Electronic Transactions on Numerical Analysis.
Volume 7, 1998, pp. 104-123.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

PRECONDITIONED EIGENSOLVERS—AN OXYMORON?*

ANDREW V. KNYAZEV†

Abstract. A short survey of some results on preconditioned iterative methods for symmetric eigenvalue problems is presented. The survey is by no means complete and reflects the author's personal interests and biases, with emphasis on author's own contributions. The author surveys most of the important theoretical results and ideas which have appeared in the Soviet literature, adding references to work published in the western literature mainly to preserve the integrity of the topic. The aim of this paper is to introduce a systematic classification of preconditioned eigensolvers, separating the choice of a preconditioner from the choice of an iterative method. A formal definition of a preconditioned eigensolver is given. Recent developments in the area are mainly ignored, in particular, on Davidson's method. Domain decomposition methods for eigenproblems are included in the framework of preconditioned eigensolvers.

Key words. eigenvalue, eigenvector, iterative methods, preconditioner, eigensolver, conjugate gradient, Davidson's method, domain decomposition.

AMS subject classifications. 65F15, 65F50, 65N25.

1. Preconditioners and Eigenvalue Problems.

In recent decades, the study of preconditioners for iterative methods for solving large linear systems of equations, arising from discretizations of stationary boundary value problems of mathematical physics, has become a major focus of numerical analysts and engineers. In each iteration step of such methods, a linear system with a special matrix, the preconditioner, has to be solved. The given system matrix can be available only in terms of a matrix-vector multiplication routine. It is known that the preconditioner should approximate the matrix of the original system well in order to obtain rapid convergence. For finite element/difference problems, it is desirable that the rate of convergence is independent of the mesh size.

Preconditioned iterative methods for eigenvalue computations are relatively less known and developed. Generalized eigenvalue problems are particularly difficult to solve.

Classical methods, such as the QR algorithm, the conjugate gradient method without preconditioning, Lanczos method, and inverse iterations are some of the most commonly used methods for solving large eigenproblems.

In quantum chemistry, however, Davidson's method, which can be viewed as a preconditioned eigenvalue solver, has become a common procedure for computing eigenpairs; e.g., [15, 16, 32, 47, 12, 81, 84]. Davidson-like methods have become quite popular recently; e.g., [12, 73, 61, 62, 74]. New results can be found in other papers of this special issue of ETNA.

Although theory can play a considerable role in development of efficient preconditioned methods for eigenproblems, a general theoretical framework is still to be developed. We note that asymptotic-style convergence rate estimates, traditional in numerical linear algebra, do not allow us to conclude much about the dependence of the convergence rate on the mesh parameter when a mesh eigenvalue problem needs to be solved. Estimates, to be useful, must take the form of inequalities with explicit constants. The *low-dimensional* techniques, developed in [37, 35, 42, 44], can be a powerful theoretical tool for obtaining estimates of this kind for symmetric eigenvalue problems.

Preliminary theoretical results are very promising. Sharp convergence estimates have been established for some preconditioned eigensolvers; e.g., [27, 21, 19, 20, 22, 35, 36]. The

*Received February 9, 1998. Accepted for publication August 24, 1998. Recommended by R. Lehoucq. This work was supported by the National Science Foundation under Grant NSF–DMS-9501507.
†Department of Mathematics, University of Colorado at Denver, P.O. Box 173364, Campus Box 170, Denver, CO 80217-3364. E-mail: aknyazev@math.cudenver.edu WWW URL: http://www-math.cudenver.edu/~aknyazev.

estimates show, in particular, that the convergence rates of preconditioned iterative methods, with an appropriate choice of preconditioner, are independent of the mesh parameter for mesh eigenvalue problems. Thus, preconditioned iterative methods for symmetric mesh eigenproblems can be approximately as efficient as the analogous solvers of symmetric linear systems of algebraic equations. They can serve as a basis for developing effective codes for many scientific and engineering applications in structural dynamics and buckling, ocean modeling, quantum chemistry, magneto-hydrodynamics, etc., which should make it possible to carry out simulations in three dimensions with high resolution. A much higher resolution than what can be achieved presently is required in many applications, e.g., in ocean modeling and quantum chemistry.

Domain decomposition methods for eigenproblems can also be analyzed in the context of existing theory for the stationary case; there are only a few results for eigenproblems, see [48, 4, 5, 11]. Recent results for non-overlapping domain decomposition [43, 39, 45, 46] demonstrate that an eigenpair can be found at approximately the same cost as a solution of the corresponding linear systems of equations.

In the rest of the paper, we first review some well-known facts for preconditioned solvers for systems. Then, we introduce preconditioned eigensolvers, separating the choice of a preconditioner from the choice of an iterative method, and give a formal definition of a preconditioned eigensolver. We present several ideas that could be used to derive formulas for preconditioned eigensolvers and show that different ideas lead to the same formulas. We survey some results, which have appeared mostly in Soviet literature. We discuss preconditioned block algorithms and domain decomposition methods for eigenproblems. We conclude with numerical results and some comments on our bibliography.

2. Preconditioned Iterative Methods for Linear Systems of Equations. We consider a linear algebraic system $Lu = f$ with a *real symmetric positive definite* matrix L. Since a direct solution of such linear systems often requires considerable computational work, iterative methods of the form

$$u^{k+1} = u^k - \gamma_k B^{-1}(Lu^k - f) \tag{2.1}$$

with a *real symmetric positive definite matrix* B, are of key importance. The matrix B is usually referred to as the preconditioner. It is common practice to use conjugate gradient type methods as accelerators in such iterations.

Preconditioned iterative methods of this kind can be very effective for the solution of systems of equations arising from discretization of elliptic operators. For an appropriate choice of the preconditioner B, the convergence does not slow down when the mesh is refined, and each iteration has a small cost.

The importance of choosing the preconditioner B so that the condition number of $B^{-1}L$ either is independent of N, the size of the system, see D'yakonov [22], or depends weakly, e.g., polylogarithmically on N, is widely recognized. At the same time, the numerical solution of the system with the matrix B should ideally require of the order of N, or $N \ln N$, arithmetic operations for single processor computers.

It is important to realize that *the preconditioned method (2.1) is mathematically equivalent to the analogous iterative method without preconditioning, applied to the preconditioned system* $B^{-1}(Lu - f) = 0$. We shall analyze such approach for eigenproblems in the next section.

In preparation for our description of preconditioned eigenvalue algorithms based on domain decomposition, we introduce some simple ideas and notation. For an overview of work on domain decomposition methods, see, e.g., [57].

Domain decomposition methods can be separated into two categories, with and without overlap of sub-domains. Methods with overlap of subdomains can usually be written in a form similar to (2.1), and the domain decomposition approach is used only to construct the preconditioner B. The class of methods without overlap can be formulated in algebraic form as follows.

Let L be a 2×2 block matrix, and consider the linear system

$$\begin{pmatrix} L_1 & L_{12} \\ L_{21} & L_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}. \tag{2.2}$$

Here, u_1 corresponds to unknowns in sub-domains, and u_2 corresponds to mesh points on the interface between sub-domains. We assume that the submatrix L_1 is "easily invertible." In domain decomposition, solving systems with coefficient matrix L_1 means solving corresponding boundary value problems on sub-domains. Then, by eliminating the unknowns u_1, we obtain a linear algebraic system of "small" size:

$$S_L u_2 = g_2, \; g_2 = f_2 - L_{21} L_1^{-1} f_1,$$

where

$$S_L = L_2 - L_{21} L_1^{-1} L_{12} = S_L^* > 0$$

is the *Schur complement* of the block L_1 in the matrix L.

Iterative methods of the form

$$u_2^{k+1} = u_2^k - \gamma_k S_B^{-1} (S_L u_2^k - g_2) \tag{2.3}$$

with some preconditioner $S_B = S_B^* > 0$ can be used for the numerical solution of the linear algebraic system. Conjugate gradient methods are often employed instead of the simple iterative scheme (2.3). The Schur complement S_L does not have to be computed explicitly in such iterative methods. With an appropriate choice of S_B, the convergence rate does not deteriorate when the mesh is refined, and each iteration has a minimal cost. The whole computational procedure can therefore be very effective.

3. Preconditioned Symmetric Eigenproblems. We now consider the partial symmetric eigenvalue problem

$$Lu = \nu u,$$

where the matrix L is symmetric positive definite, $L = L^* > 0$. Typically a couple of the smallest eigenvalues ν and corresponding eigenvectors u need to be found. Very often, e.g., when the Finite Element Method (FEM) is applied to an eigenproblem with differential operators, we have to solve resulting generalized symmetric eigenvalue problems

$$Lu = \nu M u,$$

for matrices $L = L^* > 0$ and $M = M^* > 0$. In some applications, e.g., in buckling, the matrix M may only be nonnegative, or possibly indefinite. To avoid problems with infinite eigenvalues ν, we define $\lambda = 1/\nu$ and consider the following generalized symmetric eigenvalue problem

$$Mu = \lambda L u, \tag{3.1}$$

for matrices $M = M^*$ and $L = L^* > 0$ with eigenvalues $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_{\min}$ and corresponding eigenvectors. In this problem, one usually wants to find the top part of the spectrum.

In the present paper, we will consider the problem of finding the largest eigenvalue λ_1 of (3.1), which we assume to be simple, and a corresponding eigenvector u_1. We will also mention block methods for finding a group of the first p eigenvalues and the corresponding eigenvectors, and discuss orthogonalization to previously computed eigenvectors.

In structural mechanics, the matrix L is usually referred to as *stiffness* matrix, and M as *mass* matrix. In the eigenproblem (3.1), M is still a mass matrix and L is a stiffness matrix, in spite of the fact that we put an eigenvalue on an unusual side. An important difference between the mass matrix M and the stiffness matrix L is that M is usually better conditioned than L. For example, when a simplest FEM with the mesh parameter h is applied to the classical problem of finding the main frequencies of a homogeneous membrane, which mathematically is an eigenvalue problem for the Laplace operator, then M is a FEM approximation of the identity with $\mathrm{cond} M$ bounded by a constant uniformly in h, and L corresponds to the Laplace operator with $\mathrm{cond} L$ growing as h^{-2} when $h \to 0$.

When problem (3.1) is a finite difference/element approximation of a differential eigenvalue problem, we would like to choose M and L in such a way that the finite dimensional operator $L^{-1}M$ approximates a compact operator. In structural mechanics, such choice corresponds to L being a stiffness matrix and M being a mass matrix, not *vice versa*. Then, eigenvalues λ_i of problem (3.1) tend to the corresponding eigenvalues of the continuous problem.

The ratio

$$\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_{\min}}$$

plays a major role in convergence rate estimates for iterative methods to compute λ_1, the largest eigenvalue. When this ratio is small, the convergence may be slow. It follows from our discussion above that the denominator should not increase when the mesh is refined, for mesh eigenvalue problems.

We now turn our attention to preconditioning. Let $B = B^*$ be a symmetric preconditioner. There is no consensus on whether one should use a symmetric positive definite preconditioner only, or an indefinite one, for symmetric eigenproblems. A practical comparison of symmetric eigenvalue solvers with positive definite vs. indefinite preconditioners has not yet been described in the literature.

When we apply the preconditioner to our original problem (3.1), we get

$$B^{-1}Mu = \lambda B^{-1}Lu, \tag{3.2}$$

We do not suggest applying such preconditioning explicitly. We use (3.2) to introduce preconditioned eigensolvers, similarly to that for linear system solvers.

If the preconditioner B is positive definite, we can introduce a new scalar product $(\star, \star)_B = (B\star, \star)$. In this scalar product, the matrices $B^{-1}M$ and $B^{-1}L$ in (3.2) are symmetric, and $B^{-1}L$ is positive definite. Thus, the preconditioned problem (3.2) belongs to the same class of generalized symmetric eigenproblems as our original problem (3.1). The positive definiteness of the preconditioner allows us to use iterative methods for symmetric problems, like preconditioned conjugate gradient, or Lanczos-based methods.

If the preconditioner B is not positive definite, the preconditioned eigenvalue problem (3.2) is no longer symmetric, even if the preconditioner is symmetric, For that reason, it is difficult to develop a convergence theory of eigensolvers with indefinite preconditioners, comparable with that developed for the case of positive definite preconditioners. Furthermore, iterative methods for nonsymmetric problems, e.g., based on minimization of the residual, should be employed when the preconditioner is indefinite, thus increasing computational costs. We note, however, that multigrid methods for eigenproblems, see a recent paper [10] and references there, usually use indefinite preconditioners, often implicitly, and provide tools for constructing high-quality preconditioners.

What is a *perfect* preconditioner for an eigenvalue problem? Let us assume that we already know an eigenvalue, which we assume to be simple, and want to compute the corresponding eigenvector as a nontrivial solution of the following homogeneous system of linear equations:

$$(M - \lambda L)u = 0. \tag{3.3}$$

In this case, the best preconditioner would be the one based on the pseudoinverse of $M - \lambda L$:

$$B^{-1} = (M - \lambda L)^{\dagger}.$$

Indeed, with B^{-1} defined in this fashion, the preconditioned matrix $B^{-1}(M - \lambda L)$ is an orthogonal projector on an invariant subspace, which is orthogonal to the desired eigenvector. Then, one iteration of, e.g., the Richardson method

$$u^{k+1} = w^k + \tau^k u^u, \; w^k = B^{-1}(M - \lambda L)u^k \tag{3.4}$$

with the optimal choice of the step $\tau^k = -1$ gives the exact solution.

To compute the pseudoinverse we need to know the eigenvalue and the corresponding eigenvector, which makes this choice of the preconditioner unrealistic. Attempts were made to use an approximate eigenvalue and an approximate eigenvector, and to replace the pseudoinverse with the approximate inverse. Unfortunately, this leads to the preconditioner, which is not definite, even if λ is the extreme eigenvalue as it is typically approximated from the inside of the spectrum. The only potential advantage of using this indefinite preconditioner for symmetric eigenproblems is a better handling of the case where λ lies in a cluster of unwanted eigenvalues, assuming a high quality preconditioner can be constructed.

If we require the preconditioner to be symmetric positive definite, a natural choice for B is to approximate the stiffness matrix,

$$B \approx L.$$

In many engineering applications, preconditioned iterative solvers for linear systems $Lu = f$ are already available, and efficient preconditioners $B \approx L$ are constructed. In such cases, the same preconditioner can be used to solve an eigenvalue problem $Mu = \lambda Lu$; moreover, a slight modification of the existing codes for the solution of the system $Lu = f$ can be used to solve the partial eigenvalue problem with L.

Closeness of B and L is typically understood up to scaling and is characterized by the ratio $\delta = \delta_0/\delta_1$, where δ_0 and δ_1 are constants from the operator inequalities

$$\delta_0 B \leq L \leq \delta_1 B, \; \delta_0 > 0, \tag{3.5}$$

which are defined using associated quadratic forms. We note that $1/\delta$ is the (spectral) condition number of $B^{-1}L$ with respect to the operator norm induced by the vector norm, corresponding to the B-based scalar product $(\star, \star)_B$.

Under assumption (3.5), the following simple estimates of eigenvalues μ_i of the operator $B^{-1}(M - \lambda L)$, where λ is a fixed scalar, placed between two eigenvalues of eigenproblem (3.1), $\lambda_p \geq \lambda > \lambda_{p+1}$, hold [34, 35, 36]:

$$0 \leq \delta_0(\lambda_i - \lambda) \leq \mu_i \leq \delta_1(\lambda_i - \lambda),\ i = 1, \dots, p;$$

$$-\delta_1(\lambda - \lambda_j) \leq \mu_j \leq -\delta_0(\lambda - \lambda_j) < 0,\ j > p.$$

These estimates show that the ratio $\delta = \delta_0/\delta_1$ does indeed measure the quality of a positive definite preconditioner B applied to system (3.3) as it controls the spread of the spectrum of $B^{-1}(M - \lambda L)$ as well as the gap between positive and negative eigenvalues μ.

In the rest of the paper the preconditioner is always assumed to be a positive definite matrix when approximates L in the sense of (3.5). We want to emphasize that although preconditioned eigensolvers considered here are guaranteed to converge with any positive definite preconditioner B, e.g., with the trivial $B = I$, convergence may be slow.

In some publications, e.g., in the original paper by Davidson [15], a preconditioner is essentially built-in into the iterative method, and it takes some effort to separate them. However, such a separation seems to be always possible and desirable. Thus, when comparing different eigensolvers, the same preconditioners must be used for sake of fairness. The choice of the preconditioner is distinct from a choice of the iterative solver. In the present paper, we are not concerned with the problem of constructing a good preconditioner; we concentrate on iterative solvers instead.

At a recent conference on linear algebra, it was suggested that preconditioning, as we describe it in this section, is senseless, when applied to an eigenproblem, because it does not change eigenvalues and eigenvectors of the original eigenproblem. In the next section, we shall see that some eigenvalue solvers are, indeed, invariant with respect to preconditioning, but we shall also show that some other eigenvalue solvers can take advantage of preconditioning. The latter eigensolvers are the subject of the present paper.

4. Preconditioned Eigensolvers - A Simple Example. We now address the question why some iterative methods can be called *preconditioned* methods. Let us consider the following two iterative methods for our original problem (3.1):

$$u^{k+1} = w^k + \tau^k u^k,\ w^k = L^{-1}(Mu^k - \alpha^k Lu^k),\ u^0 \neq 0, \tag{4.1}$$

and

$$u^{k+1} = w^k + \tau^k u^k,\ w^k = (Mu^k - \alpha^k Lu^k),\ u^0 \neq 0, \tag{4.2}$$

where scalars τ^k and α^k are iteration parameters. We ignore for a moment that method (4.1) does not belong to the class of methods under consideration in the present paper since it requires the inverse of L.

Though quite similar, the two methods exhibit completely different behavior when applied to the *preconditioned* eigenproblem (3.2). Namely, method (4.1) is not changed, but method (4.2) becomes the *preconditioned* one -

$$u^{k+1} = w^k + \tau^k u^k,\ w^k = B^{-1}(Mu^k - \alpha^k Lu^k),\ u^0 \neq 0. \tag{4.3}$$

We note that the eigenvectors of the iteration matrix $(M - \alpha^k L)$ appearing in (4.2) are not necessarily the same as those of problem (3.1). This is the main difficulty in the theory of methods such as (4.2), but this is also the reason why using the preconditioning makes the difference and gives hope to achieve better convergence of (4.3), as eigenvalues and eigenvectors of the matrix $B^{-1}(M - \alpha^k L)$ do actually change when we apply different preconditioners.

Let us examine closer method (4.3), which is a generalization of Richardson method (3.4). It can be used for finding λ_1 and the corresponding eigenvector u_1. Iteration parameters α^k can be chosen as a function of u^k such that $\alpha^k \to \lambda_1$, as $u^k \to u_1$. A common choice is the Rayleigh quotient for problem (3.1):

$$\alpha^k = \lambda(u^k) = (Mu^k, u^k)/(Lu^k, u^k),$$

Then, vector w^k is collinear to the gradient of the Rayleigh quotient at the point u^k in the scalar product $(\star, \star)_B = (B\star, \star)$, and methods such as (4.3) are called *gradient methods*. We will call the methods given by (4.3), for which α^k is not necessarily equal to $\lambda(u^k)$, *gradient-type methods*. It is useful to note, that the formula for the Rayleigh quotient above is invariant with respect to preconditioning, provided that the scalar product is changed, too. Iteration parameters τ^k can be chosen to maximize $\lambda(u^{k+1})$, thus, providing *steepest ascent*. Equivalently, u^{k+1} can be found by the Rayleigh–Ritz method in the trial subspace $\mathrm{span}\{u^k, w^k\}$. This steepest gradient ascent method for maximizing the Rayleigh quotient is a particular case of a general steepest gradient ascent method for maximizing a nonquadratic function. Interestingly, in our case the choice of τ^k is not limited to positive values, as a general theory of optimization would suggest. Examples were found in [38] when τ^k may be zero, or negative.

Method (4.3) is the easiest example of a preconditioner eigensolver. In the next section we attempt to describe the whole class of preconditioned eigensolvers, with a fixed preconditioner.

5. Preconditioned Eigensolvers: Definition and Ideas. We define a preconditioned iterative method for eigenvalue problem (3.1) as a polynomial method of the following kind,

$$u^n = P_{m_n}(B^{-1}M, B^{-1}L)u^0, \tag{5.1}$$

where P_{m_n} is a polynomial of the m_n-th degree of two independent variables, and B is a preconditioner.

Our simple example (4.3) fits the definition with $m_n = n$ and P_{m_n} being a product of monomials

$$B^{-1}M - \alpha^k B^{-1}L + \tau^k I.$$

Clearly, we get the best convergence on this class of methods, when we simply take u^n to be the Rayleigh–Ritz approximation on the generalized Krylov subspace corresponding to all polynomials of the degree not larger than n. Unfortunately, computing a basis of this Krylov subspace, even for a moderate value of n, is very expensive. An orthogonal basis of polynomials could help to reduce the cost, as in the standard Lanczos method, but very little is known on operator polynomials of two variables, cf. [58]. However, it will be shown in the next section, that there are some simple polynomials which provide fast convergence with small cost of every iteration.

Preconditioned iterative methods, satisfying our definition, could be constructed in many different ways. One of the most traditional ideas is to implement the classical Rayleigh quotient iterative method

$$u^{k+1} = (M - \alpha^k L)^{-1}Lu^k, \; \alpha^k = \lambda(u^k),$$

with a preconditioned iterative solver of the systems that appear on every (outer) iteration, e.g., [83]. A similar inner-outer iteration method could be based on more advanced truncated rational Krylov method, e.g., [69, 72], or on Newton's method for maximizing/minimizing the Rayleigh quotient. A homotopy method, e.g., [14, 50, 31, 86, 53], is quite analogous to the previous two methods, if a preconditioned iterative solver is used for inner iterations. When employing an inner-outer iterative method, a natural question is how many inner iterations should be performed. Our simple example (4.3) of a preconditioned eigensolver given in the previous section can be viewed as an inner/outer iterative method with only one inner iteration. In the next section, we shall consider another approach, based on an auxiliary eigenproblem, which leads to an inner/outer iterative method, and discuss the optimal number of inner iterations.

Let us also recall two other ideas of constructing preconditioned eigensolvers, as used in the previous section. Firstly, we can pretend that the eigenvalue λ is known, take a preconditioned iterative solver for the homogeneous system (3.3), and just change the value of λ on every iteration. Secondly, we can use general preconditioned optimization methods, like the steepest ascent, or the conjugate gradient method, to maximize the Rayleigh quotient. In the latter approach, we do not have to avoid local optimization problems, like the problem of finding the step τ^k in the steepest ascent method (4.3), which usually cause trouble for general nonquadratic functions in optimization, since for our function, the Rayleigh quotient, such problems can be solved easily and cheaply by the Rayleigh–Ritz method.

As far as methods fall within the same class of preconditioned eigensolvers, their origination does not matter much. They should compete against each other, and, first of all, against old well-known methods, which we discuss in the next section. With so many ideas available, it is relatively simple to design formally new methods. It is, on the other hand, difficult to design better methods and methods with better convergence estimates.

6. Theory for Preconditioned Eigensolvers. Gradient methods with a preconditioner were first considered for symmetric operator eigenvalue problems in the form of steepest ascent/descent methods by B. A. Samokish [75], who also derived asymptotic convergence rate estimates. W. V. Petryshyn [65] considered these methods for some nonsymmetric operators, using symmetrization. A. Ruhe [70, 71] clarified the connection of gradient methods (4.3) and similar iterative methods for finding a nontrivial solution of the following preconditioned system

$$B^{-1}(M - \lambda_1 L)u = 0. \tag{6.1}$$

S. K. Godunov *et al.* [27] obtained the first non-asymptotic convergence rate estimates for preconditioned gradient iterative methods (3.2); however, to prove linear convergence they assumed that

$$\delta_0^2 B^2 \leq L^2 \leq \delta_1^2 B^2,$$

which is more restrictive than (3.5). E. G. D'yakonov *et al.* [21, 18, 19] obtained the first explicit estimates of linear convergence for gradient iterative methods (3.2), including the steepest descent method, using the natural assumption (3.5). In our notations, and somewhat simplified, the main convergence rate estimate of [21, 18, 22] for method (4.3) with

$$\alpha^k = \lambda(u^k) \text{ and } \tau^k = \delta_1(\lambda^k - \lambda_{\min})$$

can be written as

$$\frac{\lambda_1 - \lambda^n}{\lambda^n - \lambda_2} \leq (1 - \xi)^n \frac{\lambda_1 - \lambda^0}{\lambda^0 - \lambda_2}, \; \xi = \frac{\delta_0}{\delta_1} \frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_{\min}}, \tag{6.2}$$

under the assumption that $\lambda^0 > \lambda_2$. The same estimate holds when τ^k is chosen to maximize the Rayleigh quotient $\lambda(u^{k+1})$, i.e. for the steepest ascent.

How sharp is estimate (6.2)? Asymptotically, when $\lambda^k \approx \lambda_1$, the convergence rate estimate of [75] is better than estimate (6.2). When $\delta \approx 1$, we have $B \approx L$ and method (4.3) becomes a standard power method with a shift. The convergence estimate of this method; e.g., [35, 36], is better than estimate (6.2) with $\delta \approx 1$. Nonasymptotically and for small/moderate values of δ it is not clear whether estimate (6.2) is sharp, or not, but it is the best known estimate.

If instead of $\lambda^0 > \lambda_2$ more general condition $\lambda_p \geq \lambda^0 > \lambda_{p+1}$ with some $p > 1$ holds, and λ^n is also between p-th and $p+1$-th eigenvalues, then estimate (6.2) is still valid [22] when we replace λ_1 and λ_2 with λ_p and λ_{p+1}, correspondingly. In numerical experiments, method (4.3) usually converges to λ_1 with a random initial guess. When $\lambda_p \geq \lambda^0 > \lambda_{p+1}$, the sequence u^k needs to pass p saddle points to reach to u_1 and can get stuck in the middle, in principle. For a general preconditioner B, there is no theory to predict whether this can happen for a given initial guess u^0.

E. G. D'yakonov played a major role in establishing preconditioned eigensolvers as asymptotically optimal methods for discrete analogs of eigenvalue problems with elliptic operators. Many of his results on the subject are collected in his recent book [22].

V. G. Prikazchikov *et al*. derived somewhat weaker convergence estimates for the steepest ascent/descent only, see [67, 66, 87, 68] and references there. The possibility of using Chebyshev parameters to accelerate convergence of (4.3) has been discussed by V. P. Il'in and A. V. Gavrilin, see [33]. A variant, combining (4.3) with the Lanczos method, is due to David Scott [77]; here $B = I$. The key idea for both papers is the same. We use it here to discuss the issue of the optimal number of inner iterations. Let the parameter λ be fixed. Consider the auxiliary eigenvalue problem

$$B^{-1}(M - \lambda L)v = \mu v. \tag{6.3}$$

If $\lambda = \lambda_1$, then there exist a zero eigenvalue μ and the corresponding eigenvector of (6.3) is also the eigenvector of the original problem (3.1) corresponding to λ_1. The eigenproblem (6.3) can be solved for a fixed $\lambda = \lambda^k$ by using inner iterations, e.g., the power method with Chebyshev acceleration, see [33], or the Lanczos method, see [77]. The new value $\lambda = \lambda^{k+1}$ is then calculated as the Rayleigh quotient of the most recent vector iterate of the inner iteration. This vector also serves as an initial guess for the next inner iteration cycle. Such inner-outer iterative method falls within our definition of preconditioned eigensolvers. If only one inner iteration is used, the method in identical to (4.3). Thus, inner-outer iteration methods can be considered as a generalization of method (4.3). One question which arises at this point is whether it is beneficial to use more than one inner iterations. We discuss below a theory, developed by the author in [34, 35, 36], that gives a somewhat positive answer to the question.

An asymptotic quadratic convergence of the outer iterations with an *infinite* number of inner iterations was proved by Scott [77]. The same result was then proved in [34, 35, 36] in the form of the following explicit estimate

$$\lambda_1 - \lambda^{k+1} \leq \left\{ 1 + \frac{4\delta}{(1-\delta)^2} \frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda^k} \right\}^{-1} (\lambda_1 - \lambda^k),\ \delta = \frac{\delta_0}{\delta_1},$$

under the assumption $\lambda^k > \lambda_2$. This estimate was used to study the case of a finite number of inner iterations.

An explicit convergence rate estimate, established in [34, 35, 36], is similar to (6.2) and shows the following: 1) *the method converges geometrically for any fixed number of inner* *iterations*; 2) *a slow, but unlimited, increase of the number of inner iterations during the process improves the convergence rate estimate, averaged with regard to the number of inner iterations.*

We do not reproduced here the estimate for a general case of an inner-outer iterative solver as it is too cumbersome. When applied to method (4.3) with $\alpha^k = \lambda(u^k)$ and $\tau^k = \delta_1(\lambda^k - \lambda_{\min})$, the estimate becomes

$$\frac{\lambda_1 - \lambda^{k+1}}{\lambda^{k+1} - \lambda_2} \leq \left(1 - (1-\sigma)\max\{\delta, \rho^k\}\right) \frac{\lambda_1 - \lambda^k}{\lambda^k - \lambda_2}, \tag{6.4}$$

where

$$\begin{aligned} \rho^k = \quad & \left(1 + \left(\frac{1}{\delta} - 1\right)\sigma\epsilon^k\right)^{-1} \left(1 + \frac{1}{4}\frac{(1-\delta)^2}{\delta}\epsilon^k\right)^{-1} \times \\ & \left(1 + \left(\frac{1}{\delta} - 1\right)\sqrt{\frac{\sigma\epsilon^k}{\delta + (\sigma - \delta)\epsilon^k}}\right)^{-1}, \end{aligned} \tag{6.5}$$

and

$$\sigma = (1-\xi)^2,\ \xi = \frac{\delta_0}{\delta_1}\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_{\min}},\ \epsilon^k = \frac{\lambda_1 - \lambda^k}{\lambda_1 - \lambda_2}.$$

Estimate (6.4) is sharp for sufficiently large δ, or small initial error $\lambda_1 - \lambda^0$, in which case it improves estimate (6.2). However, when δ is small, the estimate (6.4) is much worse than (6.2).

The general estimate of [34, 35, 36] also holds for the following method,

$$\begin{gathered} \lambda(u^{n+1}) = \max\nolimits_{u \in \mathcal{K}} \lambda(u), \\ \mathcal{K} = \operatorname{span}\{u^n, B^{-1}(M - \lambda^n L)u^n, \ldots, (B^{-1}(M - \lambda^n L))^{k_n} u^n\}, \end{gathered} \tag{6.6}$$

where k_n is the number of inner iteration steps for the n-th outer iteration step. Method (6.6) was suggested in [34, 35, 36], and was then rediscovered in [63].

Our interpretation of the well known Davidson method [15, 73, 81] has almost the same form:

$$\begin{gathered} \lambda(u^{n+1}) = \max\nolimits_{u \in \mathcal{D}} \lambda(u), \\ \mathcal{D} = \operatorname{span}\{u^0, B^{-1}(M - \lambda^0 L)u^0, \ldots, B^{-1}(M - \lambda^n L)u^n\}. \end{gathered} \tag{6.7}$$

The convergence properties of this method are still not well understood. A clear disadvantage of the method in its original form (6.7) is that the dimension of the subspace grows with every step.

Our favorite subclass of preconditioned eigensolvers is preconditioned conjugate gradient (CG) methods, e.g., [13] with $B = I$. A simplest variant of a preconditioned CG method can be written as

$$u^{k+1} = w^k + \tau^k u^k + \gamma^k u^{k-1},\ w^k = B^{-1}(Mu^k - \alpha^k Lu^k),\ \alpha^k = \lambda(u^k), \tag{6.8}$$

with properly chosen scalar iteration parameters τ^k and γ^k. The easiest choice of parameters is based on an idea of local optimality, e.g., [39], namely, we simply choose τ^k and γ^k to maximize the Rayleigh quotient of u^{k+1} by using the Rayleigh–Ritz method. We give a block version of the method in the next section.

For the locally optimal version of the preconditioned CG method, we can trivially apply convergence rate estimates (6.2) and (6.4).

ETNA
Kent State University
etna@mcs.kent.edu

It is interesting to compare theoretical properties of CG type methods for linear systems vs. eigenvalue problems.

A central fact of the theory of variational iterative methods for linear systems, with symmetric and positive definite matrices, is that the global optimum can be achieved using a local three-term recursion. In particular, the well-known locally optimal variant of the CG method, in which both parameters in the three-term recursion are chosen to minimize locally the energy norm of the error, leads to the same approximations as the global optimization if the same initial guess is used; e.g., [28, 13]. Analogous methods for symmetric eigenvalue problems are based on minimization (or, in general, on finding stationary values) of the Rayleigh quotient. The Lanczos (global optimization) and the CG (local optimization) methods are no longer equivalent for eigenproblems because the Rayleigh quotient is not quadratic and does not even have a positive definite Hessian.

Numerous papers on CG methods for eigenproblems attempt to derive sharp convergence estimates; see: [76, 26, 3]; see also [25, 24, 85, 82].

So far, we have only discussed the problem of finding an eigenvector corresponding to an extreme eigenvalue. To find the p-th largest eigenvalue, where p is not too big, a block method, which we consider in the next section, or orthogonalization to previously computed eigenvectors, can be used.

Let us consider the orthogonalization using, as a simple example, a preconditioned eigensolver (4.3) with an additional orthogonal projection onto an orthogonal complement of the subspace spanned by the computed eigenvectors:

$$u^{k+1} = P^{\perp} w^k + \tau^k u^k, \; w^k = B^{-1}(Mu^k - \alpha^k Lu^k), \; u^0 \neq 0. \tag{6.9}$$

We now discuss the choice of scalar products, when defining the orthogonal projector $P^{\perp}$.

First, we need to choose a scalar product for the orthogonal complement of the subspace spanned by the computed eigenvectors. The L scalar product, $(\star,\star)_L = (L\star,\star)$, is a natural choice here. When M is positive definite, it is common to use the M scalar product as well. Second, we need to define a scalar product, with respect to which our projector $P^{\perp}$ is orthogonal. A traditional approach is to use the same scalar product as on the first step. Unfortunately, with such choices, the iteration operator in method (6.9) is no longer symmetric with respect to the B scalar product. This makes theoretical investigation of the influence of orthogonalization to approximately computed eigenvectors quite complicated; see [22, 20], where direct analysis of perturbations is included. To preserve symmetry, we must use a B-orthogonal projector $P^{\perp}$ in spite of the fact that we use a different scalar product on the first step to define the orthogonal complement. In this case we can use the standard and simple backward error analysis, see [35, 36], instead of the direct analysis, see [22, 20]. The actual computation of $P^{\perp}w$ for a given w, however, requires special attention. Let $\hat{W}$ be the subspace, spanned by approximate eigenvectors, and find a basis for the subspace $B^{-1}L\tilde{W}$. A B-orthogonal complement to the latter subspace coincides with the L-orthogonal complement of the subspace, $\tilde{W}$. Therefore we can use the standard B-orthogonal projector onto the B-orthogonal complement of $B^{-1}L\tilde{W}$.

We note that using a scalar product associated with an ill-conditioned matrix, like L, or B, may lead to unstable methods.

Another, simpler, but in some cases more expensive, possibility of locking the converged eigenvectors is to add them in the basis of the trial subspace of the Rayleigh–Ritz method.

7. Preconditioned Subspace Iterations. Block methods, or subspace iterations, or simultaneous iterations are well known methods for the simultaneous computation of several leading eigenvalues and corresponding invariant subspaces. Preconditioned iterative methods of that kind were developed in [75, 59, 52, 9], mostly with $B = I$. The simplest example is a block version of method (4.3):

$$\hat{u}_i^{k+1} = w_i^k + \tau_i^k u_i^k,\ w_i^k = B^{-1}(Mu_i^k - \alpha_i^k L u_i^k),\ i = 1, \ldots, p, \tag{7.1}$$

where u_i^{k+1} is then computed by the Rayleigh–Ritz method in the trial subspace

$$\mathrm{span}\{\hat{u}_1^{k+1}, \ldots, \hat{u}_p^{k+1}\}.$$

The iteration operator for (7.1) is complicated and nonlinear if parameters τ_i^k and α_i^k change with i, thus making it difficult to study its convergence. The only nonasymptotic explicit convergence rate estimate of method (7.1) with a natural choice $\alpha_i^k = \lambda(u_i^k)$, has been published only recently [8]. Sharp accuracy estimates of the Rayleigh-Ritz method, see [40], plays a crucial role in the theory of block methods.

Gradient-type methods (7.1) with $\alpha_i^k = \lambda(u_p^k)$ chosen to be independent of i, have been developed by D'yakonov and Knyazev in [19, 20]. Convergence rate estimates have been established only for λ_p. These gradient-type methods are more expensive than gradient methods with $\alpha_i^k = \lambda(u_i^k)$. This is because in the first phase only the smallest eigenvalue of the group, $\lambda_p, \lambda_p > \lambda_{p+1}$, and the corresponding eigenvector can be found; see [20, 8] for details.

The following are the block version of the steepest ascent method:

$$u_i^{k+1} \in \mathrm{span}\{u_1^k, \ldots, u_p^k, w_1^k, \ldots, w_p^k\} \tag{7.2}$$

and the block version of the conjugate gradient method:

$$u_i^{k+1} \in \mathrm{span}\{u_1^{k-1}, \ldots, u_p^{k-1}, u_1^k, \ldots, u_p^k, w_1^k, \ldots, w_p^k\}, \tag{7.3}$$

where

$$w_i^k = B^{-1}(Mu_i^k - \alpha_i^k L u_i^k),\ \alpha_i^k = \lambda(u_i^k),$$

and the Rayleigh–Ritz method is used to compute u_i^{k+1} in the corresponding trial subspace. Unfortunately, the previously mentioned theoretical results for rate of convergence cannot be directly applied to methods (7.2) and (7.3). We were not able to prove theoretically that our block CG method (7.3) is the most efficient preconditioned iterative solver for symmetric eigenproblems, but it was supported by preliminary numerical results, included at the end of the present paper.

Another known idea of constructing block methods is to use methods of nonlinear optimization to minimize, or maximize, the trace of the projection matrix in the Rayleigh–Ritz method, e.g., [13, 23], here $B = I$.

To compute an eigenvalue λ_p in the middle of the spectrum, with p large, using block methods and/or finding all previous eigenvalues and eigenvectors and applying orthogonalization are computationally very expensive. In this case, the Rayleigh quotient iteration with sufficiently accurate solutions of the associated linear systems, or methods similar to those described in [64], may be more effective.

In some applications, e.g., in buckling, both ends of the spectrum should be computed. Block methods are known [13] that compute simultaneously largest and smallest eigenvalues.

8. Domain Decomposition Eigensolvers.

Domain decomposition iterative methods for eigenproblems have a lot in common with the analogous methods for linear systems. Domain decomposition also provides an interesting and important application of the preconditioned eigensolvers discussed above.

For domain decomposition with overlap, we can use any of preconditioned eigensolvers and construct a preconditioner based on domain decomposition. In this case, at every step of the iterative solver when a preconditioner is applied, we need to solve linear systems on sub-domains, in parallel if an additive Schwarz method is employed to construct the preconditioner. We strongly recommend using a positive definite preconditioner, which approximates the stiffness matrix L. Such preconditioners are widely used in domain decomposition linear solvers, and usually satisfy assumption (3.5) with constants independent of the mesh size parameter. As long as the preconditioner is symmetric positive definite, all theoretical results for general preconditioned eigensolvers apply.

This simple approach appears to be much more efficient than methods based on eigenvalue solvers on sub-domains as inner iterations of a Schwarz-like iterative method; see [56], as the cost of solving an eigenproblem on a sub-domain is often about the same as the cost of solving the original eigenproblem.

Domain decomposition without overlap requires separate treatment. Let the matrices L and M be of two-by-two block form as in (2.2). The original eigenvalue problem (3.1) can then be reduced to the problem

$$S(\lambda)u_2 = 0, \tag{8.1}$$

known as Kron's problem, e.g., [80, 78], where

$$S(\lambda) = M_2 - \lambda L_2 - (M_{21} - \lambda L_{21})(M_1 - \lambda L_1)^{-1}(M_{12} - \lambda L_{12}),$$

is the Schur complement of the matrix $M - \lambda L$. This problem was studied by A. Abramov, M. Neuhaus, and V. A. Shishov [1, 79].

In analogy with (2.3), we suggest the following method

$$u_2^{k+1} = \{-S_B^{-1}S(\lambda^k) + \tau^k I\}u_2^k, \tag{8.2}$$

where S_B is a preconditioner for $S_\infty = L_2 - L_{21}L_1^{-1}L_{12}$, the Schur complement of L and

$$\frac{1}{\lambda}S(\lambda) \to -S_\infty \text{ as } \lambda \to \infty.$$

We now discuss how to choose the parameters τ^k and λ^k. The parameter τ^k can be determined either from the formula $\tau^k = \delta_1(\lambda^k - \lambda_{\min})$, or from a variational principle, see [43, 39, 45, 46] for details.

The choice of the parameter λ^k is more complicated. If we actually compute the u_1 component as the following harmonic-like extension

$$u_1^k = (M_1 - \lambda L_1)^{-1}(M_{12} - \lambda L_{12})u_2^k,$$

then it would be natural to use the Rayleigh quotient for $\lambda^k = \lambda(u^k)$. Interestingly, this is not a good idea as $\lambda(u^k)$ may not be monotonic as a function of k. Below we discuss a better way to select λ^k as some function of u_2^k described in [43, 39, 45, 46].

Similarity of the method for the second component (8.2) and the method (4.3) was used in [43, 39, 45, 46] to establish convergence rate estimates of (8.2) similar to (6.2). Our general theory for studying the rate of convergence cannot be used directly because the Schur complement $S(\lambda)$ is a rational function of λ and λ^k cannot be computed as a Rayleigh quotient.

For regular eigenvalue problems, i.e. for $M = I$, method (8.2) with a special formula for λ^k, was proposed in [43]. It was shown that method (4.3) with the preconditioner

$$B = \lambda^k L - M + \begin{pmatrix} 0 & 0 \\ 0 & S_B - S_{\lambda^k} \end{pmatrix} \tag{8.3}$$

ETNA
Kent State University
etna@mcs.kent.edu

is equivalent to method (8.2). Using this idea and known convergence theory of method (4.3), convergence rate estimates of method (8.2) were obtained. A direct extension of the methods to generalized eigenvalue problems (3.1) was carried out in [39]; steepest ascent and conjugate gradient methods were presented, but without convergence theory.

In [45, 46], a new formula

$$B = L + \begin{pmatrix} 0 & 0 \\ 0 & S_B - S_\infty \end{pmatrix}. \tag{8.4}$$

was proposed that made it possible to estimate the convergence rate of method (8.2) for generalized eigenvalue problems (3.1).

Formulas for λ^k in [43, 39, 45, 46] depend on the choice of B and are too cumbersome to be reproduced here. We only note that the latter choice of B unfortunately leads to a somewhat more expensive formula.

Our Fortran-77 code that computes eigenvalues of the Laplacian in the L-shaped domain on a uniform mesh, using preconditioned domain decomposition Lanczos-type method, is publicly available at http://www-math.cudenver.edu/~aknyazev/software/L.

It is also possible to design inner-outer iterative methods, based on the Lanczos method, for example, as outer iterations for the operator $L^{-1}M$, and a preconditioned conjugate gradient method as inner iterations for solving linear systems with the matrix L. Another possibility is to use inner-outer iterative methods based on some sort of Rayleigh quotient iterations; e.g., [41]. We expect such methods not to be as effective as our one-stage methods.

For other results on similar methods, see, e.g., [48, 54, 55, 7]. The Kron problem for differential equations can also be recast in terms of the Poincare–Steklov operators, see [49] where a method similar to (8.2) was described.

The mode synthesis method, e.g., [11, 5, 6], is another domain decomposition method for eigenproblems. This is a discretization method, where a couple of eigenfunctions corresponding to leading eigenvalues are first computed in every sub-domain. Then, these functions are used as basis functions in the Rayleigh–Ritz method for approximating the original differential eigenproblem. The method works particularly well for eigenvalue optimization problems in mechanics; e.g., [2, 51], where a series of eigenvalue problems with similar sub-domains needs to be solved. We believe that for a singe eigenvalue problem the mode synthesis method is usually less effective than the methods considered above.

9. Numerical Results. With so many competing preconditioned eigensolvers available, it is important to have some common playground for numerical testing and comparing different methods. Existing theory is not developed enough to predict whether method “A” would always be better than method “B,” except for a few cases.

Numerical tests and comparisons are still in progress; here, we present some preliminary results.

We compare our favorite block methods: the steepest ascent (SA) (7.2) and the conjugate gradient (CG) (7.3), with the block size $p = 3$. We plot, however, errors for only two top eigenvalues, leaving the third one out of the picture. The two shades of red represent method (7.2), and the two shades of blue correspond to method (7.3). It is easy to separate the methods as the CG method converges in 100–200 iterations in all tests shown, while the SA requires at least 10 times more steps.

In all tests $M = I$, and we measure the eigenvalue error as

$$error_i = \frac{1}{\lambda_i^k} - \frac{1}{\lambda_i},\ i = 1, 2,$$

for historical reasons.

ETNA
Kent State University
etna@mcs.kent.edu

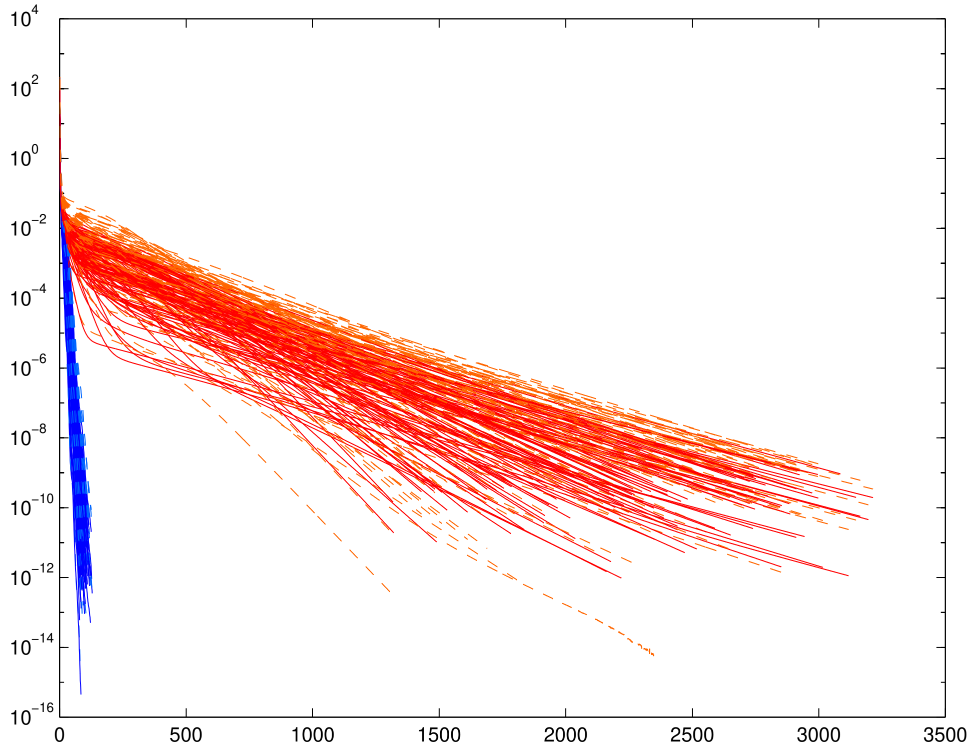

FIG. 9.1. *Error for the block Steepest Ascent and the block CG methods. Hundred runs,* $N = 400$.

In our first test, a model eigenvalue problem 400-by-400 with stiffness matrix $L = \operatorname{diag}\{1, 2, 3, \ldots, 400\}$, is solved, see Figure 9.1. For this problem,

$$\lambda_1 = 1,\ \lambda_2 = \frac{1}{2},\ \lambda_3 = \frac{1}{3},\ \lambda_{\min} = \frac{1}{400}.$$

On Figure 9.1, numerical results of 100 runs of the same codes are displayed, with a random initial guess and a *random preconditioner* B, satisfying assumption (3.5) with $\delta = 10^{-3}$.

In similar further tests, we use the same codes to solve a model N-by-N eigenvalue problem with a randomly chosen diagonal matrix L such that

$$\lambda_1 = 1,\ \lambda_2 = \frac{1}{2},\ \lambda_3 = \frac{1}{3},\ \lambda_{\min} = 10^{-10}.$$

In these tests, our goal is to check that huge condition number of L, the size of the problem N, and distribution of eigenvalues in the unwanted part of the spectrum do not noticeably affect the convergence, as we would expect from theoretical results for simpler methods. As in the previous experiments, the initial guess and the preconditioner are randomly generated, and $\delta = 10^{-3}$.

Figure 3.1 and 3.2 show that, indeed the convergence is about the same for different values of N and different choices of parameters.

On all figures, the elements of a bundle, of convergence history lines, are quite close, which suggests that our assumption (3.5) on the preconditioner, is fundamental, and that the ratio $\delta = \delta_0/\delta_1$ does predict the rate of convergence. We also observe that convergence for

ETNA
Kent State University
etna@mcs.kent.edu

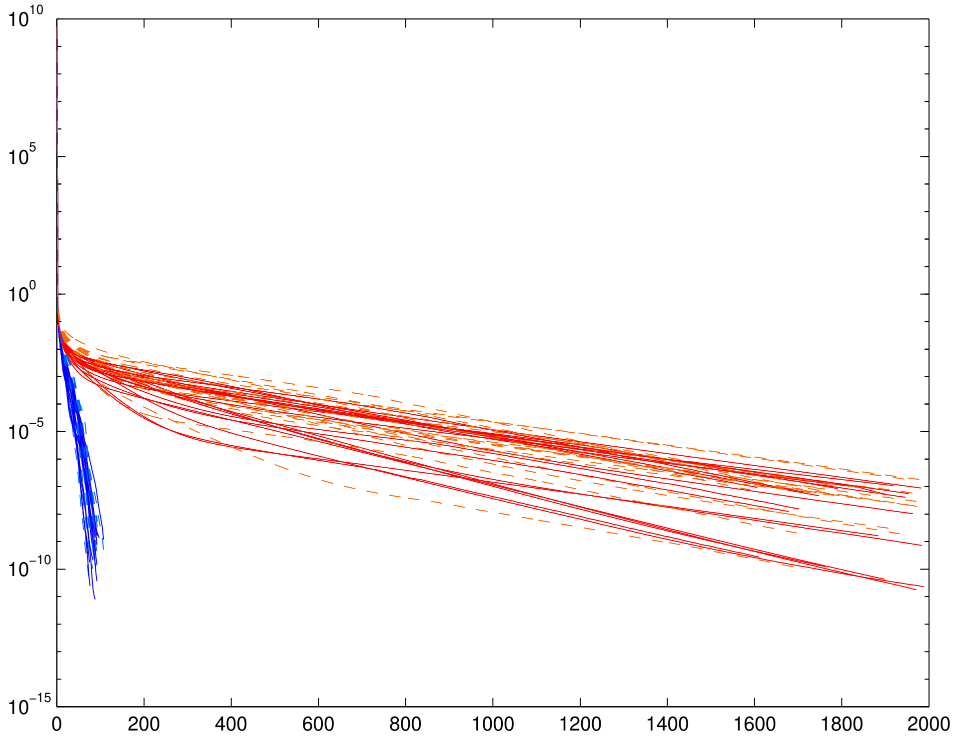

FIG. 9.2. *Twenty runs,* $N = 1200$.

the first eigenvalue, in dark colors, is typically, but not always, faster than that for the second one, in lighter colors and dashed. That is especially noticeable on SA convergence history. Finally, we can draw the conclusion that the CG method is clearly superior.

10. Conclusion. Preconditioned eigensolvers have been the subject of recent research. Hopefully, the present paper could help to point out some unusual and forgotten references, and to prevent old results and ideas from being rediscovered. An introduced systematic classification of preconditioned eigensolvers should make it easier to compare different methods.

An interactive Web page for preconditioned eigensolvers was created at

http://www-math.cudenver.edu/~aknyazev/research/eigensolvers.
All the references of the present paper were available in the BIBTEX format on this Web page, as well as some links to software.

The author would like to thank anonymous reviewers and R. Lehoucq for their numerous helpful suggestions and comments. The author was grateful to an editor at ETNA for extensive copy editing.

Finally, we wanted to comment on the bibliography. A part of the BIBTEX file was produced using MathSciNet, an electronic version of the Mathematical Reviews. That was where our references could be checked, and little-known journals, being referred to, could be identified. When citing Soviet works, we put references to English versions when available. Most often, the latter were just direct translations from Russian, but we decided not to give references to the originals, only to the translations. When citing papers in Russian, we gave English translations of the titles, usually without transliteration, but we provided only translit-

ETNA
Kent State University
etna@mcs.kent.edu

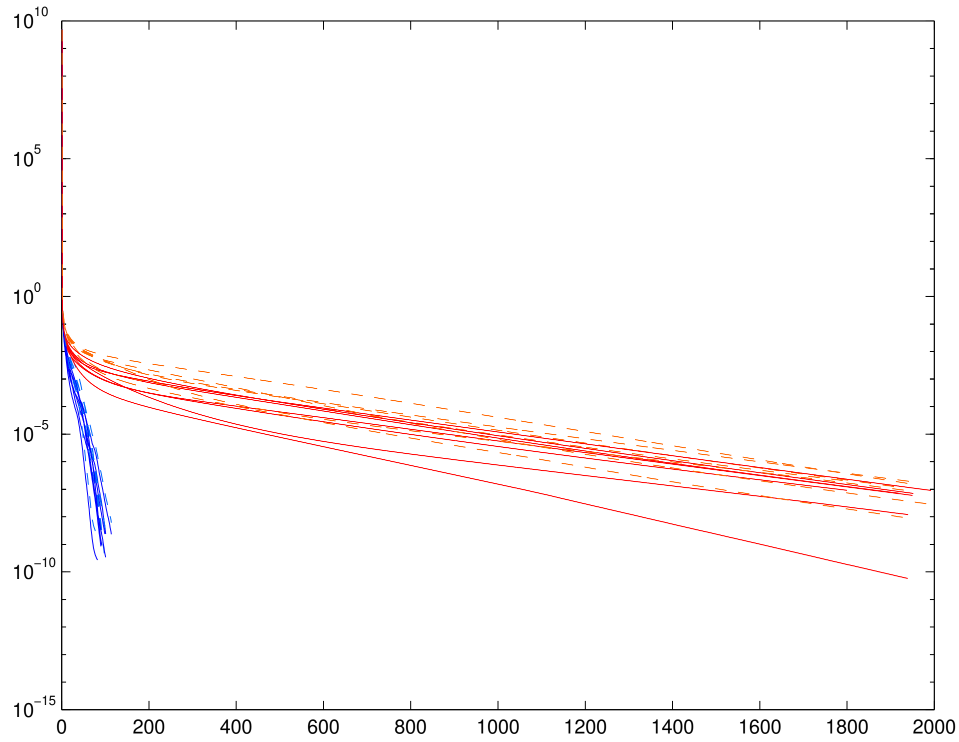

FIG. 9.3. *Ten runs,* $N = 2000$.

eration of sources, i.e. journals and publishers, and always add a note "in Russian." Several Russian titles were published by Acad. Nauk SSSR Otdel Vychisl. Mat., which stands for the Institute of Numerical Mathematics of the USSR Academy of Sciences. Many results on the subject by the author were published with detailed proofs in Russian in the monograph [35]; a survey of these results, without proofs, was published in English in [36].

A preliminary version of the present paper was published as a technical report UCD-CCM 135, 1998, at the Center for Computational Mathematics, University of Colorado at Denver.

REFERENCES

[1] A. ABRAMOW AND M. NEUHAUS, *Bemerkungen über Eigenwertprobleme von Matrizen höherer Ordnung*, in Les mathématiques de l'ingénieur, Mém, Publ. Soc. Sci. Arts Lett. Hainaut, vol. hors Série (1958), pp. 176–179.

[2] N. V. BANICHUK, *Problems and methods of optimal structural design*, Math. Concepts Methods Sci. Engng., 26, Plenum Press, New York, 1983.

[3] LUCA BERGAMASCHI, GIUSEPPE GAMBOLATI AND GIORGIO PINI, *Asymptotic convergence of conjugate gradient methods for the partial symmetric eigenproblem*, Numer. Linear Algebra Appl., 4(2) (1997), pp. 69–84.

[4] A. N. BESPALOV AND YU. A. KUZNETSOV, *RF cavity computations by the domain decomposition and fictitious component methods*, Soviet J. Numer. Anal. Math. Modeling, 2(4) (1987), pp. 259–278.

[5] F. BOURQUIN, *Analysis and comparison of several component mode synthesis methods on one-dimensional domains*, Numer. Math., 58(1) (1990), pp. 11–33.

[6] ———, *Component mode synthesis and eigenvalues of second order operators: discretization and*

algorithm, RAIRO Modél. Math. Anal. Numér., 26(3) (1992), pp. 385–423.

[7] ———, *A domain decomposition method for the eigenvalue problem in elastic multistructures*, Asymptotic methods for elastic structures (Lisbon, 1993), de Gruyter, Berlin, 1995, pp. 15–29.

[8] JAMES H. BRAMBLE, JOSEPH E. PASCIAK, AND ANDREW V. KNYAZEV, *A subspace preconditioning algorithm for eigenvector/eigenvalue computation*, Adv. Comput. Math., 6(2) (1996), pp. 159–189.

[9] ZHI HAO CAO, *Generalized Rayleigh quotient matrix and block algorithm for solving large sparse symmetric generalized eigenvalue problems*, Numer. Math. J. Chinese Univ., 5(4) (1983), pp. 342–348.

[10] ZHIQIANG CAI, JAN MANDEL AND STEVE MCCORMICK, *Multigrid methods for nearly singular linear equations and eigenvalue problems*, SIAM J. Numer. Anal., 34(1) (1997), pp. 178–200.

[11] R. R. CRAIG, JR., *Review of time domain and frequency domain component mode synthesis methods*, in Design Methods for Two-Phase Flow in Turbomachinery, New York, ASME, 1985.

[12] M. CROUZEIX, B. PHILIPPE AND M. SADKANE, *The Davidson method*, SIAM J. Sci. Comput., 15(1) (1994), pp. 62–76.

[13] J. K. CULLUM AND R. A. WILLOUGHBY, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, Birkhäuser, Boston, 1985.

[14] D. F. DAVIDENKO, *The method of variation of parameters as applied to the computation of eigenvalues and eigenvectors of matrices*, Soviet Math. Dokl., 1 (1960), pp. 364–367.

[15] E. R. DAVIDSON, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices*, J. Comput. Phys., 17 (1975), pp. 817–825.

[16] ———, *Matrix eigenvector methods*, G. H. F. Direcksen and S. Wilson, eds., Methods in Computational Molecular Physics, Reidel, Boston, 1983, pp. 95–113.

[17] I. S. DUFF, R. G. GRIMES AND J. G. LEWIS, *Sparse matrix test problems*. ACM Trans. Math. Software, 15 (1989), pp. 1–14.

[18] E. G. D'YAKONOV, *Iteration methods in eigenvalue problems*, Math. Notes, 34 (1983), pp. 945–953.

[19] E. G. D'YAKONOV AND A. V. KNYAZEV, *Group iterative method for finding lower-order eigenvalues*, Moscow University, Ser. 15, Comput. Math. Cybernetics, 2 (1982), pp. 32–40.

[20] ———, *On an iterative method for finding lower eigenvalues*, Russian J. of Numer. Anal. Math. Modeling, 7(6) (1992), pp. 473–486.

[21] E. G. D'YAKONOV AND M. YU. OREKHOV, *Minimization of the computational labor in determining the first eigenvalues of differential operators*, Math. Notes, 27(5–6) (1980), pp. 382–391.

[22] ———, *Optimization in solving elliptic problems*, CRC Press, Boca Raton, FL, 1996.

[23] T. ARIAS, A. EDELMAN, AND S. SMITH, *Curvature in Conjugate Gradient Eigenvalue Computation with Applications to Materials and Chemistry Calculations*, in Proceedings of the 1994 SIAM Applied Linear Algebra Conference, J.G. Lewis, ed., SIAM, Philadelphia, (1994), pp. 233–238.

[24] Y. T. FENG AND D. R. J. OWEN, *Conjugate gradient methods for solving the smallest eigenpair of large symmetric eigenvalue problems*, Internat. J. Numer. Methods Engrg., 39(13) (1996), pp. 2209–2229.

[25] GIUSEPPE GAMBOLATI, GIORGIO PINI AND MARIO PUTTI, *Nested iterations for symmetric eigenproblems*, SIAM J. Sci. Comput., 16(1) (1995), pp. 173–191.

[26] GIUSEPPE GAMBOLATI, FLAVIO SARTORETTO AND PAOLO FLORIAN, *An orthogonal accelerated deflation technique for large symmetric eigenproblems*, Comput. Methods Appl. Mech. Engrg., 94(1) (1992), pp. 13–23.

[27] S. K. GODUNOV, V. V. OGNEVA, AND G. K. PROKOPOV, *On the convergence of the modified steepest descent method in application to eigenvalue problems*, in American Mathematical Society Translations. Ser. 2. Vol. 105. (English) Partial differential equations in Proceedings of a symposium dedicated to Academician S. L. Sobolev, American Mathematical Society, Providence, RI, (1976), pp. 77–80.

[28] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, John Hopkins University Press, Baltimore, MD, 1983.

[29] ROGER G. GRIMES, JOHN G. LEWIS AND HORST D. SIMON, *A shifted block Lanczos algorithm for solving sparse symmetric generalized eigenproblems*, SIAM J. Matrix Anal. Appl., 15(1) (1994), pp. 228–272.

[30] S. M. GRZEGÓRSKI, *On the scaled Newton method for the symmetric eigenvalue problem*, Computing, 45(3), (1990), pp. 277-282.

[31] LIANG JIAO HUANG AND TIEN-YIEN LI, *Parallel homotopy algorithm for symmetric large sparse eigenproblems*, J. Comput. Appl. Math., 60(1-2) (1995), pp. 77–100. Linear/nonlinear iterative methods and verification of solution (Matsuyama, 1993).

[32] K. HIRAO AND H. NAKATSUJI, *A generalization of the Davidson's method to large nonsymmetric eigenvalue problem*, J. Comput. Phys, 45(2) (1982), pp. 246–254.

[33] V. P. IL'IN, *Numerical Methods of Solving Electrophysical Problems*, Nauka, Moscow, 1985, in Rus-

sian.

[34] A. V. KNYAZEV, *Some two-step methods for finding the boundaries of the spectrum of a linear matrix pencil*, Technical Report 3749/16, Kurchatov's Institute of Atomic Energy, Moscow, 1983, in Russian.

[35] ———, *Computation of eigenvalues and eigenvectors for mesh problems: algorithms and error estimates*, Dept. Numerical Math. USSR Academy of Sciences, Moscow, 1986, in Russian.

[36] ———, *Convergence rate estimates for iterative methods for mesh symmetric eigenvalue problem*, Soviet J. Numer. Anal. Math. Modeling, 2(5) (1987), pp. 371–396.

[37] ———, *Methods for the derivation of some estimates in a symmetric eigenvalue problem*, in Computational Processes and Systems, Gury I. Marchuk, ed., Computational Processes and Systems, No. 5, Nauka, Moscow, 1987, pp. 164–173, in Russian.

[38] ———, *Modified gradient methods for spectral problems*, Differ. Uravn., 23(4) (1987), pp. 715–717, in Russian.

[39] ———, *A preconditioned conjugate gradient method for eigenvalue problems and its implementation in a subspace*, in International Ser. Numerical Mathematics, v. 96, Eigenwertaufgaben in Natur- und Ingenieurwissenschaften und ihre numerische Behandlung, Oberwolfach, (1990), pp. 143–154.

[40] ———, *New estimates for Ritz vectors*, Math. Comp., 66(219) (1997), pp. 985–995.

[41] A. V. KNYAZEV AND I. A. SHARAPOV, *Variational Rayleigh quotient iteration methods for symmetric eigenvalue problem*, East-West J. of Numer. Math., 1(2) (1993), pp. 121–128.

[42] A. V. KNYAZEV AND A. L. SKOROKHODOV, *On the convergence rate of the steepest descent method in euclidean norm*, USSR J. Comput. Math. and Math. Physics, 28(5) (1988), pp. 195–196.

[43] ———, *Preconditioned iterative methods in subspace for solving linear systems with indefinite coefficient matrices and eigenvalue problems*, Soviet J. Numer. Anal. Math. Modeling, 4(4) (1989), pp. 283–310.

[44] ———, *On exact estimates of the convergence rate of the steepest ascent method in the symmetric eigenvalue problem*, Linear Algebra Appl., 154–156 (1991), pp. 245–257.

[45] ———, *The preconditioned gradient-type iterative methods in a subspace for partial generalized symmetric eigenvalue problem*, Soviet Math. Doklady, 45(2) (1993), pp. 474–478.

[46] ———, *The preconditioned gradient-type iterative methods in a subspace for partial generalized symmetric eigenvalue problem*, SIAM J. Numer. Anal., 31(4) (1994), pp. 1226–1239.

[47] N. KOSUGI, *Modification of the Liu-Davidson method for obtaining one or simultaneously several eigensolutions of a large real symmetric matrix*, J. Comput. Phys., 55(3) (1984), pp. 426–436.

[48] YU. A. KUZNETSOV, *Iterative methods in subspaces for eigenvalue problems*, in Vistas in Appl. Math., Numer. Anal., Atmospheric Sciences, Immunology, A. V. Balakrishinan, A. A. Dorodnitsyn and J. L. Lions, eds., Optimization Software, Inc., New York, 1986, pp. 96–113.

[49] V. I. LEBEDEV, *Metod kompozitsii [Method of Composition], Akad. Nauk SSSR Otdel Vychisl. Mat., Moscow, 1986, in Russian.*

[50] S. H. LUI AND G. H. GOLUB*,* Homotopy method for the numerical solution of the eigenvalue problem of self-adjoint partial differential operators*, Numer. Algorithms, 10(3-4) (1995), pp. 363–378.*

[51] V. G. LITVINOV*,* Optimization in elliptic boundary value problems with applications to mechanics*, Nauka, Moscow, 1987, in Russian.*

[52] D. E. LONGSINE AND S. F. MCCORMICK*,* Simultaneous Rayleigh-quotient minimization methods for $Ax = \lambda Bx$., Linear Algebra and Applications*, 34 (1980), pp. 195–234.*

[53] S. H. LUI, H. B. KELLER AND T. W. C. KWOK*,* Homotopy method for the large, sparse, real nonsymmetric eigenvalue problem*, SIAM J. Matrix Anal. Appl., 18(2) (1997), pp. 312–333.*

[54] JENN-CHING LUO*,* Solving eigenvalue problems by implicit decomposition*, Numer. Methods Partial Differential Equations, 7(2) (1991), pp¿ 113–145.*

[55] ———*,* A domain decomposition method for eigenvalue problems*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, Norfolk, VA, 1991, SIAM, Philadelphia, PA (1992), pp. 306–321.*

[56] S. YU. MALIASSOV*,* On the Schwarz alternating method for eigenvalue problems*, Russian J. Numer. Anal. Math. Modeling, 13(1) (1998), pp. 45–56.*

[57] JAN MANDEL, CHARBEL FARHAT, AND XIAO-CHUAN CAI*, eds.,* Domain Decomposition Methods 10*, Contemp. Math., volume 218, American Mathematical Society, Providence, RI, 1998.*

[58] V. P. MASLOV*,* Asimptoticheskie metody i teoriya vozmushcheniĭ [Asymptotic methods and perturbation theory]*, Nauka, Moscow, 1988, in Russian.*

[59] S. F. MCCORMICK AND T. NOE*,* Simultaneous iteration for the matrix eigenvalue problem*, Linear Algebra Appl., 16(1) (1977), pp. 43–56.*

[60] ARND MEYER*,* Modern algorithms for large sparse eigenvalue problems*, Math. Res., volume 34, Akademie-Verlag, Berlin, 1987.*

[61] R. B. MORGAN*,* Davidson's method and preconditioning for generalized eigenvalue problems*, J.*

Comp. Phys., 89 (1990), pp. 241–245.

[62] 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), pp. 817–825.

[63] ———, *Preconditioning the Lanczos algorithm for of sparse symmetric eigenvalue problems*, SIAM J. Sci. Comput., 14(3) (1993), pp. 585–593.

[64] YU. M. NECHEPURENKO, *The singular function method for computing the eigenvalues of polynomial matrices*, Comput. Math. Math. Phys., 35(5) (1995), pp. 507–517.

[65] 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), pp. 413–458.

[66] V. G. PRIKAZCHIKOV, *Prototypes of iteration methods of solving an eigenvalue problem*, Differential Equations, 16 (1981), pp. 1098–1106.

[67] ———, *Steepest descent in a spectral problem*, Differentsialnye Uravneniya, 22(7) (1986), pp. 1268–1271, 1288, in Russian.

[68] V. G. PRIKAZCHIKOV AND A. N. KHIMICH, *Iterative method for solving stability and oscillation problems for plates and shells*, Sov. Appl. Mech., 20 (1984), pp. 84–89.

[69] AKSEL RUÈ, *Krylov rational sequence methods for computing eigenvalues*, in Computational methods in linear algebra, Moscow, (1982), pp. 203–218, in Russian, Akad. Nauk SSSR Otdel Vychisl. Mat., Moscow, 1983, in Russian.

[70] AXEL RUHE, *SOR-methods for the eigenvalue problem with large sparse matrices*, Math. Comput., 28(127) (1974), pp. 695–710.

[71] ———, *Iterative eigenvalue algorithm based on convergent splittings*, J. Comp. Phys., 19(1) (1975), pp. 110–120.

[72] ———, *Rational Krylov: a practical algorithm for large sparse nonsymmetric matrix pencils*, SIAM J. Sci. Comput., 19(5) (1998), pp. 1535–1551, (electronic).

[73] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*. Halsted Press, NY, 1992.

[74] M. SADKANE, *Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems*, Numer. Math., 64(2) (1993), pp. 195–211.

[75] B.A. SAMOKISH, *The steepest descent method for an eigenvalue problem with semi-bounded operators*, Izvestiya Vuzov, Math., 5 (1958), pp. 105–114, in Russian.

[76] G. V. SAVINOV, *Investigation of the convergence of a generalized method of conjugate gradients for determining the extremal eigenvalues of a matrix*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI), 111 (1981), pp. 145–150, 237.

[77] D. S. SCOTT, *Solving sparse symmetric generalized eigenvalue problems without factorization*, SIAM J. Numer. Anal., 18 (1981), pp. 102–110.

[78] N. S. SEHMI, *Large order structural eigenanalysis techniques*, Ellis Horwood Series: Mathematics and its Applications. Ellis Horwood Ltd., Chichester, 1989.

[79] V. A. SHISHOV, *A method for partitioning a high order matrix into blocks in order to find its eigenvalues*, USSR Comput. Math. Math. Phys., 1(1) (1961), pp. 186–190.

[80] A. SIMPSON AND B. TABARROK, *On Kron's eigenvalue procedure and related methods of frequency analysis*, Quart. J. Mech. Appl. Math., 21 (1968), pp. 1–39.

[81] GERARD L. G. SLEIJPEN AND HENK A. VAN DER VORST, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17(2) (1996), pp. 401–425.

[82] E. SUETOMI AND H. SEKIMOTO, *Conjugate gradient like methods and their application to eigenvalue problems for neutron diffusion equation*, Annals of Nuclear Energy, 18(4) (1991), p. 205.

[83] DANIEL B. SZYLD AND OLOF B. WIDLUND, *Applications of conjugate gradient type methods to eigenvalue calculations*, in Advances in computer methods for partial differential equations, III, Proc. Third IMACS Internat. Sympos., Lehigh Univ., Bethlehem, PA, 1979, pp. 167–173.

[84] HENK VAN DER VORST, *Subspace iteration for eigenproblems*, CWI Quarterly, 9(1-2) (1996), pp. 151–160.

[85] H. YANG, *Conjugate gradient methods for the Rayleigh quotient minimization of generalized eigenvalue problems*, Computing, 51(1) (1993), pp. 79–94.

[86] T. ZHANG, K. H. LAW AND G. H. GOLUB, *On the homotopy method for perturbed symmetric generalized eigenvalue problems*, SIAM J. Sci. Comput., 19(5) (1998), pp. 1625–1645 (electronic).

[87] P. F. ZHUK AND V. G. PRIKAZCHIKOV, *An effective estimate of the convergence of an implicit iteration method for eigenvalue problems*, Differential Equations, 18 (1983), pp. 837–841, 1983.