ESTIMATE FOR THE ERRORS IN EIGENVALUE AND EIGENVECTOR APPROXIMATION BY GA. (U) MARYLAND UNIV COLLEGE PARK INST FOR PHYSICAL SCIENCE AND TECH. I BABUSKA ET AL.

UNCLASSIFIED NOV 86 :BN-1056 N00014-85-K-0169 F/G 12/1 NL
Laboratory for Numerical Analysis
Technical Note BN-1056

ESTIMATE FOR THE ERRORS IN EIGENVALUE AND EIGENVECTOR APPROXIMATION BY GALERKIN METHODS, WITH PARTICULAR ATTENTION TO THE CASE OF MULTIPLE EIGENVALUES

by
I. Babuška and J. E. Osborn

November 1986
DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.
Refined estimates for the errors in eigenvalue and eigenvector approximation by finite element, or, more generally, Galerkin methods, as they apply to self-adjoint problems, are presented. Particular attention is given to the case of multiple eigenvalues. The results are new in this case. The proof is based on a novel approach which yields the known results for simple eigenvalues in a simple way. Numerical computations are presented and analyzed in light of the theoretical results.
Estimates for the Errors in Eigenvalue and Eigenvector Approximation by Galerkin Methods, with Particular Attention to the Case of Multiple Eigenvalues

by

I. Babuška* and J.E. Osborn**

Dedicated to W.C. Rheinboldt on his 60th birthday

* Institute for Physical Science and Technology and Department of Mathematics, University of Maryland, College Park, MD 20742. The work of this author was partially supported by the Office of Naval Research under Contract N00014-85-K-0169 and by the National Science Foundation under grant DMS-85-16191.

** Department of Mathematics, University of Maryland, College Park, MD 20742. The work of this author was partially supported by the National Science Foundation under grant DMS-84-10324.
Abstract

Refined estimates for the errors in eigenvalue and eigenvector approximation by finite element, or, more generally, Galerkin methods, as they apply to self-adjoint problems, are presented. Particular attention is given to the case of multiple eigenvalues. The results are new in this case. The proof is based on a novel approach which yields the known results for simple eigenvalues in a simple way. Numerical computations are presented and analyzed in light of the theoretical results.
1. Introduction.

It is the purpose of this paper to derive some refined estimates for the errors in eigenvalue and eigenvector approximation by finite element, or, more generally, Galerkin methods, as they apply to self-adjoint problems. The results are new in the case of multiple eigenvalues. The proof is based on a novel approach which yields the known results for simple eigenvalues in a simple way.

Suppose \( k \) is an eigenvalue of multiplicity \( q \) of a self-adjoint problem and let \( M(k) \) denote the space of eigenvectors corresponding to \( k \), where \( \| \cdot \| \) denotes the energy norm for our problem. Let \( S \) be the finite dimensional approximation space employed in the Galerkin method. \( k \) will be approximated from above by \( q \) of the Galerkin approximate eigenvalue:

\[
\lambda_{S,k} \leq \lambda_{S,k+1} \leq \cdots \leq \lambda_{S,k+q-1}.
\]

If we choose the \( \lambda_{S,i} \) in increasing order then we have

\[
\lambda_{S,k} \leq \lambda_{S,k+1} \leq \cdots \leq \lambda_{S,k+q-1}.
\]

Our main estimate for the error in eigenvalue approximation is

(1.1) \[
\| \lambda_S - \lambda_k \|_2 = \inf_{u \in M(k)} \inf_{i \in S} \inf_{\| u \| = 1} \| u - i \|_2 = C \| \lambda_S - \lambda_k \|_2
\]

thus showing that the error between \( \lambda_k \) and \( \lambda_{S,k} \) the approximate eigenvalue closest to \( \lambda_k \), is, to within a multiplicative constant, the square of the minimal energy norm distance between exact eigenvectors \( u \in M(k) \) with \( \| u \| = 1 \) and \( S \), i.e., the square of the energy norm distance between \( S \) and the eigenvector.
u \in M(\ell^k) \text{ with } \|u\| = 1 \text{ that can be best approximated by } S.

For \( S, k+q-1 - \ell^k \) the error between \( \ell^k \) and \( S, k+q-1 \) the approximate eigenvalue farthest from \( \ell^k \) we prove

\begin{align*}
(1.2) \quad S, k+q-1 - \ell^k \geq \sup \left\{ \inf \left\{ \|u - v\| : u \in M(\ell^k), v \in S \right\} \right\}^2 = C k \ell^k (S)^2
\end{align*}

and for the errors \( S, k+i - \ell^k, i = 1, \ldots, q-2 \), we obtain bounds in terms of quantities intermediate in size between \( C k \ell^k (S)^2 \) and \( C k \ell^k (S)^2 \).

These results should be contrasted with those in the literature. In Babuska and Aziz [1], Fix [4], and Kolata [6], the estimates

\begin{align*}
(1.3) \quad S, k+i - \ell^k \leq C k \ell^k (S), \quad i = 0, \ldots, q-1,
\end{align*}

are proved. For \( i = 0, \ldots, q-2 \), (1.3) is weaker than the estimates stated above ((1.1) for \( i = 0 \) and those mentioned after (1.2) for \( i = 1, \ldots, q-2 \); for \( i = q-1 \), (1.3) is the same as (1.2) In Birkhoff, de Boor, Swartz, and Wendroff [2], which presents the earliest results of the general type we are discussing, the eigenvalue estimates depend on the sum of the squares of the energy norm distances between \( S \) and the unit eigenvectors associated with all the eigenvalues \( \ell^j \) not exceeding \( \ell^k \). The feature of (1.1) that is new is the dependence of \( C k \ell^k (S) \) on only one eigenvector \( u \in M(\ell^k) \), namely the one best approximated by \( S \).

Regarding the errors in the approximate eigenvectors, we show that if \( u_{S, k} \) is the Galerkin approximate eigenvector corresponding to \( S, k \), then there is a \( u_k = u_k(S) \in M(\ell^k) \) with \( \|u_k\| = 1 \)

- 4 -
such that
\[ \| u_{S,k} - u_k \| \leq C(S). \]

The error \( \| u_{S,k+q-1} - u_{k+q-1} \| \) is bounded by \( C(S) \) and the errors \( \| u_{S,k+i} - u_{k+i} \|, \ i = 1, \ldots, q-2, \) are bounded by quantities intermediate in size between \( C(S) \) and \( C(S) \). The best previously known result is
\[ \| u_{S,k+i} - u_{k+i} \| \leq C(S), \ i = 0, \ldots, q-1. \]

In Section 2 we introduce the class of variationally formulated, self-adjoint eigenvalue problems considered in the paper, define the Galerkin approximations to these problems, and in Lemmas 2.1 - 2.3 give the preliminary results which are used in the sequel. The main theoretical result of the paper is presented and proved in Section 3. The treatment is direct and self-contained, relying on a minimal amount of functional analysis background. In Section 4 we present numerical computations for a finite element approximation of a problem with double eigenvalues for which each double eigenvalue has associated eigenvectors of strikingly different approximation properties. The quantities
\[ \sup_{u \in M(k)} \inf_{i \in S} \| u - u_i \|^2 \]
and
\[ \inf_{u \in M(k)} \sup_{i \in S} \| u - u_i \|^2 \]
are thus of different sizes and we would therefore expect \( C(k) \) and \( C(k) \) to be of different sizes. This is clearly shown by
the computations. The computations also show that \( u_{s,k} \) (the approximate eigenvector belonging to the approximate eigenvalue closest to \( \lambda_k \)) converges to an exact eigenvector with good approximation properties, while \( u_{s,k+1} \) the approximate eigenvector belonging to the approximate eigenvalue farthest from \( \lambda_k \), converges to an exact eigenvector with poor approximation properties.

The literature on eigenvalue problems is extensive, with many papers bearing, at least tangentially, on the problem addressed in this paper. We have, however, mentioned only those papers that bear directly on the central theme of our results; namely, the Galerkin approximation of eigenpairs corresponding to multiple eigenvalues. For a general treatment of eigenvalue problems and their literature, we refer to the excellent and comprehensive monograph of Chatelin [3].
2. Preliminaries.

Suppose $H$ is a real Hilbert space with inner product $(\cdot, \cdot)$ and norm $\|\cdot\|$, respectively, and suppose we are given two symmetric bilinear forms $B_0(u,v)$ and $D(u,v)$ on $H \times H$. $B_0(u,v)$ is assumed to satisfy

$$(2.1) \quad |B_0(u,v)| \leq C_1 \|u\| \|v\|, \quad \forall \ u, v \in H$$

and

$$(2.2) \quad C_2 \|u\|^2 \leq B_0(u,u), \quad \forall \ u \in H, \quad \text{with} \quad C_2 > 0.$$ 

It follows from (2.1) and (2.2) that $(u,v)_{B_0} = B_0(u,v)$ and

$\|u\|_{B_0} = (B_0(u,u))^{1/2}$

are equivalent to $(u,v)$ and $\|u\|$, respectively. Regarding $D$, we suppose

$$(2.3) \quad 0 < D(u,u), \quad \forall \ 0 \neq u \in H,$$

and that

$$(2.4) \quad \|u\|_D = (D(u,u))^{1/2}$$

is compact with respect to $\|\cdot\|$, i.e., it has the property that from any subsequence which is bounded in $\|\cdot\|$, one can extract a subsequence which is Cauchy in $\|\cdot\|_D$.

We then consider the variationally formulated, self-adjoint eigenvalue problem

$$(2.5) \quad \begin{cases} \text{Seek } \lambda \text{ (real)} \text{ and } 0 \neq u \in H \text{ such that} \\ B_0(u,v) = \lambda D(u,v), \quad \forall \ v \in H. \end{cases}$$

Under the assumptions we have made, there is a sequence of eigenvalues

$$0 < \lambda_1 \leq \lambda_2 \leq \ldots < \lambda_+$$
and corresponding eigenvectors 

\[ u_1, u_2, \ldots \]

which can be chosen to satisfy

\[
B_0(u_i, u_j) = \delta_{i,j},
\]

where \( \delta_{i,j} = 1 \) for \( i = j \) and \( \delta_{i,j} = 0 \) for \( i \neq j \). Furthermore, any \( u \in \mathcal{H} \) can be written as

\[
u = \sum_{j=1}^{\infty} a_j u_j, \text{ with } a_j = B_0(u, u_j),
\]

where (2.7) converges in \( \| \cdot \|_{B_0} \). The eigenvalues \( \lambda_i \) satisfy the following well-known variational principles:

\[
\begin{align*}
\lambda_1 &= \min_{u \in \mathcal{H}} \frac{B_0(u, u)}{D(u, u)} = \frac{B(u_1, u_1)}{D(u_1, u_1)}, \\
\lambda_k &= \min_{u \in \mathcal{H}} \frac{B_0(u, u)}{D(u, u)} = \frac{B(u_k, u_k)}{D(u_k, u_k)}, \quad k = 2, 3, \ldots
\end{align*}
\]

(2.8)

(2.9)

(2.10) \( M(\lambda_k) = \{ u : u \text{ is an eigenvector of } (2.5) \text{ corresponding to } \lambda_k \} \)

We shall be interested in approximating the eigenpairs of (2.5) by finite element, or, more generally, Galerkin methods.
Toward this end we suppose we are given a (one parametric) family 
\( \{S_h\}_{0<h<1} \) of subspaces \( S_h \subseteq H \) and we consider the eigenvalue problem

\[
\begin{align*}
\text{Seek} \quad & \lambda_h \text{ (real), } \ 0 \neq u_h \in S_h \text{ such that} \\
& B_0(u_h, v) = \lambda_h D(u_h, v), \forall v \in S_h.
\end{align*}
\]

The eigenpairs \((\lambda_h, u_h)\) of (2.11) are then viewed as approximation so the eigenpairs \((\lambda, u)\) of (2.5). (2.11) is called the Galerkin method determined by the subspace \( S_h \) for the approximation of the eigenvalues and eigenvectors of (2.5). We will also sometimes refer to problem (2.11) as the Galerkin approximation of the problem (2.5). (2.11) has a sequence of eigenvalues

\[ 0 < \lambda_{h,1} < \lambda_{h,2} < \cdots < \lambda_{h,N}, \quad N = \dim S_h, \]

and corresponding eigenvectors

\[ u_{h,1}, u_{h,2}, \ldots, u_{h,N} \]

which can be chosen to satisfy

\[
(2.12) \quad B_0(u_h,i,u_h,j) = \lambda_h D(u_h,i,u_h,j) = \delta_{i,j}, \quad i,j = 1,\ldots,N.
\]

Minimum and minimum-maximum principles analogous to (2.8) and (2.9) hold for the problem (2.11); they are obtained by replacing \( H \) by \( S_h \) and letting \( k = 1,\ldots,N \). We will refer to them by (2.8*) and (2.9*), respectively. Using (2.8) and (2.9) together with (2.8*) and (2.9*) we see immediately that

\[
(2.13) \quad \lambda_k = \lambda_{h,k}, \quad k = 1,2,\ldots,N = \dim S_h.
\]

For every \( \lambda_{h,k} \) we let
\[ M(\{h, k\}) = \{u : u \text{ is an eigenvector of (2.11) corresponding to } (h, k) \}. \]

Because of (2.1)-(2.4), 0 is an eigenvalue of neither (2.5) nor (2.11). It will be convenient, however, to introduce the notation \( \lambda_0 = (h, 0) = 0. \)

In what will follow we shall assume that the family \( \{S_h\} \) satisfies the approximability assumption

\[ \inf_{\lambda = S_h} \|u - \lambda\| \to 0 \text{ as } h \to 0, \text{ for each } u \in \mathcal{H}. \]

It follows from the variational principles (2.8), (2.9), (2.8*), and (2.9*), and assumption (2.14) that \( (h, k) \to \lambda_k \) as \( h \to 0 \) for each \( k \).

Our analysis employs two functions \( \lambda(\lambda) \) and \( \lambda_h(\lambda_h) \) of the non-negative real variable \( \lambda \) which are associated with the eigenvalue of (2.5) and (2.11), respectively. We define

\[ \lambda(\lambda) = \inf_{j=1,2,...} \left| 1 - \frac{\lambda}{j} \right|, \]

and

\[ \lambda_h(\lambda_h) = \min_{j=1,...,N} \left| 1 - \frac{\lambda}{h,j} \right|. \]

It is immediate that the functions are non-negative and continuous in \( \lambda \) and that

\[ \lambda(\lambda) = 0 \text{ if and only if } \lambda = \frac{1}{j} \text{ for some } j \]

and

\[ \lambda_h(\lambda_h) = 0 \text{ if and only if } \lambda = \frac{1}{h,j} \text{ for some } j. \]

In the following lemmas we give characterization of \( \lambda \) and
which do not involve the eigenvalues \( \lambda_j \) and \( \lambda_{h,j} \) respectively. For \( 0 < \lambda < 1 \) and \( u, v \in H \) define

\[
B(\lambda, u, v) = B_0(u, v) - \lambda D(u, v).
\]

We now have

**Lemma 2.1.** For all \( 0 < \lambda < 1 \),

\[
\varphi(\lambda) = \inf_{u \in H} \sup_{v \in H} |B(\lambda, u, v)|.
\]

Suppose \( \lambda \) has multiplicity \( q \), i.e., suppose

\[
\lambda_{k-1} < \lambda = \lambda_{k+1} = \ldots = \lambda_{k+q-1} < \lambda_{k+q}.
\]

Then, for

\[
\tilde{\lambda}_k = \left( \frac{1}{k+1} - \frac{1}{k} \right)^{-1} \quad \text{and} \quad \tilde{\lambda}_{k+q} = \left( \frac{1}{k+q+1} - \frac{1}{k+q} \right)^{-1},
\]

we have

\[
\varphi(\lambda) = 1 - \frac{1}{\tilde{\lambda}_k}
\]

and

\[
\varphi(\lambda) = \sup_{v \in H} |B(\lambda, u, v)|
\]

\[
= |B(\lambda, u, u)|, \forall u \in M(\lambda), \text{with } \|u\|_{B_0} = 1.
\]

**Proof.** For \( u, v \in H \) write

\[
u = \sum_{j=1}^{\tilde{\lambda}} a_j u_j, \quad v = \sum_{j=1}^{\tilde{\lambda}} b_j v_j.
\]

Then

\[
B(\lambda, u, v) = \sum_{j=1}^{\tilde{\lambda}} a_j b_j (1 - \frac{1}{\tilde{\lambda}_j}).
\]
Thus

\[(2.23) \quad \sup_{v \in H, \|v\|_{B_0} = 1} |B(1, u, v)| = \left[ \sum_{j=1}^{\infty} a_j^2 (1 - \frac{1}{j^2}) \right]^{1/2},\]

from which we get

\[
\inf_{u \in H, \|u\|_{B_0} = 1} \sup_{v \in H, \|v\|_{B_0} = 1} |B(1, u, v)| = \inf_{j=1,2,\ldots} |1 - \frac{1}{j}| = \phi(1).
\]

This is (2.18). (2.20) follows from the definition of \(\phi(1)\) and an examination of the graphs of \(|1 - \frac{1}{j}|\) for \(j = 1, 2, \ldots\).

(2.21) follows from (2.20), (2.22), and (2.23).

In a similar way we have

**Lemma 2.2.** With \(H\) replaced by \(S_h\), \(k\) by \(h, k\) and \(u_k\) by \(u_{h,k}\), Lemma 2.1 holds for \(\phi_h(1)\). (Relationships analogous to those of Lemma 1 will be indicated by an asterisk.)

If \(\lambda_k\) is an eigenvalue of multiplicity \(g\), then \(\lambda_{h,k}, \ldots, \lambda_{h,k+q-1}\) could be multiple or simple. The graphs of \(\phi(1)\) and \(\phi_h(1)\) is given in Figure 2.1

![Figure 2.1. The graphs of \(\phi(1)\) and \(\phi_h(1)\).](image-url)
We end the section with a lemma that expresses a fundamental property of eigenvalue and eigenvector approximation.

**Lemma 2.3.** Suppose $(\lambda_i, u_i)$ is an eigenpair of (2.5) with $\|u\|_D = 1$, suppose $w$ is any vector in $H$ with $\|w\|_D = 1$, and let $\tilde{\lambda}_i = B_0(w, w)$. Then

$$\tilde{\lambda}_i - \lambda_i = \|w-u\|_{B_0}^2 - \|w-u\|_{D}^2. \tag{2.24}$$

(Note that we have assumed $u$ and $w$ are normalized with respect to $\| \cdot \|_D$ here, whereas in (2.6) and (2.12) we assumed $u_i$ and $u_{hi}$ are normalized with respect to $\| \cdot \|_{B_0}$.)

**Proof.** By an easy calculation,

$$\|w-u\|_{B_0}^2 - \|w-u\|_{D}^2 = \|w\|_{B_0}^2 - 2B_0(w, u) + \|u\|_{B_0}^2$$

$$- \|w\|_{D}^2 + 2\lambda D(w, u) - \|u\|_{D}^2.$$  

Then, since

$$\|w\|_D = \|u\|_D = 1,$$

$$\|w\|_{B_0}^2 = \tilde{\lambda}_i,$$

$$\|u\|_{B_0}^2 = \lambda_i,$$

and

$$B_0(w, u) = \lambda D(w, u),$$

we get the desired result.
3. The main result.

For \( i = 1, 2, \ldots \) suppose \( k_i \) is an eigenvalue of (2.5) of multiplicity \( q_i \), i.e., suppose

\[
\begin{align*}
&k_{i-1} < k_i = k_{i+1} = \ldots = k_{i+q_i-1} < k_{i+q_i} = k_{i+1}.
\end{align*}
\]

Here \( k_1, k_2 \) is the lowest index of the 2nd distinct eigenvalue, \( k_3 \) is the lowest index of the 3rd distinct eigenvalue, etc. Let

\[
(3.1) \quad \varepsilon_{i,j}(h) = \inf_{u \in M(k_i)} \inf_{1 \leq j \leq q_i} \|u - r\|_{B_0}, \quad j = 1, \ldots, q_i.
\]

The restrictions \( B_0(u, uh, k_i) = \ldots = B_0(u, uh, k_i + j - 2) = 0 \) is considered vacuous if \( q_i = 1 \). Note that \( \varepsilon_{i,1} = \varepsilon_{i,1}^1 \) and \( \varepsilon_{i,q_i} = \varepsilon_{i,q_i}^1 \), where \( \varepsilon_{i,1} \) and \( \varepsilon_{i,q_i} \) are the quantities introduced in Section 1. It is the purpose of this section to estimate the eigenvalue and eigenvector errors for the Galerkin method (2.11) in terms of the approximability quantities \( \varepsilon_{i,j}(h) \).

**Theorem 3.1.** There are constants \( C \) and \( h_0 \) such that

\[
(3.2) \quad k_i + j - 1 < \varepsilon_{i,j}(h) < k_i + j - 1 - C_i^2(2.2), \quad \forall 0 < h < h_0, \quad j = 1, \ldots, q_i
\]

and

\[
(3.3) \quad \|u_{k_i + j - 1} - uh, k_i + j - 1\|_{B_0} < C_i^1(2.2), \quad \forall 0 < h < h_0, \quad j = 1, \ldots, q_i
\]

To be slightly more precise, the eigenvectors \( u_1, u_2, \ldots \) of (2.5) can be chosen so that (3.3) holds (as well as (2.6)).
Proof. Overview of the Proof. The complete details of the proof, which proceeds by induction, are given below. Here we provide an overview. In Step A we give the proof for \( i = 1 \). The proof is very simple in this case and rests entirely on the minimum principle (2.8\(^*\)) and Lemma 2.3.

The central part of the proof is given in Step B. There we prove the theorem for \( i = 2 \), proving first the eigenvalue estimate (3.2) and then the eigenvector estimate (3.3). In particular, in Steps B.1 and B.2, estimates (3.2) and (3.3), respectively, are proved for \( j = 1 \). We further note that the argument used in Step B proves the main inductive step in our proof, yielding the result for \( i = i + 1 \) on the assumption that it is true for \( i \). To be somewhat more specific, in Step B.1 we prove (3.2) directly for any \( i \geq 2 \) and in B.2 we prove (3.3) for \( i = i + 1 \) under the assumption that \( \|u_{h,j} - u_j\|_{B_0} \to 0 \) as \( h \to 0 \) for \( i \).

Details of the Proof. Throughout the proof we use the fact that \( \xi_{i,j}(h) \) can also be expressed as

\[
\xi_{i,j}(h) = \inf_{\|u\|_{B_0}=1} \inf_{I = M(k_i)} \inf_{u = M(k_i)} \|u - u_{i,j}\|_{B_0}.
\]

(3.1')

\[
B_0(u, u_h, k_i) = \ldots = B_0(u, u_h, k_i + j - 2) = 0.
\]

Step A. Here we prove the theorem for \( i = 1 \).

Step A.1. Suppose \( k_1 = 1 \) is an eigenvalue of (2.5) with multiplicity \( q_1 \), i.e., suppose \( \xi_1 = \xi_2 = \ldots = \xi_{q_1} \).
In this step we estimate $h_{1, 1} - h_{1}$, the error between $h_{1}$ and the approximate eigenvalue among $h_{1, 1}, \ldots, h_{1, q}$ that is closest to $h_{1}$, i.e., we prove (3.2) for $i = j = 1$. Note that

$$
\varepsilon_{1, 1}(h) = \inf_{u : M(1)} \inf_{u \in S_{h}} \inf_{\|u\|_{B_0} = 1} \|u - u_{1}\|
$$

is the error in the approximation by elements of $S_{h}$ of the most easily approximated eigenvector associated with $h_{1}$.

From the definitions of $\varepsilon_{1, 1}(h)$ we see that there is a $\tilde{u}_{h} \in M(1)$ with $\|\tilde{u}_{h}\|_{B_0} = 1$ and an $s_{h} \in S_{h}$ such that

$$
(3.4) \quad \|\tilde{u}_{h} - s_{h}\|_{B_0} = \varepsilon_{1, 1}(h).
$$

Let

$$
\tilde{u} = \frac{\tilde{u}_{h}}{\sqrt{D(\tilde{u}_{h}, \tilde{u}_{h})}}, \quad s_{h} = \frac{s_{h}}{\sqrt{D(s_{h}, s_{h})}}.
$$

By the minimum principle (2.8*) we have

$$
(3.5) \quad h_{1, 1} - h_{1} = B_{0}(s_{h}, s_{h}) - h_{1}.
$$

Now apply Lemma 2.3 with $(\tilde{\nu}, \omega) = (1, \tilde{u}_{h})$, $w = s_{h}$, and $v = B_{0}(s_{h}, s_{h})$. This yields

$$
(3.6) \quad B_{0}(s_{h}, s_{h}) - h_{1} = \|s_{h}, \tilde{u}_{h}\|_{B_{0}} - h_{1}\|s_{h}, \tilde{u}_{h}\|_{D} / 2
$$

$$
\|s_{h}, \tilde{u}_{h}\|_{B_{0}} = C\|s_{h}, \tilde{u}_{h}\|_{B_{0}} / 2.
$$

(3.4) - (3.6) yield the desired result.

Step A.2. In this step we prove (3.3) for $i = j = 1$. Let $u_{1}, u_{2}, \ldots$ be eigenvectors of (2.5) satisfying (2.6). Write
(3.7) \[ u_{h,1} = \sum_{j=1}^{x} a^{(1)}_j u_j. \]

From (3.6) and (3.7) we have

\[
\left[ 1 - \frac{1}{q_1+1} \right] \sum_{j=q_1+1}^{x} \left[ a^{(1)}_j \right]^2 = \sum_{j=q_1+1}^{x} \left[ a^{(1)}_j \right]^2 \left( 1 - \frac{1}{q_1+1} \right)
\]

\[ = \frac{1}{q_1+1} \sum_{j=1}^{x} \left[ a^{(1)}_j \right]^2 \left( 1 - \frac{1}{q_1+1} \right)
\]

\[ = B\left( q_1+1, u_{h,1}, u_{h,1} \right)
\]

\[ = B\left( q_1+1, u_{h,1}, u_{h,1} \right) + \left( 1 - \frac{1}{q_1+1} \right) D(u_{h,1}, u_{h,1})
\]

\[ = \left( 1 - \frac{1}{q_1+1} \right) h, 1
\]

\[ = C r^2, 1, 1(h).
\]

Hence

\[
\| u_{h,1} - \sum_{j=1}^{q_1} a^{(1)}_j u_j \|_{B_0} = \left[ \sum_{j=q_1+1}^{x} \left[ a^{(1)}_j \right]^2 \right]^{1/2}
\]

\[ \leq C(1 - \frac{1}{q_1+1})^{-1/2} r, 1, 1(h).
\]

(3.9)

Redefining \( u_1 \) to be \( \frac{\sum_{j=1}^{q_1} a^{(1)}_j u_j}{q_1} \), we easily see that

\[ \| u_1 \|_{B_0} = 1, \] so that (2.6) still holds, and

(3.10) \[ \| u_{h,1} - u_1 \|_{B_0} \leq C r, 1, 1(h), \]

as desired. Note that \( u_1 \) may depend on \( h \).

Step A.3. Suppose \( q_1 = 2 \). From (3.1') we see that
\begin{equation}
\gamma_{1,2}(h) = \inf_{u \in M^1} \inf_{s \in S_h} \| u - s \|_{B_0}, \quad \| u \|_{B_0} = 1, \quad B_0(u, u_h, 1) = 0, \quad B_0(u, u_h, 1) = 0
\end{equation}

Choose \( \bar{u}_h \in M^1 \) with \( \| \bar{u}_h \|_{B_0} = 1, \quad B_0(\bar{u}_h, u_h, 1) = 0 \) and \( s_h \in S_h \) with \( B_0(s_h, u_h, 1) = 0 \) so that

\begin{equation}
\| \bar{u}_h - s_h \|_{B_0} = \gamma_{1,2}(h),
\end{equation}

and let

\[ \tilde{u}_h = \frac{\bar{u}_h}{\sqrt{D(\bar{u}_h, \bar{u}_h)}}, \quad \tilde{s}_h = \frac{s_h}{\sqrt{D(s_h, s_h)}}. \]

Since \( B_0(s_h, u_h, 1) = 0 \), from the minimum principle (2.8*), Lemma 2.3, and (3.12), we have

\begin{equation}
\gamma_{1,2} = \| \tilde{s}_h \|_{B_0} \leq C \gamma_{1,2}(h).
\end{equation}

This is (3.2) for \( i = 1 \) and \( j = 2 \).

**Step A.4.** In Step A.2 we redefined \( u_1 \). Now redefine \( u_2, \ldots, u_q \) so that \( u_1, \ldots, u_q \) are \( B_0 \)-orthogonal. Write

\[ u_{h,2} = \sum_{j=1}^{q} a_{j}^{(2)} u_j. \]

Now, proceeding as in Step A.2 and using (3.13), we have

\[ (1 - q_{1}^{-1}) \sum_{j=q_{1}+1}^{q} \left[ a_{j}^{(2)} \right]^2 - \sum_{j=1}^{q} \left[ a_{j}^{(2)} \right]^2 (1 - q_{1}^{-1}) \]

\[ = |B(\bar{u}_h, u_{h,2}, u_{h,2})| \]

\[ = (\bar{u}_h, u_{h,2}, u_{h,2})^{-1} \]

\[ - C_{1,2}(h). \]
Thus

\[ (3.14) \quad \| u_{h,2} - \sum_{j=1}^{q_1} a_j^{(2)} u_j \|_{B_0} \leq C \varepsilon_{1,2}(h). \]

But by (3.10),

\[ a_1^{(2)} = B_0(u_{h,2}, u_1) \]
\[ = B_0(u_{h,2}, u_1 - u_{h,1}) \]

\[ (3.15) \quad \leq \| u_{h,2} \|_{B_0} \| u_1 - u_{h,1} \|_{B_0} \]
\[ \leq C \varepsilon_{1,1}(h) \]
\[ \leq C \varepsilon_{1,2}(h). \]

Combining (3.14) and (3.15) we get

\[ \| u_{h,2} - \sum_{j=2}^{q_1} a_j^{(2)} u_j \|_{B_0} \leq \| u_{h,2} - \sum_{j=1}^{q_1} a_j^{(2)} u_j \|_{B_0} + \| a_1^{(2)} u_1 \|_{B_0} \]
\[ \leq C \varepsilon_{1,2}(h). \]

Redefining \( u_2 \) to be \( \frac{1}{q_1} \sum_{j=2}^{q_1} a_j^{(2)} u_j \), we see that \( \| u_2 \|_{B_0} = : \)

and \( B_0(u_1, u_2) = 0 \), so that (2.6) holds and

\[ (3.16) \quad \| u_{h,2} - u_2 \|_{B_0} \leq C \varepsilon_{1,2}(h), \]

which is (3.3) for \( i = 1, j = 2 \).

Step A.5. Continuing in the above manner we obtain the proof of (3.2) and (3.3) for \( i = 1 \) and \( j = 1, \ldots, q_1 \).

Step B. Here we prove Theorem 3.1. for \( i = 2 \).
Step B.1. Suppose \( k_2 \) \((k_2 = q_1 + 1)\) is an eigenvalue of (2.5) of multiplicity \( q_2 \). In this step we estimate \( h, k_2 - k_2 \), the error between \( k_2 \) and the approximate eigenvalue among \( h, k_2, \ldots, h, k_2 + q_2 - 1 \) that is closest to \( k_2 \). Note that

\[
(3.17) \quad \gamma_{2,1}(h) = \inf_{u \in M(k_2)} \inf_{i \in S_h} \| u - \| \leq 1
\]

Write \( k_2 - 1 = k_2 \), \( h, k_2 - 1 = h, k_2 \), and \( h, k_2 \) = \( \varphi h, k_2 \). Then \( 0 < \gamma < 1 \) and \( \varphi > 1 \). Let \( \gamma = \frac{2\varphi h}{1+r-1} \). From (2.13) and the definitions of \( \gamma_k \) and \( \gamma_{h,k} \) in Lemmas 2.1 and 2.2, respectively, we see that

\[
(3.18) \quad \gamma_{k_2} \leq \gamma_{h,k_2}.
\]

A simple calculation shows that

\[
\gamma_{k_2} = \frac{2}{1+r-1} \gamma_{k_2}
\]

and

\[
\gamma_{h,k_2} = \frac{2\varphi h}{1+r-1} \gamma_{k_2} = h, k_2.
\]

Hence (3.18) shows that

\[
\frac{2}{1+r-1} \leq \gamma_{h}.
\]

Since \( h, j \rightarrow j \) as \( h \rightarrow 0 \) (see Section 2), we see that \( h, j \rightarrow j \), \( \varphi \rightarrow 1 \), and \( \frac{2}{1+r-1} \) as \( h \rightarrow 0 \). Thus, noting that \( \frac{2}{1+r-1} < 1 \), we see that we can choose \( h_0 \) such that \( 0 < h < h_0 \) implies that

- 20 -
From (2.19), (2.20), (2.19*), and (2.20*) (see also Figure 2.1) we get

\[ 0 - \Phi_h(\bar{t}, k_2) - \Phi(\bar{t}, k_2) = \left( 1 - \frac{\bar{t}_{t}, k_2}{\Phi_h k_2} \right) - \left( 1 - \frac{\bar{t}_{t}, k_2}{k_2} \right) = \Phi_h(1 - \Phi_h^{-1}), \]

and hence, using (3.19), we get

\[ \phi_{h}(k_2) = \phi_{h}(\Phi_h^{-1}) \]

provided that

\[ \left[ \Phi_h(\bar{t}_{t}, k_2) - \Phi(\bar{t}_{t}, k_2) \right]^{-1} = \left( 1 + 3t \right)^{2} < 1. \]

We will now show that

\[ \Phi_h(\bar{t}_{t}, k_2) - \Phi(\bar{t}_{t}, k_2) = C_{2,1}(h), \quad \text{for } h \leq h_0, \]

where \( C \) depends only on \( k_2 \) and \( k_2 \), and, in particular,
is independent of $h$, and $h_0$ depends only on $'k_2-1' k_2$ and the approximability of the eigenvectors in $M('k_2)$ by $S_h$.

As in Step A.1, choose $\tilde{u}_h \in M('k_2)$ with $\|\tilde{u}_h\|_{B_0} = 1$ and $s_h \in S_h$ such that

\[ (3.23) \quad \|\tilde{u}_h - s_h\|_{B_0} \leq 2,1(h). \]

We see that $s_h$ is the $B_0$-orthogonal projection of $\tilde{u}_h$ onto $S_h$, i.e., that

\[ (3.24) \quad B_0(\tilde{u}_h - s_h, v) = 0, \forall v \in S_h. \]

Let $w_h = \tilde{u}_h - s_h$. Then

\[ (3.25) \quad \|w_h\|_{B_0} \leq 2,1(h) \]

and

\[ (3.26) \quad \|s_h\|_{B_0}^2 + \|w_h\|_{B_0}^2 = \|\tilde{u}_h\|_{B_0}^2 = 1. \]

Next we write

\[ s_h = c_h \tilde{u}_h + e_h, \]

where

\[ B_0(\tilde{u}_h, e_h) = 0, \]

i.e., we let $c_h \tilde{u}_h$ be the $B_0$-orthogonal projection of $s_h$ onto span $\{\tilde{u}_h\}$. Let $r_h = (1-c_h)\tilde{u}_h$. Then

\[ w_h = r_h - e_h, B(e_h, r_h) = 0. \]

which implies

\[ (3.27) \quad \|e_h\|_{B_0} \leq \|w_h\|_{B_0}. \]

Furthermore, using (3.24), (3.25), and (3.26) we have
\[ c_h = B_0(s_h, \bar{u}_h) = \| s_h \|_{B_0} \| \bar{u}_h \|_{B_0} - 1 \]

and

\[ c_h = B_0(s_h, \bar{u}_h) = B_0(\bar{u}_h - w_h, \bar{u}_h) \\
= \| \bar{u}_h \|_{B_0}^2 - B_0(w_h, \bar{u}_h) \\
= 1 - \| w_h \|_{B_0} \\
= 1 - h_0 \| w_h \|_{B_0} \\
> 0, \text{ for } h < h_0, \]

with \( h_0 \) sufficiently small. Thus we can assume

\[ 0 < c_h : 1. \]

Also,

(3.28) \[ \| r_h \|_{B_0} = \| w_h \|_{B_0}^2. \]

To see this we refer to Figure 3.1 and note that

\[ \| r_h \|_{B_0} = \| w_h \|_{B_0} \cos \alpha \]

and

\[ \| w_h \|_{B_0} = \| \bar{u}_h \|_{B_0} \sin \beta = \cos \alpha. \]

From (3.28) an the definition of \( r_h \) we have

(3.29) \[ c_h = 1 - \| w_h \|_{B_0}^2. \]
Figure 3.1. Configuration of $r_h$, $w_h$, $\bar{u}_h$, $\alpha$, and $\beta$.

Assume now that $u_k = \bar{u}_h$ (redefining $u_k'$, if necessary).

Write

$$e_h = \sum_{j=1}^{\chi} q_j u_j, \quad v = \sum_{j=1}^{\chi} b_j u_j.$$  

Then

$$B(\bar{r}_h, k_2, s_h, v) = c_h B_0(\bar{u}_h, v) + B_0(e_h, v)$$

$$- \bar{r}_h, k_2 c_h D(\bar{u}_h, v) = \bar{r}_h, k_2 D(e_h, v)$$

$$= c_h b_{k_2} + \sum_{j=1}^{\chi} q_j b_j - \bar{r}_h, k_2 k_2^{-1} c_h b_{k_2}$$

$$- \bar{r}_h, k_2 \sum_{j=1}^{\chi} q_j b_j^{-1}$$

$$= c_h b_{k_2} (1 - \bar{r}_h, k_2^{-1}) + \sum_{j=1}^{\chi} q_j b_j (1 - \bar{r}_h, k_2^{-1}).$$

- 24 -
and hence
\[
\sup_{v \in S_{h, k_2}, \|v\|_{B_0} = 1} |B(\tilde{h}, k_2, s_{h, k_2}, v)| \leq \sup_{v \in H} |B(\tilde{h}, k, s_{h, k}, v)|
\]
\[
= \left[ c_h^2 \left( 1 - \tilde{h}, k_2, k_2 \right)^2 + \sum_{j=1}^{1} q_j^2 \left( 1 - \tilde{h}, k_2, k_2 \right)^2 \right]^{1/2}.
\]
\[(3.30)\]

Combining (3.27), (3.29), and (3.30), we get
\[
\sup_{v \in S_{h, k_2}, \|v\|_{B_0} = 1} |B(\tilde{h}, k_2, s_{h, k_2}, v)| \leq \left[ \left( 1 - \|w_h\|_{B_0}^2 \right)^2 \left( 1 - \tilde{h}, k_2, k_2 \right)^2 \right. \]
\[
+ \sup_{j=1, 2, \ldots} \left( 1 - \tilde{h}, k_2, k_2 \right)^2 \left( \sum_{j=1}^{1} q_j^2 \right) \left. \right]^{1/2}
\]
\[
= \left[ \left( 1 - \|w_h\|_{B_0}^2 \right)^2 \left( 1 - \tilde{h}, k_2, k_2 \right)^2 \right. \]
\[
+ \sup_{j=1, 2, \ldots} \left( 1 - \tilde{h}, k_2, k_2 \right)^2 \|w_h\|_{B_0}^2 \left. \right]^{1/2}
\]
\[
= \left[ \left( 1 - \|w_h\|_{B_0}^2 \right)^2 \left( \tilde{h}, k_2 \right)^2 \right. \]
\[
+ \sup_{j=1, 2, \ldots} \left( 1 - \tilde{h}, k_2, k_2 \right)^2 \|w_h\|_{B_0}^2 \left. \right]^{1/2}
\]
\[
= Q.
\]

Thus, using (3.25) and (3.26), we have
\[
\hat{\phi}(\vec{h}, k_2) = \frac{Q}{\|\vec{h}\|^2_{\mathcal{B}_0}} \left[ (1 - \|\vec{h}\|^2_{\mathcal{B}_0})^2 \phi^2(\vec{h}, k_2) + \sup_{j=1,2,\ldots} (1 - \|\vec{h}, k_2\|^2_{\mathcal{B}_0})^2 \|w_h\|^2_{\mathcal{B}_0} (1 - \|w_h\|^2_{\mathcal{B}_0})^{-1} \right]^{1/2}
\]

and hence

\begin{equation}
\phi_h(\vec{h}, k_2) = \phi(\vec{h}, k_2) - \frac{R_{2,1}(h)}{2\phi(\vec{h}, k_2)},
\end{equation}

where

\[
R = \sup_{j=1,2,\ldots} (1 - \|\vec{h}, k_2\|^2_{\mathcal{B}_0})^2 (1 - \|w_h\|^2_{\mathcal{B}_0})^{-1} - \phi^2(\vec{h}, k_2).
\]

One can easily show that \( \frac{R}{2\phi(\vec{h}, k_2)} \) is bounded independent of \( h \); in fact, using (2.20) and (3.19), one can show that

\begin{equation}
\frac{R}{2\phi(\vec{h}, k_2)} = \frac{2^{k_2}}{2^{1}(1-\epsilon)[2,1(h)]^{1/2}} \leq \frac{1}{(1-\epsilon)},
\end{equation}

provided \( \epsilon^2_{2,1}(h) < 1/2 \), i.e., provided \( h < h_0 \) for \( h_0 \) sufficiently small. Hence, combining (3.31) and (3.32) we get

\begin{equation}
\hat{\phi}(\vec{h}, k_2) = \phi(\vec{h}, k_2) - \frac{1}{(1-\epsilon)} \epsilon^2_{2,1}(h),
\end{equation}

which is (3.22). Combining (3.20), (3.21), and (3.33) we have

\begin{equation}
h, k_2 - k_2^0 \leq C \epsilon^2_{2,1}(h), \text{ for } h - h_0,
\end{equation}

where
(3.35) \[ C = \frac{2k_2}{k_1'(1-\epsilon)}. \]

This is (3.2) for \( i = 2, j = 1. \)

Comment on Inequality (3.34). \( C \) in (3.34) clearly depends on \( k_1', k_2-1', \) and \( k_2, \) but is independent of \( h. \) Note that if we were considering a family of problems depending on a parameter \( \tau, \) we could bound \( C = C(\tau) \) above, independent of \( \tau, \) provided \( k_2 = k_2(\tau) \) was bounded above, \( k_1 = k_1(\tau) \) was bounded away from 0, and \( \tau = \tau(\tau) \) was bounded away from 0 and 1. It follows from (3.31) and (3.33), that (3.34) is valid for \( h = h_0, \) where \( h_0 \) depends only on \( k_1', k_2-1', \) and the approximability of the eigenvectors \( u_j, j = k_2, \ldots, k_2+q_2-1, \) by \( S_h. \) For a family of problems, \( h_0(\tau) \) could be bounded away from 0 if \( k_1(\tau) \) was bounded away from 0, \( k_2(\tau) \) was bounded above, \( \tau(\tau) \) was bounded away from 0 and 1, and the eigenvectors \( u_j = u_j(\tau), j = k_2, \ldots, k_2+q_2-1, \) could be approximated by \( S_h, \) uniformly in \( \tau. \)

Step B.2. Suppose, as in Step B.1, that \( k_2 \) has multiplicity \( q_2. \) We have shown in Step A.5 that we can choose the eigenvectors \( u_1, u_2, \ldots \) of (2.5) so that (2.6) holds and so that

\[
(3.36) \quad \| u_{h,j} - u_j \|_{B_0} \leq C r, j(h), j = 1, \ldots, q_1 - k_2 - 1.
\]

Write

\[
(3.37) \quad u_{h,k_2} = \sum_{j=1}^\chi a_j(k_2) u_j.
\]

From (3.37) we have
\[
\sum_{j=1}^{m} \left[ \begin{array}{c} (k_2) \\ j \end{array} \right]^2 (1 - \frac{1}{k_2/j}) = B^1 \left( k_2, u_{h}, k_2, u_{h}, k_2 \right) \\
= B^1 \left( k_2, u_{h}, k_2, u_{h}, k_2 \right) + (\begin{array}{c} k_2 - 1 \\ h, k_2 \end{array}) D^1 \left( u_{h}, k_2, u_{h}, k_2 \right) \\
= (\begin{array}{c} 1 \\ h, k_2 \end{array})^{-1}.
\]

which, together with (3.34), yields

\[
(3.38) \quad \sum_{j=1}^{k_2-1} \left[ \begin{array}{c} (k_2) \\ j \end{array} \right]^2 (1 - \frac{1}{k_2/j}) + \sum_{j=k_2+q_2}^{x} \left[ \begin{array}{c} (k_2) \\ j \end{array} \right]^2 (1 - \frac{1}{k_2/j}) = C_{2,1}^2 (h).
\]

Note that the first term inside the absolute value is negative and the second is positive. In addition

\[
C_1 \leq |1 - \frac{1}{k_2/j}| \leq C_2, \quad \forall j = k_2, k_2+1, \ldots, k_2+q_2-1,
\]

with \( C_1, C_2 \) positive numbers. Hence from (3.38) we obtain

\[
(3.39) \quad \sum_{j=1}^{k_2-1} \left[ \begin{array}{c} (k_2) \\ j \end{array} \right]^2 D_{1,2}^2 (h) + D_2 \sum_{j=k_2+q_2}^{x} \left[ \begin{array}{c} (k_2) \\ j \end{array} \right]^2 D_{1,2}^2 (h)
\]

and

\[
(3.40) \quad \sum_{j=k_2+q_2}^{x} \left[ \begin{array}{c} (k_2) \\ j \end{array} \right]^2 D_{3,2}^2 (h) + D_3 \sum_{j=1}^{k_2-1} \left[ \begin{array}{c} (k_2) \\ j \end{array} \right]^2 D_{3,2}^2 (h).
\]

Write

\[
(3.41) \quad u_{h,i} - u_{i} = \sum_{j=1}^{x} b_{i,j} u_{j}, \quad i = 1, \ldots, k_2-1 = q_1.
\]

Then, by (3.36),
Next we wish to find constants $\alpha_1, \ldots, \alpha_{k_2-1}$ so that

\begin{equation}
B_0(u_i, \sum_{j=1}^{k_2-1} \alpha_j u_{h,j}) = a_i, \quad i = 1, \ldots, k_2-1.
\end{equation}

Using (3.41), these equations can be written as

\begin{equation}
B_0\left[u_i, \sum_{j=1}^{k_2-1} (\alpha_j u_j + \alpha_j \sum_{t=1}^{\lambda} b_j, u_t)\right] = a_i + \sum_{j=1}^{k_2-1} b_j, i' = \sum_{j=1}^{k_2-1} (k_2) = a_i, \quad i = 1, \ldots, k_2-1.
\end{equation}

Since (2.14) implies $\varepsilon_2, 1(h) \to 0$ as $h \to 0$, from (3.42) we see that the $b_j, i$ are small for $h = h_0$, with $h_0$ sufficiently small, and hence the system (3.44) is uniquely solvable, and, moreover, there is a constant $L$, depending only on $k_2$, such that

\begin{equation}
\left[\sum_{j=1}^{k_2-1} \alpha_j \right]^{1/2} \leq L \left[\sum_{j=1}^{k_2-1} \left(\frac{a_i}{(k_2)}\right)^{2} \right]^{1/2}.
\end{equation}

Now, from (3.36) we obtain

\begin{equation}
|a_j| \leq |B_0(u_h, k_2, u_j)|
\end{equation}

\begin{align*}
&= |B_0(u_h, k_2, u_j - u_h, j'| \\
&= \|u_h, k_2\| B_0 \|u_j - u_h, j\| B_0 \\
&= \|u_j - u_h, j\| B_0 \\
&= C_{r_1, j}(h), \quad j = 1, \ldots, k_2-1.
\end{align*}
Letting

\[(3.46) \quad \nu_{k_2}^2(h) = \sum_{j=1}^{k_2-1} \epsilon_{i,j}^2(h),\]

we see that

\[(3.47) \quad \left[\sum_{j=1}^{k_2-1} \left(\alpha_{j}^2(k_2)\right)^{1/2}\right] = C\nu_{k_2}(h),\]

and thus, from (3.45)

\[(3.48) \quad \left[\sum_{j=1}^{k_2-1} \alpha_{i,j}^2\right]^{1/2} = LC\nu_{k_2}(h) = C\nu_{k_2}(h).\]

Now let

\[(3.49) \quad i = u_{h,k_2} - \sum_{j=1}^{k_2-1} \nu_{j}^i h_{j,i}.\]

Then \(i \in S_h\). Furthermore, from (3.43) and (3.44) we get

\[(3.50) \quad B_0(u_{i,i}) = \begin{cases} 0, & i = k_2 - 1 \\ \frac{(k_2)}{a_i} & \sum_{j=1}^{k_2-1} \nu_{j}^i b_{j,i}, & i = k_2 \end{cases} \]

From (3.48) and (3.49),

\[(3.51) \quad \| B_0 - I \| = \| u_{h,k_2} \| B_0 - \| u_{h,k_2} \| B_0 - \| u_{h,k_2} \| B_0.\]
Using (3.34), (2.21*), (3.50), and (3.51), and the fact that 
\[ \pi''_{k_2}(h) \to 0 \text{ as } h \to 0, \] 
we get

\begin{equation}
C_{2,1}(h) = \frac{(h,k_2 - k_2) - (h,k_2)}{(h,k_2)} 
= \phi_h(k_2) 
\end{equation}

(3.52)

\begin{equation}
\| h* \|_{B_0} \end{equation}

\[ = B(h,k_2 \cdot u,h,k_2, \|u\|_{B_0}) \]

\[ = C \left[ \sum_{i=k_2+q_2}^n \sum_{a_i}^{(k_2)} \left( a_i^{(k_2)} - \sum_{i=1}^{k_2-1} u_i b_i \right) \left( i - \frac{k_2}{i} \right) \right] \]

where \( C' > 0 \) and is independent of \( h \). Combining (3.42), (3.45), (3.46), and (3.52) we obtain

\begin{equation}
\sum_{i=k_2+q_2}^n \left( a_i^{(k_2)} \right)^2 = C_{2,1}(h) 
\end{equation}

\[ + \sum_{i=k_2+q_2}^n \left| a_i^{(k_2)} \right| \sum_{i=1}^{k_2-1} u_i |b_i| 
\]

\[ \leq C \left[ C_{2,1}(h) + \sum_{i=1}^{k_2-1} \left| a_i^{(k_2)} \right| \sum_{i=k_2+q_2}^n u_i |b_i| 
\]

\[ \leq C_{2,1}(h) + \sum_{i=1}^{k_2-1} \left| u_i \right| \sum_{i=k_2+q_2}^n \left| a_i^{(k_2)} \right| \]

\[ + \sum_{i=k_2+q_2}^n |b_i|^2 \left| \sum_{i=k_2+q_2}^n |b_i|^2 \right| ^{1/2} \]

(3.53)
\[ c^2 \text{I}._{2,1}(h) + \sum_{i=1}^{k_2-1} |\eta_i| \left[ \sum_{\epsilon = k_2+q_2}^{x} \left| \frac{k_2}{a_\epsilon} \right|^2 \right]^{1/2} \]

\[
\max_{i=1, \ldots, k_2-1} \left[ \frac{k_2-1}{i(h)} \sum_{\epsilon = k_2+q_2}^{x} \left| \frac{k_2}{a_\epsilon} \right|^2 \right]^{1/2} 
\]

\[ c^2 \text{I}._{2,1}(h) + \epsilon_{1,k_2-1}(h) \sqrt{k_2-1} \left[ \sum_{i=1}^{k_2-1} |\eta_i|^2 \right]^{1/2} \]

\[
\sum_{\epsilon = k_2+q_2}^{x} \left| \frac{k_2}{a_\epsilon} \right|^2 \left[ \frac{k_2}{a_\epsilon} \right]^{1/2} 
\]

\[
\sum_{\epsilon = k_2+q_2}^{x} \left| \frac{k_2}{a_\epsilon} \right|^2 \left[ \frac{k_2}{a_\epsilon} \right]^{1/2} 
\]

\[ c^2 \text{I}._{2,1}(h) + \epsilon_{1,k_2-1}(h) \sqrt{k_2-1} \sum_{i=1}^{k_2-1} \left[ \frac{k_2-1}{i(h)} \sum_{\epsilon = k_2+q_2}^{x} \left| \frac{k_2}{a_\epsilon} \right|^2 \right]^{1/2} \]

\[
\sum_{\epsilon = k_2+q_2}^{x} \left| \frac{k_2}{a_\epsilon} \right|^2 \left[ \frac{k_2}{a_\epsilon} \right]^{1/2} 
\]

(3.53) is a quadratic inequality in \[
\left[ \sum_{\epsilon = k_2+q_2}^{x} \left| \frac{k_2}{a_\epsilon} \right|^2 \right]^{1/2} \]

whose solution yields

\[ \sum_{i=1}^{k_2-1} \left[ \frac{k_2}{a_i} \right]^2 - C \cdot c^2 \text{I}._{1,k_2-1}(h) \sum_{i=1}^{k_2-1} \left[ \frac{k_2}{a_i} \right]^2 + C \cdot c^2 \text{I}._{2,1}(h). \]
Combining (3.39) and (3.54) we get

\[ \sum_{i=1}^{k_2-1} \left[ a_i \right]^2 D_{1,2,1}^2(h) + \sum_{i=1}^{k_2-1} \left[ a_i \right]^2 D_{2,1}^2(h) \]

and thus, since \( r_{k_2-1}(h) \) is small for \( h \) small,

\[ \sum_{i=1}^{k_2-1} \left[ a_i \right]^2 D_{5,2,1}^2(h) \]

Next, combining (3.40) and (3.55), we get

\[ \sum_{i=k_2+q_2}^{k_2+q_2-1} \left[ a_i \right]^2 D_{6,2,1}^2(h) \]

Finally, from (3.37), (3.55), and (3.56), we have

\[ \| u_{h,k_2} - \sum_{j=k_2}^{k_2+q_2-1} a_j u_j \|_{B_0}^2 = \sum_{j=1}^{k_2-1} \left[ a_j \right]^2 + \sum_{j=k_2+q_2}^{k_2+q_2-1} \left[ a_j \right]^2 C_{2,1}^2(h). \]

Redefining \( u_{k_2} \) to be

\[ \sum_{j=k_2}^{k_2+q_2-1} \left[ a_j \right] u_j \]

we see that \( \| u_{k_2} \|_{B_0}^2 \) is 1, so that (2.6) holds, and

\[ \| u_{h,k_2} - u_{k_2} \|_{B_0}^2 \leq C_{2,1}^2(h). \]

This is (3.3) for \( i = 2, j = 1 \).
Comment on Estimate (3.57). In the proof of (3.57) we used (3.36), which was proved in Step A. A careful examinations of the proof of (3.57) show that we did not use the full strength of (3.36), but only the weaker fact that \( \| u_h,j - u_j \|_{B_0} \to 0 \) as \( h \to 0 \) for \( j \leq k_2-1 \). (Cf. the Overview of the Proof.)

Step B.3. Suppose \( q_2 = 2 \). In Step B.1 we estimated \( \{ h, k_2 \} \). In this step we estimate \( \{ h, k_2 + 1 \} \).

We proceed by modifying problems (2.5) and (2.11) by restricting them to the space

\[
H_{h,k_2} = \{ u \in H : B_0(u, u_{h,k_2}) = 0 \},
\]

and

\[
S_{h,k_2} = \{ u \in S_h : B_0(u, u_{h,k_2}) = 0 \},
\]

respectively, i.e., we consider the problems (2.5) and (2.11) obtained when \( H \) and \( S_h \) are replaced by \( H_{h,k_2} \) and \( S_{h,k_2} \) in (2.5) and (2.11), respectively. (2.11) has the same eigenpairs \( \{ h,j, u_{h,j} \} \) as does (2.11) except that the pair \( \{ h,k_2, u_{h,k_2} \} \) is eliminated. (2.5) has eigenpairs \( \{ h,k_2, u_{h,k_2} \} \) which in general depend on \( h \). Nevertheless,

\[
(3.58)
\]

\[
\{ k_2 + r \} = \{ k_2 + r' \} = 0, \ldots, q_2 - 2,
\]

i.e., \( k_2 \), the eigenvalue under consideration, is an eigenvalue of multiplicity \( q_2 - 1 \) for problem (2.5). Its eigenspace is
M \{k_2\} = \{u : M\{k_2\} : B_0(u, u_{h,k_2}) = 0\}.

We can now apply the argument used in Step B.1 to problems \(h,k_2\) and \(h,k_2\)
(2.5) and (2.11) and we obtain (cf. (3.34))

\[ h,k_2 + 1 - k_2 + 1 = C_{2,2}(h), \text{ for } h < h_0. \]

Since \(u_{h,k_2}\) depends on \(h\), the problems (2.5) and
(2.11) depend on \(h\). It follows from the Comment on Inequality (3.34) with \(r = h\) that we can apply the argument in Step B.1 obtaining \(C\) and \(h_0\) that are independent of \(h\). To see
this, note that \(k_2 = k_2\) by (3.58), \(k_1 = k_1\) by the
minimum principle, and \(k_2 = k_2\) as \(h \to 0\). by the minimum-maximum principle and the fact that \(u_{h,k_2}\)
\(\to u_{k_2}\) (cf. (3.57)), and hence \(k_2\) is bounded from above,
\(k_2\) is bounded from below, and \(k_2\) is bounded away from \(0, 1\). Then observe that the eigenvectors in \(M \{k_2\}\) can be approximated by \(S_h\), uniformly in \(h\). Note that in Step B.1 we also used the fact that \(h,j \to k_2 - 1\) as \(h \to 0\) for \(j = k_2 - 1\) and \(k_2\). It is easy to see that the corresponding fact is true in the present context.

**Step B.4.** Suppose \(q_2 = 2\) as in Step B.3. Here we show that
\(u_{k_2+1}\) can be chosen so that \(\|u_{h,k_2+1} - u_{k_2+1}\|_{B_0} \leq C_{2,2}(h)\). We
know that
(3.60) \[ u_{h,j} - u_j^B_0 \in C_{2,1}(h), \quad j = 1, \ldots, q, \]
\[ u_{h,1} - u_1^B_0 \leq C_{2,2}(h), \quad j = q+1 = k_2. \]

(cf. (3.16), (3.14), and (3.57)). Assume that \( u_{k_2+1}^0 \) have been redefined so that (2.6) holds. Write

\[ u_{h,k_2+1} = \sum_{j=1}^{(k_2+1)} a_j u_j. \]

If we apply the argument used in Step B.2 to \( u_{h,k_2+1} \), i.e., if we let \( k_2 \) be replaced by \( k_2+1 \) and use (3.59) instead of (3.34), we obtain

\[ \| u_{h,k_2+1} - \sum_{j=k_2}^{(k_2+1)} a_j u_j \|_B_0 \leq C_{2,2}(h). \]

But, by (3.60),

\[ a_{k_2} = B_0(u_{h,k_2}, u_{k_2}). \]

\[ = B_0(u_{h,k_2}, u_{k_2}^0) - u_{h,k_2} \]
\[ = \| u_{k_2}^0 - u_{h,k_2} \|_B_0 \leq C_{2,1}(h), \]
\[ \leq C_{2,2}(h). \]

and hence

\[ \| u_{h,k_2+1} - \sum_{j=k_2+1}^{(k_2+1)} a_j u_j \|_B_0 \leq C_{2,2}(h). \]
Redefining \( u_{k_2+1} \) to be
\[
\frac{k_2+q_2+1 \begin{pmatrix} k_2' \\ a_j \end{pmatrix} \begin{pmatrix} j=k_2+1 \\ u_j \end{pmatrix}}{k_2+q_2-1},
\]
we see that
\[
\|u_{k_2+1}\|_{B_0} = 1, \quad B_0(u_{k_2+1}, u_j) = 0, \quad j = 1, \ldots, k_2,
\]
so that (2.6) holds, and
\[
(3.63) \quad \|u_{h,k_2+1} - u_{k_2+1}\|_{B_0} \leq C_{2,2}(h),
\]
which is (3.3) for \( i = j = 2 \).

**Step B.5.** Continuing in this manner we prove (3.2) and (3.3) for \( i = 2 \) and \( j = 1, \ldots, q_2 \).

**Step C.** Repeating the argument in B we get (3.2) and (3.3) for \( i = 3, 4, \ldots \). This completes the proof.

**Remark 3.1.** It is possible to use an alternate argument in Step B.1 if we introduce the so-called Riesz formulas for the spectral projections associated with an operator. We suppose the space \( H \) and the bilinear forms \( B_0 \) and \( D \) have been complexified in the usual manner. Let \( P_{2h} \), and \( P_{h,k} \) be the \( B_0 \)-orthogonal projections of \( H \) onto \( M(k,k_2') \) and \( \bigoplus_{i=0}^{q_2-1} M(h,k_2+i) \), the direct sum of the eigenspaces \( M(h,k_2+i) \), \( i = 0, \ldots, q_2-1 \), respectively.

Introduce next the operators \( T, T_h : H \rightarrow H \) defined by
\[
\begin{align*}
Tf & : H \\
B_0(Tf, v) &= D(f, v), \quad \forall v \in H
\end{align*}
\]
and

- 37 -
\[ \begin{align*}
T_h f &= S_h \\
B_0(T_h f, v) &= D(f, v), \quad \forall v \in S_h
\end{align*} \]

It follows from (2.1), (2.2), and (2.3) that \( T \) and \( T_h \) are defined and compact on \( H \). Furthermore

\[ (3.64) \quad \| (T - T_h) f \|_{B_0} \leq C \inf_{i \in S_h} \| T f - i \|_{B_0}. \]

It is immediate that \((\mu, u)\) is an eigenpair of (2.5) if and only if \((\mu = i^{-1} u)\) is an eigenpair of \( T \). Likewise \((\mu_h, u_h)\) is an eigenpair of (2.11) if and only if \((\mu_h = i^{-1} u_h)\) is an eigenpair of \( T_h \). As a consequence of (2.14), \( T_h \to T \) in the operator norm associated with \( \| \cdot \|_{B_0} \). Let \( \Gamma \) be a circle in the complex plane centered at \( \mu_{k_2} = i^{k_2} \), enclosing no other eigenvalues of \( T \).

Then for \( h \) sufficiently small, \( \Gamma \) will contain the eigenvalues \( \mu_{h,k_2} = i^{h,k_2} \), \( i = 0, \ldots, q_2 - 1 \), of \( T_h \). Also, \( P_{\frac{1}{2}} \) and \( P_{\frac{1}{2}} \), which are referred to as spectral projections associated with \( T \) and \( T_h \), respectively, can be written as

\[ (3.65) \quad P_{\frac{1}{2}} = \frac{1}{2\pi i} \int_{\Gamma} (z - T)^{-1} dz \]

and

\[ (3.66) \quad P_{h,\frac{1}{2}} = \frac{1}{2\pi i} \int_{\Gamma} (z - T_h)^{-1} dz. \]

These are the Riesz formulas. With these formulas we can derive an eigenvector error estimate which will lead to the eigenvalue estimate (3.34).
Let \( u \in M'_2 \) with \( \|u\|_0 = 1 \). Then \( v_h = P_h u \) and from the formulas (3.65) and (3.66) we obtain

\[
\|u - v_h\|_{B_0} = \| (P_{k_2} - P_{h,k_2}) u \|_{B_0}
\]

\[
= \left\| \frac{1}{2\pi i} \int_{\Gamma} (z - T_h)^{-1} (T - T_h)(2 - T)^{-1} u \, dz \right\|
\]

\[
= \frac{1}{2\pi} \left\| (z - T_h)^{-1} (T - T_h) \frac{u}{z - \mu} \, dz \right\|
\]

\[
= \frac{1}{2\pi} \left\| (z - T_h)^{-1} (T - T_h) \frac{u}{z - \mu} \, dz \right\|
\]

\[
- \frac{1}{2\pi} |2\pi \text{rad}(\Gamma)| \sup_{z \in \Gamma} \| (z - T_h)^{-1} \| \times \frac{1}{\text{rad}(\Gamma)} \| (T - T_h) u \|_{B_0}
\]

\[
= \left( -\mu k_2 + \text{rad}(\Gamma) + \mu h, k_2 + q_2 - 1 \right)^{-1} \| (T - T_h) u \|_{B_0}
\]

\[
\cdot C \| (T - T_h) u \|_{B_0}.
\]

(3.64) and (3.67) yield

\[
\|u - v_h\|_{B_0} \leq C \inf_{r \in S_h} \| r - u \|_{B_0}
\]

\[
= C \inf_{r \in S_h} \| r - u \|_{B_0}
\]

\[
- C \inf_{r \in S_h} \| u - r \|_{B_0}.
\]

This is an eigenvector estimate; it shows that starting from any \( u \in M'_2 \) with \( \|u\|_0 = 1 \) we can construct a \( v_h = v_h(u) \) that is close to \( u \). We now use (3.68) to prove (3.34).

By the minimum principle (2.8*) we have
Since \( v_h = \sum_{i=1}^{k_2-1} \delta_{v_i} \rho_{k_2+1} \) we know that \( B_0(v_h, u_{h,i}) = 0, \ i = 1, \ldots, k_2-1 \). Thus, from (3.69) we find

\[
\inf_{\|v\|_D = 1} B_0(v, v) - \|v\|_D^2
\]

Combining this with Lemma 2.3 and (3.68) we obtain

\[
\inf_{\|v\|_D = 1} B_0(v, v) - \|v\|_D^2 \geq \inf_{\|v\|_D = 1} B_0(v_h, v_h) - \|v_h\|_D^2
\]

which is (3.34). We note that the proof given here rests on equation (2.24) in Lemma 2.3 and employs formulas (3.65) and (3.66) to construct \( v_h = v_h(u) \) that is \( B_0 \)-orthogonal to \( u_{h,i}, \ i = 0, \ldots, k_2-1 \), and satisfies

\[
\int_{D_0} v_h = \int_{D_0} u_{h,i}, \quad \int_{D_0} v_h = \int_{D_0} u_{h,i}, \quad \int_{D_0} v_h = \int_{D_0} u_{h,i}
\]
\[ \|u-v_h\|_{B_0} \leq C \varepsilon_{2,1}(h). \]

We have already seen that the eigenvalue estimates in Steps A.1, A.3, ... can be based on Lemma 2.3. Proceeding as we have here, we see that all of the eigenvalue estimates (3.2) can be based on Lemma 2.3.

In the previous sections we have analyzed the errors in the Galerkin approximation of an eigenvalue problem, concentrating especially on the case of multiple eigenvalues. In this section we consider a finite element-Galerkin method for the approximation of a model, one-dimensional problem with multiple eigenvalues, presenting numerical results and their analysis in terms of the results of Section 3.

Consider the eigenvalue problem

\[
\begin{bmatrix}
-\left(\frac{1}{\varphi'(x)} u'(x)\right)' = i \varphi'(x) u, & x \in I = (-\pi, \pi), \\
\varphi(-\pi) = u(\pi), \\
\left(\frac{1}{\varphi'} u'\right)(-\pi) = \left(\frac{1}{\varphi'} u'\right)(\pi),
\end{bmatrix}
\tag{4.1}
\]

where

\[\varphi(x) = \pi^{-\alpha} |x|^{1+\alpha} \text{sgn } x, \ 0 < \alpha < 1.\]

It is easy to check that the eigenvalues and eigenfunctions are as shown in Table 4.1.

<table>
<thead>
<tr>
<th>(i)</th>
<th>(\lambda_i)</th>
<th>(u_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1.0</td>
<td>\cos \varphi(x)</td>
</tr>
<tr>
<td>2</td>
<td>1.0</td>
<td>\sin \varphi(x)</td>
</tr>
<tr>
<td>3</td>
<td>4.0</td>
<td>\cos 2\varphi(x)</td>
</tr>
<tr>
<td>4</td>
<td>4.0</td>
<td>\sin 2\varphi(x)</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
</tr>
</tbody>
</table>
1, 2, 3, 4, for \( \alpha = .4 \). These errors are plotted in Figure 4.1 in log-log scale. We clearly see the different rates of convergence, specifically seeing the rates \( h^2 \) and \( h^{1+\epsilon} = h^{1.4} \) for the errors in \( (h, i) \) for \( i = 1, 3 \) and \( i = 2, 4 \), respectively, as suggested by (4.2) and (4.3). It should be noted that the estimates presented in Theorem 3.1 are of an asymptotic nature in that they provide information only for small \( h \) (or large \( n \)), i.e., for \( h \) (or \( n \)) in the asymptotic range. From Figure 4.1 we see that for \( \alpha = .4 \) we are in the asymptotic range quite quickly, say for \( n = 16 \).

We computed \( u_{h,1} \) and \( u_{h,2} \), the approximate eigenfunctions corresponding to \( (h, 1) \) and \( (h, 2) \), respectively, normalized by \( \| \cdot \|_D = 1 \). The results of Section 3 suggest that \( u_{h,1} \) should be close to \( C \cos \varphi(x) \) and \( u_{h,2} \) close to \( C \sin \varphi(x) \) (cf. (4.2) and (4.3)), where \( C \) is such that \( C \sin \varphi(x) \) and \( C \cos \varphi(x) \) are normalized by \( \| \cdot \|_D = 1 \), i.e., \( C = n^{-1/2} \). To illustrate this point we computed \( C_1^{(i)} \) and \( C_1^{(i)} \), \( i = 1, 2 \), so that

\[
\text{(4.4)} \quad K(i) = \left\{ \begin{array}{ll}
\| u_{i,h} - C_1^{(i)} \cos \varphi(x) - C_2^{(i)} \sin \varphi(x) \|_{B_0}, & i = 1, 2 \\
\| u_{i,h} - C_1^{(i)} \cos 2\varphi(x) - C_2^{(i)} \sin 2\varphi(x) \|_{B_0}, & i = 3, 4
\end{array} \right.
\]

is minimal. We would expect that

\[
\text{(4.5)} \quad C_1^{(2)}, C_1^{(4)}, C_2^{(1)}, C_2^{(3)} = 0
\]

and

\[
\text{(4.6)} \quad C_1^{(1)} = C_2^{(2)} = C_3^{(1)} = C_4^{(4)} = C = .564189583...
\]

Table 4.2 shows some of the results for \( \alpha = .4 \). We see clearly
the results predicted in (4.5) and (4.6). The increase in $C_2^{(2)}$, $C_2^{(4)}$, $C_1^{(1)}$, and $C_2^{(3)}$ with increasing $n$ is due to the eigenvalue solver we used. Table 4.2 also shows that $K(1) < K(2)$ and $K(3) < K(4)$, as we would expect.

The last column in Table 4.2 and Figure 4.1 show that the ratios

$$\frac{\lambda_{h,i+1} - \lambda_{i+1}}{\lambda_{h,i} - \lambda_i}, \quad i = 1,3,$$

increase as $h \to 0$. This shows that in the whole $h$-range we considered, the approximate eigenvalues converging to a multiple eigenvalue are well separated.
Table 4.2
Numerical Solution of the Eigenvalue Problem (4.1) for \( \alpha = .4 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( i )</th>
<th>( h_i )</th>
<th>( X(i) )</th>
<th>( C(i) )</th>
<th>( C'(i) )</th>
<th>( \frac{h_i}{h_i-1} )</th>
<th>( \frac{i}{i-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
<td>1.0716754</td>
<td>.2704</td>
<td>0</td>
<td>.5637791</td>
<td>0</td>
<td>-.1124891</td>
</tr>
<tr>
<td>2</td>
<td>1.1115481</td>
<td>.3423</td>
<td>0</td>
<td>.4151973</td>
<td>-13</td>
<td>.5636998</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>5.0394692</td>
<td>.1075</td>
<td>+1</td>
<td>.5558919</td>
<td>0</td>
<td>.1317809</td>
<td>-12</td>
</tr>
<tr>
<td>4</td>
<td>5.2414639</td>
<td>.1191</td>
<td>+1</td>
<td>.5022638</td>
<td>-13</td>
<td>.5516234</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1.0175850</td>
<td>.1329</td>
<td>0</td>
<td>.5641633</td>
<td>0</td>
<td>.1596754</td>
</tr>
<tr>
<td>2</td>
<td>1.0352431</td>
<td>.1881</td>
<td>0</td>
<td>-.8916589</td>
<td>-12</td>
<td>.5641519</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4.2691915</td>
<td>.5259</td>
<td>0</td>
<td>.5636643</td>
<td>0</td>
<td>.1124328</td>
<td>-13</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>4.3385100</td>
<td>.5869</td>
<td>0</td>
<td>-.2689727</td>
<td>-12</td>
<td>.5637697</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>1.0043740</td>
<td>.6118</td>
<td>-1</td>
<td>.5641879</td>
<td>0</td>
<td>.6411454</td>
</tr>
<tr>
<td>2</td>
<td>1.0113741</td>
<td>.1067</td>
<td>0</td>
<td>.1323421</td>
<td>-10</td>
<td>.5641830</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4.0666055</td>
<td>.2589</td>
<td>0</td>
<td>.5641561</td>
<td>0</td>
<td>.1970954</td>
<td>-10</td>
</tr>
<tr>
<td>4</td>
<td>4.0936974</td>
<td>.5869</td>
<td>0</td>
<td>-.7375504</td>
<td>-10</td>
<td>.5641613</td>
<td>0</td>
</tr>
<tr>
<td>64</td>
<td>1</td>
<td>1.0010921</td>
<td>.3305</td>
<td>-1</td>
<td>.5641895</td>
<td>0</td>
<td>.7729760</td>
</tr>
<tr>
<td>2</td>
<td>1.0038431</td>
<td>.6202</td>
<td>-1</td>
<td>.8670648</td>
<td>-9</td>
<td>.5641883</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4.0166006</td>
<td>.1289</td>
<td>0</td>
<td>.5641875</td>
<td>0</td>
<td>.3641341</td>
<td>-10</td>
</tr>
<tr>
<td>4</td>
<td>4.0272875</td>
<td>.1653</td>
<td>0</td>
<td>.1415775</td>
<td>-8</td>
<td>.5641858</td>
<td>0</td>
</tr>
<tr>
<td>128</td>
<td>1</td>
<td>1.0002729</td>
<td>.1651</td>
<td>-1</td>
<td>.5641895</td>
<td>0</td>
<td>.4535626</td>
</tr>
<tr>
<td>2</td>
<td>1.0013431</td>
<td>.3665</td>
<td>-1</td>
<td>.3251219</td>
<td>-7</td>
<td>.5641893</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4.0041468</td>
<td>.6440</td>
<td>-1</td>
<td>.5641895</td>
<td>0</td>
<td>.4409247</td>
<td>-7</td>
</tr>
<tr>
<td>4</td>
<td>4.0083380</td>
<td>.9135</td>
<td>-1</td>
<td>-.95611</td>
<td>-8</td>
<td>.5641890</td>
<td>0</td>
</tr>
<tr>
<td>256</td>
<td>1</td>
<td>1.0000682</td>
<td>.8255</td>
<td>-2</td>
<td>.5641896</td>
<td>0</td>
<td>.8070959</td>
</tr>
<tr>
<td>2</td>
<td>1.0004811</td>
<td>.2193</td>
<td>-1</td>
<td>.7269570</td>
<td>-6</td>
<td>.5641895</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4.0010365</td>
<td>.3217</td>
<td>-1</td>
<td>.5641896</td>
<td>0</td>
<td>.6435344</td>
<td>-6</td>
</tr>
<tr>
<td>4</td>
<td>4.0026645</td>
<td>.5162</td>
<td>-1</td>
<td>-.2601000</td>
<td>-6</td>
<td>.5641895</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 4.1

The Error in the Eigenvalues \( h_{1}, h_{2} \) and \( h_{3}, h_{4} \) for \( \nu = .4 \) in Dependence on the Number of Intervals \( n \)
We next consider the case when \( \alpha = .01 \). Table 4.3 presents the same results for \( \alpha = .01 \) as Table 4.2 does for \( \alpha = .4 \).

Figure 4.2 shows the graph of

\[
\log \frac{\rho_{h,i+1} - \rho_{i+1}}{\rho_{h,i} - \rho_{i}}, \quad i = 1, 3.
\]

as a function of the number of intervals \( n \) in a semi-logarithmic scale. The computed values are indicated by o's and x's. The graphs are formed by interpolation (solid lines) and extrapolation (dotted lines). We note three related phenomena that did not occur with \( \alpha = .4 \). For small \( n \) the approximate eigenfunction associated with \( \rho_{h,1} \) is \( \rho_{h,1} - n^{-1/2} \sin \phi(x) \), in contrast to \( \rho_{h,1} - n^{-1/2} \cos \phi(x) \) when \( \alpha = .4 \). We remark that \( n^{-1/2} \cos \phi(x) \) is more easily approximated by \( S_h \) than is \( n^{-1/2} \sin \phi(x) \) for all \( 0 < \alpha < 1 \). This anomaly is present for \( n = 64 \) and for \( n = 128 \) we get results which are in agreement with the (asymptotic) results in Section 3. For \( \rho_{h,3} \) and \( \rho_{h,4} \) we have to take \( n = 256 \) to get results which agree with the asymptotic theory.

For \( \alpha = .01 \) we see that \( K(2) < K(1) \) for small \( n(n = 64) \) and \( K(2) > K(1) \) for large \( n \) and \( K(4) < K(3) \) for small \( n(n = 128) \) and \( K(4) > K(3) \) for large \( n \). Recall that \( K(2) \), \( K(1) \) and \( K(4) \) are all \( K(3) \) for all \( n \) when \( \alpha = .4 \).

Finally we note that when \( \alpha = .01 \) the ratio

\[
\frac{\rho_{h,i+1} - \rho_{i+1}}{\rho_{h,i} - \rho_{i}}, \quad i = 1, 3,
\]

first decreases as \( n \) increases, then for some \( n \) the two eigenvalue errors become equal, and then the ratio increases again.

- 48 -
This is in contrast to the case for $\alpha = .4$, in which the ratio increased over the whole range of $n$ values. We further note that the value $\bar{n}$ for which the eigenvalue errors are equal — $\bar{n} = 70$ for $i = 1$ and $\bar{n} = 160$ for $i = 2$ (see Figure 4.2) — marks a transition in each of these situations from $u_{h,1}^{-1/2} \sin \phi(x)$ to $u_{h,1}^{-1/2} \cos \phi(x)$ and $u_{h,3}^{-1/2} \sin 2\phi(x)$ to $u_{h,3}^{-1/2} \cos 2\phi(x)$, from $K(2) < K(1)$ and $K(4) < K(3)$ to $K(2) > K(1)$ and $K(4) > K(3)$, and from $\displaystyle \frac{h_{i+1} - h_i}{h_{i-1} - h_i}$, $i = 1,3$, decreasing to increasing.

We have thus seen that for $\alpha = .4$ the numerical results are in concert with the (asymptotic) results in Section 3 for the whole range of $n$ considered, while for $\alpha = .01$ they are in disagreement for small $n$, but are in agreement for large $n$. We now make an observation that further illuminates these two phases of error behavior — the pre-asymptotic and the asymptotic. Toward this end we note that if $(u_1, u_h_1)$, with $\|u_1\|_D = 1$, and $(u_{h,1}, u_{h,1})$, with $\|u_{h,1}\|_D = 1$, are first eigenpairs of (2.5) and (2.11), respectively, then

$$0 - \left| \begin{array}{c} h_1 - u_1 \end{array} \right| = \left\| u_{h,1} - u_1 \right\|_{B_0}^2 - \left| \begin{array}{c} u_{h,1} - u_1 \end{array} \right| \left\| u_1 \right\|_D^2$$

(4.7)

$$\leq \inf_{i : S_{h_1}} \left\| u_1 \right\|_{B_0}^2 - \left| \begin{array}{c} u_1 \end{array} \right| \left\| u_1 \right\|_D^2$$

If $\frac{h_{i+1} - h_i}{h_{i-1} - h_i}$ is a multiple eigenvalue, then the $u_1$ in (4.7) can be any corresponding eigenvector with $\|u_1\|_D = 1$. (Note that we are here assuming $u_1$ and $u_{h,1}$ have ($\cdot$)-length equal $1$, whereas in (2.6) and (2.12) they are assumed to have ($\cdot$)-length equal $1$. )
The first inequality in (4.7) follows from the minimum principle (2.8*) and has already been stated in (2.13). The first equality in (4.7) follows immediately from Lemma 2.3 with \((\cdot, u) = (\cdot, u_1), w = u_{h, 1}\), and \(\tilde{u} = B_0(u_{h, 1}, u_{h, 1}) = \tilde{u}_{1, 1}\). If \(x \in S_h\) with \(\|x\|_D = 1\), then from the minimum principle (2.8*),

\[
(4.8) \quad \tilde{u}_{h, 1} = \tilde{u}_{1, 1} - B_0(\tilde{u}_{1, 1}) - \tilde{u}_{1, 1}.
\]

Again from Lemma 2.3, this time with \((\cdot, u) = (\cdot, u_1), w = \tilde{u}_{1, 1}\), and \(\tilde{u} = B_0(\tilde{u}_{1, 1})\), we have

\[
(4.9) \quad B_0(\tilde{u}_{1, 1}) = \tilde{u}_{1, 1} - \|\tilde{u}_{1, 1}\|_{B_0}^2 - \|\tilde{u}_{1, 1}\|_D^2.
\]

The second equality in (4.7) follows from (4.8) and (4.9). It is clear from the above discussion that \(u_1\) can be any eigenvector corresponding to \(\lambda_1\).

From (4.7) we have

\[
(4.10) \quad \tilde{u}_{h, 1} - \tilde{u}_{1, 1} = \|\tilde{u}_{1, 1}\|_{B_0}^2 - \|\tilde{u}_{1, 1}\|_D^2, \forall \tilde{u} \in S_h \text{ with } \|\tilde{u}\|_D = 1.
\]

If \(\tilde{u}\) is \(\|\cdot\|_{B_0}\)-close to \(u_1\), to be more precise, if \(\tilde{u}\) is taken to be the \(B_0\)-projection of \(u_1\) onto \(S_h\), then the second term as the right side of (4.10) is negligible with respect to the first term. This follows from the compactness assumption made on \(\|\cdot\|_D\) in Section 2. On the other hand, if \(\|u_1 - \tilde{u}\|_{B_0}\) is not small, \(\tilde{u}_{h, 1} - \tilde{u}_{1, 1}\) may still be small because of cancellation between the two terms on the right side of (4.10). Regarding the case \(\tilde{u} = .01\), this explains why for \(h\) large (the pre-asymptotic phase), we can have \(u_{h, 1} = h^{-1/2} \sin \varphi (x)\) and \(K(1) > K(2)\), and yet have \(\tilde{u}_{h, 1}\) the approximate eigenvalue associated with \(u_{h, 1}\).
closer to \( h,1 \) than is \( h,2 \) the approximate eigenvalue associated with \( u_{h,2}^{-1/2} \cos \varphi(x) \), while for \( h \) small (the asymptotic phase), we have \( u_{h,1}^{-1/2} \cos \varphi(x) \), \( K(1) < K(2) \), and \( h,1 \) closer to \( h,1 \) than is \( h,2 \), showing that the eigenvalue error, \( h,1 - h,1 \), is governed by \( \inf_{S_h} \| u_1 \|^2 \).

This situation is very similar to the situation with \( n \) fixed and \( \omega \) varying, as can be seen from Table 4.4 where computations for the case \( n = 4 \) are shown. We see that the characteristics observed in Table 4.3 regarding dependence on \( n \) are present in Table 4.4 regarding dependence on \( \omega \). Namely, the abrupt switch in the values of \( C_1^{(1)}, C_2^{(1)} \), the abrupt switch from \( K(2) < K(1) \) and \( K(4) < K(3) \) to \( K(2) > K(1) \) and \( K(4) > K(3) \), and the abrupt switch from decreasing to increasing ratio of errors near the parameter value corresponding to \( h,1 = h,2 \). We mention this situation — \( \omega \) varying and \( n \) fixed — since it is easier to understand in terms of perturbation theory (cf. Katz [5]) than is our original situation — \( n \) varying and \( \omega \) fixed.
Table 4.3

Numerical Solution of the Eigenvalue Problem (4.1) for $\mu = 0.01$

<table>
<thead>
<tr>
<th>n</th>
<th>i</th>
<th>$h_i$</th>
<th>$\omega(i)$</th>
<th>$C_1(i)$</th>
<th>$C_2(i)$</th>
<th>$\frac{h_{i+1} - h_i}{h_i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1.0520268</td>
<td>.2338</td>
<td>.00.8181940</td>
<td>-11</td>
<td>.5634386 0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.0529172</td>
<td>.2268</td>
<td>.56459665</td>
<td>0</td>
<td>-.2916448 -11</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.8576239</td>
<td>.9593</td>
<td>-.9346720</td>
<td>-13</td>
<td>.5597529 0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.8717141</td>
<td>.9615</td>
<td>.56045333</td>
<td>0</td>
<td>.1167277 -1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>1.0128661</td>
<td>.1223</td>
<td>.8717399  -10</td>
<td>0</td>
<td>.5635957 0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.0130098</td>
<td>.1052</td>
<td>.5647369</td>
<td>0</td>
<td>-.8480131 -9</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.2088367</td>
<td>.4650</td>
<td>.2507177  -10</td>
<td>0</td>
<td>.1068703 0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.2106542</td>
<td>.4577</td>
<td>.5642694</td>
<td>0</td>
<td>-.3101833 -10</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1.0032139</td>
<td>.7274</td>
<td>-.9345818</td>
<td>-9</td>
<td>.5636031 0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.0032360</td>
<td>.3568</td>
<td>.5647430</td>
<td>0</td>
<td>.1273043 -7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.0515675</td>
<td>.2384</td>
<td>.3745461</td>
<td>-9</td>
<td>.5638078 0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.0518629</td>
<td>.2205</td>
<td>.5644172</td>
<td>0</td>
<td>-.4115544 -3</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>1.0008063</td>
<td>.5369</td>
<td>-.1311961</td>
<td>-5</td>
<td>.5636032 0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.0008077</td>
<td>.3398</td>
<td>.5647430</td>
<td>0</td>
<td>.2462939 -7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.0128623</td>
<td>.1343</td>
<td>.2743681</td>
<td>-7</td>
<td>.5638240 0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.0129086</td>
<td>.9792</td>
<td>.5644235</td>
<td>0</td>
<td>.3196172 -2</td>
</tr>
<tr>
<td>64</td>
<td>1</td>
<td>1.0002018</td>
<td>.4196</td>
<td>.5647430</td>
<td>0</td>
<td>.3556056 -5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.0002031</td>
<td>.4775</td>
<td>.7414162</td>
<td>-6</td>
<td>.5636032 0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.0032196</td>
<td>.9166</td>
<td>.2379072</td>
<td>-6</td>
<td>.5638239 0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.0032230</td>
<td>.9745</td>
<td>.5644235</td>
<td>0</td>
<td>.1197135 -5</td>
</tr>
<tr>
<td>128</td>
<td>1</td>
<td>1.0000504</td>
<td>.4372</td>
<td>.5647429</td>
<td>0</td>
<td>.1061527 -4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.0000515</td>
<td>.4614</td>
<td>.1553659</td>
<td>-4</td>
<td>.5636031 0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.0008054</td>
<td>.5011</td>
<td>.5644234</td>
<td>0</td>
<td>-.2123279 -4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.0008079</td>
<td>.7741</td>
<td>.1165012</td>
<td>-5</td>
<td>.5638238 0</td>
</tr>
</tbody>
</table>
Figure 4.2

The Graphs of \(\log \left(\frac{h_{2-1}}{h_{1-2}}\right)\) and \(\log \left(\frac{h_{4-4}}{h_{3-3}}\right)\) for \(\sigma = 0.01\) in Dependence on the Number of Intervals \(n\).
Table 4.4
Numerical Solution of the Eigenvalue Problem (4.1) for \( \lambda = 4 \) and for Various \( \omega \)

<table>
<thead>
<tr>
<th>( \omega )</th>
<th>( i )</th>
<th>( h, i )</th>
<th>( K(i) )</th>
<th>( C_1^{(i)} )</th>
<th>( C_2^{(i)} )</th>
<th>( \frac{h, i}{\omega} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1</td>
<td>1</td>
<td>1.2030785</td>
<td>.4689</td>
<td>- .2438763</td>
<td>-13</td>
<td>.5594106</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2496444</td>
<td>.5161</td>
<td>- .5594467</td>
<td>0</td>
<td>.000</td>
</tr>
<tr>
<td>.2</td>
<td>1</td>
<td>1.2238488</td>
<td>.5017</td>
<td>- .8324194</td>
<td>-14</td>
<td>.5562571</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2679299</td>
<td>.5366</td>
<td>- .5585072</td>
<td>0</td>
<td>.000</td>
</tr>
<tr>
<td>.25</td>
<td>1</td>
<td>1.2474367</td>
<td>.5322</td>
<td>- .8324194</td>
<td>-14</td>
<td>.5562571</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2713772</td>
<td>.5414</td>
<td>- .5577904</td>
<td>0</td>
<td>.000</td>
</tr>
<tr>
<td>.275</td>
<td>1</td>
<td>1.2628455</td>
<td>.5506</td>
<td>- .1979358</td>
<td>-12</td>
<td>.5526908</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2718555</td>
<td>.5426</td>
<td>- .5577904</td>
<td>0</td>
<td>.000</td>
</tr>
<tr>
<td>.28750</td>
<td>1</td>
<td>1.2715137</td>
<td>.5605</td>
<td>- .7895498</td>
<td>-11</td>
<td>.5520158</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2718136</td>
<td>.5428</td>
<td>- .55768520</td>
<td>0</td>
<td>.5301689</td>
</tr>
<tr>
<td>.28779</td>
<td>1</td>
<td>1.2717254</td>
<td>.5607</td>
<td>.2170355</td>
<td>-10</td>
<td>.5519997</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2718105</td>
<td>.5429</td>
<td>.5576828</td>
<td>0</td>
<td>- .2773092</td>
</tr>
<tr>
<td>.28794</td>
<td>1</td>
<td>1.2718089</td>
<td>.5429</td>
<td>.5576816</td>
<td>0</td>
<td>.1753561</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2718309</td>
<td>.5608</td>
<td>.1125467</td>
<td>-10</td>
<td>.5519997</td>
</tr>
<tr>
<td>.28809</td>
<td>1</td>
<td>1.2718072</td>
<td>.5429</td>
<td>.5576804</td>
<td>0</td>
<td>.1631249</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2719375</td>
<td>.5610</td>
<td>.2645923</td>
<td>-11</td>
<td>.5519836</td>
</tr>
<tr>
<td>.28868</td>
<td>1</td>
<td>1.2718005</td>
<td>.5429</td>
<td>.5576757</td>
<td>0</td>
<td>.2249782</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2723627</td>
<td>.5614</td>
<td>.3464439</td>
<td>-12</td>
<td>.5519513</td>
</tr>
<tr>
<td>.29</td>
<td>1</td>
<td>1.2717839</td>
<td>.5429</td>
<td>.5576649</td>
<td>0</td>
<td>.2249782</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2733271</td>
<td>.5625</td>
<td>.8906887</td>
<td>-13</td>
<td>.5519836</td>
</tr>
<tr>
<td>.29735</td>
<td>1</td>
<td>1.2717263</td>
<td>.5429</td>
<td>.5576289</td>
<td>0</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2760979</td>
<td>.5661</td>
<td>.2000056</td>
<td>-13</td>
<td>.5519513</td>
</tr>
<tr>
<td>.30</td>
<td>1</td>
<td>1.2715965</td>
<td>.5430</td>
<td>.5575862</td>
<td>0</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.2808526</td>
<td>.5790</td>
<td>.1073461</td>
<td>-13</td>
<td>.5513203</td>
</tr>
<tr>
<td>.40</td>
<td>1</td>
<td>1.2646804</td>
<td>.5382</td>
<td>.5570759</td>
<td>0</td>
<td>.2249782</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.3830892</td>
<td>.6704</td>
<td>.1185635</td>
<td>-13</td>
<td>.5451205</td>
</tr>
<tr>
<td>.60</td>
<td>1</td>
<td>1.2364746</td>
<td>.5093</td>
<td>.5576799</td>
<td>0</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.8239095</td>
<td>.9695</td>
<td>.2249822</td>
<td>-16</td>
<td>.5304168</td>
</tr>
</tbody>
</table>
References


The Laboratory for Numerical analysis is an integral part of the Institute for Physical Science and Technology of the University of Maryland, under the general administration of the Director, Institute for Physical Science and Technology. It has the following goals:

- To conduct research in the mathematical theory and computational implementation of numerical analysis and related topics, with emphasis on the numerical treatment of linear and nonlinear differential equations and problems in linear and nonlinear algebra.

- To help bridge gaps between computational directions in engineering, physics, etc., and those in the mathematical community.

- To provide a limited consulting service in all areas of numerical mathematics to the University as a whole, and also to government agencies and industries in the State of Maryland and the Washington Metropolitan area.

- To assist with the education of numerical analysts, especially at the postdoctoral level, in conjunction with the Interdisciplinary Applied Mathematics Program and the programs of the Mathematics and Computer Science Departments. This includes active collaboration with government agencies such as the National Bureau of Standards.

- To be an international center of study and research for foreign students in numerical mathematics who are supported by foreign governments or exchange agencies (Fulbright, etc.)

Further information may be obtained from Professor I. Babuška, Chairman, Laboratory for Numerical Analysis, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742.
END

4-87

DTIC