NEW *A PRIORI* FEM ERROR ESTIMATES FOR EIGENVALUES

ANDREW V. KNYAZEV [*] AND JOHN E. OSBORN [†]

Abstract. We analyze the Ritz method for symmetric eigenvalue problems and prove *a priori* eigenvalue error estimates. For a single eigenvalue, we prove an error estimate that depends mainly on just the approximability of the corresponding eigenfunction and provide explicit values for all constants. For a multiple eigenvalue we prove, in addition, apparently the first truly *a priori* error estimates that show the levels of the eigenvalue errors depending on approximability of eigenfunctions in the corresponding eigenspace. These estimates reflect a known phenomenon that different eigenfunctions in the corresponding eigenspace may have different approximabilities, thus resulting in different levels of errors for the approximate eigenvalues. For clustered eigenvalues, we derive eigenvalue error bounds that do not depend on the width of the cluster. Our results are readily applicable to the classical Ritz method for compact symmetric integral operators and to finite element method eigenvalue approximation for symmetric positive definite differential operators.

Key words. eigenvalue problem, operator, invariant subspace, multiple eigenvalues, clustered eigenvalues, approximation, Ritz method, Ritz value, finite element method, a priori error estimates, angles between subspaces

AMS(MOS) subject classifications. 65F35
August 6, 2004

1. Introduction. We revisit the classical subject of *a priori* eigenvalue error estimates for the Ritz–Galerkin approximation of symmetric eigenvalue problems, with applications to the Finite Element Method (FEM) eigenvalue approximations.

Early examples of *a priori* eigenvalue error estimates can be found, e.g., in [16]. Later, it became clear that the eigenvalue error is governed by the approximability of the exact eigenfunctions by the approximation space. In [5], Birkhoff, de Boor, Swartz, and Wendroff showed that the error for the jth eigenvalue is bounded by a constant times the sum of the norms squared of the approximation errors of the all eigenfunctions corresponding to the first j eigenvalues. In [21], Weinberger improved this result, showing that in the estimate for the relative eigenvalue error the constant simply equals to one; see Remark 2.3 for the exact formulation. Knyazev in [11], see also [8], further improved this result by replacing the norms of the approximation errors of individual eigenfunctions with the angle that measures the approximability of the invariant subspace spanned by these eigenfunctions. We reproduce this latter result by Knyazev in the present paper, in Theorem 2.4, and show that it is sharp.

The estimates of [5, 11, 21] suggest that the jth eigenvalue error depends on the approximability of all the eigenfunctions in the corresponding eigenspace, as well as of all the eigenfunctions corresponding to the previous eigenvalues. In reality, this is not the case, e.g., the first two eigenfunctions of the L-shaped membrane eigenvalue problem are singular because of the reentrant corner, but the third eigenfunction is analytic because of symmetry, and hence easily approximated, especially by the p-method (see [1]), thus resulting in a significantly better accuracy of approximation for the third eigenvalue compared to the first two. Vainikko [15] and Chatelin [7] derived estimates of the eigenvalue error mainly in terms of just the approximability

[*] Department of Mathematics, University of Colorado at Denver. Electronic mail address: knyazev@na-net.ornl.gov. WWW home page URL: http://www-math.cudenver.edu/~aknyazev. Partially supported by the National Science Foundation award DMS 0208773.

[†] Department of Mathematics, University of Maryland. Electronic mail address: jeo@math.umd.edu. WWW home page URL: http://www.wam.umd.edu/~jeo. Partially supported by the National Science Foundation award DMS 0341982.

of the eigenfunctions in the corresponding eigenspace. Moreover, they showed that the multiplicative constant in the estimate of the relative eigenvalue error approaches 1 under the approximability assumption on the family of the approximating spaces; see Section 3.3 for details. In [3], Babuška and Osborn determined that the closeness of the constant to 1 depends on the approximability of the operator of the original problem by the Ritz method; again, see Section 3.3.

Our first main results — Theorems 2.7 and 3.2 — clarify the estimate of [3] and improve the constant. Our proof is simpler, more transparent, and leads to an estimate with all constants explicitly given.

When the eigenvalue of interest is of multiplicity $q > 1$, different eigenfunctions in the corresponding eigenspace may have different approximabilities, thus resulting in different levels of error for the approximate eigenvalues. In other words, the q Ritz values, corresponding to the multiple eigenvalue, may approach the eigenvalue with different rates. It is important to have eigenvalue error estimates that capture this phenomenon.

The error bounds of Vainikko [15] and Chatelin [7] effectively require approximability of all eigenfunctions in the corresponding eigenspace, providing then an estimate for the largest eigenvalue error. In [2–4], Babuška and Osborn perform analysis that differentiate levels of eigenvalue error depending on approximability of different eigenfunctions in the eigenspace, but their estimates are not truly *a priori*, except for the estimate for the smallest eigenvalue error, which depends mainly on the approximability of the most easily approximated eigenfunction within the eigenspace.

Our results for multiple eigenvalues — Theorems 2.11 and 3.3 — clarify and improve these results of [2–4]. For example, if the eigenspace is spanned by three eigenfunctions of different approximation qualities, our results estimate the corresponding quality of each of the three Ritz values.

Error estimates for clustered eigenvalues are not well studied in the literature. The results presented in this paper are valid for clustered eigenvalues, as well as for multiple eigenvalues, and give error estimates that do not depend on the width of the cluster. Ovtchinnikov in [18] independently derives similar but somewhat cumbersome, estimates, which he calls “cluster robust.” Our estimates, compared to those of [18], are more compact and use less information.

In our proofs, we significantly use approximation error estimates for eigenspaces and invariant subspaces obtained by Knyazev in [13].

The paper is organized as follows: In Section 2, we formulate and prove in the abstract setting of a compact symmetric operator on a Hilbert space our first main result — Theorem 2.7 — an error estimate for a j-th eigenvalue mainly in terms of the approximation error of the corresponding eigenfunction, and we discuss the special features of the case of multiple eigenvalues and prove a generalization of Theorem 2.7 — Theorem 2.11 — that provides estimates for multiple and clustered eigenvalues. In Section 3, we apply our abstract results from Section 2, first briefly for integral operators in Subsection 3.1, and then, in Subsection 3.2, to eigenvalue error analysis in a context applicable for FEM eigenvalue approximation by variational Ritz method for second order symmetric positive definite differential operators. Our last main results — Theorem 3.2 and Theorem 3.3 — are proved in this subsection. Finally, in Subsection 3.3, we compare our results with those earlier known and clarify claims made in the introduction.

Preliminary results of this paper were presented at the meeting State Of The Art In Finite Element Method at the City University of Hong Kong in 1998.

2. Estimates for a compact symmetric operator.

2.1. An abstract eigenvalue problem.

We consider in this section a compact symmetric positive definite operator T defined on a real separable Hilbert space H, with inner product (u, v) and norm $\|u\| = \sqrt{(u,u)}$. The spectral theory of such operators is well known; see e.g. [9]. The spectrum consists of nonzero eigenvalues of finite multiplicity, together with 0, which is in the continuous spectrum. The eigenvectors can be chosen to be orthonormal. We denote the eigenvalues and corresponding eigenvectors of T by

$$\mu_1 \geq \mu_2 \geq \cdots > 0,$$

$$u_1, u_2, \ldots, \quad (u_i, u_j) = \delta_{i,j}.$$

We are interested in approximating the eigenpairs (μ_i, u_i) of T by the Ritz method. Given a finite dimensional subspace $\tilde{U}$ of H, referred to as the trial subspace, the Ritz approximation to T is the operator

$$\tilde{T} = (\tilde{Q}T)|_{\tilde{U}},$$

where $\tilde{Q}$ is the orthogonal projector onto $\tilde{U}$. The operator $\tilde{T}$ is symmetric positive definite. The eigenpairs of $\tilde{T}$ are called the Ritz pairs of T; we regard them as approximations of the eigenpairs of T. We denote the eigenvalues and corresponding eigenvectors of $\tilde{T}$ by

$$\tilde{\mu}_1 \geq \tilde{\mu}_2 \geq \cdots \geq \tilde{\mu}_n > 0, \text{ where } n = \dim \tilde{U},$$

$$\tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_n, \quad (\tilde{u}_i, \tilde{u}_j) = \delta_{i,j}.$$

The numbers $\tilde{\mu}_i$ are called the Ritz values and the vectors $\tilde{u}_i$ are called the Ritz vectors. In this paper we are specifically concerned with approximating the eigenvalues of T by Ritz values: $\mu_i \approx \tilde{\mu}_i$.

It is an immediate consequence of the max-min characterization of eigenvalues that

$$\tilde{\mu}_i \leq \mu_i, \quad i = 1, \ldots, n. \tag{2.1}$$

2.2. Principal angles between subspaces.

If M and N are nontrivial finite dimensional subspaces of H, we will quantify the approximability of M by N using the sine of the largest principal angle from M to N, which is defined by

$$\sin \angle\{M; N\} = \sup_{u \in M, \|u\|=1} \text{dist } (u, N) = \sup_{u \in M, \|u\|=1} \inf_{v \in N} \|u - v\|. \tag{2.2}$$

For nonzero vectors u and v, if $M = \text{span}\{u\}$, we write $\sin \angle\{u; N\}$ for $\sin \angle\{M; N\}$; and if $M = \text{span}\{u\}$ and $N = \text{span}\{v\}$, we write $\sin \angle\{u; v\}$ for $\sin \angle\{M; N\}$.

It is immediate that $0 \leq \sin \angle\{M; N\} \leq 1$ and that $\sin \angle\{M; N\} = 0$ if and only if $M \subseteq N$. If $\dim M > \dim N$, then $\sin \angle\{M; N\} = 1$. If $\dim M = \dim N < \infty$, then $\sin \angle\{M; N\} = \sin \angle\{N; M\}$. In the remainder of this paper, we will typically have $\dim M \leq \dim N$.

If P and Q are orthogonal projectors onto M and N, respectively, then the sine of the angle, $\sin\angle\{M;N\}$, can be expressed by

$$\begin{aligned}\sin\angle\{M;N\} &= \sup_{u\in M,\|u\|=1}\|(I-Q)u\| \\ &= \|(I-Q)P\| \\ &= \|P(I-Q)\|.\end{aligned} \tag{2.3}$$

The quantity $\sin\angle\{M;N\}$ is also denoted by $\delta(M,N)$, and

$$\hat{\delta}(M,N)=\max[\delta(M,N),\delta(N,M)]$$

is called the gap between M and N. It is easily seen that $\hat{\delta}(M,N)=\|P-Q\|$ and that $0\le\hat{\delta}(M,N)\le 1$. See [9, 14] for a discussion of $\hat{\delta}\{M;N\}$ and $\sin\angle\{M;N\}$.

We will need the following simple observations, cf. Lemma 3.4 of [6].

Lemma 2.1. *Let the subspace M be split into an orthogonal sum of subspaces $M=M_1\oplus M_2$. Then*

$$\sin^2\angle\{M;N\}\le\sin^2\angle\{M_1;N\}+\sin^2\angle\{M_2;N\}. \tag{2.4}$$

Proof. By (2.3), $\sin^2\angle\{M;N\}=\|P(I-Q)\|^2$. Let P_1 and P_2 denote the orthogonal projectors on subspaces M_1 and M_2, correspondingly, so that $P=P_1+P_2$, then by the definition of an operator norm and using the Pythagorean theorem,

$$\begin{aligned}\|P(I-Q)\|^2 &= \|(P_1+P_2)(I-Q)\|^2 \\ &= \sup_{x\in H,\|x\|\le 1}\|(P_1+P_2)(I-Q)x\|^2 \\ &= \sup_{x\in H,\|x\|\le 1}\left(\|P_1(I-Q)x\|^2+\|P_2(I-Q)x\|^2\right) \\ &\le \sup_{x\in H,\|x\|\le 1}\|P_1(I-Q)x\|^2+\sup_{x\in H,\|x\|\le 1}\|P_2(I-Q)x\|^2 \\ &= \|P_1(I-Q)\|^2+\|P_2(I-Q)\|^2.\end{aligned}$$

Now, using (2.3) for M_1 and M_2, we prove (2.4). □

Applying (2.4) recursively, we immediately obtain

Corollary 2.2. *Let vectors $\{u_i, i=1,\ldots,\dim M\}$ form an orthogonal basis for the subspace M. Then*

$$\sin^2\angle\{M;N\}\le\sum_i\sin^2\angle\{u_i;N\}. \tag{2.5}$$

We call angle $\angle\{M;N\}$ the largest since it is also well known, e.g., [14], that smaller angles between subspaces can be defined as follows. Using P and Q, the orthogonal projectors onto M and N, respectively, the sine of the largest angle by (2.3) equals to the largest singular value of the operator $(I-Q)P$. Introducing the notation $s_1((I-Q)P)\le s_2((I-Q)P)\le\ldots\le s_{\dim M}((I-Q)P)$ for the $\dim M$ *largest* singular values of the operator $(I-Q)P$, we define the ith angle from subspace M to subspace N using its sine: $\sin\angle_i\{M;N\}=s_i((I-Q)P)$, $i=1,\ldots,\dim M$, assuming that all angles lie on the closed interval $[0,\pi/2]$. The complete set of $\dim M$ angles from subspace M to subspace N gives detailed information of approximability of M

by N, e.g., if the smallest angle vanishes, the subspaces M and N have a nontrivial intersection.

Later in the paper we use the following property of angles (see [14])

$$\angle_j\{M;N\} = \inf_{L \subseteq M,\, \dim L = j} \angle\{L;N\},\ j = 1,\ldots,\dim M. \tag{2.6}$$

Finally, we will also need the following generalization of Corollary 2.2.

LEMMA 2.3. *Let vectors $\{u_i, i = 1,\ldots,\dim M\}$ form an orthogonal basis for the subspace M and are arranged in such a way that*

$$\angle\{u_1;N\} \leq \ldots \leq \angle\{u_{\dim M};N\}.$$

Then

$$\sin^2 \angle_j\{M;N\} \leq \sum_{i=1,\ldots,j} \sin^2 \angle\{u_i;N\},\ j = 1,\ldots,\dim M. \tag{2.7}$$

Proof. We deduce from (2.6) that

$$\sin^2 \angle_j\{M;N\} \leq \sin^2 \angle\{\operatorname{span}\{u_1,\ldots,u_j\};N\}.$$

Now, the statement of the lemma, (2.7), immediately follows from (2.5) applied to $M = \operatorname{span}\{u_1,\ldots,u_j\}$. □

2.3. Estimates based on the approximability of all previous eigenvectors. Sharp eigenvalue error estimates are usually derived under the assumption that the eigenvector corresponding to the eigenvalue being estimated is well approximated by the trial subspace.

We derive an estimate for the error in approximating μ_j, the jth eigenvalue of T, by $\tilde{\mu}_j$, the jth Ritz value of T, i.e., the jth eigenvalue of $\tilde{T}$. Let $U_{1,\ldots,j}$ denote the span of the eigenvectors $u_1,\ldots,u_j$, and let $P_{1,\ldots,j}$ be the orthogonal projector onto $U_{1,\ldots,j}$. For $u \neq 0$, let

$$\mu(u) = \frac{(Tu,u)}{(u,u)} = \frac{(u,u)_T}{(u,u)} \tag{2.8}$$

be the Rayleigh quotient associated with T. Here $(\cdot,\cdot)_T$ is a second inner product on H. We will refer to orthogonality in $(\cdot,\cdot)_T$ as T-orthogonality. Note that $\mu(u) > 0$ since T is positive definite.

Our first theorem is known; it was proved in [11] and reproduced in [8]. For the particular case $j = \dim \tilde{U}$, a different proof was then suggested in [10, 12]. Our proof is a modification of the latter proof for the general case $j \leq \dim \tilde{U}$ and is different from that of [11]. We provide it here since we use a similar, but more sophisticated approach, to prove our main results later in the paper.

THEOREM 2.4. *For $j = 1,2,\ldots,n = \dim \tilde{U}$ we have*

$$0 \leq \frac{\mu_j - \tilde{\mu}_j}{\mu_j} \leq \sin^2 \angle\{U_{1,\ldots,j};\tilde{U}\} = \|(I-\tilde{Q})P_{1,\ldots,j}\|^2. \tag{2.9}$$

Proof. We first note that (2.9) is trivially true if $\sin \angle\{U_{1,\ldots,j};\tilde{U}\} = 1$. Now suppose

$$\sin \angle\{U_{1,\ldots,j};\tilde{U}\} < 1. \tag{2.10}$$

Note that $\dim U_{1,\dots,j} = j \le \dim \tilde{U} < \infty$ by definition. From (2.3) we see that

$$\sin \angle\{U_{1,\dots,j}; \tilde{U}\} = \|(I - \tilde{Q})P_{1,\dots,j}\|.$$

Thus, with assumption (2.10), it follows from Theorem 6.34 in Chapter I in [9], that

$$\dim \tilde{Q}U_{1,\dots,j} = \dim U_{1,\dots,j} = j,$$

that $\tilde{Q}$ is an isomorphism between $U_{1,\dots,j}$ and $\tilde{Q}U_{1,\dots,j}$, where $\tilde{Q}U_{1,\dots,j}$ denotes the image of the subspace $U_{1,\dots,j}$ under the mapping $\tilde{Q}$, and finally that

$$\sin \angle\{\tilde{Q}U_{1,\dots,j}; U_{1,\dots,j}\} = \sin \angle\{U_{1,\dots,j}; \tilde{Q}U_{1,\dots,j}\} = \sin \angle\{U_{1,\dots,j}; \tilde{U}\}. \tag{2.11}$$

We choose a normalized vector $\bar{u} \in \tilde{Q}U_{1,\dots,j}$ such that

$$\mu(\bar{u}) = \min_{w \in \tilde{Q}U_{1,\dots,j} \setminus \{0\}} \mu(w),$$

where $\mu(\cdot)$ is the Rayleigh quotient introduced in (2.8), and consider the orthogonal decomposition,

$$\bar{u} = u + v, \;\; u \in U_{1,\dots,j} \text{ and } v \in U_{1,\dots,j}^{\perp}, \tag{2.12}$$

where $M^{\perp}$ denotes an orthogonal complement of subspace M. This decomposition is also T-orthogonal since $U_{1,\dots,j}$ is an invariant subspace of T.

Since $\bar{u} \in \tilde{Q}U_{1,\dots,j}$, $\|\bar{u}\| = 1$, and u is the orthogonal projection of $\bar{u}$ onto $U_{1,\dots,j}$, we see from (2.12) that

$$\begin{aligned} \|v\| &= \sin \angle\{\bar{u}; U_{1,\dots,j}\} \\ &\le \sin \angle\{\tilde{Q}U_{1,\dots,j}; U_{1,\dots,j}\}. \end{aligned} \tag{2.13}$$

It now follows from (2.10), (2.11) and (2.13) that $\|v\| < 1$; thus, $u \neq 0$ and $\mu(u)$ is defined.

We next establish the following chain of inequalities:

$$\mu(\bar{u}) \le \tilde{\mu}_j \le \mu_j \le \mu(u). \tag{2.14}$$

Using the max-min characterization of eigenvalues of $\tilde{T}$ we have

$$\begin{aligned} \mu(\bar{u}) &= \min_{w \in \tilde{Q}U_{1,\dots,j} \setminus \{0\}} \frac{(Tw, w)}{(w, w)} \\ &\le \max_{\substack{W \subset \tilde{U} \\ \dim W = j}} \min_{w \in W \setminus \{0\}} \frac{(\tilde{T}w, w)}{(w, w)} \\ &= \tilde{\mu}_j, \end{aligned}$$

which is the fist inequality in (2.14). The second inequality in (2.14), $\tilde{\mu}_j \le \mu_j$, is just the estimate (2.1) stated above. To prove the last inequality in (2.14), we write $u = \sum_{i=1}^{j} \alpha_i u_i$, and observe that

$$\mu(u) = \frac{(Tu, u)}{(u, u)} = \frac{\sum_{i=1}^{j} \alpha_i^2 \mu_i}{\sum_{i=1}^{j} \alpha_i^2} \ge \mu_j.$$

We deduce immediately from (2.12) that

$$\mu(\bar{u}) = \frac{(Tu,u)+(Tv,v)}{(u,u)+(v,v)}.$$

Using this identity, a direct calculation shows that

$$\mu(u) - \mu(\bar{u}) = \begin{cases} [\mu(\bar{u}) - \mu(v)]\dfrac{(v,v)}{(u,u)},\ v \neq 0 \\ 0,\ v = 0 \end{cases}. \tag{2.15}$$

For $v \neq 0$, it follows directly from (2.14) and (2.15) that

$$\begin{array}{rcll} 0 \leq \mu_j - \tilde{\mu}_j & \leq & \mu(u) - \mu(\bar{u}) & \\ & = & [\mu(\bar{u}) - \mu(v)]\dfrac{(v,v)}{(u,u)} & \\ & \leq & \tilde{\mu}_j \dfrac{\|v\|^2}{\|u\|^2} & \qquad\qquad (\text{since } \mu(v) > 0); \end{array}$$

hence, since $\|v\|^2 + \|u\|^2 = 1$,

$$0 \leq \frac{\mu_j - \tilde{\mu}_j}{\mu_j} \leq \|v\|^2. \tag{2.16}$$

If $v = 0$, then from (2.15) we see that $\mu(u) = \mu(\bar{u})$, which, together with (2.14), shows that $\tilde{\mu}_j = \mu_j$. Thus (2.16) is also valid for $v = 0$.

Finally, combining (2.11), (2.13) and (2.16), we obtain (2.9). □

Remark 2.1. *The estimate (2.9) in Theorem 2.4 can be also written in any of the following equivalent ways:*

$$0 \leq \frac{\mu_j - \tilde{\mu}_j}{\tilde{\mu}_j} \leq \tan^2 \angle\{U_{1,\dots,j}; \tilde{U}\}, \tag{2.17}$$

$$\tilde{\mu}_j \leq \mu_j \leq \frac{\tilde{\mu}_j}{1 - \sin^2 \angle\{U_{1,\dots,j}; \tilde{U}\}}. \tag{2.18}$$

The latter inequality, (2.18), exemplifies how our eigenvalue error estimate can be used to obtain two–sided bounds for eigenvalues, demonstrating the importance of having explicit constants in error estimates.

Remark 2.2. *The estimate of Theorem 2.4 is sharp in the following sense. For a given operator T and a fixed scalar $a, 0 \leq a < 1$, there exist a sequence of trial subspaces $\tilde{U}^{(k)}$, $k = 1, 2, \dots,$ such that*

$$0 \leq \frac{\mu_j - \tilde{\mu}_j^{(k)}}{\mu_j} \to a = \sin^2 \angle\{U_{1,\dots,j}; \tilde{U}^{(k)}\},\ j = 1, \dots, n,\ \text{when } k \to \infty.$$

Indeed, such subspaces are

$$\tilde{U}^{(k)} = \text{span}\{u_1 + \alpha u_{1+n+k}; u_2 + \alpha u_{2+n+k}; \dots; u_n + \alpha u_{2n+k}\},$$

where $\alpha^2 = a/(1-a)$. Then it is easy to see that

$$\frac{\mu_j - \tilde{\mu}_j^{(k)}}{\mu_j - \mu_{j+n+k}} = a,$$

and the statement follows since $\mu_{j+n+k} \to 0$ as $k \to \infty$. Thus, the statement of Theorem 2.4 cannot be improved without making additional assumptions or using additional terms.

REMARK 2.3. *Since by Corollary 2.2 we have*

$$\sin^2 \angle\{U_{1,\ldots,j}; \tilde{U}\} \le \sum_{i=1}^{j} \sin^2 \angle\{u_i; \tilde{U}\} = \sum_{i=1}^{j} \|(I - \tilde{Q})u_i\|^2,$$

the estimate

$$\frac{\mu_j - \tilde{\mu}_j}{\mu_j} \le \sum_{i=1}^{j} \|(I - \tilde{Q})u_i\|^2 \tag{2.19}$$

follows directly from Theorem 2.4. Estimate (2.19) is well-known (see, e.g., [19, 21]); on the right-hand side we have the sum of the squares of the approximation errors for the eigenvectors $u_1, \ldots, u_j$. If $j = 1$, the estimates (2.9) and (2.19) are identical.

2.4. Estimates based mainly on the approximability of the target eigenvector. Theorem 2.4 has a major weakness; namely, the right-hand side of estimate (2.9) for the target eigenvalue μ_j involves the approximability of all functions in $U_{1,\ldots,j}$. The result thus suggests that the eigenvalue error $(\mu_j - \tilde{\mu}_j)/\mu_j$ depends on the approximation errors for all eigenfunctions $u_1, \ldots, u_j$. We now mention two results implying that this is not the case; that, in fact, the ratio $(\mu_j - \tilde{\mu}_j)/\mu_j$ depends mainly on just the approximation error for u_j, the target eigenfunction. First, consider the following

LEMMA 2.5. *For $j = 1, 2, \ldots, n = \dim \tilde{U}$, the estimate holds:*

$$\begin{aligned}\frac{\mu_j - \tilde{\mu}_j}{\mu_j} &= \|(I - \tilde{P}_j)u_j\|^2 - \frac{1}{\mu_j}((I - P_j)\tilde{u}_j, T(I - P_j)\tilde{u}_j) \\ &\le \sin^2 \angle\{u_j, \tilde{u}_j\},\end{aligned} \tag{2.20}$$

where $\tilde{P}_j$ is the orthogonal projector onto span$\{\tilde{u}_j\}$.

Proof. Evidently, $\|(I - P_j)\tilde{u}_j\| = \|(I - \tilde{P}_j)u_j\| = \sin \angle\{u_j, \tilde{u}_j\}$; we remind the reader that $\|u_j\| = \|\tilde{u}_j\| = 1$. The first line of (2.20) follows from the chain of identities, taken from the proof of Lemma 3.5 of [6],

$$\begin{aligned}\frac{\mu_j - \tilde{\mu}_j}{\mu_j} &= (\tilde{u}_j, (I - \frac{1}{\mu_j}T)\tilde{u}_j) \\ &= ((I - P_j)\tilde{u}_j, (I - \frac{1}{\mu_j}T)(I - P_j)\tilde{u}_j) \\ &= \|(I - P_j)\tilde{u}_j\|^2 - \frac{1}{\mu_j}((I - P_j)\tilde{u}_j, T(I - P_j)\tilde{u}_j).\end{aligned}$$

Since $\mu_j > 0$ and the additional term $((I - P_j)\tilde{u}_j, T(I - P_j)\tilde{u}_j)$ is nonnegative, the second line of (2.20) follows immediately. □

Next consider

LEMMA 2.6. *If $(\tilde{u}_j, u_j) \neq 0$, the estimate holds:*

$$\begin{aligned}\frac{\mu_j - \tilde{\mu}_j}{\mu_j} &= \|(I - \tilde{Q})u_j\|^2 + \frac{1}{\mu_j}\left(T(I - \tilde{Q})u_j, \frac{(I - P_j)\tilde{u}_j}{\|P_j\tilde{u}_j\|}\right) \\ &\le \left(1 + \frac{\|(I - \tilde{Q})T\|}{\mu_j}\frac{\tan \angle\{u_j, \tilde{u}_j\}}{\sin \angle\{u_j, \tilde{U}\}}\right)\sin^2 \angle\{u_j, \tilde{U}\},\end{aligned} \tag{2.21}$$

where P_j is the orthogonal projector onto span$\{u_j\}$.

Proof. We use the argument from the proof of Theorem 4.1 in [3] (see also [17]) to establish the identity in the first line of (2.21). Denote $\overline{u}_j = P_j\tilde{u}_j$. By the assumption of the lemma, $\overline{u}_j \neq 0$, thus, $\overline{u}_j$ is an jth eigenvector of T, with length not larger than 1. For each j we have

$$\begin{aligned}
(\mu_j - \tilde{\mu}_j)(\overline{u}_j, \tilde{u}_j) &= (\mu_j\overline{u}_j, \tilde{\mu}_j) - (\overline{u}_j, \tilde{\mu}_j\tilde{u}_j)\\
&= (T\overline{u}_j, \tilde{u}_j) - (\overline{u}_j, \tilde{T}\tilde{u}_j)\\
&= (\overline{u}_j, (T - \tilde{T})\tilde{u}_j)\\
&= (\overline{u}_j, (I - \tilde{Q})T\tilde{u}_j)\\
&= (T(I - \tilde{Q})\overline{u}_j, \tilde{u}_j)\\
&= (T(I - \tilde{Q})\overline{u}_j, \overline{u}_j) + (T(I - \tilde{Q})\overline{u}_j, \tilde{u}_j - \overline{u}_j)\\
&= ((I - \tilde{Q})\overline{u}_j, T\overline{u}_j) + (T(I - \tilde{Q})\overline{u}_j, \tilde{u}_j - \overline{u}_j)\\
&= \mu_j\|(I - \tilde{Q})\overline{u}_j\|^2 + (T(I - \tilde{Q})\overline{u}_j, \tilde{u}_j - \overline{u}_j). \qquad (2.22)
\end{aligned}$$

Using

$$(\overline{u}_j, \tilde{u}_j) = (P_j\overline{u}_j, \tilde{u}_j) = (\overline{u}_j, P_j\tilde{u}_j) = \|\overline{u}_j\|^2,$$

and noting that $\overline{u}_j/\|\overline{u}_j\| = u_j$, (2.22) leads to identity in the first line of (2.21). The inequality in the second line of (2.21) follows directly from

$$\begin{aligned}
\left(T(I - \tilde{Q})u_j, \frac{(I - P_j)\tilde{u}_j}{\|P_j\tilde{u}_j\|}\right) &\leq \|T(I - \tilde{Q})\|\|(I - \tilde{Q})u_j\|\frac{\|(I - P_j)\tilde{u}_j\|}{\|P_j\tilde{u}_j\|}\\
&= \|(I - \tilde{Q})T\| \sin\angle\{u_j, \tilde{U}\} \tan\angle\{u_j, \tilde{u}_j\}.
\end{aligned}$$

□

It is informative to compare (2.20) with (2.21). The first term on the right-hand side of the first line of (2.20) is larger than that of (2.21). However, the second term in the first line of (2.20) is negative, and thus is dropped in the second line of (2.20). The second term on the right-hand side in the first line of (2.21), while generally not negative, in typical applications (when $\|(I - \tilde{Q})T\|$ is small) is of a smaller order compared the first term, in other words, the term added to 1 in the second line of (2.21) in such applications is small because of the multiplier $\|(I - \tilde{Q})T\|$. We conclude that both (2.20) and (2.21) suggest that $(\mu_j - \tilde{\mu}_j)/\mu_j$ depends mainly on the approximation error for u_j.

Both estimates (2.20) and (2.21), in addition to being dependent on the eigenfunction u_j, depend explicitly on the approximate eigenfunction $\tilde{u}_j$: (2.20) in the main term and (2.21) in the constant. Our next theorem is based on a novel alternative technique, where the approximate eigenfunction $\tilde{u}_j$ is not used in the proof and does not appear in the theorem statement.

Theorem 2.7. *For a fixed index j such that $1 \leq j \leq n = \dim\tilde{U}$, suppose that*

$$\min_{i=1,\dots,j-1} |\tilde{\mu}_i - \mu_j| \neq 0. \tag{2.23}$$

Then

$$\begin{aligned}
0 \leq \frac{\mu_j - \tilde{\mu}_j}{\mu_j} &\leq \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\|^2 \qquad (2.24)\\
&\leq \left(1 + \frac{\|(I - \tilde{Q})T\tilde{P}_{1,\dots,j-1}\|^2}{\min_{i=1,\dots,j-1}|\tilde{\mu}_i - \mu_j|^2}\right)\sin^2\angle\{u_j; \tilde{U}\},
\end{aligned}$$

where $\tilde{P}_{1,\dots,j-1}$ *is the orthogonal projector onto* $\tilde{U}_{1,\dots,j-1} = \mathrm{span}\{\tilde{u}_1, \dots, \tilde{u}_{j-1}\}$ *(if* $j = 1$*, we define* $\tilde{P}_{1,\dots,j-1} = 0$ *and do not use (2.23)).*

Proof. The case $j = 1$ is already covered by Theorem 2.4. Let $j > 1$. The operator $I - \tilde{Q} + \tilde{P}_{1,\dots,j-1}$ is an orthogonal projector and the vector u_j is normalized, so, $\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\| \le 1$. If $\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\| = 1$, the first estimate in (2.24) is trivially true since the relative eigenvalue error cannot be larger than one. Now we suppose

$$\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\| < 1. \tag{2.25}$$

Letting $U_j = \mathrm{span}\{u_j\}$, since $\dim U_j = 1$, the subspace $(\tilde{Q} - \tilde{P}_{1,\dots,j-1})U_j$ is also one dimensional by (2.25), and hence

$$\begin{aligned} \sin\angle\{U_j; (\tilde{Q} - \tilde{P}_{1,\dots,j-1})U_j\} &= \sin\angle\{u_j; (\tilde{Q} - \tilde{P}_{1,\dots,j-1})u_j\} \\ &= \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\| < 1 \end{aligned} \tag{2.26}$$

(this also follows from Theorem 6.34 in Chapter I of [9] applied to one dimensional subspaces).

We now choose a normalized vector $\bar{u} \in (\tilde{Q} - \tilde{P}_{1,\dots,j-1})U_j$, and introduce the orthogonal and T-orthogonal decomposition as in (2.12),

$$\bar{u} = u + v,\ u \in U_{1,\dots,j} \text{ and } v \in U_{1,\dots,j}^{\perp}.$$

We note that $(\tilde{Q} - \tilde{P}_{1,\dots,j-1})U_j$ is simply the span of the vector $\bar{u}$ and that

$$\begin{aligned} \|v\| &= \sin\angle\{\bar{u}; U_{1,\dots,j}\} \\ &\le \sin\angle\{\bar{u}; u_j\} \qquad\qquad (\text{since } u_j \in U_{1,\dots,j}) \\ &= \sin\angle\{(\tilde{Q} - \tilde{P}_{1,\dots,j-1})u_j; u_j\}. \end{aligned} \tag{2.27}$$

It now follows from (2.26) and (2.27) that $\|v\| < 1$; thus, $u \neq 0$ and $\mu(u)$ is defined.

We next establish the following chain of inequalities:

$$\mu(\bar{u}) \le \tilde{\mu}_j \le \mu_j \le \mu(u). \tag{2.28}$$

To prove the first inequality in (2.28), we proceed as follows. Let us introduce the space

$$\tilde{U}_{j,\dots,n} = \mathrm{span}\{\tilde{u}_j, \dots, \tilde{u}_n\}.$$

It is immediate that $\tilde{T}$ maps $\tilde{U}_{j,\dots,n}$ into itself, $\tilde{T}|_{\tilde{U}_{j,\dots,n}} : \tilde{U}_{j,\dots,n} \to \tilde{U}_{j,\dots,n}$ is symmetric, and the eigenvalues of $\tilde{T}|_{\tilde{U}_{j,\dots,n}}$ are

$$\tilde{\mu}_j \ge \tilde{\mu}_{j+1} \ge \dots \ge \tilde{\mu}_n.$$

Thus, using the max characterization of $\tilde{\mu}_j$, the largest eigenvalue of $\tilde{T}|_{\tilde{U}_{j,\dots,n}}$, we have

$$\begin{aligned} \mu(\bar{u}) &= \frac{(\tilde{T}\bar{u}, \bar{u})}{(\bar{u}, \bar{u})} \\ &= \frac{(\tilde{T}|_{\tilde{U}_{j,\dots,n}}\bar{u}, \bar{u})}{(\bar{u}, \bar{u})} \qquad\qquad (\text{since } \bar{u} \in \tilde{U}_{j,\dots,n}) \\ &\le \max_{w \in \tilde{U}_{j,\dots,n} \setminus \{0\}} \frac{(\tilde{T}|_{\tilde{U}_{j,\dots,n}} w, w)}{(w, w)} \\ &= \tilde{\mu}_j, \end{aligned}$$

as desired. The second inequality in (2.28), $\tilde{\mu}_j \le \mu_j$, is just the estimate (2.1). To prove the last inequality in (2.28), we write $u = \sum_{i=1}^{j} \alpha_i u_i$, and observe that

$$\mu(u) = \frac{(Tu,u)}{(u,u)} = \frac{\sum_{i=1}^{j} \alpha_i^2 \mu_i^2}{\sum_{i=1}^{j} \alpha_i^2} \ge \mu_j.$$

Returning back to the orthogonal decomposition of $\bar{u}$, we get

$$\mu(\bar{u}) = \frac{(Tu,u) + (Tv,v)}{(u,u) + (v,v)},$$

and hence, as in (2.15),

$$\mu(u) - \mu(\bar{u}) = \left\{ \begin{array}{l} [\mu(\bar{u}) - \mu(v)]\dfrac{(v,v)}{(u,u)},\ v \ne 0 \\ 0,\ v = 0 \end{array} \right. . \tag{2.29}$$

For $v \ne 0$ it follows immediately from (2.29) and (2.28) that

$$\begin{aligned} 0 \le \mu_j - \tilde{\mu}_j &\le \mu(u) - \mu(\bar{u}) \\ &= [\mu(\bar{u}) - \mu(v)]\frac{(v,v)}{(u,u)} \\ &\le \tilde{\mu}_j \frac{\|v\|^2}{\|u\|^2} \qquad\qquad \text{(since } \mu(v) > 0); \end{aligned}$$

therefore, since $\|v\|^2 + \|u\|^2 = 1$, we have

$$\frac{\mu_j - \tilde{\mu}_j}{\mu_j} \le \|v\|^2. \tag{2.30}$$

If $v = 0$, then from (2.29) we see that $\mu(u) = \mu(\bar{u})$, which, together with (2.28), shows the $\tilde{\mu}_j = \mu_j$. Thus (2.30) is also valid for $v = 0$.

Estimates (2.26), (2.27), and (2.30) prove the first estimate in (2.24) under assumption (2.25).

It remains to estimate

$$\begin{aligned} \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\|^2 &= \|(I - \tilde{Q})u_j + \tilde{P}_{1,\dots,j-1}u_j\|^2 \\ &= \|(I - \tilde{Q})u_j\|^2 + \|\tilde{P}_{1,\dots,j-1}u_j\|^2. \end{aligned} \tag{2.31}$$

Now, using Theorem 3.2 in [13] we have

$$\|\tilde{P}_{1,\dots,j-1}u_j\| \le \frac{\|(I - \tilde{Q})T\tilde{P}_{1,\dots,j-1}\|}{\tilde{d}} \|(I - \tilde{Q})u_j\|, \tag{2.32}$$

where

$$\tilde{d} = \inf_{\tilde{\nu} \in \sigma(\tilde{P}_{1,\dots,j-1}T\tilde{P}_{1,\dots,j-1}|_{\tilde{U}_{1,\dots,j-1}})} |\tilde{\nu} - \mu_j|.$$

It is easily shown that

$$\tilde{P}_{1,\dots,j-1}T\tilde{P}_{1,\dots,j-1}|_{\tilde{U}_{1,\dots,j-1}} = \tilde{T}|_{\tilde{U}_{1,\dots,j-1}},$$

and hence

$$\sigma(\tilde{P}_{1,\ldots,j-1}T\tilde{P}_{1,\ldots,j-1}|_{\tilde{U}_{1,\ldots,j-1}}) = \{\tilde{\mu}_1,\ldots,\tilde{\mu}_{j-1}\},$$

where $\tilde{U}_{1,\ldots,j-1}$ is the image of the orthogonal projector $\tilde{P}_{1,\ldots,j-1}$. Thus

$$\tilde{d} = \min_{i=1,\ldots,j-1} |\tilde{\mu}_i - \mu_j|. \tag{2.33}$$

Finally, combining the first estimate in (2.24) with (2.31), (2.32), and (2.33), and using $\|(I-\tilde{Q})u_j\| = \sin\angle\{u_j;\tilde{U}\}$, we obtain the second estimate in (2.24). $\square$

Since $\|(I-\tilde{Q}+\tilde{P}_{1,\ldots,j-1})u_j\| \le \|(I-\tilde{P}_j)u_j\|$, our new estimate (2.24) clearly improves (2.20). A direct comparison of the constants in (2.21) and (2.24) in a general case does not appear to be simple because of the unresolved dependence of (2.21) on $\tilde{u}_j$. However, we have

$$\begin{aligned}\frac{\|(I-\tilde{Q})T\tilde{P}_{1,\ldots,j-1}\|^2}{\min_{i=1,\ldots,j-1}|\tilde{\mu}_i-\mu_j|^2} &\le \frac{\|(I-\tilde{Q})T\|^2}{\min_{i=1,\ldots,j-1}|\tilde{\mu}_i-\mu_j|^2}\\ &\le \frac{\|(I-\tilde{Q})T\|}{\mu_j},\end{aligned}$$

assuming

$$\|(I-\tilde{Q})T\| \le \frac{\min_{i=1,\ldots,j-1}|\tilde{\mu}_i-\mu_j|^2}{\mu_j}. \tag{2.34}$$

Since $\tan\angle\{u_j,\tilde{u}_j\} \ge \sin\angle\{u_j,\tilde{U}\}$, we can conclude that our estimate (2.24) is sharper than (2.20) under the assumption (2.34). We note that in the FEM context assumption (2.34) is realistic as for typical problems $\|(I-\tilde{Q})T\|$ approaches 0 when the mesh parameter tends to 0.

Remark 2.4. *In the proof of Theorem 2.7, we can use alternative arguments instead of (2.32):*

$$\begin{aligned}\|\tilde{P}_{1,\ldots,j-1}u_j\|^2 &= \|\sum_{i=1}^{j-1}\tilde{P}_iu_j\|^2\\ &= \sum_{i=1}^{j-1}\|\tilde{P}_iu_j\|^2,\end{aligned}$$

where by Theorem 2.1 in [13]

$$\|\tilde{P}_iu_j\| \le \frac{\|T\tilde{u}_i - \tilde{\mu}_i\tilde{u}_i\|}{|\tilde{\mu}_i-\mu_j|}\sin\angle\{u_j;\tilde{U}\},$$

so we get

$$0 \le \frac{\mu_j-\tilde{\mu}_j}{\mu_j} \le \left(1+\sum_{i=1}^{j-1}\frac{\|T\tilde{u}_i-\tilde{\mu}_i\tilde{u}_i\|^2}{|\tilde{\mu}_i-\mu_j|^2}\right)\sin^2\angle\{u_j;\tilde{U}\},$$

which in some cases may provide an improvement of (2.24).

Remark 2.5. *A careful examination of the proof of Theorem 2.7 shows that in the proof of the first estimate in (2.24) we do not need to assume that the vector u_j*

is an eigenvector; i.e. if u_j is replaced with any normalized vector $u \in U_{1,\dots,j}$ the argument still holds and the first estimate in (2.24) turns into

$$0 \leq \frac{\mu_j - \tilde{\mu}_j}{\mu_j} \leq \inf_{u \in U_{1,\dots,j},\, \|u\|=1} \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u\|^2. \tag{2.35}$$

This constitutes a potential improvement of (2.24) — provided one can estimate the right-hand side of (2.35) using terms similar to those of the second estimate in (2.24).

Let us derive a simple estimate of the right-hand side of (2.35) based on the observation, which follows from dimensionality arguments, that there exist a nontrivial intersection $(\tilde{U}_{1,\dots,j-1})^\perp \cap U_{1,\dots,j}$. Clearly,

$$\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u\| = \|(I - \tilde{Q})u\|,\ u \in (\tilde{U}_{1,\dots,j-1})^\perp \cap U_{1,\dots,j}.$$

Restricting in such a way the choice of u in (2.35), we immediately obtain

$$0 \leq \frac{\mu_j - \tilde{\mu}_j}{\mu_j} \leq \inf_{u \,\in\, (\tilde{U}_{1,\dots,j-1})^\perp \cap U_{1,\dots,j},\ \|u\| = 1} \sin^2 \angle\{u_j; \tilde{U}\}, \tag{2.36}$$

that constitutes a clear improvement of (2.9). We note that (2.36) is not truly an a priori estimate since the right-hand side of it depends on the Ritz vectors $\tilde{u}_1, \dots, \tilde{u}_{j-1}$ that are not known a priori.

2.5. Corollaries of Theorems 2.4 and 2.7 for multiple eigenvalues. Here we address in details the case when the eigenvalue μ_j is multiple of multiplicity $q > 1$. Our Theorems 2.4 and 2.7 hold for multiple eigenvalues since we never assumed the eigenvalues are simple. However, the case of multiple eigenvalues has special features, which we want to highlight. Let us start with the simplest case, where we are interested only in estimates for the largest eigenvalue μ_1. We have from Theorem 2.4

Corollary 2.8. *Let*

$$\mu_1 = \mu_2 = \dots = \mu_q > \mu_{q+1}$$

and $q \leq n = \dim \tilde{U}$. For $j = 1, 2, \dots, q$ we have

$$\begin{aligned} 0 \leq \frac{\mu_1 - \tilde{\mu}_j}{\mu_1} &\leq \inf_{\substack{U_{1,\dots,j} \subset U_{1,\dots,q} \\ \dim U_{1,\dots,j} = j}} \sin^2 \angle\{U_{1,\dots,j}; \tilde{U}\} \\ &= \sin^2 \angle_j\{U_{1,\dots,q}; \tilde{U}\}. \end{aligned} \tag{2.37}$$

Proof. By the multiplicity assumption,

$$\frac{\mu_j - \tilde{\mu}_j}{\mu_j} = \frac{\mu_1 - \tilde{\mu}_j}{\mu_1}.$$

In Theorem 2.4, the subspace $U_{1,\dots,j}$ is the invariant subspace corresponding to the first j eigenvalues. Since μ_1 is of multiplicity q and $j \leq q$, $U_{1,\dots,j}$ can be viewed as an arbitrary j dimensional subspace of the eigenspace $U_{1,\dots,q}$. Thus, we have the freedom to choose $U_{1,\dots,j}$ to minimize the right hand side of the estimate (2.37). The final equality follows from (2.6). □

Estimate (2.37) has two important properties. First, it controls the error for *every* Ritz values corresponding to the first eigenvalue μ_1. Second, it shows that different Ritz values may have different approximation qualities, depending on approximability of the eigenspace $U_{1,\dots,q}$ by the trial subspace $\tilde{U}$ of the Ritz method, where the approximability is measured by the angles from $U_{1,\dots,q}$ to $\tilde{U}$ and, thus, can be estimated *a priori*.

In general, the multiple eigenvalue of interest may not be the largest:

$$\mu_{p-1} > \mu_p = \mu_{p+1} = \cdots = \mu_j = \cdots = \mu_{p+q-1} > \mu_{p+q}. \tag{2.38}$$

Applying Theorem 2.4, we obtain

Corollary 2.9. *Suppose (2.38) is satisfied and $p+q-1 \le n$. For any index $j = p, p+1, \dots, p+q-1$ we have*

$$0 \le \frac{\mu_p - \tilde{\mu}_j}{\mu_p} \le \inf_{\substack{U_{1,\dots,p-1} \subset U_{1,\dots,j} \subset U_{1,\dots,p+q-1} \\ \dim U_{1,\dots,j} = j}} \sin^2 \angle\{U_{1,\dots,j}; \tilde{U}\}.$$

Proof. The subspace $U_{1,\dots,j}$ has a fixed part $U_{1,\dots,p-1} \subset U_{1,\dots,j}$, but the rest of it we can choose within $U_{p,\dots,\min\{p+q-1,n\}}$ as we like. □

Corollary 2.9 preserves the desired properties of Corollary 2.8, i.e. it provides different estimates for every Ritz value of interest, but it requires approximability of all previous eigenvectors.

Let us now turn our attention to Theorem 2.7. The only relevant assumption in Theorem 2.7 is that (2.23) is satisfied so that the denominator in the constant in Theorem 2.7 is not zero. Let us analyze the likely behavior of this constant for the particular case $q = 2$ so that

$$\mu_{p-1} > \mu_p = \mu_{p+1} > \mu_{p+2}. \tag{2.39}$$

There are two relevant possibilities for j in Theorem 2.7: $j = p$ and $j = p+1$. Assuming that all Ritz values $\tilde{\mu}_i$ approximate the corresponding eigenvalues μ_i, which is typical for FEM applications (see Section 3.3 for details), we observe that in (2.23)

$$\min_{i=1,\dots,j-1} |\tilde{\mu}_i - \mu_j| \approx \mu_{j-1} - \mu_j.$$

Thus, if $j = p$, the denominator is asymptotically positive; specifically, it is asymptotically equal to $\mu_{p-1} - \mu_p$, and the estimate of Theorem 2.7 is asymptotically valid; while if $j = p+1$, the denominator in the constant in Theorem 2.7 asymptotically vanishes. This discussion demonstrates that Theorem 2.7 provides an asymptotically valid estimate only for one out of the $q = 2$ Ritz values. On the positive side, however, we can freely chose the eigenvector u_j within the eigenspace corresponding to μ_p to minimize the right hand side of (2.24). Let us reformulate Theorem 2.7 to reflect these observations.

Corollary 2.10. *Suppose that the eigenvalue μ_p, where $p > 1$, has multiplicity $q > 1$ so that (2.38) holds, and that $p+q-1 \le n$, and denote the corresponding eigenspace by $U_{p,\dots,p+q-1}$. As in Theorem 2.7, suppose that*

$$\min_{i=1,\dots,p-1} |\tilde{\mu}_i - \mu_p| \neq 0.$$

Then

$$\begin{aligned}0 \le \frac{\mu_p - \tilde{\mu}_p}{\mu_p} &\le \min_{u \in U_{p,\dots,p+q-1},\, \|u\|=1} \|(I - \tilde{Q} + \tilde{P}_{1,\dots,p-1})u\|^2 \\ &\le \left(1 + \frac{\|(I-\tilde{Q})T\tilde{P}_{1,\dots,p-1}\|^2}{\min_{i=1,\dots,p-1} |\tilde{\mu}_i - \mu_p|^2}\right) \min_{u \in U_{p,\dots,p+q-1},\, \|u\|=1} \sin^2 \angle\{u; \tilde{U}\} \\ &= \left(1 + \frac{\|(I-\tilde{Q})T\tilde{P}_{1,\dots,p-1}\|^2}{\min_{i=1,\dots,p-1} |\tilde{\mu}_i - \mu_p|^2}\right) \sin^2 \angle_1\{U_{p,\dots,p+q-1}; \tilde{U}\}. \qquad (2.40)\end{aligned}$$

Proof. We take $j = p$ in Theorem 2.7 and notice that we can choose u_j to be any normalized vector in the eigenspace $U_{p,\dots,p+q-1}$ and finally use (2.6). □

It is useful to compare Corollary 2.9 with Corollary 2.10. Corollary 2.9 gives different estimates for every Ritz value out of the q Ritz values corresponding to the multiple eigenvalue μ_p, but requires approximability of all previous eigenvectors. In Corollary 2.10, the approximability of previous eigenvectors appears only in the constant, but it gives an estimate only for the largest Ritz value out of the q.

We want to obtain a result that combines the advantages of Corollary 2.9 and Corollary 2.10 and removes their weaknesses. E.g., if $q = 3$ and the eigenspace corresponding to the triple eigenvalue μ_p is spanned by eigenfunctions of different approximation quality, we want to have three error estimates for μ_p reflecting it and not depending strongly on approximability of previous eigenfunctions.

2.6. A new estimate that covers multiple and clustered eigenvalues.

Our new result is a generalization of Theorem 2.7 that gives us the desired estimates for a multiple eigenvalue corresponding to an eigenspace spanned by eigenfunctions of different approximation quality. In addition, the new estimate also covers the case of clustered eigenvalues, i.e., the constant in the new estimate does not depend on the width of the eigenvalue cluster.

THEOREM 2.11. *For fixed indexes j and m satisfying $1 \le j \le n$ and $1 \le m \le j$, let $U_{j-m+1,\dots,j}$ be the m-dimensional invariant subspace corresponding to eigenvalues $\mu_{j-m+1} \le \dots \le \mu_j$ and $P_{j-m+1,\dots,j}$ be the orthogonal projector on $U_{j-m+1,\dots,j}$. If*

$$\min_{i=1,\dots,j-m} |\tilde{\mu}_i - \mu_j| \ne 0, \qquad (2.41)$$

then

$$\begin{aligned}0 \le \frac{\mu_j - \tilde{\mu}_j}{\mu_j} &\le \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-m})P_{j-m+1,\dots,j}\|^2 \\ &\le \left(1 + \frac{\|(I-\tilde{Q})T\tilde{P}_{1,\dots,j-m}\|^2}{\min_{i=1,\dots,j-m} |\tilde{\mu}_i - \mu_j|^2}\right) \|(I-\tilde{Q})P_{j-m+1,\dots,j}\|^2, \qquad (2.42)\end{aligned}$$

where $\tilde{P}_{1,\dots,j-m}$ is the orthogonal projector onto $\tilde{U}_{1,\dots,j-m} = \mathrm{span}\{\tilde{u}_1, \dots, \tilde{u}_{j-m}\}$ (if $j = m$ we set $\tilde{P}_{1,\dots,j-m} = 0$ and do not use (2.41). If $m = j$, the present theorem turns into Theorem 2.4; if $m = 1$, it turns into Theorem 2.7.

Proof. The operators $I - \tilde{Q} + \tilde{P}_{1,\dots,j-m}$ and $P_{j-m+1,\dots,j}$ are orthogonal projectors; thus, $\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-m})P_{j-m+1,\dots,j}\| \le 1$. If $\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-m})P_{j-m+1,\dots,j}\| = 1$, the first estimate in (2.42) is trivially true. Now we suppose

$$\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-m})P_{j-m+1,\dots,j}\| < 1. \qquad (2.43)$$

Then, since $\dim U_{j-m+1,\dots,j} = m$, the subspace $(\tilde{Q} - \tilde{P}_{1,\dots,j-m})U_{j-m+1,\dots,j}$ is also m dimensional by Theorem 6.34 in Chapter I in [9].

We choose a normalized vector $\bar{u}$ such that

$$\bar{u} \in (\tilde{Q} - \tilde{P}_{1,\dots,j-m})U_{j-m+1,\dots,j}, \ \mu(\bar{u}) = \min_{w \in (\tilde{Q}-\tilde{P}_{1,\dots,j-m})U_{j-m+1,\dots,j} \setminus \{0\}} \mu(w),$$

and introduce the orthogonal and T-orthogonal decomposition

$$\bar{u} = u + v, \ u \in U_{1,\dots,j}, \ v \in U_{1,\dots,j}^{\perp}.$$

Since $\bar{u} \in (\tilde{Q} - \tilde{P}_{1,\dots,j-m})U_{j-m+1,\dots,j}$, $\|\bar{u}\| = 1$, and $u = \bar{u} - v$ is the orthogonal projection of $\bar{u}$ onto $U_{1,\dots,j}$, we see similarly to (2.13) and (2.27) and using again Theorem 6.34 in Chapter I in [9] that

$$\begin{array}{rcl} \|v\| & = & \sin\angle\{\bar{u}; U_{1,\dots,j}\} \\ & \leq & \sin\angle\{\bar{u}; U_{j-m+1,\dots,j}\} \\ & \leq & \sin\angle\{(\tilde{Q} - \tilde{P}_{1,\dots,j-m})U_{j-m+1,\dots,j}; U_{j-m+1,\dots,j}\} \\ & = & \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-m})P_{j-m+1,\dots,j}\|. \end{array} \tag{2.44}$$

It now follows from (2.43) and (2.44) that $\|v\| < 1$; thus, $u \neq 0$ and $\mu(u)$ is defined.

We next prove the following chain of inequalities (cf. (2.14) and (2.28)):

$$\mu(\bar{u}) \leq \tilde{\mu}_j \leq \mu_j \leq \mu(u). \tag{2.45}$$

Indeed, the first inequality,

$$\mu(\bar{u}) = \min_{w \in (\tilde{Q}-\tilde{P}_{1,\dots,j-m})U_{j-m+1,\dots,j} \setminus \{0\}} \mu(w) \leq$$

$$\max_{\substack{W \subseteq \operatorname{Im}(\tilde{Q}-\tilde{P}_{1,\dots,j-m}) \\ \dim W = m}} \min_{w \in W \setminus \{0\}} \mu(w) = \tilde{\mu}_j,$$

follows from the min-max principle for Ritz values, since the dimension of the subspace $(\tilde{Q}-\tilde{P}_{1,\dots,j-m})U_{j-m+1,\dots,j}$ is m. The second inequality, $\tilde{\mu}_j \leq \mu_j$, is simply the estimate (2.1).The third inequality, $\mu_j \leq \mu(u)$, follows from the fact that $u \in U_{1,\dots,j}$ exactly as in the proof of (2.14) and (2.28).

The identity

$$\mu(\bar{u}) = \frac{(Tu, u) + (Tv, v)}{(u, u) + (v, v)}$$

can be rewritten (cf. (2.15) and (2.29)) as

$$\mu(u) - \mu(\bar{u}) = \left\{ \begin{array}{l} [\mu(\bar{u}) - \mu(v)]\dfrac{(v, v)}{(u, u)}, \ v \neq 0 \\ 0, \ v = 0 \end{array} \right. . \tag{2.46}$$

For $v \neq 0$, it follows directly from (2.45) and (2.46) that

$$\begin{array}{rcll} 0 \leq \mu_j - \tilde{\mu}_j & \leq & \mu(u) - \mu(\bar{u}) & \\ & = & [\mu(\bar{u}) - \mu(v)]\dfrac{(v, v)}{(u, u)} & \\ & \leq & \tilde{\mu}_j \dfrac{\|v\|^2}{\|u\|^2} & \text{(since } \mu(v) > 0); \end{array}$$

hence,

$$0 \le \frac{\mu_j - \tilde{\mu}_j}{\mu_j} \le \|v\|^2. \tag{2.47}$$

If $v = 0$, then from (2.46) we see that $\mu(u) = \mu(\bar{u})$, which, together with (2.45), shows that $\tilde{\mu}_j = \mu_j$. Thus, estimate (2.47) is also valid for $v = 0$.

Combining estimates (2.44) and (2.47), we obtain the first estimate in (2.42).

Finally, by Lemma 2.1,

$$\|(I-(\tilde{Q}-\tilde{P}_{1,\dots,j-m}))P_{j-m+1,\dots,j}\|^2 \le \|(I-\tilde{Q})P_{j-m+1,\dots,j}\|^2+\|\tilde{P}_{1,\dots,j-m}P_{j-m+1,\dots,j}\|^2.$$

The second term can be estimated using Theorem 3.2 of [13]:

$$\|\tilde{P}_{1,\dots,j-m}P_{j-m+1,\dots,j}\| \le \frac{\|(I-\tilde{Q})T\tilde{P}_{1,\dots,j-m}\|}{\min_{i=1,\dots,j-m}|\tilde{\mu}_i - \mu_j|}\|(I-\tilde{Q})P_{j-m+1,\dots,j}\|.$$

Combining the first estimate in (2.42) with the last two inequalities completes the proof. □

Alternatively, the arguments of Remark 2.4 with the help of Lemma 2.1 can be used to estimate the term $\|\tilde{P}_{1,\dots,j-m}P_{j-m+1,\dots,j}\|$, which results in a constant similar to that of the error estimate in Remark 2.4.

Remark 2.6. *Similarly to Remark 2.5, we note that the proof of of the first estimate in (2.42) of Theorem 2.11 allows replacing the orthoprojector $P_{j-m+1,\dots,j}$ with an orthoprojector P_L to any m–dimensional subspace L of $U_{1,\dots,j}$, so that the first estimate in (2.42) can be improved:*

$$0 \le \frac{\mu_j - \tilde{\mu}_j}{\mu_j} \le \inf_{L \subseteq U_{1,\dots,j},\ \dim L = m} \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-m})P_L\|^2. \tag{2.48}$$

It is not yet clear to how to use this fact to improve the second estimate in (2.42).

As in Remark 2.5, we can derive a simple estimate of the right-hand side of (2.48), using the fact, which follows from dimensionality arguments, that

$$\dim(\tilde{U}_{1,\dots,j-m})^{\perp} \cap U_{1,\dots,j} \ge m.$$

Evidently,

$$\|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-m})u\| = \|(I - \tilde{Q})u\|,\ u \in (\tilde{U}_{1,\dots,j-m})^{\perp} \cap U_{1,\dots,j},$$

so we derive from (2.48) that

$$0 \le \frac{\mu_j - \tilde{\mu}_j}{\mu_j} \le \inf_{L \subseteq (\tilde{U}_{1,\dots,j-m})^{\perp} \cap U_{1,\dots,j},\ \dim L = m} \|(I - \tilde{Q})P_L\|^2. \tag{2.49}$$

In FEM applications typically (because of the approximability assumption) we have $\dim\left((\tilde{U}_{1,\dots,j-m})^{\perp} \cap U_{1,\dots,j}\right) = m$ so the $\inf$ in (2.49) is then redundant.

We note that m is a free parameter in (2.49) and can be chosen arbitrarily, $1 \le m \le j$. We also note that (2.49) is not truly an a priori estimate since the right-hand side of it depends on the Ritz vectors $\tilde{u}_1, \dots, \tilde{u}_{j-m}$ that are not known a priori.

Let us now reformulate Theorem 2.11 in the context of the multiple eigenvalue in order to obtain a generalization of Corollary 2.10. Theorem 2.11 gives us enough flexibility to establish a different error estimate for every of q Ritz values corresponding to the multiple eigenvalue of multiplicity q:

COROLLARY 2.12. *Suppose that the eigenvalue μ_p, where $p > 1$, has multiplicity $q > 1$, so that (2.38) holds, and that $p + q - 1 \leq n$. Suppose that*

$$\min_{i=1,\ldots,p-1} |\tilde{\mu}_i - \mu_p| \neq 0.$$

Then, for $j = p, \ldots, p + q - 1$, we have

$$\begin{aligned} 0 \leq \frac{\mu_p - \tilde{\mu}_j}{\mu_p} &\leq \|(I - \tilde{Q} + \tilde{P}_{1,\ldots,p-1})P_{p,\ldots,j}\|^2 \\ &\leq \left(1 + \frac{\|(I - \tilde{Q})T\tilde{P}_{1,\ldots,p-1}\|^2}{\min_{i=1,\ldots,p-1} |\tilde{\mu}_i - \mu_p|^2}\right) \|(I - \tilde{Q})P_{p,\ldots,j}\|^2, \end{aligned} \tag{2.50}$$

where $\tilde{P}_{1,\ldots,p-1}$ is the orthogonal projector onto $\tilde{U}_{1,\ldots,p-1} = \operatorname{span}\{\tilde{u}_1, \ldots, \tilde{u}_{p-1}\}$ and $P_{p,\ldots,j}$ is the orthogonal projector onto any $j - p + 1$ dimensional subspace of the eigenspace $U_{p,\ldots,p+q-1}$ corresponding to the eigenvalue μ_p. The optimal choice of the projector $P_{p,\ldots,j}$ allows us to replace the term $\|(I - \tilde{Q})P_{p,\ldots,j}\|^2$ in estimate (2.50) with $\sin^2 \angle_{j-p+1}\{U_{p,\ldots,p+q-1}, \tilde{U}\}$.

Proof. We simply take $m = j - p + 1$ in Theorem 2.11. □

To see the improvement of Corollary 2.12 over Theorem 2.7, consider the following situation. Suppose μ_2 has multiplicity 2, so $p = q = 2$. Then

$$\min_{i=1,\cdots,p-1} |\tilde{\mu}_i - \mu_p| \approx \mu_1 - \mu_2 > 0,$$

provided $\tilde{\mu}_1$ is very close to μ_1. Taking $j = 2$ in Corollary 2.12 yields

$$\frac{\mu_2 - \tilde{\mu}_2}{\mu_2} \lesssim \left(1 + \frac{\|(I - \tilde{Q})T\tilde{P}_1\|^2}{(\mu_1 - \mu_2)^2}\right) \|(I - \tilde{Q})P_2\|^2; \tag{2.51}$$

while taking $j = 3$ yields

$$\frac{\mu_3 - \tilde{\mu}_3}{\mu_3} = \frac{\mu_2 - \tilde{\mu}_3}{\mu_2} \lesssim \left(1 + \frac{\|(I - \tilde{Q})T\tilde{P}_1\|^2}{(\mu_1 - \mu_2)^2}\right) \|(I - \tilde{Q})P_{2,3}\|^2. \tag{2.52}$$

In (2.51), the eigenvalues error is bounded by a constant that is slightly larger than 1 times the square of the best approximation error for u_2; while in (2.52), we have the square of the best approximation error for $\operatorname{span}\{u_2, u_3\}$ = the eigenspace for $\mu_2 = \mu_3$. Note that estimating $(\mu_3 - \tilde{\mu}_3)/\mu_3$ with Theorem 2.7 yields no asymptotically valid estimate (cf. the discussion preceding Corollary 2.10).

Results giving different estimates for $(\mu_p - \tilde{\mu}_j)/\mu_p$, $j = p, \ldots, p+q-1$ (cf. Corollaries 2.9 and 2.12) were first proved in [2], see also [3, 4]. Our presentation simplifies and clarifies the analysis in [2, 3], and provides explicit constants. In Section 3.3 we compare these results in details. For an example of a multiple eigenvalue with eigenvector of differing approximabilities, see [2, 4].

Let us finally highlight the opportunities that Theorem 2.11 provides for error estimates of clustered eigenvalues in the following situation. Let

$$\mu_1 > \mu_2 \approx \mu_3 > \mu_4,$$

and suppose we are interested in error estimates for μ_2 and μ_3, assuming that $\tilde{\mu}_1 \approx \mu_1$ and $\tilde{\mu}_2 \approx \mu_2$. We do not even need Theorem 2.11 to estimate the error for μ_2: Theorem 2.7 with $j = 2$ already gives us an asymptotically valid estimate (2.51), and the fact that μ_2 is clustered (or multiple as above) is irrelevant. Theorem 2.7 with $j = 3$ does not provide an asymptotically valid estimate for the error in μ_3 since the term $|\mu_3 - \tilde{\mu}_2| \approx 0$ appears in the denominator.

Applying Theorem 2.11 with $j = 3$ we have an option to choose the free parameter $m = 1$, 2, or 3. Taking $m = 1$ reduces Theorem 2.11 to Theorem 2.7, which does not work well in this situation as we just discussed. Taking $m = 2$ yields a good estimate

$$\frac{\mu_3 - \tilde{\mu}_3}{\mu_3} \lesssim \left(1 + \frac{\|(I - \tilde{Q})T\tilde{P}_1\|^2}{(\mu_1 - \mu_3)^2}\right) \|(I - \tilde{Q})P_{2,3}\|^2. \tag{2.53}$$

Taking $m = 3$ reduces Theorem 2.11 to Theorem 2.4,

$$\frac{\mu_3 - \tilde{\mu}_3}{\mu_3} \leq \|(I - \tilde{Q})P_{1,2,3}\|^2. \tag{2.54}$$

Comparing the right-hand sides of (2.53) and (2.54), we see that (2.53) provides a sharper estimate than (2.54) if $\mu_1 - \mu_3$ is large enough and u_1 cannot be well approximated by the trial subspace. To summarize, choosing different m in Theorem 2.11 allows us to reduce the constants in estimating errors for clustered eigenvalues at the cost of enlarging the invariant subspace that needs to be well approximated by the trial subspace. Note that in nether (2.53) nor (2.54) does the constant depend on the width of the eigenvalue cluster $\mu_2 \approx \mu_3$. Ovtchinnikov in [18] calls such estimates “cluster robust.”

3. Application of our abstract results.

We now consider the previous abstract results in two important contexts.

3.1. The classical Ritz method for integral operators.

Suppose we have an eigenvalue problem for a symmetric positive compact integral operator T defined on $H = L_2$. All our results apply immediately and provide relative eigenvalue error estimates for the largest eigenvalues in terms of L_2 approximability of the corresponding eigenfunctions.

3.2. The variational Galerkin method.

Suppose, as above, that H is a real separable Hilbert space with inner produce (u, v) and norm $\|u\| = \sqrt{(u, u)}$, and suppose we are given two symmetric bilinear forms $B(u.v)$ and $D(u, v)$ on $H \times H$. The bilinear form $B(u, v)$ is assumed to satisfy

$$|B(u, v)| \leq C_1\|u\|\|v\|, \text{ for all } u, v \in H \tag{3.1}$$

and

$$C_0\|u\|^2 \leq B(u.u), \text{ for all } u \in H, \text{ with } C_0 > 0. \tag{3.2}$$

It follows from (3.1) and (3.2) that $\|u\|_B = \sqrt{B(u, u)}$ and $\|u\|$ are equivalent norms on H. For the remainder of this section we use $B(u, v)$ and $\|u\|_B$ as the inner product and norm, respectively, on H, and denote the resulting space by H_B. We also measure all angles in H_B, i.e. with respect to $B(u, v)$. Regarding $D(u, v)$ we assume that

$$0 < D(u, u), \text{ for all nonzero vectors } u \in H \tag{3.3}$$

and that the norm

$$\|u\|_D = \sqrt{D(u,u)} \tag{3.4}$$

is compact with respect to $\|u\|$ or, equivalently, $\|u\|_B$, i.e., from any sequence that is bounded in $\|\cdot\|_B$, one can extract a subsequence that is Cauchy in $\|\cdot\|_D$.

We then consider the variationally formulated symmetric eigenvalue problem

$$\left\{ \begin{array}{l} \text{Seek } \lambda \in R \text{ and } 0 \neq u \in H_B \text{ satisfying} \\ B(u,v) = \lambda D(u,v), \text{ for all } v \in H_B \end{array} \right. . \tag{3.5}$$

Under the assumptions we have made, (3.5) has a sequence of eigenvalues

$$0 < \lambda_1 \leq \lambda_2 \leq \ldots \nearrow +\infty$$

and corresponding eigenvectors

$$u_1, u_2, \ldots,$$

which satisfy

$$B(u_i, u_j) = \lambda_i D(u_i, u_j) = \delta_{ij},\ i,j = 1,2,\ldots. \tag{3.6}$$

We will be interested in approximating the eigenpairs of (3.5) by the variational Ritz method. Toward this end, we suppose we are given a finite dimensional subspace $\tilde{U}$ of H_B, and consider the finite dimensional, variationally formulated eigenvalue problem

$$\left\{ \begin{array}{l} \text{Seek } \tilde{\lambda} \in R \text{ and } 0 \neq \tilde{u} \in \tilde{U} \text{ satisfying} \\ B(\tilde{u},v) = \lambda D(\tilde{u},v), \text{ for all } v \in \tilde{U} \end{array} \right. . \tag{3.7}$$

Problem (3.7), being a finite dimensional eigenvalue problem, has eigenvalues

$$0 < \tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \ldots \leq \tilde{\lambda}_n, \quad n = \dim \tilde{U},$$

and corresponding eigenvectors

$$\tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_n,$$

which satisfy

$$B(\tilde{u}_i, \tilde{u}_j) = \tilde{\lambda}_i D(\tilde{u}_i, \tilde{u}_j) = \delta_{ij}, \quad i,j = 1,\ldots,n.$$

We then view $\tilde{\lambda}_i$ as an approximation to $\lambda_i : \lambda_i \approx \tilde{\lambda}_i, \quad i = 1,\ldots,n$. It is a consequence of the min-max characterization of eigenvalues that

$$\lambda_i \leq \tilde{\lambda}_i, \quad i = 1,\ldots,n. \tag{3.8}$$

Next we introduce the operator $T : H_B \to H_B$ defined by

$$\left\{ \begin{array}{l} Tf \in H_B \\ B(Tf,v) = D(f,v), \text{ for all } v \in H_B \end{array} \right. \tag{3.9}$$

and the operator $\tilde{T} : \tilde{U} \to \tilde{U}$ defined by

$$\left\{ \begin{array}{l} \tilde{T}f \in \tilde{U},\ f \in \tilde{U}, \\ B(\tilde{T}f,v) = D(f,v), \text{ for all } v \in \tilde{U} \end{array} \right. . \tag{3.10}$$

The operator T is the solution operator for the "boundary value problem" corresponding to the eigenvalue problem (3.5). It follows immediately from our assumption that $\|\cdot\|_D$ is compact with respect to $\|\cdot\|_B$, that T is compact in H_B. Of course, $\tilde{T}$, being an operator on a finite dimensional space, is also compact. It follows directly from the definition (3.9) that T is symmetric and positive definite on H_B and from the definition (3.10) that $\tilde{T}$ is symmetric and positive definite on $\tilde{U}$ (with respect to $B(u,v)$). It is easily seen that, if, as above, $\tilde{Q}$ is the orthogonal projector of H_B onto $\tilde{U}$, then $\tilde{T} = (\tilde{Q}T)|_{\tilde{U}}$.

The eigenvalues of problem (3.5) and of the operator T are reciprocals:

$$\lambda_i = 1/\mu_i,\ i = 1, 2, \dots; \tag{3.11}$$

problem (3.5) and the operator T have the same eigenvectors u_i. Likewise, the eigenvalues of problem (3.7) and of the operator $\tilde{T}$ are reciprocals:

$$\tilde{\lambda}_i = 1/\tilde{\mu}_i,\ i = 1, 2, \dots, n; \tag{3.12}$$

problem (3.7) and the operator $\tilde{T}$ have the same eigenvectors $\tilde{u}_i$. As in the previous section, we choose $\{u_i\}$ and $\{\tilde{u}_i\}$ to be orthonormal systems, in the context of the present section, that is in H_B.

The FEM approximation of eigenvalue problems for symmetric differential operators can be viewed as a variational Ritz method; and the FEM eigenvalue errors can be estimated using the theorems of the previous section.

Because of (3.11) and (3.12), we can utilize Theorems 2.4 and 2.7, applied to T and $\tilde{T}$ on H_B, to estimate the eigenvalue error $(\tilde{\lambda}_i - \lambda_i)/\tilde{\lambda}_i$. Here $U_{1,\dots,j}$ denotes the span of the eigenvectors $u_1, \dots, u_j$ and $P_{1,\dots,j}$ is the H_B orthogonal projector onto $U_{1,\dots,j}$.

THEOREM 3.1. *For $j = 1, \dots, n = \dim \tilde{U}$ we have*

$$0 \le \frac{\tilde{\lambda}_j - \lambda_j}{\tilde{\lambda}_j} \le \sin^2 \angle_B \{U_{1,\dots,j}; \bar{U}\} = \|(I - \tilde{Q})P_{1,\dots,j}\|_B^2. \tag{3.13}$$

Proof. From (3.11) and (3.12) we have

$$\frac{\tilde{\lambda}_j - \lambda_j}{\tilde{\lambda}_j} = \frac{\mu_j - \tilde{\mu}_j}{\mu_j}. \tag{3.14}$$

Using (3.14) and applying Theorem 2.4, we obtain

$$0 \le \frac{\tilde{\lambda}_j - \lambda_j}{\tilde{\lambda}_j} \le \sin^2 \angle_B \{U_{1,\dots,j}; \tilde{U}\},\ j = 1, \dots, n,$$

which is the inequality in (3.13). The equality in (3.13) follows from (2.3) since we are working in the space H_B. □

REMARK 3.1. *By analogy with Remark 2.3, from Theorem 3.1 we get the following estimate, mathematically equivalent to estimate (2.19):*

$$0 \le \frac{\tilde{\lambda}_j - \lambda_j}{\tilde{\lambda}_j} \le \sum_{i=1}^{j} \|(I - \tilde{Q})u_i\|_B^2,$$

which can be rewritten as

$$0 \le \frac{\tilde{\lambda}_i - \lambda_i}{\lambda_i} \le \frac{\sum_{i=1}^{j} \|(I-\tilde{Q})u_i\|_B^2}{1-\sum_{i=1}^{j} \|(I-\tilde{Q})u_i\|_B^2}, \tag{3.15}$$

assuming that the denominator in the latter expression is positive. Estimate (3.15) is well-known (see, e.g., Theorem 2.1, Chapter 4 of [21]); a similar estimate is proved in [5].

Remark 3.2. *In our notation system, Strang and Fix in [20] Lemma 6.1 prove the following. By the min-max characterization of the Ritz values*

$$\begin{aligned} \tilde{\lambda}_j &= \min_{\tilde{W}\subset\tilde{U}, \dim\tilde{W}=j} \max_{\tilde{w}\in\tilde{W}\setminus\{0\}} \frac{(\tilde{w},\tilde{w})_B}{(\tilde{w},\tilde{w})_D} \\ &\le \max_{\tilde{w}\in\tilde{Q}U_{1,\ldots,j}\setminus\{0\}} \frac{(\tilde{w},\tilde{w})_B}{(\tilde{w},\tilde{w})_D} \\ &= \max_{u\in U_{1,\ldots,j}\setminus\{0\}} \frac{(\tilde{Q}u,\tilde{Q}u)_B}{(\tilde{Q}u,\tilde{Q}u)_D} \\ &\le \max_{u\in U_{1,\ldots,j}\setminus\{0\}} \frac{(u,u)_B}{(\tilde{Q}u,\tilde{Q}u)_D} \;(\textit{as } \tilde{Q} \textit{ is an orthoprojector in } H_B) \\ &\le \left(\max_{u\in U_{1,\ldots,j}\setminus\{0\}} \frac{(u,u)_B}{(u,u)_D}\right)\left(\max_{u\in U_{1,\ldots,j}\setminus\{0\}} \frac{(u,u)_D}{(\tilde{Q}u,\tilde{Q}u)_D}\right) \\ &= \lambda_j \max_{u\in U_{1,\ldots,j}\setminus\{0\}} \frac{(u,u)_D}{(\tilde{Q}u,\tilde{Q}u)_D}. \end{aligned} \tag{3.16}$$

A difficulty in estimating the last term is that the projector $\tilde{Q}$ is orthogonal in H_B, while the scalar products use the bilinear form D. However, in the FEM context, Strang and Fix [20] use (3.16) to obtain the correct order of convergence estimates for the eigenvalue errors in [20].

Knyazev in [11] uses similar arguments, but replaces the projector $\tilde{Q}$ with the projector $P_{1,\ldots,j}$, which is orthogonal with respect to both bilinear forms, $B(u,v)$ and $D(u,v)$, i.e.,

$$B(P_{1,\ldots,j}u,v) = B(u,P_{1,\ldots,j}v) \textit{ and } D(P_{1,\ldots,j}u,v) = D(u,P_{1,\ldots,j}v),$$

since it projects onto the span of eigenvectors of problem (3.5). The first step is the same as above:

$$\begin{aligned} \tilde{\lambda}_j &= \min_{\tilde{W}\subset\tilde{U}, \dim\tilde{W}=j} \max_{\tilde{w}\in\tilde{W}\setminus\{0\}} \frac{(\tilde{w},\tilde{w})_B}{(\tilde{w},\tilde{w})_D} \\ &\le \max_{\tilde{w}\in\tilde{Q}U_{1,\ldots,j}\setminus\{0\}} \frac{(\tilde{w},\tilde{w})_B}{(\tilde{w},\tilde{w})_D}. \end{aligned}$$

Now,

$$\begin{aligned} \frac{(\tilde{w},\tilde{w})_B}{(\tilde{w},\tilde{w})_D} &= \frac{(\tilde{w},P_{1,\ldots,j}\tilde{w})_D}{(\tilde{w},\tilde{w})_D} \frac{(\tilde{w},P_{1,\ldots,j}\tilde{w})_B}{(\tilde{w},P_{1,\ldots,j}\tilde{w})_D} \frac{(\tilde{w},\tilde{w})_B}{(\tilde{w},P_{1,\ldots,j}\tilde{w})_B} \\ &\le \lambda_j \frac{(\tilde{w},\tilde{w})_B}{(\tilde{w},P_{1,\ldots,j}\tilde{w})_B}, \end{aligned}$$

since the first fraction in the product of three fractions is bounded by one, and the second fraction is bounded by λ_j. *Finally,*

$$\frac{(\tilde{w}, P_{1,\dots,j}\tilde{w})_B}{(\tilde{w}, \tilde{w})_B} = 1 - \sin^2 \angle_B\{\tilde{w}; U_{1,\dots,j}\},$$

and, since $\tilde{w} \in \tilde{Q}U_{1,\dots,j}$, *using (2.11), we have*

$$\sin \angle_B\{\tilde{w}; U_{1,\dots,j}\} \le \sin \angle_B\{\tilde{Q}U_{1,\dots,j}; U_{1,\dots,j}\} = \sin \angle_B\{U_{1,\dots,j}; \tilde{U}\}.$$

Putting these results together, we obtain

$$\tilde{\lambda}_j \le \frac{\lambda_j}{1 - \sin^2 \angle_B\{U_{1,\dots,j}; \tilde{U}\}},$$

which is equivalent to (3.13). Thus, this gives a different proof of Theorem 3.1, see [8, 11] for additional information.

To formulate the next theorem — an analog of Theorem 2.7 — we recall that $\tilde{P}_{1,\dots,j-1}$ is the orthogonal projector of H_B onto $\tilde{U}_{1,\dots,j-1} = span\{\tilde{u}_1, \dots, \tilde{u}_{j-1}\}$, where $\tilde{u}_i$ are eigenvectors of (3.7).

THEOREM 3.2. *For a fixed index* j *such that* $1 \le j \le n = \dim \tilde{U}$, *suppose*

$$\min_{1,\dots,j-1} |\tilde{\lambda}_i - \lambda_j| \ne 0. \tag{3.17}$$

Then

$$\begin{aligned} 0 \le \frac{\tilde{\lambda}_j - \lambda_j}{\tilde{\lambda}_j} &\le \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\|_B^2 \\ &\le \left(1 + \max_{i=1,\dots,j-1} \frac{\tilde{\lambda}_i^2 \lambda_j^2}{|\tilde{\lambda}_i - \lambda_j|^2} \|(I - \tilde{Q})T\tilde{P}_{1,\dots,j-1}\|_B^2\right) \sin^2 \angle_B\{u_j; \tilde{U}\}. \end{aligned} \tag{3.18}$$

Proof. We can apply Theorem 2.7, obtaining

$$\begin{aligned} 0 \le \frac{\tilde{\lambda}_j - \lambda_j}{\tilde{\lambda}_j} &\le \|(I - \tilde{Q} + \tilde{P}_{1,\dots,j-1})u_j\|_B^2 \\ &\le \left(1 + \frac{\|(I - \tilde{Q})T\tilde{P}_{1,\dots,j-1}\|_B^2}{\min_{i=1,\dots,j-1} |\tilde{\mu}_i - \mu_j|^2}\right) \sin^2 \angle_B\{u_j; \tilde{U}\}. \end{aligned}$$

Finally, noting that

$$\frac{1}{\tilde{\mu}_i - \mu_j} = \frac{\lambda_j \tilde{\lambda}_i}{\lambda_j - \tilde{\lambda}_i},$$

we get the desired result. □

REMARK 3.3. *The arguments of Remark 2.4 can evidently be adopted to (3.18).*

Similarly, we can apply Theorem 2.11 to obtain

THEOREM 3.3. *For fixed indexes* j *and* m *satisfying* $1 \le j \le n$ *and* $1 \le m \le j$, *let* $U_{j-m+1,\dots,j}$ *be the* m*-dimensional invariant subspace corresponding to eigenvalues*

$\lambda_j \leq \ldots \leq \lambda_{j-m+1}$ and $P_{j-m+1,\ldots,j}$ be the H_B orthogonal projector on $U_{j-m+1,\ldots,j}$. If

$$\min_{i=1,\ldots,j-m} |\tilde{\lambda}_i - \lambda_j| \neq 0, \tag{3.19}$$

then

$$\begin{aligned} 0 \leq \frac{\tilde{\lambda}_j - \lambda_j}{\tilde{\lambda}_j} &\leq \|(I - \tilde{Q} + \tilde{P}_{1,\ldots,j-m})P_{j-m+1,\ldots,j}\|_B^2 \\ &\leq \left(\max_{i=1,\ldots,j-m} \frac{\tilde{\lambda}_i^2 \lambda_j^2}{|\tilde{\lambda}_i - \lambda_j|^2} \|(I - \tilde{Q})T\tilde{P}_{1,\ldots,j-m}\|_B^2 \right) \|(I - \tilde{Q})P_{j-m+1,\ldots,j}\|_B^2, \end{aligned} \tag{3.20}$$

where $\tilde{P}_{1,\ldots,j-m}$ is the H_B orthogonal projector onto $\tilde{U}_{1,\ldots,j-m} = \operatorname{span}\{\tilde{u}_1, \ldots, \tilde{u}_{j-m}\}$ (if $j = m$ we set $\tilde{P}_{1,\ldots,j-m} = 0$ and do not use (3.19)). If $m = j$, the present theorem turns into Theorem 3.1; if $m = 1$, it turns into Theorem 3.2.

Let us finally reformulate Theorem 3.3 in the context of the multiple eigenvalue by analogy with Corollary 2.12.

Corollary 3.4. *Suppose that the eigenvalue λ_p, where $p > 1$, has multiplicity $q > 1$, so that*

$$\lambda_{p-1} < \lambda_p = \lambda_{p+1} = \cdots = \lambda_{p+q-1} < \lambda_{p+q} \tag{3.21}$$

holds, and that $p + q - 1 \leq n$. Suppose that

$$\min_{i=1,\ldots,p-1} |\tilde{\lambda}_i - \lambda_p| \neq 0.$$

Then, for $j = p, \ldots, p + q - 1$, we have

$$\begin{aligned} 0 \leq \frac{\tilde{\lambda}_j - \lambda_p}{\tilde{\lambda}_j} &\leq \|(I - \tilde{Q} + \tilde{P}_{1,\ldots,p-1})P_{p,\ldots,j}\|_B^2 \\ &\leq \left(1 + \max_{i=1,\ldots,p-1} \frac{\tilde{\lambda}_i^2 \lambda_p^2}{|\tilde{\lambda}_i - \lambda_p|^2} \|(I - \tilde{Q})T\tilde{P}_{1,\ldots,p-1}\|_B^2 \right) \|(I - \tilde{Q})P_{p,\ldots,j}\|_B^2, \end{aligned} \tag{3.22}$$

where $\tilde{P}_{1,\ldots,p-1}$ is the H_B orthogonal projector onto $\tilde{U}_{1,\ldots,p-1} = \operatorname{span}\{\tilde{u}_1, \ldots, \tilde{u}_{p-1}\}$ and $P_{p,\ldots,j}$ is the H_B orthogonal projector onto any $j - p + 1$ dimensional subspace of the eigenspace $U_{p,\ldots,p+q-1}$ corresponding to the eigenvalue λ_p. The multiplier $\|(I - \tilde{Q})T\tilde{P}_{1,\ldots,p-1}\|_B^2$ in (3.22) can be replaced with $\sin^2 \angle_{j-p+1}\{U_{p,\ldots,p+q-1}, \tilde{U}\}$ by choosing the projector $P_{p,\ldots,j}$ in the optimal way, where the angle is measured in H_B.

3.3. Comparison with known asymptotic estimates for eigenvalues.

Estimate (3.18) should be compared with estimates of Vainikko [15], Chatelin [7], and Babuška and Osborn [3], which address a slightly different context that we now describe.

In addition to all assumptions of the previous subsection, let $\{U^h\}$ be a family of finite dimensional subspaces of H_B, depending on a parameter $h > 0$ called the mesh parameter. For a fixed h, we use $U^h = \tilde{U}$ as the trial subspace for the variational Ritz method. Let $Q^h = \tilde{Q}$ be the H_B orthogonal projector on U^h. We make the following approximability assumption on the family $\{U^h\}$:

$$\|(I-Q^h)u\|_{H_B} = \inf_{v^h\in U^h} \|u-v^h\|_{H_B} \to 0 \text{ as } h\to 0, \text{ for each } u \in H_B. \tag{3.23}$$

To be consistent with our new h-based notation, we denote the approximate eigenvalues by $\lambda_j^h = \tilde{\lambda}_j$ and the corresponding eigenvectors by $u_j^h = \tilde{u}_j$. It is well known that under assumption (3.23) we have

$$\lambda_j^h \to \lambda_j \text{ as } h \to 0 \text{ for each fixed } j.$$

Estimates of [3, 7, 15] that we refer to below are asymptotic in the usual sense, i.e. it is assumed that $h\to 0$ and negligible terms are dropped. These estimates are asymptotic upper bounds for the ratio $(\lambda_j^h-\lambda_j)/\lambda_j$, while our results are (nonasymptotic) inequalities involving the ratio $(\lambda_j^h-\lambda_j)/\lambda_j^h$ with a slightly different denominator. Since

$$\frac{\lambda_j^h-\lambda_j}{\lambda_j} = \frac{\lambda_j^h-\lambda_j}{\lambda_j^h} + \frac{(\lambda_j^h-\lambda_j)^2}{\lambda_j\lambda_j^h},$$

where the second term in the sum on the right can be asymptotically ignored, the results of [3, 7, 15] asymptotically estimate the same eigenvalue error as our results. In order to highlight the asymptotic nature of estimates of [3, 7, 15], we formulate them here using the $\lesssim$ rather than $\leq$.

We start our discussion with the case of a simple eigenvalue λ_j and later turn our attention to the case of multiple eigenvalues. The convergence rate for a simple eigenvalue is bounded by the following well known estimate

$$0 \leq \frac{\lambda_j^h-\lambda_j}{\lambda_j} \lesssim \left(1+r_j^h\right)\|(I-Q^h)u_j\|_B^2, \tag{3.24}$$

where

$$r_j^h \to 0 \text{ as } h \to 0;$$

see Subsection 18.6 (pp. 285–286) of [15] and Subsection 6.2 (pp. 315–317) of [7]. Babuška and Osborn [3] showed that

$$|r_j^h| \lesssim d_j \sup_{\|g\|_D=1} \|(I-Q^h)Tg\|_B^2 \to 0 \tag{3.25}$$

and that, cf. (2.21),

$$|r_j^h| \lesssim d_j \sup_{\|g\|_B=1} \|(I-Q^h)Tg\|_B \to 0, \tag{3.26}$$

where $d_j > 0$ are unknown generic constants.

Our present estimate (3.18) using the h notation takes the form

$$0 \leq \frac{\lambda_j^h-\lambda_j}{\lambda_j^h} \leq \left(1+r_j^h\right)\|(I-Q^h)u_j\|_B^2$$

with

$$r_j^h = \max_{i=1,\dots,j-1} \frac{(\lambda_i^h)^2\lambda_j^2}{|\lambda_i^h-\lambda_j|^2}\|(I-Q^h)TP_{1,\dots,j-1}^h\|_B^2. \tag{3.27}$$

The first multiplier in r_j^h in (3.27) is asymptotically (as $h \to 0$) a constant,

$$\frac{\lambda_{j-1}^2 \lambda_j^2}{|\lambda_{j-1} - \lambda_j|^2},$$

provided that the eigenvalue λ_j is simple. The second multiplier is bounded by

$$\begin{aligned} \|(I - Q^h) T P^h_{1,\dots,j-1}\|_B^2 &\leq \|(I - Q^h) T\|_B^2 \\ &= \sup_{\|g\|_B = 1} \|(I - Q^h) T g\|_B^2 \\ &\leq \frac{1}{\lambda_1} \sup_{\|g\|_D = 1} \|(I - Q^h) T g\|_B^2; \end{aligned}$$

thus, our estimate (3.27) is an improvement of both estimates (3.25) and (3.26) of [3]. Let us note that the denominator $|\lambda_{j-1} - \lambda_j|^2$ may be small, but the term in the numerator is bounded from above by a constant times $\sup_{\|g\|_D=1} \|(I - Q^h)Tg\|_B^2 \to 0$ as $h \to 0$.

Now suppose eigenvalue λ_p has multiplicity q, so that (3.21) holds, and let $P_{p,\dots,p+q-1}$ be the H_B orthogonal projector on the q dimensional eigenspace, corresponding to $\lambda_p = \lambda_{p+1} = \cdots = \lambda_{p+q-1}$ as in Corollary 3.4. Vainikko in Subsection 18.6 (pp. 285–286) of [15] and Chatelin in Subsection 6.2 (pp. 315–317) [7] prove that

$$(3.28)\; 0 \leq \frac{\lambda_j^h - \lambda_p}{\lambda_p} \lesssim \left(1 + r_p^h\right) \frac{\|(I - Q^h) P_{p,\dots,p+q-1} u_j^h\|_B^2}{\|P_{p,\dots,p+q-1} u_j^h\|_B^2},\; j = p, \dots, p+q-1,$$

where $r_p^h \to 0$ as $h \to 0$. An evident difficulty in using estimate (3.28) for *a priori* error analysis is that the approximate eigenfunctions u_{j+i-1}^h are not known *a priori*. If we consider the worst case, it leads to the following estimate, which is the same for all q Ritz values:

$$\begin{aligned} 0 \leq \frac{\lambda_j^h - \lambda_p}{\lambda_p} &\lesssim \left(1 + r_p^h\right) \|(I - Q^h) P_{p,\dots,p+q-1}\|_B^2 \\ &= \left(1 + r_p^h\right) \sin^2 \angle\{U_{p,\dots,p+q-1}; U_h\},\; j = p, \dots, p+q-1. \end{aligned} \tag{3.29}$$

Let us remind the reader that an angle without an index denotes the largest angle, according to our agreement in Subsection 2.2, and that in this and the previous subsections all angles are measured in H_B.

In some cases, see [2, 4] for an example, the eigenspace may be spanned by eigenfunctions of different approximation qualities, and it is interesting to analyze how this affects the error for different Ritz values. As mentioned in the Introduction, such results were first proved in by Babuška and Osborn in [2]. In [3], they completed such analysis for the smallest of the q Ritz values, proving the following error bound:

$$\begin{aligned} 0 \leq \frac{\lambda_p^h - \lambda_p}{\lambda_p} &\lesssim \left(1 + r_p^h\right) \inf_{u \in U_{p,\dots,p+q-1},\, \|u\|_B = 1} \|(I - Q^h) u\|_B^2 \\ &= \left(1 + r_p^h\right) \sin^2 \angle_1\{U_{p,\dots,p+q-1}; U_h\}, \end{aligned} \tag{3.30}$$

where

$$|r_p^h| \lesssim d_p \sup_{\|g\|_B = 1} \|(I - Q^h) T g\|_B, \tag{3.31}$$

with a generic constant $d_p > 0$. Estimate (3.30) depends mainly on the approximability of the most easily approximated eigenfunction in the eigenspace. Thus, estimates (3.29) and (3.30) represent two extremes: (3.29) uses the largest angle and estimates the largest error (thus effectively all q errors at once), while (3.30) uses the smallest angle and estimates only one, the smallest, eigenvalue error.

For the intermediate multiple eigenvalue error, Babuška and Osborn in [3] established the following estimate for $j = p, \dots, p+q-1$:

$$0 \le \frac{\lambda_j^h - \lambda_p}{\lambda_p} \lesssim \left(1 + r_p^h\right) \inf_{\substack{u \in U_{p,\dots,p+q-1}, \\ u \in \left(U^h_{p,\dots,j-1}\right)^{\perp_B}, \\ \|u\|_B = 1}} \|(I - Q^h)u\|_B^2, \tag{3.32}$$

with r_p^h given by (3.31), where the orthogonal complement $\left(U^h_{p,\dots,j-1}\right)^{\perp_B}$ is taken in H_B. In [4], estimate (3.32) appears in a slightly weaker form, without (3.31).

We note that the constraints on u in (3.32) are similar to those in (2.36) except that (2.36) involves orthogonalization to all previous Ritz vectors, while (3.32) only needs orthogonalization to previous Ritz vectors corresponding to the multiple eigenvalue under the consideration. Both (2.36) and (3.32) are not truly *a priori* estimates since their right-hand sides depends on Ritz vectors that are not known *a priori*.

In contrast, our estimate (3.22) is in the form

$$0 \le \frac{\lambda_j^h - \lambda_p}{\lambda_j^h} \le \left(1 + r_p^h\right) \sin^2 \angle_{j-p+1}\{U_{p,\dots,p+q-1}, U^h\}, \tag{3.33}$$

where $j = p, \dots, p+q-1$ and

$$r_p^h = \max_{i=1,\dots,p-1} \frac{(\lambda_i^h)^2 \lambda_p^2}{|\lambda_i^h - \lambda_p|^2} \|(I - Q^h) T P^h_{1,\dots,p-1}\|_B^2.$$

We have already discussed that our ratio $(\lambda_j^h - \lambda_p)/\lambda_j^h$ is asymptotically the same the ratio $(\lambda_j^h - \lambda_p)/\lambda_p$ used in (3.32) and shown that our expression for r_p^h is better that that given by estimate (3.31): the constant is explicitly written and the h-dependent part is smaller. Let us turn our attention to the main term of the right-hand side of (3.33), namely, the $\sin^2 \angle_{j-p+1}\{U_{p,\dots,p+q-1}, U^h\}$ multiplier.

We first highlight again that this multiplier can be estimated *a priori* since it does not depend on Ritz vectors, contrary to main term of the estimate (3.32). Second, we can directly compare the main terms in (3.32) and (3.33). Indeed, by (2.6), and since $\dim\{\left(U^h_{p,\dots,j-1}\right)^\perp \cap U_{p,\dots,p+q-1}\} \ge j - p + 1$, we have for $j = p, \dots, p+q-1$:

$$\begin{aligned}
\angle_{j-p+1}\{U_{p,\dots,p+q-1}, U^h\} &= \inf_{L \subseteq U_{p,\dots,p+q-1},\, \dim L = j-p+1} \angle\{L; U^h\} \\
&\le \angle\{\left(U^h_{p,\dots,j-1}\right)^\perp \cap U_{p,\dots,p+q-1}; U^h\} \\
&= \inf_{\substack{u \in U_{p,\dots,p+q-1}, \\ u \in \left(U^h_{p,\dots,j-1}\right)^\perp, \\ \|u\|_B = 1}} \|(I - Q^h)u\|_B^2,
\end{aligned}$$

so our estimate (3.33) is sharper than (3.32).

Using the term $\sin^2 \angle_{j-p+1}\{U_{p,\dots,p+q-1}, U^h\}$ has yet another advantage: namely, it permits the application of (2.7). Suppose the vectors $\{u_i, i = p, \dots, p+q-1\}$ form an orthogonal basis for the subspace $U_{p,\dots,p+q-1}$ and are arranged in such a way that

$$\angle\{u_p; U^h\} \leq \ldots \leq \angle\{u_{p+q-1}; U^h\}.$$

Then, by (2.7),

$$\sin^2 \angle_{j-p+1}\{U_{p,\dots,p+q-1}; U^h\} \leq \sum_{i=p,\dots,j} \sin^2 \angle\{u_i; U^h\},\ j = p, \dots, p+q-1.$$

In other words, if the eigenspace $U_{p,\dots,p+q-1}$ is spanned by eigenfunctions of different approximation qualities, our result assesses the quality of each of the Ritz values corresponding to the multiple eigenvalue.

Conclusions. We derive eigenvalue error bounds for the Ritz method that have several novel features:

- For a single eigenvalue, our estimates improve those previously known and provide explicit values for all constants.
- For a multiple eigenvalue we prove, in addition, apparently the first truly *a priori* error estimates that show the levels of the eigenvalue errors depending on approximability of eigenfunctions in the corresponding eigenspace.
- For clustered eigenvalues, our results provide elegant eigenvalue error bounds that do not depend on the width of the cluster.

In the FEM eigenvalue approximation context, our results allow one to *a priori* choose the mesh size properly and to intelligently predict *a priori* the optimal mesh refinement using information on the eigenfunctions smoothness determined by coefficients discontinuities or irregularities in the computational domain.

References.

[1] I. Babuška, B. Q. Guo, and J. E. Osborn. Regularity and numerical solution of eigenvalue problems with piecewise analytic data. *SIAM J. Numer. Anal.*, 26(6):1534–1560, 1989.

[2] I. Babuška and J. E. Osborn. Estimates for the errors in eigenvalue and eigenvector approximation by Galerkin methods, with particular attention to the case of multiple eigenvalues. *SIAM J. Numer. Anal.*, 24(6):1249–1276, 1987.

[3] I. Babuška and J. E. Osborn. Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems. *Math. Comp.*, 52(186):275–297, 1989.

[4] I. Babuška and J. E. Osborn. Eigenvalue problems. In *Handbook of Numerical Analysis, Vol. II*, pages 641–787. North-Holland, Amsterdam, 1991.

[5] G. Birkhoff, C. de Boor, B. Swartz, and B. Wendroff. Rayleigh-Ritz approximation by piecewise cubic polynomials. *SIAM J. Numer. Anal.*, 3:188–203, 1966.

[6] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. An iterative method for elliptic problems on regions partitioned into substructures. *Math. Comp.*, 46(174):361–369, 1986.

[7] F. Chatelin. *Spectral approximations of linear operators.* Academic Press, New York, 1983.

[8] E. G. D'yakonov. *Optimization in solving elliptic problems.* CRC Press, Boca Raton, FL, 1996. Translated from the 1989 Russian original; translation edited and with a preface by Steve McCormick.

[9] T. Kato. *Perturbation Theory for Linear Operators.* Springer–Verlag, New–York, 1976.

[10] A. V. Knyazev. *Computation of eigenvalues and eigenvectors for mesh problems: algorithms and error estimates.* Dept. Numerical Math. USSR Academy of Sciences, Moscow, 1986. (In Russian).

[11] A. V. Knyazev. Sharp a priori error estimates of the Rayleigh-Ritz method without assumptions of fixed sign or compactness. *Mathematical Notes*, 38(5–6):998–1002, 1986.

[12] A. V. Knyazev. Convergence rate estimates for iterative methods for mesh symmetric eigenvalue problem. *Soviet J. Numerical Analysis and Math. Modelling*, 2(5):371–396, 1987.

[13] A. V. Knyazev. New estimates for Ritz vectors. *Math. Comp.*, 66(219):985–995, 1997.

[14] A. V. Knyazev and Merico E. Argentati. Principal angles between subspaces in an A-based scalar product: Algorithms and perturbation estimates. *SIAM J. Sci. Comput.*, 23(6):2009–2041, 2002.

[15] M. A. Krasnosel'skii, G. M. Vainikko, P. P. Zabreiko, Ya. B. Rutitskii, and Y. Ya. Stetsenko. *Approximate Solutions of Operator Equations.* Wolters-Noordhoff, Groningen, 1972. Translated from Russian.

[16] M. N. Krylov. Les méthodes de solution approachée des problèmes de la physique mathématique. *Mémr. Sci Math. Gauthier-Villars, Paris*, XLIX:68, 1931.

[17] J. E. Osborn. Spectral approximation for compact operators. *Math. Comput.*, 29:712–725, 1975.

[18] E. Ovtchinnikov. Cluster robust error estimates for the Rayleigh–Ritz approximation II: Estimates for eigenvalues. Published as CU-Denver CCM report 210, $http://math.cudenver.edu/ccm/reports/rep210.pdf.gz$ Accepted to Linear Algebra and Applications, 2004.

[19] B. N. Parlett. *The symmetric eigenvalue problem.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Corrected reprint of the 1980 original.

[20] G. Strang and G. Fix. *An Analysis of the Finite Element Method.* Prentice-Hall, 1973.

[21] H. F. Weinberger. *Variational methods for eigenvalue approximation.* Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974. Based on a series of lectures presented at the NSF-CBMS Regional Conference on Approximation of Eigenvalues of Differential Operators, Vanderbilt University, Nashville, Tenn., June 26–30, 1972, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 15.