SOLVING MESH EIGENPROBLEMS WITH MULTIGRID EFFICIENCY

KLAUS NEYMEYR

ABSTRACT. Multigrid techniques can successfully be applied to mesh eigenvalue problems for elliptic differential operators. They allow to compute a few of the smallest eigenvalues and the eigenfunctions within optimal or quasi-optimal computational complexity. A review is given on three multigrid solvers for elliptic eigenproblems. Numerical results are presented for a simple model problem.

1. INTRODUCTION

Consider the eigenvalue problem for an elliptic partial differential operator. Numerical approximations of its eigenvalues and eigenfunctions can be gained by a finite element discretization of the variational problem associated with the differential operator. The discretization results in a generalized matrix eigenvalue problem of the form

$$Ax = \lambda Mx. \tag{1.1}$$

The discretization matrix A and the mass matrix M are sparse with only a small number of nonzero elements per row. In typical applications, the dimension of these matrices may exceed $10^6 - 10^9$ and, at the same time, only a few of the smallest eigenvalues (in modulus) together with the eigenvectors are to be determined. These smallest eigenvalues characterize, for example, the base frequencies of a vibrating mechanical structure modeled by an eigenvalue problem for an elliptic partial differential operator.

Most of the "classical" methods of numerical linear algebra for solving a generalized matrix eigenvalue problem cannot be applied to (1.1) because typically, the computer storage for a full representation of A and M or any factorization of these matrices is not available. The QR method, the Rayleigh quotient iteration or inverse iteration, for example, would all demand to generate matrix factorizations and to store the more dense matrix factors. In any case, none of these iterations would allow to determine an eigenpair within $\mathcal{O}(n)$ floating point operations for A and M being $n \times n$ matrices. In contrast to this, *boundary value problems* for the same elliptic partial differential operators can be solved with optimal complexity or quasi-optimal computational complexity (i.e. $\mathcal{O}(n)$ or $\mathcal{O}(n \log n)$) by means of multigrid or domain decomposition methods.

The aim of this paper is to rivet on eigensolvers for *elliptic eigenvalue problems* with "multigrid efficiency", i.e. with optimal or quasi-optimal complexity. In Section 2 we first point out common elements of eigensolvers for mesh eigenvalue problems and then we briefly review and classify the following eigensolvers:

(I) The *Rayleigh quotient multigrid minimization*, [4, 5, 16],
(II) *Direct multigrid eigensolvers*, [3, 7, 8, 10],
(III) *Eigensolvers using multigrid as a linear solver*, [14, 19].

Finally, Section 3 contains numerical results for the Laplacian in 2D.

2. MULTIGRID EIGENSOLVERS

In order to give a concise description of multigrid eigensolvers we first separate those parts which are common to any such iteration:

(1) *Subspace extension*: Any eigensolver for computing only a single eigenvalue/eigenvector pair (typically the smallest eigenvalue), can simply be extended to a subspace iteration in order to determine a modest number of the smallest eigenvalues and eigenvectors. One can implement a subspace eigensolver by applying a given vector iteration to each of the Ritz vectors spanning an actual approximating subspace. Subsequent application of the Rayleigh-Ritz procedure serves to compute the new Ritz values and Ritz vectors. Alternatively, subspace computations can be realized by using the less stable deflation technique.

(2) *Nested iteration and adaptivity*: The concepts of *nested iteration* and *adaptivity*, which are well known from boundary value problem solvers, can be transferred to the eigenvalue problem. Then nested iteration starts with the solution of a coarse grid eigenproblem. The following steps are successive prolongations of the actual eigenvector approximations to refined grids in combination with sufficiently many iterations of the eigensolver on each grid level.

Nested iteration can be combined with the concept of adaptivity. By means of adaptively generated grids, numerical approximation of eigenvalues and eigenvectors within a prescribed tolerance can often be gained with only a small portion of the work necessary if uniform grid refinement is employed. An adaptive algorithm requires appropriate error estimators for the iteration error and for the discretization error. The first estimator is used to define a stopping criterion for the iterative eigensolver on the actual grid whereas the second estimator controls the mesh refinement. Realizations of such estimators are, for example, suggested in [5, 11, 20].

Let us turn to the multigrid eigensolvers I–III, whose discussion because of (1) can be restricted to a single-vector iteration.

I. *Rayleigh quotient multigrid minimization (RQMG)*: We consider the eigenvalue problem for a *self-adjoint* elliptic partial differential operator. Then the discrete eigenproblem (1.1) can be treated as an optimization problem for the Rayleigh quotient

$$\lambda(x) = \frac{(x, Ax)}{(x, Mx)}. \tag{2.1}$$

By the Courant-Fischer principle the minimum of (2.1) equals the smallest eigenvalue of (A, M) and is taken in the corresponding eigenvector. Hence iterative minimization of (2.1) can serve as an eigensolver.

This minimization can be realized by means of a multigrid procedure. One applies a coordinate relaxation scheme, i.e. for each coordinate direction d_i^k (which is associated with the i-th finite element function on a certain grid level k) one computes the minimum

$$\lambda(x + \vartheta^* d_i^k) = \min_{\vartheta \in \mathbb{R}} \frac{(x + \vartheta d_i^k, A(x + \vartheta d_i^k))}{(x + \vartheta d_i^k, M(x + \vartheta d_i^k))}, \tag{2.2}$$

which is at the same time the smallest Ritz value of (A, M) in the 2D space $\mathrm{span}\{x, d_i^k\}$. The new iterate is $x + \vartheta^* d_i^k$. A multigrid cycle of RQMG consists in a successive minimization of $\lambda(x)$ for all finite element functions on all grid levels [4, 16, 18]. Grid independent convergence estimates for RQMG have been presented in [18]. For a generalization of RQMG to non-selfadjoint elliptic operators see [5].

II. *Direct multigrid eigensolvers*: Direct multigrid eigensolvers and the third class of eigensolvers (see III) are related to approximate variants of "classical" iterative eigensolvers, namely inverse iteration and the Rayleigh quotient iteration [23]. Inverse iteration applied to (1.1) generates a sequence of iterates $x^{(k)}$ by solving the linear systems

$$(A - \sigma M)x^{(k+1)} = Mx^{(k)}, \qquad k = 0, 1, 2, \ldots, \tag{2.3}$$

where σ denotes an eigenvalue approximation. In practice, the iterates are normalized after each step. The main hurdle emerging from (2.3) is that if σ is near to an eigenvalue, then $A - \sigma M$ is almost singular, cf. [17]. For instance, for the Rayleigh quotient iteration it holds $\sigma = \lambda(x^{(k)})$. Solving such equations with multigrid methods is a hard task [21, 25, 26]. To overcome this difficulty, one can either bound away σ from the eigenvalues of (A, M), see [1]. Alternatively, one can solve (instead of (2.3)) a non-singular *coarse grid correction equation* within the *orthogonal complement* of the actual eigenvector approximation. The latter approach provides the basis for the *direct multigrid eigensolver* introduced in [7, 8]. In order to realize such a coarse grid projection on the orthogonal complement, the algorithm is closely related to nested iteration which serves to provide sufficiently accurate eigenfunctions for building reliable coarse grid projections.

The resulting *two-grid method* maps a given iterate x having the Rayleigh quotient $\lambda(x)$ to the new eigenvector approximations x'. It is given by

$$\begin{aligned}
\tilde{x} &= Sx, && \text{(Smoothing step)},\\
d_c &= R(A - \lambda(x)M)\tilde{x}, && \text{(Coarse grid projection of the residual)},\\
d_c^{\perp} &= Q_c d_c, && (M\text{-orthogonal projection}),\\
v_c &= (A_c - \lambda(x)M_c)^{-1} d_c^{\perp}, && \text{(Solution of correction equation)},\\
x' &= x - PQ_c v_c, && \text{(Prolongation and correction)}.
\end{aligned} \tag{2.4}$$

Therein, the index c denotes coarse grid quantities. R is a restriction operator, P is a prolongation and Q_c is the orthogonal projection operator to the M-orthogonal complement of the actual eigenvector approximation. The coarse grid problem (2.4) in the orthogonal complement is quite different from (2.3) and can be solved by a multigrid scheme for singular equations [8].

III. *Eigensolvers using multigrid as a linear solver*: Let us consider an alternative way of solving (2.3) without getting into trouble due to singularity of $A - \sigma M$. We can do this by ignoring the shift parameter, i.e. $\sigma = 0$. Additionally, we assume an eigenproblem for a coercive and self-adjoint partial differential operator so that A and M are symmetric positive definite matrices. Now we apply multigrid preconditioning for A in order to determine an approximate solution of the linear system

$$Ax^{(k+1)} = \lambda(x^{(k)})Mx^{(k)}, \qquad k = 0, 1, 2, \ldots. \tag{2.5}$$

In comparison to (2.3), we have introduced the scaling constant $\lambda(x^{(k)})$ on the right hand side of (2.5) in order to achieve its stationarity in eigenvectors. Such a scaling prepares the approximate solution of (2.5) by means of preconditioning.

The multigrid preconditioner B^{-1} is an approximate inverse of A which is assumed to be a symmetric positive definite operator and which is characterized by a spectral equivalence

$$\gamma_0(x, Bx) \le (x, Ax) \le \gamma_1(x, Bx), \qquad \forall x \in \mathbb{R}^n,$$

for positive constants γ_0 and γ_1. It is often more convenient to assume a properly scaled preconditioner in such a way that

$$\|I - B^{-1}A\|_A \leq \gamma, \tag{2.6}$$

for a constant $\gamma \in [0,1)$; therein $\|\cdot\|_A$ denotes the operator norm induced by A. This assumption is typical of multigrid and domain decomposition preconditioners. The best preconditioners satisfy (2.6) with γ bounded away from 1 independently of the mesh size and hence the number of unknowns [2]. The approximate solution of (2.5) for $x = x^{(k)}$ and by using B^{-1} as a preconditioner yields a new iterate x' approximating $x^{(k+1)}$

$$x' = x - B^{-1}(Ax - \lambda(x)Mx). \tag{2.7}$$

The iteration (2.7) can be considered as the most simple eigensolver embodying the idea of *multigrid as a linear solver*. It can be interpreted as a (multigrid) **p**reconditioned variant of **inv**erse **it**eration (PINVIT) [19]. In principle, its behavior can be understood by the associated *error propagation equation*

$$x' - \lambda(x)A^{-1}Mx = (I - B^{-1}A)(x - \lambda(x)A^{-1}Mx), \tag{2.8}$$

which maps the error of the initial iterate $x - \lambda(x)A^{-1}Mx$ to the error of the new iterate $x' - \lambda(x)A^{-1}Mx$ by premultiplying the error propagation matrix $I - B^{-1}A$. We can also derive from (2.8) that

$$\|x' - \lambda(x)A^{-1}Mx\| \leq r\,\|A^{-1} - B^{-1}\|,$$

which shows the Lipschitz continuity in A^{-1} of the mapping from the set of admissible preconditioners, as characterized by (2.6), to the set of possible new iterates. The Lipschitz constant r equals the norm of the residual vector $Ax - \lambda(x)Mx$.

Keeping these interpretations in mind, it is not surprising that the convergence estimates for (2.7) are closely related to inverse iteration. Sharp convergence estimates in terms of the Rayleigh quotients of the iterates are available. If the Rayleigh quotient $\lambda(x)$ is located between the smallest eigenvalue λ_1 and the next larger eigenvalue λ_2 then it holds

$$\frac{\lambda(x') - \lambda_1}{\lambda_2 - \lambda(x')} \leq \left(\gamma + (1-\gamma)\frac{\lambda_1}{\lambda_2}\right)^2 \frac{\lambda(x) - \lambda_1}{\lambda_2 - \lambda(x)}, \tag{2.9}$$

see [14, 19]. This estimate can be applied recursively and guarantees linear convergence of the eigenvalue approximations $\lambda(x^{(k)})$ to the eigenvalue λ_1. The case $\gamma = 0$ corresponding to an exact solution of (2.5) leads to the convergence factor $(\lambda_1/\lambda_2)^2$ of inverse iteration. For the other extreme case of poorest preconditioning, i.e. $\gamma \to 1$, the convergence factor tends to 1.

We have assumed in (2.7) a symmetric positive definite preconditioner B^{-1}. But there is no consent in the literature whether positive definite or indefinite preconditioners lead to more efficient eigensolvers, cf. [12]. This has also opened the question of how accurately (2.5) is to be solved in order to achieve reasonable convergence estimates [6, 15, 22, 24]. In any way the cubic convergence of the Rayleigh quotient iteration cannot be transferred to the preconditioned multigrid case.

A significant acceleration of (2.7) can be achieved by minimizing the Rayleigh quotient with respect to a certain search subspace. The idea of the Locally Optimal Block Preconditioned Conjugate Gradient method (LOBPCG, [13]) is to determine the new eigenvalue/vector approximation as the smallest Ritz value/vector with respect to the three-dimensional space

$$\operatorname{span}\{x, B^{-1}(Ax - \lambda(x)Mx), x^-\}, \tag{2.10}$$

where x^- denotes the iterate preceding x. To implement the latter iteration only one additional vector (compared to (2.7)) is to be stored (i.e. the old iterate) and few additional inner products are to be computed to determine the Rayleigh-Ritz projections. It has been observed that these additional costs are more than compensated by a characteristic speedup of convergence. Practically, the LOBPCG scheme behaves like a *preconditioned conjugate gradient iteration* for boundary value problems [13].

3. NUMERICAL EXPERIMENTS

All numerical experiments covered in this section treat the eigenproblem for the Laplacian on $[0,1]^2$ with homogeneous Dirichlet boundary conditions as a simple model problem. See [5, 20] for adaptive multigrid eigensolvers on non-convex domains and with discontinuous coefficient functions. The eigenproblem is discretized by using linear finite elements on uniform triangle meshes with mesh sizes $h_l = 2^{-l}$. The initial grid consists of 9 interior nodes for $l = 2$. Computations are presented up to $l = 12$ corresponding to a grid with 16769025 interior nodes.

Table 1 lists the convergence history of RQMG, PINVIT and LOBPCG for the smallest eigenvalue and the corresponding residuals for $l = 6$ (3969 inner nodes). We use a unigrid version of RQMG, which is easy to code but which is computationally expensive, see [16] for the details. The coordinate relaxation scheme is applied to a black-white ordering of the finite element functions. The preconditioned eigensolvers are equipped with a multigrid $V(2,2)$-cycle preconditioner working on the sequence of grids for $l = 2, \ldots, 6$. Pre- and postsmooting is done by each 2 steps of the weighted Jacobi method. For all computations the initial function in each node (x_1, x_2) is given by $x_1^2 + x_2^2$. Table 1 contains the eigenvalue approximations, i.e. the Rayleigh quotients of the iterates, and the corresponding Euclidean norms of the residuals $Ax - \lambda(x)Mx$ with $\|x\|_M = 1$. RQMG and LOBPCG converge within 5 iterations to a six-digit accuracy, whereas more than 10 iterations of PINVIT are required to reach the same accuracy. All these iterations are guaranteed to generate a sequence of monotone decreasing Rayleigh quotients.

	Rayleigh quotients $\lambda(x^{(k)})$			Residuals $\|Ax^{(k)} - \lambda(x^{(k)})Mx^{(k)}\|$		
k	RQMG	PINVIT	LOBPCG	RQMG	PINVIT	LOBPCG
0	426.5873	426.5873	426.5873	$2.07 \cdot 10^{1}$	$2.07 \cdot 10^{1}$	$2.07 \cdot 10^{1}$
1	19.787335	23.621502	22.703719	$1.97 \cdot 10^{-3}$	$5.32 \cdot 10^{-1}$	$6.79 \cdot 10^{-1}$
3	19.751232	19.897647	19.755509	$9.53 \cdot 10^{-5}$	$4.50 \cdot 10^{-2}$	$8.62 \cdot 10^{-2}$
5	19.751101	19.760942	19.751101	$5.21 \cdot 10^{-6}$	$9.08 \cdot 10^{-3}$	$5.80 \cdot 10^{-4}$
10	19.751101	19.751117	19.751101	$4.18 \cdot 10^{-9}$	$3.48 \cdot 10^{-4}$	$5.03 \cdot 10^{-8}$

TABLE 1. *RQMG, PINVIT and LOBPCG convergence history for $l = 6$.*

Mesh independence of the convergence of (2.7) is illustrated by Table 2. On each grid with $h = 1/2^l$, $l = 4, \ldots, 12$, the iteration is started with the restriction of $x_1^2 + x_2^2$ to the nodes. Then 25 iterations of (2.7) with $V(2,2)$-Jacobi preconditioning have been applied resulting in a final iterate x_h. The final eigenvalue approximations $\lambda(x_h)$ are listed in Table 2. They reflect that $\lambda(x_h) - \lambda_1 = \mathcal{O}(h^2)$, whereas λ_1 denotes the smallest eigenvalue of the continuous problem. Moreover, the norms of the residuals (defined as above) are shown, too. Note that the M^{-1} norm of the residual is an upper estimate for the error of the eigenvalue approximations. As the condition number of M is bounded independently of h, the last column of Table 2 indicates the mesh independent convergence, since the final

Level l	Grid	Inner nodes	$\lambda(x_h)$	$\|Ax_h - \lambda(x_h)Mx_h\|$
4	17×17	225	19.9297898	$7.14 \cdot 10^{-8}$
5	33×33	961	19.7867923	$4.53 \cdot 10^{-8}$
6	65×65	3969	19.7511008	$2.41 \cdot 10^{-8}$
7	129×129	16129	19.7421816	$1.23 \cdot 10^{-8}$
8	257×257	65025	19.7399520	$6.20 \cdot 10^{-9}$
9	513×513	261121	19.7393946	$3.12 \cdot 10^{-9}$
10	1025×1025	1046529	19.7392553	$1.56 \cdot 10^{-9}$
11	2049×2049	4190209	19.7392204	$7.85 \cdot 10^{-10}$
12	4097×4097	16769025	19.7392117	$2.08 \cdot 10^{-10}$

TABLE 2. *Eigenvalues and residuals after 25 iterations of* (2.7) *for* $h_l = 2^{-l}$, $l = 4, \ldots, 12$.

residuals do not deteriorate as the mesh size gets smaller. Let us finally mention that it is a nontrivial task to give a fair comparison (in terms of total computational costs and efficiency) of the direct multigrid eigensolver [7, 8] with those eigensolvers discussed above. This is the case since the direct multigrid scheme is inseparably connected with nested iteration in order to generate the required projection operators. Therefore, we only refer to [9], where the direct multigrid eigensolver is applied to the biharmonic plate equation.

4. SUMMARY

A review has been given on three quite different multigrid techniques for solving the eigenvalue problem for a (coercive and self-adjoint) elliptic partial differential operator. Each of these solvers achieves optimal or quasi-optimal complexity for the *partial* eigenproblem, i.e. to compute only a small number of eigenpairs. These multigrid eigensolvers can be considered as the eigenproblem-counterparts of the celebrated, highly efficient multigrid solvers for boundary value problems.

Eigensolvers using multigrid as a linear solver are conceptually very simple, and recent results allow a new theoretical understanding. They can be implemented easily: For the multigrid part one can adopt any existing program code realizing multigrid preconditioning. The remaining task is only to implement the linear algebra of (2.7) or (2.10).

REFERENCES

[1] R.E. Bank. Analysis of a multilevel inverse iteration procedure for eigenvalue problems. *SIAM J. Numer. Anal.*, 19:886–898, 1982.

[2] J.H. Bramble, J.E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55(191):1–22, 1990.

[3] A. Brandt, S. McCormick, and J. Ruge. Multigrid methods for differential eigenproblems. *SIAM J. Sci. Statist. Comput.*, 4(2):244–260, 1983.

[4] Z. Cai, J. Mandel, and S. McCormick. Multigrid methods for nearly singular linear equations and eigenvalue problems. *SIAM J. Numer. Anal.*, 34:178–200, 1997.

[5] T. Friese, P. Deuflhard, and F. Schmidt. A multigrid method for the complex Helmholtz eigenvalue problem. In *Eleventh International Conference on Domain Decomposition Methods (London, 1998)*, pages 18–26 (electronic). DDM.org, Augsburg, 1999.

[6] G.H. Golub and Q. Ye. Inexact inverse iteration for generalized eigenvalue problems. *BIT*, 40(4):671–684, 2000.

[7] W. Hackbusch. On the computation of approximate eigenvalues and eigenfunctions of elliptic operators by means of a multi-grid method. *SIAM J. Numer. Anal.*, 16:201–215, 1979.

[8] W. Hackbusch. *Multi-grid methods and applications*. Springer series in computational mathematics 4. Springer, Berlin, 1985.
[9] W. Hackbusch and G. Hofmann. Results of the eigenvalue problem for the plate equation. *Z. Angew. Math. Phys.*, 31(6):730–739, 1980.
[10] V. Heuveline and C. Bertsch. On multigrid methods for the eigenvalue computation of nonselfadjoint elliptic operators. *East-West J. Numer. Math.*, 8(4):275–297, 2000.
[11] V. Heuveline and R. Rannacher. A posteriori error control for finite element approximations of elliptic eigenvalue problems. *Adv. Comput. Math.*, 15(1-4):107–138, 2001.
[12] A.V. Knyazev. Preconditioned eigensolvers—an oxymoron? *Electron. Trans. Numer. Anal.*, 7:104–123, 1998.
[13] A.V. Knyazev. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. *SIAM J. Sci. Comp.*, 23:517–541, 2001.
[14] 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 Appl.*, 2001.
[15] Y. Lai, K. Lin, and W. Lin. An inexact inverse iteration for large sparse eigenvalue problems. *Numer. Linear Algebra Appl.*, 4:425–437, 1997.
[16] J. Mandel and S.F. McCormick. A multilevel variational method for $Au = \lambda Bu$ on composite grids. *J. Comput. Phys.*, 80:442–452, 1989.
[17] S.F. McCormick. A mesh refinement method for $Ax = \lambda Bx$. *Math. Comp.*, 36(154):485–498, 1981.
[18] S.F. McCormick. Multilevel adaptive methods for elliptic eigenproblems: a two-level convergence theory. *SIAM J. Numer. Anal.*, 31:1731–1745, 1994.
[19] K. Neymeyr. A geometric theory for preconditioned inverse iteration. I: Extrema of the Rayleigh quotient. *Linear Algebra Appl.*, 322:61–85, 2001.
[20] K. Neymeyr. A posteriori error estimation for elliptic eigenproblems. *Numer. Linear Algebra Appl.*, 9:263–279, 2002.
[21] M.G. Neytcheva and P.S. Vassilevski. Preconditioning of indefinite and almost singular finite element elliptic equations. *SIAM J. Sci. Comput.*, 19(5):1471–1485, 1998.
[22] Y. Notay. Convergence analysis of inexact Rayleigh quotient iterations. Technical report, Service de Métrologie Nucléaire, Université Libre de Bruxelles, 2001.
[23] B.N. Parlett. *The symmetric eigenvalue problem*. Prentice Hall, Englewood Cliffs New Jersey, 1980.
[24] P. Smit and M. Paardekooper. The effects of inexact solvers in algorithms for symmetric eigenvalue problems. *Linear Algebra Appl.*, 287:337–357, 1999.
[25] J. Xu. A new class of iterative methods for nonselfadjoint or indefinite problems. *SIAM J. Numer. Anal.*, 29:303–319, 1992.
[26] H. Yserentant. Preconditioning indefinite discretization matrices. *Numer. Math.*, 54(6):719–734, 1989.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT TÜBINGEN, AUF DER MORGENSTELLE 10, 72076 TÜBINGEN, GERMANY.
E-mail address: neymeyr@na.uni-tuebingen.de