NEW ESTIMATES FOR RITZ VECTORS

ANDREW V. KNYAZEV *

Abstract. The following estimate for the Rayleigh–Ritz method is proved:

$$|\tilde\lambda - \lambda||(\tilde u, u)| \le \|A\tilde u - \tilde\lambda \tilde u\| \sin \angle\{u; \tilde U\}, \ \|u\| = 1.$$

Here A is a bounded self-adjoint operator in a real Hilbert/euclidian space, $\{\lambda, u\}$ one of its eigenpairs, $\tilde U$ a trial subspace for the Rayleigh–Ritz method, and $\{\tilde\lambda, \tilde u\}$ a Ritz pair. This inequality makes it possible to analyze the fine structure of the error of the Rayleigh–Ritz method, in particular, it shows that $|(\tilde u, u)| \le C\epsilon^2$, if an eigenvector u is close to the trial subspace with accuracy ϵ and a Ritz vector $\tilde u$ is an ϵ approximation to another eigenvector, with a different eigenvalue. Generalizations of the estimate to the cases of eigenspaces and invariant subspaces are suggested, and estimates of approximation of eigenspaces and invariant subspaces are proved.

Key words. eigenvalue problem, Rayleigh–Ritz method, approximation, error estimate

AMS(MOS) subject classifications. 65F35

1. Introduction.

Let A be a bounded self-adjoint operator in a real Hilbert (or euclidian) space, and $\{\lambda, u\}$ be an eigenpair,

$$Au = \lambda u, \ \|u\| = 1.$$

A Ritz pair, $\{\tilde\lambda_j, \tilde u_j\}$, is, by definition, an eigenpair of the operator $\tilde A = (\tilde Q A)\mid_{\tilde U}$, where $\tilde Q$ is an orthoprojector on the *trial* subspace $\tilde U$ and $\mid_{\tilde U}$ means the restriction of the operator to its invariant subspace $\tilde U$;

$$\tilde A \tilde u = \tilde\lambda \tilde u, \ \tilde u \in \tilde U, \ \|\tilde u\| = 1.$$

The behavior of a Ritz vector as a function of the trial subspace is complicated and still not completely studied. For example, let an eigenvector u be close to the trial subspace with accuracy ϵ and a Ritz vector $\tilde u$ be an ϵ approximation to another eigenvector, with a different eigenvalue. Either of the two assumptions leads to the trivial estimate

$$|(\tilde u, u)| \le C\epsilon.$$

Do they together give

$$|(\tilde u, u)| \le C\epsilon^2?$$

The following basic estimate gives the positive answer to this question.

THEOREM 1.1. *If* $\lambda \ne \tilde\lambda$, *then*

$$|(\tilde u, u)| \le \frac{\|A\tilde u - \tilde\lambda \tilde u\|}{|\tilde\lambda - \lambda|} \|(I - \tilde Q)u\|.$$

* Department of Mathematics, University of Colorado at Denver. Electronic mail address: knyazev@na-net.ornl.gov. This research was supported by the National Science Foundation under grant NSF-CCR-9204255 and was performed while the author was visiting the Courant Institute.

In the next section a somewhat more general formulation of the basic estimate is presented in Theorem 2.1 along with two dual proofs.

We note that it is often important to analyze several eigenspace components together rather than just one component, as in Theorem 1.1 and Theorem 2.1. Such analysis is carried out in section 3. The purpose is to make clear a behavior of Ritz vectors when $\|(I-\tilde{Q})A\tilde{Q}\|$ is small, e.g. for a case when the trial subspace is close to an invariant subspace of the operator A. Interestingly, generalizations of the two dual proofs of Theorem 2.1 lead to two different statements, Theorem 3.1 and Theorem 3.2. Theorem 3.1 shows that the approximation error of an eigenvector by the corresponding Ritz vector is essentially orthogonal to this invariant subspace. Theorem 3.2 leads to a dual statement that the orthoprojection of any eigenvector from the invariant subspace onto the trial subspace essentially coincides with a Ritz vector; see also Saad [12].

Asymptotic statements of this kind were formulated in [9]; the proof, based on a perturbation theory, was published in Russian in Knyazev [7]. They now turn out to be direct consequences of the simple estimates of Theorem 3.1 and Theorem 3.2.

There is a classical fact that the Ritz procedure, applied to a linear equation with a selfadjoint positive operator, produces an approximation in a trial subspace which is just the orthogonal projection of the solution onto the subspace with respect to the "energy" scalar product. Perhaps, by analogy with this fact, there was a well-known naive statement that the Rayleigh–Ritz procedure, applied to an eigenvalue problem, gives an approximation (a Ritz vector) to an eigenvector which must be an orthogonal projection of the eigenvector onto the trial subspace, because of "the optimality" of the Rayleigh–Ritz procedure; see a discussion in Parlett [11]. Such a point of view was popular among specialists of structural analysis in seventies, who used subspace iterations – the method of computing a sequence of subspaces that tends to an invariant subspace. The present paper shows that the engineers were right, in their own way, though strictly speaking, the statement is not mathematically correct, see Corollaries 3.2 and 4.2.

In the section 4, further extensions, of the main Theorem 2.1, are made for the case of invariant subspaces of A and $\tilde{A}$ instead of eigenspaces. Such generalizations are of particular importance for the analysis of approximations of an invariant subspace, corresponding to a cluster of eigenvalues, or even an interval of the continuous spectrum of A; see [4]. Two estimates of approximations of an invariant subspace are proved, of Corollary 4.1 and Theorem 4.3, as simple consequences of the main results. The statement of Corollary 4.1 is already known, see Davis and Kahan [2] and cf. also Theorem 11.7.11 of [11]. Theorem 4.3 is a generalization of the Saad's estimate [12, 11] for the case of invariant subspaces.

Let us finally mention that $\|(I-\tilde{Q})A\tilde{Q}\|$ can be small not only in the case when the trial subspace is close to an invariant subspace of the operator A. If the operator A is compact and for a sequence of trial subspaces the corresponding orthogonal projectors $\tilde{Q}$ strongly converge to the identity operator, then $\|(I-\tilde{Q})A\tilde{Q}\|$, in fact even $\|(I-\tilde{Q})A\|$, tends to zero [3]. Such a situation is typical for the Rayleigh–Ritz method applied for approximation of low eigenpairs of differential operators, e.g. using finite elements [13, 1]. However, if the operator A is not compact, then $\|(I-\tilde{Q})A\tilde{Q}\|$ is not necessarily small, even for $\tilde{Q}$ strongly converging to the identity [4]. The importance of $\|(I-\tilde{Q})A\tilde{Q}\|$ is based on the representation

$$A = \tilde{Q}A\tilde{Q} + (I-\tilde{Q})A(I-\tilde{Q}) + (I-\tilde{Q})A\tilde{Q} + \tilde{Q}A(I-\tilde{Q}),$$

and the fact that

$$\|(I-\tilde{Q})A\tilde{Q} + \tilde{Q}A(I-\tilde{Q})\| \leq \|(I-\tilde{Q})A\tilde{Q}\| = \|\tilde{Q}A(I-\tilde{Q})\|.$$

Considering the last two terms in the representation of A as a perturbation and using Theorem 4.10, p. 291 of [6] lead to the following important estimate, cf. (11.5.1) of [11],

$$\text{dist}\left\{\sigma(A), \sigma(\tilde{A}) \bigcup \sigma\left(\{(I-\tilde{Q})A(I-\tilde{Q})\}|_{\tilde{U}^{\perp}}\right)\right\} \leq \|(I-\tilde{Q})A\tilde{Q}\|. \tag{1}$$

Here $\sigma(\star)$ is the spectrum of an operator $\star$. In particular, this shows, that there is no *spectral pollution* [4] if $\|(I-\tilde{Q})A\tilde{Q}\|$ is small. Examples of polluting and non-polluting approximations of the continuous spectrum of operators related to the MHD equations can be found in [4].

The results of the paper were presented at the Fifth SIAM Conference on Applied Linear Algebra, June, 1994, in Snowbird, Utah.

2. The main theorem. Let P and $\tilde{P}$ be orthoprojectors on eigenspaces of the operators A and $\tilde{A}$, corresponding to the eigenvalues λ and $\tilde{\lambda}$:

$$AP = \lambda P, \ \tilde{A}\tilde{P} = \tilde{\lambda}\tilde{P}. \tag{2}$$

It is not required that P and $\tilde{P}$ be orthoprojectors on complete eigenspaces; and there are no restrictions on the dimensions of their images.

There are several important equalities that stem from the definition (2)

$$AP = PA, \ \tilde{A}\tilde{P} = \tilde{P}\tilde{A}, \ \tilde{Q}\tilde{P} = \tilde{P}\tilde{Q} = \tilde{P}, \ \tilde{Q}(A-\tilde{\lambda})\tilde{P} = 0. \tag{3}$$

If, in particular, P and $\tilde{P}$ are orthoprojectors on the one-dimensional subspaces span$\{u\}$ and span$\{\tilde{u}\}$, respectively, then the following equalities hold

$$\|P\tilde{P}\| = \|P\tilde{u}\| = |(\tilde{u}, u)| = \|\tilde{P}u\| = \|\tilde{P}P\|, \text{ for } \|u\| = \|\tilde{u}\| = 1,$$

and the estimate of the basic Theorem 1.1 becomes a particular case of an estimate of the next Theorem.

Theorem 2.1. *If* $\lambda \neq \tilde{\lambda}$, *then*

$$\|\tilde{P}P\| = \|P\tilde{P}\| \leq \frac{\|(A-\tilde{\lambda})\tilde{P}\|}{|\tilde{\lambda}-\lambda|}\|(I-\tilde{Q})P\|. \tag{4}$$

Proof. [First] We have

$$|\lambda-\tilde{\lambda}|\|P\tilde{P}\| = \|(A-\tilde{\lambda})P\tilde{P}\| = \|P(A-\tilde{\lambda})\tilde{P}\| =$$

$$\|P(I-\tilde{Q})(A-\tilde{\lambda})\tilde{P}\| \leq \|P(I-\tilde{Q})\|\|(A-\tilde{\lambda})\tilde{P}\| =$$

$$\|(I-\tilde{Q})P\|\|(A-\tilde{\lambda})\tilde{P}\|.$$

Here the first three equalities are based on (2) and (3). The final equality is a particular case of a general equality,

$$\|FG\| = \|(FG)^*\| = \|GF\|$$

for linear bounded selfadjoint operators F and G. (Below such equalities will be used without special references.) $\square$

Proof. [Second] By exchanging of P and $\tilde{P}$ in the previous arguments, we obtain

$$|\tilde{\lambda} - \lambda| \|\tilde{P}P\| = \|\tilde{P}(A - \lambda)\tilde{P}P\| =$$

$$\|\tilde{P}(A - \lambda)\tilde{Q}P\| = \|\tilde{P}A(I - \tilde{Q})P\| \leq$$

$$\|\tilde{P}A(I - \tilde{Q})\| \|(I - \tilde{Q})P\| = \|(I - \tilde{Q})A\tilde{P}\| \|(I - \tilde{Q})P\| =$$

$$\|(A - \tilde{\lambda})\tilde{P}\| \|(I - \tilde{Q})P\|,$$

which gives the second proof of the theorem. $\square$

REMARK 2.1. *It is useful to note, that*

$$(I - \tilde{P})A\tilde{P} = (A - \tilde{\lambda})\tilde{P} = (I - \tilde{Q})(A - \tilde{\lambda})\tilde{P} = (I - \tilde{Q})A\tilde{P} \tag{5}$$

due to $(I - \tilde{Q})\tilde{P} = 0$ *and (3).*

3. Accuracy estimates for eigenspaces. Let R and $\tilde{R}$ be orthoprojectors on invariant subspaces of the operators A and $\tilde{A}$, such that $\mathrm{Im}\tilde{R} \subseteq \mathrm{Im}\tilde{Q}$. Then

$$AR = RAR = RA, \ \tilde{A}\tilde{R} = \tilde{R}A\tilde{R} = \tilde{R}\tilde{A}, \ \tilde{Q}\tilde{R} = \tilde{R}\tilde{Q} = \tilde{R}. \tag{6}$$

Replacing R by P in Theorem 2.1 and using the first proof, we obtain the following more general statement.

THEOREM 3.1. *If*

$$d = \inf_{\nu \in \sigma(\{RAR\}|_{\mathrm{Im}R})} |\nu - \tilde{\lambda}| > 0,$$

then

$$\|R\tilde{P}\| \leq \frac{\|(I - \tilde{Q})A\tilde{P}\|}{d} \|(I - \tilde{Q})R\|.$$

Proof. The spectrum $\sigma(\{RAR\}|_{\mathrm{Im}R})$ does not contain $\tilde{\lambda}$, therefore the operator $\{R(A - \tilde{\lambda})R\}|_{\mathrm{Im}R}$ has a bounded inverse and

$$d\|R\tilde{P}\| \leq \|(A - \tilde{\lambda})R\tilde{P}\| = \|R(A - \tilde{\lambda})\tilde{P}\| =$$

$$\|R(I - \tilde{Q})(A - \tilde{\lambda})\tilde{P}\| \leq \|R(I - \tilde{Q})\| \|(I - \tilde{Q})(A - \tilde{\lambda})\tilde{P}\| =$$

$$\|(I - \tilde{Q})R\| \|(A - \tilde{\lambda})\tilde{P}\|$$

by (6). $\square$

Let us now consider the particular case $R = I - P$, in view of the fact $\|(I-\tilde{Q})R\| \leq 1$.

Corollary 3.1. *If*

$$d = \inf_{\nu \in \sigma(\{RAR\}|_{\mathrm{Im}R})} |\nu - \tilde{\lambda}| > 0, \; R = I - P,$$

then

$$\|(I - P)\tilde{P}\| \leq \frac{\|(I - \tilde{Q})A\tilde{P}\|}{d}.$$

Remark 3.1. *If the eigenvalue λ is simple, and P and $\tilde{P}$ are orthoprojectors on the one-dimensional subspaces* $\mathrm{span}\{u\}$ *and* $\mathrm{span}\{\tilde{u}\}$*, respectively, then the estimate of Corollary 3.1 converts into a wellknown estimate due to Kato [5]*

$$\sin \angle\{u; \tilde{u}\} \leq \frac{\|A\tilde{u} - \tilde{\lambda}\tilde{u}\|}{\inf_{\nu \in \sigma(A)\setminus\lambda} |\nu - \tilde{\lambda}|}, \; \|\tilde{u}\| = 1.$$

By replacing $\tilde{R}$ by $\tilde{P}$ in the second proof of Theorem 2.1 we get

Theorem 3.2. *If*

$$\tilde{d} = \inf_{\tilde{\nu} \in \sigma(\{\tilde{R}A\tilde{R}\}|_{\mathrm{Im}\tilde{R}})} |\tilde{\nu} - \lambda| > 0,$$

then

$$\|\tilde{R}P\| \leq \frac{\|(I - \tilde{Q})A\tilde{R}\|}{\tilde{d}} \|(I - \tilde{Q})P\|.$$

Proof. The spectrum $\sigma(\{\tilde{R}A\tilde{R}\}|_{\mathrm{Im}\tilde{R}})$ does not contain λ, therefore the operator $\{R(A - \tilde{\lambda})R\}|_{\mathrm{Im}R}$ has a bounded inverse and

$$\tilde{d}\|\tilde{R}P\| \leq \|\tilde{R}(A - \lambda)\tilde{R}P\| =$$

$$\|\tilde{R}(A - \lambda)\tilde{Q}P\| = \|\tilde{R}A(I - \tilde{Q})P\| \leq$$

$$\|\tilde{R}A(I - \tilde{Q})\| \|(I - \tilde{Q})P\| = \|(I - \tilde{Q})A\tilde{R}\| \|(I - \tilde{Q})P\|$$

by (6). $\square$

Remark 3.2. *It is clear that*

$$\|(I - \tilde{Q})A\tilde{P}\| \leq \|(I - \tilde{Q})A\tilde{Q}\|, \; \|(I - \tilde{Q})A\tilde{R}\| \leq \|(I - \tilde{Q})A\tilde{Q}\| \tag{7}$$

because of

$$\mathrm{Im}\tilde{P} \subseteq \mathrm{Im}\tilde{Q}, \; \mathrm{Im}\tilde{R} \subseteq \mathrm{Im}\tilde{Q}.$$

In the particular case $\tilde{R} = \tilde{Q} - \tilde{P}$, taking into account of the previous remark, we can make the following conclusion from Theorem 3.2.

COROLLARY 3.2. *If*

$$\tilde{d} = \inf_{\tilde{\nu} \in \sigma(\{\tilde{R}A\tilde{R}\}|_{\mathrm{Im}\tilde{R}})} |\tilde{\nu} - \lambda| > 0, \ \ \tilde{R} = \tilde{Q} - \tilde{P},$$

then

$$\|(\tilde{Q} - \tilde{P})P\| \leq \frac{\|(I - \tilde{Q})A\tilde{Q}\|}{\tilde{d}} \|(I - \tilde{Q})P\|.$$

Finally, using the inequality

$$\|(I - \tilde{P})P\|^2 \leq \|(\tilde{Q} - \tilde{P})P\|^2 + \|(I - \tilde{Q})P\|^2,$$

that follows from the equalities

$$I - \tilde{P} = (\tilde{Q} - \tilde{P}) + (I - \tilde{Q}), \ (\tilde{Q} - \tilde{P})(I - \tilde{Q}) = 0,$$

we obtain the theorem, that was proved by Saad [12] for the case of a simple eigenvalue λ of a matrix A.

THEOREM 3.3. *If*

$$\tilde{d} = \inf_{\tilde{\nu} \in \sigma(\{\tilde{R}A\tilde{R}\}|_{\mathrm{Im}\tilde{R}})} |\tilde{\nu} - \lambda| > 0, \ \ \tilde{R} = \tilde{Q} - \tilde{P},$$

then

$$\|(I - \tilde{P})P\|^2 \leq \left[1 + \frac{\|(I - \tilde{Q})A\tilde{Q}\|^2}{\tilde{d}^2}\right] \|(I - \tilde{Q})P\|^2.$$

REMARK 3.3. *It was shown in [9, 7] that this estimate is stronger than the classical one of Vainikko [10], even though it looks much simpler and uses less information. Also, for fixed positive numbers d and r examples of operators A and projectors $\tilde{P}$ were constructed in [9, 7], such that $\|(I - \tilde{Q})A\tilde{Q}\| = r$ and the inequality of Theorem 3.3 becomes an equality. Therefore the estimate of Theorem 3.3 cannot be improved without new information.*

4. Accuracy estimates for invariant subspaces. We now redefine P and $\tilde{P}$ as orthoprojectors on invariant subspaces of the operators A and $\tilde{A}$, corresponding to the spectrum of A in the interval $[\lambda - \delta, \lambda + \delta]$ and of $\tilde{A}$ in the interval $[\tilde{\lambda} - \tilde{\delta}, \tilde{\lambda} + \tilde{\delta}]$

$$AP = PA, \ \|P(A - \lambda)P\| \leq \delta, \ \tilde{A}\tilde{P} = \tilde{P}\tilde{A}, \ \|\tilde{P}(\tilde{A} - \tilde{\lambda})\tilde{P}\| \leq \tilde{\delta}.$$

Then we do not require, that these subspaces incorporate all and/or the complete eigenspaces of the spectrum in the intervals. There are no restrictions on the dimensions of their images.

In this section, we generalize all the statements of the previous section to the case of invariant subspaces instead of eigenspaces, using the new definitions of P and $\tilde{P}$ given above.

THEOREM 4.1. *If*

$$d = \inf_{\nu \in \sigma(\{RAR\}|_{\mathrm{Im} R})} |\nu - \tilde{\lambda}| > \tilde{\delta},$$

then

$$\|R\tilde{P}\| \le \frac{\|(I - \tilde{Q})A\tilde{P}\|}{d - \tilde{\delta}} \|(I - \tilde{Q})R\|.$$

Proof. We have

$$d\|R\tilde{P}\| \le \|(A - \tilde{\lambda})R\tilde{P}\| = \|R(A - \tilde{\lambda})\tilde{P}\| \le$$

$$\|R(I - \tilde{Q})(A - \tilde{\lambda})\tilde{P}\| + \|R\tilde{Q}(A - \tilde{\lambda})\tilde{P}\| \le$$

$$\|R(I - \tilde{Q})\| \|(I - \tilde{Q})(A - \tilde{\lambda})\tilde{P}\| + \|R\tilde{P}\| \|\tilde{P}(A - \tilde{\lambda})\tilde{P}\| =$$

$$\|(I - \tilde{Q})R\| \|(I - \tilde{Q})A\tilde{P}\| + \tilde{\delta}\|R\tilde{P}\|$$

The theorem is proved. □

We can, considering the particular case of $R = I - P$ and using the trivial inequality $\|(I - \tilde{Q})R\| \le 1$, in analogy with an argument of the previous section, give the statement, which was proved by Davis and Kahan, see [2] and cf. also Theorem 11.7.11 of [11].

COROLLARY 4.1. *If*

$$d = \inf_{\nu \in \sigma(\{RAR\}|_{\mathrm{Im} R})} |\nu - \tilde{\lambda}| > \tilde{\delta}, \ R = I - P,$$

then

$$\|(I - P)\tilde{P}\| \le \frac{\|(I - \tilde{Q})A\tilde{P}\|}{d - \tilde{\delta}}.$$

We now prove an analog of Theorem 3.2 for invariant subspaces.

THEOREM 4.2. *If*

$$\tilde{d} = \inf_{\tilde{\nu} \in \sigma(\{\tilde{R}A\tilde{R}\}|_{\mathrm{Im} \tilde{R}})} |\tilde{\nu} - \lambda| > \delta,$$

then

$$\|\tilde{R}P\| \le \frac{\|(I - \tilde{Q})A\tilde{R}\|}{\tilde{d} - \delta} \|(I - \tilde{Q})P\|.$$

Proof.

$$\tilde{d}\|\tilde{R}P\| \le \|\tilde{R}(A - \lambda)\tilde{R}P\| =$$

$$\|\tilde{R}(A-\lambda)\tilde{Q}P\| \le \|\tilde{R}A(I-\tilde{Q})P\| + \|\tilde{R}(A-\lambda)P\| \le$$

$$\|\tilde{R}A(I-\tilde{Q})\|\|(I-\tilde{Q})P\| + \|P(A-\lambda)P\|\|\tilde{R}P\|$$

$$\le \|(I-\tilde{Q})A\tilde{R}\|\|(I-\tilde{Q})P\| + \delta\|\tilde{R}P\|.$$

The theorem is proved. □

Choosing $\tilde{R} = \tilde{Q} - \tilde{P}$ in Theorem 4.2 and taking into account Remark 3.2 we can make the following conclusion.

COROLLARY 4.2. *If*

$$\tilde{d} = \inf_{\tilde{\nu}\in\sigma(\{\tilde{R}A\tilde{R}\}|_{\mathrm{Im}\tilde{R}})} |\tilde{\nu}-\lambda| > \delta, \;\; \tilde{R} = \tilde{Q} - \tilde{P},$$

then

$$\|(\tilde{Q}-\tilde{P})P\| \le \frac{\|(I-\tilde{Q})A\tilde{Q}\|}{\tilde{d}-\delta}\|(I-\tilde{Q})P\|.$$

Now, using the same arguments as in the previous section, we prove the following generalization of the Saad's Theorem 3.3 for the case of invariant subspaces instead of eigenspaces.

THEOREM 4.3. *If*

$$\tilde{d} = \inf_{\tilde{\nu}\in\sigma(\{\tilde{R}A\tilde{R}\}|_{\mathrm{Im}\tilde{R}})} |\tilde{\nu}-\lambda| > \delta, \;\; \tilde{R} = \tilde{Q} - \tilde{P},$$

then

$$\|(I-\tilde{P})P\|^2 \le \left[1 + \frac{\|(I-\tilde{Q})A\tilde{Q}\|^2}{(\tilde{d}-\delta)^2}\right]\|(I-\tilde{Q})P\|^2.$$

REMARK 4.1. *It is wellknown, see [7], that*

- $\|(I-P)\tilde{P}\| \le 1$ *measures the proximity of* $\mathrm{Im}\tilde{P}$ *to* $\mathrm{Im}P$;
- $\|\tilde{P}-P\| = \max\{\|(I-P)\tilde{P}\|, \|(I-\tilde{P})P\|\} \le 1$ *measures the proximity (the gap) between* $\mathrm{Im}P$ *and* $\mathrm{Im}\tilde{P}$;
- $\|(I-\tilde{P})P\| \le 1$ *measures the proximity of* $\mathrm{Im}P$ *to* $\mathrm{Im}\tilde{P}$.

The proximity, for example, of $\mathrm{Im}\tilde{P}$ *to* $\mathrm{Im}P$ *means by Theorem 6.34 of [6], p. 56, the proximity between* $\mathrm{Im}\tilde{P}$ *and a subspace of* $\mathrm{Im}P$, *i.e. if* $\|(I-P)\tilde{P}\| < 1$, *then there is a subspace in* $\mathrm{Im}P$ *with an associated orthogonal projector* P' *such that* $\|(I-P)\tilde{P}\| = \|\tilde{P}-P'\| < 1$.

If

$$\dim \mathrm{Im}P = \dim \mathrm{Im}\tilde{P} < \infty,$$

then

$$\|(I-\tilde{P})P\| = \|\tilde{P}-P\| = \|(I-P)\tilde{P}\|, \tag{8}$$

as follows from Theorem 6.34 of [6]. Therefore, Corollaries 3.1 and 4.1 estimate the proximity of $\mathrm{Im}\tilde{P}$ *to* $\mathrm{Im}P$, *while Theorems 3.3 and 4.3 estimate the proximity of* $\mathrm{Im}P$

to $\mathrm{Im}\tilde{P}$*. In the case* $\dim \mathrm{Im} P = \dim \mathrm{Im} \tilde{P} < \infty$ *all of them estimate the gap between* $\mathrm{Im} P$ *and* $\mathrm{Im}\tilde{P}$*.*

REMARK 4.2. *By choosing subspaces properly, and by using statement (1), we can derive a priori estimates of* d *and* $\tilde{d}$ *in the denominators of estimates.*

Assume that the points $\lambda - \delta$ *and* $\lambda + \delta$ *do not belong to the spectrum of* A*. Let* $E(\nu)$ *be a spectral family associated with* A*, i. e.*

$$A = \int_{-\infty}^{+\infty} \nu \, dE(\nu).$$

Then it is possible to define the projector P *as*

$$P = \int_{\lambda-\delta}^{\lambda+\delta} dE(\nu) = E(\lambda + \delta) - E(\lambda - \delta).$$

With such a definition, P *is the orthogonal projector on the invariant subspace of the operator* A*, which corresponds to the spectrum of* A *in the interval* $[\lambda - \delta, \lambda + \delta]$*; this subspace incorporates all and complete eigenspaces of* A *with the spectrum in the interval. Further, let for a number* $\Delta > 0$ *the closed intervals* $[\lambda - \delta - \Delta, \lambda - \delta]$ *and* $[\lambda + \delta, \lambda + \delta + \Delta]$ *contain no points of the spectrum of* A*, i. e. there is at least a* Δ *gap in the spectrum of* A *around the interval* $[\lambda - \delta, \lambda + \delta]$*. Let*

$$r \stackrel{\text{def}}{=} \|(I - \tilde{Q})A\tilde{Q}\| < \frac{\Delta}{2}.$$

Then, by statement (1), $\lambda - \delta - r$ *and* $\lambda + \delta + r$ *do not belong to the spectrum of the operator* $\tilde{A}$*, while the interval* $[\lambda - \delta - r, \lambda + \delta + r]$ *may contain points of the spectrum of the operator* $\tilde{A}$*. Let* $\tilde{E}(\nu)$ *be a spectral family associated with* $\tilde{A}$*, and define the projector* $\tilde{P}$ *as*

$$\tilde{P} = \tilde{E}(\lambda + \delta + r) - \tilde{E}(\lambda - \delta - r).$$

This makes $\tilde{P}$ *the orthogonal projector on the invariant subspace of the operator* $\tilde{A}$*, corresponding to the spectrum of* $\tilde{A}$ *in the interval* $[\lambda - \delta - r, \lambda + \delta + r]$*; this subspace incorporates all and complete eigenspaces of* $\tilde{A}$ *with the spectrum in the interval. To comply with the previous definition of* $\tilde{P}$*, we set*

$$\tilde{\lambda} = \lambda, \ \tilde{\delta} = \delta + r.$$

Now it is clear, that in Corollary 4.1 $d = \lambda + \delta + \Delta - \tilde{\lambda} = \delta + \Delta$*, and the estimate of Corollary 4.1 takes the form*

$$\|(I - P)\tilde{P}\| \leq \frac{\|(I - \tilde{Q})A\tilde{P}\|}{\Delta - r} \leq \frac{r}{\Delta - r} < 1 \tag{9}$$

because of the condition $\Delta > 2r$ *introduced above.*

To estimate $\tilde{d}$ *of Theorem 4.3, we have to use statement (1) again to conclude that inside the interval* $[\lambda - \delta - \Delta + r, \lambda + \delta + \Delta - r]$ *there are no points of spectrum of the operator* $\tilde{A}$ *except these of the spectrum of the operator* $\{\tilde{P}\tilde{A}\tilde{P}\}|_{\mathrm{Im}\tilde{P}}$ *. But*

$$\sigma(\tilde{A}) = \sigma\left(\{\tilde{P}\tilde{A}\tilde{P}\}|_{\mathrm{Im}\tilde{P}}\right) \bigcup \sigma\left(\{\tilde{R}\tilde{A}\tilde{R}\}|_{\mathrm{Im}\tilde{R}}\right), \ \tilde{R} = \tilde{Q} - \tilde{P}.$$

Therefore, $\tilde{d} = \lambda + \delta + \Delta - r - \lambda = \delta + \Delta - r$, *and the estimate of Theorem 4.3 converts into*

$$\|(I - \tilde{P})P\|^2 \leq \left[1 + \frac{r^2}{(\Delta - r)^2}\right] \|(I - \tilde{Q})P\|^2. \tag{10}$$

Taking the previous remark into account, we can draw the conclusion that small r *affords the proximity of the Ritz vectors to an eigenspace of* A *by (9), but not necessarily ensures a good approximation of the complete eigenspace, in contrast to (10), except for the case when the dimension of the eigenspace is known a priory (and is finite). We can then use (8) if the dimension of the approximation subspace, spanned by Ritz vectors, happens to be the same.*

REMARK 4.3. *There is also a sharp accuracy estimate [15, 14] for a finite dimensional invariant subspace corresponding a group of leading eigenvalues just in terms of the accuracy approximation of the eigenvalues. The simplified proof for the particular case of an eigenspace instead of an invariant subspace can be found in [8].*

Acknowledgments. The author thanks Olof Widlund for his help. Without his support this paper would have been never written.

REFERENCES

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

[2] C. Davis and W. M. Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.*, 7(1):1–46, 1970.

[3] J. Descloux, N. Nassif, and J. Rappaz. On spectral approximation. Part 1. The problem of convergence. *RAIRO, Numerical Analysis*, 12(2):97–112, 1978.

[4] R. Gruber and J. Rappaz. *Finite Element Methods in Linear Ideal Magnetohydrodynamics.* Springer Series in Computational Physics. Springer-Verlag, New York, 1985.

[5] T. Kato. On the upper and lower bounds on eigenvalues. *J. Phys. Soc. Japan*, 4:334–339, 1949.

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

[7] 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.

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

[9] 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.

[10] 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.

[11] B. N. Parlett. *The Symmetric Eigenvalue problem.* Prentice-Hall, 1980.

[12] Y. Saad. *Numerical Methods for Large Eigenvalue Problems.* Halsted Press, New York, 1992.

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

[14] H. F. Weinberger. Error bounds in the Rayleigh-Ritz approximation of eigenvectors. *J. Res. Nat. Bur. Standards*, 64 B(4):217–225, 1960.

[15] H. F. Weinberger. *Variational Methods for Eigenvalue Approximation.* SIAM, 1974.