A SUBSPACE PRECONDITIONING ALGORITHM FOR EIGENVECTOR/EIGENVALUE COMPUTATION

James H. Bramble, Andrew V. Knyazev, and Joseph E. Pasciak

Advances in Computational Mathematics, 6 (1996), no. 2, 159–189.

We consider the problem of computing a modest number of the smallest eigenvalues along with orthogonal bases for the corresponding eigenspaces of a symmetric positive definite operator A defined on a finite dimensional real Hilbert space V. In our applications, the dimension of V is large and the cost of inverting A is prohibitive. In this paper, we shall develop an effective parallelizable technique for computing these eigenvalues and eigenvectors utilizing subspace iteration and preconditioning for A. Estimates will be provided which show that the preconditioned method converges linearly when used with a uniform preconditioner under the assumption that the approximating subspace is close enough to the span of desired eigenvectors.

1. Introduction.

In this paper, we shall be concerned with computing a modest number of the smallest eigenvalues and their corresponding eigenvectors of a large symmetric ill-conditioned system. More explicitly, let A be a symmetric and positive definite linear operator on a N-dimensional real vector space V with inner product $(\cdot,\cdot)$ and norm $\|\cdot\| = (\cdot,\cdot)^{1/2}$. The distinct eigenvalues, $\{\lambda_i\}$, of A are positive real numbers which along with their respective eigenvectors $\{v_i\}$, satisfy

$$Av_i = \lambda_i v_i.$$

The eigenvalues $\{\lambda_i\}$ are ordered to be increasing with i and are assumed to be simple. The eigenvectors $\{v_i\}$ are orthogonal with respect to the inner product $(\cdot,\cdot)$ and chosen so that (Av_i, v_i) =1. The algorithms and results of this paper carry over to the case of Hermitian operators on complex inner product spaces without change.

We consider the case where the first s eigenvalues are simple and well separated. It seems possible to extend the analysis to the case of eigenvalues of higher

1991 *Mathematics Subject Classification*. Primary 65N30; Secondary 65F10.

This manuscript has been authored under contract number DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. This work was also supported in part under the National Science Foundation Grants No. DMS-9007185, DMS-9501507 and NSF-CCR-9204255, and by the U.S. Army Research Office through the Mathematical Sciences Institute, Cornell University.

multiplicity. The extensions of the results to the case of eigenvalues with little separation is somewhat more tedious. If the operator A is a mesh analog of a PDE with multiple eigenvalues, then A has clusters of eigenvalues and this is one of the most interesting practical examples of a bad separation of eigenvalues. However, for this case the operator A can be viewed as a perturbation of an operator with well separated multiple eigenvalues, see [24]. The analysis of such perturbation appears possible for the method described below, but will not be addressed in this paper.

We shall be interested in iterative schemes for computing $\{\lambda_i\}$ and $\{v_i\}$, for $i = 1, \ldots, s$, where s is small compared to N. We will only assume that a procedure for evaluating the action of A applied to vectors in V is available. Given a basis for V the corresponding matrix representing A is often large and sparse and a full computer representation, including the zero elements, is not feasible since its size would be too large to manage. We will also avoid the computation of the action of A^{-1}.

There are many methods for obtaining eigenvalues and eigenvectors for symmetric positive definite matrices (see, [44]). Methods like the QR–algorithm work extremely well for relatively small matrices. Classical iterative methods involve subspaces of vectors which result from applying $A - \lambda I$ or its inverse with a non-negative parameter λ which may change from iteration to iteration. Because A is ill-conditioned and we seek the eigenspaces corresponding to the smaller eigenvalues, to be effective, the classical methods require inversions of $A - \lambda I$.

In this paper, we shall study an iterative eigenvalue scheme which utilizes subspace iteration and preconditioning. A preconditioner B is a symmetric positive definite operator on V which "approximates" the inverse of A. For our purposes, we shall assume that B is scaled so that the operator $I - BA$ is a reducer. This means that there is a number γ in (0,1) satisfying

$$|(A(I - BA)v, v)| \leq \gamma(Av, v) \quad \text{for all } v \in V. \tag{1.1}$$

Note that (1.1) implies that

$$(1 - \gamma)(Av, v) \leq (BAv, Av) \leq (1 + \gamma)(Av, v) \quad \text{for all } v \in V. \tag{1.2}$$

There is a vast literature of techniques for developing preconditioners for symmetric positive definite problems, especially in the case when the operator A is a discretization of an elliptic partial differential equation (see, e.g., [1], [2], [15], [16], [27], [28]). The best preconditioners satisfy (1.1) with γ bounded away from one (independently of N). In addition, a good preconditioner is economical to evaluate. This means that the cost of computing the action of B applied to an arbitrary vector should not be much greater than that of applying A. When A corresponds to a discretization of a partial differential equation, often low cost preconditioners are known for which (1.1) holds with γ independent of the mesh size and hence the number of unknowns (see, e.g., [3]–[6], [12], [21], [52]). Multigrid and domain decomposition are two examples of effective techniques for developing preconditioners for the discrete systems arising from approximations to elliptic boundary value problems (see, e.g., [4]–[13]).

The use of preconditioned iterations for computing the eigenvectors and eigenvalues has been first studied by L. V. Kantorovitch's graduate student B. A. Samokish

[50], and then by W. V. Petryshyn [45]. Axel Ruhe [46], [47] clarified the asymptotic connection of the such methods and similar preconditioned iterative methods for finding a nontrivial solution of the corresponding Helmholtz system.

Explicit convergence rate estimates, independent of the number of unknowns, for preconditioned iterative methods in the case of one eigenvector were first obtained in the Russian literature, by S. K. Godunov *et. al.* [29] and by E. G. D'yakonov *et. al.* [25], [26] (see also [22] and the included references). In particular, the base iteration used in our preconditioned subspace iteration was used in [29] and was further analyzed in [25].

There are a number of alternative preconditioned schemes which have been proposed to further improve convergence of the base method. The possibility of using Chebyshev parameters to accelerate the convergence of two stage preconditioned eigenvalue iterations was discovered by V. P. Il'in [30]. An analogous idea of using Lanczos method in the inner stage is due to D. Scott *et. al.* [51], [42]. Explicit convergence rate estimates, independent of the number of unknowns, for these methods have been established in [33],[32]. The convergence estimates for the two stage method are better than those for the base method when high accuracy is required. The locally optimal preconditioned conjugate gradient method was suggested in [35]. In [35], [36], a new preconditioned variant especially suited for the domain decomposition approach was presented and the corresponding convergence rate estimates were proved.

Davidson proposed a diagonally preconditioned subspace method [18] where the dimension of the subspace increased each step of the iteration. Although the original Davidson method is likely to converge fast with a good preconditioner, the cost per iteration and storage requirements grow with the iteration number. This makes the method unacceptable in many large scale applications. Nevertheless, the method became popular in computational molecular physics and was further developed (see, e.g., [19], [37], [41], [40], [43], [49], [20]). Theoretically Davidson-like methods are still not well studied, for example, it is not proved that their convergence does not depend on the number of unknowns with a proper preconditioner, though it seems to be the case.

Generalizations of the preconditioned methods for the simultaneous computation of several leading eigenvalues and the corresponding invariant subspaces by using subspace iterations were suggested in [50], [39], [38], [14]. The first and, in fact, the only explicit estimates on the convergence behavior, independent of the number of unknowns, were obtained by Dyakonov and Knyazev (DK) in [23], [24], [22] where simplified methods with the same iteration operator for all vectors in a subspace were developed and analyzed. This simplification, however, leads to a method which computes only one eigenvalue, the largest in the group, at a time. To find another, smaller, eigenvalue, the method has to be used again with an initial subspace of a smaller dimension and with an orthogonalization to the previously computed eigenvector. For the DK method, the convergence estimates do not depend on the separation of the eigenvalues.

In the present paper we analyze a preconditioned subspace iteration technique. This involves a recursively generated sequence of subspaces. Given a subspace in this sequence, we compute the approximate eigenvectors by applying the Rayleigh-Ritz procedure. The next subspace in the sequence is defined as the linear space spanned by the vectors which result from the application of a simple preconditioned eigenvector iteration procedure to the Ritz eigenvectors. In contrast to the DK method, our iteration operator is different for different Ritz eigenvectors. On the other hand, our method differs from that of [39], [38] by using a fixed-shift basic algorithm, whereas [39], [38] suggest a block version of the steepest descent method [50]. Our method could also be thought of as a truncated version of a block Davidson method.

We present a theorem which guarantees convergence of the preconditioned subspace method provided that the starting subspace is sufficiently accurate. As in the classical block inverse power method, the convergence rate estimate for the smaller eigenvalues and their corresponding eigenvectors is better than for eigenpairs whose eigenvalues are closer to λ_{s+1}. The rates only depend on the first $s+1$ eigenvalues, δ and is independent of the largest eigenvalue λ_m and/or the number of unknowns. This is crucial for our applications where we seek the eigenvalues of a discrete second order elliptic operator. The largest eigenvalue of such an operator and the number of unknowns tend to infinity like h^{-2} where h is the mesh parameter. The only disadvantage of our theory is that the accuracy condition on the initial subspace depends on the gap in first $s+1$ eigenvalues, in contrast to the theory of the classical block inverse power method and of the simplified preconditioned subspace DK method. Numerical experiments suggest that the actual convergence of our method is, in fact, independent of the gap and that its overall performance is much better than the DK method. We have chosen the DK method for numerical comparison as it was, before the present paper, the only preconditioned subspace iteration method with proven explicit convergence rate estimates, independent of the number of unknowns.

The form of the algorithm proposed was motivated by a need for developing a parallelizable eigenvalue/eigenvector algorithm which would be effective in computing a number of the smallest eigenvalues and their corresponding eigenvectors for large symmetric systems. The scheme is currently being applied to the computation of several hundred eigenfunctions for a problem with several thousand unknowns which arises in first principle electronic structure computations [17].

The outline of the remainder of the paper is as follows. We describe the algorithm in Section 2 and give a theorem which bounds its convergence. Section 3 contains several useful estimates for the Rayleigh-Ritz method, partially based on results of [48], [33], [34]. The proof of the theorem is given in Section 4. Finally, the results of numerical experiments involving the preconditioned subspace technique are given in Section 5.

2. The subspace preconditioning algorithm.

In this section, we describe the subspace preconditioning algorithm which will be studied in this paper. This algorithm involves the development of a sequence of subspaces $V_s^n \subset V$, for $n = 1, 2, \ldots$, approximating the eigenspace

$$V_s = \operatorname{span}\{v_1, \ldots, v_s\}.$$

The initial approximation subspace V_s^0 of dimension s is assumed given.

Given an approximation subspace V_s^n of dimension s, the eigenvectors and eigenvalues of A are approximated by computing the Ritz eigenvectors $\{v_i^n\} \subset V_s^n$ along with their corresponding eigenvalues λ_i^n satisfying

$$(Av_i^n, w) = \lambda_i^n (v_i^n, w) \quad \text{for all } w \in V_s^n. \tag{2.1}$$

Here $\lambda_1^n \leq \lambda_2^n \leq \cdots \leq \lambda_s^n$ and $\{v_i^n\}$ are mutually orthogonal and normalized so that $(Av_i^n, v_i^n) = 1$, for $i = 1, \ldots, s$.

The new subspace V_s^{n+1} is generated from a basis which is defined by applying a simple preconditioned eigenvalue iteration scheme to the Ritz vectors of the previous subspace V_s^n. Specifically, let

$$\hat{v}_i^{n+1} = v_i^n - B(Av_i^n - \lambda_i^n v_i^n) \text{ for } i = 1, \ldots, s, \tag{2.2}$$

and define

$$V_s^{n+1} = \text{span}\{\hat{v}_1^{n+1}, \ldots, \hat{v}_s^{n+1}\}.$$

The iteration (2.2) has been studied as a stand alone iterative scheme for computing the smallest eigenvalue λ_1 and a vector in the corresponding eigenspace [25].

The above method could also be thought of as the simplest truncated version of a block Davidson method which defines

$$V_s^{n+1} = V_s^n + \text{span}\{\hat{v}_1^{n+1}, \ldots, \hat{v}_s^{n+1}\}.$$

We shall need some additional notation to state our convergence estimates. Let $(\cdot, \cdot)_A$ denote the inner product defined by

$$(v, w)_A = (Av, w) \quad \text{for all } v, w \in V.$$

The corresponding norm will be denoted by

$$||v||_A \equiv (v, v)_A^{1/2} \quad \text{for all } v \in V.$$

Let Q_s and Q_s^n be the $(\cdot, \cdot)_A$ orthogonal projection operators onto V_s and V_s^n, respectively. We measure the gap between the subspaces V_s and V_s^n by considering the norm of the difference of these projectors:

$$\theta^n \equiv ||Q_s^n - Q_s||_A. \tag{2.3}$$

We also define

$$\theta_i^n \equiv ||v_i - Q_s^n v_i||_A, \text{ for } i = 1, \ldots, s,$$

and

$$\Delta \equiv \max_{i=1,\ldots,s} \frac{\lambda_{i+1} + \lambda_i}{\lambda_{i+1} - \lambda_i}.$$

The main result of this paper is given in the following theorem.

Theorem 2.1. *Suppose that the initial approximation subspace V_s^0 satisfies*

$$\sum_{i=1}^{s} (\theta_i^0)^2 \leq \frac{(1-\gamma)^2 (\lambda_1)^4 (1 - \lambda_s/\lambda_{s+1})^4}{1999 \Delta^2 (\lambda_s)^4} \tag{2.4}$$

where γ satisfies (1.1). Then the dimension of V_s^n for $n > 0$ is equal to s. Moreover, we have

$$||(I - P_i)v_i^n||_A^2 \leq \frac{1.03}{1 - \lambda_i/\lambda_{s+1}} \bar{\delta}_i^{2n} (\theta_i^0)^2 \tag{2.5}$$

and

$$0 \leq 1 - \lambda_i/\lambda_i^n \leq \frac{1.03}{1 - \lambda_i/\lambda_{s+1}} \bar{\delta}_i^{2n} (\theta_i^0)^2 \tag{2.6}$$

where

$$\delta_i = \gamma + (1-\gamma)\lambda_i/\lambda_{s+1} < 1, \tag{2.7}$$

$$\bar{\delta}_i = \delta_i + \frac{(1-\delta_s)}{2} \left(\frac{\lambda_{s+1} - \lambda_s}{\lambda_{s+1} - \lambda_i} \right)^{1/2} \frac{\lambda_i}{\lambda_s} < 1. \tag{2.8}$$

In (2.5), P_i denotes the orthogonal projector onto the one dimensional eigenspace spanned by v_i.

Remark 2.1. Note that $\bar{\delta}_i$ is less than or equal to $\delta_i + (1-\delta_i)/2$ which is clearly less than one. In addition, the convergence rate estimate $\bar{\delta}_i$ is closer to δ_i for smaller eigenvalues since λ_i is smaller than λ_s in that case.

Remark 2.2. Note that $||v_i^n||_A = 1$ and hence (2.5) implies that the eigenvector v_i^n converges to the eigenvector v_i exponentially with n.

Remark 2.3. The above theorem can be applied to the problem of the approximation of relatively few of the lowest eigenvalues and their corresponding eigenvectors of an elliptic boundary value problem defined on a bounded domain in the case when the corresponding lower eigenvalues are simple. It is well known that these eigenvalues are distinct. Moreover, these eigenvalues and eigenvectors can be approximated by those which result from finite element discretization. The eigenvalues of the discrete system exhibit behavior similar to those of the continuous problem. They are well separated provided that the mesh parameter h is suitably small since the discrete eigenvalues converge to those of the continuous problem. If $A = A_h$ is the system corresponding to the discretization with mesh size h then the parameters λ_s, λ_1, Δ can be bounded independently of the mesh parameter h. If one uses a preconditioner $B = B_h$ which satisfies (1.1) with $\gamma < 1$ (not depending h) then the above proposition provides asymptotic rates of convergence which do not depend h and hence the dimension of the underlying system.

Remark 2.4. With a slightly more complicated proof, it is possible to reduce the constant 1999 to 500 in the above theorem. Of course, even the latter constant is still too large to convince someone theoretically that our method is useful. However, the method appears to be much more robust in practice. In our experience, the method converges with initial subspaces consisting of vectors with random entries (see Section 5). Such an initial subspace fails to satisfy the accuracy condition (2.4).

The following corollary show that the rates of convergence for the given eigenvalue/eigenvector tend asymptotically to δ_i as the V_s^n converges to V_s.

Corollary 2.1. *Assume that the hypothesis of Theorem 2.1 hold. Then for any* $\epsilon \in (0,1]$ *there exists an* $m = m(\epsilon)$ *such that for* $n \geq m$,

$$||(I - P_i)v_i^n||_A^2 \leq \frac{1.03}{1 - \lambda_i/\lambda_{s+1}} (\delta_i + \epsilon(\bar{\delta}_i - \delta_i))^{2n-2m} \bar{\delta}_i^{2m} (\theta_i^0)^2$$

and

$$0 \leq 1 - \lambda_i/\lambda_i^n \leq \frac{1.03}{1 - \lambda_i/\lambda_{s+1}} (\delta_i + \epsilon(\bar{\delta}_i - \delta_i))^{2n-2m} \bar{\delta}_i^{2m} (\theta_i^0)^2$$

where δ_i *and* $\bar{\delta}_i$ *are as in (2.7) and (2.8).*

3. Properties of the Ritz approximation

In this section, we give some lemmas which describe the approximation properties of the eigenvalues and eigenvectors resulting from the Ritz subspace method. These lemmas only depend on the distribution of the eigenvalues of A and the approximation properties of the subspace. Thus, we shall state them in terms of an arbitrary approximation subspace $\tilde{V} \subset V$ of dimension less than or equal to s.

The Ritz eigenvectors $\{\tilde{v}_i\} \subset \tilde{V}$ along with their corresponding eigenvalues $\tilde{\lambda}_i$ satisfy

$$(A\tilde{v}_i, w) = \tilde{\lambda}_i(\tilde{v}_i, w) \quad \text{for all } w \in \tilde{V}.$$

Here $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \cdots$ and the vectors in $\{\tilde{v}_i\}$ are mutually orthogonal and normalized so that $(A\tilde{v}_i, \tilde{v}_i) = 1$ for each i.

Let $\tilde{Q}$ be the $(\cdot,\cdot)_A$ orthogonal projection operator onto $\tilde{V}$ and set

$$\tilde{\theta} \equiv ||\tilde{Q} - Q_s||_A.$$

It will also be important to measure how well the actual eigenvectors can be approximated in the subspace $\tilde{V}$. To this end, we define the quantities

$$\tilde{\theta}_i \equiv ||v_i - \tilde{Q}v_i||_A, \text{ for } i = 1, \ldots, s.$$

The eigenvalue accuracy can be estimated in terms of the above parameters by the following lemma.

Lemma 3.1. *If*

$$\tilde{\theta} < 1 \tag{3.1}$$

then the dimension of $\tilde{V}$ *is* s *and for* $i = 1, \ldots, s$,

$$0 \leq 1 - \lambda_i/\tilde{\lambda}_i \leq \tilde{\theta}^2.$$

The proof of the above lemma follows [33] and uses two additional lemmas. The first is essentially given in [31] (see Theorem 6.34).

Lemma 3.2. *Let P and Q be two orthogonal projectors with $M = Range(P)$, $N = Range(Q)$ such that $dim(N) \leq dim(M)$ and*

$$\|(I - Q)P\| = \delta < 1.$$

Then $dim(N) = dim(M)$ and

$$\|P - Q\| = \|(I - P)Q\| = \delta.$$

Lemma 3.3. *Let V_i be the space spanned by the eigenvectors v_j, for $j = 1, \ldots, i$. We define the additional operators Q_i and $\tilde{Q}_i$ to be the $(\cdot,\cdot)_A$ orthogonal projectors onto the spaces V_i and $\tilde{V}_i \equiv Range(\tilde{Q}V_i)$. Assume that $\tilde{\theta} < 1$. Then for $i \leq s$, the maps*

$$\tilde{Q}_i : V_i \mapsto \tilde{V}_i,$$
$$Q_i : \tilde{V}_i \mapsto V_i$$

are isomorphisms. Moreover,

$$||\tilde{Q}_i - Q_i||_A = ||(I - \tilde{Q}_i)Q_i||_A = ||(I - Q_i)\tilde{Q}_i||_A \leq \tilde{\theta}. \tag{3.2}$$

Proof. By definition of $\tilde{Q}_i$,

$$\tilde{Q}_i v = \tilde{Q}v \quad \text{for all } v \in V_i.$$

Thus,

$$||(I - \tilde{Q}_i)Q_i||_A = \max_{v \in V_i, ||v||_A = 1} ||v - \tilde{Q}_i v||_A =$$

$$\max_{v \in V_i, ||v||_A = 1} ||v - \tilde{Q}v||_A \leq \max_{v \in V_s, ||v||_A = 1} ||v - \tilde{Q}v||_A = \tilde{\theta}.$$

This proves the last equality of (3.2). The lemma now follows from Lemma 3.2.

Proof (of Lemma 3.1). Let i be less than or equal to s and $v^* \in \tilde{V}_i$ be the eigenvector with maximal eigenvalue λ^* eigenproblem:

$$(Av, w) = \lambda(v, w) \quad \text{for all } w \in \tilde{V}_i.$$

By Lemma 3.3, the dimension of $\tilde{V}_i$ is equal to i. By the Courant–Fischer Theorem,

$$\lambda_i \leq \tilde{\lambda}_i \leq \lambda^* \tag{3.3}$$

from which the lower bound of the lemma follows. For the upper bound, it suffices to bound the quantity

$$1 - \lambda_i/\lambda^*.$$

We may assume that $||v^*||_A = 1$. We decompose

$$v^* = Q_i v^* + Q_i^{\perp} v^*.$$

Let $\mathbf{Rq}(w) = (Aw, w)/\|w\|^2$ denote the Rayleigh quotient. Then

$$1 - \lambda_i/\lambda^* = (\lambda^* - \lambda_i)\|v^*\|^2.$$

Now it follows from (3.2) that $Q_i v^* \neq 0$. Clearly,

$$\begin{aligned} 1 &= \mathbf{Rq}(Q_i v^*)\|Q_i v^*\|^2 + ||Q_i^\perp v^*||_A^2 \\ &\leq \lambda_i \|Q_i v^*\|^2 + ||Q_i^\perp v^*||_A^2. \end{aligned} \tag{3.4}$$

If $Q_i^\perp v^* = 0$ then $v^* \in V_i$. In this case, it follows that $\lambda^* = \lambda_i$ and there is nothing more to prove. If $Q_i^\perp v^* \neq 0$ then (3.4) implies

$$\begin{aligned} 1 &\leq \lambda_i \|Q_i v^*\|^2 + \mathbf{Rq}(Q_i^\perp v^*)\left\|Q_i^\perp v^*\right\|^2 \\ &= \lambda_i \|v^*\|^2 + (\mathbf{Rq}(Q_i^\perp v^*) - \lambda_i)\left\|Q_i^\perp v^*\right\|^2. \end{aligned}$$

Consequently,

$$\begin{aligned} 1 - \lambda_i/\lambda^* &\leq (\mathbf{Rq}(Q_i^\perp v^*) - \lambda_i)\left\|Q_i^\perp v^*\right\|^2 \\ &= [1 - \lambda_i/\mathbf{Rq}(Q_i^\perp v^*)]\ ||Q_i^\perp v^*||_A^2 \leq ||Q_i^\perp v^*||_A^2 \\ &\leq ||(I - Q_i)\tilde{Q}_i||_A^2 \leq \tilde{\theta}^2 \end{aligned}$$

where we used Lemma 3.3 for the last inequality. This completes the proof of Lemma 3.1

The next lemma provides a relationship between $\tilde{\theta}$ and $\tilde{\theta}_i$.

Lemma 3.4.

$$||(I - \tilde{Q})Q_s||_A \leq \left(\sum_{j=1}^{s} \tilde{\theta}_j^2\right)^{1/2}.$$

Moreover, if

$$||(I - \tilde{Q})Q_s||_A < 1 \tag{3.5}$$

then $dim(\tilde{V}) = s$ *and for* $i = 1, \ldots, s$,

$$\tilde{\theta}_i \leq \tilde{\theta} = ||\tilde{Q} - Q_s||_A = ||(I - \tilde{Q})Q_s||_A = ||(I - Q_s)\tilde{Q}||_A. \tag{3.6}$$

Proof. Let v be in V_s and write

$$v = \sum_{j=1}^{s} c_j v_j.$$

Clearly

$$\begin{aligned} ||(I - \tilde{Q})v||_A &= ||\sum_{j=1}^{s} c_j (I - \tilde{Q}) v_j||_A \\ &\leq \sum_{j=1}^{s} |c_j|\ ||(I - \tilde{Q})v_j||_A \\ &\leq \Big(\sum_{j=1}^{s} c_j^2\Big)^{1/2}\Big(\sum_{j=1}^{s} \tilde{\theta}_j^2\Big)^{1/2} = ||v||_A \Big(\sum_{j=1}^{s} \tilde{\theta}_j^2\Big)^{1/2}. \end{aligned}$$

It follows that

$$||(I - \tilde{Q})Q_s||_A \leq \big(\sum_{j=1}^{s} \tilde{\theta}_j^2\big)^{1/2}.$$

The lemma follows from Lemma 3.2.

Lemma 3.1 implies that for sufficiently small $\tilde{\theta}$, $\tilde{\lambda}_j \neq \lambda_i$ whenever $i, j \in \{1 \ldots s\}$ and $i \neq j$. Thus, we can define the quantities

$$\tilde{\sigma}_{ij} = \frac{\lambda_i}{|1 - \lambda_i/\tilde{\lambda}_j|} \text{ for } i \neq j,\ i, j = 1, \ldots, s.$$

Let

$$\tilde{\sigma}_i = \max_{j=1,\ldots,s,\, j\neq i} (\tilde{\sigma}_{ij}), \tag{3.7}$$

The next lemma provides bounds for the eigenvector error in terms of the approximation parameter $\tilde{\theta}_i$. Its proof is essentially based on the analysis given in [48] and is included for completeness.

Lemma 3.5. *Assume that (3.5) holds and that for a fixed index $i \in [1, \ldots, s]$, $\tilde{\lambda}_j \neq \lambda_i$ for $j \neq i$ and $j \in [1, \ldots, s]$. Then,*

$$||(I - P_i)\tilde{v}_i||_A^2 \leq \left[1 + \frac{\tilde{\sigma}_i^2}{(\lambda_1)^2}\tilde{\theta}^2\right]\tilde{\theta}_i^2 \equiv \tilde{C}_1\tilde{\theta}_i^2.$$

Here P_i is the projection onto the i'th eigenvector v_i.

Proof. Let $\tilde{P}_j$ denote the $(\cdot,\cdot)_A$ orthogonal projector onto the subspace generated by $\tilde{v}_j$. Note that

$$\begin{aligned} ||(I - P_i)\tilde{v}_i||_A &= ||(I - P_i)\tilde{P}_i||_A = ||(I - \tilde{P}_i)P_i||_A \\ &= ||(I - \tilde{P}_i)v_i||_A \end{aligned}$$

and that

$$\begin{aligned} ||(I - \tilde{P}_i)v_i||_A^2 &= ||(I - \tilde{Q})v_i||_A^2 + ||(\tilde{Q} - \tilde{P}_i)v_i||_A^2 \\ &= \tilde{\theta}_i^2 + ||(\tilde{Q} - \tilde{P}_i)v_i||_A^2. \end{aligned} \tag{3.8}$$

We estimate the last term above as follows: Note that

$$\tilde{Q} = \sum_{j=1}^{s} \tilde{P}_j \text{ and } \tilde{Q}A^{-1}\tilde{Q} = \sum_{j=1}^{s} \frac{1}{\tilde{\lambda}_j}\tilde{P}_j.$$

Thus,

$$\begin{aligned} (\frac{1}{\lambda_i} - \frac{1}{\tilde{\lambda}_i})\tilde{P}_i v_i + \sum_{j=1,\ j\neq i}^{s} (\frac{1}{\lambda_i} - \frac{1}{\tilde{\lambda}_j})\tilde{P}_j v_i &= \left(\frac{1}{\lambda_i} - \tilde{Q}A^{-1}\tilde{Q}\right)\tilde{Q}v_i \\ &= \tilde{Q}A^{-1}(I - \tilde{Q})v_i. \end{aligned} \tag{3.9}$$

The two terms on the left of (3.9) are $(\cdot,\cdot)_A$ orthogonal and hence

$$
\begin{aligned}
\frac{1}{\tilde\sigma_i}||(\tilde Q-\tilde P_i)v_i||_A &\le ||\sum_{j=1,\ j\neq i}^{s}(\frac{1}{\lambda_i}-\frac{1}{\tilde\lambda_j})\tilde P_j v_i||_A\\
&\le ||\tilde Q A^{-1}(I-\tilde Q)v_i||_A\\
&\le ||\tilde Q A^{-1}(I-\tilde Q)||_A||(I-\tilde Q)v_i||_A.
\end{aligned}
\tag{3.10}
$$

In the last product the second term is $\tilde\theta_i$ by definition and it remains now to estimate the first term. Since the projector Q_s commutes with A^{-1}, Lemma 3.4 and the triangle inequality give

$$
\begin{aligned}
||\tilde Q A^{-1}(I-\tilde Q)||_A &= ||\tilde Q(A^{-1}-\frac{1}{2\lambda_1}I)(I-\tilde Q)||_A\\
&\le ||\tilde Q(I-Q_s)(A^{-1}-\frac{1}{2\lambda_1}I)(I-\tilde Q)||_A\\
&\quad + ||\tilde Q(A^{-1}-\frac{1}{2\lambda_1}I)Q_s(I-\tilde Q)||_A \le \frac{\tilde\theta}{\lambda_1}.
\end{aligned}
\tag{3.11}
$$

Combining (3.10) and (3.11) gives

$$
||(\tilde Q-\tilde P_i)v_i||_A^2 \le \frac{\tilde\sigma_i^2}{\lambda_1^2}\tilde\theta^2\tilde\theta_i^2. \tag{3.12}
$$

The lemma follows combining (3.8) and (3.12).

The error in the corresponding eigenvalue is second order in $\tilde\theta_i$ and is bounded by the following lemma.

Lemma 3.6. *Assume that the hypotheses of Lemma 3.5 hold. Then,*

$$
0 \le 1-\lambda_i/\tilde\lambda_i \le ||(I-P_i)\tilde v_i||_A^2 \le \tilde C_1\tilde\theta_i^2.
$$

Proof. Using the identity $\tilde\lambda_i = \|\tilde v_i\|^{-2}$ gives

$$
1-\lambda_i/\tilde\lambda_i = ((A-\lambda_i)\tilde v_i,\tilde v_i).
$$

We clearly have

$$
\begin{aligned}
((A-\lambda_i)\tilde v_i,\tilde v_i) &= (A(I-P_i)\tilde v_i,(I-P_i)\tilde v_i)-\lambda_i((I-P_i)\tilde v_i,(I-P_i)\tilde v_i)\\
&\le ||(I-P_i)\tilde v_i||_A^2.
\end{aligned}
$$

This completes the proof of Lemma 3.6.

Both the eigenvectors $\{v_i\}$ and the corresponding approximate eigenvectors $\{\tilde v_i\}$ are orthonormal with respect to the $(\cdot,\cdot)_A$ inner product. The following lemma, based on [34], shows that for $i\neq j$, $|(v_i,\tilde v_j)_A|$ is small with θ since

$$
|(v_i,\tilde v_j)_A| = |((Q_s-P_j)v_i,\tilde v_j)_A| \le ||(Q_s-P_j)\tilde v_j||_A.
$$

Lemma 3.7. *Assume that (3.5) holds and that for a fixed index* $j \in [1, \dots, s]$, $\tilde{\lambda}_j \neq \lambda_i$ *for* $i \neq j$ *and* $j \in [1, \dots, s]$. *Then,*

$$||(Q_s - P_j)\tilde{v}_j||_A^2 \leq \tilde{C}_2 \tilde{\theta}^2 \tilde{\theta}_j^2$$

where

$$\tilde{C}_2 = \frac{(\tilde{\sigma}_j^*)^2 \tilde{C}_1}{(\lambda_1)^2}$$

and

$$\tilde{\sigma}_j^* = \max_{i=1,\dots,s,\, i\neq j} (\tilde{\sigma}_{ij}).$$

Proof. We clearly have

$$\frac{1}{\tilde{\sigma}_j^*} ||(Q_s - P_j)\tilde{v}_j||_A \leq ||(Q_s - P_j)(A^{-1} - \tilde{\lambda}_j^{-1} I)\tilde{v}_j||_A.$$

Since

$$((A^{-1} - \tilde{\lambda}_j^{-1} I)\tilde{v}_j, w)_A = 0 \quad \text{for all } w \in \tilde{V}, \tag{3.13}$$

it follows that

$$\begin{aligned} ||(Q_s - P_j)(A^{-1} - \tilde{\lambda}_j^{-1} I)\tilde{v}_j||_A &= ||(Q_s - P_j)(I - \tilde{Q})(A^{-1} - \tilde{\lambda}_j^{-1} I)\tilde{v}_j||_A \\ &\leq ||(Q_s - P_j)(I - \tilde{Q})||_A ||(A^{-1} - \tilde{\lambda}_j^{-1} I)\tilde{v}_j||_A. \end{aligned}$$

Now by Lemma 3.4,

$$||(Q_s - P_j)(I - \tilde{Q})||_A \leq ||Q_s(I - \tilde{Q})||_A = ||(I - \tilde{Q})Q_s||_A = \tilde{\theta}.$$

In addition, (3.13) implies that

$$\begin{aligned} ||(A^{-1} - \tilde{\lambda}_j^{-1} I)\tilde{v}_j||_A &\leq ||(A^{-1} - \lambda_j^{-1} I)\tilde{v}_j||_A = ||(A^{-1} - \lambda_j^{-1} I)(I - P_j)\tilde{v}_j||_A \\ &\leq ||(A^{-1} - \lambda_j^{-1} I)||_A ||(I - P_j)\tilde{v}_j||_A. \end{aligned}$$

The lemma follows combining the above estimates with Lemma 3.5.

The final lemma of this section shows that the approximation parameter $\tilde{\theta}_i$ can be bounded in terms of the orthogonal component of the approximate eigenvector.

Lemma 3.8. *If* $\tilde{C}_2 \tilde{\theta}^2 < 1$ *then*

$$\tilde{\theta}_i^2 \leq \left(1 - \tilde{C}_2 \tilde{\theta}^2\right)^{-1} ||Q_s^{\perp} \tilde{v}_i||_A^2.$$

Proof. It follows that

$$\begin{aligned} \tilde{\theta}_i^2 &= ||(I - \tilde{Q})v_i||_A^2 \leq ||(I - \tilde{P}_i)v_i||_A^2 \\ &= ||(I - P_i)\tilde{v}_i||_A^2 \\ &= ||Q_s^{\perp} \tilde{v}_i||_A^2 + ||(Q_s - P_i)\tilde{v}_i||_A^2. \end{aligned}$$

Thus, applying Lemma 3.7 gives

$$\tilde{\theta}_i^2 \leq ||Q_s^{\perp} \tilde{v}_i||_A^2 + \tilde{C}_2 \tilde{\theta}^2 \tilde{\theta}_i^2.$$

The lemma immediately follows.

4. Convergence analysis.

In this section, we will prove the theorem and corollary stated in Section 2. Their proofs are based on three additional lemmas as well as the lemmas for the Ritz eigenpairs.

We shall require some additional notation for this section. If θ^n is sufficiently small, Lemma 3.1 implies that $\lambda_j^n \neq \lambda_i$ whenever $i, j \in \{1 \ldots s\}$ and $i \neq j$. Following Section 3, we define the quantities

$$\sigma_{ij}(n) = \frac{\lambda_i}{|1 - \lambda_i/\lambda_j^n|} \text{ for } i \neq j, \; i, j = 1, \ldots, s$$

and set

$$\sigma_i(n) = \max_{j=1,\ldots,s,\, j\neq i}(\sigma_{ij}(n)) \text{ and } \sigma_j^*(n) = \max_{i=1,\ldots,s,\, i\neq j}(\sigma_{ij}(n)).$$

Let $V_s^\perp$ denote the A orthogonal complement of V_s. For each i in $[1, 2, \ldots, s]$, let $[\cdot,\cdot]_i$ denote the inner product

$$[v, w]_i = ((A - \lambda_i I)v, w) \text{ for all } v, w \in V_s^\perp.$$

The corresponding norm (on $V_s^\perp$) will be denoted by $|||\cdot|||_i \equiv [\cdot,\cdot]_i^{1/2}$. It is equivalent to the original norm,

$$|||v|||_i \leq ||v||_A \leq \left(\frac{\lambda_{s+1}}{\lambda_{s+1} - \lambda_i}\right)^{1/2} |||v|||_i, v \in V_s^\perp. \tag{4.1}$$

We start by considering the effect of a related iteration operator on the complement of V_s. For each i in $[1, 2, \ldots, s]$, let

$$E_i = I - Q_s^\perp B(A - \lambda_i I) \tag{4.2}$$

Note that E_i is symmetric on $V_s^\perp$ with respect to the inner product $[\cdot,\cdot]_i$. Moreover, it is a reducer as guaranteed by the following lemma.

Lemma 4.1. *For all w in $V_s^\perp$,*

$$|||E_i w|||_i \leq \delta_i |||w|||_i \tag{4.3}$$

where δ_i is given by (2.7).

Proof. From (1.2) it follows that

$$[1 - \gamma](A^{-1}w, w) \leq (Bw, w) \leq [1 + \gamma](A^{-1}w, w) \text{ for all } w \in V.$$

For $u \in V_s^\perp$,

$$\begin{aligned}[1 - \lambda_i/\lambda_{s+1}]((A - \lambda_i)u, u) &\leq (A^{-1}(A - \lambda_i)u, (A - \lambda_i)u) \\ &= ((I - \lambda_i A^{-1})u, (A - \lambda_i)u) \leq ((A - \lambda_i)u, u).\end{aligned}$$

Combining the above inequalities gives

$$[1-\gamma][1-\lambda_i/\lambda_{s+1}]((A-\lambda_i)u,u) \le (B(A-\lambda_i)u,(A-\lambda_i)u) \\ \le [1+\gamma]((A-\lambda_i)u,u) \ \text{ for all } u \in V_s^{\perp}.$$

This implies that

$$|[(I-Q_s^{\perp}B(A-\lambda_i I))u,u]_i| \le \delta_i |||u|||_i^2 \ \text{ for all } u \in V_s^{\perp}.$$

Inequality (4.3) follows from the symmetry of $(I-Q_s^{\perp}B(A-\lambda_i I))$ with respect to the $[\cdot,\cdot]_i$ inner product.

The next lemma shows that θ^{n+1} can be controlled in terms of θ^n.

Lemma 4.2. *Assume that*

$$\theta^n < \frac{\lambda_1}{(1+\gamma)\lambda_s^n}.$$

Define

$$C_0 = \left[1 + \frac{\lambda_s^n(1+\gamma)}{\lambda_1}\right]^2.$$

Then $dim V_s^{n+1} = s$ *and*

$$(\theta^{n+1})^2 \le C_0(\theta^n)^2. \tag{4.4}$$

Proof. By the triangle inequality,

$$\theta^{n+1} = ||Q_s^{n+1} - Q_s||_A \le ||Q_s^{n+1} - Q_s^n||_A + ||Q_s^n - Q_s||_A \\ = ||Q_s^{n+1} - Q_s^n||_A + \theta^n. \tag{4.5}$$

By construction, $\dim V_s^{n+1} \le \dim V_s^n$. We now estimate

$$||(I-Q_s^{n+1})Q_s^n||_A = \max_{v^n \in V_s^n, ||v^n||_A=1} ||(I-Q_s^{n+1})v^n||_A.$$

Let v^n be an arbitrary element of V_s^n with $||v^n||_A = 1$, $v^n = \sum_{j=1}^s \alpha_j v_j^n$ and consider $w = \sum_{j=1}^s \alpha_j \hat{v}_j^{n+1}$. We clearly have

$$||(I-Q_s^{n+1})v^n||_A \le ||v^n - w||_A \\ \le ||\sum_{j=1}^s \alpha_j(v_j^n - \hat{v}_j^{n+1})||_A.$$

By definition,

$$\sum_{j=1}^s \alpha_j(v_j^n - \hat{v}_j^{n+1}) = B\sum_{j=1}^s \alpha_j(Av_j^n - \lambda_j^n v_j^n).$$

It follows by (1.2) and the identity

$$((A^{-1}-(\lambda_j^n)^{-1}I)v_j^n, w)_A = 0 \quad \text{for all } w \in V_s^n$$

that

$$\begin{aligned}||\sum_{j=1}^{s} \alpha_j(v_j^n - \hat{v}_j^{n+1})||_A &\le (1+\gamma)||\sum_{j=1}^{s}\alpha_j(v_j^n - \lambda_j^n A^{-1} v_j^n)||_A \\ &= (1+\gamma)||\sum_{j=1}^{s}\alpha_j(I-Q_s^n)(v_j^n - \lambda_j^n A^{-1}v_j^n)||_A \\ &= (1+\gamma)||(I-Q_s^n)A^{-1}Q_s^n\sum_{j=1}^{s}\lambda_j^n\alpha_j v_j^n||_A\end{aligned}$$

Applying (3.11) gives

$$\begin{aligned}||\sum_{j=1}^{s}\alpha_j(v_j^n - \hat{v}_j^{n+1})||_A &\le (1+\gamma)\lambda_s^n||(I-Q_s^n)A^{-1}Q_s^n||_A \\ &\le (1+\gamma)\frac{\lambda_s^n}{\lambda_1}\theta^n.\end{aligned}$$

Finally, by the assumptions of the lemma, the last value is less than one, therefore, by Lemma 3.2, we have $\dim V_s^{n+1} = s$ and

$$||Q_s^{n+1} - Q_s^n||_A = ||(I-Q_s^{n+1})Q_s^n||_A \le (1+\gamma)\frac{\lambda_s^n}{\lambda_1}\theta^n.$$

The lemma follows combining (4.5) and the last estimate.

We will need the following technical lemma.

Lemma 4.3. *Let Δ be as in Section 2 and suppose that*

$$(\theta^n)^2 \le \frac{\alpha}{\Delta}$$

for some positive parameter $\alpha \le 1$. Then Lemmas 3.5 – 3.8 apply with $\tilde{V} = V_s^n$. Moreover, the quantities $\tilde{\sigma}_{ij} = \sigma_{ij}(n)$ are well defined, and

$$\lambda_i^n \le \lambda_i\big(1-\alpha\big)^{-1}, \tag{4.6}$$

$$\lambda_1 \le \sigma_i(n) \le \frac{\lambda_s\Delta}{2-\alpha}, \tag{4.7}$$

$$\lambda_1 \le \sigma_i^*(n) \le \frac{\lambda_s\Delta}{2-\alpha}. \tag{4.8}$$

Proof. Lemma 3.1 shows that for $i = 1, \dots, s$,

$$1 - \lambda_i/\lambda_i^n \le \alpha \frac{\lambda_{i+1} - \lambda_i}{\lambda_{i+1} + \lambda_i}.$$

A simple manipulation gives

$$\frac{1}{\lambda_i^n} \ge \frac{1}{\lambda_i}\left(1 - \alpha \frac{\lambda_{i+1} - \lambda_i}{\lambda_{i+1} + \lambda_i}\right). \tag{4.9}$$

Inequality (4.6) immediately follows.

The left inequalities in (4.7) and (4.8) are based on the simple estimate

$$\lambda_1 \le \sigma_{11}(n).$$

We next prove the right inequalities. Suppose that j is greater than i. Then $\lambda_i \le \lambda_{j-1} < \lambda_j \le \lambda_j^n$ and hence

$$\begin{aligned} \frac{\lambda_i}{|1 - \lambda_i/\lambda_j^n|} &= \frac{\lambda_i}{1 - \lambda_i/\lambda_j^n} \le \frac{\lambda_{j-1}}{1 - \lambda_{j-1}/\lambda_j} \\ &= \frac{\lambda_{j-1}\lambda_j}{\lambda_j - \lambda_{j-1}} \le \frac{\lambda_j(\lambda_{j-1} + \lambda_j)}{2(\lambda_j - \lambda_{j-1})}. \end{aligned} \tag{4.10}$$

If $j < i \le s$ then $\lambda_i \ge \lambda_{j+1} > \lambda_j^n$ and (4.9) imply

$$\begin{aligned} \frac{\lambda_i}{|1 - \lambda_i/\lambda_j^n|} &\le \frac{\lambda_{j+1}}{\lambda_{j+1}/\lambda_j^n - 1} \\ &\le \frac{\lambda_j\lambda_{j+1}}{\lambda_{j+1}(1 - \alpha(\lambda_{j+1} - \lambda_j)/(\lambda_{j+1} + \lambda_j)) - \lambda_j} \\ &\le \frac{\lambda_{j+1}(\lambda_j + \lambda_{j+1})}{(2 - \alpha)(\lambda_{j+1} - \lambda_j)}. \end{aligned} \tag{4.11}$$

The inequality (4.7) follows from (4.10) and (4.11). We have also shown (4.8).

The next lemma provides a perturbation estimate which will be used to develop bounds for convergence of the preconditioned subspace method.

Lemma 4.4. *Assume that* $\theta^n \le \Delta^{-1/2}, \theta^{n+1} \le \Delta^{-1/2}$ *and define*

$$\begin{aligned} C_1 &= 1 + \frac{\sigma_i^2(n)}{(\lambda_1)^2}(\theta^n)^2, \\ C_2 &= \frac{(\sigma_i^*(n))^2 C_1}{(\lambda_1)^2}, \\ C_3 &= \left[(1 - C_2(\theta^n)^2)(1 - \lambda_i/\lambda_{s+1})\right]^{-1}, \\ C_4 &= 1 + \frac{\sigma_i^2(n+1)}{(\lambda_1)^2}(\theta^{n+1})^2, \\ C_5 &= \sqrt{C_3 C_1}\left(1 + \frac{\lambda_i^n(1+\gamma)}{\lambda_1}\right)\left(1 - C_1(\theta_i^n)^2\right)^{-1/2}\left(\sqrt{C_4} + 1\right), \\ C_6 &= \left(\frac{C_3}{1 - C_1(\theta_i^n)^2}\right)^{1/2} \frac{(1+\gamma)}{\lambda_1}(\lambda_i^n C_1 + \lambda_i\sqrt{C_2}). \end{aligned}$$

If $C_2(\theta^n)^2 < 1$ *then*

$$|||Q_s^\perp v_i^{n+1}|||_i \le \tilde{\delta}_i |||Q_s^\perp v_i^n|||_i. \tag{4.12}$$

where

$$\tilde{\delta}_i = (1 - C_1(\theta_i^n)^2)^{-1/2}\delta_i + \theta^{n+1}C_5 + \theta^n C_6.$$

Proof. Let us first notice that Lemma 4.3 guarantees that Lemmas 3.5 – 3.8 can be applied with $\tilde{V} = V_s^n$ and with $\tilde{V} = V_s^{n+1}$, the quantities $\sigma_i(n)$, $\sigma_i^*(n)$, and $\sigma_i^2(n+1)$ are well defined and the assumption $C_2(\theta^n)^2 < 1$ of the lemma assures $C_1(\theta_i^n)^2 < 1$ because $C_2 \ge C_1$ by (4.8). Therefore, all constants above are well defined.

We start the proof with two technical estimates. By (1.2), it follows that

$$||B(Av_i^n - \lambda_i^n v_i^n)||_A \le (1+\gamma)||v_i^n - \lambda_i^n A^{-1} v_i^n||_A.$$

For $i = 1, \dots, s$,

$$(v_i^n - \lambda_i^n A^{-1} v_i^n, v_i^n)_A = 0$$

and hence

$$\begin{aligned} ||v_i^n - \lambda_i^n A^{-1} v_i^n||_A &\le \lambda_i^n/\lambda_i ||v_i^n - \lambda_i A^{-1} v_i^n||_A \\ &= \lambda_i^n/\lambda_i ||(I - \lambda_i A^{-1})(I - P_i)v_i^n||_A \\ &\le \lambda_i^n/\lambda_i ||I - \lambda_i A^{-1}||_A ||(I - P_i)v_i^n||_A \\ &\le \lambda_i^n/\lambda_1 ||(I - P_i)v_i^n||_A. \end{aligned} \tag{4.13}$$

Consequently, by Lemma 3.5

$$||B(Av_i^n - \lambda_i^n v_i^n)||_A \le \frac{(1+\gamma)\lambda_i^n}{\lambda_1}\sqrt{C_1}\theta_i^n. \tag{4.14}$$

Using the estimate above, we now prove that θ_i^n controls θ_i^{n+1}. Indeed,

$$\theta_i^{n+1} \le ||v_i - (v_i^n, v_i)_A^{-1}\hat{v}_i^{n+1}||_A = \frac{1}{|(v_i^n, v_i)_A|}||\hat{v}_i^{n+1} - P_i v_i^n||_A$$

Applying Lemma 3.5, we can estimate the first factor by

$$|(v_i^n, v_i)_A|^{-1} = ||P_i v_i^n||_A^{-1} \le (1 - C_1(\theta_i^n)^2)^{-1/2}. \tag{4.15}$$

We clearly have

$$\hat{v}_i^{n+1} - P_i v_i^n = (I - P_i)v_i^n + \hat{v}_i^{n+1} - v_i^n = (I - P_i)v_i^n - B(Av_i^n - \lambda_i^n v_i^n) \tag{4.16}$$

and thus (4.15), the triangle inequality, and (4.14) give

$$\theta_i^{n+1} \le (1 - C_1(\theta_i^n)^2)^{-1/2}\sqrt{C_1}\left(1 + \frac{\lambda_i^n(1+\gamma)}{\lambda_1}\right)\theta_i^n. \tag{4.17}$$

We are now ready to prove the lemma. Define

$$\begin{aligned} u_i^n &= \frac{v_i^n}{(v_i^n, v_i)_A}, \\ u_i^{n+1} &= \frac{v_i^{n+1}}{(v_i^{n+1}, v_i)_A}, \\ \hat{u}_i^{n+1} &= \frac{\hat{v}_i^{n+1}}{(v_i^n, v_i)_A}, \\ w_i^{n+1} &= E_i Q_s^\perp u_i^n. \end{aligned}$$

Then

$$|||Q_s^\perp v_i^{n+1}|||_i = |||Q_s^\perp u_i^{n+1}|||_i |(v_i^{n+1}, v_i)_A|.$$

Applying the triangle inequality to

$$Q_s^\perp u_i^{n+1} = w_i^{n+1} + Q_s^\perp(u_i^{n+1} - \hat{u}_i^{n+1}) + (Q_s^\perp \hat{u}_i^{n+1} - w_i^{n+1})$$

and Lemma 4.1 with $w = Q_s^\perp u_i^n$ gives

$$\begin{aligned} |||Q_s^\perp u_i^{n+1}|||_i \le \delta_i |||Q_s^\perp u_i^n|||_i &+ |||Q_s^\perp(u_i^{n+1} - \hat{u}_i^{n+1})|||_i \\ &+ |||Q_s^\perp \hat{u}_i^{n+1} - w_i^{n+1}|||_i. \end{aligned}$$

By multiplying both sides by $|(v_i^{n+1}, v_i)_A|$ and using the estimate $|(v_i^{n+1}, v_i)_A| \le 1$ for all terms on the right but the second one, it follows that

$$\begin{aligned} |||Q_s^\perp v_i^{n+1}|||_i \le \delta_i |||Q_s^\perp u_i^n|||_i &+ |||Q_s^\perp(u_i^{n+1} - \hat{u}_i^{n+1})|||_i |(v_i^{n+1}, v_i)_A| \\ &+ |||Q_s^\perp \hat{u}_i^{n+1} - w_i^{n+1}|||_i. \end{aligned} \tag{4.18}$$

We estimate the three terms in (4.18) separately. For the first, by (4.15)

$$|||Q_s^\perp u_i^n|||_i = \frac{|||Q_s^\perp v_i^n|||_i}{|(v_i^n, v_i)_A|} \le (1 - C_1(\theta_i^n)^2)^{-1/2} |||Q_s^\perp v_i^n|||_i. \tag{4.19}$$

By (4.1), the norm part of the second term of (4.18) can be bounded by

$$\begin{aligned} |||Q_s^\perp(u_i^{n+1} - \hat{u}_i^{n+1})|||_i &\le ||Q_s^\perp(u_i^{n+1} - \hat{u}_i^{n+1})||_A \\ &= ||Q_s^\perp Q_s^{n+1}(u_i^{n+1} - \hat{u}_i^{n+1})||_A \\ &\le \theta^{n+1} ||u_i^{n+1} - \hat{u}_i^{n+1}||_A. \end{aligned} \tag{4.20}$$

By (4.16) and the fact that $P_i u_i^n = v_i$,

$$u_i^{n+1} - \hat{u}_i^{n+1} = u_i^{n+1} - v_i - (I - P_i)u_i^n + B(Au_i^n - \lambda_i^n u_i^n)$$

and the triangle inequality gives

$$\begin{aligned} ||u_i^{n+1} - \hat{u}_i^{n+1}||_A &\le ||(I - P_i)u_i^{n+1}||_A + ||(I - P_i)u_i^n||_A + ||B(Au_i^n - \lambda_i^n u_i^n)||_A \\ &= \frac{||(I - P_i)v_i^{n+1}||_A}{|(v_i^{n+1}, v_i)_A|} + \frac{||(I - P_i)v_i^n||_A + ||B(Av_i^n - \lambda_i^n v_i^n)||_A}{|(v_i^n, v_i)_A|}. \end{aligned}$$

Multiplying both sides by $|(v_i^{n+1}, v_i)_A|$ and using the estimates $|(v_i^{n+1}, v_i)_A| \le 1$ and (4.15) for the last term give

$$||u_i^{n+1} - \hat{u}_i^{n+1}||_A |(v_i^{n+1}, v_i)_A| \le ||(I - P_i)v_i^{n+1}||_A + \frac{||(I - P_i)v_i^n||_A + ||B(Av_i^n - \lambda_i^n v_i^n)||_A}{\sqrt{1 - C_1(\theta_i^n)^2}}.$$

Applying Lemma 3.5 with $\tilde{V} = V_s^{n+1}$ and $\tilde{V} = V_s^n$ and (4.14) gives

$$||u_i^{n+1} - \hat{u}_i^{n+1}||_A |(v_i^{n+1}, v_i)_A| \le \sqrt{C_4}\theta_i^{n+1} + \left[1 + \frac{(1+\gamma)\lambda_i^n}{\lambda_1}\right] \frac{\sqrt{C_1}}{\sqrt{1 - C_1(\theta_i^n)^2}} \theta_i^n. \tag{4.21}$$

By Lemma 3.8 and (4.1),

$$(\theta_i^n)^2 \le \left[1 - C_2(\theta^n)^2\right]^{-1} ||Q_s^\perp v_i^n||_A^2 \le C_3 |||Q_s^\perp v_i^n|||_i^2. \tag{4.22}$$

Combining (4.20), (4.21), (4.17) and (4.22) gives

$$|||Q_s^\perp (u_i^{n+1} - \hat{u}_i^{n+1})|||_i |(v_i^{n+1}, v_i)_A| \le C_5 \theta^{n+1} |||Q_s^\perp v_i^n|||_i. \tag{4.23}$$

For the last term in (4.18), we use the fact that

$$Q_s^\perp \hat{u}_i^{n+1} - w_i^{n+1} = (\lambda_i^n - \lambda_i) Q_s^\perp B u_i^n - Q_s^\perp B(A - \lambda_i I)(Q_s - P_i) u_i^n$$

and hence by (4.1), (1.2), by Lemmas 3.6, 3.4, and 3.7, and by (4.15),

$$\begin{aligned} |||Q_s^\perp \hat{u}_i^{n+1} - w_i^{n+1}|||_i &\le (1+\gamma)|(v_i^n, v_i)_A|^{-1} \big\{ \lambda_i^n (1 - \lambda_i/\lambda_i^n) ||A^{-1} v_i^n||_A \\ &\quad + ||(I - \lambda_i A^{-1})(Q_s - P_i) v_i^n||_A \big\} \\ &\le \frac{(1+\gamma)}{\lambda_1} \frac{\lambda_i^n C_1 + \lambda_i \sqrt{C_2}}{\sqrt{1 - C_1(\theta_i^n)^2}} \theta^n \theta_i^n \\ &\le C_6 \theta^n |||Q_s^\perp v_i^n|||_i. \end{aligned} \tag{4.24}$$

We used (4.22) for the last inequality above. Combining (4.18), (4.19), (4.23) and (4.24) verifies (4.12). This completes the proof of the lemma.

Proof of Theorem 2.1. Assume that (2.4) holds. Let β be defined by

$$\beta = \frac{(1-\gamma)^2(\lambda_1)^2(1 - \lambda_s/\lambda_{s+1})^3}{1941(\lambda_s)^2}. \tag{4.25}$$

We use mathematical induction to show that for $k \ge 0$,

$$(\theta^k)^2 \le \beta \frac{(\lambda_1)^2}{\Delta^2 (\lambda_s)^2} \tag{4.26}$$

and

$$|||Q_s^{\perp} v_i^k|||_{\,i} \le \bar{\delta}_i^k |||Q_s^{\perp} v_i^0|||_{\,i} \tag{4.27}$$

where $\bar{\delta}_i$ is given by (2.8). It follows from (2.4) and Lemma 3.4 that (4.26) holds for $k=0$. The inequality (4.27) holds trivially for $k=0$.

Assume that (4.26) and (4.27) hold for $k=n$, we show that they also hold for $k=n+1$. The induction assumption (4.26) with $k=n$, (4.6), (4.7), $\gamma<1$, (4.25) and elementary manipulations give

$$\begin{aligned} C_0 &\le 9.0062(\lambda_s)^2/(\lambda_1)^2,\\ C_1 &\le 1+\beta/(2-\beta)^2 \le 1.00013. \end{aligned} \tag{4.28}$$

Thus by (4.4) and (4.26) with $k=n$,

$$(\theta^{n+1})^2 \le 9.0062(\lambda_s)^2/(\lambda_1)^2(\theta^n)^2 \le 9.0062\Delta^{-2}\beta. \tag{4.29}$$

By (4.8),

$$\sigma_i^2(n+1) \le \frac{(\lambda_s)^2\Delta^2}{(2-9.0062\beta)^2}.$$

We also have

$$\begin{aligned} C_1(\theta^n)^2 &\le \beta(1+\beta/(2-\beta)^2) \le .00052,\\ C_2(\theta^n)^2 &\le \beta(1+\beta/(2-\beta)^2)/(2-\beta)^2 \le .00013,\\ \Big(1-C_2(\theta^n)^2\Big)^{-1} &\le 1.00013,\\ C_3 &\le \frac{1.00013}{1-\lambda_i/\lambda_{s+1}},\\ C_4 &\le 1+\frac{9.0062(\lambda_s)^2\beta}{(2-9.0062\beta)^2(\lambda_1)^2} \le 1.00117,\\ \theta^{n+1}C_5 &\le 18.025\sqrt{\beta}\frac{\lambda_i}{\lambda_1},\\ \theta^n C_6 &\le 3.0031\Big(\frac{\beta}{1-\lambda_i/\lambda_{s+1}}\Big)^{1/2}\frac{\lambda_i}{\lambda_1}. \end{aligned}$$

We used the inequality $1-\lambda_i/\lambda_{s+1}\le\Delta$ for the next to last inequality above and (4.6) in the last two inequalities above.

Applying Lemma 4.3 gives that (4.12) holds with

$$\begin{aligned} \tilde{\delta}_i &= (1+1.00026\sqrt{\beta})\delta_i + 18.025\sqrt{\beta}\frac{\lambda_i}{\lambda_1}\\ &\quad + 3.0031\Big(\frac{\beta}{1-\lambda_i/\lambda_{s+1}}\Big)^{1/2}\frac{\lambda_i}{\lambda_1}\\ &\le \delta_i + 22.028\Big(\frac{\beta}{1-\lambda_i/\lambda_{s+1}}\Big)^{1/2}\frac{\lambda_i}{\lambda_1}. \end{aligned} \tag{4.30}$$

It then follows from (4.25) that

$$\tilde{\delta}_i \leq \delta_i + \frac{(1-\gamma)(1-\lambda_s/\lambda_{s+1})}{2}\left(\frac{\lambda_{s+1}-\lambda_s}{\lambda_{s+1}-\lambda_i}\right)^{1/2}\frac{\lambda_i}{\lambda_s} = \bar{\delta}_i$$

and hence (4.27) holds for $k = n+1$.

The bound (4.29) is not strong enough to imply (4.26) for $k = n+1$. We get a better bound as follows. The above inequalities imply that

$$(\theta^{n+1})^2 \leq 9.0062\Delta^{-2}\beta \leq .00464\frac{(\lambda_1)^2}{(\lambda_s)^2\Delta^2},$$

$$\frac{(\sigma_i^*(n+1))^2}{(\lambda_1)^2}C_4(\theta^{n+1})^2 \leq .00117.$$

Applying Lemma 3.4, (4.27) and the inequality analogous to (4.22) gives

$$\begin{aligned}(\theta^{n+1})^2 \leq \sum_{i=1}^{s}(\theta_i^{n+1})^2 &\leq \frac{1.00117}{1-\lambda_i/\lambda_{s+1}}\sum_{i=1}^{s}|||Q_s^\perp v_i^{n+1}|||_i^2 \\ &\leq \frac{1.00117}{1-\lambda_i/\lambda_{s+1}}\sum_{i=1}^{s}||Q_s^\perp v_i^0||_A^2.\end{aligned}$$

Now

$$Q_s^\perp v_i^0 = (I-P_i)v_i^0 - (Q_s - P_i)v_i^0 \tag{4.31}$$

and hence applying Lemmas 3.5 and 3.7 gives

$$\begin{aligned}(\theta^{n+1})^2 &\leq \frac{1.00117}{1-\lambda_i/\lambda_{s+1}}\sum_{i=1}^{s}\big((1+\sqrt{\beta})C_1 + (1+1/\sqrt{\beta})C_2(\theta^0)^2\big)(\theta_i^0)^2 \\ &\leq \frac{1.02984}{1-\lambda_i/\lambda_{s+1}}\sum_{i=1}^{s}(\theta_i^0)^2.\end{aligned}$$

In the above inequality, the constants C_1 and C_2 correspond to $n = 0$. The inequality (4.26) for $k = n+1$ then follows from (2.4). This completes the induction part of the theorem.

We now show that (4.26) and (4.27) imply (2.5) and (2.6). Lemmas 3.5 and 3.7, (4.22) and (4.31) give

$$\begin{aligned}||(I-P_i)v_i^n||_A^2 &\leq C_1C_3|||Q_s^\perp v_i^n|||_i^2 \leq C_1C_3\bar{\delta}_i^{2n}|||Q_s^\perp v_i^0|||_i^2 \\ &\leq C_1C_3\big((1+\sqrt{\beta})C_1 + (1+1/\sqrt{\beta})C_2(\theta^0)^2\big)\bar{\delta}_i^{2n}(\theta_i^0)^2 \\ &\leq \frac{1.03}{1-\lambda_i/\lambda_{s+1}}\bar{\delta}_i^{2n}(\theta_i^0)^2.\end{aligned}$$

A similar argument using Lemma 3.6 proves (2.6). This completes the proof of the theorem.

Proof of Corollary 2.1. It follows from (3.6), (4.27) and Lemma 3.8 that for any $m > 0$,

$$(\theta^m)^2 \le C_3 \bar{\delta}_s^{2m} \sum_{i=1}^{s} |||Q_s^\perp v_i^0|||_i^2.$$

Let

$$M = \epsilon^{-2}$$

and choose m so large that

$$\sum_{i=1}^{s} (\theta_i^m)^2 \le \frac{(1-\gamma)^2 (\lambda_1)^4 (1-\lambda_s/\lambda_{s+1})^4}{1999 M \Delta^2 (\lambda_s)^4}.$$

Following the proof of the theorem, we prove by mathematical induction, that for $k \ge m$,

$$(\theta^k)^2 \le \beta \frac{(\lambda_1)^2}{M \Delta^2 (\lambda_s)^2}$$

and

$$|||Q_s^\perp v_i^k|||_i \le (\delta_i + \epsilon(\bar{\delta}_i - \delta_i))^{k-m} |||Q_s^\perp v_i^m|||_i.$$

It then follows that

$$||(I - P_i) v_i^n||_A^2 \le \frac{1.03}{1 - \lambda_i/\lambda_{s+1}} \big(\delta_i + \epsilon(\bar{\delta}_i - \delta_i)\big)^{2n-2m} \bar{\delta}_i^{2m} (\theta_i^0)^2$$

and

$$0 \le 1 - \lambda_i/\lambda_i^n \le \frac{1.03}{1 - \lambda_i/\lambda_{s+1}} \big(\delta_i + \epsilon(\bar{\delta}_i - \delta_i)\big)^{2n-2m} \bar{\delta}_i^{2m} (\theta_i^0)^2.$$

This completes the proof of the corollary.

5. Numerical Experiments.

The results of numerical experiments involving the preconditioned subspace iteration technique are given in this section. We first give results for the case when the eigenvalues are well separated. We also consider an example where a group of eigenvalues are distinct but much closer together. For comparison, we include the results of numerical experiments involving a block preconditioned method analyzed in [23], [24], [22]. All of our experiments include subspaces of multiplicity greater than one.

The model problem which we shall consider is the one dimensional eigenvalue problem

$$\begin{aligned} u - \frac{\partial^2 u}{\partial x^2} &= \lambda u \quad \text{on } (0,1) \\ u(0) = u(1) \text{ and } \frac{\partial u}{\partial x}(0) &= \frac{\partial u}{\partial x}(1). \end{aligned} \tag{5.1}$$

The smallest eigenvalue for (5.1) is one and its eigenspace is equal to the space of constants. The larger eigenvalues are given by $\lambda_j = 1 + 4\pi^2 j^2$ for $j = 1, 2, \ldots$. The corresponding eigenspace has dimension two and is spanned by the vectors

$$\tilde{v}_j^+ = \cos(2\pi j x), \quad \tilde{v}_j^- = \sin(2\pi j x).$$

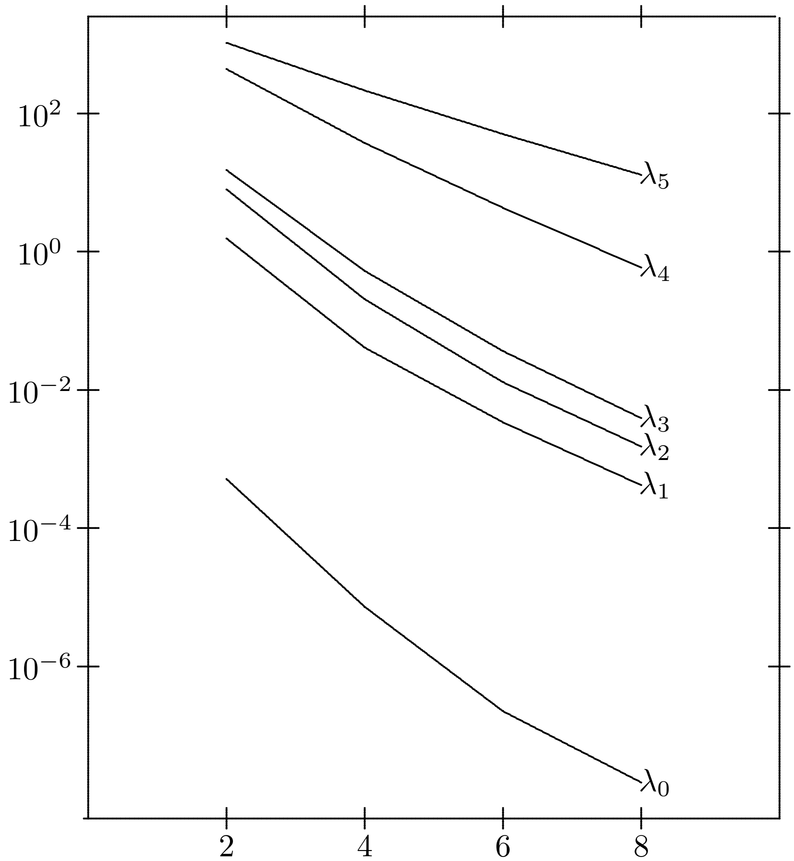

Figure 1.
Eigenvalue error: The case of well separated eigenvalues.

We approximate (5.1) by using a spectral approximation. Let n be given and define $h = (2n)^{-1}$ and $x_i = ih$. V is the space of $2n$ dimensional vectors with inner product

$$(v, w) = h \sum_{i=1}^{2n} v_i w_i.$$

The i'th component of a vector $v \in V$ corresponds to the nodal value at x_i. Consider the functions

$$\{\theta_i\} = \{1, \tilde{v}_1^+, \tilde{v}_1^-, \ldots, \tilde{v}_n^+, \tilde{v}_n^-, \tilde{v}_{n+1}^+\} \tag{5.2}$$

and let $\tilde{V}$ be the corresponding linear span. Given a vector $v \in V$ there exists a unique function $\tilde{v} \in \tilde{V}$ such that

$$\tilde{v}(x_i) = v_i.$$

In fact, the transformation from $v \in V$ to the coefficients of $\tilde{v}$ in the basis (5.2) can be rapidly computed by use of the fast Fourier transform. The operator $A : V \mapsto V$ is defined by

$$(Av, w) = \int_0^1 (\tilde{v}\tilde{w} + \frac{\partial \tilde{v}}{\partial x} \frac{\partial \tilde{w}}{\partial x})\ dx.$$

The eigenvalues of A coincide with those of (5.1). In addition, the eigenvectors of A are of the form $(1, \dots, 1)$, $(v_j^+)_i = \tilde{v}_j^+(x_i)$ and $(v_j^-)_i = \tilde{v}_j^-(x_i)$.

It is well known that different numerical approximations can often be used as preconditioners for each other. Accordingly, we consider the finite difference approximation

$$(\mathcal{A}v)_i = v_i + h^{-2}(2v_i - v_{i-1} - v_{i+1}). \tag{5.3}$$

We interpret $i-1$ and $i+1$ modulo n above. The use of $\mathcal{A}^{-1}$ as a preconditioner for A would be a bit too trivial since $\mathcal{A}$ and A share the same eigenvectors. Instead, we define B by applying an additive overlapping domain decomposition preconditioner for (5.3) [21], [12]. This is a more realistic choice since the resulting preconditioner B does not commute with A.

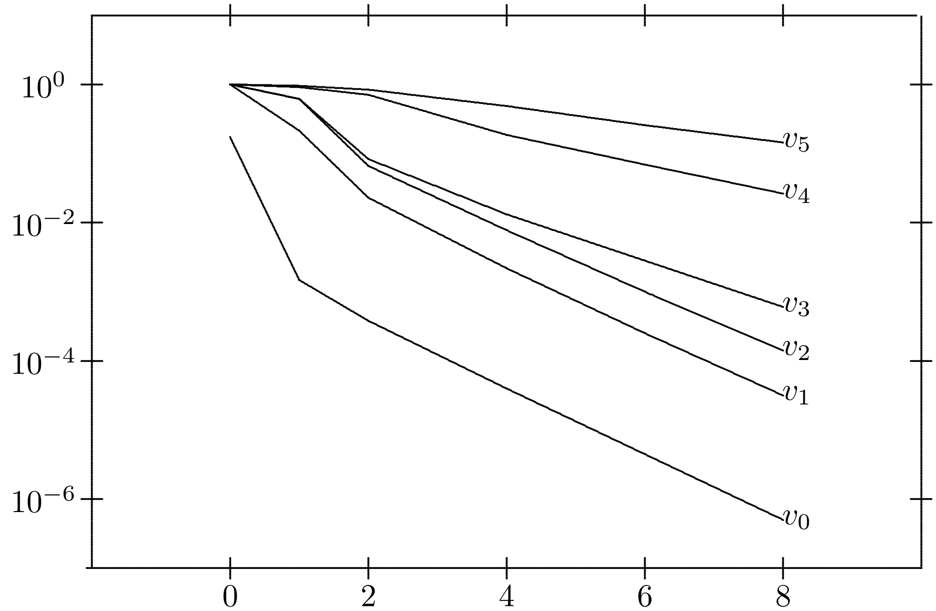

Figure 2.
Eigenvector error: The case of well separated eigenvalues.

For all of the numerical results reported, we use eight overlapping subdomains (of size 1/4) and a coarse grid of size $H = 1/8$. For this problem, the additive preconditioner was scaled so that (1.1) holds with $\gamma = 2/3$ (independently of n). Such a value of γ is not unusual in many applications. For example, a well designed multigrid V-cycle gives rise to $\gamma \approx 1/10$. In all of our runs, we took $n = 256$. The results for other n were qualitatively the same.

Figure 1 illustrates the eigenvalue convergence for the eigenvalues

$$\begin{aligned} \lambda_0 &= 1 \\ \lambda_1 &= 40.48 \\ \lambda_2 &= 158.9 \\ \lambda_3 &= 356.3 \\ \lambda_4 &= 632.7 \\ \lambda_5 &= 988.0. \end{aligned}$$

We applied the above algorithm with an eleven dimensional subspace. The initial subspace was chosen to be the space spanned by 11 vectors with entries consisting of random numbers uniformly distributed in the interval [0,1]. No attempt was made to generate a sufficiently accurate starting subspace as required by the theory. Nevertheless, the method converged extremely well. Note the faster rate of convergence for the smaller eigenvalues. Figure 1 reports the actual error. Thus after eight iterations the error in the fifth eigenvalue was less than 1.5 percent.

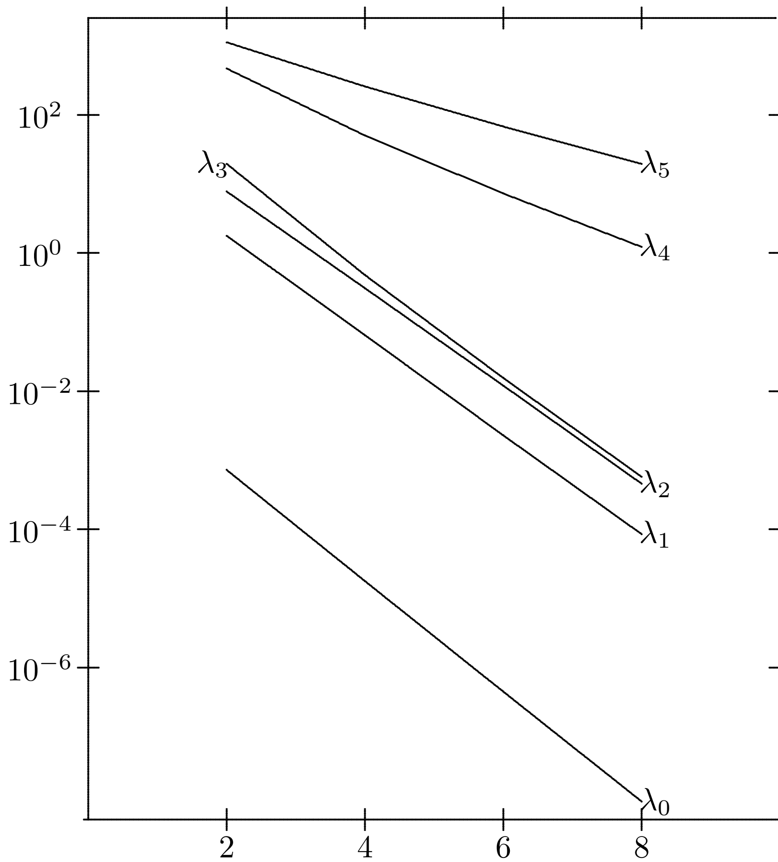

Figure 3.
Eigenvalue error: The case of clustered eigenvalues.

We did not report the eigenvalue error for the initial subspace. The reason for this is that these errors are very large and depend on the largest eigenvalue of the system. For the reported example ($n = 256$), the eigenvalue error in the initial subspace was on the order of 10^6. This is to be expected since the initial subspace is chosen at random and is dominated by high frequency components. It is somewhat surprising that the method reduces these errors so rapidly. The eigenvector convergence behavior is illustrated in Figure 2. Here we report the eigenvector error as measured by

$$e_i^n = \frac{\left\| v_i - \tilde{Q}_i^n v_i \right\|}{\|v_i\|}$$

where $\tilde{Q}_i^n$ is the $(\cdot,\cdot)$ orthogonal projector onto the the space V_i^n spanned by the approximate eigenvectors. The first eigenvalue is of multiplicity zero and hence

V_0^n is one dimensional. The remaining eigenvalues are of multiplicity two and V_i^n involves the span of subsequent pairs of approximate eigenvectors. For the multiplicity two case, we report the larger of the two values of e_i^n. As predicted by the theory, more rapid convergence is observed for the eigenvectors associated with the smaller eigenvalues.

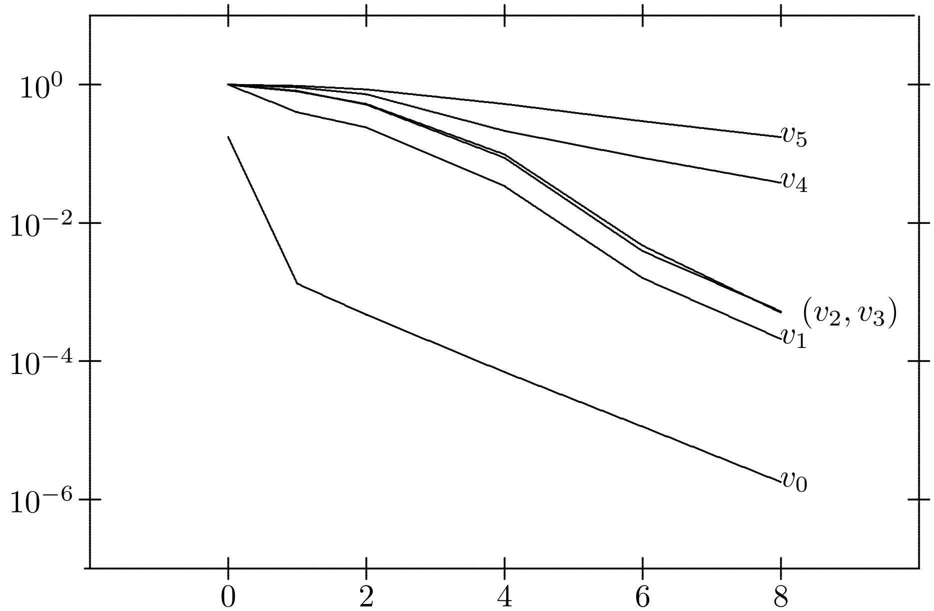

Figure 4.
Eigenvector error: The case of clustered eigenvalues.

The estimates developed earlier in this manuscript deteriorate in the case of eigenvalues with little separation. The next example suggests that the preconditioned subspace method still works well even when the eigenvalues cluster. For this example, we no longer use (5.1) but still preserve the same eigenspaces. We simply move the second and fourth eigenvalues so that they are significantly closer to the third. Specifically, we apply the preconditioned subspace method to the problem with eigenvalues

$$\begin{aligned}
\lambda_0 &= 1 \\
\lambda_1 &= 157.7 \\
\lambda_2 &= 158.9 \\
\lambda_3 &= 160.9 \\
\lambda_4 &= 632.7 \\
\lambda_5 &= 988.0.
\end{aligned}$$

For this problem, there is a six dimensional eigenspace with eigenvalues which are separated by less than two percent. We used the same preconditioner as in the first example but in this case (since A has changed) we have $\delta \leq .7$.

Figure 3 illustrates that we still get rapid eigenvalue convergence. In fact, there is not a lot of difference in the performance when compared to the well separated case in Figure 1. Figure 4 gives the eigenvector convergence measured the same way as reported in the first example (Figure 2). Note that the eigenspaces corresponding to the three clustered eigenvalues still converge quite rapidly.

As our final example, we consider the DK block iterative method studied in [23], [24], [22]. We choose this method because it was the only other method available with a fully developed convergence theory. This method uses the shift corresponding to the largest eigenvalue in the subspace for all vectors in the space. Thus,

$$V_s^{n+1} = \text{span}\{\bar{v}_1^{n+1}, \dots, \bar{v}_s^{n+1}\}. \tag{5.4}$$

where

$$\bar{v}_i^{n+1} = v_i^n - B(Av_i^n - \lambda_s^n v_i^n) \text{ for } i = 1, \dots, s \tag{5.5}$$

The above scheme is somewhat easier to implement than that studied in this paper. This is because the Ritz eigenvectors $\{v_i^n\}$ in (5.5) can be replaced by any spanning set for the subspace V_s^n, e.g., $\{\bar{v}_i^n\}$.

It is shown in [23], [24], [22] that convergence is acheived for the eigenspace corresponding to the largest eigenvector in the subspace V_s. In contrast to the theory of the present paper, the convergence estimates for the DK method do not depend on the clustering of eigenvalues.

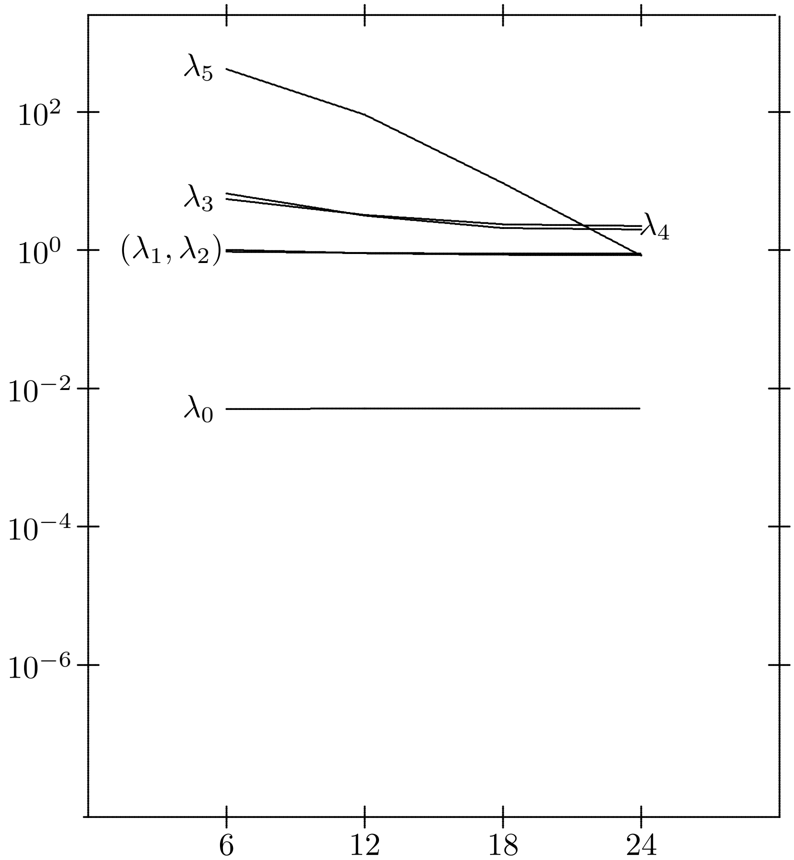

Figure 5.
Eigenvalue error: DK method, clusered eigenvalues.

To compare the method of the paper with the DK method, we ran the DK method for the clustered eigenvalue example. The eigenvalue convergence is illustrated in Figure 5. The figure shows that the DK method does exhibit a rate of convergence for the largest eigenvalue λ_5. However, it should be noted that the accuracy achieved by the DK method for this eigenvalue using 18 iterations was not as good as that by the method of this paper using 8 iterations. Note that the approximations for the lower eigenvalues are not improving. This is what one would expect as the theory does not suggest any convergence for these eigenvalues.

The initial steps of the DK method show good convergence on the smaller eigenvalues and eigenvectors. The eigenvector convergence is illustrated in Figure 6. Although a uniform convergence rate is achieved for the eigenspace corresponding to the largest eigenvector, the method stops converging the remaining eigenspaces. The method of this paper gives rise to a faster convergent approximation to the eigenspace corresponding to the largest eigenvalue λ_5 while converging the smaller eigenspaces at much better rates.

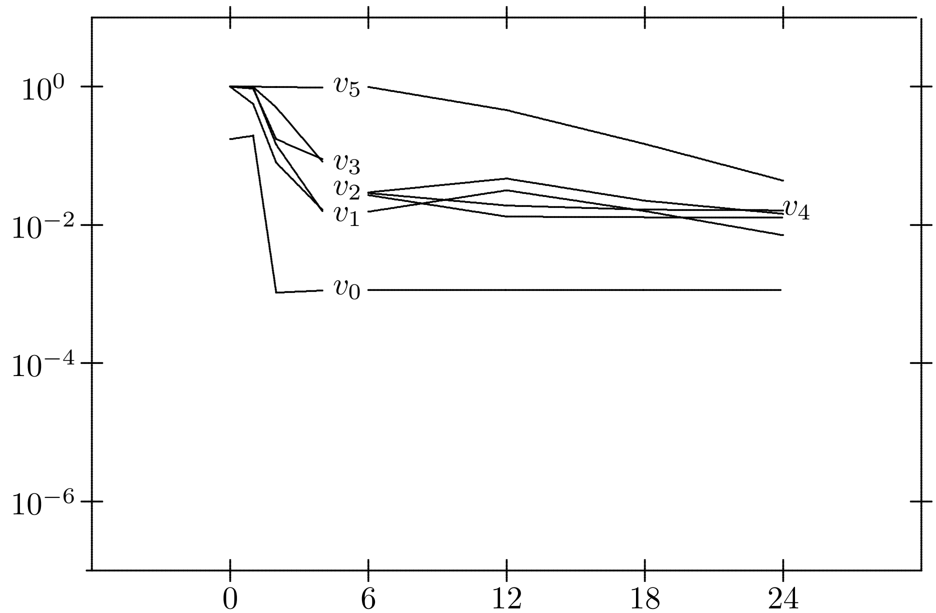

Figure 6.
Eigenvector error: DK method, clustered eigenvalues.

References

1. O. Axelsson, *A generalized conjugate gradient, least squares method*, Numer. Math. **51** (1987), 209–228.
2. G. Birkhoff and A. Schoenstadt, Eds, *Elliptic Problem Solvers II*, Academic Press, New York, 1984.
3. P.E. Bjørstad and O.B. Widlund, *Solving elliptic problems on regions partitioned into substructures*, Elliptic Problem Solvers II (G. Birkhoff and A. Schoenstadt, eds.), Academic Press, New York, 1984, pp. 245–256.
4. J.H. Bramble, *Multigrid Methods*, Pitman Research Notes in Mathematics Series, Longman Scientific & Technical, London. Copublished with John Wiley & Sons, Inc., New York, 1993.
5. J.H. Bramble and J.E. Pasciak, *New convergence estimates for multigrid algorithms*, Math. Comp. **49** (1987), 311–329.

6. J.H. Bramble and J.E. Pasciak, *New estimates for multigrid algorithms including the V-cycle*, Math. Comp. **60** (1993), 447–471.
7. J.H. Bramble, J.E. Pasciak and A.H. Schatz , *An iterative method for elliptic problems on regions partitioned into substructures*, Math. Comp. **46** (1986), 361–369.
8. J.H. Bramble, J.E. Pasciak and A.H. Schatz , *The construction of preconditioners for elliptic problems by substructuring, I*, Math. Comp. **47** (1986), 103–134.
9. J.H. Bramble, J.E. Pasciak and A.H. Schatz , *The construction of preconditioners for elliptic problems by substructuring, II*, Math. Comp. **49** (1987), 1–16.
10. J.H. Bramble, J.E. Pasciak and A.H. Schatz , *The construction of preconditioners for elliptic problems by substructuring, III*, Math. Comp. **51** (1988), 415–430.
11. J.H. Bramble, J.E. Pasciak and A.H. Schatz , *The construction of preconditioners for elliptic problems by substructuring, IV*, Math. Comp. **53** (1989), 1–24.
12. J.H. Bramble, J.E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp. **57** (1991), 1–21.
13. J.H. Bramble, J.E. Pasciak and J. Xu, *Parallel multilevel preconditioners*, Math. Comp. **55** (1990), 1–22.
14. Zhihao Cao, *Generalized Rayleigh quotient matrix and bloc algorith for solving large sparse symmetric generalized eigenvalue problems*, Numerical Math., J. of China Univ. **5** (1983), 342–348.
15. T.F. Chan, R. Glowinski, J. Periaux, and O.B. Widlund, Eds., *Domain Decomposition Methods*, SIAM, Phil. PA, 1989.
16. T. Chan, R. Glowinski, J. Periaux and and O.B. Widlund, Eds., *Third Inter. Symp. on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Phil. PA, 1990.
17. N. Chetty, M. Weinert, T.S. Rahman, and J.W. Davenport, *Vacancies and impurities in aluminum and magnesium*, Pys. Rev. B, 6313–6326.
18. J. Davidson, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices*, Jour. Comp. Phys. **17** (1975), 87–94.
19. J. Davidson, *Matrix eigenvector methods*, Methods in Computational Molecular Physics (1983), Reidel, Boston, 95–113.
20. A. Stathopoulos, Y. Saad and C.F. Fischer, *Robust Preconditioning of Large, Sparse, Symmetric Eigenvalue Problems*, Journal of Computational and Applied Mathematics (to appear, Also report of AHPCRC, U of Minneapolis, TR-93-093).
21. M. Dryja and O. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*, Technical Report, Courant Institute of Mathematical Sciences **339** (1987).
22. E.G. D'yakonov, *Optimization in Solving Elliptic Problems*, CRC, Boca Raton, 1995.
23. E.G. D'yakonov and A.V. Knyazev, *Group iterative method for finding lower-order eigenvalues*, Moscow University, Ser. 15, Computational Math. and Cybernetics **2** (1982), 32–40.
24. E.G. D'yakonov and A.V. Knyazev, *On an iterative method for finding lower eigenvalues*, Russian J. of Numerical Analysis and Math. Modelling **7** (1992), 473–486.
25. E.G. D'yakonov and M.Yu. Orekhov, *Minimization of the computational labor in determining the first eigenvalues of differential operators*, Math. Notes **27** (1980), 382–391.
26. E.G. D'yakonov and M.Yu. Orekhov, *On some iterative methods for eigenvalue problems*, Math. Notes **34** (1983), 879–912.
27. R. Glowinski, Y.A. Kuznetzov, G. Meurant, J. Periaux and O.B. Widlund, Eds., *Fourth Inter. Symp. on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Phil. PA, 1991, pp. 263–289.
28. R. Glowinski, G.H. Golub, G.A. Meurant, and J. Periaux, Eds., *First Inter. Symp. on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Phil. PA, 1988, pp. 144–172.
29. S.K. Godunov, V.V. Ogneva and G.P. Prokopov, *On the convergence of the modified steepest descent method in application to eigenvalue problems*, Am. Math. Soc. Trans. **2** (1976), 105..
30. V.P. Il'in, *Numerical Methods of Solving Electrophysical Problems*, Nauka, Moscow, 1985, In Russian..
31. T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1976.

32. A.V. Knyazev, *Convergence rate estimates for iterative methods for a mesh symmetric eigenvalue problem.*, Sov. J. Num. Anal. Math. Modeling, **2** (1987), 371–396.
33. A.V. Knyazev, *Computation of Eigenvalues and Eigenvectors for Mesh Problems: The Algorithms and Error Estimates* (1986), Dept. Numer. Math., USSR Acad. Sci., Moscow, USSR, In Russian..
34. A.V. Knyazev, *New estimates for Ritz vectors*, CIMS NYU, New York (to appear, Math. Comp.) **677** (1994).
35. A. V. Knyazev, *A preconditioned conjugate gradient method for eigenvalue problems and its implementation in a subspace*, International Ser. Numerical Mathematics, v. 96, Eigenwertaufgaben in Natur- und Ingenieurwissenschaften und ihre numerische Behandlung, Oberwolfach, 1990. (1991), Birkhauser, Basel, 143–154.
36. A. V. Knyazev and A. L. Skorokhodov, *The preconditioned gradient-type iterative methods in a subspace for partial generalized symmetric eigenvalue problem*, SIAM J. Numerical Analysis **31** (1994), 1226.
37. N. Kosugi, *Modification of the Liu-Davidson method for obtaning one or simultaneously several eigensolutions of a large real symmetric The preconditioned gradient-type iterative methods in a subspace for partial generalized symmetric eigenvalue problem*, J. Comp. Phys. **55** (1984), 426–436.
38. D. E. Longsine and S. F. McCormick, *Simultaneous Raileigh-quotient minimization methods for* $Ax = \lambda Bx$., Linear Algebra and Applications **34** (1980), 195–234.
39. S. F. McCormick and T. Noe, *Simultaneous iteration for the matrix eigenvalue problem*, Linear Algebra and Applications **16** (1977), 43–56.
40. R. B. Morgan, *Davidson's method and preconditioning for generalized eigenvalue problems*, J. Comp. Phys. **89** (1990), 241–245.
41. R. B. Morgan and D. S. Scott, *Generalizations of Davidson's method for computing eigenvalues of sparse symmetric matrices*, SIAM J. Sci. Statist. Comput. **7** (1986), 817–825.
42. R. B. Morgan and D. S. Scott, *Preconditioning the Lanczos algorithm for of sparse symmetric eigenvalue problems*, SIAM J. Sci. Comput. **14** (1993), 585–593.
43. C. W. Murray, S. C. Racine and E. R.Davidson, *Improved algorithms for the lowest few eigenvalues and associated eigenvectors of large matrices*, J. Comp. Phys. **103** (1992), 382–389.
44. B.N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1980.
45. W.V. Petryshyn, *On the eigenvalue problem* $Tu - \lambda Su = 0$ *with unbounded and non-symmetric operators* T *and* S, Philos. Trans. R. Soc. Math. Phys. Sci. **262** (1968), 413–458.
46. Axel Ruhe, *SOR-methods for the eigenvalue problem with large sparse matrices*, Math. Comput. **28** (1974), 695–710.
47. Axel Ruhe, *Iterative eigenvalue algorithm based on convergent splittings*, J. Comp. Phys. **19** (1975), 110–120.
48. Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Halsted Press, New-York, 1992.
49. M. Sadkane, *Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems*, Numerische Mathematik **64** (1993), 195–211.
50. B. A. Samokish, *The steepest descent method for an eigenvalue problem with semi-bounded operators*, Izvestiya Vuzov, Math. **5** (1958), 105–114, In Russian..
51. D. S. Scott, *Solving sparse symmetric generalized eigenvalue problems without factorization*, SIAM J. Numer. Analysis **18** (1981), 102–110.
52. P. Vassilevski, *Hybrid V-cycle algebraic multilevel preconditioners* (1987), (preprint) Bulgarian Academy Sciences, Sofia Bulgaria.