiple eigenvalues, we give a (necessarily) somewhat weakened result in Theorem 4.3. As a direct application of these results, we then obtain a generalization of the basic convergence theorems of Birkhoff, de Boor, Swartz and Wendroff [2], Ciarlet, Schultz and Varga [4] and Schultz [15]. This generalization eliminates the assumption of simple eigenvalues, and indicates how one can obtain improved L^2 and uniform norm error estimates for the approximate eigenfunctions. We shall consider the application of the Rayleigh–Ritz method to subspaces of spline functions, and develop the corresponding error estimates in a subsequent paper (cf. Pierce and Varga [14]).

Throughout this paper, we shall use K to denote a positive generic constant. The parameters which determine K will be made clear in each case.

2. The basic problem. Let Ω be an open subset of R^N, $N \geqq 1$, with boundary $\partial\Omega$. We consider the eigenvalue problem (1.1)–(1.2), where $\mathfrak{N}$ and $\mathfrak{M}$ are the formally self-adjoint elliptic differential operators

$$\mathfrak{N}u(x) \equiv \sum_{|\alpha| \leqq n} (-1)^{|\alpha|} D^\alpha[p_\alpha(x) D^\alpha u(x)], \qquad x \in \Omega,$$
$$\mathfrak{M}u(x) \equiv \sum_{|\alpha| \leqq r} (-1)^{|\alpha|} D^\alpha[q_\alpha(x) D^\alpha u(x)], \qquad x \in \Omega, \tag{2.1}$$

where $0 \leqq r < n$, and where we use the usual multi-index notation. We assume that the coefficient functions $p_\alpha(x)$ and $q_\alpha(x)$ are real-valued functions of class $C^{|\alpha|}(\Omega)$. The homogeneous boundary conditions of (1.2) will consist of n linearly independent conditions of the form

$$B_j u(x) \equiv \sum_{|\alpha| \leqq 2n-1} m_{j,\alpha}(x) D^\alpha u(x) = 0, \qquad 1 \leqq j \leqq n, \quad x \in \partial\Omega. \tag{2.2}$$

In the case $N = 1$ with $\Omega = (a, b)$, we shall also allow the more general (possibly coupled) boundary conditions of the form

$$B_j u(x) \equiv \sum_{k=1}^{2n} \{m_{j,k} D^{k-1} u(a) + n_{j,k} D^{k-1} u(b)\} = 0, \qquad 1 \leqq j \leqq 2n. \tag{2.2'}$$

In addition, we may assume in this case that

$$p_n(x) \text{ and } q_r(x) \text{ do not vanish on } [a, b]. \tag{2.3}$$

Next, let $W^{s,2}(\Omega)$ denote, for $s = 0$, the space $L^2(\Omega)$, and, for $s \geqq 1$, the Sobolev space of all real-valued functions $u(x)$, defined on Ω, such that $D^\alpha u(x) \in L^2(\Omega)$, for all $|\alpha| \leqq s$, where we consider equivalence classes of functions which are equal almost everywhere in Ω. It is well known that $W^{s,2}(\Omega)$ is a Hilbert space with inner product

$$(u, v)_s \equiv \int_\Omega \left\{ \sum_{|\alpha| \leqq s} D^\alpha u(x) D^\alpha v(x) \right\} dx \quad \text{for all } u, v \in W^{s,2}(\Omega),$$

and we take $\|\cdot\|_s$ to denote the corresponding norm. Finally, we shall take $\|u\|_\infty \equiv \sup_{x \in \Omega} |u(x)|$ for all $u(x)$ defined on Ω.

Let $\mathfrak{D}$ be the linear space of all real-valued functions $u(x) \in C^{2n}(\Omega)$ satisfying the boundary conditions of (2.2). We assume that

$$(\mathfrak{N}u, v)_0 = (u, \mathfrak{N}v)_0 \quad \text{for all } u, v \in \mathfrak{D}, \tag{2.4}$$

and that

$$(\mathfrak{M}u, v)_0 = (u, \mathfrak{M}v)_0 \quad \text{for all } u, v \in \mathfrak{D}. \tag{2.5}$$

In addition, we assume that there exist positive constants K such that

$$(\mathfrak{N}u, u)_0 \geqq K(\mathfrak{M}u, u)_0 \quad \text{and} \quad (\mathfrak{M}u, u)_0 \geqq K(u, u)_0 \quad \text{for all } u \in \mathfrak{D}. \tag{2.6}$$

We remark that if

$$(\mathfrak{N}u, u)_0 \geqq K(u, u)_0 \quad \text{and} \quad (\mathfrak{M}u, u)_0 \geqq K(u, u)_0 \quad \text{for all } u \in \mathfrak{D}, \tag{2.6'}$$

then condition (2.6) is valid (cf. Brauer [3], or [4]) after a slight modification of the problem (1.1)–(1.2).

Following Ciarlet, Schultz and Varga [4], we define the inner products

$$(u, v)_D \equiv (\mathfrak{M}u, v)_0 \quad \text{for all } u, v \in \mathfrak{D}, \tag{2.7}$$

and

$$(u, v)_N \equiv (\mathfrak{N}u, v)_0 \quad \text{for all } u, v \in \mathfrak{D}. \tag{2.8}$$

As a consequence of assumption (2.6), $\|u\|_D \equiv (u, u)_D^{1/2}$, and $\|u\|_N \equiv (u, u)_N^{1/2}$ are norms on $\mathfrak{D}$, and we denote the Hilbert space completions of $\mathfrak{D}$ with respect to $\|\cdot\|_D$ and $\|\cdot\|_N$, respectively, as H_D and H_N. It follows from (2.6) that

$$H_N \subset H_D. \tag{2.9}$$

An *eigenvalue* of (1.1)–(1.2) is a value of λ for which there exists a nontrivial solution, or *eigenfunction* $u(x)$, of (1.1)–(1.2). We now state some basic results guaranteeing the existence of eigenvalues and eigenfunctions of (1.1)–(1.2).

THEOREM 2.1 (Gould [7]). *With the assumptions of* (2.4)–(2.6), *assume that bounded sets in H_N are precompact in H_D. Then, the eigenvalue problem* (1.1)–(1.2) *has countably many real eigenvalues* $0 < \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_k \leqq \lambda_{k+1} \leqq \cdots$, *having no finite limit point, and a corresponding sequence of eigenfunctions* $\{f_j(x)\}_{j=1}^\infty$, *such that*

$$\mathfrak{N}f_j(x) = \lambda_j \mathfrak{M}f_j(x), \qquad j \geqq 1. \tag{2.10}$$

The eigenfunctions can be chosen to be orthonormal in H_D, i.e.,

$$(2.11) \qquad (f_i, f_j)_D = \delta_{i,j} \quad \textit{for all } i, j \geqq 1,$$

and $\{f_j\}_{j=1}^{\infty}$ is complete in H_D.

We remark that the conclusions of Theorem 2.1 are always valid in the case $N = 1$ and $\Omega = (a, b)$ for the more general boundary conditions of (2.2′) provided only that the assumptions (2.3)–(2.6) are satisfied. Moreover, in this case, $f_j \in C^{2n}[a, b]$ for all $j \geqq 1$ (cf. F. Brauer [3] and Kamke [10], [11]).

We shall assume throughout this paper that bounded sets in H_N are precompact in H_D, and hence, that the existence of discrete eigenvalues and the corresponding eigenfunctions is guaranteed. It is then well known (cf. Collatz [5], Gould [7], Mikhlin [12]) that the eigenvalues and eigenfunctions of (1.1)–(1.2) can be characterized as the extreme values and critical points, respectively, of the Rayleigh quotient $R[w]$, where

$$(2.12) \qquad R[w] \equiv \|w\|_N^2 / \|w\|_D^2 \quad \text{for } w \in H_N, \quad w \neq 0.$$

More precisely, for each $k \geqq 1$, we have the following characterizations:

$$\text{(i)} \quad \lambda_k = \min\{R[w] : w \in H_N, w \neq 0, (w, f_l)_D = 0, \; 1 \leqq l \leqq k-1\} = R[f_k],$$

$$\text{(ii)} \quad \lambda_k = \max_{\substack{c_1, c_2, \cdots, c_k \\ \text{not all zero}}} \left\{ R\left[\sum_{i=1}^{k} c_i f_i\right]\right\},$$

$$(2.13) \quad \text{(iii)} \quad \lambda_k = \max_{\substack{v_1(x), \cdots, v_{k-1}(x) \\ \text{linearly independent}}} [\min\{R[w] : w \in H_N, w \neq 0, (w, v_l)_D = 0, \; 1 \leqq l \leqq k-1\}],$$

$$\text{(iv)} \quad \lambda_k = \min_{\substack{v_1(x), \cdots, v_k \in H_N \\ \text{linearly independent}}} \left(\max_{\substack{c_1, \cdots, c_k \\ \text{not all zero}}} R\left[\sum_{i=1}^{k} c_i v_i\right]\right).$$

Moreover, it follows from Theorem 2.1 and (2.9) that $w = \sum_{k=1}^{\infty} (f_k, w)_D f_k$ for all $w \in H_N$, and hence that

$$(2.14) \qquad (f_j, w)_N = \lambda_j (f_j, w)_D, \qquad j \geqq 1, \quad \text{for all } w \in H_N.$$

We remark that, by continuity, the inequalities of (2.6) hold for all $w \in H_N$, i.e.,

$$(2.15) \qquad \|w\|_N^2 \geqq K \|w\|_D^2 \quad \text{for all } w \in H_N,$$

$$(2.16) \qquad \|w\|_D^2 \geqq K \|w\|_0^2 \quad \text{for all } w \in H_N.$$

3. The Rayleigh–Ritz method. We now let S_M be a finite-dimensional subspace of dimension M in H_N. The Rayleigh–Ritz method for computing approximate eigenvalues and eigenfunctions consists of finding the extreme values and critical points, respectively, of the Rayleigh quotient (cf. (2.12)) over S_M.

Let $\{w_i(x)\}_1^M$ be a basis for S_M. Then any function $w(x)$ in S_M can be written as $w(x) = \sum_{i=1}^M u_i w_i(x)$. Hence,

$$(3.1)\quad \begin{aligned} R[w] &= R\left[\sum_{i=1}^{M} u_i w_i(x)\right] = \mathscr{R}[\mathbf{u}] \\ &= \frac{\left\|\sum_{i=1}^{M} u_i w_i(x)\right\|_N^2}{\left\|\sum_{i=1}^{M} u_i w_i(x)\right\|_D^2} = \frac{\mathscr{A}(\mathbf{u})}{\mathscr{B}(\mathbf{u})}, \qquad \mathbf{u} \equiv (u_1, u_2, \cdots, u_M)^T. \end{aligned}$$

The stationary values of $R[w]$ are then found from the equations

$$(3.2)\quad \frac{\delta\mathscr{A}(\mathbf{u})}{\delta u_i} = \lambda\frac{\delta\mathscr{B}(\mathbf{u})}{\delta u_i}, \qquad 1 \leqq i \leqq M,$$

which yield the matrix eigenvalue problem,

$$(3.3)\quad A_M\mathbf{u} = \lambda B_M\mathbf{u}.$$

The $M \times M$ matrices $A_M = (\alpha_{i,j}^M)$ and $B_M = (\beta_{i,j}^M)$ have entries given by

$$(3.4)\quad \alpha_{i,j}^M = (w_i, w_j)_N, \qquad \beta_{i,j}^M = (w_i, w_j)_D, \qquad 1 \leqq i,j \leqq M.$$

From assumptions (2.4)–(2.6) it follows that A_M and B_M are real, symmetric, and positive definite. The matrix eigenvalue problem of (3.3) can thus be written as

$$(3.5)\quad C_M\mathbf{v} = \lambda\mathbf{v},$$

where $C_M \equiv B_M^{-1/2}A_MB_M^{-1/2}$ and $\mathbf{v} \equiv B_M^{1/2}\mathbf{u}$. The matrix C_M is also real, symmetric and positive definite, and hence, (3.5) has M positive eigenvalues $0 < \hat{\lambda}_1 \leqq \cdots \leqq \hat{\lambda}_M$, and M corresponding eigenvectors $\hat{\mathbf{v}}_1, \cdots, \hat{\mathbf{v}}_M$, which can be chosen to be orthonormal, i.e., so that $\hat{\mathbf{v}}_i^T\hat{\mathbf{v}}_j = \delta_{i,j}$, $1 \leqq i,j \leqq M$. We call $\hat{\lambda}_k$, $1 \leqq k \leqq M$, the *approximate eigenvalues*, and $\hat{f}_k(x) \equiv \sum_{i=1}^M \hat{u}_{k,i}w_i(x)$, $1 \leqq k \leqq M$, the corresponding *approximate eigenfunctions*, where $\hat{u}_{k,i}$ is the ith component of $\hat{\mathbf{u}}_k \equiv B^{-1/2}\hat{\mathbf{v}}_k$. Since $(\hat{f}_i, \hat{f}_j)_D = \hat{\mathbf{u}}_i^TB_M\hat{\mathbf{u}}_j = \hat{\mathbf{v}}_i^T\hat{\mathbf{v}}_j = \delta_{i,j}$, the functions $\{\hat{f}_j(x)\}_1^M$ are therefore orthonormal in H_D; i.e.,

$$(3.6)\quad (\hat{f}_i, \hat{f}_j)_D = \delta_{i,j}, \qquad 1 \leqq i,j \leqq M.$$

For each $1 \leqq k \leqq M$, we then have the following analogue of (2.13) as characterizations of $\hat{\lambda}_k$ and $\hat{f}_k(x)$;

$$\text{(i)}\quad \hat{\lambda}_k = R[\hat{f}_k] = \max_{\substack{c_1,\cdots,c_k \\ \text{not all zero}}} R\left[\sum_{i=1}^{k} c_i\hat{f}_i\right],$$

$$(3.7)\quad \text{(ii)}\quad \hat{\lambda}_k = \min\{R[w] : w \in S_M, w \neq 0, (w, \hat{f}_i)_D = 0, \text{for all } 1 \leqq i \leqq k-1\},$$

$$\text{(iii)}\quad \hat{\lambda}_k = \min_{\substack{v_1(x),\cdots,v_k(x)\in S_M \\ \text{linearly independent}}} \left(\max_{\substack{c_1,\cdots,c_k \\ \text{not all zero}}} R\left[\sum_{i=1}^{k} c_iv_i\right]\right).$$

It is clear from (2.13(iv)) and (3.7(iii)) that, for any subspace S_M,

$$(3.8)\quad \lambda_k \leqq \hat{\lambda}_k \quad \text{for all } 1 \leqq k \leqq M.$$

Finally, since $A_M\hat{\mathbf{u}}_k = \hat{\lambda}_k B_M \hat{\mathbf{u}}_k$, $1 \leqq k \leqq M$, it follows that

$$\sum_{j=1}^{M} (w_i, w_j)_N \hat{u}_{k,j} = \hat{\lambda}_k \sum_{j=1}^{M} (w_i, w_j)_D \hat{u}_{k,j}, \qquad 1 \leqq i \leqq M, \quad 1 \leqq k \leqq M.$$

Thus, since $\hat{f}_k(x) = \sum_{j=1}^{M} \hat{u}_{k,j} w_j(x)$,

$$(\hat{f}_k, w_i)_N = \hat{\lambda}_k (\hat{f}_k, w_i)_D, \qquad 1 \leqq i \leqq M, \quad 1 \leqq k \leqq M. \tag{3.9}$$

Because (3.9) is valid for the basis elements of S_M, it is therefore valid for all elements of S_M, i.e.,

$$(\hat{f}_k, w)_N - \hat{\lambda}_k (\hat{f}_k, w)_D = 0 \quad \text{for all } w \in S_M \quad \text{and all } 1 \leqq k \leqq M. \tag{3.10}$$

4. Convergence results. With the assumptions of (2.4)–(2.6), let j be a fixed positive integer, and $\{S_{M_t}\}_{t=1}^{\infty}$ be a given (not necessarily nested) sequence of finite-dimensional subspaces of H_N, with $\dim S_{M_t} \equiv M_t \geqq j$ for all $t \geqq 1$. The Rayleigh–Ritz method applied to S_{M_t} yields M_t approximate eigenvalues $\{\hat{\lambda}_{k,t}\}_{k=1}^{M_t}$ and M_t approximate eigenfunctions $\{\hat{f}_{k,t}(x)\}_{k=1}^{M_t}$, which are chosen to be orthonormal in the D-norm, i.e., $(\hat{f}_{i,t}, \hat{f}_{j,t})_D = \delta_{i,j}$, $1 \leqq i, j \leqq M_t$. Let λ_j be the jth eigenvalue, and $f_j(x)$ a corresponding eigenfunction of (1.1)–(1.2). We are then interested in demonstrating convergence of $\hat{\lambda}_{j,t}$ to λ_j, and $\hat{f}_{j,t}(x)$ to $f_j(x)$ under suitable assumptions on the asymptotic properties of the sequence $\{S_{M_t}\}_{t=1}^{\infty}$.

We state the result of [2] relating to eigenvalue convergence.

THEOREM 4.1. *If* $\lim_{t\to\infty} \{\inf_{w \in S_{M_t}} \|w - f_k\|_N\} = 0$ *for each* $1 \leqq k \leqq j$, *then the sequence* $\{\hat{\lambda}_{j,t}\}_{t=1}^{\infty}$ *converges to* λ_k *(from above). In fact, for all t sufficiently large (say* $t \geqq t_j$*), there exist j functions* $\{\tilde{f}_{k,t}\}_{k=1}^{j}$ *in* S_{M_t} *for which* $\sum_{k=1}^{j} \|\tilde{f}_{k,t} - f_k\|_D^2 < 1$, *and*

$$\lambda_j \leqq \hat{\lambda}_{j,t} \leqq \lambda_j + \frac{\sum_{k=1}^{j} \|\tilde{f}_{k,t} - f_k\|_N^2}{\left(1 - \left(\sum_{k=1}^{j} \|\tilde{f}_{k,t} - f_k\|_D^2\right)^{1/2}\right)^2} \quad \textit{for all } t \geqq t_j. \tag{4.1}$$

The above result is believed to be best possible, in a sense which will be clarified in the subsequent work [14]. The corresponding results developed in [4] for the approximate eigenfunctions, however, can be significantly improved.

We note that the approximate eigenfunction $\hat{f}_{j,t}(x)$ is not uniquely defined, i.e., both $+\hat{f}_{j,t}(x)$ and $-\hat{f}_{j,t}(x)$ are approximate eigenfunctions, as defined above. To expect convergence of the $\{\hat{f}_{j,t}(x)\}_{t=1}^{\infty}$, therefore, we assume that $(\hat{f}_{j,m}, \hat{f}_{j,n})_D \geqq 0$ for all $m, n \geqq 1$. This assumption is easily satisfied by multiplying $\hat{f}_{j,t}(x)$ by ± 1, accordingly. We assume initially that λ_j is a simple eigenvalue of (1.1)–(1.2), i.e., that $\lambda_{j-1} < \lambda_j < \lambda_{j+1}$ if $j > 1$, and $\lambda_1 < \lambda_2$ if $j = 1$. If, in fact, λ_j is a multiple eigenvalue of (1.1)–(1.2), a (necessarily) somewhat weaker result will be obtained.

For each $t \geqq 1$, let $\bar{f}_{j,t}(x)$ be the N-norm projection of $f_j(x)$ on the subspace S_{M_t}, i.e.,

$$(f_j - \bar{f}_{j,t}, w)_N = 0 \quad \text{for all } w \in S_{M_t}. \tag{4.2}$$

Since S_{M_t} is a finite-dimensional subspace of the Hilbert space H_N, such an element $\bar{f}_{j,t}$ always exists, and is in fact unique. Equivalently, we have that

$$\|\bar{f}_{j,t} - f_j\|_N = \inf_{w \in S_{M_t}} \|w - f_j\|_N. \tag{4.3}$$

We remark that $\bar{f}_{j,t}$ can be alternatively viewed as the *Galerkin approximation* on S_{M_t} to the solution of the boundary value problem $\mathfrak{N}u = \mathfrak{N}f_j$ with the boundary conditions of (1.2). The main result of this section is the following theorem.

THEOREM 4.2. *With the assumptions of* (2.4)–(2.6), *let* λ_j *and* $f_j(x)$ *be the j-th eigenvalue and eigenfunction, respectively, of* (1.1)–(1.2). *Let* $\{S_{M_t}\}_{t=1}^{\infty}$ *be any (not necessarily nested) sequence of finite-dimensional subspaces of* H_N, *such that* $\dim S_{M_t} \equiv M_t \geqq j$ *for all* $t \geqq 1$, *and such that*

$$\lim_{t \to \infty} \{\inf_{w \in S_{M_t}} \|w - f_k\|_N\} = 0 \quad \textit{for all } l \leqq k \leqq j. \tag{4.4}$$

For each $t \geqq l$, *let* $\hat{f}_{j,t}(x)$ *be the j-th approximate eigenfunction of* (1.1)–(1.2), *obtained by an application of the Rayleigh–Ritz method to* S_{M_t}, *and let* $\bar{f}_{j,t}(x)$ *be the N-norm projection of* $f_j(x)$ *onto* S_{M_t}. *Suppose* λ_j *is a simple eigenvalue. Then, there exists a positive integer* t_j *and a positive constant K, depending only on j, such that*

$$\|\bar{f}_{j,t} - \hat{f}_{j,t}\|_N \leqq K\|f_j - \bar{f}_{j,t}\|_D \quad \textit{for all } t \geqq t_j. \tag{4.5}$$

Since λ_j is assumed to be a simple eigenvalue, as a consequence of Theorem 4.1, (2.15) and the assumption of (4.4), there exists a positive integer t_j, depending only on j, such that for all $t \geqq t_j$,

$$\lambda_j - \hat{\lambda}_{j-1,t} \geqq (\lambda_j - \lambda_{j-1})/2 > 0 \quad \text{if } j > 1, \tag{4.6}$$

$$\lambda_{j+1} - \hat{\lambda}_{j,t} > 0, \tag{4.7}$$

and

$$\|\bar{f}_{j,t} - f_j\|_D < 1/2 \quad \text{and} \quad \|\bar{f}_{j,t} - f_j\|_N < 1. \tag{4.8}$$

Making use of the fact that the $\{\hat{f}_{k,t}(x)\}_{k=1}^{M_t}$ are orthonormal in the D-norm (and, from (3.10), also orthogonal in the N-norm), we expand $\bar{f}_{j,t}$ in the following manner:

$$\bar{f}_{j,t}(x) = \sum_{k=1}^{j-1} (\bar{f}_{j,t}, \hat{f}_{k,t})_D \hat{f}_{k,t}(x) + (\bar{f}_{j,t}, \hat{f}_{j,t})_D \hat{f}_{j,t}(x) + \bar{f}''_{j,t}(x) \quad \text{if } j > 1, \tag{4.9}$$

and

$$\bar{f}_{1,t}(x) = (\bar{f}_{1,t}, f_{1,t})_D \hat{f}_{1,t}(x) + \bar{f}''_{1,t}(x). \tag{4.10}$$

In order to prove Theorem 4.2, we first obtain estimates of

$$\|\bar{f}''_{j,t}\|_N, \quad j \geqq 1, \quad \text{and} \quad \left\| \sum_{k=1}^{j-1} (\bar{f}_{j,t}, \hat{f}_{k,t})_D \hat{f}_{k,t} \right\|_N \quad \text{for } j > 1.$$

LEMMA 4.1. *Suppose* $\{S_{M_t}\}_{t=1}^{\infty}$ *satisfies assumption* (4.4) *and that* $\lambda_j < \lambda_{j+1}$. *Write* $\bar{f}_{j,t}(x)$ *as in* (4.9)–(4.10). *Then, there exists a positive constant K, depending only on j, such that*

$$\|\bar{f}''_{j,t}\|_N^2 \leqq K\|f_j - \bar{f}_{j,t}\|_D^2 \quad \textit{for all } t \geqq 1. \tag{4.11}$$

Proof. We consider the quantity $\|\bar{f}''_{j,t}\|_N^2 - \lambda_j\|\bar{f}''_{j,t}\|_D^2$. From (4.9)–(4.10) and (3.6), we have that $(\bar{f}''_{j,t}, \hat{f}_{k,t})_D = 0, 1 \leqq k \leqq j$. Moreover, by (3.10), $(\bar{f}''_{j,t}, \hat{f}_{k,t})_N = \hat{\lambda}_{k,t}(\bar{f}''_{j,t}, \hat{f}_{k,t})_D = 0, 1 \leqq k \leqq j$. Thus,

$$\|\bar{f}''_{j,t}\|_N^2 - \lambda_j\|\bar{f}''_{j,t}\|_D^2 = (\bar{f}_{j,t}, \bar{f}''_{j,t})_N - \lambda_j(\bar{f}_{j,t}, \bar{f}''_{j,t})_D. \tag{4.12}$$

Since $\bar{f}''_{j,t} \in H_N$, it follows from (2.14) that $(f_j, \bar{f}''_{j,t})_N - \lambda_j(f_j, \bar{f}''_{j,t})_D = 0$. Subtracting this from the right-hand side of (4.12), we obtain

$$\|\bar{f}''_{j,t}\|_N^2 - \lambda_j\|\bar{f}''_{j,t}\|_D^2 = (\bar{f}_{j,t} - f_j, \bar{f}''_{j,t})_N - \lambda_j(\bar{f}_{j,t} - f_j, \bar{f}''_{j,t})_D. \tag{4.13}$$

But since $\bar{f}''_{j,t} \in S_{M_t}$, from (4.2) and the Cauchy–Schwarz inequality,

$$\|\bar{f}''_{j,t}\|_N^2 - \lambda_j\|\bar{f}''_{j,t}\|_D^2 \leqq \lambda_j\|f_j - \bar{f}_{j,t}\|_D\|\bar{f}''_{j,t}\|_D. \tag{4.14}$$

If $\|\bar{f}_{j,t}\|_D \neq 0$, then by (3.7(ii)),

$$\|\bar{f}''_{j,t}\|_N^2 \geqq \hat{\lambda}_{j+1}\|\bar{f}''_{j,t}\|_D^2 \geqq \lambda_{j+1}\|\bar{f}''_{j,t}\|_D^2,$$

and (4.14) becomes

$$\left(\frac{\lambda_{j+1} - \lambda_j}{\lambda_{j+1}}\right)\|\bar{f}''_{j,t}\|_N^2 \leqq \frac{\lambda_j}{\lambda_{j+1}^{1/2}}\|f_j - \bar{f}_{j,t}\|_D\|\bar{f}''_{j,t}\|_N.$$

Since $\lambda_{j+1} > \lambda_j$ by assumption, then either $\|\bar{f}''_{j,t}\|_N = 0$ or

$$\|\bar{f}''_{j,t}\|_N \leqq \left(\frac{\lambda_{j+1}}{\lambda_{j+1} - \lambda_j}\right)\frac{\lambda_j}{\lambda_{j+1}^{1/2}}\|f_j - \bar{f}_{j,t}\|_D, \tag{4.15}$$

and (4.11) follows. If $\|\bar{f}''_{j,t}\|_D = 0$, then $\|\bar{f}''_{j,t}\|_N = 0$ from (4.14), and again (4.11) is valid. This completes the proof.

We now obtain an estimate of $\|\sum_{k=1}^{j-1}(\bar{f}_{j,t}, \hat{f}_{k,t})_D\hat{f}_{k,t}\|_N$ for $j > 1$.

LEMMA 4.2. *Suppose that* $\{S_{M_t}\}_{t=1}^{\infty}$ *satisfies assumption* (4.4) *and that* $j > 1$ *and* $\lambda_j > \lambda_{j-1}$. *Write* $\bar{f}_{j,t}(x)$ *as in* (4.9), *and define*

$$\bar{f}'_{j,t}(x) \equiv \sum_{k=1}^{j-1}(\bar{f}_{j,t}, \hat{f}_{k,t})_D\hat{f}_{k,t}(x). \tag{4.16}$$

Then, there exists a positive constant K, *depending only on* j, *such that*

$$\|\bar{f}'_{j,t}\|_N^2 \leqq K\|f_j - \bar{f}_{j,t}\|_D^2 \quad \textit{for all } t \geqq t_j, \tag{4.17}$$

with t_j *as in* (4.6)–(4.8).

Proof. We consider the quantity $-\|\bar{f}'_{j,t}\|_N^2 + \lambda_j\|\bar{f}'_{j,t}\|_D^2$. From (4.9), (3.6) and (3.10), it follows that

$$(\bar{f}_{j,t}, \bar{f}'_{j,t})_N = \sum_{k=1}^{j-1}\hat{\lambda}_{k,t}(\bar{f}_{j,t}, \hat{f}_{k,t})_D^2 = (\bar{f}'_{j,t}, \bar{f}'_{j,t})_N = \|\bar{f}'_{j,t}\|_N^2,$$

and that

$$(\bar{f}_{j,t}, \bar{f}'_{j,t})_D = \sum_{k=1}^{j-1}(\bar{f}_{j,t}, \hat{f}_{k,t})_D^2 = (\bar{f}'_{j,t}, \bar{f}'_{j,t})_D = \|\bar{f}'_{j,t}\|_D^2.$$

Therefore,

$$-\|\bar{f}'_{j,t}\|_N^2 + \lambda_j\|\bar{f}'_{j,t}\|_D^2 = -(\bar{f}_{j,t}, \bar{f}'_{j,t})_N + \lambda_j(\bar{f}_{j,t}, \bar{f}'_{j,t})_D. \tag{4.18}$$

From (2.14), $(f_j, \bar{f}'_{j,t})_N - \lambda_j(f_j, \bar{f}'_{j,t})_D = 0$ since $\bar{f}'_{j,t} \in H_N$. Adding this to the right-hand side of (4.18) we obtain

$$(4.19) \qquad -\|\bar{f}'_{j,t}\|_N^2 + \lambda_j\|\bar{f}'_{j,t}\|_D^2 = -(\bar{f}_{j,t} - f_j, \bar{f}'_{j,t})_N + \lambda_j(\bar{f}_{j,t} - f_j, \bar{f}'_{j,t})_D.$$

Since $\bar{f}'_{j,t} \in S_{M_t}$, from (4.2) and the Cauchy–Schwarz inequality,

$$(4.20) \qquad -\|\bar{f}'_{j,t}\|_N^2 + \lambda_j\|\bar{f}'_{j,t}\|_D^2 \leqq \lambda_j\|\bar{f}_{j,t} - f_j\|_D\|\bar{f}'_{j,t}\|_D.$$

If $\|\bar{f}'_{j,t}\|_D \neq 0$, then from (3.7(i)),

$$(4.21) \qquad \|\bar{f}'_{j,t}\|_N^2 \leqq \hat{\lambda}_{j-1}\|\bar{f}'_{j,t}\|_D^2,$$

and using (2.15), equation (4.20) becomes

$$\left(\frac{\lambda_j - \hat{\lambda}_{j-1}}{\hat{\lambda}_{j-1}}\right)\|\bar{f}'_{j,t}\|_N^2 \leqq \frac{\lambda_j}{K^{1/2}}\|\bar{f}_{j,t} - f_j\|_D\|\bar{f}'_{j,t}\|_N.$$

Since from (4.6), $\lambda_j - \hat{\lambda}_{j-1} \geqq (\lambda_j - \lambda_{j-1})/2 > 0$ for all $t \geqq t_j$, then

$$\|\bar{f}'_{j,t}\|_N^2 \leqq \left(\frac{2\lambda_j}{\lambda_j - \lambda_{j-1}}\right)\frac{\lambda_j}{K^{1/2}}\|f_j - \bar{f}_{j,t}\|_D\|\bar{f}'_{j,t}\|_N,$$

from which (4.17) easily follows. If $\|\bar{f}'_{j,t}\|_D = 0$, then (4.19) implies that $\|\bar{f}'_{j,t}\|_N = 0$, and (4.17) again is valid. This completes the proof.

The jth eigenfunction of (1.1)–(1.2) is not uniquely defined, i.e., both $+f_j(x)$ and $-f_j(x)$ are eigenfunctions corresponding to the eigenvalue λ_j. With the previous assumptions on the $\{\hat{f}_{j,t}\}_{t=1}^{\infty}$, therefore, we assume that the jth eigenfunction $f_j(x)$ is selected so that $(f_j, \hat{f}_{j,t})_D > 0$ for all t sufficiently large. To show this normalization can be affected, it suffices to show that

$$\lim_{t\to\infty} \|f_j - (f_j, \hat{f}_{j,t})_D\hat{f}_{j,t}\|_D = \lim_{t\to\infty} \|-f_j - (-f_j, \hat{f}_{j,t})_D\hat{f}_{j,t}\|_D = 0.$$

But this fact is a rather easy consequence of the results of Lemmas 4.1 and 4.2 and the assumption of (4.4) (cf. Pierce [13]). It then follows that $(f_j, \hat{f}_{j,t})_D \to 1$ as $t \to \infty$, and we redefine the positive t_j's such that (4.6)–(4.8) are satisfied, and that

$$(4.22) \qquad (\bar{f}_j, \hat{f}_{j,t})_D > 1/2 \quad \text{for all } t \geqq t_j.$$

We now proceed with the proof of Theorem 4.2.

Proof of Theorem 4.2. By the triangle inequality,

$$(4.23) \qquad \|\bar{f}_{j,t} - \hat{f}_{j,t}\|_N \leqq \left\|\frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} - \hat{f}_{j,t}\right\|_N + \left\|\bar{f}_{j,t} - \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D}\right\|_N.$$

We write

$$(4.24) \qquad \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} = \left(\frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D}, \hat{f}_{j,t}\right)_D \hat{f}_{j,t} + e_{j,t} \quad \text{for all } t \geqq t_j,$$

where, from (4.9) and (4.10),

$$(4.25) \qquad e_{j,t} \equiv \frac{1}{\|\bar{f}_{j,t}\|_D}\left[\sum_{k=1}^{j-1} (\bar{f}_{j,t}, \hat{f}_{k,t})_D\hat{f}_{k,t} + \bar{f}''_{j,t}\right] \quad \text{if } j > 1,$$

and

$$e_{1,t} \equiv \bar{f}''_{1,t}/\|\bar{f}_{1,t}\|_D. \tag{4.26}$$

From (3.10) and (4.24), $(\hat{f}_{j,t}, e_{j,t})_N = \hat{\lambda}_{j,t}(\hat{f}_{j,t}, e_{j,t})_D = 0$ and therefore,

$$\begin{aligned} \left\| \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} - \hat{f}_{j,t} \right\|_N^2 &= \left\| \left(\frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D}, \hat{f}_{j,t} \right)_D \hat{f}_{j,t} - \hat{f}_{j,t} \right\|_N^2 + \|e_{j,t}\|_N^2 \\ &= \hat{\lambda}_{j,t}\left[1 - \left(\frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D}, \hat{f}_{j,t} \right)_D \right]^2 + \|e_{j,t}\|_N^2. \end{aligned} \tag{4.27}$$

Now, by (4.24),

$$\left\| \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} \right\|_D^2 = 1 = \left(\frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D}, \hat{f}_{j,t} \right)_D^2 + \|e_{j,t}\|_D^2, \tag{4.28}$$

and therefore,

$$\left(\frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D}, \hat{f}_{j,t} \right)_D = \pm[1 - \|e_{j,t}\|_D^2]^{1/2}. \tag{4.29}$$

It is easily shown that $(\bar{f}_{j,t}/\|\bar{f}_{j,t}\|_D, \hat{f}_{j,t})_D > 0$ for all $t \geqq t_j$ as a consequence of (4.22). Hence, (4.27) becomes

$$\left\| \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} - \hat{f}_{j,t} \right\|_N^2 = [1 - [1 - \|e_{j,t}\|_D^2]^{1/2}]^2 \hat{\lambda}_{j,t} + \|e_{j,t}\|_N^2 \quad \text{for all } t \geqq t_j. \tag{4.30}$$

However $[1 - \sqrt{1 - x}]^2 \leqq x$ for $0 \leqq x \leqq 1$, and since from (4.28), $0 \leqq \|e_{j,t}\|_D^2 \leqq 1$, then

$$\left\| \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} - \hat{f}_{j,t} \right\|_N^2 \leqq \hat{\lambda}_{j,t}\|e_{j,t}\|_D^2 + \|e_{j,t}\|_N^2.$$

By (2.15) and (4.7), therefore,

$$\left\| \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} - \hat{f}_{j,t} \right\|_D^2 \leqq \left[\frac{\lambda_{j+1}}{K} + 1 \right] \|e_{j,t}\|_N^2. \tag{4.31}$$

It follows from (4.9), (4.10), (3.6) and (3.10) that

$$\|e_{j,t}\|_N^2 = \frac{1}{\|\bar{f}_{j,t}\|_D^2} \left[\left\| \sum_{k=1}^{j-1} (\bar{f}_{j,t}, \hat{f}_{k,t})_D \hat{f}_{k,t} \right\|_N^2 + \|\bar{f}''_{j,t}\|_N^2 \right] \quad \text{if } j > 1,$$

and that

$$\|e_{1,t}\|_N^2 = \frac{1}{\|\bar{f}_{1,t}\|_D^2} \|\bar{f}''_{1,t}\|_N^2.$$

From Lemmas 4.1 and 4.2, therefore,

$$\|e_{j,t}\|_N^2 \leqq \frac{K}{\|\bar{f}_{j,t}\|_D^2} \|f_j - \bar{f}_{j,t}\|_D^2 \quad \text{for all } t \geqq t_j.$$

Equation (4.31) then becomes

$$(4.32)\qquad \begin{aligned} \left\| \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} - \hat{f}_{j,t} \right\|_N &\leqq \frac{1}{\|\bar{f}_{j,t}\|_D}\left[K\left(\frac{\lambda_{j+1}}{K} + 1\right)\right]^{1/2} \|f_j - \bar{f}_{j,t}\|_D \\ &\equiv \frac{K}{\|\bar{f}_{j,t}\|_D}\|f_j - \bar{f}_{j,t}\|_D \quad \text{for all } t \geqq t_j. \end{aligned}$$

Now,

$$(4.33)\qquad \left\| \bar{f}_{j,t} - \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} \right\|_N = \frac{|\,\|\bar{f}_{j,t}\|_D - 1|}{\|\bar{f}_{j,t}\|_D}\|\bar{f}_{j,t}\|_N.$$

From the triangle inequality, (4.8) and the fact that $\|f_j\|_N^2 = \lambda_j$,

$$\|\bar{f}_{j,t}\|_N \leqq \|f_j\|_N + \|f_j - \bar{f}_{j,t}\|_N \leqq \lambda_j^{1/2} + 1 \quad \text{for all } t \geqq t_j.$$

Moreover,

$$|\,\|\bar{f}_{j,t}\|_D - 1| = |\,\|\bar{f}_{j,t}\|_D - \|f_j\|_D| \leqq \|f_j - \bar{f}_{j,t}\|_D.$$

Equation (4.33) then becomes

$$(4.34)\qquad \left\| \bar{f}_{j,t} - \frac{\bar{f}_{j,t}}{\|\bar{f}_{j,t}\|_D} \right\|_N \leqq \left(\frac{\lambda_j^{1/2} + 1}{\|\bar{f}_{j,t}\|_D}\right)\|f_j - \bar{f}_{j,t}\|_D \quad \text{for all } t \geqq t_j.$$

But

$$(4.35)\qquad \frac{1}{\|\bar{f}_{j,t}\|_D} \leqq \frac{1}{|\,\|f_j\|_D - \|f_j - \bar{f}_{j,t}\|_D|} \leqq 2 \quad \text{for all } t \geqq t_j.$$

Finally, using (4.32), (4.34), and (4.35) in (4.23), we obtain

$$(4.36)\qquad \|\bar{f}_{j,t} - \hat{f}_{j,t}\|_N \leqq 2(K + \lambda_j^{1/2} + 1)\|f_j - \bar{f}_{j,t}\|_D \quad \text{for all } t \geqq t_j,$$

which proves the result of Theorem 4.2.

If λ_j is a multiple eigenvalue of (1.1)–(1.2), the result of Theorem 4.2 must be weakened somewhat. For fixed $j \geqq 1$, let λ_j be an eigenvalue of multiplicity $p + 1$, $p \geqq 1$, of (1.1)–(1.2), i.e., $\lambda_{j-1} < \lambda_j = \lambda_{j+1} = \cdots = \lambda_{j+p} < \lambda_{j+p+1}$, where we set $\lambda_0 \equiv 0$. The associated eigenfunctions $f_j(x), \cdots, f_{j+p}(x)$ can again be chosen to be orthonormal in the D-norm, and span a linear $(p + 1)$-dimensional subspace $\mathfrak{D}_{p+1}$ of $\mathfrak{D}$. Let $f(x)$ be any element of $\mathfrak{D}_{p+1}$. The $f(x) = \sum_{k=j}^{j+p}(f, f_k)_D f_k(x)$, and $f(x)$ is an eigenfunction of (1.1)–(1.2) associated with λ_j. Let $\{\hat{f}_{k,t}\}_{k=j}^{j+p}$ be the corresponding approximate eigenfunctions in S_{M_t}, chosen to be orthonormal in the D-norm, and let $\hat{f}_t(x) \equiv \sum_{k=j}^{j+p}(f, \hat{f}_{k,t})_D \hat{f}_{k,t}(x)$ be the *D-norm projection* of $f(x)$ onto the $(p + 1)$-dimensional subspace of S_{M_t} spanned by the $\{\hat{f}_{k,t}\}_{k=j}^{j+p}$. Finally, let $\bar{f}_t(x)$ be the N-norm projection of $f(x)$ onto S_{M_t}. Then the following result is valid.

THEOREM 4.3. *With the assumptions of* (2.4)–(2.6), *let* λ_j *be an eigenvalue of multiplicity* $p + 1$ *of* (1.1)–(1.2), *and* $f(x)$ *be any associated eigenfunction. Let* $\{S_{M_t}\}_{t=1}^{\infty}$ *be any sequence of finite-dimensional subspaces of* H_N *such that* $\dim S_{M_t} \equiv M_t \geqq j + p$ *for all* $t \geqq 1$, *and such that*

$$(4.37)\qquad \lim_{t\to\infty}\{\inf_{w \in S_{M_t}} \|w - f_k\|_N\} = 0 \quad \textit{for all } 1 \leqq k \leqq j + p.$$

Let $\hat{f}_t(x)$ and $\bar{f}_t(x)$ be as defined above. Then there exists a positive integer t_j and a positive constant K, depending only on j and p, such that

$$\|\bar{f}_t - \hat{f}_t\|_N^2 \leqq K\|f - \bar{f}_t\|_D^2 \quad \textit{for all } t \leqq t_j. \tag{4.38}$$

Proof. Write $\bar{f}_t(x)$ in the form

$$\bar{f}_t(x) = \sum_{k=1}^{j-1} (\bar{f}_t, \hat{f}_{k,t})_D \hat{f}_{k,t}(x) + \sum_{k=j}^{j+p} (\bar{f}_t, \hat{f}_{k,t})_D \hat{f}_{k,t}(x) + \bar{f}_t''(x) \quad \text{if } j > 1, \tag{4.39}$$

and if $j = 1$,

$$\bar{f}_t(x) = \sum_{k=1}^{p+1} (\bar{f}_t, \hat{f}_{k,t})_D \hat{f}_{k,t}(x) + \bar{f}_t''(x). \tag{4.40}$$

Since $\lambda_{j+p} < \lambda_{j+p+1}$ and since $f(x)$ is an eigenfunction of (1.1)–(1.2) corresponding to the eigenvalue λ_{j+p}, it follows as in Lemma 4.1 that

$$\|\bar{f}_t''\|_N^2 \leqq K\|f - \bar{f}_t\|_D^2 \quad \text{for all } t \geqq 1. \tag{4.41}$$

Similarly, since $\lambda_{j-1} < \lambda_j$ for $j > 1$, and $f(x)$ is an eigenfunction of (1.1)–(1.2) corresponding to the eigenvalue λ_j, it follows as in Lemma 4.2 that

$$\left\| \sum_{k=1}^{j-1} (\bar{f}_t, \hat{f}_{k,t})_D \hat{f}_{k,t} \right\|_N^2 \leqq K\|f - \bar{f}_{j,t}\|_D^2 \quad \text{for all } t \geqq t_j, \tag{4.42}$$

with t_j as defined in (4.6)–(4.8).

We now compute from (4.39), (3.10), and the definition of $\hat{f}_t$ that

$$\|\bar{f}_t - \hat{f}_t\|_N^2 = \left\| \sum_{k=1}^{j-1} (\bar{f}_t, \hat{f}_{k,t})_D \hat{f}_{k,t} \right\|_N^2 + \left\| \sum_{k=j}^{j+p} (\bar{f}_t - f, \hat{f}_{k,t})_D \hat{f}_{k,t} \right\|_N^2 + \|\bar{f}_t''\|_N^2 \tag{4.43}$$

if $j > 1$, with the obvious analogue if $j = 1$. In either case, from (4.41) and (4.42), we have that

$$\|\bar{f}_t - \hat{f}_t\|_N^2 \leqq K\|f - \bar{f}_t\|_D^2 + \left\| \sum_{k=j}^{j+p} (f - \bar{f}_t, \hat{f}_{k,t})_D \hat{f}_{k,t} \right\|_N^2 \quad \text{for all } t \geqq t_j. \tag{4.44}$$

But

$$\begin{aligned} \left\| \sum_{k=j}^{j+p} (f - \bar{f}_t, \hat{f}_{k,t})_D \hat{f}_{k,t} \right\|_N^2 &= \sum_{k=j}^{j+p} \hat{\lambda}_{k,t} [(f - \bar{f}_t, \hat{f}_{k,t})_D]^2 \\ &\leqq (p + 1)\hat{\lambda}_{j+p,t} \|f - \bar{f}_t\|_D^2. \end{aligned}$$

Since $\lambda_{j+p+1} > \lambda_{j+p}$, then from (4.7), $\lambda_{j+p+1} > \hat{\lambda}_{j+p,t}$ for all $t \geqq t_{j+p}$. Therefore,

$$\begin{aligned} \|\bar{f}_t - \hat{f}_t\|_N^2 &\leqq [K + (p + 1)\lambda_{j+p+1}]\|f - \bar{f}_t\|_D^2 \\ &\equiv K\|f - \bar{f}_t\|_D^2 \quad \text{for all } t \geqq \max(t_j, t_{j+p}). \end{aligned}$$

This completes the proof.

The results of Ciarlet, Schultz and Varga [4, Theorems 4 and 5] regarding the convergence of the sequence of approximate eigenfunctions $\{\hat{f}_{j,t}\}_{t=1}^{\infty}$ to f_j in the uniform norm require the assumption that $\lambda_1 < \lambda_2 < \cdots < \lambda_j$. That this assumption is unnecessary is an immediate consequence of Theorems 4.2 and 4.3. We may assume in the following that there exists a positive constant K such that

$$\|w\|_\infty \leqq K\|w\|_N \quad \text{for all } w \in H_N. \tag{4.45}$$

THEOREM 4.4. *With the assumptions of* (2.4)–(2.6), *let* λ_j, *the j-th eigenvalue of* (1.1)–(1.2), *be of multiplicity* $p + 1$, $p \geqq 0$. *Let* $\{f_k\}_{k=1}^{j+p}$ *be a set of D-orthonormal eigenfunctions corresponding to the eigenvalues* $\lambda_1, \cdots, \lambda_j$. *Let* $\{S_{M_t}\}_{t=1}^{\infty}$ *be any sequence of (not necessarily nested) subspaces of* H_N, *with* $\dim S_{M_t} \equiv M_t \geqq j + p$ *for all* $t \geqq 1$, *and such that*

$$\lim_{t\to\infty} \{\inf_{w\in S_{M_t}} \|w - f_k\|_N\} = 0, \qquad 1 \leqq k \leqq j + p. \tag{4.46}$$

For each $t \geqq 1$, *let* $\{\hat{f}_{k,t}\}_{k=j}^{j+p}$ *be a set of D-orthonormal approximate eigenfunctions corresponding to the eigenvalue* λ_j. *Then the following statements are valid.*

(a) *If* $p = 0$, *i.e., if* λ_j *is simple, then there exists a positive integer* t_j, *and a positive constant K, depending only on j, such that*

$$\|f_j - \hat{f}_{j,t}\|_N \leqq K\|f_j - \bar{f}_{j,t}\|_N \quad \text{for all } t \geqq t_j, \tag{4.47}$$

where $\bar{f}_{j,t}$ *is the N-norm projection of* f_j *onto* S_{M_t} *for each* $t \geqq 1$. *Thus, if* (4.45) *is satisfied, then* $\{\hat{f}_{j,t}\}_{t=1}^{\infty}$ *converges uniformly to* f_j.

(b) *If* $p > 0$, *let* $f(x)$ *be any element of the* $(p + 1)$-*dimensional subspace of* H_N *spanned by the* $\{f_k\}_{k=j}^{j+p}$. *Let* $\hat{f}_t \equiv \sum_{k=j}^{j+p} (f, \hat{f}_{k,t})_D \hat{f}_{k,t}$ *be the D-norm projection of f onto the subspace spanned by the* $\{\hat{f}_{k,t}\}_{k=j}^{j+p}$ *for each* $t \geqq 1$. *Then there exists a positive integer* t_j *and a positive constant K, depending only on j and p, such that*

$$\|f - \hat{f}_t\|_N \leqq K\|f\|_D^{1/2} \sum_{k=j}^{j+p} \|f_k - \bar{f}_{k,t}\|_N. \tag{4.48}$$

Thus, if (4.45) *is satisfied, then* $\{\hat{f}_t\}_{t=1}^{\infty}$ *converges uniformly to f.*

Proof. We shall prove only the result of (4.48). The result of (4.47) follows in the same fashion. From Theorem 4.3 and (2.15), we have that

$$\|\bar{f}_t - \hat{f}_t\|_N \leqq K\|f - \bar{f}_t\|_D \leqq K\|f - \bar{f}_t\|_N,$$

where $\bar{f}_t$ is the N-norm projection of f onto S_{M_t}. But, by the triangle inequality,

$$\|f - \hat{f}_t\|_N \leqq \|f - \bar{f}_t\|_N + \|\bar{f}_t - \hat{f}_t\|_N,$$

and hence,

$$\|f - \hat{f}_t\|_N \leqq (1 + K)\|f - \bar{f}_t\|_N.$$

But $f = \sum_{k=j}^{j+p} (f, f_k)_D f_k$, and it follows easily that $\bar{f}_t = \sum_{k=j}^{j+p} (f, f_k)_D \bar{f}_{k,t}$. Thus

$$\|f - \bar{f}_t\|_N = \left\| \sum_{k=j}^{j+p} (f, f_k)_D (f_k - \bar{f}_{k,t}) \right\|_N,$$

and (4.48) follows with an application of Hölder's inequality. This completes the proof.

The results of this section have been stated for *sequences* $\{S_{M_t}\}_{t=1}^{\infty}$ of subspaces of H_N for which $\lim_{t\to\infty} \|f_k - \bar{f}_{k,t}\|_N = 0$, $1 \leqq k \leqq j$. It is easily seen that these results are valid for a *fixed* finite-dimensional subspace $S_M \equiv S_{M_t}$, provided that $\| f_k - \bar{f}_{k,t} \|_N$, $1 \leqq k \leqq j$, are all sufficiently small. In particular, then, given a finite-dimensional subspace S_M of H_N whose degree of approximation in H_N depends essentially on a *single* parameter (e.g., $\bar{\Delta}$, in the case of polynomial splines on a mesh of maximum size $\bar{\Delta}$), these results are valid whenever this parameter is sufficiently small.

The asymptotic error estimates developed in Ciarlet, Schultz and Varga [4, Theorems 7 and 9] for polynomial subspaces and subspaces of L-splines can thus be immediately extended by the results of this section to eliminate the assumption that $\lambda_1 < \lambda_2 < \cdots < \lambda_j$. Moreover, when λ_j is a multiple eigenvalue, it is clear from (4.48) that the "eigensubspace" error estimates are of the same order.

It is clear from approximation-theoretic results that these estimates are best possible in the N-norm with respect to exponents. But, the uniform norm estimates are *not* best possible. If, however, one knew estimates of $f_j - \bar{f}_{j,t}$ and its derivatives in the L^2- and L^∞-norms, it is clear that the corresponding estimates for $f_j - \hat{f}_{j,t}$ could be significantly improved. We shall be concerned with this in our subsequent paper [14]. As a particular example, however, consider the one-dimensional Sturm–Liouville eigenvalue problem. If we assume that the eigenfunctions are in class $C^4[a, b]$, and use a subspace of cubic polynomial spline functions, we shall show that

$$\| f_j - \bar{f}_{j,t} \|_\infty = O(\bar{\Delta}_t^4),$$

and hence, that, if λ_j is a simple eigenvalue,

$$\| f_j - \hat{f}_{j,t} \|_\infty = O(\bar{\Delta}_t^4),$$

provided that $\bar{\Delta}_t$, the maximum mesh spacing, is sufficiently small. This result is best possible with respect to the exponent of $\bar{\Delta}_t$. From Theorem 4.3, it follows that the "eigensubspace" accuracy is also of fourth order. Thus, we can answer in the affirmative the conjecture to this effect in [2].

The same techniques carry over to multidimensional problems; i.e., known estimates for Galerkin approximation as, for example, in Schultz [15] and Strang and Fix [16] can be used to obtain improved error estimates for the Rayleigh–Ritz method. We shall consider several such examples in [14]. In general, however, if $r = 0$ in (2.1), it follows from (2.15) and Theorems 4.2 and 4.3 that

$$\| f_j - \hat{f}_{j,t} \|_{L^2} \leqq K \| f_j - \bar{f}_{j,t} \|_{L^2}.$$

That is, whatever L^2-estimates are available or can be proved for Galerkin approximation are then directly valid for the Rayleigh–Ritz method. The case of uniform norm estimates is, of necessity, more complicated.

REFERENCES

[1] G. Birkhoff and C. deBoor, *Piecewise polynomial interpolation and approximation*, Approximation of Functions, H. L. Garabedian, ed., Elsevier, Amsterdam, 1965, pp. 164–190.

[2] G. Birkhoff, C. deBoor, B. Swartz and B. Wendroff, *Rayleigh–Ritz approximation by piecewise cubic polynomials*, this Journal, 3 (1966), pp. 188–203.

[3] F. Brauer, *Singular self-adjoint boundary value problems for the differential equation* $Lx = \lambda Mx$, Trans. Amer. Math. Soc., 88 (1958), pp. 331–345.

[4] P. G. CIARLET, M. H. SCHULTZ AND R. S. VARGA, *Numerical methods of high-order accuracy for nonlinear boundary value problems. III: Eigenvalue problems*, Numer. Math., 12 (1968), pp. 120–133.

[5] L. COLLATZ, *The Numerical Treatment of Differential Equations*. 3rd ed., Springer-Verlag, Berlin–Göttingen-Heidelberg, 1960.

[6] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics. I*, Interscience, New York, 1953.

[7] S. H. GOULD, *Variational Methods for Eigenvalue Problems*, University of Toronto Press, Toronto, 1966.

[8] O. HALD AND O. WIDLUND, *On the eigenvalue problem for the finite element variational method*, to appear.

[9] O. JOHNSON, *Error bounds for Sturm–Liouville eigenvalue approximations by several piecewise cubic Rayleigh–Ritz methods*, this Journal, 6 (1969), pp. 317–333.

[10] A. A. KAMKE, *Über die definiten selbstadjungierten Eigenwertaufgaben bei gewöhnlichen linearen Differentialgleichungen. II, III*, Math. Z., 46 (1940), pp. 231–286.

[11] ———, *Über die definiten selbstadjungierten Eigenwertaufgaben bei gewöhnlichen linearen Differentialgleichungen. IV*, Ibid., 48 (1942), pp. 67–100.

[12] S. G. MIKHLIN, *Variational Methods in Mathematical Physics*. Macmillan, New York, 1964.

[13] J. G. PIERCE, *Higher order convergence results for the Rayleigh–Ritz method applied to a special class of eigenvalue problems*, Thesis, Case Western Reserve University, Cleveland, Ohio, 1969.

[14] J. G. PIERCE AND R. S. VARGA, *Higher order convergence results for the Rayleigh–Ritz method applied to eigenvalue problems. II. Improved error bounds for eigenfunctions*, to appear.

[15] M. H. SCHULTZ, *Multivariate spline functions and elliptic problems*, Approximation with Special Emphasis on Spline Functions, I. J. Schoenberg, ed., Academic Press, New York, 1969, pp. 279–347.

[16] GILBERT STRANG AND GEORGE FIX, *An analysis of the finite element method*, to appear.

[17] B. WENDROFF, *Bounds for eigenvalues of some differential operators by the Rayleigh–Ritz method*, Math. Comp., 19 (1965), pp. 218–224.