A posteriori error estimation for elliptic eigenproblems

Klaus Neymeyr

Mathematisches Institut der Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany; neymeyr@na.uni-tuebingen.de

SUMMARY

An *a posteriori* error estimator is presented for a subspace implementation of Preconditioned Inverse Iteration, which derives from the well–known Inverse Iteration in such a way that the associated system of linear equations is solved approximately by using a preconditioner. The error estimator is integrated in an adaptive multigrid algorithm to compute approximations of a modest number of the smallest eigenvalues together with the eigenfunctions of an elliptic differential operator.

Error estimation is applied both within the actual finite element space (in order to estimate the iteration error) as well as in its hierarchical refinement of higher order elements (to estimate the discretization error) which gives rise to a balanced reduction of the iteration error and of the discretization error in the adaptive multigrid algorithm.

KEY WORDS: Symmetric eigenvalue problem; Preconditioning; Multigrid; A posteriori error estimation.

1. Introduction

Using adaptive discretization methods for the numerical solution of partial differential equations arising from problems in physics and engineering, numerical approximation within a prescribed tolerance can often be gained with only a small portion of the work which is necessary when uniform grid refinement is employed. Whereas *a priori* error estimation gives asymptotic rates of convergence as the mesh parameter h tends to zero, one needs for adaptive methods *a posteriori* error estimators to provide practical criteria to control the adaptive mesh refinement and to construct a stopping condition for the iterative solver.

Adaptive discretization methods and a posteriori error estimators can nevertheless be applied to construct efficient solvers for the eigenproblem. The structure of the eigenvalue problem leads to a modified approach and to some extensions, e.g. for treating invariant subspaces of eigenvectors in an appropriate manner.

A recent survey on *a posteriori* error estimators has been given by Verfürth [25] discussing the most frequently used error estimators which have been developed in the last two decades. For example, in the software package PLTMG [2] for solving elliptic partial differential equations, the triangle oriented error estimator of Bank and Weiser [3] is implemented, which is based on the solution of small Neumann problems on each triangle. For edge oriented local error estimation, see for instance Deuflhard, Leinen and Yserentant [8] and Zienkiewicz et. al. [28].

For the generalized eigenvalue problem Friberg [13] constructed an a posteriori error estimator describing the relative change in an eigenvalue for a hierarchical refinement. The error estimator presented here is based on a similar quantity, namely the first order term of a Taylor expansion of the Rayleigh quotient around the actual iterate.

This work presents the mathematical concept underlying the *a posteriori* error estimation for a subspace implementation of *Preconditioned Inverse Iteration* on a sequence of adaptively generated grids. Preconditioned inverse iteration (PINVIT) derives from inverse iteration (INVIT or also called inverse power method) in a way that the associated system of linear equations is solved approximately by using a preconditioner [21, 22]. The extension of preconditioned inverse iteration to a subspace algorithm is analyzed in [20].

The remaining part of this paper is organized as follows: In Section 2 we introduce preconditioned inverse iteration and show how to transfer the convergence theory to the generalized eigenvalue problem, which appears necessary since the results in [21, 20] are only given for the standard eigenvalue problem. In Section 3 an error estimator for eigenvalue problems is presented. In Section 4 the iteration error estimator as well as the discretization error estimator are derived, which prepare the ground for the construction for an adaptive multigrid solver. Finally, the results of some numerical experiments are given in Section 5.

2. A preconditioned eigensolver for the generalized eigenvalue problem

2.1. The eigenproblem and its discretization

Let $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, be a connected, bounded, polygonal domain with mixed homogeneous Dirichlet and Neumann boundary conditions and consider the eigenvalue problem

$$a(u, v) = \bar{\lambda}\ (u, v), \qquad\qquad v \in H(\Omega), \tag{1}$$

where $H(\Omega)$ consists of all functions of a certain Hilbert space V vanishing on the Dirichlet boundary. The bilinear form $a(\cdot, \cdot)$ corresponds to a self–adjoint, coercive, elliptic differential operator and $(\cdot, \cdot)$ is the inner product in V. E.g. for the Laplacian, V is a subspace of the Sobolev space $H^{1,2}(\Omega)$ and $(\cdot, \cdot)$ is the usual $L^2(\Omega)$ inner product. Here, our aim is to determine approximations of a modest number of the smallest eigenvalues $\bar{\lambda}$ together with its eigenfunctions u.

Therefore, we consider a finite element discretization of (1) which leads to the following generalized matrix eigenvalue problem

$$Ax_i = \lambda_i M x_i, \qquad i = 1, \ldots, n, \tag{2}$$

where the finite element space is assumed to have the dimension n. Usually, $A \in \mathbb{R}^{n\times n}$ is called the stiffness matrix and $M \in \mathbb{R}^{n\times n}$ the mass matrix. Both matrices are assumed to be symmetric and positive definite. Moreover, they are typically very large and sparse. For simple test problems the dimension n may exceed 10^6 even on standard present–day personal computers. Obviously, these matrices cannot be stored in the computer as full matrices but program routines are provided to compute the matrix–vector products Ax and Mx. Moreover, A typically has a large spectral condition number; e.g. for the Laplacian the condition number behaves like h^{-d} in the mesh parameter h.

The eigenvalues of the generalized matrix eigenvalue problem (A, M) may have arbitrary multiplicity and are put in an order so that $0 < \lambda_1 \leq \ldots \leq \lambda_n$. The corresponding eigenvectors

x_i are assumed to satisfy $\|x_i\|_M^2 = (x_i, Mx_i) = 1$, where $(\cdot,\cdot)$ denotes the Euclidean norm. Then the Rayleigh quotient associated with the generalized matrix eigenvalue problem is given by

$$\lambda(x) = \frac{(x, Ax)}{(x, Mx)}. \tag{3}$$

2.2. Gradient type solvers for the eigenproblem

Our goal is a *partial* solution of the eigenvalue problem (2), i.e. we are interested in computing some of the smallest eigenvalues together with their invariant subspace. The large number of unknowns rule out the use of any eigensolvers that rely on dense matrices or their factorizations like the QR or Jacobi method. We cannot even factor A or M by the LU or Cholesky decomposition because of a lack of the necessary computer storage. In other words, we are forced to construct an iterative eigensolver in a so–called *matrix–free environment.*

The most simple eigensolver satisfying the mentioned restriction is based on the idea to reformulate the eigenvalue problem (2) as an optimization problem for the Rayleigh quotient (3). Following the idea of Kantorovich [15] as well as Hestenes and Karush [14] one can correct a given iterate x in the direction of the negative gradient of the Rayleigh quotient in order to decrease the Rayleigh quotient of the new iterate.

Since the gradient of the Rayleigh quotient (3) for a given iterate x is given by

$$\nabla\lambda(x) = \frac{2}{(x, Mx)}(Ax - \lambda(x)Mx), \tag{4}$$

the so–called gradient method for the eigenproblem has the form

$$x' := x - \omega(Ax - \lambda(x)Mx). \tag{5}$$

The new iterate x' has a decreased Rayleigh quotient for a proper choice of the scaling parameter ω. In the best case ω is determined in a way that the Rayleigh quotient of the new iterate x' takes its minimum. For the latter choice the scheme is called the *steepest descent method.*

As long as the residual vector $Ax - \lambda(x)Mx$ is nonzero, which means that x is no eigenvector of A, the Rayleigh quotients of the gradient method (steepest descent) form a strictly monotone decreasing sequence of real numbers, which (usually) tends to the smallest eigenvalue λ_1. Then the iterates themselves converge to an associated eigenvector. Unfortunately the gradient scheme (5) is known to suffer from poor convergence for ill–conditioned A [6]. Preconditioning can provide some cure.

2.3. Preconditioned gradient eigensolvers

Applying preconditioning to the gradient eigensolver (5) can improve its convergence properties. For the best multigrid preconditioners one can even guarantee grid–independent convergence and thus (quasi–) optimal computational complexity for the iterative solver. The preconditioner B^{-1} (which is also often called an approximate inverse of A) is assumed to satisfy the estimate

$$\delta_0(x, Bx) \le (x, Ax) \le \delta_1(x, Bx), \qquad x \in \mathbb{R}^n, \tag{6}$$

where δ_0 and δ_1 are positive constants. Here, we prefer the more convenient assumption of a properly scaled preconditioner B^{-1} for which

$$\|I - B^{-1}A\|_A \leq \gamma. \tag{7}$$

Therein γ is a positive constant, $\gamma \in [0,1)$, and $\|\cdot\|_A$ denotes the operator norm induced by A, cf. also Section 2.5.

Now *preconditioned gradient eigensolvers* can be derived by premultiplying the residual in (5) by the preconditioner. This leads to the scheme

$$x' := x - \omega B^{-1}(Ax - \lambda(x)Mx). \tag{8}$$

The assumption (7) of the properly scaled preconditioner B^{-1} allows us from now on to set $\omega = 1$. There is a vast literature on the convergence theory of (8), predominantly written by Russian authors; but the estimates contained in these works are only asymptotically sharp. See D'yakonov [9] and Knyazev [16] for a review.

A new approach to the convergence analysis of (8) has been presented in [21, 22], where also sharp and non–asymptotic convergence estimates have been derived. The idea of this analysis is based on an alternative derivation of (8) by means of approximatively solving the system of linear equations associated with inverse iteration, i.e. for a given iterate x the new iterate $\hat{x}$ is the solution of the linear system

$$A\hat{x} = \lambda Mx, \tag{9}$$

for some $\kappa \neq 0$. First observe that the choice of κ is immaterial, since inverse iteration is invariant with respect to the scaling of the iterates. Setting $\kappa = \lambda(x)$ paves the way for applying preconditioning to solve (9), since then the residual of (9) vanishes for x converging to an eigenvector. Writing down the error propagation equation for the approximate solution of (9) for this choice of κ we obtain

$$x' - \lambda(x)A^{-1}Mx = (I - B^{-1}A)(x - \lambda(x)A^{-1}Mx). \tag{10}$$

In this equation the initial error $x - \lambda(x)A^{-1}Mx$ is premultiplied by the error propagation matrix resulting in the final error $x' - \lambda(x)A^{-1}Mx$, where x' approximates $\hat{x}$. Now the key point is that (10) can be rewritten in the simple form (containing no inverse of A)

$$x' = x - B^{-1}(Ax - \lambda(x)Mx). \tag{11}$$

To sum up, we have derived the iteration (8) by approximating inverse iteration. For this reason we also call (11) the scheme of *preconditioned inverse iteration* (PINVIT). The new approach to its convergence analysis in [21, 22] exploits the fact that (10) has a simple geometric interpretation, i.e. the set of all the possible new iterates (under the assumption (7)) is a ball whose center $\lambda(x)A^{-1}Mx$ is defined by the result of inverse iteration. This geometric view turns out as a valuable tool for the analysis.

2.4. Subspace implementation of preconditioned inverse iteration

In the following we introduce the subspace implementation of preconditioned inverse iteration. Therefore, let an s–dimensional subspace of the $\mathbb{R}^n$ be spanned by the vectors v_i, for $i = 1, \ldots, s$, which are assumed to be the Ritz vectors of (A, M). Then for $V = [v_1, \ldots, v_s] \in \mathbb{R}^{n\times s}$ it holds

$$V^TAV = \Theta = \operatorname{diag}(\theta_1, \ldots, \theta_s) \quad \text{and} \quad V^TMV = I, \tag{12}$$

where $I \in \mathbb{R}^{s\times s}$ is the identity matrix. The Ritz vectors are assumed to be in an order so that the positive Ritz values θ_i are increasing with i.

In order to generalize preconditioned inverse iteration to a subspace method, which is to determine an invariant subspace of eigenvectors corresponding to some of the smallest eigenvalues, we rewrite (10) in matrix terms

$$\tilde{V} = A^{-1}MV\Theta + (I - B^{-1}A)(V - A^{-1}MV\Theta). \tag{13}$$

Its simplified representation (containing no inverse of A) in analogy to (11) reads

$$\tilde{V} = V - B^{-1}(AV - MV\Theta). \tag{14}$$

Therein the preconditioner B^{-1} is assumed to satisfy (7).

Once again, we have the two representations (13) and (14) of the preconditioned subspace iteration. The first representation (13) shows its relation to inverse iteration (INVIT). Inserting the best possible preconditioner, i.e. $B = A$, in Equation (13) one obtains

$$\tilde{V} = A^{-1}MV\Theta. \tag{15}$$

Then $\tilde{V}$ is the result of inverse iteration (for the generalized eigenvalue problem) if applied to the subspace spanned by the columns of V. The convergence of inverse iteration in analyzed for instance in [7, 23]. In the standard representation of INVIT the (nonsingular) diagonal matrix Θ in (15) equals the identity matrix. But Θ has no influence on the Ritz values and Ritz vectors of $\tilde{V}$, since it only gives rise to a column scaling of $\tilde{V}$. Approximate solution of (15) for $\tilde{V}$ immediately leads to Equation (13), which simply is the corresponding error propagation equation. Therefore, $\tilde{V}$ in (13) equals the result of inverse iteration $A^{-1}MV\Theta$ plus a perturbation term given by the product of the error propagation matrix $I - B^{-1}A$ and the correction $V - A^{-1}MV\Theta$ made by inverse iteration. The correction term $V - A^{-1}MV\Theta$ tends to the zero matrix if the column space of V converges to an invariant subspace of eigenvectors of (A, M).

Once computed $\tilde{V}$ by (14), the Rayleigh–Ritz procedure is applied to determine the approximate eigenvectors and eigenvalues. The Ritz vectors v_i', $i = 1, \ldots, s$, of the column space span$(\tilde{V})$ define the columns of V', and the corresponding Ritz values θ_i' are the diagonal elements of the matrix Θ'. Thus the preconditioned subspace eigensolver iterates the transformation $V,\ \Theta \longrightarrow\ V',\ \Theta'$.

Different attempts have been made to analyze (14) in its simplified form with $M = I$. Bramble et al. [5] have analyzed Equation (14) and have shown convergence under restrictive assumptions on the initial space; but the algorithm was not related to inverse iteration. For an analysis of some simplified preconditioned subspace schemes see [10, 11, 9, 16]. In all these works the convergence theory is based on a direct analysis of Equation (14), which is commonly related with preconditioned gradient methods for the eigenvalue problem, see Section 2.2.

In contrast to that, the analysis in [20], for $M = I$, attacks Equation (13) and derives sharp convergence estimates for the Ritz values. The convergence theory is mainly based on an analysis of the geometry underlying preconditioned inverse iteration, see [21, 22] for the analysis of the preconditioned vector ($s = 1$) scheme. In [20], treating the subspace algorithm (14), it is shown that the convergence estimates for each Ritz value are the same as those which have been derived for the Rayleigh quotient of the vector scheme (11) in [22].

In the next theorem the results given in [20] are extended to the generalized matrix eigenvalue problem (2) emerging from the finite element discretization, while the operator representation

in [5, 20, 21, 22] only treats the standard eigenvalue problem. Theorem 2.1 does not consist in a trivial substitution of the Euclidean inner product by the one induced by M, but recurs on the error propagation equation underlying (13).

Theorem 2.1. *Consider the generalized eigenproblem for (A, M) given by (2) with the eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ of arbitrary multiplicity. The preconditioner B^{-1} is assumed to satisfy (7) for a $\gamma \in [0, 1)$.*

Let an s–dimensional subspace V be defined by (12). Then the Ritz values $\theta_1, \ldots, \theta_s$ converge linearly to eigenvalues of (A, M). For each of these θ_i it holds that:

If $\theta_i \in [\lambda_{k_i}, \lambda_{k_i+1})$ for some index k_i, then either $\theta_i' < \lambda_{k_i}$ (unless $\theta_i' < \lambda_{i+1}$) or it holds that

$$\Delta_{k_i,k_i+1}(\theta_i') \leq \left(\gamma + (1-\gamma)\frac{\lambda_{k_i}}{\lambda_{k_i+1}}\right)^2 \Delta_{k_i,k_i+1}(\theta_i), \quad i = 1, \ldots, s, \tag{16}$$

where

$$\Delta_{k_i,k_i+1}(\kappa) = \frac{\kappa - \lambda_{k_i}}{\lambda_{k_i+1} - \kappa}.$$

Proof. First we show that full rank of V in (13) implies full rank of $\tilde{V}$. For any nonzero $y \in \mathbb{R}^s$ we have

$$\begin{aligned}
\|\tilde{V}y\|_A &= \|A^{-1}MV\Theta y + (I - B^{-1}A)(V - A^{-1}MV\Theta)y\|_A \\
&\geq \|A^{-1}MV\Theta y\|_A - \|(I - B^{-1}A)(V - A^{-1}MV\Theta)\|_A \\
&\geq \|A^{-1}MV\Theta y\|_A - \|(V - A^{-1}MV\Theta)y\|_A \\
&= \frac{\|A^{-1}MV\Theta y\|_A^2 - \|(V - A^{-1}MV\Theta)y\|_A^2}{\|A^{-1}MV\Theta y\|_A + \|(V - A^{-1}MV\Theta)y\|_A} \\
&= \frac{\|Vy\|_A^2}{\|A^{-1}MV\Theta y\|_A + \|(V - A^{-1}MV\Theta)y\|_A} > 0,
\end{aligned}$$

where we have used the relation $V^TA(V - A^{-1}MV\Theta) = 0$ in the last step. Since $\operatorname{rank}(V) = s$, we have $\|Vy\|_A > 0$ and $\|A^{-1}MV\Theta y\|_A > 0$ so that $\operatorname{rank}(\tilde{V}) = s$.

To show (16), a Cholesky decomposition of $M = CC^T$ is used to transform (2) to the standard eigenvalue problem for $\bar{A} = C^{-1}AC^{-T}$. Then (2) with $y_i = C^Tx_i$ reads

$$\bar{A}y_i = \lambda_i y_i, \qquad i = 1, \ldots, n, \tag{17}$$

while the Rayleigh quotient (3) transforms to

$$\lambda_{\bar{A}}(y) = \frac{(y, \bar{A}y)}{(y, y)}.$$

The scheme (13) with $\bar{V} = C^TV$ reads

$$C^T\tilde{V} = \bar{A}^{-1}\bar{V}\Theta + (I - \bar{B}^{-1}\bar{A})(\bar{V} - \bar{A}^{-1}\bar{V}\Theta), \tag{18}$$

where the symmetric, positive definite $\bar{B}^{-1} = C^TB^{-1}C$ fulfills

$$\|I - \bar{B}^{-1}\bar{A}\|_{\bar{A}} \leq \gamma,$$

in the operator norm induced by $\bar{A}$. Hence Equation (18) is nothing else than the preconditioned eigensolver for standard eigenvalue problem (17) and Theorem 3.3 in [20] as well as Theorem 1 in [17] can be applied. Finally, the Ritz values of $C^T\tilde{V}$ with respect to $\bar{A}$ equal those of $\tilde{V}$ with respect to (A, M); this proves the bound (16) for the Ritz values. □

Theorem 2.1 guarantees at least linear convergence for each Ritz values toward a smaller eigenvalue. We emphasize that the convergence estimates for θ_i only depend on the nearest eigenvalues λ_{k_i} and λ_{k_i+1} enclosing θ_i. Hence, the convergence estimate is independent of the mesh size and so of the number of the unknowns. This guarantees a grid independent convergence as known from preconditioned solvers for the linear systems arising from discretizing boundary value problems.

2.5. More general preconditioners

In practice, assumption (7) on the quality of the preconditioner is often satisfied for scaled preconditioners only. Whenever we assume (6) to hold for positive constants γ_0 and γ_1, we have as the smallest bound for the scaled preconditioner ϑB^{-1}

$$\|I - \vartheta B^{-1} A\|_A \leq \frac{\gamma_1 - \gamma_0}{\gamma_0 + \gamma_1} =: \gamma < 1, \tag{19}$$

with the scaling constant $\vartheta = 2/(\gamma_0 + \gamma_1)$. In general ϑ is not available computationally. But the situation is not hopeless since optimal subspace approximations can be determined from

$$[V, B^{-1}(AV - MV\Theta)] \in \mathbb{R}^{n \times 2s}$$

by means of the Rayleigh–Ritz method where the results do not depend on the scaling of the preconditioner. This iteration is called preconditioned steepest descent and is guaranteed to converge faster than PINVIT (the faster decrease of the Rayleigh quotients is assured by the Courant–Fischer principle).

3. A posteriori error estimation

A posteriori error estimation for the given eigenproblem is based on an estimate from above for the deviation of each Ritz value from the nearest eigenvalue of problem (2). This estimate derives from evaluating a suitable norm of the gradient of the Rayleigh quotient within the given Ritz vector. The estimator is easily computable from known quantities. In contrast to a posteriori error estimation for the numerical solution of boundary value problems, our a posteriori error estimator works on a *subspace* of eigenvector approximations. This gives us the freedom to couple error estimation to a specific single eigenfunction or to a subset of eigenfunctions in order to compute these with an increased accuracy. We will discuss some strategies in Section 5.

We do not try to bound the error of the Ritz vectors. Such an error, for instance, can be measured by the canonical angles enclosed by the subspace of exact eigenfunctions and its numerical approximation. The main difficulty with such a bound is to be seen in the fact that the acute angles enclosed by the eigenvector approximations generated by PINVIT and the eigenspace to the smallest eigenvalues are not monotone decreasing, see Sec. 3.2 in [22].

Given a certain Ritz pair (v_i, θ_i), we employ a Taylor expansion of the Rayleigh quotient in order to estimate the distance of θ_i to the nearest eigenvalue of A. If $\tilde{v}_i$ is the i–th column of $\tilde{V}$, as defined by (14), then the first order Taylor expansion of (3) reads

$$\lambda(\tilde{v}_i) = \lambda(v_i - d_i) = \lambda(v_i) - (\nabla\lambda(v_i), d_i) + O(\|d_i\|^2).$$

Since $\nabla\lambda(v) = \frac{2}{(v,Mv)}\,(Av - \lambda(v)Mv)$, we can rewrite the negative first order term F_i in the form

$$F_i := (\nabla\lambda(v_i), d_i) = \frac{2}{(v_i, Mv_i)}(r_i, d_i) = 2\|d_i\|_B^2. \tag{20}$$

Therein, $r_i = Av_i - \lambda(v_i)Mv_i$ denotes the residual of v_i. To compute F_i only a single dot product is required, since the residual r_i and the preconditioned residual d_i are known while computing $B^{-1}(AV - MV\Theta)$ and $\tilde{V}$ by Equation (14). Therefore, F_i can serve as a simple estimate for $\lambda(v_i) - \lambda(\tilde{v}_i)$, i.e. the decrease of the Rayleigh quotient in the i–th column of V for a step of (14).

In Theorem 3.1 we show that F_i can even serve as an error estimator for the larger distance of the Ritz value θ_i to the nearest eigenvalue of A.

Theorem 3.1. *Let v_i be the i–th Ritz vector, i.e. the i–th column of V, and let m be the index so that $\theta_i \in [\lambda_m, \lambda_{m+1})$. Then it holds that*

$$(\theta_i - \lambda_m)(\lambda_{m+1} - \theta_i) \leq \frac{\lambda_m\lambda_{m+1}}{2\theta_i(1-\gamma)}F_i \leq \frac{\lambda_{m+1}}{2(1-\gamma)}F_i, \tag{21}$$

where the error estimator F_i is given by Equation (20).

Proof. It is first shown that for any v, $\|v\|_M = 1$, with $\lambda = \lambda(v) \in [\lambda_m, \lambda_{m+1})$ and $r = Av - \lambda Mv$ the following Temple–type inequality holds (see D'yakonov and Orekhov [12])

$$\frac{\lambda(\lambda - \lambda_m)(\lambda_{m+1} - \lambda)}{\lambda_m\lambda_{m+1}} \leq \|r\|_{A^{-1}}^2. \tag{22}$$

To prove (22) one starts from a basic inequality

$$(\lambda_m - \lambda_j)(\lambda_{m+1} - \lambda_j) \geq 0, \qquad j = 1, \ldots, n,$$

which leads to

$$\frac{1}{\lambda_j} \geq \frac{\lambda_m + \lambda_{m+1} - \lambda_j}{\lambda_m\lambda_{m+1}}.$$

Expanding v in eigenvectors of A, that is $v = \sum_{j=1}^n c_j x_j$, one directly obtains $\|v\|_M^2 = \sum_{j=1}^n c_j^2 = 1$ and thus

$$\sum_{j=1}^n \frac{1}{\lambda_j}c_j^2 \geq \sum_{j=1}^n \frac{\lambda_m + \lambda_{m+1} - \lambda_j}{\lambda_i\lambda_{i+1}}c_j^2 = \frac{\lambda_m + \lambda_{m+1} - \lambda}{\lambda_m\lambda_{m+1}},$$

or

$$\lambda\left(\lambda\sum_{j=1}^n \frac{c_j^2}{\lambda_j} - 1\right) \geq \frac{\lambda(\lambda - \lambda_m)(\lambda_{m+1} - \lambda)}{\lambda_m\lambda_{m+1}}.$$

Showing that the left hand side of the last inequality is equal to $\|r\|_{A^{-1}}^2$ completes the proof of (22). Inserting v_i and θ_i one obtains by using (7)

$$\frac{\theta_i(\theta_i - \lambda_m)(\lambda_{m+1} - \theta_i)}{\lambda_m\lambda_{m+1}} \leq \|r_i\|_{A^{-1}}^2 \leq \frac{1}{1-\gamma}\|r_i\|_{B^{-1}}^2 = \frac{1}{1-\gamma}\|d_i\|_B^2.$$

□

To form a judgment on the quality of the error estimator one has to distinguish between the influence of the preconditioner, which is fully controlled by (7) and γ, and the error introduced by the inequality (22). Let us show that $\|r\|_{A^{-1}}$ in Equation (22) is also bounded from above. By using Theorem 2.1 in [22] (the proof is based on the Lagrange multiplier method) one has for $\lambda \in [\lambda_m, \lambda_{m+1})$

$$B_{m,m+1}(\lambda) \leq \|r\|_{A^{-1}} \leq B_{1,n}(\lambda),$$

where $B_{i,j}^2(\lambda) := \lambda(\lambda/\lambda_i - 1)(1 - \lambda/\lambda_j)$. It is easy to show that the bound $B_{i,j}$ is attained if the vector v, whose Rayleigh quotient is assumed to be equal to λ and whose residual is given by $r = Av - \lambda v$, is spanned only by eigenvectors to the eigenvalues λ_i and λ_j. But in general there is a wide corridor between the curves $B_{i,i+1}(\lambda)$ and $B_{1,n}(\lambda)$. We obtain for $\lambda \in [\lambda_i, \lambda_j]$

$$\frac{\partial}{\partial \lambda_i} B_{i,j}^2(\lambda) < 0, \qquad \frac{\partial}{\partial \lambda_j} B_{i,j}^2(\lambda) > 0,$$

and $\lim_{\lambda_n \to \infty} B_{1,n}^2(\lambda) = \lambda^2/\lambda_1 - 1$ so that $B_{1,n}$ appears as an extremely bad bound. But $B_{1,n}$ is too pessimistic, since the more the actual iterate v approximates some eigenvector of A (which will be the case, sooner or later), the better the influence of eigenvectors to other (larger) eigenvalues vanishes and the error estimator (21) will provide a more reliable bound, cf. the numerical results in Section 5.

4. The iteration and discretization error estimators

To construct an adaptive multigrid algorithm for the approximate solution of the elliptic eigenproblem (1) one requires both an *iteration error* estimator as well as a *discretization error* estimator. While the iteration error estimator should provide practical criteria for the termination of the preconditioned eigensolver (with respect to the actual grid), the discretization error estimator should make available local error indicators to control the mesh refinement.

The *iteration error* estimator is simply based on the evaluation of F_i within the given finite element space $\mathcal{S}$, in which the matrix eigenvalue problem (2) is the discretization of the given elliptic eigenproblem. Let us now assume that the Ritz value θ_i is of sufficient quality so that $\theta_i \in [\lambda_i, \lambda_{i+1})$. Then Equation (21) for $m = i$ reads

$$\frac{2\theta_i(1-\gamma)}{\lambda_i \lambda_{i+1}} (\theta_i - \lambda_i)(\lambda_{i+1} - \theta_i) \leq F_i. \tag{23}$$

We emphasize that the error estimator F_i results from known quantities (the residual and the preconditioned residual) by evaluating a simple dot product. Aside from the constants on the left hand side of (23) a small estimate F_i indicates that actual Ritz value θ_i is near to the eigenvalue λ_i.

To estimate the *discretization error* we introduce an enlarged subspace $\mathcal{Q} \supset \mathcal{S}$ of higher order finite elements. The eigenvalues and the Ritz values in these spaces are denoted as follows: The eigenvalues $\bar{\lambda}_i$ are those of the continuous problem (1) and the eigenvalues $\lambda_{i,\mathcal{S}}$ ($\lambda_{i,\mathcal{Q}}$) are defined by (2) if the continuous eigenproblem is discretized with respect to $\mathcal{S}$ ($\mathcal{Q}$). Finally, the Ritz values θ_i are approximations of the eigenvalues $\lambda_{i,\mathcal{S}}$. Then we have

$$\bar{\lambda}_i \leq \lambda_{i,\mathcal{Q}} \leq \lambda_{i,\mathcal{S}} \leq \theta_i,$$

since $\mathcal{S} \subset \mathcal{Q} \subset H(\Omega)$.

In the following we use a saturation assumption in terms of eigenvalues

$$\lambda_{i,\mathcal{Q}} - \bar{\lambda}_i \leq \beta(\lambda_{i,\mathcal{S}} - \bar{\lambda}_i) \tag{24}$$

with a positive $\beta < 1$. The Courant–Fischer principles guarantee that $\beta \leq 1$ and the theory presented in Chapter 6 of [24] shows that β is bounded away from 1 if we assume, for instance, $\mathcal{S}$ ($\mathcal{Q}$) to consist of piecewise linear (quadratic) finite elements and if the triangulation is fine enough. The next lemma (cf. Lemma 2.1 in [4]) shows that the total error $\theta_i - \bar{\lambda}_i$ and the discretization error $\lambda_{i,\mathcal{S}} - \bar{\lambda}_i$ can be estimated from above by $\theta_i - \lambda_{i,\mathcal{Q}}$. To estimate the latter difference we later apply our error estimator.

Lemma 4.1. *The saturation assumption (24) is equivalent to each of the following inequalities:*

$$\theta_i - \bar{\lambda}_i \leq \frac{1}{1-\beta}(\theta_i - \lambda_{i,\mathcal{Q}}), \tag{25}$$

$$\lambda_{i,\mathcal{S}} - \bar{\lambda}_i \leq \frac{1}{1-\beta}(\lambda_{i,\mathcal{S}} - \lambda_{i,\mathcal{Q}}). \tag{26}$$

Proof. First we show that the saturation assumption implies (25) and (26). Since

$$\theta_i - \lambda_{i,\mathcal{Q}} = \theta_i - \bar{\lambda}_i + \bar{\lambda}_i - \lambda_{i,\mathcal{Q}} \geq \theta_i - \bar{\lambda}_i - \beta(\lambda_{i,\mathcal{S}} - \bar{\lambda}_i) \geq (1-\beta)(\theta_i - \bar{\lambda}_i),$$

we have (25). For $\theta_i = \lambda_{i,\mathcal{S}}$ the inequality (26) follows immediately. To see that the saturation assumption follows from (26) we calculate

$$\lambda_{i,\mathcal{S}} - \bar{\lambda}_i = \lambda_{i,\mathcal{Q}} - \bar{\lambda}_i + \lambda_{i,\mathcal{S}} - \lambda_{i,\mathcal{Q}} \geq \lambda_{i,\mathcal{Q}} - \bar{\lambda}_i + (1-\beta)(\lambda_{i,\mathcal{S}} - \bar{\lambda}_i)$$

from which we obtain (24). □

To be concrete we assume $\mathcal{S}$ to consist of piecewise linear finite elements and realize the discretization error estimator in $\mathcal{Q}$, which is assumed to be the finite element space of piecewise quadratic elements. A hierarchical decomposition of $\mathcal{Q}$ is given by the direct sum

$$\mathcal{Q} = \mathcal{S} \oplus \mathcal{V}, \tag{27}$$

where the piecewise quadratic functions in $\mathcal{V}$ vanish on all nodes defined by the grid associated with $\mathcal{S}$. For a more detailed description of $\mathcal{Q}$ see [8].

To estimate the discretization error within a given Ritz pair (v_i, θ_i) we write down the discretization of the continuous eigenvalue problem (1) with respect to the hierarchical extension (27). We obtain, instead of (2), the following generalized eigenvalue problem for the 2×2 block matrices $A_\mathcal{Q}$ and $M_\mathcal{Q}$

$$\begin{pmatrix} A & A_{\mathcal{S}\mathcal{V}} \\ A_{\mathcal{V}\mathcal{S}} & A_{\mathcal{Q}\mathcal{Q}} \end{pmatrix} x_{i,\mathcal{Q}} = \lambda_{i,\mathcal{Q}} \begin{pmatrix} M & M_{\mathcal{S}\mathcal{V}} \\ M_{\mathcal{V}\mathcal{S}} & M_{\mathcal{Q}\mathcal{Q}} \end{pmatrix} x_{i,\mathcal{Q}}.$$

The $(1,1)$ submatrices A (M) of $A_\mathcal{Q}$ ($M_\mathcal{Q}$) coincide with those in (2). We further define $v_{i,\mathcal{Q}} = (v_i, 0)^T$, whose residual $r_{i,\mathcal{Q}}$ in $\mathcal{Q}$ is given by

$$r_{i,\mathcal{Q}} = A_\mathcal{Q} v_{i,\mathcal{Q}} - \theta_i M_\mathcal{Q} v_{i,\mathcal{Q}} = \begin{pmatrix} Av_i - \theta_i M v_i \\ A_{\mathcal{V}\mathcal{S}} v_i - \theta_i M_{\mathcal{V}\mathcal{S}} v_i \end{pmatrix} \tag{28}$$

Now assume a symmetric positive definite preconditioner $B_\mathcal{Q}$ to be given so that the spectral radius of $I - B_\mathcal{Q}^{-1} A_\mathcal{Q}$, in analogy to (7), is equal to $\gamma_\mathcal{Q} < 1$. Applying the error estimator (21) within $\mathcal{Q}$ results in

$$F_{i,\mathcal{Q}} := 2\|d_{i,\mathcal{Q}}\|^2_{B_\mathcal{Q}}, \tag{29}$$

where $d_{i,\mathcal{Q}} = (B_\mathcal{Q})^{-1} r_{i,\mathcal{Q}}$ is the preconditioned residual.

If $\theta_i \in [\lambda_{i,\mathcal{Q}}, \lambda_{i+1,\mathcal{Q}})$, then the error estimator $F_{i,\mathcal{Q}}$ allows to estimate the term $\theta_i - \lambda_{i,\mathcal{Q}}$ on the right hand side of (25) and thus implicitly the discretization error in the form

$$\frac{2\theta_i(1-\gamma_\mathcal{Q})}{\lambda_{i,\mathcal{Q}}\lambda_{i+1,\mathcal{Q}}}(\theta_i - \lambda_{i,\mathcal{Q}})(\lambda_{i+1,\mathcal{Q}} - \theta_i) \le F_{i,Q}. \tag{30}$$

5. Numerical results

The *a posteriori* error estimators have been integrated within an adaptive multigrid algorithm to solve elliptic eigenproblems [19]. Starting with an initial coarse grid, the subspace variant of PINVIT is applied on each grid of a sequence of adaptively generated refined grids. Especially for only weakly differentiable eigenfunctions we obtain a highly nonuniform grid, maximally refined in the critical region near the origin, exemplifying the efficacy of the adaptive scheme.

The iteration error estimator is used to define a stopping criterion for (14) (with respect to the actual subspace) and the discretization error estimator is employed to derive a mesh refinement strategy. The central idea, which motivates the derivation of the stopping condition, is that the iteration error and the discretization error on the sequence of adaptively generated grids should be reduced in a balanced ratio, cf. [8] who use this concept for the numerical solution of *boundary value problems*. Such a balancing is reasonable, since solving the eigenproblem on a given mesh with high accuracy is ineffective if the discretization error is still large. If on the other hand the iteration error with respect to the actual grid is large, then the local error indicators, which are responsible for the mesh refinement, are expected to be of low accuracy. For the details of this implementation see the technical report [19]. In [18] the concept is applied to differential operators in the $\mathbb{R}^3$.

As a test problem we consider the weak form of the Laplacian eigenproblem

$$-\Delta u = \lambda u,$$

on the unit circle, centered at the origin, with a slit along the positive x axis. Homogeneous Dirichlet boundary conditions are supposed on the boundary of the circle and on the top of the slit, while homogeneous Neumann boundary conditions are given on the bottom of the slit.

The numerical results can be compared with the exact eigenfunctions and eigenvalues. In polar coordinates with $\theta \in [0, 2\pi)$ and $r \in [0, 1]$ these are given by

$$u_{k,l}(r,\theta) = c\sin(\alpha\theta) J_\alpha(\omega_{\alpha,l} r), \qquad \lambda_{k,l} = (\omega_{\alpha,l})^2,$$

where $\alpha = \frac{1}{4} + \frac{k}{2}$ and for integer separation constants $k \ge 0$ and $l \ge 1$. Therein, $J_\alpha(r)$ is the Bessel function of first kind and fractional order α, see [1], whose positive zeros $\omega_{\alpha,l}$ are assumed to increase with l. Table I gives the 8 smallest eigenvalues. We further note that for $k = 0$ and $k = 1$ the derivative $\frac{\partial}{\partial r} J_\alpha$ is unbounded at $r = 0$ so that the eigenfunctions $u_{0,l}$ and $u_{1,l}$ are only in $H^1(\Omega)$.

	k	l	$\lambda_{k,l}$	$\frac{\partial}{\partial r}J_\alpha$ unbounded
1.	0	1	7.73333	+
2.	1	1	12.18714	+
3.	2	1	17.35078	-
4.	3	1	23.19939	-
5.	4	1	29.71453	-
6.	0	2	34.88252	+
7.	5	1	36.88189	-
8.	1	2	44.25756	+

Table I. *Separation constants k, l and eigenvalues rounded to five digits after the period. For $k = 0, 1$ the derivative of the Bessel function $J(\alpha, r)$ is unbounded at the origin.*

	$P = \{1\}$			$P = \{3\}$			$P = \{1,2,3\}$		
i	θ_i	λ_i	$\lambda_{i,\mathcal{Q}}$	θ_i	λ_i	$\lambda_{i,\mathcal{Q}}$	θ_i	λ_i	$\lambda_{i,\mathcal{Q}}$
1	7.777	7.769	7.744	8.219	8.116	7.934	7.796	7.786	7.751
2	12.244	12.243	12.197	12.244	12.237	12.194	12.230	12.230	12.192
3	17.462	17.460	17.365	17.422	17.418	17.363	17.422	17.422	17.359

Table II. *Final Ritz values θ_i in $\mathcal{S}$ and stationary values λ_i ($\lambda_{i,\mathcal{Q}}$) in $\mathcal{S}$ ($\mathcal{Q}$).*

All test calculations have been performed within an eight dimensional subspace ($s = 8$). As a preconditioner the (scaled) hierarchical basis preconditioner is used, see Yserentant [26, 27]. The spectral properties of $B_\mathcal{Q}$ in $\mathcal{Q}$ are described in [8].

Since PINVIT is applied to an s dimensional subspace, the termination criterion with respect to the actual grid as well as the grid refinement strategy can be coupled to a certain subset $P \subseteq \{1, \ldots, s\}$ containing indexes of critical eigenfunctions of low regularity. Then the adaptive multigrid eigensolver should give best results for those eigenfunctions whose indexes are contained in P. To demonstrate the dependence of the generated grid on the index set P, we make three different choices.

1. $P = \{1\}$. To approximate $u_{0,1}$ with an unbounded partial derivative at the origin a highly non–uniform triangulation is expected.
2. $P = \{3\}$. The derivatives of $u_{2,1}$ are bounded. Thus the triangulation will be more uniform.
3. $P = \{1, 2, 3\}$. The final triangulation should reflect the properties of both cases above. The influence of the second eigenfunction $u_{1,1}(r, \theta)$ at the critical origin is insignificant since $u_{1,0}(r, \theta)$ is the steeper function in r in the neighborhood of the origin.

In Figure 1 the initial triangulation consisting of 40 triangles is shown where on only 10 inner nodes the function is not fixed to zero by the boundary conditions. Additionally, contour plots of the first three eigenfunctions are given. Table II lists the three smallest final Ritz values θ_i on the final grid for which the subspace eigensolver is stopped. Table II also contains the eigenvalue approximations λ_i ($\lambda_{i,\mathcal{Q}}$) for which the iteration is stationary with respect to the final finite element space of piecewise linear (quadratic) finite elements. Figure 2 shows the

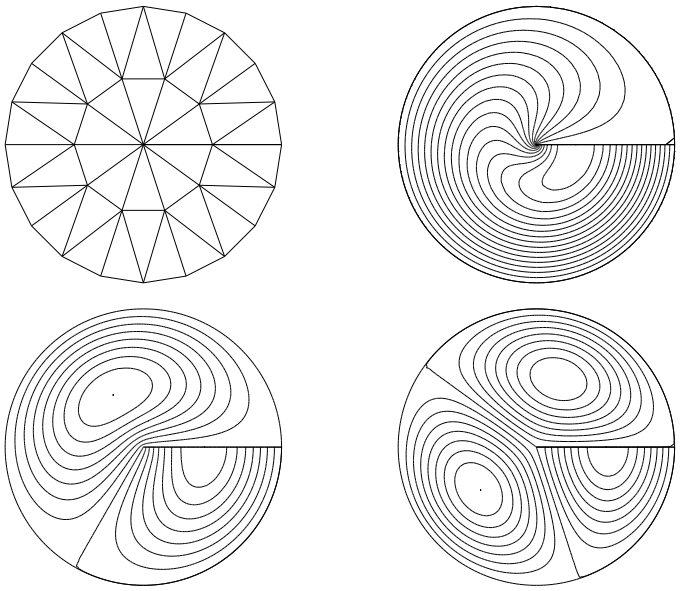

Figure 1. $\boxed{\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}}$ *a) Initial triangulation of the unit circle consisting of* 40 *triangles and 10 inner nodes. b)–d) Contour plots of* $u_{0,1}$, $u_{1,1}$, $u_{2,1}$.

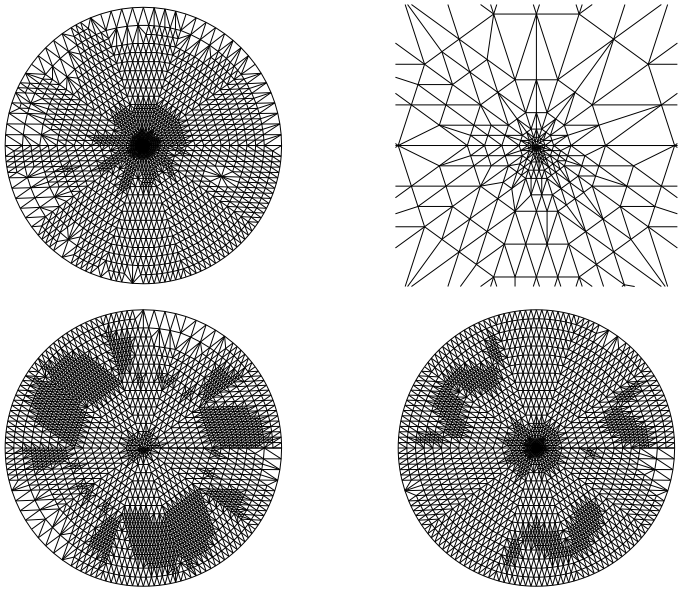

Figure 2. $\boxed{\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}}$ *Final triangulations. a)* $P = \{1\}$, *2385 nodes. b) Zoom of* $[-2^{-10}, 2^{-10}]^2$ *in (a). c)* $P = \{3\}$, *2374 nodes. d)* $P = \{1, 2, 3\}$, *2381 nodes.*

	# nodes	# inner nodes	$t_{\min}$	$t_{\max}$
$P = \{1\}$	2385	2237	3	16
$P = \{3\}$	2374	2215	4	6
$P = \{1, 2, 3\}$	2381	2201	4	14

Table III. *Number of nodes of the final triangulation and number of nodes that are not on the Dirichlet boundary.* $t_{\min}$ *is the number of uniform grid refinements with respect to the initial triangulation and* $t_{\max}$ *the maximal depth of the triangulation.*

final triangulations for each choice of P. The corresponding grid parameters are listed in Table III. Highly non–uniform triangulations with a final depth of 16 (14) are gained for the first (third) choice of P giving evidence for the efficacy of the adaptive scheme. As anticipated, the grid is maximally refined near to the origin.

Comparing the values λ_1 in Table II one observes that for $P = \{1\}$ this stationary value is about 4% smaller than for $P = \{3\}$ and only about 0.2% smaller than for $P = \{1, 2, 3\}$. Similar results for λ_3 and $P = \{3\}$ in comparison with the other choices of P are gained. Apparently, $P = \{i\}$ is the best strategy to generate a grid on which λ_i achieves its minimum. The final Ritz values θ_i are approximations of the eigenvalues of the continuous problem within the 1 percent range. The remaining error is mainly caused by the discretization error. The calculation with $P = \{1, 2, 3\}$ leads to a grid which allows comparatively good approximations of *all* three eigenvalues. The depth of the final triangulation is close to the depth in the case $P = \{1\}$.

The a posteriori error estimators in the case $P = \{1\}$ are now studied in more detail. Figure 3 documents the iteration error of the Ritz values by plotting $\theta_i^{(k)} - \lambda_i$, for $i = 1, 2, 3$, for the computations described above. The Ritz values $\theta_i^{(k)}$ are those which are generated by the PINVIT subspace scheme within the adaptive process. The iteration error of the third eigenvalue is the largest one (despite of the final grid). Each stair in the plot corresponds to the appearance of a new grid and on each grid the iteration error is reduced by the rate of a geometric progression, see [20].

The reduction of the discretization error is illustrated by plotting the difference $\theta_1^{(k)} - \lambda_{1,\mathcal{Q}}$ (solid line) in Figure 4. Additionally, the dotted line in Figure 4 represents the error estimator $F_{i,\mathcal{Q}}$, see Equation (29), which roughly behaves like the error in $\mathcal{Q}$ but tends to underestimate the real error. To make sure that the error of all iterates is measured with the same norm, all Ritz vectors generated in the course of the adaptive process have been prolongated to the final grid. After this, error estimation is applied to all iterates. But note that this is an artificial situation since in the adaptive process the error estimator is only called with respect to the actual finite element space and will then provide more accurate estimates.

ACKNOWLEDGEMENT

The author is grateful to P. Leinen and W. Lembach for providing several parts of the adaptive eigensolver code.

REFERENCES

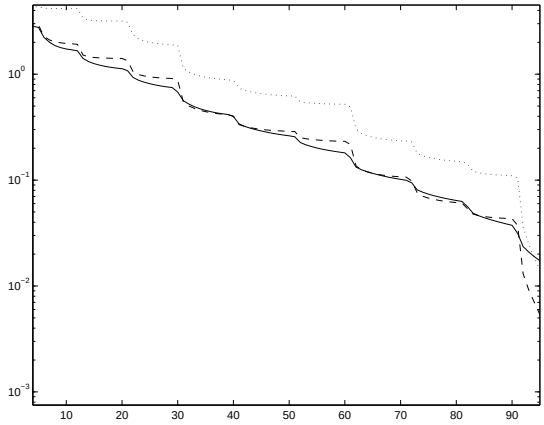

Figure 3. *Convergence of the eigenvalues* $\theta_i^{(k)} - \lambda_i$ *for* $P = \{1\}$. *Abscissa: Iteration number of (14). Ordinate: Solid line* $i = 1$ *solid line; broken line* $i = 2$*; dotted line* $i = 3$.

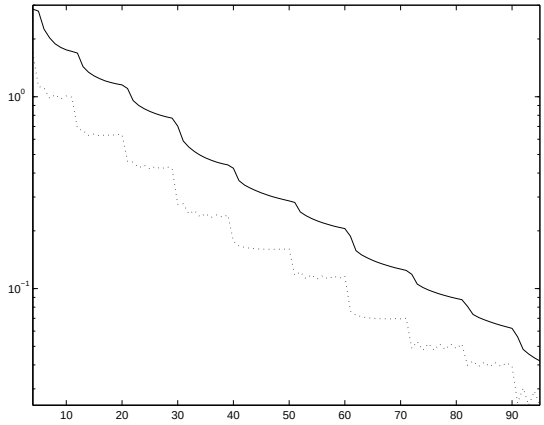

Figure 4. *Error estimation for the case* $P = \{1\}$ *against the iteration number of (14). Estimator* $F_{i,\mathcal{Q}}$ *dotted line. Solid line* $\theta_i^{(k)} - \lambda_{1,\mathcal{Q}}$.

1. G. Arfken. *Mathematical methods for physicists.* Academic Press, San Diego, 1985.
2. R.E. Bank. *PLTMG: A software package for solving elliptic partial differential equations. Users' guide 7.0.* SIAM Books, Philadelphia, 1994.
3. R.E. Bank and A. Weiser. Some a–posteriori error estimators for elliptic partial differential equations. *Mathematics of Computation*, 44:283–301, 1985.
4. F.A. Bornemann, B. Erdmann, and R. Kornhuber. A posteriori error estimation for elliptic problems in two and three space dimensions. *SIAM Journal on Numerical Analysis*, 33:1188–1204, 1996.
5. J.H. Bramble, J.E. Pasciak, and A.V. Knyazev. A subspace preconditioning algorithm for eigenvector/eigenvalue computation. *Advances in Computational Mathematics*, 6:159–189, 1996.
6. W.W. Bradbury and R. Fletcher. New iterative methods for solution of the eigenproblem. *Numerische Mathematik*, 9:259–267, 1966.
7. F. Chatelin. *Eigenvalues of matrices.* Wiley, Chichester, 1993.

8. P. Deuflhard, P. Leinen, and H. Yserentant. Concepts of an adaptive hierarchical finite element code. *Impact of Computing in Science and Engineering*, 1:3–35, 1989.
9. E.G. D'yakonov. *Optimization in solving elliptic problems.* CRC Press, Boca Raton, Florida, 1996.
10. E.G. D'yakonov and A. V. Knyazev. Group iterative method for finding lower-order eigenvalues. *Moscow University Computational Mathematics and Cybernetics*, 2:32–40, 1982.
11. E.G. D'yakonov and A. V. Knyazev. On an iterative method for finding lower eigenvalues. *Russian Journal of Numerical Analysis and Mathematical Modelling* , 7(6):473–486, 1992.
12. E.G. D'yakonov and M.Y. Orekhov. Minimization of the computational labor in determining the first eigenvalues of differential operators. *Mathematical Notes*, 27:382–391, 1980.
13. P.O. Friberg. An error indicator for the generalized eigenvalue problem using the hierarchical finite element method. *International Journal for Numerical Methods in Engineering*, 23:91–98, 1986.
14. M.R. Hestenes and W. Karush. A method of gradients for the calculation of the characteristic roots and vectors of a real symmetric matrix. *Journal of Research of the National Bureau of Standards*, 47:45–61, 1951.
15. L. V. Kantorovich. *Functional analysis and applied mathematics.* U. S. Department of Commerce National Bureau of Standards, Los Angeles, California, 1952. Translated by C. D. Benster.
16. A.V. Knyazev. Preconditioned eigensolvers -an oxymoron? *Electronic Transactions on Numerical Analysis*, 7:104–123, 1998.
17. A.V. Knyazev and K. Neymeyr. A geometric theory for preconditioned inverse iteration. III: A short and sharp convergence estimate for generalized eigenvalue problems. Accepted for Linear Algebra and Its Applications, 2001.
18. W. Lembach. *Adaptive Mehrgitterverfahren zur Lösung von Eigenwertproblemen elliptischer Differentialoperatoren.* PhD thesis, Universität Tübingen, 1998.
19. K. Neymeyr. A posteriori error estimation for a preconditioned algorithm to solve elliptic eigenproblems. Technical Report 77, Sonderforschungsbereich 382, Universitäten Tübingen und Stuttgart, 1997.
20. K. Neymeyr. A geometric theory for preconditioned inverse iteration applied to a subspace. Sonderforschungsbereich 382, Universität Tübingen, Report 131, 1999, Accepted for Mathematics of Computation, 1999.
21. K. Neymeyr. A geometric theory for preconditioned inverse iteration. I: Extrema of the Rayleigh quotient. *Linear Algebra and Its Applications*, 322:61–85, 2001.
22. K. Neymeyr. A geometric theory for preconditioned inverse iteration. II: Convergence estimates. *Linear Algebra and Its Applications*, 322:87–104, 2001.
23. B.N. Parlett. *The symmetric eigenvalue problem.* Prentice Hall, Englewood Cliffs New Jersey, 1980.
24. G. Strang and G.J. Fix. *An analysis of the finite element method.* Prentice–Hall, Englewood Cliffs, N.J., 1973.
25. R. Verfürth. *A review of a posteriori error estimation and adaptive mesh–refinement techniques.* Wiley and Teubner, New York and Stuttgart, 1995.
26. H. Yserentant. On the multi–level splitting of finite element spaces. *Numerische Mathematik*, 49:379–412, 1986.
27. H. Yserentant. Hierarchical bases. In O'Malley and R.E., editors, *ICIAM 91*, pages 256–276, Washington, DC, 1992. SIAM, Philadelphia.
28. O.C. Zienkiewicz, D.W. Kelley, S.R. Gago, and I. Babuška. Hierarchical finite element approaches, error estimates and adaptive refinement. In *The mathematics of finite elements and applications IV*, pages 313–346. Academic Press, New York, 1982.